Face a este aumento de competitividade, os gestores necessitaram de tomar decisões cada vez mais rápidas com vista a uma resposta atempada às exigências da procura. Neste contexto, as aplicações informáticas integradas para o meio empresarial (ERP) surgiram como uma evolução dos primeiros sistemas MRP (pela necessidade intrínseca de integração total dos fluxos das organizações) e tornaram-se ferramentas imprescindíveis no auxílio ao processo de tomada de decisão (ao auxiliar a uniformização dos processos de negócio). No entanto, a abordagem dos sistemas ERP tem sido colocada em causa devido à sua falta de orientação para a gestão por processos (principalmente, dos pedidos de cotação e gestão de encomendas de cliente).

Este artigo inicia-se com uma avaliação do processo tradicional de gestão de encomendas levado a cabo pela maioria dos sistemas ERP. Posteriormente, é apresentada uma análise (sob vários indicadores) de alguns dos actuais sistemas ERP existentes no mercado, com vista à avaliação das suas metodologias de gestão de processos de resposta às necessidades dos gestores (no processo de tomada de decisão).

Este artigo apresenta igualmente um protótipo aplicacional (eventualmente para ser integrado num sistema ERP) que se assuma como um sistema de apoio à decisão para os gestores de produção. Este protótipo pretende disponibilizar um conjunto de critérios aos utilizadores, permitindo a sua manipulação e ajuste, de forma a criar um ambiente de simulação para possíveis prazos de entrega. Estas simulações deverão basear-se na capacidade produtiva real da empresa industrial, de forma a garantir um maior grau de confiança dos valores propostos e tornarem-se numa importante ferramenta de auxílio à tomada de decisão.

2. Sistemas ERP e a Gestão de Capacidades

A abordagem tradicional de um sistema ERP para o tratamento de pedidos de encomenda engloba os passos mostrados na Figura 1.

Na Figura 1 denota-se que os sistemas ERP apresentam uma falha no conceito de colaboração e “direccionamento à encomenda”, pois o sistema é totalmente baseado nos requisitos de cada uma das partes (do produto) e não da encomenda de cliente. Um sistema direccionado à encomenda permitiria o rastreio total da mesma, garantindo a sua gestão completa e um controlo mais apurado.

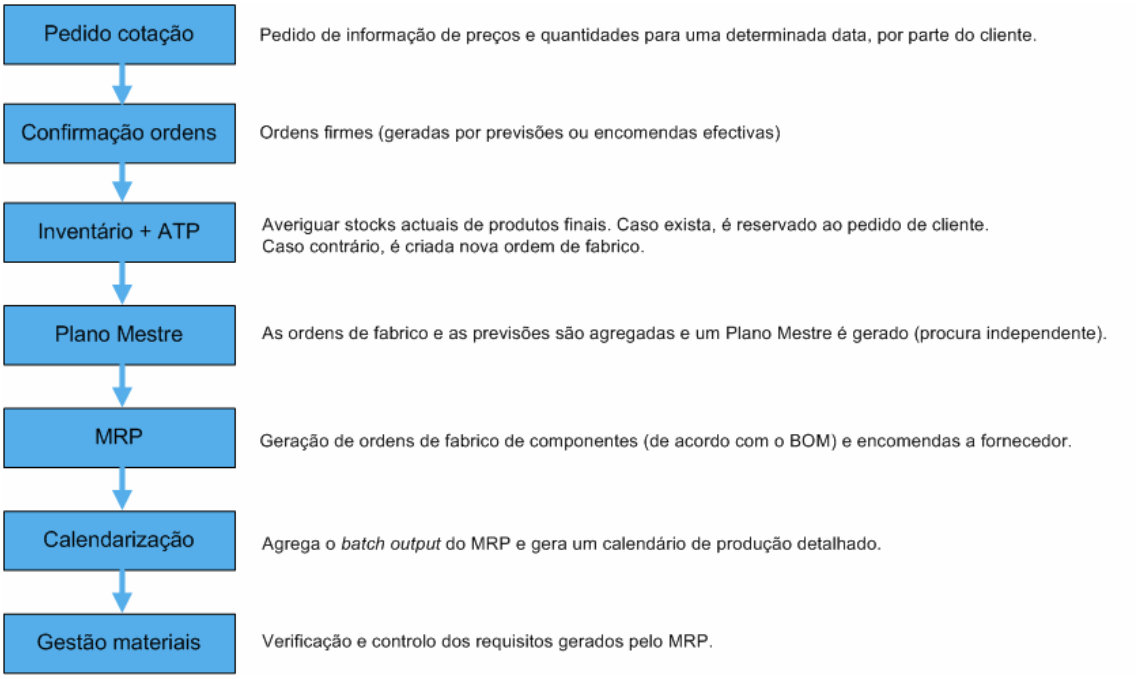


Figura 1 – Passos tradicionais de um sistema ERP para tratamento de pedidos de encomenda

(adaptado de [Makatsoris 2004])

Uma das maiores lacunas deste tipo de sistemas é a falta de uma gestão dinâmica de alterações correntes (falhas nas entregas dos fornecedores, faltas imprevistas de funcionários, avarias extraordinárias das máquinas, etc). Desta forma, os gestores de produção não conseguem avaliar eficazmente as alterações provocadas por estas "variáveis" que influenciam o seu planeamento.

Para validar se os actuais sistemas ERP possuem uma nova abordagem (mais abrangente), foi preparado um estudo que envolveu a caracterização de alguns sistemas ERP existentes no mercado. Foram seleccionados três sistemas ERP (dois deles desenvolvidos por empresas internacionais, e um desenvolvido por uma empresa nacional) com "expressão" no mercado nacional. Os critérios utilizados para a sua selecção foram: a garantia de integração da Gestão de Produção com a Gestão Comercial e Financeira, não serem específicos para um determinado negócio (ou seja, serem sistemas de índole *standard*) e possuírem um número de implementações superior a 60 empresas industriais, sobre as mais diversas áreas.

Os sistemas de informação seleccionados para o estudo foram: o software Manufactor (integrado com software PHC), Microsoft Navision 4.0 e SAP R/3. A análise dos sistemas de informação envolveu a sua utilização prática, com a criação de um exemplo real (comum a todos os sistemas ERP), para um conjunto de competências apresentadas na Figura 2.

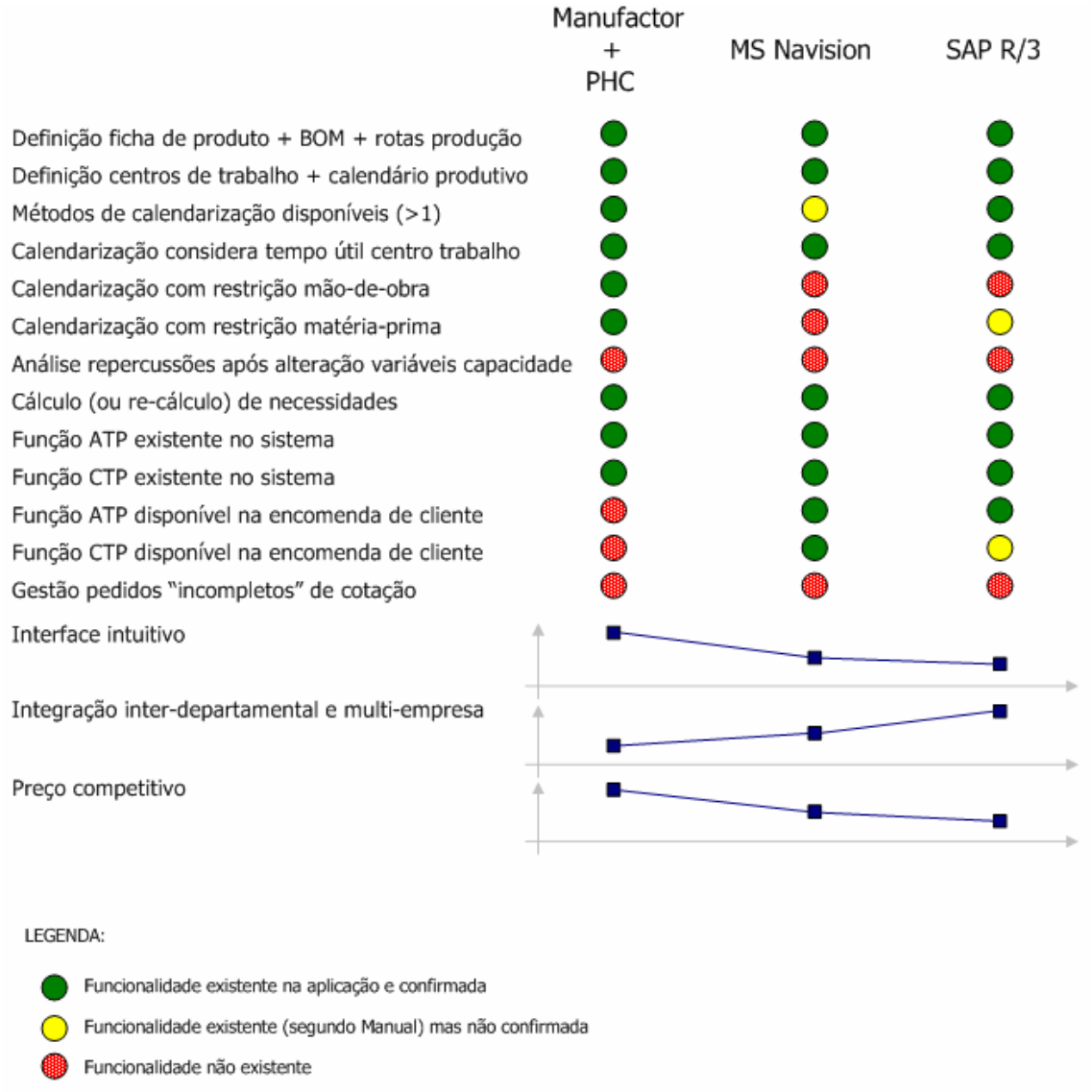


Figura 2 – Comparação de funcionalidades (negociação de encomenda) entre os sistemas ERP

Esta comparação foi elaborada pela experiência prática e directa nos sistemas de informação empresariais em causa (recorrendo a versões de demonstração), sem recorrer a especialistas ou consultores destes sistemas.

Este estudo comparativo permitiu, assim, concluir que ainda existem requisitos (análise de repercussões com recurso à capacidade produtivo e a gestão de pedidos incompletos) aos quais os sistemas de informação actuais ainda não conseguem responder.

3. Arquitectura proposta

Como tentativa de resposta aos requisitos estudados no capítulo anterior (nomeadamente, a análise de repercussões após alteração das variáveis de capacidade) optamos pelo desenho de um novo sistema que permitisse uma gestão dinâmica de capacidade.

Pretende-se um sistema que auxilie o gestor na tomada de decisão, no acto de negociação de encomenda de cliente, e que englobe o desenvolvimento de funcionalidades de parametrização e simulação de cenários de calendarização possíveis (para servir de resposta aos pedidos de cotação e/ou encomendas de cliente). Pretende-se ainda "fortalecer" o modelo de capacidades do sistema, ao disponibilizar análises de repercussões e a flexibilizar a gestão dos recursos da empresa.

Um sistema deste tipo caracteriza-se pela sua enorme abrangência e necessidade de integração com os demais processos internos de uma organização, nomeadamente a gestão de encomendas de clientes, gestão de ordens de fabrico e métodos de calendarização, com vista a um completo diagnóstico das necessidades de recursos e materiais.

O sistema proposto foi dividido em quatro subsistemas, agrupados com base nas suas funcionalidades, nomeadamente *Gestão de Encomendas*, *Calendarização de Ordens de Fabrico*, *Modelo de Capacidades* e *Análise de Repercussões*. A figura 3 apresenta o diagrama de pacotes de casos de uso (esquema desenvolvido em UML) para o sistema proposto.

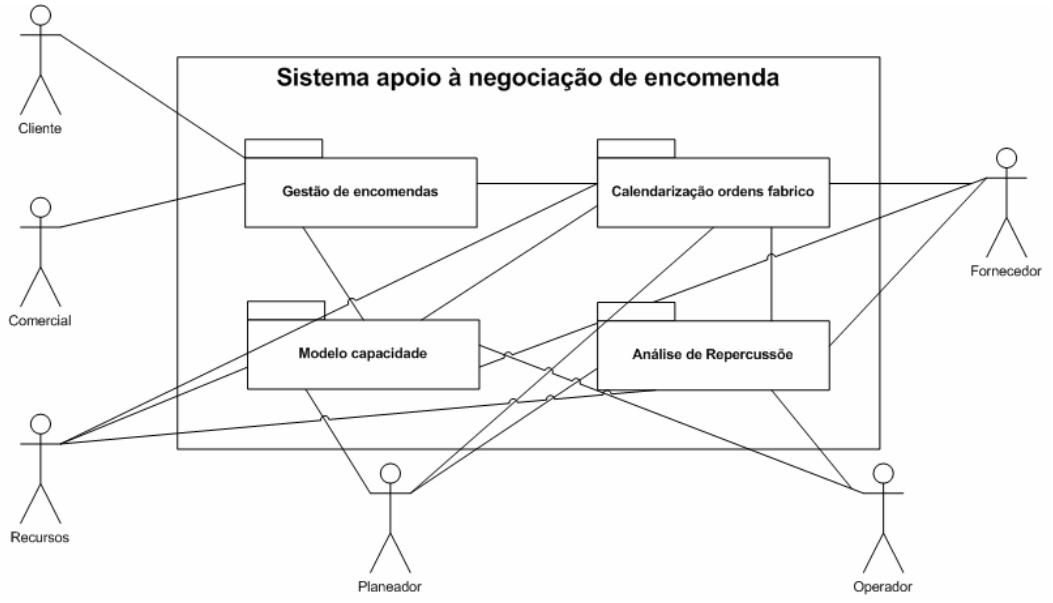


Figura 3 – Diagrama de pacotes (UML) do sistema proposto

O subsistema *Gestão de Encomendas* representa todas as funcionalidades relativas à gestão de encomendas do sistema. Tem como objectivo principal fornecer uma data de entrega possível e fiável ao utilizador (recorrendo ao stock previsto ou validando a sua capacidade produtiva, ou seja, recorrendo aos subsistemas de Calendarização de Ordens de Fabrico e Modelo de Capacidades).

O subsistema *Calendarização de Ordens de Fabrico* representa todas as funcionalidades relativas à calendarização das operações das ordens de fabrico. Tem como objectivo definir a ordem de prioridade e o método de calendarização das ordens de fabrico.

O subsistema *Modelo de Capacidades* é responsável pela análise da capacidade produtiva. Tem como objectivo principal "auxiliar" todos os subsistemas que efectuem a calendarização de ordens de fabrico, ao considerar a ocupação dos centros de trabalho, mão-de-obra e matérias-primas no cálculo das datas de início ou fim das operações das ordens de fabrico.

O subsistema *Análise de Repercussões* é responsável pela análise de todas as variações que possam ocorrer nas variáveis e parâmetros que afectem o modelo de capacidades. Face à sua complexidade, tornou-se no objectivo fulcral deste artigo! Tem como objectivo fornecer, ao utilizador, uma análise completa das alterações ocorridas às variáveis e parâmetros críticos do modelo de capacidade produtiva. Deverá permitir efectuar acções de re-planeamento para adaptar a calendarização global às novas condições.

A opção escolhida foi o desenho do protótipo para integração com o software Manufactor. A linguagem de programação sugerida para o desenvolvimento deste sistema foi o Microsoft Visual FoxPro 9.0 (devido à sua especialização e facilidade de tratamento de grandes quantidades de dados e devido à utilização da mesma no desenvolvimento do software Manufactor) e, para sistema de gestão de base de dados cliente-servidor, a escolha recaiu sobre o Microsoft SQL Server 2005.

O sistema proposto deverá ser desenhado (idealmente) para um sistema de 3 camadas, por razões de baixa de custo numa futura mudança da lógica de negócios ou de sistema de bases de dados.

A arquitectura lógica é representada na figura 4.

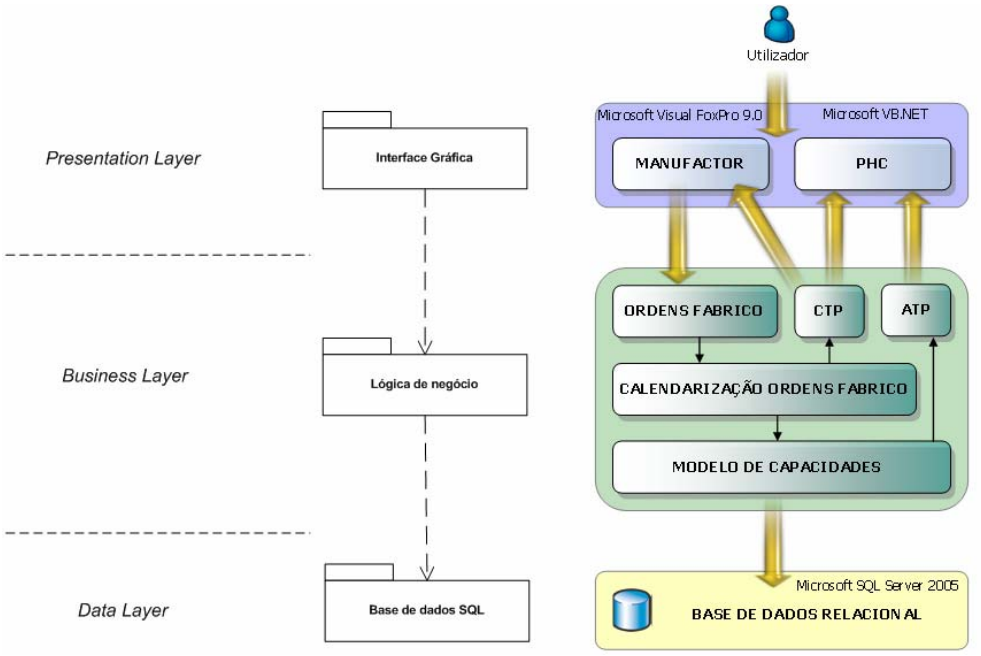


Figura 4 – Arquitectura lógica do sistema

4. Desenvolvimento do protótipo

Como já referido anteriormente, o novo sistema deverá fornecer (ao gestor de produção) uma interface visual totalmente manipulável, promovendo uma integração com o modelo de capacidades.

Neste capítulo serão apresentados vários exemplos de interfaces possíveis para uma análise detalhada dos recursos (centros de trabalho, mão-de-obra e matéria-prima).

4.1. Análise de recursos (centros de trabalho)

Esta funcionalidade pretende disponibilizar um conjunto de indicadores básicos para auxiliar o gestor a determinar uma data de entrega possível, para responder a uma encomenda de cliente. É essencial que o utilizador possa avaliar (segundo vários critérios) a capacidade dos centros de trabalho críticos, num interface elucidativo como o ilustrado graficamente na Figura 5.

Desta forma, o utilizador tem uma visão geral e imediata da disponibilidade produtiva, das reservas efectuadas para pedidos de encomenda, das paragens para manutenção e, até mesmo, recorrer a dados provindos de histórico (anos anteriores).

Figura 5 - Análise de recursos: centros de trabalho

5.2. Análise de recursos (mão-de-obra)

Uma das variáveis que também influencia o cálculo da capacidade é a gestão de afectação da mão-de-obra às operações das ordens de fabrico. A figura 6 ilustra uma interface de análise da afectação da mão-de-obra, horas extraordinárias e faltas inesperadas dos seus funcionários mais críticos, para um determinado período. Este tipo de análise é fundamental para antecipar alguns comportamentos imprevistos de alguns funcionários e que afectam directamente o planeamento de produção.

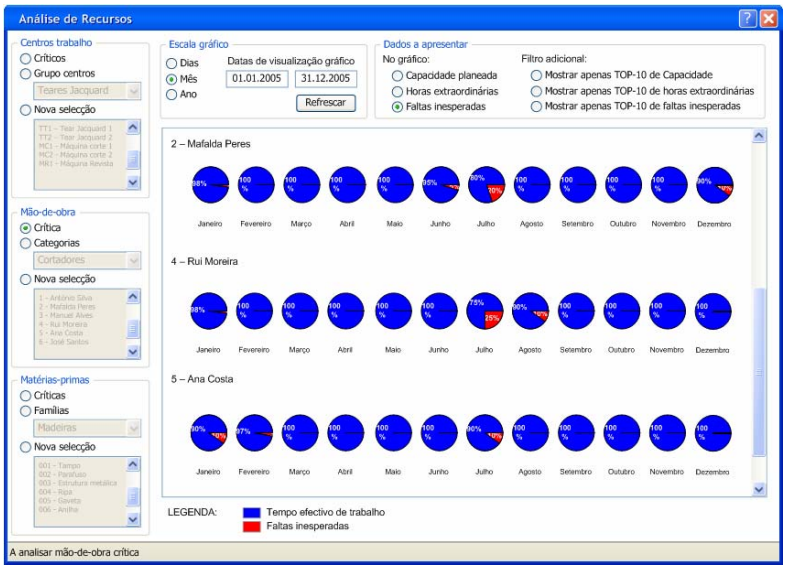


Figura 6 - Análise de recursos: mão-de-obra

5.3. Análise de recursos (matéria-prima)

Uma outra variável que influencia o cálculo da capacidade é a gestão de afectação da matéria-prima às operações das ordens de fabrico. Para estes cenários, é também necessário estar atento às falhas de entrega por parte dos fornecedores para antecipar alguns problemas na gestão de stocks. A análise ilustrada na figura 7 permite avaliar as variações de stock para um determinado período, bem como o stock reservado para efeitos de capacidade, tempos de entrega médios, etc.

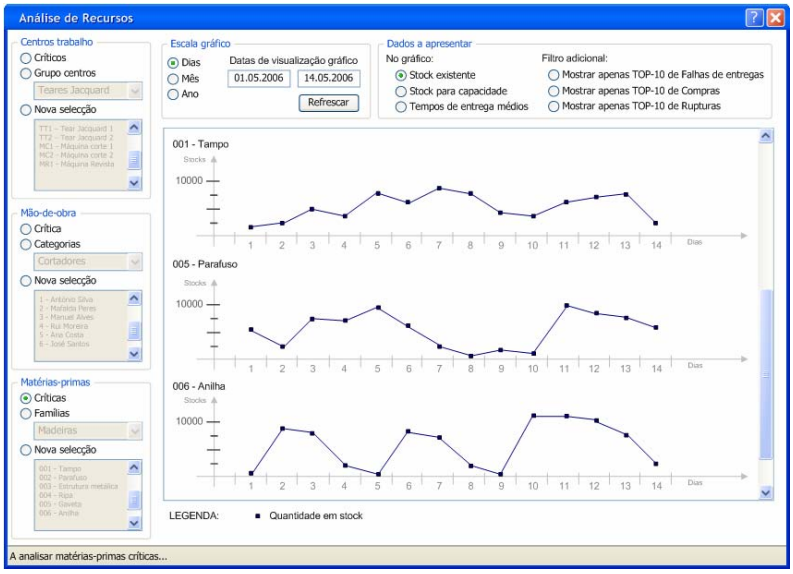


Figura 7 - Análise de recursos: matérias-primas

5.4. Análise de repercussões

Pretende-se que, com esta funcionalidade, o utilizador seja avisado de qualquer alteração que ocorra a qualquer variável que influencie a capacidade produtiva, e que lhe permita tomar uma acção correctiva para repor o planeamento segundo um conjunto de parâmetros aceitáveis. A ideia base é disponibilizar as variáveis que foram recentemente alteradas e que carecem da análise e rectificação por parte do utilizador. A figura 8 ilustra uma interface de decisão de elevada potencialidade, pois estão disponíveis as variáveis que influenciam directamente o planeamento de produção e capacidade produtiva, permitindo uma análise global em poucos segundos. O sistema indica, numa única interface e por intermédios de ícones elucidativos, os problemas e situações detectadas, permitindo analisar recursos alternativos (antes de optar por uma solução de aumento de capacidade).

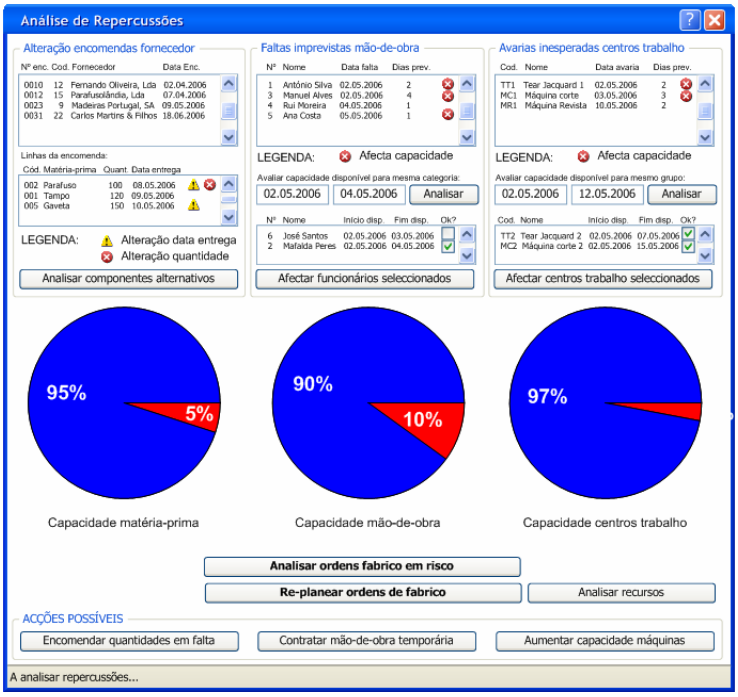


Figura 8 - Análise de repercussões

O sistema permite uma análise rápida das ordens de fabrico em risco (de não serem entregues nas datas confirmadas) e, acedendo à interface de análise de recursos, tomar acções de aquisição de capacidade e re-planeamento de ordens de fabrico.

5. Conclusões e trabalhos futuros

A evolução dos mercados tem sido marcada por um crescimento acentuado dos modelos de negócio MTO em virtude de ser exigida, cada vez mais, uma especialização nas necessidades do cliente. Esta "vertente de personalização", associada à imprevisibilidade da procura, obriga as empresas a um controlo apurado de custos e aproveitamento dos seus recursos, recorrendo a técnicas de planeamento que incluam várias restrições para a gestão da capacidade. No entanto, uma questão foi tomada como mais complexa: a análise de repercussões de alterações da capacidade produtiva.

Ao optar por técnicas de planeamento com inclusão de variáveis de capacidade finita, as empresas ficam "mais sujeitas" às variações que possam ocorrer (como alterações de datas de entrega de matérias-primas, avarias inesperadas de centros de trabalho, etc). A este nível, os sistemas de informação existentes deveriam disponibilizar um maior controlo das variáveis que influenciam a capacidade produtiva, e permitir ao gestor tomar acções correctivas que

permitissem um re-planeamento eficaz. Após a análise (segundo diversos critérios) de três sistemas ERP existentes no mercado, foram encontradas algumas lacunas na resposta a estes requisitos.

Como resposta a esta lacuna dos sistemas ERP, este artigo apresenta um novo sistema baseado na construção de um modelo de capacidades inovador, de elevada integração e "nervosismo" (atento a qualquer alteração das suas variáveis), que fornece ao gestor várias panorâmicas da sua capacidade produtiva e permite, em poucos passos, efectuar um conjunto de acções rectificativas. O sistema proposto (a integrar no software Manufactor) garantiria uma maior abrangência e segurança na tomada de decisão dos gestores, ao disponibilizar uma análise completa das repercussões e um processo bem mais facilitado de ajustamento dos recursos da organização, com vista à satisfação plena da capacidade produtiva.

O próximo passo será o desenvolvimento deste protótipo, com vista à sua validação e aplicação real em organizações de índole industrial.

6. Referências

Higgins, P & Le Roy, P & Tierney, L 1996, *Manufacturing Planning and Control Beyond MRP II*, 1ª edição, Chapman & Hall

Xiong M, Tor, S, Khoo, L, Bhatnagar R 2004, "Fulfillment of Rush Customer Orders under Limited Capacity", *Singapore-MIT Alliance*

Xiong M & Tor, S & Khoo, L 2003, "WebATP: a Web-based flexible available-to-promise computation system", *Prodution Planning & Control*, vol. 14, no. 7, pp. 662-672

Makatsoris, C & Chang, S 2004, "Design of a demand-driven collaborative supply-chain planning and fulfilment system for distributed enterprises", *Prodution Planning & Control*, vol. 15, no. 3, pp. 256-269

Skok, W & Legg, M 2002, "Evaluating Enterprise Resource Planning (ERP) Systems using an Interpretive Approach", *Knowledge and Process Management*, vol. 9, no. 2, pp. 72-82

Wortman, J, Euwe, M, Tall M, Wiers V 1996, "A review of capacity planning techniques within standard software packages", *Production Planning & Control*, vol. 7, no. 2, pp.117-128

Courtois, A & Pillet, M & Martin-Bonnefous, C 2006, *Gestão da produção*, 5ª edição, Lidel

Apoio à decisão na gestão dinâmica da capacidade produtiva

Ricardo João Costa Almeida[1], Américo Lopes de Azevedo[2]

1) Universidade de Aveiro, Aveiro, Portugal

almeida.ricardo@gmail.com

2) Faculdade de Engenharia da Universidade do Porto, Porto, Portugal
ala@fe.up.pt

Resumo

A evolução dos mercados tem sido caracterizada por um crescimento acentuado dos modelos de negócio *"make-to-order"* (MTO) em virtude de ser exigida, cada vez mais, uma especialização nas necessidades do cliente. Esta vertente, associada à imprevisibilidade da procura, obriga as empresas a avaliar e afectar constantemente os seus recursos (desistindo dos tradicionais sistemas MRP – *"Material Requirement Planning"*) e a recorrer a técnicas de planeamento que incluam várias restrições para a gestão da capacidade.

Alguns sistemas ERP (*"Enterprise Resource Planning"*) existentes no mercado foram analisados sobre diversos critérios e todos eles apresentaram lacunas na análise de repercussões, requisito este fundamental para empresas com mais do que uma variável de restrição para capacidade produtiva.

Neste artigo é proposto um novo sistema para a construção de um modelo de capacidades inovador, de elevada integração e grau de "nervosismo" (atento a qualquer alteração das suas variáveis), que forneça ao gestor várias panorâmicas da sua capacidade produtiva. Este sistema deverá permitir, em pouco passos e num único ecrã, visualizar a globalidade da capacidade produtiva segundo vários parâmetros e efectuar um conjunto de acções rectificativas.

Palavras chave: Negociação de encomenda, Análise de repercussões, Capacidade produtiva, ATP, CTP, ERP

1. Introdução

As últimas décadas têm-se revelado, para o mercado industrial, como épocas de "ebulição" constante [Higgins 1996]. Se na década de 70 a massificação da produção era a doutrina predominante, o início do século XXI tem comprovado uma completa reviravolta na gestão da procura, denotando uma total orientação ao cliente e numa expectativa de fornecimento de produtos e serviços de qualidade indiscutível. Cada vez mais se optaram por modelos de negócio de produção por encomenda, face aos pedidos de elevada especialização de cada cliente, adicionando um elevado factor de incerteza ao seu negócio.

TEMA 1 – CICLO DE VIDA DOS SISTEMAS DE INFORMAÇÃO

1.2. Concepção e Desenvolvimento de Sistemas de Informação

Benchmark de bases de dados de suporte a serviços de informação

José Novais[1], Leonel Duarte do Santos [2]

1) Universidade do Minho – DSI, Guimarães, Portugal

josenovais@josenovais.com

2) Universidade do Minho – DSI, Guimarães, Portugal
leonel@dsi.uminho.pt

Resumo

A necessidade de fazer chegar informação a uma audiência cada vez mais vasta estimula o aparecimento de serviços com capacidades de recolha, armazenamento e difusão de informação, com o objectivo de facilitar a sua gestão, numa perspectiva de divulgação e de recolha, isto é, serviços de informação *online*. Para o bom funcionamento desses serviços, o armazenamento eficiente de informação assume um papel de grande relevância, pelo que se torna necessária a execução de testes para a selecção dos modelos mais adequados para esta tarefa. Neste trabalho são propostos um modelo e um sistema de testes que deverão ser capazes de permitir tirar conclusões sobre o modelo de armazenamento de informação (relacional ou XML) a adoptar pelo repositório de dados de suporte a serviços de informação *online*.

Palavras chave: Repositórios, *benchmarks*, BDs, serviços de informação *online*

1. Introdução

Actualmente, com o aumento da utilização da Internet e das tecnologias como um meio de comunicação privilegiado e com a diversificação das aplicações que vão sendo colocadas online, tem-se assistido à disponibilização de cada vez maiores quantidades de informação a um cada vez maior número de utilizadores. Nesta perspectiva, as entidades vão disponibilizando serviços com capacidades de recolha, armazenamento e difusão de informação com vista a facilitar a gestão da informação, quer numa perspectiva de divulgação quer de recolha. Estes serviços implicam normalmente a manipulação de grandes quantidades de informação e de sistemas capazes de executar esta tarefa eficazmente, aos quais pode ser dado o nome de repositórios. Desta forma, os repositórios são o que alimenta os serviços de informação. Para o armazenamento desta informação, e assumindo que tal é feito com recurso a bases de dados (BDs), existem actualmente dois modelos bastante divulgados: o relacional e XML. Com este artigo, pretende-se comparar estes modelos para o armazenamento de informação, tendo sido feito para isso um conjunto de testes cujo modelo e resultado são apresentados.

O artigo está organizado da seguinte forma. Inicialmente, é feita uma abordagem ao conceito de repositório, sendo exemplificadas áreas onde este é utilizado. Faz-se também uma análise a formas de armazenamento de informação que poderão ser utilizadas nos repositórios. A secção seguinte (a 3) apresenta ferramentas utilizadas para o teste de performance – *benchmarks*. De seguida, na secção 4 são definidos os cenários a considerar neste trabalho e na secção 5 é apresentado um possível modelo para a execução de testes. Na secção 6 é descrito um sistema no qual serão conduzidos os testes. Por fim, na secção 7 são apresentados resultados da execução dos testes e na secção 8 são delineadas as conclusões e o trabalho futuro.

2. Repositórios e armazenamento de informação

Repositório é um termo utilizado de várias formas em vários contextos. Um repositório pode ser visto de várias perspectivas como por exemplo espaço de armazenamento, simples ficheiros com informação, BDs ou sistemas complexos de gestão de informação. Como exemplos de áreas onde é utilizado o conceito de repositório temos o armazenamento digital de documentos, nomeadamente bibliotecas digitais, e o comércio electrónico. Actualmente, um sistema com bastante notoriedade na área das bibliotecas digitais é o DSpace [DSpace 2004]. É definido como sendo um repositório institucional *open source* que tem como objectivo armazenar e gerir material de investigação e educacional produzido por organizações ou instituições.

O conceito de repositório está também presente no comércio electrónico. Nesta área têm sido criados standards de comunicação baseados em XML [W3C 2004]. Um standard bastante divulgado nesta área é o ebXML [OASIS 2004]. No ebXML os repositórios armazenam vários tipos de informação como perfis de organizações (que são consultados pelas outras organizações com vista a encontrar um possível parceiro de negócio), definições de processos, componentes básicos, etc.

Em suma, o termo repositório pode ter várias utilizações em vários contextos. De uma forma geral, um repositório pode ser visto como um sistema onde a informação é armazenada. Poderá ser uma camada inferior onde assentam todas as outras que constituem as aplicações – camada de armazenamento – ou mesmo uma aplicação completa. Um aspecto importante é o facto de a informação não ser acedida directamente por utilizadores (humanos ou aplicações informáticas). O seu acesso e gestão são feitos exclusivamente através de uma camada superior, a qual abarca toda a lógica de acesso, atribuindo significado à informação armazenada. Desta forma, a abrangência do termo repositório pode variar, podendo ser considerado apenas como um depósito de dados ou conjuntamente com quaisquer outras camadas. Independente das partes constituintes de um repositório, algo que é comum a todas as abordagens de repositório é a sua capacidade básica de armazenar informação, o que implica a utilização de um modelo. Assumindo que isto é feito com recurso a BDs, existem várias opções: utilização de um modelo como o relacional, por objectos, ou outro, bem como a utilização de XML.

O modelo relacional, introduzido por Codd em 1970 [Codd 1970] é o modelo mais divulgado actualmente nos sistemas de gestão de BD. Com a massificação da utilização da Internet, o XML tem-se assumido como um standard com um papel preponderante na troca de informação. Esta tendência actual leva a que actualmente muitos fabricantes de software estejam a incorporar nos seus produtos suporte para XML, como é o caso das BDs. Esta nova forma de representação de informação estruturada apresenta novos desafios para o seu armazenamento exigindo conversões para a utilização com os modelos convencionais.

Existem dois modelos de documentos XML [Bourret 2005b]: centrado nos dados e centrado no documento. Os documentos do primeiro são caracterizados por possuírem uma estrutura regular e normalmente a ordem pela qual os vários elementos com o mesmo nível na estrutura são representados é irrelevante ao passo que os restantes caracterizam-se por possuir uma estrutura irregular e uma maior granulosidade na informação, isto é as unidades mais elementares de informação podem estar ao nível de elementos com conteúdo misto e a ordem com que os elementos são representados é normalmente relevante. Estes modelos influenciam a forma como o XML é armazenado de forma persistente em BD.

A tradução referida anteriormente poderá ser executada internamente na BD. A estas dá-se o nome de BDs com suporte para XML (XML-*enabled*) [Bourret 2005b]. Desta forma, estas são BDs convencionais que incluem internamente um mecanismo de tradução, dispensando assim a utilização componentes de software adicionais. Este tipo de abordagem está normalmente associado à utilização do modelo de documentos XML centrado nos dados. Actualmente, todos os grandes fabricantes de BDs incluem nos seus produtos suporte para o armazenamento e manipulação de XML.

Uma outra perspectiva para o armazenamento de XML em BDs é um novo tipo de BD, denominadas nativas XML. Neste tipo de BDs, a entrada, saída e armazenamento de informação é sempre no formato XML, não existindo qualquer tipo de conversões intermédias como nas soluções descritas anteriormente. Uma das grandes vantagens é o facto de os documentos serem armazenados e recuperados intactos, pelo que este tipo de BDs é apropriado para aplicações centradas nos documentos. No entanto, também são apropriadas para aplicações centradas nos dados, pois permitem a recuperação de fragmentos dos documentos, caso seja necessário. A utilização deste tipo de BDs considerada por alguns autores [Schmidt et al. 2001] como a melhor solução para o armazenamento e manipulação de documentos XML. A utilização deste tipo de BD permite a manipulação de qualquer tipo de documento mas de forma especialmente eficaz os documentos orientados ao documento.

Apesar de ambas manipularem XML, as BDs XML nativas e com suporte para XML apresentam diferenças [Bourret 2005b]:

- As nativas preservam a estrutura física dos documentos o que nem sempre é feito nas com suporte para XML.
- As nativas guardam qualquer ficheiro, mesmo sem conhecer o seu *schema*, o que pode não ser possível nas com suporte para XML.
- Nas nativas, a informação apenas é acessível em formato XML, ao passo nas com suporte para XML poderá existir outra forma de aceder à informação.

3. Benchmarks

Existe uma grande diversidade de sistemas que manipulam e armazenam informação, baseada em diversos modelos como o relacional, orientado por objectos ou XML, com desempenhos e características variados. Desta forma, a necessidade de comparar desempenhos surge de uma forma natural.

Um *benchmark* é um programa utilizado para testar a performance de software, hardware ou um sistema [Collin 2002]. Mais concretamente, um *benchmark* para BDs pode ser visto como um conjunto de instruções utilizadas para medir e comparar o desempenho de dois ou mais sistemas de gestão de base de dados. Isto é feito recorrendo à execução de experiências bem definidas cujas medidas de desempenho serão usadas para prever o desempenho do sistema [Seng et al. 2005]. Desta forma, na especificação de um *benchmark* são considerados 3 componentes principais [Menascé 2002]: o sistema a ser testado (SUT[1]), a carga de trabalho submetida ao SUT (workload), que consiste nas operações de teste, e uma ou mais métricas que são resultantes da monitorização e avaliação do desempenho do SUT o qual inclui a BD de teste (Figura 1). Exemplos de métricas são *throughput*, tempo de resposta, tamanho da BD e relação performance / custos de manutenção.

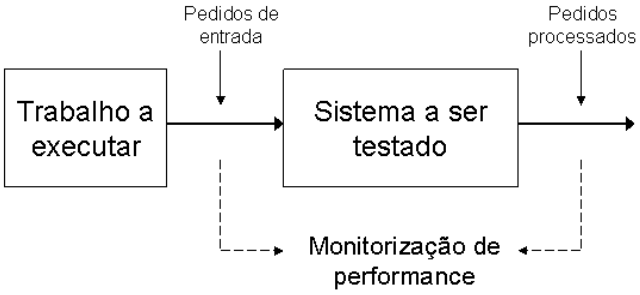


Figura 1 – Estrutura de um *benchmark* (adaptado de [Menascé 2002])

Os sistemas, que serão alvo de teste dos *benchmarks*, são construídos de forma a darem resposta a problemas específicos de determinados domínios de aplicação. A utilização de *benchmarks* genéricos, em sistemas de vários domínios de aplicação, não conduz a resultados fiáveis dada a natureza dos sistemas, já que uns estarão mais aptos para certo tipo de tarefas que outros. Desta forma em [Gray 1993] defende-se o uso de *benchmarks* específicos para os diversos domínios, os quais deverão respeitar os princípios básicos: relevância (deverá capturar as características do sistema a ser medido), portabilidade (deverá facilmente ser implementado em diferentes sistemas), escalabilidade (deverá ter a possibilidade de testar várias BDs em diferentes sistemas) e simplicidade (deverá ser perceptível de forma a ser credível).

Os *benchmarks* podem ainda classificar-se como sintéticos ou empíricos [Seng et al. 2005] ou uma mistura entre ambos. *Benchmarks* sintéticos emulam aplicações típicas de um determinado domínio, tanto a nível de operações como de BDs de teste, ao passo que os empíricos utilizam operações de testes e informação reais. Para sistemas baseados em modelos mais comuns como o relacional, foram desenvolvidos ao longo do tempo diversos testes de performance que estimularam a comparação de performance e consequentemente ao aperfeiçoamento dos sistemas. Com o aparecimento de sistemas capazes de lidar com XML, foi necessário criar testes que tenham em consideração novos desafios colocados por este modelo. Estes, em [Schmidt et al. 2002] são especificados como:

- Preservação de ordem textual das várias estruturas que compõem os documentos manipulados.
- Utilização de strings como tipo de dados básico cujo armazenamento e manipulação podem levantar problemas aos sistemas e podem entrar em conflito com a forma como os tipos de dados são tratados pelas linguagens de *query*.
- *Queries* que envolvem a manipulação de estruturas hierárquicas complexas e a preservação de ordem requerem a execução de operações dispendiosas principalmente quando o XML está armazenado numa estrutura relacional.
- Utilização de *Schemas* pouco rígidos de validação dos documentos. Isto não só não facilita a tarefa de escrita de *queries* complexos por parte dos utilizadores, como também pode levantar problemas a nível de optimização do armazenamento da informação, introduzindo valores NULL que podem fazer aumentar o tamanho da BD.

Existem vários *benchmarks* utilizados em BDs relacionais, tais como Wisconsin [DeWitt 1993], AS3AP [Turbyfill et al. 1989], mais orientados para o teste de aspectos importantes do servidor de BD, sendo constituídos por uma BD de teste e alguns *queries*.

Outros *benchmarks* importantes são os definidos pelo *Transaction Processing Performance Council* (TPC), um consórcio de vários fabricantes de *software* e *hardware* sem fins lucrativos. Este consórcio define *benchmarks* para vários fins, como processamento de transacções online (OLTP), comércio electrónico, apoio à decisão, etc. Os resultados da utilização destes *benchmarks* são publicados regularmente. Dois *benchmarks* definidos pelo TPC são o TPC-C [TPC 2005] e o TPC-W [TPC 2002]. Estes *benchmarks*, que como referem as suas especificações poderiam ser utilizados por qualquer SGBD, não só SGBD relacionais, têm grande aceitação na indústria.

Para BDs XML existem dois tipos de *benchmarks*: micro-*benchmarks* e aplicacionais. Todos eles consistem num conjunto de informação (BD de teste) que pode ter várias versões com tamanhos diferentes e sobre o qual serão executados um conjunto predefinido de *queries*.

Os micro-*benchmarks*, como o Michigan Benchmark [Runapongsa et al. 2003], são desenhados de forma a testarem componentes específicos do sistema com vista a isolar e ajudar a corrigir

[1] Da língua inglesa: *System under test*.

certos problemas. Pretendem explorar o impacto na performance do sistema das características mais importantes do XML, dispondo de uma BD de teste heterogénea, não inspirado numa qualquer aplicação real, sobre o qual são especificados *queries* especialmente desenhados para testar componentes elementares da linguagem de *query* (como selecção, joins, etc). Com um *benchmark* deste tipo torna-se possível aperfeiçoar as operações mais básicas ao nível do SGBD.

Os *benchmarks* aplicacionais, por seu lado funcionam a um nível mais elevado, pretendendo medir a performance do sistema como um todo e não questões específicas. Cada um deles dispõe de uma BD de teste, que pode ser ou não inspirada numa aplicação real, sobre a qual são definidos *queries* que pretendem abranger o maior número possível de características da linguagem de *query*. Exemplos destes *benchmarks* são o XOO7 [Li et al. 2001], Xmark [Schmidt et al. 2002], XBench [Yao et al. 2004] e o XMach-1 [Böhme e Rahm 2001].

4. Cenários

A informação poderá estar modelada segundo o modelo relacional, devidamente inserida em registos de tabelas numa BD. No entanto, poder-se-á optar por a representar não utilizando o modelo relacional mas XML.

A informação representada em XML poderá ser encarada de duas formas distintas, tal como referido anteriormente: centrado nos dados ou centrado nos documentos. Apesar de não existirem regras rígidas, cada uma destas hipóteses poderá estar mais adequada para utilização para diferentes tipos de BDs. Desta forma, uma visão centrada nos dados poderá ser mais adequada para utilização numa BD com suporte para XML (XML *enabled*) ao passo que uma visão centrada no documento poderá ter melhor desempenho numa BD XML nativa [Vakali et al. 2005].

No contexto do presente trabalho, terá interesse estudar cenários onde sejam utilizadas para o armazenamento físico de informação do repositório de suporte a um serviço de informação, BDs relacional (para a representação da informação no modelo relacional), XML *enabled* ou XML nativa (para suportar informação representada em XML). O estudo destes cenários deverá ter como base uma grande quantidade de informação de forma a atribuir o maior realismo possível ao estudo a efectuar. Desta forma, a quantidade de informação será outro aspecto com interesse no estudo, além do tipo de BD, podendo ser considerados cenários com quantidades distintas de informação.

Este trabalho centra-se numa perspectiva de divulgação de informação (isto é, consultas) deixando-se de parte os aspectos de recolha de informação. Desta forma, tomou-se como ponto de partida uma BD (relacional) real com informação respeitante a currículos de investigadores, cuja informação foi posteriormente convertida para XML e inserida nas respectivas BDs.

5. Modelo de testes

Para tirar conclusões sobre os modelos mais adequados para a utilização em repositórios torna-se necessária a execução de testes, isto é de *benchmarking*. Torna-se necessário conceber e implementar um sistema, executar testes e interpretar os resultados.

Trabalhos como [Böhme e Rahm 2001; Li et al. 2001; Schmidt et al. 2002; Yao et al. 2004] abordam questões relacionadas com o processamento de XML nomeadamente o processamento de *queries* que pretendem abranger o maior número possível de características da linguagem de *query*, pretendendo medir a performance do sistema como um todo, existindo um foco exclusivo em BDs de XML ao passo que outros como [Turbyfill et al. 1989; DeWitt 1993] abordam questões de processamento de *queries* em sistemas relacionais. Na presente situação o objectivo não é o mesmo que o dos trabalhos referidos uma vez que não se está particularmente

interessado em explorar questões de processamento de XML ou características das linguagens de *query* ou ainda a capacidade de processamento destes pelos sistemas. A atenção foca-se antes na escolha de um modelo, XML ou relacional, que melhor resposta às necessidades de um serviço de informação online para divulgação de currículos de investigadores. Desta forma, para os testes de performance, parte-se deste cenário real previamente definido e irá tentar-se definir os aspectos principais a testar. É uma abordagem diferente de outros trabalhos (por exemplo [Böhme e Rahm 2001; Schmidt et al. 2002; TPC 2002; TPC 2005] dado que não é feito com vista a simular uma dada aplicação de um determinado domínio com o objectivo de abordar o maior número de aspectos relevantes possível, mas é feito com os condicionalismos inerentes à aplicação definida pelo contexto. O maior destes condicionalismos é, como referido, o facto de se estar a considerar dois modelos distintos, XML e relacional.

Com vista à realização de testes é necessário criar pesquisas que se pretendem o mais simples possível. Cada pesquisa tem 3 versões, uma para cada tipo de BD envolvida, na respectiva linguagem de *query* utilizada. Uma questão importante na criação destas pesquisas é o facto de os modelos a testar serem bastante diferentes. Aspectos como a recuperação de documentos na sua forma original, travessias em documentos estruturados ou a composição de novos documentos com base na informação da BD, apenas fazem sentido quando se trata de XML, pelo que as pesquisas terão de abordar aspectos comuns, como pesquisas em texto integral, manipulação de resultados de grandes dimensões ou a utilização de *joins*.

Ao invés de tentar criar pesquisas que abrangessem uma lista completa de aspectos comuns aos modelos em causa, optou-se por definir um conjunto de requisitos para as pesquisas, as quais ao serem implementadas poderão incluir obrigatoriamente pelo menos uma das características referidas.

Desta forma, podemos considerar que um serviço de informação *online* com um perfil centrado na consulta de informação deverá permitir as seguintes funcionalidades básicas:

1. Acesso directo aos itens por chave (pesquisa por chave)
2. Obtenção de resultados ordenados segundo determinados critérios.
3. Obtenção de indicadores com base na informação armazenada.
4. Suporte eficiente para resultados de grandes dimensões (elevado número de itens).
5. Pesquisas textuais

Com base nestas características básicas foram construídas as pesquisas.

A execução de testes pressupõe a medição de parâmetros relevantes. Um parâmetro importante, e que será medido na execução dos testes, é o *thoughput* do sistema que se pode definir como o número de pesquisas, baseadas nas características propostas, para o qual se obteve uma resposta por segundo.

6. Sistema de testes

Para a execução de testes torna-se necessária a criação de um sistema que tenha em atenção as funcionalidades básicas descritas. A sua arquitectura (Figura 2) é inspirada nas propostas em [Böhme e Rahm 2001; TPC 2002] sendo igualmente baseada numa aplicação *web*. É composta por três componentes: BD, servidor aplicacional e clientes, não existindo a necessidade do componentes adicionais externos aos SUT como o que faz recolha e upload de informação para a BD (loader) [Böhme e Rahm 2001] ou gateway de pagamentos [TPC 2002].

Na presente situação o sistema irá utilizar, em testes distintos, BDs para a manipulação de XML (*enabled* e nativa) e relacional, todas elas contendo informação equivalente representada nos respectivos modelos.

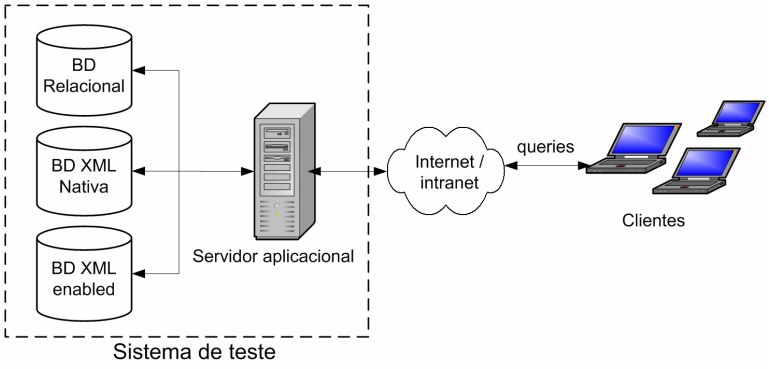


Figura 2 – Arquitectura do sistema de teste

No servidor aplicacional (que tal como o servidor de BD poderia ser na realidade constituído por vários servidores) é executada uma aplicação *web* com base num servidor *web* (http) bem como outros componentes de software necessários para acesso às BDs. O conjunto do servidor aplicacional e de BD constitui o sistema de teste, excluindo os clientes pelo que não se considera nos resultados dos testes atrasos devidos a comunicações e a processamentos feitos a nível dos clientes. Desta forma, os clientes conectam-se ao sistema de teste via o protocolo http e todos os resultados são medidos a nível do servidor aplicacional tendo apenas em conta o intervalo de tempo entre a chegada do pedido e o final do seu processamento (ignorando o tempo que o resultado demoraria a chegar ao cliente). Para evitar atrasos relativos à rede, os servidores aplicacional e de BDs serão os mesmos, com dados armazenados localmente. Com vista a evitar inconsistências nos resultados, não é permitida *cache* ao nível do servidor. Para simular diferentes cenários são utilizadas diferentes BDs com diferente número de registos, bem com é alterado o número de clientes e consequentemente o número de pedidos ao servidor aplicacional. Os valores vão de 25 a 125 em intervalos de 25, tendo sido criada uma aplicação que simula os pedidos simultâneos à aplicação. Ao *schema* da BD original foi feita uma simplificação, mas mantendo a estrutura básica de um currículo. O XML *Schema* foi criado com base neste *schema* relacional simplificado. Apesar de terem sido utilizadas ferramentas automáticas para esta tarefa, a versão final teve uma revisão manual, tendo sido adoptada uma abordagem baseada em atributos para representar as colunas das tabelas e usando os nomes das colunas no nome dos atributos [Williams et al. 2000]. Após a criação do XML *Schema* foi feita a conversão da informação relacional para o formato XML, respeitando este *schema*. Esta conversão foi feita inteiramente com recurso a ferramentas automáticas.

Para responder aos vários cenários idealizados para os testes é necessário que o tamanho da BD de teste seja constituído por um número de registos variável. Desta forma a BD de teste inicial foi completada utilizando ferramentas especialmente desenvolvidas. Os valores são 115.000, 517.000 e 1 milhão de registos.

7. Resultados dos testes

Facilmente se observa nos gráficos da Figura 3, que a BD relacional apresenta um comportamento quase sempre superior nesta carga a que foi submetida. Isto apenas não acontece na BD de maiores dimensões com o menor número de clientes por uma margem pequena relativamente às outras BDs, pelo que se poderá considerar pouco significativo. Os testes realizados revelam assim que existem notoriamente 2 tipos de desempenho: alto para a BD relacional e baixo para as BDs de XML.

No caso das BDs pequena e média, a diferença de desempenho entre a BD relacional e as de XML é significativa chegando a ser mais de 6 vezes superior no caso da BD média. Isto é mais relevante com o aumento de clientes. Nota-se claramente dois grupos de desempenho: alto para a relacional e baixo para as BD de XML. Entre as BDs de XML, nota-se um desempenho que se

pode considerar idêntico no caso das BDs médias. No caso da BD pequena a com suporte para XML leva vantagem chegando a ter um desempenho cerca de 50% superior à nativa XML.

Na BD grande esta diferença de desempenho relativa entre os vários tipos de BDs não é tão acentuada, estando todos os valores relativamente próximos, mas novamente com a relacional em vantagem. Nota-se no entanto uma queda de desempenho bastante acentuada relativamente às BDs média e pequena.

A análise de variância dos resultados revela que os factores que mais influenciaram os resultados foram o tamanho da BD (33,52%) e o tipo (31,97%), ao passo que o número de clientes a influência foi menor (5,3%). Um outro aspecto que é observável nos valores obtidos é uma quebra acentuada de desempenho na BD de maiores dimensões, para todos os tipos. Mesmo assim a relacional está em vantagem. Esta quebra poderá estar relacionada com configurações ao nível do servidor onde foram executados os testes. No entanto, não foi possível executar testes que permitissem comprovar esta hipótese, pelo que estudos adicionais poderão de futuro clarificar este aspecto.

8. Conclusão e trabalho futuro

Neste documento foi introduzido um conceito de repositório bem como diferentes abordagens ao armazenamento de informação. Foi também proposto um modelo e um sistema de testes com os quais se pretende averiguar qual o modelo mais adequado para o armazenamento de informação nos repositórios. Para isto torna-se necessário a execução de testes, isto é, *benchmarking*. Estes testes terão de respeitar duas condições fundamentais para que possam ser credíveis. Deverão ter como base cenários de aplicação reais e não deverão abordar aspectos que não façam sentido em ambos os modelos considerados.

Os resultados obtidos nestes testes demonstram uma superioridade inegável da BD relacional. A escalabilidade de uma solução baseada no modelo relacional é bastante superior comparativamente com uma solução baseada em XML. Isto tem duas faces. Por um lado, as BDs relacionais apresentam décadas de desenvolvimento ao passo que a tecnologia XML é bastante mais recente. Desta forma, as BDs XML ainda não apresentam uma maturidade suficiente para possam ser colocadas ao nível das relacionais e para que o modelo XML possa ser uma alternativa ao modelo relacional. Por outro lado, a informação utilizada neste trabalho inicialmente estava numa BD relacional, tendo sido convertida para XML. Apresenta um perfil que se enquadra perfeitamente no modelo relacional, isto é, representação em tabelas e onde não existe uma noção de hierarquia. Estas situações são portanto mais favoráveis à utilização do modelo relacional [Lapis 2005], o que terá penalizado a BD de XML. É importante ter em atenção que existem situações onde uma BD de XML é mais apropriada, como gestão de informação orientada a documentos, integração de informação (integração entre aplicações, com informação obtida de diversas fontes), gestão de informação não estruturada ou evolução do *schema* [Bourret 2005a]. Desta forma, seria bastante interessante repetir os testes deste trabalho mas com um tipo de informação mais apropriada ao modelo XML (com características como a possibilidade de alterações ao seu *schema* ou com hierarquias). Uma solução que adopte ambos os modelos aproveita o melhor de cada um deles: a rapidez associada ao modelo relacional (como demonstrado nos testes efectuados) e a capacidade de manipulação de documentos e lidar com *schemas* complexos ou que mudam frequentemente. Neste caso, parte da informação está armazenada em tabelas relacionais e outra parte em documentos XML, por exemplo a informação não estruturada cujo *schema* é pouco rígido e que poderá mudar mais frequentemente. Nesta perspectiva, o uso dos modelos relacional ou XML não pode ser visto numa perspectiva competitiva, mas sim complementar.

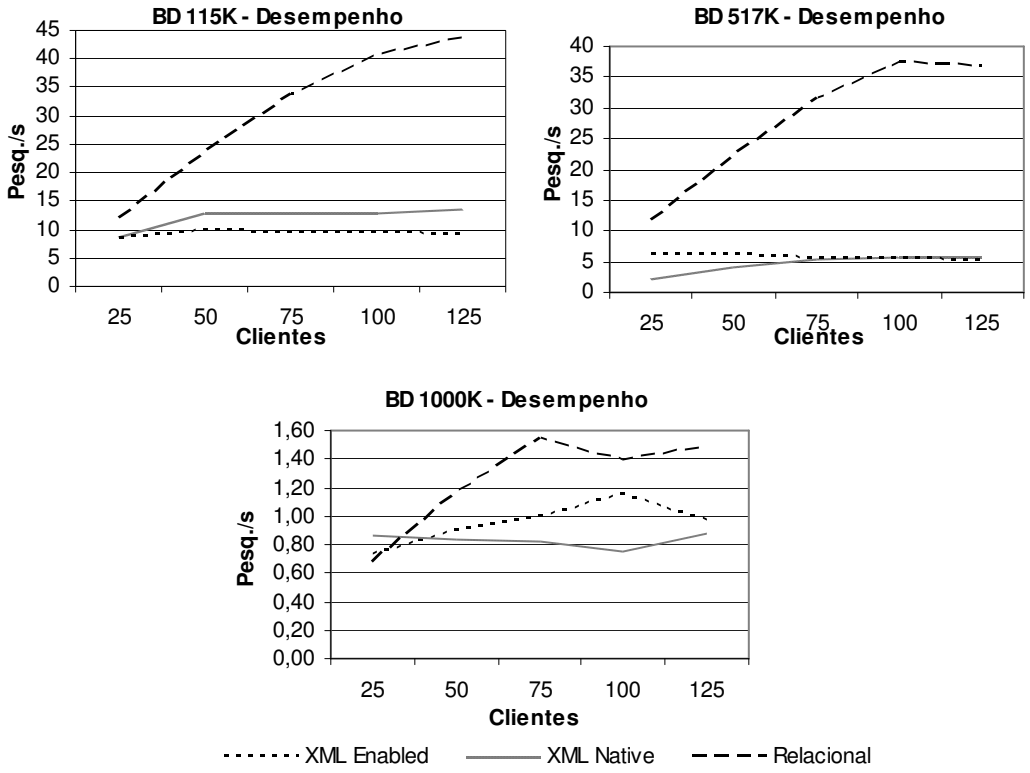


Figura 3 – Resultados dos testes

9. Referências

Böhme, T. e E. Rahm (2001). XMach-1: A Benchmark for XML Data Management. Datenbanksysteme in Büro, Technik und Wissenschaft (BTW), Oldenburg, Germany, Springer-Verlag.

Bourret, R. (2005a). Native XML Databases in the Real World. XML 2005 Conference & Exposition, Atlanta, USA.

Bourret, R., XML and Databases, http://www.rpbourret.com/xml/XMLAndDatabases.htm, (15/01/2006), 2005b

Codd, E. F. (1970). "A relational model of data for large shared data banks." Communications of the ACM 13(6): 377 - 387.

Collin, S. M. H. (2002). Dictionary of Information Technology, Third Edition, Peter Collin Publishing.

DeWitt, D. J. (1993). The Wisconsin Benchmark: Past, Present, and Future. The Benchmark Handbook. J. Gray (Eds.), Morgan Kaufmann Publishers, Inc.

DSpace, DSpace Federation, http://www.dspace.org/, (01/02/2004), 2004

Gray, J. (1993). Database and Transaction Processing Performance Handbook. The Benchmark Handbook, Morgan Kaufmann Publishers, Inc.

Lapis, G. (2005). XML and Relational Storage - Are they mutually exclusive? XTech 2005, Amsterdam, The Netherlands.

Li, Y. G., S. Bressan, G. Dobbie, Z. Lacroix, M. L. Lee, U. Nambiar e B. Wadhwa (2001). XOO7: Applying OO7 Benchmark to XML Query Processing Tools. Conference on Information and Knowledge Management, Atlanta, Georgia, USA, ACM Press, New York, NY, USA.

Menascé, D. A. (2002). TPC-W: A Benchmark for E-Commerce. IEEE Internet Computing. 6**:** 83 - 87.

OASIS, ebXML, http://www.ebxml.org/, (01/02/2004), 2004

Runapongsa, K., J. M. Patel, H. V. Jagadish, Y. Chen e S. Al-Khalifa (2003). The Michigan Benchmark: Towards XML Query Performance Diagnostics. 29th VLDB Conference, Berlin, Germany.

Schmidt, A., F. Waas, M. Kersten, M. J. Carey, I. Manolescu e R. Busse (2002). XMark: A benchmark for XML Data Management. 28th VLDB Conference, Hong Kong, China.

Schmidt, A., F. Waas, M. Kersten, D. Florescu, M. J. Carey, I. Manolescu e R. Busse (2001). "Why and how to benchmark XML databases." ACM SIGMOD Record 30(3): 27 - 32.

Seng, J.-L., S. B. Yao e A. R. Hevner (2005). "Requirements-driven database systems benchmark method." Decision Suppot Systems 38(4): 629 - 648.

TPC (2002) Transaction Processing Performance Council - TPC Benchmark W, ver. 1.8,

TPC (2005) Transaction Processing Performance Council - TPC Benchmark C, ver. 5.4,

Turbyfill, C., C. Orji e D. Bitton (1989). AS3AP:a comparative relational database benchmark. 34th IEEE Computer Society International Conference, San Francisco, CA.

Vakali, A., B. Catania e A. Maddalena (2005). "XML Data Stores: Emerging Practices." Internet Computing, IEEE 9(2): 62-69.

W3C, Extensible Markup Language (XML) 1.0 (Third Edition), http://www.w3.org/TR/REC-xml, (13/02/2005), 2004

Williams, K., P. Dengler, J. Gabriel, A. Hoskinson, M. Kay, M. Brundage, T. Maxwell, M. Ochoa, J. Papa e M. Vanmane (2000). Professional XML Databases, Wrox Press Ltd.

Yao, B. B., M. T. Özsu e N. Khandelwal (2004). XBench Benchmark and Performance Testing of XML DBMSs. 20th International Conference on Data Engineering, Boston, Massachusetts, U.S.A., IEEE Computer Society.

Validação de Sistemas Computorizados na Indústria Farmacêutica: O caso BIAL

Rui Rodrigues[1], Paulo Marques[1], João Nuno Oliveira[2]

1) BIAL, Trofa, Portugal

rui.rodrigues@bial.com paulo.marques@bial.com

2) Dept. Sistemas de Informação, Universidade do Minho, Guimarães, Portugal

jno@dsi.uminho.pt

Resumo

A indústria farmacêutica vive num contexto fortemente regulado, onde entidades como a FDA (Food and Drug Administration), nos Estados Unidos, e a EMEA (European Medicines Agency), determinam um conjunto de boas práticas, sejam elas boas práticas de laboratório, clínicas ou fabrico, a seguir pelas as empresas que operam neste sector, que visam assegurar os elevadíssimos padrões de qualidade exigíveis a qualquer medicamento vendido nesses países.

No momento em que o Grupo BIAL se prepara para submeter aos mercados de todo o mundo dois novos medicamentos, totalmente desenvolvidos pelo Grupo, em Portugal, o cumprimento das referidas boas práticas é imperioso para o processo de acreditação da empresa nos diversos mercados (sobretudo os já referidos, que são os mais exigentes em termos regulamentares).

Em todas as boas práticas que regulam a indústria farmacêutica é constante a exigência de validação de sistemas computorizados, que consistem num conjunto de actividades destinadas a comprovar, de forma devidamente documentada, que os sistemas computorizados cumprem os seus requisitos, de modo a suportar processos de trabalho regulados. No entanto, não é unânime definir que sistemas devem ser validados e qual a extensão dessa validação. No contexto de BIAL, as directivas emanadas pela FDA foram seguidas, sendo estas consideradas como as mais restritivas do mundo.

Este artigo procurará apresentar a estratégia de validação adoptada em BIAL, através da descrição da sua framework de validação de sistemas computorizados. Uma vez que este processo de validações já se encontra em curso de forma mais alargada e sistemática, há mais de um ano, é possível apresentar e discutir indicadores e resultados do esforço desenvolvido. Será feita uma análise do impacto desta actividade ao nível organizacional e ao nível da função específica de gestão de sistemas de informação.

Palavras chave: Validação de sistemas computorizados, garantia de qualidade de sistemas, testes de sistemas

1 A validação de sistemas computorizados

A indústria farmacêutica vive num contexto fortemente regulado, onde entidades como a FDA (Food and Drug Administration), nos Estados Unidos, e a EMEA (European Medicines Agency), determinam um conjunto de boas práticas, sejam elas boas práticas de laboratório,

clínicas ou fabrico, a seguir pelas as empresas que operam neste sector, que visam assegurar os elevadíssimos padrões de qualidade exigíveis a qualquer medicamento vendido nesses países.

O processo de demonstração de adesão às boas práticas regulamentadas é conhecido como validação e implica "estabelecer a evidência documental que demonstra com um elevado grau de garantia que um processo específico produzirá consistentemente um produto nas suas especificações predeterminadas e atributos da qualidade" [FDA 1987] e "demonstrar que um sistema computorizado é adequado para a sua utilização pretendida" [OECD 1995].

As grandes questões que uma organização enfrenta, ao aplicar um processo de validação aos seus sistemas, são: Definir que sistemas deve validar, e para estes, qual a dimensão do processo de validação. Não há divisões claras entre sistemas a validar e sistemas a não validar, e muito menos certezas no esforço de validação que deve ser dispendido, como pode ser constatado em [Wingate 2003] e [McDowall 2005]. Um esforço diminuto de validação pode implicar uma situação de incumprimento de boas práticas, com prejuízos avultados para a empresa, ao passo que um esforço exagerado aumenta em muito o seu custo e diminui o valor acrescentado. A análise de risco assume aqui um papel importante, pois permite dotar a organização de uma ferramenta para aferir a resposta a essas duas questões, e essa importância é reconhecida em vários referenciais de boas práticas [McDowall 2005].

Apesar de não existirem muitos estudos publicados que apresentem relações de custo/benefício dos processos de validação, o retorno do investimento advirá de custos inferiores de manutenção dos sistemas e da maior produtividade do processo que o sistema suporta [Wingate 2003]. No entanto, [Farkas 2006] aponta para o atraso na decisão de implementar novos sistemas na indústria farmacêutica quando comparado com outros sectores, algo que é atribuído à forte regulação.

É importante que as empresas reconheçam que a validação de sistemas é um processo contínuo, e as necessidades neste campo no futuro serão ainda maiores, na medida em que as organizações necessitarão de mais sistemas, mais rápidos e mais abrangentes, e as entidades reguladoras terão imposições mais sofisticadas [Budihandojo et al. 2001].

2 O Problema

O Grupo Bial desenvolve, há 80 anos, soluções terapêuticas na área da Saúde, com um espírito de qualidade e inovação, trabalhando com a máxima eficiência no sentido de satisfazer plenamente todos os seus clientes e antecipando as necessidades de um mercado cada vez mais exigente e competitivo.

Bial assume a I&D e a internacionalização como áreas estratégicas para a expansão e autonomia do Grupo e daí a aposta no desenvolvimento de novos medicamentos através dos mais exigentes processos científicos e no desenvolvimento de um número crescente de mercados com vista a uma presença global.

Embora a qualidade e rigor tenham sido sempre apanágio de Bial, a abordagem a mercados regulados pela FDA e a o registo centralizado na EMEA colocam novos desafios a este nível, colocando Bial ao nível de exigência das maiores multinacionais farmacêuticas.

É neste contexto que a validação dos sistemas computorizados assume uma importância crescente, considerando que se trata de uma componente determinante nos processos de investigação, produção e qualidade.

Porém, numa organização, em particular num ambiente competitivo, os recursos são limitados e são normalmente absorvidos pelas actividades correntes e pelos processos de melhoria e crescimento, sendo particularmente difícil disponibilizar os meios para fazer face a tarefas tão complexas e exigentes como as de validação.

Por outro lado, sendo os benefícios deste processo claros para a organização e seus principais responsáveis, nem sempre é linear a percepção a todos os níveis da organização, pelo que era fundamental uma sensibilização generalizada, à semelhança de qualquer outro processo de qualidade.

Perante este cenário, Bial identificou a necessidade de estabelecer uma política de validação de sistemas computorizados, com o objectivo de garantir o máximo de qualidade e desempenho dos sistemas classificados como GxP[1] na respectiva análise de risco, i.e., sistemas que suportem directamente actividades ou processos no âmbito das áreas regulamentares regidas por boas práticas de laboratório, clínicas, de fabrico, ou de distribuição.

Teria, no entanto, de ser identificada a forma adequada de desdobrar a política num conjunto de procedimentos que garantisse, por um lado, a sua correcta implementação através da definição de um conjunto de linhas estratégicas comuns, e por outro, proporcionar um enquadramento adequado na metodologia de planeamento estratégico do Grupo e desta forma possibilitar a correcta programação de actividades e recursos para validação sem colidir, e se possível complementando, as restantes actividades de negócio.

[1] GxP – Good X Practice (FDA *compliance*; "x" pode significar: *Clinical, Laboratory, Manufacturing, Distribution*, etc.)

Era também necessário criar um modelo de trabalho que proporcionasse os modelos e guias para correcta implementação da estratégia estabelecida e que constituísse a ferramenta de trabalho a adoptar por todos os órgãos do Grupo no esforço de validação.

3 A abordagem

A definição normalmente aceite para "sistema computorizado" é: Um conjunto de equipamentos, software, pessoas e procedimentos, que em conjunto viabilizam a execução de um determinado processo de negócio [Stokes 1998]. Bial adoptou esta definição de sistema computorizado na sua metodologia de validação.

Os sistemas adquiridos em Bial, quando classificados como GxP, são objecto de um longo e exaustivo processo de validação e todos os sistemas computorizados de Bial são inventariados anualmente, aquando da elaboração do plano e orçamento anuais, sendo efectuada uma análise de risco que determina a revisão da sua classificação GxP.

Com base na lista de sistemas existentes ou a adquirir, e respectiva classificação, é estabelecido um plano de garantia de qualidade de sistemas onde se programam as actividades de validação e os recursos associados. Desta forma há um perfeito enquadramento de todas as actividades de planeamento, sejam de validação de sistemas ou de processos de negócio, possibilitando racionalização de recursos, identificação de pontos críticos e maiores sinergias operacionais.

A abordagem de validação de sistemas computorizados, em Bial, tem por base três procedimentos, que definem os conceitos e a metodologia a aplicar, para que se possa considerar qualquer sistema computorizado como validado.

O procedimento geral aprovado, intitulado "Validação de Sistemas Computorizados" [Bial DSIPG002 2006], descreve o procedimento standard a seguir para validação de um sistema computorizado identificado como crítico no âmbito regulamentar (GxP) em Bial, com base numa avaliação de risco efectuada ao sistema em causa. Descreve ainda o conteúdo de um dossier de validação elaborado em conformidade com a política de validações de Bial e com os aspectos regulamentares aplicáveis, nomeadamente os que dizem respeito exclusivamente a registos electrónicos e assinaturas electrónicas [FDA 1997] e sistemas computorizados [EMEA 1998].

A validação de um sistema é indissociável do seu ciclo de vida [EMEA 1998]. Assim, no momento em que é detectada a necessidade de existir um sistema computorizado para apoiar ou suportar um ou mais processos regulamentados, também se deve imediatamente avaliar o risco, que decorre da utilização desse sistema, para a segurança, eficácia e qualidade (SEQ) do

produto final (medicamento). Essa avaliação deve ser refinada ao longo da selecção do sistema, podendo ser efectuada por grupos de requisitos, ou em cada processo suportado.

Várias condicionantes determinam a extensão da validação de um sistema. Destacamos as três seguintes: Análise de risco [Stokes 1998]; Categoria do software e do hardware [ISPE 2001]; Fornecedor seleccionado e grau de cumprimento com normas de qualidade, internacionalmente aceites, relativas a desenvolvimento, teste e manutenção de software (ex.: ISO 9126).

A validação de um sistema computorizado deve iniciar-se com a elaboração e aprovação de um plano de validação, em que é constituída a equipa de validação do sistema. Esta equipa deve conter pelo menos as seguintes 5 competências obrigatórias: Responsável funcional; Líder de equipa; Gestor de documentos; Coordenador de testes; Garantia de qualidade. A validação de um sistema só deve terminar quando o sistema for definitivamente descontinuado, pelo que a equipa de validação tem que desempenhar a sua função até essa altura.

O plano de validação é o documento que estabelece a extensão da validação, as actividades a executar e os critérios a verificar para que o sistema se possa considerar validado. Entre essas actividades, consta sempre a elaboração de documentação que permita acompanhar todas as fases do ciclo de vida do sistema.

Note-se que Bial não dispõe das competências adequadas para desenvolvimento de software, pelo que o ciclo de vida de um sistema pressupõe sempre a existência do fornecedor, estando claramente definidas as actividades que a este compete executar, em parceria com Bial.

Para além do plano de validação, a documentação mais importante que é desenvolvida durante a validação de um sistema computorizado é a seguinte:

- Especificação de requisitos de utilizador (ou URS, de User Requirements Specification): Define os objectivos a atingir com o sistema, do ponto de vista do utilizador;
- Especificação funcional (ou FRS, de Functional Requirements Specification): Define o modo de funcionamento do sistema, para atingir cada um dos requisitos de utilizador;
- Especificação de desenho (ou DRS, de Design Requirements Specification): Determina o enquadramento técnico necessário à execução das funcionalidades especificadas;
- Plano de testes: Define a estratégia de teste, relacionando os requisitos especificados com os testes a executar e pode ser dividida em três grupos (designados como qualificações), dirigidos a cada uma das especificações anteriores, conforme a figura seguinte [ISPE 2001]:

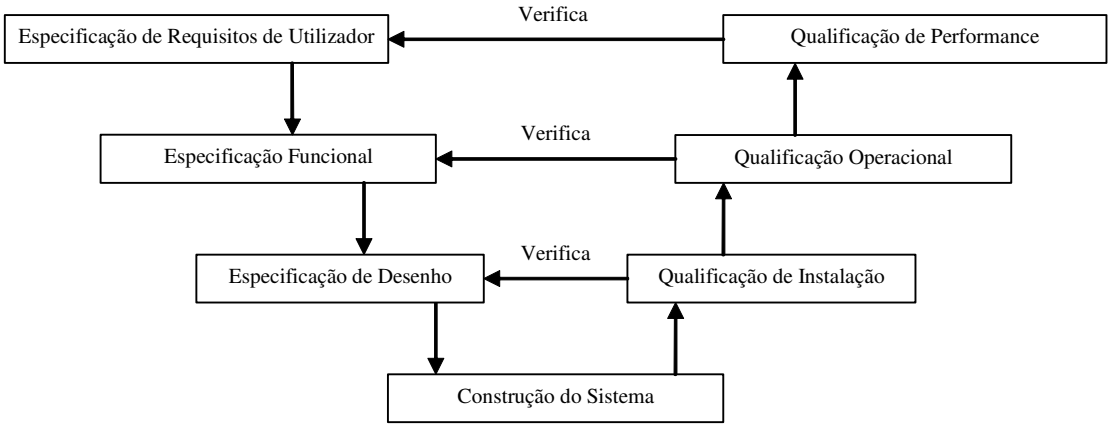


Figura 1 – Modelo de Validação em V

- Relatório de testes: Avalia as actividades de teste de um sistema computorizado face ao respectivo plano de testes e justifica os desvios verificados;
- Relatório de validação: Descreve as actividades realizadas, os resultados observados das tarefas planeadas, os desvios que se verificaram ao plano de validação inicial e apresenta a decisão sobre o estado de validação do sistema, em função dos critérios estabelecidos no plano de validação;
- Validação simplificada: Para o caso de sistemas classificados como simplificados, estes podem ser validados recorrendo a um impresso próprio, onde é possível registar as principais actividades de validação. Nesse caso, nenhum dos documentos anteriormente descritos é obrigatório.

Para além do procedimento geral referido, existem algumas instruções específicas que determinam procedimentos a seguir no âmbito das validações de sistemas computorizados. Uma dessas instruções é "Regras de Testes Formais a Sistemas Computorizados" [Bial DSIIE008 2006], que descreve as práticas standard a serem usadas pelas equipas de validação.

A instrução anterior define também as regras de desenvolvimento dos guiões de teste previstos no plano de testes. Devem ser nomeados alguns elementos da equipa de validação para executar todas as tarefas ligadas a testes (planificação, elaboração, aprovação, execução, testemunha e relatório). Esta equipa de testes é liderada por um coordenador e de resto, dela fazem parte os desenhadores e aprovadores dos guiões, o fornecedor do sistema, os testadores e as testemunhas de execução de testes.

Os guiões de teste devem ser preparados com sentido crítico sobre a verificação de cumprimento dos requisitos especificados. Devem incluir descrições detalhadas sobre os procedimentos necessários à execução do teste, os resultados esperados em cada passo e espaço

para registar os resultados obtidos. Os guiões impressos e colocados dentro de um envelope etiquetado constituem o envelope de teste, que é levado junto do testador, que o executa. A testemunha deve acompanhar essa execução, verificando se a mesma cumpre a instrução específica referida e se essa execução é documentada da melhor forma possível (no mínimo, com o material de suporte previsto no guião).

Por outro lado, a instrução específica "Gestão de Alterações a Sistemas Computorizados" [Bial DSIIE010 2006] tem por objectivo descrever o procedimento a seguir para que um sistema computorizado, após ser declarado validado, se mantenha nesse estado, durante todo o seu ciclo de vida.

A ocorrência de alterações ao sistema, quer sejam só de configuração, ou mesmo funcionais, devem ser analisadas pela equipa de validação e avaliado o seu impacto no sistema. Os tipos de alterações definidos e previstos na instrução são: Significativa, Moderada, Mínima e Rotina. Consoante o tipo de cada alteração, a abordagem necessária para manter o sistema validado é prevista nesta instrução, sendo tanto mais extensa quanto mais significativa for a alteração.

4 Resultados

Uma vez que este processo de validações já se encontra em curso de forma mais alargada e sistemática, em Bial, há mais de um ano, seguindo a abordagem anteriormente descrita, já existem resultados mensuráveis, que se resumem em 2 sistemas validados dentro dos prazos internamente previstos e vários outros com validação em curso.

Algumas validações aconteceram de forma retroactiva, isto é, com os sistemas já em funcionamento normal. Nesses casos, a respectiva validação destina-se a comprovar que o sistema sempre cumpriu os requisitos de utilizador, que presidiram à sua selecção, implementação e manutenção, isto é, durante todo o seu ciclo de vida.

Para dar uma ideia mais clara da dimensão dos processos de validação, apresentamos na tabela 1 alguns indicadores relativos aos dois sistemas já validados e a outro em fase adiantada de validação: WMS (Warehouse Management System), MES (Manufacturing Execution System) e ERP (Enterprise Resource Planning) utilizado na empresa.

De resto, alguns sistemas foram classificados como simplificados e a respectiva validação já está em curso. Nestes sistemas, estão sobretudo incluídos os casos de equipamentos de laboratório ou de controlo de qualidade, ou outros que são adquiridos com documentação de qualificação, ou de validação, preparada pelo fornecedor.

Indicador \ Sistema	**WMS**	**MES**	**ERP [2]**
Nº de requisitos de utilizador (de âmbito GXP)	189	129	283
Nº de envelopes de teste elaborados e executados	80	78	94
Nº de procedimentos/iterações de teste	129	184	281
Nº total de pessoas da equipa de validação	22	16	24
Nº de pessoas da equipa de testes (incluídas no anterior)	16	9	17
Duração da tarefa de execução dos testes (em meses)	1	1,5	2
Duração total do projecto de validação (em meses)	3	8	10

Tabela 1- Indicadores do processo de validação

Foi visível que, sendo o modelo implementado tal como previsto, foi possível planear correctamente as actividades de validação e, porque foram alocados alguns recursos específicos, cumprir com os objectivos.

5 Impacto da validação de sistemas na organização

De uma forma geral, é visível o impacto na organização na medida em que é mensurável, através dos planos anuais de garantia de qualidade de sistemas, a evolução dos processos de validação em todas as áreas da empresa e da análise individual dos sistemas onde se detectam oportunidades de melhoria, que de outra forma, passariam despercebidas.

A resposta a auditorias, não só no âmbito GxP, como também ISO 9000 ou financeiras, é fortemente melhorada.

Naturalmente, ao nível das actividades operacionais, é visível nos órgãos envolvidos o acréscimo de tarefas, pois a validação de um sistema, para além de um esforço inicial, é também uma actividade em contínuo, que exige uma preocupação constante de documentação, testes, gestão de alterações, etc.

No entanto, é também notório que, com o acumular de experiência e com a normalização de procedimentos adequados a cada tipo de sistema, o esforço tende a reduzir-se e a constituir parte

[2] Indicadores do ERP são estimativa

natural da actividade diária, o que, associado aos benefícios de qualidade obtidos e à confiança acrescida, deverá proporcionar um saldo positivo.

Sendo o Departamento de Sistemas de Informação (DSI) de Bial o órgão gestor do modelo de validações e também responsável pela validações dos sistemas transversais, logo de maior complexidade e quantidade de requisitos, é visível o esforço acrescido pela percentagem de ocupação que representa no conjunto de actividades do Departamento.

É nossa convicção que, também no DSI, esta actividade tende a ser parte natural de todos os projectos, pelo que o esforço de validação tenderá a diluir-se nos projectos e, com o crescimento da equipa, alguns dos recursos de validação poderão tornar-se recursos de projecto. Não é previsível, no entanto, a curto prazo, considerando o número de sistemas e a evolução constante, deixar de ser necessário alocar recursos específicos para esta actividade, essencialmente devido à gestão de alterações e validação em contínuo.

6 Evolução dos procedimentos de validação

No decorrer dos vários projectos de validação, têm sido introduzidas alterações ao modelo, que visam simplificar os procedimentos, de forma a proporcionar resultados de melhor qualidade, com o menor esforço possível. Temos previsto realizar uma auditoria interna, eventualmente com ajuda de uma entidade externa competente, para comprovar se estamos a cumprir os nossos procedimentos internos de validação e também para identificar oportunidades de melhoria nesses procedimentos. Julgamos que essas oportunidades podem surgir no sentido da simplificação (por exemplo, não obrigar a testemunha em todos os testes), mas também no alinhamento com as boas práticas internacionais de garantida de qualidade de sistemas.

Para além disso, neste momento estão em estudo ferramentas integradas para que o DSI possa melhorar o controlo sobre os seus processos, nomeadamente gestão de incidentes, problemas e alterações, que podem representar uma forte sobrecarga, em termos operacionais.

7 Conclusões

Um ano após a implementação formal do modelo de validação de sistemas computorizados e da sua aplicação aos sistemas anteriormente referidos, é possível avaliar a sua eficácia e os benefícios obtidos e também o seu impacto na organização e em particular nos órgãos mais directamente envolvidos.

Através dos resultados obtidos internamente, com os projectos iniciais, podemos afirmar que o modelo é adequado aos fins e à realidade da organização, sendo no entanto reconhecida a oportunidade de melhoria.

É perceptível também o aumento do nível de confiança nos sistemas validados, o que possibilita a simplificação de processos que visavam controlos de qualidade redundantes. Um dos exemplos dessa simplificação foi a eliminação de etiquetas de identificação coloridas, cuja utilização obrigava a uma sobrecarga de trabalho no processo logístico. Isso só foi possível após validação do Sistema WMS.

Temos a convicção que, já hoje em dia, estamos preparados para qualquer inspecção das entidades regulamentares (por exemplo, da FDA), uma vez que avaliámos o risco associado aos nossos sistemas e planeámos o caminho que temos que percorrer para mitigar esse risco.

Em termos gerais, julgamos poder assumir que estamos a caminhar no bom sentido, no que respeita ao alinhamento de Bial com os requisitos dos órgãos que regulam o sector, relativos aos sistemas computorizados, e como tal, encarar com confiança os desafios que se avizinham a curto prazo.

8 Referências

Bial, *Validação de Sistemas Computorizados*, Procedimento Geral DSIPG002, Trofa, 2006

Bial, *Regras de Testes Formais a Sistemas Computorizados*, Instrução Específica DSIIE008, Trofa, 2006

Bial, *Gestão de Alterações a Sistemas Computorizados*, Instrução Específica DSIIE010, Trofa, 2006

Budihandojo, R., Bergeson, D, D'Eramo, P., Huber, L., *The future state of computer validation - part I*, Pharmaceutical Technology North America, Vol.25, nº 7, 2001.

EMEA, *Good Manufacturing Practices – Computerised Systems*, European Medicines Agency, Volume 4 – Annex 11, 1998.

Farkas, J., *The correlation of the FDA's computer system validation enforcement with the rate of adoption of ERP technology by the medical device industry*, Ph.D., Northcentral University, 2006.

FDA, *Electronic Records; Electronic Signatures; Final Rule*, Food and Drug Administration, Federal Register, 21 CFR – part 11, Vol. 62, No. 54, 1997.

FDA, *General Principles of Validation,* Food and Drug Administration, Center for Drug Evaluation and Research, Rockville, MD, 1987.

ISPE, *GAMP Guide for Validation of Automated Systems*, International Society for Pharmaceutical Engineering, GAMP 4 (Good Automated Manufacturing Practice), 2001.

McDowall, R., *Effective and Practical Risk Management Options for CSV*, Quality Assurance Journal, 2005, 9, 196-227.

OECD, *The Application of the Principles of GLP to Computerised Systems*, No. 10, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, GLP Consensus Document, Environmental Monograph No. 116. Organisation for Economic Co-operation and Development Environmental Directorate, Paris, 1995.

Stokes, T., *The Survive and Thrive Guide to Computer Validation*, Interpharm Press, 1998.

Wingate, G., *Computer Systems Validation: Quality Assurance, Risk Management, and Regulatory Compliance for Pharmaceutical and Healthcare Companies*, Taylor & Francis, 2003.

Avaliação de Dados em Sistemas de Data Warehousing

Alexandre Costa [1], Orlando Belo [2]

1) Departamento de Informática, Escola de Engenharia, Universidade do Minho
Campus de Gualtar, 4710-057 Braga, Portugal
alexandre.costa@portugalmail.com

2) Departamento de Informática, Escola de Engenharia, Universidade do Minho
Campus de Gualtar, 4710-057 Braga, Portugal
obelo@di.uminho.pt

Resumo

Em *Sistemas de Data Warehousing* (SDW), a eficácia das decisões tomadas encontra-se fortemente condicionada pelos dados divulgados aos consumidores finais. Neste contexto, é possível medir o sucesso alcançado por estes sistemas pela qualidade dos dados aí residentes. A qualidade dos dados tende a assumir-se como uma preocupação cimeira a enfrentar pelas organizações e por isso, interessa identificar os problemas de qualidade dos dados verificáveis na arquitectura dos SDWs, diagnosticar as razões da sua existência e promover os modos adequados de resolução desses problemas. O presente artigo apresenta um caso prático duma organização real e analisa a qualidade dos dados do DW da organização. Nesse sentido, é descrita uma taxionomia própria das anomalias nos dados verificáveis em SDWs e recorre-se à utilização de software de análise e limpeza dos dados que possibilite a recolha de indicadores da presença dessas anomalias. Baseados nestes resultados é possível a enunciação de algumas recomendações a implementar pela organização.

Palavras chave: Qualidade dos dados em SDW, Dimensões da qualidade dos dados, Taxionomia de problemas nos dados, *Data profiling*, Limpeza dos dados, Duplicação de valores.

1. Introdução

Os SDW assumem um papel de relevo no domínio dos processos de tomada de decisão, ocupando claramente a dianteira, relativamente aos sistemas de suporte à decisão anteriormente disponíveis, como uma plataforma tecnologicamente capaz de disponibilizar um conjunto de meios potenciadores de ir ao encontro das preocupações, necessidades e ambições mais essenciais reveladas pelas organizações, em particular pelos seus agentes de tomada de decisão. Os SDW são uma ferramenta muito útil para suporte às actividades quotidianas dos agentes de decisão, essencialmente na condução táctica das organizações. Este auxílio revela-se importante, na medida em que torna o processo de tomada de decisão mais rápido e efectivo, flexibilizando o acesso a mais dados, melhores e mais bem organizados, e garantindo uma maior confiança aos agentes sobre a credibilidade da informação que disponibiliza. Estes são alguns dos vectores basilares que justificam a integração e a exploração de SDW no seio das organizações. Porém, a sua implementação não corresponde, só por si, ao sucesso da sua implementação – os casos de insucesso são muitos, conforme se pode comprovar através de diversos estudos sobre esta realidade [Kenyon et al. 2004, Neely 1998]. O insucesso dos SDW corresponde muitas vezes, e paradoxalmente, à escassa ou inexistente obtenção das vantagens enunciadas e possível de ser justificada por uma variedade de razões, como seja a qualidade dos

dados. Esta questão, tantas vezes negligenciada pelas organizações, tem vindo a assumir nos últimos tempos um relativo protagonismo, revelando-se cada vez mais como uma fonte da origem desses eventuais fracassos (ou razoáveis sucessos) na construção destes sistemas. As experiências excessivamente aterradoras, em termos de custos (directos e ocultos), infringidas às organizações levam-nas a estarem alerta para as reais vantagens em garantir a melhor qualidade dos dados e por isso, começam a assumir a qualidade dos dados como uma das prioridades a resolver [Eckerson 2002, Wang 2004]. A qualidade dos dados é geralmente definida como a adequação ao uso, isto é, a capacidade dos dados satisfazerem ou superarem os anseios dos consumidores [Cappiello et al. 2004, Wang 1998]. Todavia, a qualidade dos dados permanece uma ciência inexacta em termos de validação e benchmark [Kahn 2002]. Estas dificuldades devem-se sobretudo a aspectos de natureza multidimensional e evolutiva dos próprios dados. A definição do nível apropriado da qualidade dos dados está dependente do seu contexto [Pipino et al. 2002]. Ora, este pressuposto configura a constituição dos dados com aspectos de natureza factual e percepcional. A visão factual ou imparcial dos dados permite aferir quantitativamente da qualidade dos mesmos, ou seja, possibilita a avaliação da qualidade dos dados independentemente de outros factores, como sejam os utilizadores ou a decisão a tomar. A visão percepcional ou contextual faz depender a qualidade dos dados da utilidade que estes possuem na satisfação das necessidades dos utilizadores ou da decisão a tomar [Skriletz 2002, Shankaranarayan 2005]. Assim, a avaliação dos dados constantes no repositório do DW mostra-se como uma iniciativa a implementar tendo em vista a obtenção dos dados adequados aos utilizadores finais.

2. Problemas nos Dados em SDW

A presença de dados de fraca qualidade em SDW deve-se a um conjunto alargado de factores, como sejam, entre outros, a combinação de dados que não foram projectados para serem integrados, a recolha de dados irregulares e a pouca atenção dos responsáveis organizacionais em assegurar uma boa qualidade dos dados [Ballou & Tayi, 1998]. A arquitectura de um SDW pode ser perspectivada em torno de camadas essenciais: o Sistema Operacional (SO), a *Área de Retenção de Dados* (ARD), o repositório de dados do DW e os Data Marts (DMs) que interagem com os utilizadores [Chaudhuri & Dayal 1997]. Cada uma destas camadas é fonte de um conjunto de problemas em relação à qualidade dos dados.

2.1 Sistema operacional

Os problemas que se encontram nesta camada derivam geralmente de dois factores. Um primeiro factor respeita à dificuldade na interpretação dos dados constantes nas fontes. A heterogeneidade e diversidade das fontes dificultam a reunião dos melhores dados a integrar no DW (e.g. diferentes níveis de história dos dados, colunas sobrecarregadas). Os repositórios de dados alvos de recolha são, muitas vezes, antigos e escassos no fornecimento de metadados reveladores das características dos dados. Um segundo factor deriva da recolha de dados irregulares pelo SO. Esta questão constitui-se como o principal foco de origem de defeitos nos dados [Kimball et al. 1998] e que são motivados por: a) aplicações de recolha de dados complexas e que induzem a inserção de erros; b) dificuldades na inserção completa e correcta dos dados por limitações temporais na recolha dos dados; e c) inserção involuntária de dados válidos mas incorrectos (e.g. o operador digita a idade igual a 43 em vez de 34) ou voluntária de valores errados (e.g. inserção sistemática de valores por defeito ou valores fictícios).

2.2 Área de Retenção dos dados

A ARD assume importância capital na garantia da boa qualidade dos dados em SDWs, essencialmente, por dois motivos. Em primeiro lugar pela inexistência ou escassez de iniciativas de promoção da melhoria da qualidade dos dados no SO. Em segundo lugar, a circunstância de existirem múltiplas fontes para os mesmos dados ou que permitem a sua complementaridade, determina que seja este, vulgarmente, o primeiro local de encontro desses dados dispersos. Os

dados provenientes do SO apresentam-se geralmente com pouca qualidade, quando submetidos às exigências de um consumidor de informação do DW e por isso, é comum designá-los por dados sujos. Os dados sujos compreendem uma variedade de defeitos (ruídos, inconsistências e dados incompletos). Os dados incompletos revelam a ausência de valores, a ausência de atributos de interesse ou respeitam somente a valores agregados. Os dados manifestam ruído quando são introduzidos valores aberrantes, erros aleatórios e erros provocados pelo sistema. A inconsistência dos dados surge da introdução incorrecta de valores, de representações diferentes sobre os mesmos dados ou de erros oriundos da integração de várias bases de dados. No futuro, as tarefas executadas neste local tendem a tratar problemas de muito maior complexidade e que resultam do crescimento exponencial da captura de dados de diferente natureza. Assim, além dos dados estruturados comummente integrados nos SDWs, outra classe de dados carente de estruturação, pela sua própria natureza, será igualmente alvo de carregamento nos DWs (e.g., imagens, mensagens de correio electrónico) [Inmon 2006]. Esta realidade, elevada ainda mais a importância vital deste local na arquitectura dos SDWs.

2.3 DW/DMs

A ocorrência de problemas na camada do DW/DMs relacionados com a qualidade dos dados pode ficar a dever-se ao incumprimento cabal das tarefas na ARD. Muitas vezes as tarefas na ARD não executam a totalidade de operações necessárias na reparação e preparação dos dados a integrar no DW. Em casos mais graves assiste-se ao curto-circuito do próprio sistema, isto é, os dados provenientes do SO são vertidos directamente no DW. Uma outra gama de problemas comuns de ocorrer neste local resulta da permanência de dados históricos no sistema. Estes dados possuem significados diferentes daqueles que possuíam anteriormente, na altura da sua introdução, ou seja, os dados recentes e veteranos que são registados num mesmo campo tendem a não se referirem à mesma característica (e.g. pode ser observável que antes duma determinada data a representação do sexo de um aluno fosse ‘M’ ou ‘F’ e que após essa data o sexo passe a ser representado por ‘1’ e ‘0’). Estes dados vincados por um peso histórico que os descaracterizam, não são, geralmente, alvo de interrogações, e quando aliados a outros que também não são acedidos pela irrelevância que mostram junto dos decisores, constituem um bloco de dados inertes ou monos, vulgarmente designados por dados dormentes [Inmon et al. 1998]. Assiste-se ainda, neste local, a problemas resultantes da difusão de informações de qualidade desajustada em termos de interpretação, prontidão temporal, ou credibilidade das fontes.

3. Taxionomia dos Defeitos nos Dados

A catalogação dos erros encontrados em sistemas de dados é um tema estudado em diversas investigações. Em [Olson 2003] é abordada a identificação das anomalias nos dados de modo conciso e pragmático, mas mostra-se demasiado simplista na tipificação dos erros existentes. Em [Adelman et al. 2005] são descritas mais detalhadamente as regras para a validação dos dados, em torno das dimensões mais representativas. As investigações apresentadas em [Kim et al. 2003] e [Oliveira et al. 2005a, 2005b] abordam os defeitos nos dados baseando-se num processo de refinamentos sucessivos de busca do problema alvo de observação. Nestes estudos, um problema nos dados é conduzido por patamares sucessivos de avaliação, de acordo com a sua tipologia, até ser encontrado o problema específico. As investigações consideradas encontram-se orientadas para a generalidade dos repositórios de dados e como tal, mostram-se exaustivas na identificação das anomalias verificadas nesses locais, sendo por isso, adequadas para repositórios de dados ainda não sujeitos a imposições de níveis de qualidade. Todavia, no âmbito dos SDWs, muitas acções de melhoramento dos dados encontram-se geralmente já concretizadas (e.g. a adopção de chaves de substituição). Assim, iremos adoptar as terminologias consideradas válidas para o caso em estudo e adoptar outras que visem verificar a ocorrência de outros defeitos ao nível da qualidade dos dados, além dos considerados nessas investigações. É o caso da relevância dos dados que se encontram ou não armazenados no

repositório. Por outras palavras, podemos analisar a qualidade dos dados existentes em torno da qualidade do modelo multidimensional adoptado [Calero et al. 2001] e em termos do cumprimento dos requisitos dos utilizadores.

Nº	Anomalia	Descrição
1	Valor ausente	São os valores relevantes para os consumidores finais, mas que não se encontram presentes na estrutura multidimensional.
1.1	Linha incompleta	Consiste num significativo número de colunas, referentes a uma linha, que não se encontram devidamente preenchidas.
1.2	Coluna incompleta	Verifica-se quando um expressivo número de linhas, referentes a uma coluna, não se encontram preenchida.
1.3	Inexistência de coluna relevante	Corresponde à ausência de uma coluna considerada importante para a produção de resultados.
1.4	Inexistência de linha relevante	Corresponde à ausência de uma linha de dados considerada importante para a produção de resultados.
2	Valores irrelevantes	Refere-se à materialização de valores insignificantes (e.g. linhas ou colunas duplicadas).
3	Valores inválidos	São aqueles que violam um conjunto ou intervalo de valores possíveis.
4	Valores válidos	São os valores que contêm um conteúdo aceite pelo sistema.
4.1	Correctos	Apresentam o conteúdo com o valor correspondente ao mundo real.
4.1.1	Representação certa	São os valores que apresentam o conteúdo e o formato definidos pelo sistema.
(...)	(...)	(...)
4.2	Errados	Correspondem aos valores cujo conteúdo se apresenta incorrecto.
4.2.1	Erros ortográficos	Ocorre quando são inseridos erros ortográficos.
4.2.2	Incorrectos	Os valores que não correspondem à realidade (e.g. a moeda deixa de ser escudo e passa a euro).
4.2.3	Duplicação de valores	Ocorrência de diferentes valores para a mesma entidade.
4.2.4	Além do contexto da coluna	A coluna está sobrecarregada de valores além dos adequados para a coluna (e.g. inserção de símbolos).
4.2.5	Integridade referencial	Quando uma coluna que é chave estrangeira numa tabela contém um valor que não existe como chave primária na tabela relacionada.
4.2.6	Inconsistência de valores	Quando existem contradições entre os valores das colunas para uma mesma entidade.
(...)	(...)	(...)

Tabela 1 – Excerto da taxionomia de anomalias nos dados verificáveis em DMs

Assim, importa aferir sobre a pertinência de determinadas colunas ou a ausência de outras passíveis de auxiliarem os decisores. Estas questões apesar de importantes não são alvo de análise na sua plenitude, especialmente, no que diz respeito à ausência de colunas, no modelo multidimensional, importantes para os consumidores finais (engenharia de requisitos). Tanto por razões de objecto de estudo, como de natureza logística, coloca este assunto além do âmbito deste trabalho. Relativamente, à manutenção de colunas no modelo multidimensional, aparentemente, sem utilidade para as interrogações comuns por parte dos consumidores, iremos relacioná-las com a completude dos dados existentes (e.g., coluna com elevado índice de valores ausentes). Apesar deste não ser um critério único, mostra-se uma condição necessária. O dinamismo associado ao processo de tomada de decisão condiciona as colunas a serem alvo de apreciação ou cálculo para indicadores. Por isso, importa abordar as colunas existentes que apresentam pouca probabilidade de acesso num futuro próximo. Um outro critério passível de ser avaliado é relativo à inconsistência de representação de valores de todas as colunas da tabela e que se pode relacionar com a interpretação dos dados, isto é, devem realizar-se operações de transformação idênticas sobre os valores (e.g. uma coluna mantém valores 'Y' e 'N', enquanto outra alberga os valores 'S' e 'N').

Nº	Descrição	Nº	Descrição

1	Decomposição dos dados em vista a obtenção de elementos atómicos.	5	Enriquecimento do conteúdo dos dados.
2	Operações de transformação:	6	A resolução do problema dos valores duplicados nos dados das fontes.
2.1	Estandardização (maiúsculas e minúsculas, acrónimos e abreviaturas).	7	Intervenção de perito.
2.2	Normalização (e.g. estabelecer regras de negócio).	8	Restringir a entrada de valores.
2.3	Correcção.	9	Reforçar a obrigatoriedade da entrada de valores.
3	Preenchimento de valores ausentes.	10	Alterar tipo de dados.
4	A aplicação das regras de integridade referencial.	11	Reformular o modelo multidimensional.

Tabela 2 – Métodos de resolução das anomalias nos dados

Esta observação permite igualmente aferir por um lado, sobre a qualidade dos processos de transformação dos dados e por outro lado, sobre o modo de produção do software. Tendo em vista a classificação inequívoca das irregularidades verificadas nos dados constantes no DM em estudo, adoptamos uma taxionomia própria, fundamentada nas propostas enunciadas anteriormente e que se encontra parcialmente apresentada na tabela 1. A resolução das anomalias verificadas deve prever a sua solução da forma mais adequada, conforme se pode verificar na tabela 2.

Figura 1 – Esquema simplificado referente ao DM das vendas dos artigos

4. Um Caso de Aplicação Prática

O foco de atenção do nosso trabalho definiu-se em torno da análise dos dados de algumas dimensões de um DM de vendas que sustentou parte das actividades decorrentes da gestão comercial de uma organização, que não será revelada por razões de confidencialidade. Para tal, a organização disponibilizou uma amostra representativa de dados. Em resultado das primeiras observações sobre os dados constantes nesse DM, verificaram-se algumas imperfeições nos dados comuns de ocorrer na ARD. O transporte de uma significativa quantidade de dados, entre o SO e os DMs, é realizada quase directamente ("curto-circuito" do sistema) e não são facultadas informações sobre os dados que se baseiem em metadados. Neste contexto, interessa perceber o estado dos dados no DM alvo de estudo e antever a sua influência nas respostas às interrogações dos consumidores. A figura 1 ilustra simplificadamente o DM em questão.

4.1 Indicadores sobre os dados

O entendimento sobre o estado dos dados assentou na recolha de indicadores objectivos sobre os dados e que se encontrem consignados à taxionomia apresentada. Nesse sentido, efectuou-se a avaliação dos dados com recurso às potencialidades oferecidas pelas ferramentas de análise de dados (*data profiling*, limpeza – tratamento de valores duplicados e auditoria). Estas ferramentas são capazes de transmitir além de medidas objectivas e estatísticas sobre os dados, também conhecimentos sobre as características destes e que são possíveis de se consubstanciarem, no futuro, em metadados a integrar no sistema. Os resultados obtidos pelo software considerado e os acessos directos aos dados servem de base para a concretização do objectivo central proposto e que consiste em reconhecer a problemática da qualidade dos dados

no DM alvo de estudo. Os resultados obtidos revelam a existência de factos sobre os dados (e.g. a coluna *cod_zona_nac*, da tabela unidade funcional apresenta 54% de valores omissos). A junção destes factos pode revelar assuntos sobre os dados e posteriormente ser a base de iniciativas de melhoramento [Olson 2003]. Os indicadores obtidos pela execução dos programas focam algumas dimensões concretas dos dados, em especial, aquelas que apresentam um condão intrínseco factual e independente das tarefas [Pipino et al. 2002]. Assim, são colhidas medidas que permitem avaliar os dados sobre: a ausência de valores nos dados, a relevância dos dados disponibilizados pelo sistema, a correcção dos valores armazenados em relação à realidade, a facilidade na interpretação dos dados divulgados e a consistência da representação dos valores dos dados. Conscientes da existência de outras medidas capazes de permitir a aferição sobre a qualidade dos dados, os critérios adoptados parecem os mais adequados e de possível averiguação, tendo em conta o nível de maturidade da organização e a concretização do objectivo proposto. Neste contexto, o estudo será realizado pela enunciação dos defeitos identificados nos dados, a sua catalogação e posterior indicação de resolução por um método adequado para o efeito.

4.2 Sobre o software

As aplicações de software consideradas para a realização deste trabalho consistem em dois tipos: *data profiling* e detecção de valores duplicados (tabela 3). A justificação para a utilização destas aplicações consiste no facto da sua acção conjunta possibilitar uma recolha mais completa de informações relativas aos dados constantes no repositório. Nesse sentido, a aplicação de *data profiling* proporciona a captação de informações relativas ao conteúdo e estrutura da qualidade dos dados existentes nas tabelas. A ferramenta de detecção de linhas duplicadas procura, baseada num conjunto de parâmetros, a identificação de linhas similares sobre uma entidade, numa ou em várias tabelas. Neste caso, dado o caso em estudo incidir sobre um DM, será efectuada a pesquisa apenas sobre uma única tabela. Adicionalmente, recorreu-se a consultas directas sobre os dados como forma de comprovar tendências ou problemas identificados pela utilização das aplicações.

Aplicação	Tipo	Versão	Limitações	Site
Datiris	Data profiling	Profiler Professional 1.2 – Trial	15 dias	www.datiris.com
Wizsame	Identificação de duplicados	1.02 Demo	Máximo 1000 linhas	www.wizsoft.com

Tabela 3 – Características do software utilizado no caso em estudo

Consciente da existência de outras aplicações, certamente, dotadas de maiores potencialidades, importa referir que a opção pelo software utilizado se deve ao escasso número de ferramentas disponíveis em versões de teste e de demonstração. É igualmente importante salientar que o recurso às aplicações consideradas consiste apenas como meio de trabalho na identificação de defeitos a nível da qualidade dos dados sobre o caso em estudo. Assim, apenas são expostas parte das capacidades patenteadas por estas aplicações, visto que o objectivo passa pela enunciação das anomalias dos dados e possíveis recomendações de tratamento e não a resolução dos problemas ou validação das ferramentas usadas.

4.3 Análise dos dados

4.3.1 Considerações gerais sobre as dimensões

As dimensões: fornecedor, promoção, artigo e unidade funcional, justificam pela sua natureza serem consideradas as prioritárias em termos de análise dos seus dados. Por imposição das ferramentas utilizadas (e.g. limite de linhas a avaliar) e do volume de dados a analisar, o modo de abordagem, em termos de análise, mereceu necessariamente uma devida reflexão. Sempre que possível as primeiras análises efectuadas reflectiram-se sobre a totalidade das linhas das tabelas. Os resultados obtidos servem de ponto de partida para análises mais específicas, mais manuseáveis e capazes de melhor revelarem anomalias nos dados. A definição de análises mais

específicas efectuou-se segundo dois critérios: as colunas apresentarem o campo *status_key* igual a 'A' e sobre tranches de mil linhas. A justificação para o estabelecimento de lotes de mil linhas deve-se ao facto das ferramentas consideradas serem versões limitadas nas suas capacidades de utilização. De qualquer modo, esta situação é perfeitamente possível à luz de [Kimball & Caserta 2004], porque inclusivamente faculta uma maior flexibilidade dos dados a examinar. O propósito que residiu à definição das vistas com a coluna *status_key* igual a 'A' consistiu no facto de considerar estas linhas como as activas (maior probabilidade de interrogações) e mais recentes e por isso, supostamente, as que apresentam maiores índices de qualidade, em resultado de operações de tratamento e limpeza mais efectivas sobre os dados. Esta constatação deriva de uma maior homogeneização dos valores dos dados registados (e.g. apenas maiúsculas, sem linhas duplicadas, sem espaços em branco antes ou após os valores dos dados e sem caracteres especiais). Adicionalmente, é possível observar a existência de duas colunas, *data_activacao* e *data_inactivacao*, que balizam o período durante o qual as linhas se encontram activas no sistema. Verifica-se que esta abordagem sobre a manutenção das linhas nas dimensões segue a aproximação de tipo 2 para as dimensões de variação lenta [Kimball et al. 1998]. As linhas com a coluna *status_key* igual a 'I', apresentam a coluna *data_inactivacao* com o preenchimento duma data já passada. Esta situação induz que se tratam de linhas actualmente inactivas, mas que são mantidas no sistema de modo a preservar a história e as interrogações temporais dos dados. Saliente-se o peso destes dados representar nas dimensões artigo, fornecedor e unidade funcional, cerca de 41%, 53% e 24% respectivamente do volume total de linhas. As características gerais das dimensões podem ser observadas na tabela 4.

Dimensão	**Tamanho (em linhas)**	**Linhas activas (*status_key*='A')**		**Linhas inactivas (*status_key*='I')**		**Nº colunas**
Artigo	1 434 126	852 659	*(59,5%)*	581 467	*(40,5%)*	51
Unidade funcional	1 183	897	*(75,8%)*	286	*(24,2%)*	46
Promoção	19 154	19 154	*(100%)*	-	*(0%)*	13
Fornecedor	33 442	15 901	*(47,5%)*	17 541	*(52,5%)*	18

Tabela 4 – Propriedades gerais das tabelas a analisar

De referir que a análise sobre a dimensão artigo não será apresentada por duas ordens de razões. Primeiro, o elevado número de linhas da tabela inviabiliza uma análise eficaz, atendendo às restrições impostas pelo software utilizado. Segundo, os problemas verificáveis são em tudo semelhantes aos registados nas restantes dimensões, com excepção da presença de inconsistências na representação de valores. Do mesmo modo, a dimensão fornecedor não será alvo de análise em virtude dos dados apresentarem incorrecções idênticas às outras dimensões.

4.3.2 Dimensão unidade funcional

A análise de *data profiling* sobre a vista que apresenta a coluna *status_key*='I' revela elevados índices de valores ausentes num número significativo de colunas, encontrando-se muitas linhas semi-vazias (tabela 5). Cerca de 50% das colunas apresentam elevados índices de valores nulos, zeros ou brancos. Destas, cerca de metade revelam taxas de ausência acima dos 50%. A análise detalhada sobre a vista contendo as linhas mais recentes (*status_key*='A'), aponta para taxas de ausência de valores nulos e brancos, muito superiores. Cerca de dois terços das colunas mostram taxas de ausência superiores a 50%. Dado que a incidência de valores nulos, zeros e brancos se agrava na vista que contem os registos mais recentes, podemos deduzir que estas colunas não se mostram críticas para a organização e consequentemente tendem, aparentemente, no presente e no futuro, a não serem alvo de interrogações. Além desta observação é possível constatar que determinadas colunas contêm valores válidos, mas incorrectos e que inviabilizam qualquer tipo de análise. É o caso da coluna *area_venda* que apresenta o mesmo valor '1' em 92% das linhas. Logo, podemos considerar estes dados (coluna) como dormentes porque aparentam uma certa irrelevância ou inutilidade quando colocados no DM afim de serem alvo de interrogações. Relativamente, aos valores por defeito predefinidos encontram-se diferenças

de representação. Alguns dos valores por defeito utilizados são: ? e ???. Acresce que os valores por omissão atribuídos representam uma minoria face aos valores nulos, brancos e zeros. As consultas directas realizadas sobre a dimensão revelam uma variedade de incorrecções sobre os dados de acordo com a taxionomia adoptada.

4.3.3 Dimensão promoção

Uma análise geral sobre esses resultados revela a coluna *descricao* como aquela que apresenta o maior grau de probabilidade da ocorrência de defeitos nos valores dos dados. Esta observação resulta da heterogeneidade dos valores existentes na coluna e que são originados pela ausência de regras de negócio e cumprimento de critérios na inserção destes valores.

Coluna	Distintos		Nulos	Zeros	Brancos	Padrões	Observações
Cod_und_funcional	897	76%					
Descr_und_funcional	948	80%				554	
Sigla_8			1,4%		0,5%	60	8 caracteres
Área_venda				3,8%			1 = 1106 (93,5%)
Descr_zona			36%				? = 121
Primeiro_mes				43%			0 = 507
Primeiro_ano_cc				57%			0 = 674
Descr_cc			53%			148	
Data_abertura_cc			57%				

Tabela 5 – Características dos dados do extracto da vista com a coluna status_key='I' da dimensão unidade funcional (Datiris)

As informações obtidas pelo processo de *data profiling* corroboram esta análise, conforme é observável na tabela 6. Cerca de 27% das linhas apresentam valores iguais na coluna descrição e muitas outras linhas mostram valores similares ou sinónimos. Esta questão é merecedora de maior atenção na medida em que, praticamente, metade dos valores distintos (6970) exibe padrões de construção diversos. Esta tendência confirma a despreocupação na orientação sobre a implementação de padrões regulares na construção dos valores introduzidos. É, igualmente, possível observar a elevada taxa de valores ausentes nas colunas *evento*, *prom_type* e *old,* registando 47%, 57% e 100% respectivamente. Ainda no domínio da ausência de valores dos dados verifica-se a indistinção entre valores nulos e ausentes. A presença de valores idênticos sobre uma mesma coluna (*descricao*) pode indiciar a ocorrência de valores duplicados na tabela. Esta situação merece uma análise mais cuidada em vista aferir sobre a possibilidade de duplicação de linhas. A análise apenas sobre esta coluna não permite concluir sobre a duplicação de linhas, outras colunas necessitam ser confrontadas (*data_inicio* e *data_fim*). Para a elucidação desta questão recorreu-se à aplicação wizsame, com o intuito de aferir sobre a duplicação de linhas. A execução da aplicação implicou a partição da tabela em tranches de 1000 linhas. Tendo em vista determinar a duplicação de linhas, a aplicação foi parametrizada para as colunas: *descricao*, *data_inicio* e *data_fim*. Dado que a coluna *descricao* apresenta um significativo número de linhas com conteúdo igual ou similar a 'promoção', optou-se por submeter uma outra amostra à aplicação wizsame. Tendo em vista a obtenção de resultados mais objectivos, parametrizou-se a aplicação para considerar linhas duplicadas quando a coluna descricao apresente um conteúdo similar a 'promoção' e as colunas *data_inicio* e *data_fim* sejam exactamente iguais. Obtiveram-se cerca de 80 ocorrências que satisfizeram os requisitos predefinidos e como tal indiciando a presença de valores duplicados. Uma possível justificação para a presença de linhas duplicadas deve-se à introdução de uma ocorrência por cada unidade funcional que efectue uma promoção. Pelo que é dado a observar, uma mesma promoção pode estar a ocorrer simultaneamente em unidades funcionais distintas. Esta realidade inviabiliza uma análise mais flexível e de maior grau de profundidade nos acessos realizados pelos consumidores dos dados. Por exemplo, como se determina a influência da linha 'Saldo montanha 50%', em cada unidade funcional ou artigo, se existem dezasseis linhas iguais?

Coluna	Distintos		Nulos	Zeros	Brancos	Padrões	Observações
Descricao	13 916	73%				6 970	
Evento			48%				
Prom_type			57%				
Old			100%				Sem valores

Tabela 6 – Características da vista 1, da dimensão promoção (datiris)

5. Conclusões e Trabalho Futuro

Neste artigo, perspectivámos a qualidade dos dados ao longo das quatro camadas da arquitectura de um SDW, de modo a um mais fácil entendimento dos motivos para as circunstâncias da fraca qualidade dos dados em cada camada, bem como os problemas daí inerentes. Esta constatação permitiu estabelecer uma taxionomia das irregularidades verificáveis em DWs e simultaneamente antever as estratégias mais adequadas da resolução desses problemas. A validação desta taxionomia realizou-se com o recurso ao estudo de um caso real. Para o efeito, recorremos a algumas ferramentas de software de qualidade dos dados e que assumem um papel preponderante porque disponibilizam informações preciosas relativamente às características dos dados. As informações recolhidas permitiram enquadrar a organização num panorama em que pouca atenção é prestada à qualidade dos seus repositórios de dados ou em que a estratégia de inspecção e correcção dos dados actua de modo descoordenado e descontinuado e fica aquém das exigências organizacionais. Assim, exige-se uma mudança estratégica e cultural que institucionalize uma visão preventiva dos dados de modo a efectuar uma gestão equilibrada da qualidade destes. Por outras palavras, o tratamento da qualidade dos dados o mais a montante possível no fluxo circulatório dos dados pelos SDWs. No futuro, pretendemos enriquecer e testar a taxionomia das irregularidades dos dados em outros casos e simultaneamente desenvolver estudos sobre a análise dos custos provocados pela fraca qualidade dos dados, a análise de ROI a obter pela implementação de iniciativas de melhoria da qualidade dos dados e a definição de critérios a satisfazer junto dos agentes de decisão.

6. Referências

Adelman, S., Moss, L. e Abai, M. “Data Strategy”. Addison-Wesley Professional. 2005.

Ballou, D. e Tayi, G. “Examining Data Quality”. Communications of the ACM, vol. 41, nº2, pp. 54-57. Fevereiro, 1998.

Calero, C., Piattini, M., Pascual, C. e Serrano, M. “Towards Data Warehouse Quality Metrics”. Proceedings of the 3th International Workshop on Design and Management of Data Warehouses (DMDW'01). Interlaken, Suiça. Junho, 2001.

Cappiello, C., Francalanci, C. e Pernici, B. “Data quality assessment from the user’s perspective”. Communications of the ACM, pp. 59-67. Maison de la Chimie, Paris, França. 2004.

Chaudhuri, S. e Dayal, U. “An Overview of Data Warehousing and OLAP Technology”. SIGMOD Record, vol. 26, nº1. Março, 1997.

Eckerson, W. “Data Quality and The Bottom Line: Achieving Business Success Through a Commitment to High Quality Data”. The Data Warehousing Institute. www.dwinstitute.com. Seattle, EUA. 2002.

Inmon, B. “An Introduction to DWs 2.0”. Business Intelligence Network. Março, 2006.

Inmon, W., Rudin, K., Buss, C. e Sousa, R. “Data Warehouse Performance”. John Wiley & Sons, Inc. New York, EUA. 1998.

Kahn, B., Strong, D. e Wang, R. ”Information Quality Benchmarks: Product and Service Performance”. Communications of the ACM, vol. 45, nº4, pp. 184-192. Abril, 2002.

Kenyon, H., Smith, P., Jordaan, W., Marinos, G. e Benge, J. “Data Quality Management”. PriceWaterhouseCoopers. www.pwc.com. 2004.

Kim, W., Choi, B., Hong, E., Kim, S. e Lee, D. “A Taxonomy of Dirty Data”. Data Mining and Knowledge Discovery, nº7, pp. 81-99. Kluwer Publishers. 2003.

Kimball, R. e Caserta, J. “The Data Warehouse ETL Toolkit: Practical Techniques for Extracting and Cleaning”. John Wiley & Sons, Inc. New York, EUA. 2004.

Kimball, R., Reeves, L., Ross, M. e Thornthwaite, W. “The Data Warehouse Lifecycle Toolkit”. John Wiley & Sons, Inc. New York, EUA. 1998.

Neely, M. “Data Quality Tools for Data Warehousing – A Small Sample Survey”. Center for Technology, University of Albany. EUA. 1998.

Oliveira, P., Rodrigues, F. e Henriques, P. “A Framework for Detection and Correction of Data Quality Problems”. Proceedings of Data Gadgets' 2005 Workshop – Bringing Up Emerging Solutions for Data Warehousing Systems. pp. 50-74. Granada, Espanha. Setembro, 2005.

Oliveira, P., Rodrigues, F., Henriques, P. e Galhardas, H. “A Taxonomy of Data Quality Problems”. Proceedings of 2nd International Workshop on Data and Information Quality. (in conjuction with CAISE'05), pp. 219-233. Porto, Portugal. Junho, 2005.

Olson, J. “Data Quality: The Accuracy Dimension”. Morgan Kaufmann Publishers. Elsevier Science. 2003.

Pipino, L., Lee, Y. e Wang. R. “Data Quality Assessment”. Communications of the ACM, vol.45, nº4, pp. 211-218. Abril, 2002.

Shankaranarayan, G. “Towards Implementing Total Data Quality Management in a Data Warehouse”. Journal of Information Technology Management, vol. 16, nº 1. 2005.

Skriletz, R. “Strategic Insight: Data Quality: Fact and Perception”. DM Review Magazine. Novembro, 2002.

Wang, R. “A Product Perspective on Total Quality Management”. Communications of the ACM, vol.41, nº2, pp.58-65. Fevereiro, 1998.

Wang, R. “Data Quality: Theory in Practice”. EPA 23rd Annual National Conference on Managing Environmental Quality Systems. 2004.

UML to OWL Mapping Overview
An analysis of the translation process and supporting tools

David de Almeida Ferreira [1], Alberto Manuel Rodrigues da Silva [2].

1) INESC-ID/Instituto Superior Técnico, Lisbon, Portugal
david.ferreira@inesc-id.pt

2) INESC-ID/Instituto Superior Técnico, Lisbon, Portugal
alberto.silva@acm.org

Abstract

This paper presents a brief overview of the translation techniques between the OMG's UML modeling language and ontological engineering languages such as the W3C's standard for Semantic Web, the Web Ontology Language (OWL). The main objective of this paper is to analyze the feasibility and easiness of modeling an ontology by using common and mature UML CASE tools, thus reducing the gap between software engineering practitioners and AI techniques for ontology creation. The analysis and practical application of a proposed MOF-based Ontology Definition Metamodel (ODM) and the respective Ontology UML Profile (OUP) allows the straightforward comparison between MOF and OWL concepts.

Keywords: Modeling, Translation, Semantic Web, Ontology, UML, MOF, ODM, OUP, OWL.

1 Introduction

Nowadays we are observing a significant effort of introducing meaning (semantics) to data stored in the Internet. The W3C started to develop an ambitious framework called "Semantic Web" to define common formats for knowledge data interchange for overcoming the original Web limitation of only interchanging data through documents [Miller 2004], which constitute a coarse-grain level of knowledge bundles. To achieve this goal, the Semantic Web effort also engages the definition of a XML-based language for recording how data relates to real world objects allowing a human or a machine to store, access, and interchange that knowledge with a wide range of databases, which are connected through a lattice of concepts.

The main objective of the Semantic Web framework is to enable data to be shared and reused across application, enterprise, and community boundaries in a straightforward manner. Domain ontologies are the formal organization of domain-specific knowledge, thus constituting the most important part of Semantic Web applications. Besides the advantages in terms of knowledge interchange through heterogeneous environments, the developed ontology-based framework reduces the gap between the wider software engineering population and the disregarded, or even unknown, potential of Artificial Intelligence (AI) techniques, like reasoning.

The main purpose of this paper is to gather information about the current state of the art in this investigation area, and present some initiatives that allow one to create an ontology by using Unified Modeling Language (UML) CASE tools, and further refinement with ontology-specific tools. One of these initiatives, which provided a strong and consolidate theoretical background on the field, proposes an Ontology Definition Metamodel (ODM) [Djuric 2004]. This MOF-based metamodel was used as the starting point to defining the Ontology UML Profile (OUP) [Djuric 2004], which allows the creation of an ontology by using the UML modeling language.

Initially, in Section 2, the reader will be presented with a brief overview of Semantic Web Architecture and the Web Ontology Language (OWL). Section 3 contains the theory that supports the conceptualization of an Ontology Definition Metamodel (ODM) by using the OMG's Model Driven Architecture (MDA). Section 4 provides some insight on the peculiarities of the translation process and ontology engineering tools, and Section 5 describes a practical case study. Finally, the last section presents the main conclusions of this research.

2 Semantic Web Architecture's Overview

The World Wide Web "next big step" is the broad adoption of the W3C's Semantic Web framework, enabling the interchange of computer-understandable data over the Internet [Miller 2004]. The adoption of a XML-based language provides the interoperability needed and, simultaneously, enables information enrichment with metadata tags, thus endorsing it with meaning through a precise semantics specification. Semantic Web architecture consists in a functional, non-fixed, three-level architecture. Each one of these levels incrementally introduces new primitives providing more expressive power, namely:

– **Metadata layer:** since most of RDF [Herman 2006a] usage is for describing data about data ("metadata"), in which assertions about other assertions are made, this layer justifies its existence by providing a set of basic constructs that fulfill this low-level functionality.

– **Schema layer:** it's required to declare the existence of new properties and constraining their usage. This type of meta-assertions enables one to perform elementary validations on a document. At the schema layer, the language allows simple assertions about the acceptable combinations of lower-level constructs. These basic constraints are easily expanded into more powerful logical layer expressions; however, the expressive power has to be restricted to guaranty it is still logically well-defined, hence enabling further semantics addition in the next level without compromising the usage of inference-engines at the logical layer.

– Logical layer: provides the mechanisms for writing logic into documents by defining logical rules. This allows document type's inheritance deduction, documents validation against a set of self-consistency rules, and query resolution by conversion from unknown terms into known ones; thus the potential usages of RDF are only limited by imagination.

Figure 1 presents the languages that support Semantic Web and the respective place of OWL [Herman 2006b] in the overall architecture. As depicted, XML is responsible for the syntactical level, whereas the semantic level is provided by RDF, RDF Schema, and especially by OWL.

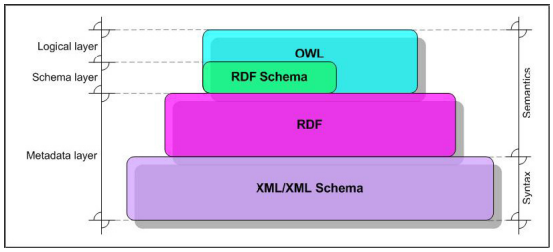


Figure 1 – Semantic Web Architecture (adapted from [Djuric 2004])

The OWL has three sub-languages which are a subset of one another. Each one of them has more expressive power than the inner sub-language, allowing one to model problems at different complexity levels. This OWL's partitioning was introduced as a countermeasure for the trade-offs of using more expressive languages instead of simpler ones: (1) by using an unrestricted language as OWL Full (which provides capabilities for unconstrained representation of knowledge), we lose the guaranty of the OWL DL support for calculations and reasoning in finite time with the use of inference-engines; however (2) the usage of OWL DL

may constitute a complexity overkill for modeling a simple ontology. Therefore, with this clear separation, one can choose the OWL subset that best fits for a specific problem at hands. A brief description of each one of these OWL's sub-languages and their main features is provided by [McGuinness 2004].

3 ODM's Adoption of MDA Standards

The initiative adopted as a reference for mapping UML to OWL was [Djuric 2004]. Its main goal is to provide a widely adopted, industry standard, graphical modeling language for easy ontology creation. This initiative was chosen since it presented a solid work on this area, strongly consolidated by a robust theoretical groundwork. To understand the followed approach one has to bear in mind some of the main concepts in which it is based on. To inherit the main advantages of OMG's previous work (in terms of standard-based definitions, interoperability by intra-level bridging, and also metamodeling and translation mechanisms) the authors specified an Ontology Definition Metamodel (ODM), founded on OMG's Model Driven Architecture (MDA). Only after this brief MDA's main concepts overview one can fully comprehend the proposed approach.

3.1 OMG's MDA Paradigm

Model Driven Architecture (MDA) is a framework for software development presented and maintained by the Object Management Group (OMG). The core of MDA paradigm is based on the importance of models in the software development process. Within MDA the software development process is driven by the activity of modeling a specific software system. It separates domain knowledge from platform knowledge. The more clearly this separation of concerns (between these two logically independent fields of knowledge) is made, the more fundamental MDA's advantages are emphasized, namely: (1) separation between domain and platform expertise, enabling reusability; (2) quick adaptability to domain and technology changes, ensured by efficient maintenance and system's documentation processes; (3) high productivity by generating working, high-quality applications; (4) easiness of integration with legacy or other third party systems by enhancing portability and interoperability; and (5) leveraging by reuse of existing IT technologies.

As presented in Figure 2, MDA's metamodeling architecture presents a four-layer organization and entails several OMG's additional standards. The mentioned MDA's layers are: (1) meta-metamodel (M3) layer, (2) metamodel (M2) layer, (3) model (M1) layer, and (4) instance (M0) layer. The previously referred complementary standards, which MDA requires, are: (1) Meta-Object Facility (MOF), (2) Unified Modeling Language (UML), (3) and XML Metadata Interchange (XMI). Once again, by looking at Figure 2, one can observe that on top of MDA's architecture is the OMG's Meta-Object Facility (MOF) meta-metamodel.

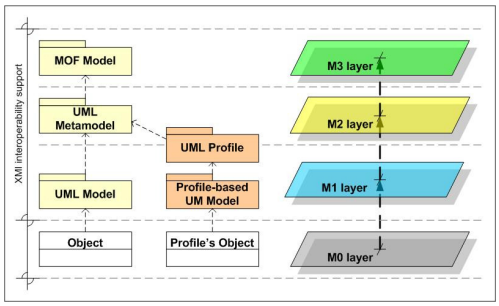


Figure 2 – OMG's MDA Metamodeling MOF-based Architecture

MOF defines an abstract language and framework for specifying, constructing and managing technology neutral metamodels, being the foundation for defining any modeling language, even allowing MOF's self-definition. MOF also defines a framework for implementing repositories that hold metadata described by metamodels. The main goal of having this four-level stratified structure with a common meta-metamodel is to support multiple models and metamodels, their extensibility, integration, and also generic model and metamodel management.

3.2 A MOF-based Ontology Definition Model (ODM)

The creation of an Ontology Definition Metamodel (ODM) [Djuric 2004] based on the MDA's principles and defined by the Meta-Object Facility (MOF) allows the definition of an Ontology UML Profile that enhances the ontological engineering creation process with commonly used UML notation and, simultaneously, allows the comparison between the MOF and OWL concepts in a straightforward manner. Another desired advantage, inherited from adopting MDA's as the base framework, is a bidirectional mapping between OWL and ODM, ODM and Ontology UML Profile (OUP), and even OUP and other UML profiles, through XSL Transformations.

The proposed ODM comprehends common ontology concepts, namely the ones present in OWL (the reference W3C standard), thus providing a starting point to this metamodel. To understand it, one has to clarify the main differences between metamodeling based on MDA, and the functional architecture used for ontology languages definition. While MDA has fixed and well-defined four-level architecture, the schema layer (RDFS) constitutes a non-standard and non-fixed-layer metamodeling architecture, making certain elements in model to have dual roles. This ambiguity has two major implications: (1) the impossibility of using modelers and, since it occurs in a low-level layer, (2) it propagates this problem to all the languages defined on top of it, namely OWL. In OWL DL, these problems are partially solved by introducing new modeling elements that are used for defining an ontology. However, since OWL Full allows unconstrained use of RDFS constructs it completely inherits RDFS's problems. Therefore, ODM for OWL Full cannot be modeled directly by using MOF. Having in mind the previous enunciated premises, ODM was primarily designed to support OWL DL, whereas OWL Full is only supported partially – concepts that do not break the MDA's fixed-layer architecture.

3.3 The Ontology UML Profile

In order to make use of graphical UML's modeling capabilities, an ODM must provide a corresponding UML Profile. This profile is required for graphical editing of ontologies using standard UML diagrams as well as benefiting of other advantages of using established UML CASE tools, such as Poseidon for UML, which was used in the practical case study in Section 5. Both UML models and ODM models are serialized in the XMI format so the two-way transformation between them can be done by using XSL Transformations. Since OWL is also based on a layered XML-based platform, as illustrated in Figure 1, it is possible to define another pair of XSL Transformations through which we can attain a bidirectional mapping between UML and OWL, having ODM as intermediary representation. This enables the definition of an ontology by simply using a common UML modeler and a XSLT processor, thus producing the required OWL file that can be further refined by using a specific Ontology CASE tool as Protégé. Having a solid theoretical background that supports the previously described ODM, which provides the foundation to define this OUP, the potential of using additional well-defined bidirectional XSL Transformations of OUP into another technology-specific UML Profiles is unlimited, enabling the usage of ontology engineering techniques in the design of other domains and vice versa.

As depicted in Figure 2 one can observe that UML profiles are a workaround to support new languages without introducing new elements at the metamodel layer (M2). Therefore, UML

profiles conceptually situate in the middle of M2 and M1 layers, since they are defined by M2 layer primitive constructs but used at M1 layer as M2 elements.

Figure 3 presents the OMG's UML superstructure specification [OMG 2005]. One can verify that formally UML and OWL are quite different. Some relevant observations [Hart 2004] to keep in mind when comparing them are: (1) UML model's content belongs mostly to the M1 layer and the M0 model just requires obeying to the M1 model's specification; (2) the linkage between a `Class` and an `Association` must be mediated by a `Property`; (3) the M1 model is built from structural elements; (4) both `Class` and `Association` are classifiers, having by inheritance a specific type; (5) a `Class` always has a `Property` which is the structural element that implements it; and (6) a `Property` may or may not be owned by one or more `Classes`. A `Property` owned by at least one `Class` is called navigable (memberEnd), whereas a `Property` not owned by any `Class` is called not navigable (ownedEnd).

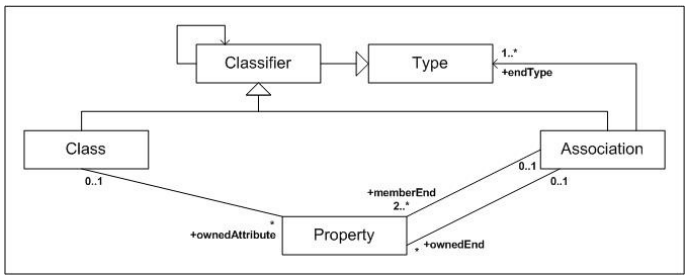


Figure 3 – OMG's UML Superstructure Specification (adapted from [OMG 2005])

This brief overview of the UML superstructure will be useful when analyzing the Ontology UML Profile (OUP) available at [Djuric 2004], which describes in detail the main concepts related to ODM and the Ontology UML Profile (OUP).

4 Ontology Engineering Tools

This section gives an overview of the tools used during the UML-to-OWL mapping process analyzed in the practical case study of transforming a UML domain model into an ontology.

4.1 XSLT Processor

The XSLT Processor initially used during the practical case study (to support simple XSL Transformation tests) was Altova XMLSpy [Altova 2006], which is a powerful tool to aid the creation of advanced XML-based applications. It revealed to be at the same time flexible enough to allow one to work with large and complex XML-based documents by providing multiple views and options that best suit to one's working preferences. This tool significantly increases productivity by allowing one to quickly develop higher-quality, standards-conformant XML-based applications. Its end-to-end development process' coverage range includes support for: (1) project management; (2) XML document and DTD creation and editing; (3) XPath development; (4) schema-aware XQuery and XSLT 1.0/2.0 development and debugging; (5) graphical WSDL and Web Services development; (6) SOAP development and debugging; (7) database interaction; and (8) VS.NET and Eclipse integration with code generation options.

Besides the Altova XMLSpy, the Microsoft .NET Framework [Microsoft 2006a] was also used to develop a simple program to process the XSL Transformations provided, in a later phase of the undertaken case study. The .NET Framework Class Library offers a rich set of features to support XSL Transformation, namely the *XslCompiledTransform* [Microsoft 2006a] class which implements the XSLT processor in this framework, supporting the W3C XSLT 1.0 recommendations. Moreover, it also provides an extremely flexible XPath-Navigator Class [Microsoft 2006b], which enables one to easily navigate and manipulate XML documents.

4.2 UML Modeler

Poseidon for UML [Gentleware 2006] offers a rich Graphical User Interface (GUI) and, besides the portability offered by the Java Virtual Machine (JVM), it has the advantage of providing a free, non-commercial, single-user Community Edition, which contributed for its mainstream adoption. Despite not being open source, this freely available version supports all UML diagrams, reverse engineering for Java code, and code generation. Poseidon is a sophisticated UML drawing program with a rich set of useful usability features and natively supports UML 2.0. Its main features are: (1) comprehensive suite of powerful features; (2) inherently simple functionality; and (3) Eclipse integration (for professional edition only).

4.3 Ontology IDE

Protégé [SMIOSC 2006] is a free, open-source platform supported by a growing user community, encompassing wide-ranging expertise areas. This application offers a coherent set of tools to design domain models and to build knowledge-based applications by using ontologies. Protégé implements a rich set of knowledge-modeling structures and actions that support the creation, visualization, and manipulation of ontologies in various representation formats. Protégé can be extensively customized and provides a good support for creating knowledge models and entering data for populating the knowledge-base. Through its plug-in architecture and a well-defined Java-based Application Programming Interface (API), Protégé can be easily extended for building more complex (or integrate with external third-party) knowledge-based tools and applications.

The Protégé tool supports two different approaches for modeling an ontology:

– **Protégé-Frames Editor:** allows one to easily manage frame-based ontologies, in accordance with the Open Knowledge Base Connectivity (OKBC) protocol (Figure 4).

– **Protégé-OWL Editor:** allows one to create ontologies for the Semantic Web, namely by using the W3C's Web Ontology Language (OWL), which allows deriving logical consequences through its well-defined formal semantics specification, i.e. to deduce facts not literally present in the ontology, but implicitly entailed by the constructs' semantics.

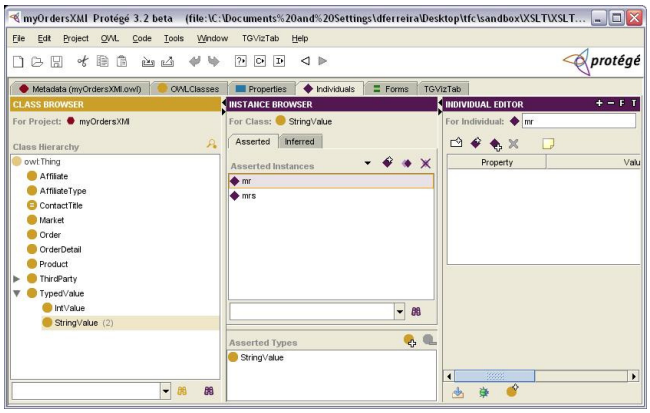


Figure 4 –Protégé-Frames Editor Screenshot

4.4 OWL Reasoners

FaCT++ [Tsarkov 2006] is the new generation of the well-known FaCT OWL-DL reasoner, using the established FaCT algorithms, but with a different internal architecture and presenting enhanced features. This application is implemented using C++ in order to create a more efficient software tool, and to maximize portability. It supports the standardized XML interface

to Description Logics systems developed by the DL Implementation Group (DIG), which allows it to run in server mode to support reasoning features required by other third-party tools.

Jena [HPLabs 2006a] is a Java framework for building Semantic Web applications. It provides a programmatic environment for RDF, RDFS and OWL, SPARQL and includes a rule-based inference engine [HPLabs 2006b].

5 A Practical Case Study

This section's main goal is to demonstrate the practical appliance of the concepts and techniques introduced in the previous sections through a concrete case study. This example, besides serving as a proof-of-concept, provided a good theoretical insight, consolidated by experience. The feedback gathered was vital to realize the pending idiosyncrasies in this expertise field of UML-to-OWL transformations through an Ontology Definition Metamodel (ODM) and Ontology UML Profile (OUP).

The case study consists in the analysis of the ontological engineering process involved in the creation of an ontology for a typical information system. The presented system, MyOrders, is an academic modeling exercise. Its domain model class diagram is illustrated in Figure 5.

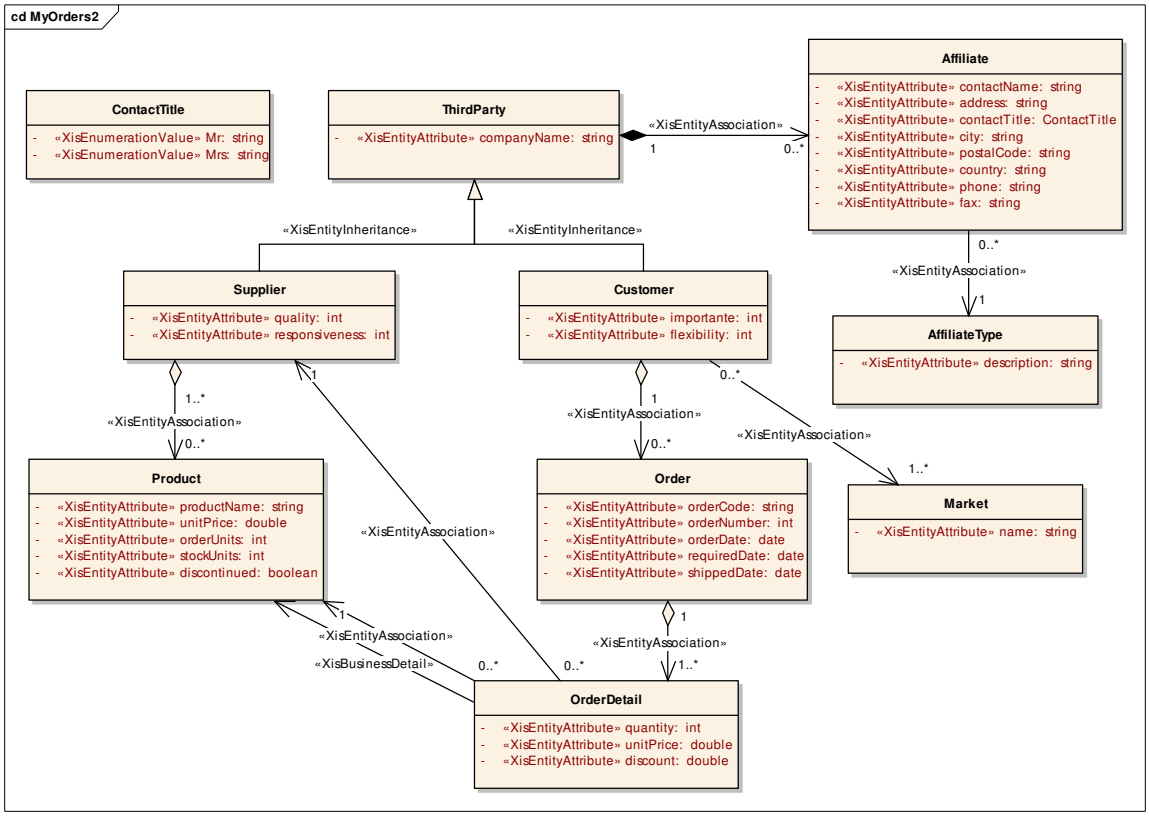


Figure 5 – MyOrders System's Domain Model Class Diagram

The main activities involved in the ontological engineering process, presented in Figure 6, were:

– The first step consisted in analyzing existing XLST processors, namely XML Spy and .NET Framework Class Library features. At this stage it was noticed that there were some problems when generating OWL files with XML Spy having as input the UMLtoOWL.xslt file available at [Gasevic 2006]. Simply, it generated header errors when importing it with

Protégé 3.2 beta. Therefore, it was implemented a simple application that generates parameterized OWL files by consuming as input the mentioned XSLT file with the mapping rules and a XMI file generated by Poseidon enhanced with OUP.

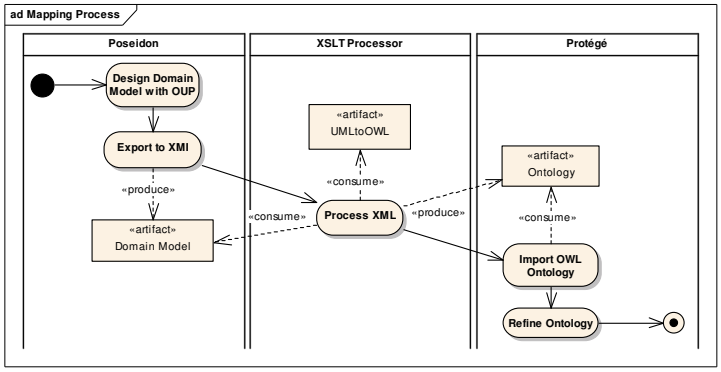


Figure 6 – UML-to-OWL Mapping Process

– The next step consisted in understanding existing OUP examples, to explore its true potentialities, since there are some significant paradigm differences when comparing MDA-based and Ontology-based approaches, as discussed in Section 3.

– Bearing in mind all the required concepts, the concrete modeling task started. This activity revealed most of the idiosyncrasies pointed out in the theoretical sections, namely the diagram cluttering (consult Figures 7 and 8) by explicitly defining UML attributes with the `ObjectProperty` and `DataTypeProperty` stereotypes.

– The process to achieve an .owl file consists simply in developing the model with OUP stereotypes in Poseidon, export it to XMI, run the application, and import the produced .owl file with Protégé. The main objective of this practical case study was achieved. However, there were serious difficulties to understand some unexpected errors related to the XSL Transformations provided (available at [Gasevic 2006]) which do not fully support OUP stereotypes, namely it offered a limited set of `DataType` properties that had to be enhanced to attain the case study goal, and some UML Associations' stereotypes that where not possible to transpose to Protégé, thus the produced ontology had to be further refined with Protégé to obtain the expected result.

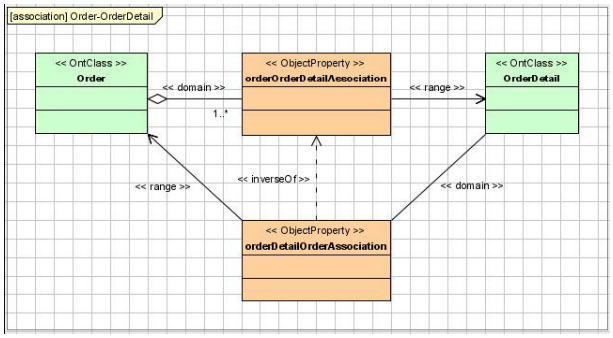


Figure 7 – Technique for Mapping an UML Association with OUP

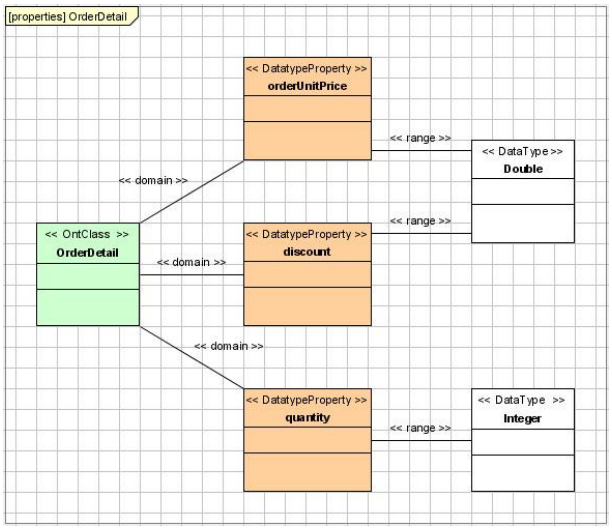


Figure 8 – Technique for Mapping an UML Attributes with OUP

6 Conclusions

An ontology describes the most important concepts and relationships in a particular domain, providing a vocabulary for it as well as a computer-based meaning's specification of terms used in the vocabulary. Ontologies range from taxonomies and classifications, database schema, to fully axiomatized theories. Recently, ontologies have been adopted in many business and scientific communities as a way of sharing, reuse, and process domain knowledge. Ontologies are now central to many applications, namely scientific knowledge portals, information management and integration systems, electronic commerce, and semantic web services.

To be widely adopted by the software development community and to succeed in real-world applications, knowledge engineering and ontology modeling must catch up with mainstream software trends, namely OMG's standards. The introduction of an Ontology Modeling Architecture and its related techniques will provide a good support and ease the integration between existing or forthcoming software tools and applications, adding value to both sides.

The Ontology Definition Metamodel (ODM) and Ontology UML Profile (OUP) analyzed in this paper are allegedly in accordance with the OMG's RFP initiative for ontology modeling, constituting a strong bet in this scientific research area, and can be considered as a part of the effort to specify a standard ontology metamodel. The possibilities of defining other AI metamodels in MOF should and probably will be explored in a near future. This means that MDA and MOF will most likely constitute the integration/bridge point between common and AI-related metamodels.

Despite of the enormous value of this contribution to both software engineering and ontological engineering areas, this proposal has several issues to solve, namely the adoption of graphical icons or forms to represent OUP stereotypes, instead the usage of color schemes to enhance diagram's readability, and to further develop the XSL Transformation rules to fully support OUP mapping to OWL and, eventually, bidirectional transformations.

Finally, the reader was presented with a practical proof-of-concept application of the presented theory and techniques for modeling a typical Information System by using XSL Transformations and CASE tools (Poseidon and Protégé). This case study allowed pointing out some idiosyncrasies of the theoretical techniques, emphasizing the need for future work.

7 References

Altova, "XML Editor for Modeling, Editing, Transforming, and Debugging XML Technologies", (2006). http://www.altova.com/products/xmlspy/xml_editor.html, [accessed on 04/06/2006].

Djuric, D., "MDA-based Ontology Infrastructure", *International Journal on Computer Science and Information Systems*, 1, 1, (2004), 91–116.

Gasevic, D., "UMLtoOWL: Converter from UML to OWL", (2006), http://afrodita.rcub.bg.ac.yu/~gasevic/projects/UMLtoOWL, [accessed on 02/06/2006].

Gentleware, "Poseidon for UML", (2006), http://gentleware.com/, [accessed on 06/06/2006].

Hart, L., et al., "OWL Full and UML 2.0 Compared", (2004).

Herman, I., Swick, R., Brickley, D., "Resource Description Framework (RDF)", (2006), http://www.w3.org/RDF, [accessed on 04/06/2006].

Herman, I., Hendler, J., "Web Ontology Language (OWL), (2006), http://www.w3.org/2004/OWL, [accessed on 04/06/2006].

HP Labs Semantic Web Programme and Open Source Community, "Jena: A Semantic Web Framework for Java", (2006), http://jena.sourceforge.net, [accessed on 04/06/2006].

HP Labs Semantic Web Programme and Open Source Community, "Jena 2 Inference Support", (2006), http://jena.sourceforge.net/inference, [accessed on 04/06/2006].

McGuinness, D., Harmelen, F., "OWL Web Ontology Language Overview", (2004). http://www.w3.org/TR/owl-features, [accessed on 06/06/2006].

Microsoft, "XSLT Transformations: System.Xml.Xsl.XslCompiledTransform Class", (2006), http://msdn2.microsoft.com/en-us/library/14689742.aspx, [accessed on 03/06/2006].

Microsoft, ".NET Framework Class Library: System.Xml.XPath.XPathNavigator Class", (2006), http://msdn.microsoft.com/library/default.asp?url=/library/en-us/cpref/html/frlrfsystemxmlxpathxpathnavigatorclasstopic.asp, [accessed on 03/06/2006].

Miller, E., "The Semantic Web", (2004), http://www.w3.org/2004/Talks/0120-semweb-umich, [accessed on 28/10/2006].

Object Management Group (OMG), "Unified Modeling Language: Superstructure", (2005). http://www.omg.org/docs/formal/05-07-04.pdf, [accessed on 06/06/2006], version 2.0.

Stanford Medical Informatics and Open Source Community, "Protégé, an Ontology Editor and Knowledge Acquisition System", (2006), http://protege.stanford.edu, [accessed on 06/06/2006].

Tsarkov, D., "FaCT OWL-DL external reasoner", (2006), http://owl.man.ac.uk/factplusplus, [accessed on 06/06/2006].

Compressão de Estruturas Multidimensionais de Dados com vista ao seu Armazenamento em Plataformas Móveis de Pequena Dimensão

Jorge Ribeiro [1], Orlando Belo [2]

1) Departamento de Ciências Básicas e Computacionais, Escola Superior de Gestão e Tecnologia
Viana do Castelo, Portugal
jribeiro@estg.ipvc.pt

2) Departamento de Informática, Escola de Engenharia, Universidade do Minho
Campus de Gualtar, Braga, Portugal
obelo@di.uminho.pt

Resumo

Nos dias de hoje, os gestores necessitam cada vez mais de informação sobre as várias actividades de negócio das suas organizações, para tomarem decisões estratégicas com base em toda a informação coerente das várias vertentes de negócio. O histórico da informação destas organizações é armazenado em repositórios de informação, sobre os quais os servidores de processamento analítico trabalham todas as combinações possíveis de análise, com base em "todas" as possíveis *queries* que os agentes de decisão costumam lançar. Com a evolução dos computadores de bolso, e a sua utilização em grande escala em termos de comunicações e acesso aos dados, os agentes de decisão têm a possibilidade de aceder à informação em qualquer altura e em qualquer lugar. Seria extremamente útil para estes agentes a utilização da sua informação analítica em dispositivos móveis. Este trabalho debruça-se, essencialmente, sobre o problema de processar, comprimir e migrar estruturas multidimensionais de dados para plataformas móveis de pequena dimensão, analisando os algoritmos Dwarf, Cubo Supresso, Cubo Prefixo e Quotient Cube.

Palavras chave: Sistemas de Processamento Analítico, Estruturas Multidimensionais de Dados, Processamento de Cubos e Compressão de Estruturas Multidimensionais.

1. Introdução

Os sistemas móveis "reduziram" drasticamente o espaço e o tempo entre dois lugares, entre duas pessoas. Hoje, disponibilizam-nos um leque de serviços tão vasto que muitas das coisas que fazemos diariamente, nos nossos postos de trabalho, podem ser feitas à distância, sem qualquer problema. Seguindo de muito perto este advento da mobilidade, os fabricantes de plataformas computacionais de pequena dimensão integraram vários dispositivos, que lhes garantem não só a sua autonomia em termos de serviço como também lhes asseguram novos meios de comunicação com outros dispositivos. Como consequência, é frequente encontrarmos já integrados em pequenas plataformas computacionais móveis não só os serviços usualmente disponíveis em plataformas fixas pessoais como também os serviços assegurados pelos telefones móveis actuais mais sofisticados. Os sistemas de processamento analítico, vulgarmente definidos como sistemas *OnLine Analytical Processing* (OLAP) [Codd et al. 1993], operam normalmente sobre a informação contida nos repositórios dos *Sistemas de Data Warehousing* (SDW) e alargam significativamente os horizontes de análise dos agentes de decisão das organizações, ao disponibilizarem meios muito expeditos de análise sobre as várias áreas de actividade das organizações. Estes sistemas assentam sobre a informação armazenada nos SDW, pre-calculando e processando todas as combinações dos agrupamentos de cada entidade dos SDW, e, por fim, materializando-as em *Estruturas Multidimensionais de Dados* (EMD) ou, como vulgarmente são designadas, em Cubo de Dados [Gray et al. 1997]. Os

agentes de decisão mais activos costumam ter o hábito de analisar todos os tipos de informação que têm à sua disposição, desenvolvendo processos de procura constante, em busca deste ou daquele indício, que lhes tragam algumas vantagens no desenvolvimento das suas actividades e no posicionamento das suas organizações nos mercados em que estão inseridas. Hoje são poucos os agentes de decisão que estão a trabalhar num único lugar, gerindo os negócios a partir daí. Todavia, apesar dos avanços enormes das tecnologias da informação e das comunicações, os agentes têm, por vezes, algumas dificuldades em aceder aos seus sistemas de informação e, em particular, às estruturas de dados de suporte à decisão que habitualmente utilizam nas suas actividades diárias. Isto porque, embora as limitações destes dispositivos móveis estejam a ser eliminadas dia após dia, estes continuam a ter algumas dificuldades quando nos reportamos às suas características de mobilidade, portabilidade e comunicação sem fios. É o caso da capacidade de armazenamento, a limitação da largura de banda, as reduzidas dimensões dos ecrãs para análise da informação, as quebras de comunicação frequentes, entre outros. Será possível, assim, que um dispositivo móvel tenha a capacidade de processar informação analítica volumosa? Haverá a possibilidade de armazenar informação analítica num dispositivo móvel em que à partida não seria possível o seu armazenamento dada a limitação física de espaço? Na realidade tudo isso é possível actualmente. Recentemente várias técnicas [Sismanis et al. 2002] [Wang et al. 2002] vocacionadas para o processamento e compressão de EMD foram apresentadas, conseguindo comprimir de forma eficiente este tipo de estruturas sem ter a necessidade de proceder à descompressão da informação quando as *queries* são solicitadas. Em vários cenários reais, consegue-se já aplicar técnicas de processamento analítico e de compressão de EMD com vista à sua exploração em dispositivos móveis, tendo em conta as várias limitações associadas a estes dispositivos, relativamente ao armazenamento e consequentemente visualização da informação que a priori não seria possível ser nele armazenada. Desta maneira, os gestores de negócio conseguem aceder a informação analítica para a tomada de decisões estratégicas, tornando-os, deste modo, independentes das estações de trabalho a que estavam habituados para analisarem a informação e abrindo novas formas de visualização da informação de que necessitam, de forma simples e rápida. Neste trabalho, apresentamos o POCKET CUBES, um sistema protótipo especificamente orientado para o processamento analítico e compressão de EMD, utilizando algoritmos como Cubo Supresso e Cubo Prefixo. Aqui, demonstramos que é possível reduzir o espaço de armazenamento em cerca de 70% de maneira a conseguir armazenar uma EMD em plataformas móveis (ou não) que à partida não teriam capacidade de a armazenar.

2. Processamento e Compressão de Estruturas Analíticas

O processamento de EMD é, em regra geral, uma tarefa computacional bastante difícil, uma vez que, a partir da informação armazenada nos repositórios analíticos, são processadas todas as combinações possíveis de agregação entre as várias dimensões de análise das suas estruturas de dados [Kimball et al. 1998]. Vários estudos sobre o processamento de EMD [Harinarayan et al. 1996] [Han et al. 2001] [Xin et al 2003] foram apresentados na área do processamento de EMD, especialmente sobre a adopção de metodologias para a selecção de partições de EMD a serem processadas seguindo uma determinada orientação: Top-Down [Zhao et al. 1997] e Bottom-up Computation [Beyer & Ramakrishan 1999]. Para analisar essa dificuldade do processamento, consideremos genericamente uma relação R com n atributos de dimensão (D1, D2, ..., Dn) e um atributo de medida M. A EMD processada a partir desta relação R em todas as dimensões, aplicando o operador de agrupamento (GROUP-BY) irá gerar 2n agrupamentos distintos constituindo desta forma a EMD final [Gray et al. 1997]. Consideremos, também, a relação R (A,B,C,M) da tabela 1 como exemplo de uma simples relação (base) típica em repositórios analíticos. Esta relação R tem três dimensões {A}, {B}, {C} e uma medida M, sendo adoptada a função SOMA como função de agregação. Para facilitar a análise de resultados com designações reais, podemos considerar que as dimensões {A}, {B} e {C} correspondem respectivamente, ao Dia da semana, aos Produtos e às Lojas, enquanto que o atributo de medida

corresponde ao total de Vendas efectuado ao longo do tempo pelas várias filiais de uma hipotética organização.

	A	B	C	M
t1	1	8	7	6
t2	1	3	7	12
t3	2	8	3	9

(a)

ou

DIAS	Produtos	Lojas	Vendas
Sabado	Computador	Lisboa	6
Sabado	Teclado	Lisboa	12
Domingo	Computador	Braga	9

(b)

Tabela 1 – Tabela base de um *Data Warehouse*

Esta relação R contém três tuplos ou células (t1, t2 e t3) e irá processar 23 agrupamentos (figura 1 (a)), de maneira a garantir todas as possibilidades de análise da informação. Cada aplicação do operador de agrupamento corresponde a um conjunto de tuplos (ou células), podendo ser também descritos como tuplos sobre as dimensões. As células do processamento da EMD são apresentadas na figura 1 (a), sendo a interligação das células de cada agrupamento apresentadas na figura 1 (d). O símbolo "*", como valor de uma dimensão, denota a generalização do valor da dimensão, ou seja, um valor especial para denotar o valor "ALL" em que qualquer valor pode ser assumido. Para exemplificar basicamente estas operações, consideremos a figura 1 (d) por exemplo que, a partir da célula (1,8,7:6) se efectua a operação de refinamento (aplicando o operador de agrupamento na dimensão {C}) para a célula (1,8,*:6), e a partir deste para (1,*,*:18) aplicando o operador de agrupamento na dimensão {B}. Posteriormente a agregação pode ser efectuada para a célula (1,3,*:12), seleccionando a dimensão {C} no operador de agrupamento. Esta navegação sobre os dados e sobre os agrupamentos, pode ser facilmente visualizada nas figuras 1 (b) e na figura 1 (d) . A implementação realizada do processamento da EMD está apresentada na figura 1 (a), em que cada uma das células apresentadas pertence a um sub-agrupamento, sendo o valor da coluna "ID-Agrupamento" a identificação do sub-agrupamento correspondente à célula. O valor desta coluna tem uma correspondente relação com a dimensão da tabela 1, sendo descrita esta equivalência na figura 1 (c). Esta numeração deve-se ao facto do processamento das dimensões (independentemente das designações a atribuir a cada uma delas) ter de ser numerado consoante a ordem do processamento do algoritmo na orientação da selecção das dimensões a serem particionadas. Uma das particularidades do processamento analítico é o facto de, através da interligação das várias células da EMD, se poder formar o denominado reticulado de células da EMD (figura 1 (d)), vulgarmente designado por "cube lattice". A interligação das células dos vários agrupamentos permitem semanticamente executar as operações de agrupamento ou refinamento entre as células através de algum atributo de dimensão necessárias para as ferramentas de processamento analítico. Ao analisar o reticulado das células da EMD e tendo em atenção o aumento do número de células interligadas entre si, surgem entre estas ligações algumas características peculiares que merecem ser analisadas. Uma delas corresponde à existência de várias células com o mesmo valor agregado e a outra refere-se à redundância semântica que ocorre nas interligações entre as várias células. Esta última característica é aquela pela qual os algoritmos mais recentes [Sismanis et al. 2002] [Wang et al. 2002] de processamento e compressão de EMD dão especial atenção, tentando conseguir aglomerar células com partilha de atributos de dimensão, por forma a armazenar apenas alguns tuplos e assim, reduzir o espaço de armazenamento da EMD. Esta característica da redundância entre as células da EMD pode ser dividida em dois tipos: redundância Prefixa e redundância Sufixa [Sismanis et al. 2002]. A generalidade dos algoritmos Dwarf, Cubo Supresso e Quotient Cube exploram o fenómeno da redundância Prefixa e Sufixa de modo eficiente, tentando criar estruturas com referências entre as células da EMD de maneira, a armazenar fisicamente menos células da EMD, mantendo no entanto, todos os requisitos funcionais das ferramentas OLAP. Na maioria dos casos, estes repositórios analíticos (DW) armazenam grandes volumes de dados e como consequência, o processamento de todos os agrupamentos é na generalidade dos casos muito custoso, quer em termos computacionais, quer em termos de espaço de armazenamento. Neste caso, em particular, a geração dos vários agrupamentos ocupam muito espaço extra e uma das formas de

minimizar estas dificuldades é a utilização de técnicas de optimização, que podem ser caracterizadas resumidamente em: vistas materializadas [Gupta & Mumick 2005], utilização de técnicas de indexação, como é o caso das Cube Forest [Johnson & Shasha 1997], Cubetrees [Roussopoulos et al. 1997], QC-Trees [Lackshmanan et al. 2003] e DC-Trees [Ester et al. 2000], processamento dos dados de forma paralela e distribuída e as técnicas de compressão de dados [Sismanis et al. 2002] [Wang et al. 2002] [Lackshaman et al. 2002].

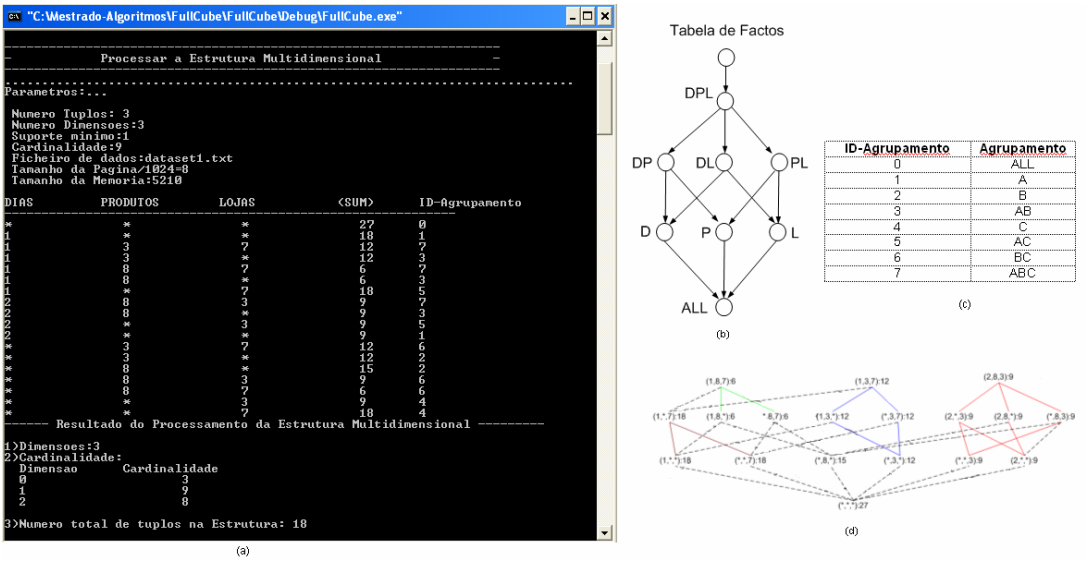


Figura 1 – Resultado do processamento da EMD a partir do exemplo da Tabela 1

As técnicas de redução de dados tentam adoptar estratégias de redução de dados, como por exemplo o somatório dos dados ou modelos de aproximação de respostas às *queries*, seleccionando estimativas de valores para disponibilizarem rapidamente a informação. Estes estudos centram-se no desenvolvimento de técnicas estatísticas em que produzem somatórios genéricos através da representação de esquemas matemáticos (matrizes multidimensionais), de maneira a armazenar os valores e aplicar técnicas matemáticas sobre eles, de forma a produzir representações reduzidas dos espaços multidimensionais. Contudo este tipo de técnicas de redução tem pouca aplicação em ambientes de redução de EMD, uma vez que os agentes de decisão necessitam de respostas exactas às *queries* invocadas e não de aproximações. Por esta razão, focaremos o nosso trabalho nas técnicas de processamento e compressão de EMD. Neste sentido, o estudo de Sismanis et al. (2002) apresenta uma estrutura denominada por "Dwarf" (Figura 2 (a)), para armazenar a EMD eliminando os dois tipos de redundância encontrados entre as células da EMD. Conseguem criar uma estrutura altamente compressa de durante o processamento da EMD. O trabalho desenvolvido por Long-gang et al. (2004) (Figura 2 (b)) verifica que embora a estrutura Dwarf elimine tanto a redundância Prefixa como a Sufixa, na realidade em alguns casos não se consegue eliminar totalmente a redundância Sufixa, criando adicionalmente nodos redundantes. Wang et al. (2002) apresentam o Cubo Supresso, baseando-se no conceito Tuplo Base Simples "Base Single Tuple (BST)", suprimindo tuplos da EMD num único tuplo representativo. Conseguem com esta técnica evitar o armazenamento de todos os tuplos da EMD. Neste estudo são apresentados também além do Cubo Supresso (BST - Cubo Supresso), o Cubo Mínimo, o BU - Cubo Supresso e o BU - Cubo Supresso Reordenado. Feng et al. (2004) melhoraram o Cubo Supresso detectando que o mesmo continuava a não eliminar totalmente a redundância Prefixa, propondo uma nova estrutura de armazenamento, partilhando os prefixos entre os tuplos, denominando este novo algoritmo de Cubo Prefixo ("PrefixCube") (figura 2 (c)).

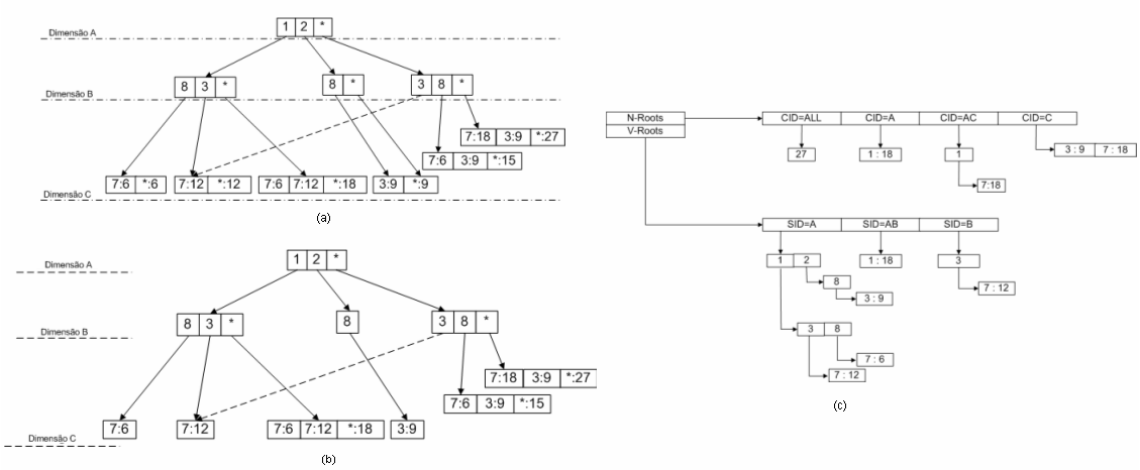


Figura 2 – Resultado dos algoritmos: Dwarf, Dwarf Supresso e Cubo Prefixo

Lakshmanan et al. (2002), baseando-se na ideia de agrupar células com o mesmo valor agregado em classes de células, apresentam o "Quotient Cube". Neste algoritmo de processamento e compressão de EMD, em vez de, armazenar todas as células da EMD apenas armazena os tuplos limite superior e inferior de cada uma das classes. O mesmo autor num outro estudo [Lakshmanan et al. 2003] apresenta uma estrutura denominada de QC-Trees, para implementar o algoritmo QC.

3. Desempenho dos Algoritmos de Processamento e de Compressão

Neste trabalho implementámos os algoritmos BU - Cubo Supresso e Cubo Prefixo para processamento de EMD. Para este efeito utilizamos vários conjuntos de dados caracterizados, essencialmente, em dois tipos: dados sintéticos e dados reais, tendo sido utilizada a SOMA como função de agregação. Todas as experiências foram realizadas em máquinas com processador Pentium IV 3Ghz com 512 MB DDR RAM, disco de 40 GB com 72 RPM, sistema operativo Windows XP Profissional e o conjunto de dados lido a partir de um ficheiro. Os algoritmos foram implementados em Visual C++ 6.0. Os resultados obtidos para a análise do rácio de compressão de EMD estão apresentados na figura 3 (a), (b) e (c). Como podemos analisar, a eficiência do algoritmo Cubo Prefixo face ao BU - Cubo Supresso é, em média, cerca de 25.5% para conjuntos de dados reais. Por outro lado, com o aumento da densidade do conjunto de dados, a taxa de compressão do algoritmo BU – Cubo Supresso aumenta, ao contrário do obtido pelo algoritmo Cubo Prefixo, que efectivamente diminuiu. A distribuição Zipf [Zipf 1949] foi utilizada para criar conjuntos de dados de maneira a analisar a dispersão dos dados e como podemos constatar, assim que o factor Zipf aumenta, o Cubo Prefixo tem melhor desempenho no rácio de compressão sobre o BU - Cubo Supresso. Em termos gerais, o algoritmo Cubo Prefixo é em média cerca de 30% melhor que o BU – Cubo Supresso. Outra observação que se pode analisar é o facto da curva do Cubo Prefixo ser regular (sem grandes variações), indicando que o Cubo Prefixo não ser sensível a variações na dispersão dos dados. Para a análise do tempo de processamento e compressão de cada um dos algoritmos (figura 3 (d) (e) e (f)), e embora consideremos os resultados obtidos como satisfatórios, temos consciência de que, se as características do equipamento fossem melhores em termos de processador e em termos de memória e rapidez de armazenamento, conseguiríamos obter melhores resultados. Como podemos analisar nessas figuras, para conjuntos com poucas dimensões, o tempo de processamento é bastante inferior do que o obtido para conjuntos com maior número de dimensões. Por outro lado, embora a maior parte dos conjuntos o algoritmo Cubo Prefixo seja mais eficiente em termos de tempo de processamento comparativamente com o BU – Cubo Supresso, para conjuntos de dados com menor número de dimensões, o algoritmo BU – Cubo Supresso é mais eficiente, ocorrendo o inverso com o aumento do número de

dimensões. A razão prende-se com o facto do algoritmo Cubo Prefixo despender mais tempo na análise de prefixos comuns entre os tuplos por ter menos dimensões, a eficiência da análise da partilha de prefixos comuns no tempo de armazenamento não se nota, notando-se mais à medida que o número de dimensões aumenta no conjunto de dados.

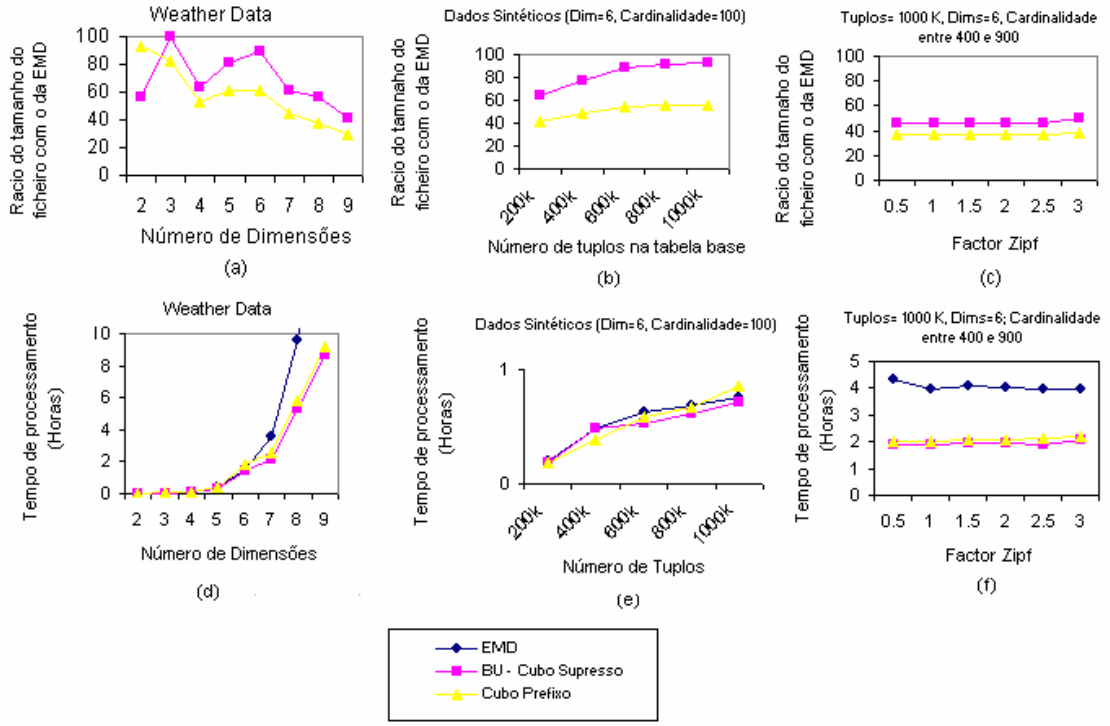


Figura 3 – Avaliação do rácio de compressão e tempo de processamento de EMD

Analisando a globalidade dos algoritmos de processamento e compressão de estruturas multidimensionais abordados neste nosso estudo e embora não tivéssemos implementado a estrutura QC-Tree (para implementação do algoritmo Quotient Cube), consideramos que em termos de rácio de redução esta estrutura deve superar os resultados obtidos por qualquer um dos outros algoritmos de compressão abordados, mesmo quando comparada com os resultados obtidos pelo algoritmo Cubo Prefixo. Isto porque a estrutura QC-Tree segue a semântica do algoritmo Quotient Cube, aglomerando células da estrutura multidimensional em classes necessitando apenas de armazenar as células limite inferior e superior de cada uma das classes, enquanto que os outros algoritmos necessitam de armazenar mais células. Por exemplo, o algoritmo Dwarf necessita de armazenar todas as células ao passo que os algoritmos BU – Cubo Supresso e Cubo Prefixo embora suprimam alguns tuplos (não necessitando de serem armazenados) e optando ou não pela técnica de partilha de prefixos entre as células, necessitam de armazenar sempre mais células do que a estrutura QC-Tree. Manter nos dias de hoje uma EMD actualizada é um requisito em ambientes de processamento analítico, uma vez que é necessário disponibilizar a informação mais recente aos agentes de decisão. Na generalidade dos casos, o volume da informação nas EMD é elevado e as actualizações são normalmente frequentes e volumosas. Vários estudos [Mumick et al. 1997] definem e debruçam-se sobre esta problemática da manutenção incremental da informação em repositórios analíticos. Todos eles têm o objectivo de evitar o reprocessamento de toda a informação analítica, apresentando metodologias e esquemas de manutenção de maneira eficiente e transparente para os agentes de decisão. O processo de manutenção de vistas também é um problema bastante estudado, havendo alguns estudos realizados [Hurtado et al. 1999]. Cada um dos algoritmos apresentados em epígrafe, como o Dwarf, o Cubo Supresso e o Quotient Cube, para actualizarem a informação contida nas suas estruturas apresentam alguns métodos eficientes para o efeito.

4. Um Caso de Aplicação Prática

4.1 Hardware Utilizado e Conjunto de Dados

O sistema POCKET CUBES permite processar e comprimir estruturas multidimensionais e, posteriormente, migrá-las para dispositivos de armazenamento com restrições de espaço, passíveis de serem utilizados em plataformas móveis de pequena dimensão. Para a análise do desempenho deste protótipo e tendo em consideração a restrição do limite de espaço de armazenamento da informação em dispositivos móveis, utilizamos um cartão *Secure Card* de 1GB Ultra II com capacidade de transferência de 9 Mbps. Para efectuar o processamento da EMD recorremos a uma estação de trabalho tal como a apresentada na secção 2.1 e com uma base de dados em Microsoft SQL Server 2000. As restrições que se colocam no caso em estudo referem-se ao espaço máximo de armazenamento, ao tempo disponível para se efectuar o processamento da EMD e ao tempo de transferência da informação para dispositivos de armazenamento. Assim, a restrição de espaço de armazenamento do dispositivo foi definido como 1 GB, a velocidade (média) de 9 Mbps referente à transferência da informação para o dispositivo de armazenamento e o tempo disponível para o processamento e compressão da EMD foi de 11 Horas. O conjunto de dados deste caso de estudo contém um milhão de registos referentes às Vendas dos últimos cinco anos registadas por uma organização que se dedica ao comércio de artigos e serviços de Optometria, dispondo de cinco lojas e trinta vendedores. Esta organização possui 5869 artigos caracterizados por dez categorias e 128 marcas, estando registados na base de dados neste momento 9202 clientes. Os requisitos de análise da informação por parte dos gestores desta organização centram-se no somatório das Vendas por períodos (ano e mês), por marca de artigo, pelo artigo, pela categoria dos artigos, por vendedor, por loja e finalmente por cliente.

4.2 Processamento e Compressão

O conjunto de dados estava armazenado, tal como referimos, numa base de dados *Microsoft SQL Server 2000* e o fluxo da componente de processamento do POCKET CUBES foi implementado com o recurso a Serviços de Transformação de Dados (*Data Transformation Services* (DTS)), também do *SQL Server 2000*. Este fluxo de tarefas foi dividido em vários de menor dimensão, cada um associado a um processamento específico, nomeadamente: análise das restrições do tempo disponível e do espaço máximo de armazenamento e para a selecção das dimensões e o número de registos para a selecção do conjunto de dados. Depois desta primeira tarefa, de maneira a seleccionar o mais adequado para garantir o processamento com base nas restrições e com base nas restrições definidas, o sistema com base no número de dimensões e no número de registos do conjunto de dados, analisa um conjunto de estimativas em termos de espaço de armazenamento, de tempo de processamento para os algoritmos BU Cubo Supresso e Cubo Prefixo, e tempo de transferência da informação para o dispositivo de armazenamento. Para este exemplo, os valores estimados para o conjunto de dados foram de 5.3 horas para o tempo de processamento do Cubo BU-Supresso e 5.81 Horas para o Cubo Prefixo. Em termos de espaço de armazenamento estimado foi de 795.2 MB para o BU - Cubo Supresso e 535.54 MB para o Cubo Prefixo. Com base nos valores estimados para o armazenamento e considerando a restrição da velocidade média de transferência da EMD para o dispositivo, os valores foram de 20 minutos para o BU-Cubo Supresso e de 14 minutos para o Cubo Prefixo. Com base nestes valores da velocidade de transferência, a estimativa para o tempo de processamento (e transferência) dos algoritmos BU - Cubo Supresso e Cubo Prefixo é de 5.5 Horas e 5.95 Horas, respectivamente. Desta maneira, ambos os algoritmos conseguem processar e transferir a EMD para o dispositivo de armazenamento dentro da restrição de tempo disponível. Em termos de limite máximo para o espaço de armazenamento, foi estipulado em 1 GB, donde ambos os algoritmos (com base nas estimativas) conseguem garantir esta restrição. Caso a restrição de espaço máximo de armazenamento fosse por exemplo de 700 MB, então o fluxo de decisão iria seleccionar o algoritmo Cubo Prefixo, uma vez que o algoritmo BU - Cubo Supresso (com base nas estimativas) não conseguiria garantir a restrição de espaço disponível.

Tal como referimos no estudo dos algoritmos de processamento e compressão Cubo Supresso e Cubo Prefixo, na selecção de um determinado agrupamento, o protótipo implementa o método de "Expansão", extraindo os tuplos supressos a partir dos tuplos BST. Na eventualidade de nenhum dos algoritmos conseguir garantir ambas as restrições, o protótipo contempla um "plano de contingência" para definir um conjunto de possíveis novas dimensões e um novo número de tuplos para construir um novo conjunto de dados a ser processado. Este "plano" é definido pelos administradores do sistema tendo como base as necessidades dos agentes de decisão. Por exemplo, se não for possível processar a EMD na sua totalidade então, os agentes poderão delinear algumas alternativas, como por exemplo, as dimensões vendas por ano, mês e por vendedor. Desta maneira, o conjunto de dados será mais pequeno o mesmo acontecendo com a EMD. Assim, o fluxo de decisão e com base nas restrições definidas, o algoritmo seleccionado foi o Cubo Supresso, demorando cerca de quatro horas e trinta minutos em que a EMD ocupou cerca de 580 MB. Opcionalmente, com o objectivo de analisar o desempenho do protótipo, procedemos ao processamento da EMD com o algoritmo Cubo Prefixo e um outro teste efectuando, apenas, o processamento da EMD sem proceder à compressão. Para o processamento do Cubo Prefixo, obtivemos um espaço total de 714 MB (748.714.920 bytes) demorando cerca de 5 Horas o seu processamento. Quanto ao processamento da EMD, sem efectuar a tarefa de compressão obtivemos um espaço total de 2.230 MB demorando cerca de 4 horas e 94 minutos. Com efeito, o rácio de compressão de 0.35 obtido pelo algoritmo Cubo Supresso e um rácio de 0.24 obtido pelo algoritmo Cubo Prefixo. O algoritmo que conseguiu processar e comprimir a EMD, ocupando um menor espaço de armazenamento, foi o Cubo Prefixo. Com base nestes resultados, o protótipo foi capaz de processar e comprimir a EMD garantindo as restrições de tempo disponível para o processamento e o espaço máximo de armazenamento. Com efeito, podemos observar os seguintes pontos: o protótipo foi capaz de processar e comprimir a EMD, garantindo as restrições pré-definidas. Os valores estimados estiveram próximos dos realmente obtidos, após o processamento, considerando que o valor de cerca de 37% como margem de erro poder ser aceitável com base nas condições do equipamento de suporte. Conseguiu-se reduzir a EMD em cerca de 75%, armazenando-a em dispositivos que à partida, não teriam capacidade de o fazer.

4.3 Visualização da Informação

Desejavelmente, a visualização e interrogação da informação analítica são uma necessidade cada vez maior por parte dos gestores. Todavia, a implementação deste protótipo apenas contempla a visualização dos registos de um agrupamento previamente seleccionado, não criando métodos para explorar a informação, tanto a nível dos vários tipos de *query*, como, disponibilizar operações de navegação sobre a informação. Como referimos ao longo deste estudo, por um lado, este será um dos trabalhos futuros e, por outro, as operações de navegação podem ser disponibilizadas por ferramentas de processamento analítico que poderiam importar os vários ficheiros processados pelo protótipo e fornecer aos gestores meios de análise mais expedita. Na generalidade dos casos, os agentes de decisão transportam na sua vida quotidiana computadores de bolso que lhes permitem aceder à informação da sua organização. Porém, dado o seu tamanho tão pequeno, limitam o potencial de visualizar toda a informação, obrigando em vários casos quando o volume de informação é enorme, a uma constante selecção de páginas de maneira a se conseguir visualizar toda a informação. Perante este facto, a visualização de grandes volumes de dados neste tipo de dispositivos deverá ser ponderada pela sua importância e necessidade dos agentes de decisão, sendo por vezes preferível a visualização de somatórios genéricos de informação em vez de analisar a informação com mais detalhe. Isto porque, à medida que se vai seleccionando outros agrupamentos com vários níveis de agregação, o número de registos aumenta. Para este caso de estudo e como referimos, o protótipo processou 256 agrupamentos, sendo disponibilizada a informação analítica (Figura 4) em binário ou em XML.

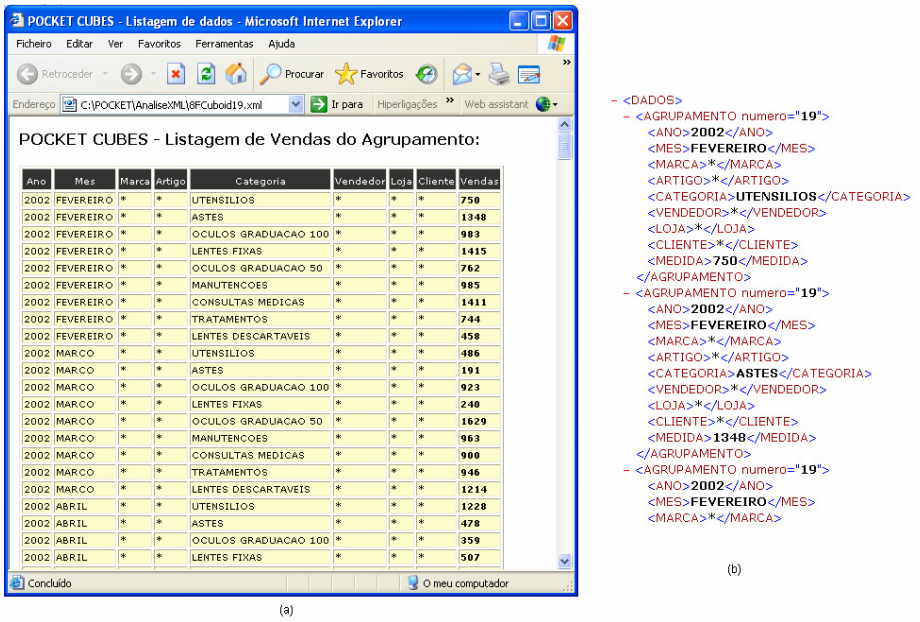


Figura 4 – Visualização da Informação da EMD para o caso de estudo

5. Conclusões

Os estudos que foram desenvolvidos durante a realização deste trabalho, sobre os algoritmos de processamento e compressão de estruturas multidimensionais de dados, permite-nos, agora, ter uma ideia concreta acerca da forma como migrar repositórios analíticos para plataformas móveis de pequena dimensão. Este problema de migração já não é algo que nos levanta grandes dificuldades se tivermos em conta os algoritmos estudados e as suas características de aplicação. Apesar disso, sabemos que não podemos migrar toda e qualquer estrutura multidimensional devido a problemas de espaço de armazenamento, tempo de processamento e tempo de transferência da informação para dispositivos de armazenamento com algumas limitações em termos de velocidade de transferência da informação. Com o conhecimento adquirido neste trabalho de dissertação, podemos, com certeza, preparar e migrar com sucesso estruturas multidimensionais de pequena e médio porte para plataformas móveis de pequena dimensão. Desta forma, conseguimos garantir o tão desejado objectivo de colocar num POCKET PC (ou noutro dispositivo semelhante) uma estrutura de suporte à decisão a ser manipulada, eventualmente, por uma folha de cálculo, disponibilizando desta maneira aos agentes de decisão a Em suma, neste nosso estudo foram abordados áreas e temas de grande importância para quem implementa, ou pensa implementar, uma aplicação de processamento analítico numa plataforma computacional de pequena dimensão. Entendemos que os assuntos abordados e o trabalho desenvolvido abrem caminho a novos projectos, podendo desta maneira, disponibilizar aos agentes de decisão, ferramentas de visualização de informação analítica directamente nos seus computadores de bolso.

6. Referências

Beyer K. e R. Ramakrishanan Bottom-up computation of sparse and iceberg cubes, in SIGMOD, 1999.

Codd E.F, S.B. Codd e C.T. Salley. Providing OLAP (OnLine Analytical Processing) to User-Analysts: An IT Mandate. Technical Report, E.F. Codd and Associates, 1993.

Ester M., J. Kohlhammer, e H.-P. Kriegel. The DC-Tree: A Fully Dynamic Index Structure for Data Warehouses. In Proc. of ICDE, páginas: 379–388, San Diego, California, 2000.

Gray J. e S. Chaudhuri, et al. Data Cube: A Relational Aggregation Operator Generalizing Group-by, Cross-tables and Sub-Totals, in Data Mining and Knowledge Discovery, Vol.1, No.1, páginas: 29-53, 1997.

Gupta H. e I. Mumick. Selection of views to materialize in a data warehouse. In IEEE Transcations on Knowledge and Data Engineering, vol. 17, páginas 24-43, Janeiro 2005.

Han J., J. Pei, G. Dong, e K. Wang. Eficient computation of iceberg cubes with complex measures. In SIGMOD'01. Harinarayan, V., Rajaraman, A., e Ullman, J. Implementing Data Cube Efficiently. In Proc. Of the 1996 ACM-SIGMOD Conference, 1996.

Hurtado C., A. O. Mendelzon e A. Vaisman. Updating OLAP dimensions. In DOLAP'99, 1999.

Johnson T. e D. Shasha. Some Approaches to Index Design for Cube Forests. Data Engineering Bulletin, 20(1), March 1997.

Kimball, R. L. Reeves et al. The Data Warehouse Lifecycle ToolKit – Expert Methods for Designing, Developing, and Deploying Data Warehouses. Jonh Wiley & Sons, 1998.

Mueller J. P.. Visual C++ 6.0 - From the ground up – The accelerated Track for Professional Programmers, 2nd ed., McGraw-Hill, 1998.

Mumick I. D. Quass e B. Mumick. Maintenance of Data Cubes and Summary Tables in a Warehouse, In Proc. ACM SIGMOD, Tucson, páginas: 100-111, 1997.

Roussopoulos N., Y. Kotidis, e M. Roussopoulos. Cubetree: Organization of and Bulk Incremental Updates on the Data Cube. In SIGMOD 1997, Tucson.

Sismanis Y., A. Deligiannakis, N. Roussopoulos e Y. Kotidis. Dwarf: Shrinking the PetaCube, In Proc. Of ACM SIGMOD, páginas: 464-475, Madison, Wisconsin, USA, 2002.

Wang W. J. Feng, H. Lu and J. Yu. "Condensed Cube: An Effective Approach to Reducing Data Cube Size". In Proc. Of ICDE, páginas. 155-165, San Jose, California, USA, 2002.

Zhao Y., P.M.Deshpande and J.F. Naughton. An array-based algorithm for simultaneous multidimensional aggregates. In SIGMOD'97, 1997.

Zipf G. K.. Human behaviour and the principle of least effort. Addison-Wesley, MA, 1949. Xin D. J. Han, X. Li e B. Wah.Star-cubing: computing iceberg cubes by top-down and bottomup integrations. In VLDB, 2003.

A Utilização de Sistemas *Groupware*/*Workflow* para Suportar o Desenvolvimento de Software em Equipa

Marco Aurélio Martins Ribeiro[1], José Luís Mota Pereira[2]

1) Finibanco, Porto, Portugal
marco.ribeiro@finibanco.pt

2) Universidade do Minho, Guimarães, Portugal
jlmp@dsi.uminho.pt

Resumo

Este artigo pretende fazer um mapeamento entre as necessidades sentidas pelas equipas de desenvolvimento de software e as funcionalidades oferecidas pelos actuais sistemas *workflow* e *groupware* (também designados colaborativos). A ideia é identificar, para cada tipologia possível de trabalho em equipa (nomeadamente, quanto à distribuição dos seus elementos no espaço e no tempo) e abordagem ao desenvolvimento de software (distinguindo entre metodologias ágeis e clássicas), quais as principais ferramentas *groupware*/*workflow* que poderão ser utilizadas para suportar devidamente o seu trabalho.

Palavras chave: Engenharia de Software, Desenvolvimento de Software em Equipa, Sistemas *Groupware*, Sistemas *Workflow*

1. Introdução

O software transformou profundamente a forma como vivemos e trabalhamos. Desenvolver software de qualidade é hoje um desafio importante enfrentado diariamente por inúmeras equipas de desenvolvimento, espalhadas pelo mundo. A engenharia de software define um conjunto de processos, métodos e técnicas que, quando correctamente aplicados num projecto, possibilitam obter melhorias na qualidade dos produtos desenvolvidos.

Existem vários processos de desenvolvimento possíveis que definem abordagens e estratégias distintas para um projecto. A motivação para a adopção de um processo de software nasce da convicção que a qualidade do produto final está intimamente ligada ao seu processo de produção, logo a adopção de um processo adequado resultará num desenvolvimento mais eficiente e eficaz [Fugetta 2000].

Recentemente, têm vindo a surgir novas teorias que advogam que os ciclos de desenvolvimento modernos são hoje incompatíveis com processos demasiado rígidos e formais. Estas abordagens

defendem o uso de processos denominados ágeis, com ciclos de desenvolvimento mais curtos, menos formalismos e uma maior interacção entre a equipa e o cliente.

Apesar das duas correntes, clássica ou ágil, um processo de desenvolvimento caracteriza-se por um conjunto de actividades e produtos resultantes, cujo resultado final é o software [Sommerville 2001]. O processo de desenvolvimento de software define *quem* faz o *quê*, *quando* o faz e *como* o faz [Fayad 1997].

Genericamente, um processo de desenvolvimento de software é composto pelas seguintes actividades: análise de requisitos, concepção, implementação e validação. Um sistema de suporte ao desenvolvimento de software deve ser capaz de acomodar o trabalho realizado nestas quatro fases. Este sistema deve também ser capaz de suportar processos mais burocráticos, bem como processos ágeis de desenvolvimento.

A adaptação entre o sistema de suporte ao desenvolvimento de software e o tipo de processo adoptado consistirá num balanceamento dos pesos das componentes de *workflow* e *groupware*. Assim, se o processo adoptado é um processo mais pesado e formal, a componente *workflow* do sistema de suporte será preponderante, no entanto se o processo adoptado é um processo mais ágil então a componente principal do sistema de suporte será o *groupware*.

2. Equipas de desenvolvimento de software

O software é concebido por pessoas e estas constituem o factor mais importante para o sucesso de um projecto de software. Os projectos de software são executados por equipas de especialistas, com competências e experiências distintas. Uma abordagem individual já não é válida para desenvolver grandes sistemas de software com qualidade.

Nem todas as equipas são capazes de produzir software com a mesma qualidade. São vários os factores que determinam a eficácia da equipa. Entenda-se por eficácia da equipa a sua capacidade para desenvolver software de acordo com as especificações, dentro das restrições de tempo e recursos do projecto.

Deminors identifica quatro aspectos que influenciam a eficácia de uma equipa: dimensão, estrutura, composição e processo [Deminors 1997]:

- **Dimensão**. Compreende-se por dimensão da equipa o número de membros que a compõem. Normalmente, uma equipa pequena tende a ser mais eficaz do que uma equipa de maior dimensão porque os seus custos de coordenação são menores [Kraul et al. 1995], [Deminors 1997]. Um sistema de suporte ao desenvolvimento de software deve ser

adaptável a equipas de dimensões diferentes. Este sistema deve possuir funcionalidades de suporte ao trabalho colaborativo.

- **Estrutura**. A estrutura de uma equipa corresponde à sua organização interna, isto é, a forma como cada membro se posiciona em relação aos outros de acordo com critérios de fluxo de informação (estrutura de comunicação) e fluxo de autoridade (estrutura de tomada de decisão) [Deminors 1997]. A eficácia de uma dada estrutura de equipa depende da complexidade da tarefa, dos objectivos da equipa e do ambiente. Um sistema de apoio ao desenvolvimento de software deve ser capaz de se adaptar a diferentes estruturas e potenciar a comunicação directa. As ferramentas de *groupware* como, por exemplo as mensagens instantâneas ou o correio electrónico são boas escolhas para suportar a comunicação síncrona e assíncrona, respectivamente.
- **Composição**. A composição de uma equipa reflecte os factores que determinam o desempenho individual dos membros da equipa. Alguns destes factores são: as suas competências, a experiência, a formação e a estratégia de resolução de problemas [Deminors 1997]. Um sistema de apoio ao desenvolvimento de software deve potenciar a partilha e recolha do conhecimento, espalhados pelos vários membros. Se cada membro possui competências e especializações técnicas distintas é importante preservar as experiências de cada um, para que os outros membros possam dispor desse conhecimento e aplicá-lo. É também importante que face a um problema específico seja possível identificar e contactar da forma mais directa possível o membro mais preparado para o resolver.
- **Processo**. O processo adoptado por uma equipa determina a forma como as tarefas serão executadas. Ele constitui a abordagem adoptada pela equipa para resolver o problema. Processo é o conjunto de actividades a realizar e o critério de avaliação do sucesso da execução dessas mesmas actividades [Deminors 1997]. Um sistema de suporte ao desenvolvimento de software pode incluir um sistema de *workflow*. O sistema de *workflow* permite modelar e controlar a execução do processo de desenvolvimento de software. O sistema de *workflow* pode atribuir, em tempo de execução, as tarefas aos seus executores.

Nem todas as equipas de desenvolvimento de software são eficazes, no entanto uma boa equipa que funcione eficientemente, é um factor essencial para o sucesso do projecto. Apresentam-se na Tabela 1 um conjunto de aspectos que caracterizam aquilo que se entende por uma equipa eficaz e uma equipa ineficaz [Miguel 2003].

Equipa Eficaz	Equipa Ineficaz
Elevado desempenho e eficiência na tarefa	Desempenho pobre
Comportamento inovador/criativo	Comportamento passivo
Comprometimento com os objectivos do projecto	Baixo comprometimento com os objectivos do projecto
Objectivos profissionais dos membros da equipa coincidentes com os requisitos do projecto	Objectivos do projecto pouco claros e níveis de compromisso fluidos por parte dos elementos chave
Membros da equipa muito interdependentes com interfaces eficazes	Acções improdutivas, manipulação dos outros, sentimentos escondidos, evitar o conflito a todo o custo
Capacidade de resolução de conflitos, embora o conflito seja encorajado quando pode conduzir a resultados benéficos	Confusão, conflito, ineficiência
Comunicação eficaz	Sabotagem subtil, medo, desinteresse
Níveis de confiança elevados	Constituição de grupos de influência, conluios, isolamento de membros
Orientação pelos resultados	Letargia/falta de capacidade de resposta
Interesse em ser membro da equipa	
Elevados níveis de energia e entusiasmo	
Moral elevado	
Orientação para a mudança	

Tabela 1 - Características de eficácia/ineficácia das equipas

Tipos de equipas

As equipas de desenvolvimento de software podem assumir várias configurações ou tipologias. Estas podem ser classificadas, genericamente, numa matriz espaço/tempo (Tabela 2):

	Trabalho síncrono	**Trabalho assíncrono**
Espaço partilhado	Equipa co-localizada	Equipa co-localizada a trabalhar em turnos
Espaço não partilhado	Equipa distribuída a trabalhar ao mesmo tempo	Equipa distribuída a trabalhar em turnos

Tabela 2 – Classificação espaço/tempo das equipas de desenvolvimento de software

- **Equipa co-localizada** é a configuração mais usual numa equipa de desenvolvimento de software. Numa equipa co-localizada os seus vários elementos partilham um espaço físico de trabalho e trabalham todos nos mesmos momentos. A comunicação informal é frequente e facilitada pela proximidade física das pessoas. Este facto não deve ser desprezado, pois está demonstrado que a discussão informal, não planeada e *ad-hoc* com colegas é uma das formas mais eficazes de coordenação e colaboração [Kraul et al. 1995], [Grinter et al. 1999].

- **Equipa co-localizada a trabalhar em turnos** implica que os seus membros partilhem um espaço físico, não podendo trabalhar todos no mesmo horário. Na realidade, uma equipa com esta organização não é frequente, no entanto existem funções dentro de uma equipa de desenvolvimento de software, que têm que ser realizadas, fora das horas normais de trabalho da equipa. Funções como a administração de base de dados, é frequentemente executada fora do horário normal de trabalho com o objectivo de minimizar os períodos de indisponibilidade do ambiente de desenvolvimento. Neste tipo de organização de equipa, a comunicação informal entre os membros do mesmo turno possui um papel muito importante para a tomada de decisões e de coordenação do trabalho. É também habitual, o recurso a sistemas de mensagens assíncronas que facilitem a comunicação entre os membros que trabalham em horários diferentes.

- **Equipa distribuída a trabalhar ao mesmo tempo** possibilita que os seus membros se encontrem geograficamente dispersos, embora todos eles trabalhem no mesmo horário. A difusão de novos meios electrónicos de comunicação potencia e facilita a execução de trabalho remoto. Uma equipa distribuída potencia a capacidade de captação de recursos. Uma equipa geograficamente distribuída dificulta a comunicação informal que é extremamente importante em trabalhos intelectuais como o desenvolvimento de software [Grinter et al. 1999]. No entanto as ferramentas de *groupware*, como por exemplo: as mensagens instantâneas, a videoconferência, a edição conjunta, etc., podem facilitar e suprir em parte esta lacuna.

- **Equipa distribuída a trabalhar em turnos** reflecte uma situação em que os seus diversos membros se encontrem geograficamente distribuídos e que não trabalhem normalmente durante o mesmo horário. Com este tipo de organização é possível distribuir actividades diferentes, como por exemplo a análise e a codificação, por equipas geograficamente dispersas e a trabalhar em turnos não sobrepostos. A comunicação em tempo real é limitada devido à diferença horária de trabalho. O recurso a sistemas de comunicação assíncrona será mais frequente.

Um sistema de suporte ao desenvolvimento de software deve ser capaz de acomodar as necessidades de comunicação, colaboração e coordenação do trabalho das distintas organizações possíveis de uma equipa.

3. Requisitos do desenvolvimento de software em equipa

O desenvolvimento de software em equipa implica que o respectivo sistema de suporte seja capaz de acomodar duas vertentes: a do processo que determina a organização do trabalho; e do trabalho colaborativo que reúne as necessidades de comunicação e colaboração entre os membros da equipa.

Um sistema de suporte ao desenvolvimento de software baseado em tecnologias de *groupware* e *workflow* útil deverá suportar os seguintes requisitos:

Workflow	• Definição de diferentes processos • Controlo da execução do processo • Modificação do processo • Tratamento de situações de excepção do processo • Atribuição de tarefas aos membros da equipa • Listas de tarefas de cada membro • Consulta do estado actual de execução do processo • Definição de prazos para cada actividade do processo • Controlar desvios na execução das actividades
Groupware	• Repositório dos documentos do processo • Facilidades de comunicação • Suporte para comunicação formal e informal • Suporte para diferentes meios de comunicação • Suportar comunicação síncrona e assíncrona • Disponibilidade actual de cada membro da equipa • Trabalho colaborativo numa actividade • Apoio ao processo de decisão em grupo • Suporte para reuniões

Tabela 3 – Requisitos de um sistema de suporte ao desenvolvimento de software

Na tabela anterior, estes requisitos encontram-se mapeados para as componentes base do sistema, *workflow* e *groupware*. As características das tecnologias *workflow* e *groupware* são distintas e, portanto, adequam-se a requisitos distintos. Enquanto a tecnologia *workflow* possibilita suportar a parte do processo de desenvolvimento, a atribuição de tarefas e o controlo da sua execução, a tecnologia *groupware* facilita a interacção e a comunicação entre os membros da equipa.

4. Sistemas de Gestão de *Workflow*

A *Workflow Management Coalition* (WfMC), como organismo internacional que tem por missão criar normas universais que permitam a generalização do uso da tecnologia de *workflow*, define *workflow* como "*a automação de procedimentos em que documentos, informação ou tarefas são passadas entre participantes de um processo, respeitando um conjunto de regras que contribuem para atingir um objectivo de negócio*" [WfMC 1995].

A tecnologia de *workflow* possibilita uma nova abordagem ao desenvolvimento de sistemas, fornecendo a plataforma tecnológica adequada para a modelação, instanciação e controlo de processos. O papel dos sistema de gestão de *workflow* no sistema de suporte ao desenvolvimento de software consiste na modelação, execução e controlo do processo de software adoptado no projecto.

5. Sistemas *Groupware*

A designação *groupware* surgiu em meados da década de oitenta para designar esta nova classe de software de suporte ao trabalho em equipa [Grudin 1994]. Podemos definir *groupware* como "*sistemas informáticos que apoiam grupos de pessoas envolvidas numa tarefa ou objectivo comum e que fornecem uma interface para um ambiente partilhado*" [Ellis 1991].

O trabalho em equipa possui várias dimensões que podem ser eficientemente suportadas por ferramentas de *groupware*: interacção entre as pessoas, troca de ideias, discussão de pontos de vista, tomadas de decisão conjuntas, manipulação e acesso a documentos comuns, conjugação de esforços e coordenação de acções.

As ferramentas *groupware* podem revelar-se valiosas quando adoptadas pelas equipas dos projecto de software. Desenvolver software é um trabalho em equipa que implica uma intensa comunicação, colaboração e coordenação de diversas pessoas com competências distintas. Durante um projecto de desenvolvimento de software os membros da equipa interagem, com o objectivo de tomarem decisões importantes sobre o sistema ou esclarecerem dúvidas.

A comunicação entre os membros da equipa tem um papel de coordenação do trabalho. Numa equipa de software ocorrem dois tipos de comunicação: formal e informal. Ambos os tipos são necessários embora se possam adequar a diferentes géneros de tarefas. Investigação levada a cabo demonstrou a importância dos fluxos de comunicação informal e da facilidade de aquisição da informação como um factor de sucesso [Kraul et al. 1995].

A qualidade da informação aumenta com a diminuição da distância entre os intervenientes, isto é, cada membro da equipa deve ser capaz de comunicar directamente com a pessoa mais capaz

de o auxiliar. Demonstrou-se que com o aumento do número de intermediários a probabilidade da comunicação ser bem sucedida diminuía [Curtis et al. 1988]. As ferramentas de *groupware* podem ser valiosas neste contexto porque facilitam o contacto directo e informal.

As ferramentas de *groupware* facilitam o trabalho colaborativo de uma equipa, cabendo aos utilizador combinar a sua utilização de forma a suprir as suas necessidades de interacção. As vantagens da utilização de sistemas de *groupware* estão intimamente ligadas à dimensão do universo dos seus utilizadores, sendo tanto maiores os benefícios quanto maior for o número de utilizadores.

6. Ambiente de Suporte ao Desenvolvimento de Software em Equipa

Existem diversas configurações possíveis para uma equipa de desenvolvimento de software. Cada configuração possui as suas próprias características e necessidades. O objectivo de um ambiente de suporte ao desenvolvimento de software é maximizar o desempenho da equipa.

O ambiente de suporte pode ser construído por duas componentes tecnológicas: *workflow* e *groupware*. A componente de *workflow* permite definir, executar e controlar o processo adoptado pelo projecto. A componente de *groupware* fornece as funcionalidades de comunicação e colaboração necessárias.

Não existe um ambiente de suporte genérico que se possa adequar a todas as configurações de equipas. Cada tipo de equipa tem as suas particularidades e necessidades. Por este facto, será descrito um possível ambiente de suporte ao desenvolvimento, baseado em tecnologias de *workflow* e de *groupware*, para cada uma das configurações de equipa.

Elementos tecnológicos comuns do ambiente de desenvolvimento

Usando um processo de desenvolvimento clássico:

- Sistema de *workflow* para a definição, execução e controlo do processo de desenvolvimento adoptado;
- Correio electrónico para troca de mensagens assíncronas entre os elementos da equipa;
- Agendas electrónicas partilhadas que permitam aos elementos da equipa consultar a disponibilidade de cada pessoa e facilitar o agendamento de reuniões;
- Gestão e partilha electrónica de documentos.

Usando um processo de desenvolvimento ágil:

- Correio electrónico para troca de mensagens assíncronas entre os elementos da equipa;
- Mensagens instantâneas para fácil comunicação entre os membros da equipa;

- Agendas electrónicas partilhadas que permitam aos elementos da equipa consultar a disponibilidade de cada pessoa e facilitar o agendamento de reuniões;
- Gestão e partilha electrónica de documentos.

Elementos tecnológicos específicos do ambiente de desenvolvimento para equipas em turnos

Usando um processo de desenvolvimento ágil:

- Calendário partilhado pelos elementos da equipa e que permita a qualquer pessoa consultar a existência de eventos importantes para o projecto;

Uma equipa em turnos tenderá a utilizar de forma mais intensiva as ferramentas de *groupware* do que uma equipa co-localizada que favorecerá a comunicação face a face. A execução do trabalho por turnos diferentes com dependência entre tarefas coloca novos desafios de coordenação do trabalho.

Elementos tecnológicos específicos do ambiente de desenvolvimento para equipas distribuídas a trabalhar sincronamente

Usando um processo de desenvolvimento clássico:

- Mensagens instantâneas;
- Fóruns de discussão;
- Vídeo-conferência.

Usando um processo de desenvolvimento ágil:

- Fóruns de discussão;
- Vídeo-conferência.

Elementos tecnológicos específicos do ambiente de desenvolvimento para equipas distribuídas a trabalhar em diferentes momentos

Usando um processo de desenvolvimento clássico:

- Mensagens instantâneas;
- Grupos de discussão;
- Vídeo-conferência.

Usando um processo de desenvolvimento ágil:

- Grupos de discussão;
- Vídeo-conferência.

Este tipo de equipa colocará uma ênfase muito maior na utilização das ferramentas de *groupware* que facilitam a comunicação assíncrona.

7. Conclusões

Com este trabalho identificaram-se as diversas situações de trabalho em equipa, nomeadamente no que se refere ao desenvolvimento de software, e quais as funcionalidades disponibilizadas pelas actuais ferramentas *groupware* e *workflow* que poderão ser utilizadas para suportar as suas necessidades. O grande objectivo é identificar as condições tecnológicas necessárias para que as equipas de desenvolvimento de software melhorem o seu desempenho, quer quando usam metodologias de desenvolvimento ágeis ou clássicas, quer quando se encontram co-localizadas ou fisicamente distribuídas, a trabalhar em simultâneo ou em momentos diferentes.

Neste artigo, pretendeu-se descrever o ambiente ideal de suporte ao desenvolvimento de software em equipa. Seria útil, em trabalho futuro, avaliar as alternativas tecnológicas existentes e conceber uma arquitectura coesa do sistema. A consolidação das funcionalidades numa única interface potenciaria, naturalmente, ganhos de produtividade da equipa de desenvolvimento.

Embora o ambiente descrito apenas pretenda apoiar as equipas no desenvolvimento de software propriamente dito, a verdade é que o sucesso deste tipo de projectos está também fortemente dependente da forma como é gerido, tirando o melhor partido possível dos recursos disponíveis e respeitando as restrições orçamentais e temporais. Seria útil explorar os requisitos de gestão de projecto e integrá-los no ambiente de desenvolvimento descrito.

8. Referências

Curtis, B. Krasner H. e Iscoe, N., "A Field Stufy of the Software Design Process for Large Systems", *Communications of the ACM*, 1988.

Deminörs, E., "The Role of Teamwork in Software Development: Microsoft Case Study", *IEEE*, 1997.

Ellis, C. A., "Groupware Some Issues and Experiences", *Communications of the ACM*, 1991.

Fayad, M. E., "Software Development Process: A Necessary Evil", *Communications of the ACM*, 1997.

Fugetta, A., "Software Process: A Roadmap", Proceedings of the 22nd Conference on Software Engineering, ACM Press, 2000.

Grinter, R. E. e Herbsleb J. D., "The Geography of Coordination: Dealing with distance in R&D work", GROUP 1999.

Grudin, J., "Groupware and Social Dynamics. Eight Challenges for Developers", *Communications of the ACM*, 1994.

Kraul, R. E. e Streeter L. A., "Coordination in Software Development", *Communications of the ACM*, 1995.

Miguel, A., *Gestão de Projectos de Software*, FCA, 2003.

Sommerville, I., *Software Engineering* (6th ed.), Addison-Wesley. 2001.

WFMC, *WfMC Reference Model*, Workflow Management Coalition, 1995.

Modelo para Integração de Aplicações na Dielmar

Octávio Pereira, Pedro António

Escola Superior de Tecnologia, Castelo Branco, Portugal

octavio@est.ipcb.pt, pantonio@est.ipcb.pt

, Eduardo Valente, Hugo Ferreira

Escola Superior de Tecnologia, Castelo Branco, Portugal

eduardo@est.ipcb.pt, hj_ferreira@portugalmail.pt

Resumo

Descreve-se um modelo de integração de aplicações, aplicado a um caso prático numa empresa. O EAI (*Enterprise Application Integration*) é uma nova abordagem aos sistemas de informação, pelo que conduz ao desenvolvimento de novas soluções estratégicas empresariais competitivas, integrando funcionalidades e dados de aplicações díspares, aproveitando os investimentos do passado no sistema de informação e potenciando a abertura do mesmo ao ambiente externo da empresa. O sistema de informação em causa é constituído por diversas aplicações legadas, por sua vez suportadas em bases de dados separadas e interface do utilizador distinta. As limitações que estas apresentam com o crescimento dos dados e incapacidade para integrar e uniformizar o seu interface, bem como relatórios produzidos em diferentes tipos, conduziram a uma solução de projecto EAI. O modelo de desenvolvimento proposto procura acomodar vários aspectos de evolução do sistema de informação, no sentido de permitir a evolução das bases de dados das aplicações para novas versões, uniformizar o interface com as aplicações via web e, essencialmente, criar uma infra-estrutura de aplicações idêntica, que permita consultar, cruzar e transferir dados harmoniosamente entre si, permitindo configurar uma arquitectura de EAI como sendo a espinha dorsal do seu sistema de informação. Descreve-se um modelo para integração das várias aplicações existentes e das aplicações que futuramente possam vir a ser criadas, para facilitar a sua integração e desenvolvimento. Não menos importante é a possibilidade de poder configurar posteriormente uma Intranet com uma interface web uniforme para as várias aplicações e uma solução de back-office para administração das mesmas, para o qual o protótipo aqui descrito foi desenvolvido com esse propósito.

Palavras chave: EAI; sistemas legados; integração de aplicações; back-office; Intranet

1. Introdução

Durante as últimas décadas, as empresas focaram-se nas tecnologias de informação e várias aplicações executadas para automatizar os seus processos do negócio. Estas aplicações não foram desenvolvidas de uma forma coordenada, mas evoluíram em consequência da tecnologia mais avançada do momento. Em diversas organizações, o sistema de informação consiste apenas em soluções autónomas e muitas vezes heterogéneas. Esta situação causou vários problemas de integração, pois as aplicações não poderiam cooperar e assim, não poderia ser construída uma ponte entre aplicações díspares, para cruzar dados e funcionalidades ou a apresentação destes (Cunha, 2001).

Recentemente, uma categoria de integração chamada Integração de Aplicações Empresariais (EAI) ou simplesmente integração de aplicações tentou dirigir-se eficazmente aos muitos problemas da integração, que assim, resultam no desenvolvimento e manutenção de sistemas de informação integrados e flexíveis. A integração de aplicações é uma área nova de investigação e, consequentemente, pesquisa científica, sendo que a literatura existente em torno dela permanece limitada. Themistocleos advoga que para integrar eficazmente as n aplicações constituintes de uma organização é preciso criar n(n-1)/2 ligações entre elas (Themistocleos, 2002).

Não obstante a simplicidade conceptual da arquitectura de integração ponto-a-ponto, existe já uma arquitectura para a integração de aplicações consubstanciada, para onde o sistema poderá convergir igualmente com o mesmo número de ligações, conhecida na literatura como *Hub-and-Spoke*. Mais recentemente, surgiu outro conceito, impulsionado pelo aparecimento da tecnologia de *web services* que utiliza o protocolo de invocação remota SOAP (*Simple Object Access Protocol*), a arquitectura SOA (*Service Oriented Arquitecture*), onde se pretende decompor todas as funcionalidades possíveis das aplicações do sistema de informação, para ser invocados como serviços de uma forma distribuída, por qualquer outra aplicação interna ou externa à organização (Neves et al, 2004).

A integração de aplicações empresariais é uma área recente de intervenção em sistemas de informação sobre os planos, métodos, e ferramentas pensadas para a modernização, consolidação e coordenação das aplicações informáticas numa empresa. Tipicamente, uma empresa tem um legado de aplicações e bases de dados existentes e quer continuar a usá-las ao adicionar ou ao migrar para um conjunto novo de aplicações que exploram a internet e extranet ou o comércio electrónico. O EAI pretende ser uma visão totalmente nova do sistema de informação, observando como as aplicações existentes se enquadram no novo cenário, com o planeamento de novas formas de reutilização eficiente do que já existe, ao adicionar novos dados e aplicações. Abrange metodologias de projecto de sistemas de informação, bem como tecnologias de *middleware* desde a programação orientada por objectos, aplicações distribuídas, comunicação entre plataformas diferentes usando CORBA, COM+ e *web services*, o planeamento dos recursos empresariais (ERP – *Enterprise Resource Planning*) para os novos objectivos, gestão do relacionamento com o cliente (CRM – *Customer Relationship Management*) e distribuição empresarial em larga escala (SCM – *Supply Chain Management*) da informação usando bases de dados comuns e ajustamento da semântica com XML (*Extensible Markup Language*), ou sistemas de comunicação por mensagens (Boiling et al, 2006).

A tendência em sistemas de informação é adoptar um sistema de planeamento de recursos da empresa global integrado (ERP) ou por outro lado, projectar o sistema de informação com análise e modelação integrada, para manter uma percepção holística do negócio da empresa. O mito do ERP único é sempre difícil de atingir, pelo que aparece sempre mais qualquer mudança e novamente dá lugar a uma aplicação periférica. Novos métodos surgem para se adaptar cada vez melhor às inovações tecnológicas e de mercado, sistematicamente (Delen, D. et al).

O projecto descrito no restante deste artigo tem como objectivo principal criar um modelo normalizado para integração das várias aplicações existentes e das aplicações que futuramente possam vir a ser criadas na empresa Dielmar. Descreve-se o controlo de acessos para as aplicações existentes, com a mesma interface de acesso e o mesmo sistema de autenticação e autorização, bem como o modelo de acesso aos servidores de Base de Dados, comuns a todas as aplicações. As aplicações que vierem a ser criadas futuramente, poderão assim ter uma base para a sua integração e desenvolvimento ser mais suave, uma vez que já existe um modelo normalizado a partir do qual se começa a trabalhar.

2. Descrição do problema

Existem várias aplicações a funcionar na Dielmar que foram feitas de forma independente umas das outras, ou seja, bases de dados em diferentes sistemas de gestão de bases de dados, tais como SQLServer, Access, FoxPro ou Excel, métodos de acesso à base de dados e de autenticação diferentes, o que causa muita entropia na utilização das várias aplicações e dificulta o processo de integração das mesmas, assim como de novas aplicações futuras (Mourinho, 2002). Pretende-se então, minimizar estas dificuldades de integração através da criação de um modelo normalizado, na construção de novas aplicações e modificação gradual das já existentes (Oliveira e Silva, 2003) (Galguinho e Tavares, 2004) (Martins e Marques, 2004).

Um dos problemas é a autenticação dos utilizadores nas aplicações. O método de autenticação actual para aceder às aplicações é feito através da introdução de pares *login/password*, como ilustra a figura 1. Cada aplicação tem a sua própria informação de autenticação sobre os utilizadores. Um utilizador começa por autenticar-se no domínio de rede e, posteriormente, tem que efectuar o processo de autenticação em cada uma das aplicações que pretende aceder. A quantidade de vezes que um utilizador tem de introduzir informação de autenticação para aceder às várias aplicações pode ser fastidioso, incómodo e desnecessário.

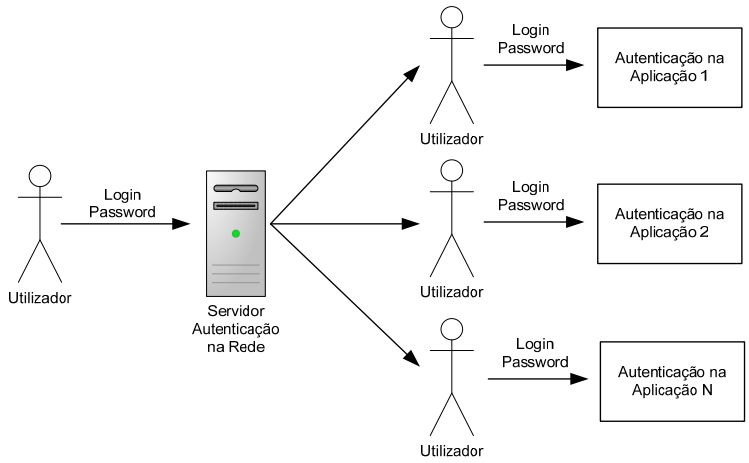


Figura 1 – Modelo de autenticação existente

Assim, uma possibilidade de resolver o problema, seria a autenticação das aplicações estar integrada no domínio da rede, de forma a facilitar o processo de autenticação em todas as aplicações. Assim, passa a existir apenas um único ponto de introdução de informação de autenticação (*login/password*), como é apresentado na figura 2 (Ferreira, H., 2005).

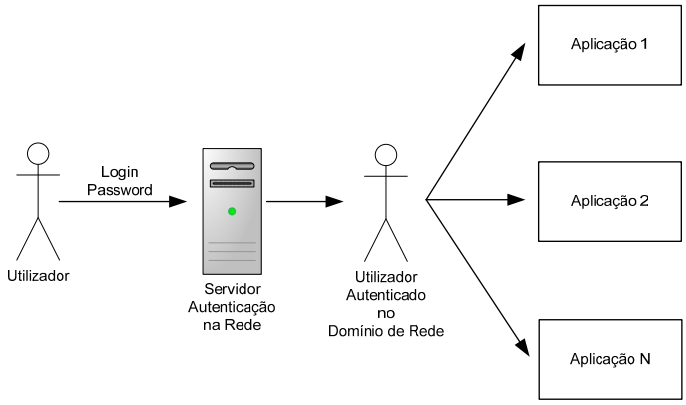


Figura 2 – Modelo de autenticação integrado no domínio da rede

Outro problema existe ao nível dos privilégios dos utilizadores para aceder às aplicações. Os utilizadores apenas devem aceder às aplicações que estão relacionados com a sua área de intervenção e não devem aceder às restantes aplicações, por razões de segurança e/ou produtividade.

3. Modelo de desenvolvimento

A maioria das aplicações informáticas, independentemente do seu papel ou da tecnologia que as implementa, possuem três componentes gerais de funcionalidade. Um componente (camada de apresentação) é a interface com o utilizador ou com outro sistema que permita à aplicação comunicar com outra entidade. Noutro componente (camada de aplicação) reside a lógica de negócio, ou regras de negócio, que contém a parte do processo informático que as aplicações automatizam. Por fim, o componente do acesso aos dados (camada de dados), que tem a finalidade de realizar operações de armazenamento, procura e recolha de dados garantindo a sua integridade (Linthicum, 2003).

O modelo normalizado de desenvolvimento de novas aplicações proposto é baseado numa arquitectura com estes três componentes (ou camadas), como se pode ver na figura 3. A interface do utilizador com a aplicação deve ser desenvolvida em ambiente *web*. A principal vantagem desta abordagem é, praticamente, não existirem requisitos de instalação ao nível dos utilizadores, pois actualmente a maioria dos computadores possuem *web browsers*. As interfaces do utilizador devem ser desenvolvidas com uma linguagem que tenha a capacidade de gerar páginas *web* dinâmicas. A lógica de negócio da aplicação deve ser desenvolvida recorrendo a uma linguagem de programação que permita codificar os vários componentes de negócio. Os serviços são oferecidos por um servidor *web* com suporte de interpretação de uma linguagem que gera páginas *web* dinâmicas. A componente de dados pode ser um simples ficheiro ou um motor de base de dados.

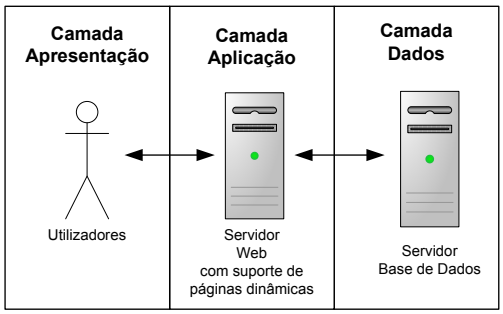


Figura 3 – Modelo de desenvolvimento de aplicações baseado na arquitectura das 3 camadas

Com base neste modelo desenvolveu-se um sistema para centralizar o acesso às diversas aplicações existentes na Dielmar (descrito na secção 4). A arquitectura de funcionamento deste sistema é ilustrado na figura 4, onde o acesso uniforme às várias aplicações é centralizado pela aplicação designada por INTRANET Portal, que concentra a autenticação, segurança e permissão do acesso às outras aplicações (Ferreira, H., 2005). Pretende-se também que o modelo desenvolvido esteja vocacionado para integrar as aplicações legadas monolíticas, onde se verificam características tipicamente pouco flexíveis, quando comparadas com uma arquitectura a três camadas, onde se pode operar separadamente em cada uma das três. Tipicamente, nas aplicações monolíticas, o código que implementa a lógica do negócio, a interface e os dados está num único bloco ou misturado no mesmo corpo, sendo este, por vezes, muito extenso. Estas eram também, frequentemente desenvolvidas usando linguagens e ferramentas diferentes, com objectivos operacionais específicos, para servirem entidades particulares e não existia a necessidade de interacção entre aplicações.

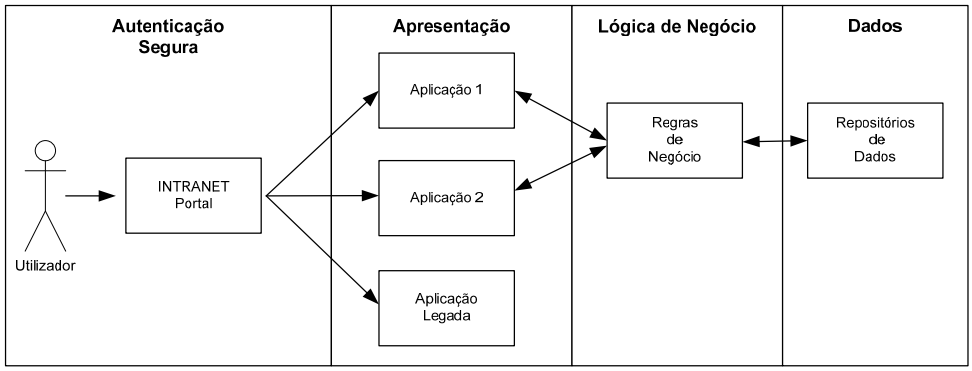


Figura 4 – Arquitectura funcional do sistema de controlo de acessos

4. Sistema de Controlo de Acessos

O sistema permite realizar um controlo de acessos, para os vários utilizadores registados, de forma a filtrar o acesso a várias aplicações, consoante o utilizador. Cada utilizador poderá aceder a aplicações conforme as permissões que lhe foram atribuídas. O sistema faculta a existência de três tipos de utilizadores, com diferentes níveis de utilização, para uma maior integridade do sistema. Eles são o ADMINISTRADOR (com privilégios totais de utilização, podendo realizar mais tarefas), o UTILIZADOR NORMAL (com privilégios mais reduzidos de utilização, podendo realizar menos tarefas que o administrador) e o UTILIZADOR ANÓNIMO (com privilégios ainda mais baixo de utilização). As funções ou requisitos funcionais que o sistema implementa e que serão desempenhadas pelo administrador, utilizador normal e pelo utilizador anónimo são descritas de seguida.

O administrador tem disponíveis as funcionalidades de gestão de todo o sistema de controlo de acessos. O sistema é composto por vários componentes e permite gerir utilizadores, grupos de utilizadores, as aplicações a disponibilizar aos grupos de utilizadores e mensagens.

- O componente de gestão de utilizadores permite efectuar as operações tradicionais como inserir novos utilizadores, consultar informação dos utilizadores, apagar utilizadores, alterar dados dos utilizadores e associar utilizadores a grupos;
- O componente de gestão de grupos de utilizadores tem a finalidade de criar, consultar, apagar e alterar grupos de utilizadores (por exemplo, Recursos Humanos, Produção, etc.). O sistema tem flexibilidade elevada para a manipulação de grupos de utilizadores, permitindo a qualquer instante criar um novo grupo de utilizadores;
- No componente de gestão de aplicações é realizada a introdução de novas aplicações que se pretendem que sejam acedidas a partir do sistema e associar as aplicações com grupos de utilização. Permite também remover aplicações obsoletas, ou seja, que deixaram de ser utilizadas. Neste componente, ainda existe a possibilidade de aceder a todas as aplicações disponibilizadas;
- Por fim, o componente de gestão de mensagens destina-se a oferecer um canal de comunicação entre os utilizadores integrados neste sistema. Aqui podem ser enviadas mensagens entre os utilizadores do sistema. O cariz das mensagens pode ser bastante diversificado, como por exemplo solicitar determinadas tarefas a um utilizador, pedido de esclarecimento de dúvidas sobre determinado assunto, etc.

Um utilizador normal do sistema de controlo de acessos, após se autenticar, tem acesso imediato às aplicações que lhe foi concedido acesso. Para isso, o utilizador apenas tem de estar associado a pelo menos um grupo de utilização e que estes grupos tenham aplicações associadas. Assim, são apresentadas apenas as aplicações que são relevantes a este utilizador

numa única interface. Para além desta funcionalidade, o utilizador tem também a possibilidade de enviar e receber mensagens.

Os utilizadores anónimos são os utilizadores que ainda não estão registados no sistema de controlo de acessos. Estes apenas podem efectuar o registo no sistema, que ficará pendente até aprovação, ou não, por parte de um administrador.

O sistema considera alguns aspectos relevantes tidos em conta durante a implementação do sistema como a flexibilidade, a usabilidade e a sustentabilidade. O sistema foi implementado para ser flexível ao nível de operações para adicionar, remover ou alterar os elementos. A usabilidade do sistema desenvolvido teve o objectivo de permitir que o utilizador se sinta confortável na utilização do sistema, nomeadamente com uma interface utilizador simples e flexível, permitindo eficácia na realização das várias operações. A sustentabilidade permite que o sistema seja melhorado de uma forma sistemática sem que se percam as funcionalidade principais do sistema. Estes aspectos não são necessariamente impostos no próprio sistema (produto final), mas possivelmente nos processos usados para criar o sistema ou nos artefactos (relatórios, código fonte) produzidos no decorrer desse processo. O protótipo é ilustrado na figura 5 (Ferreira, H., 2005).

Figura 5 – Protótipo da aplicação IA@DI

5. Análise das tecnologias utilizadas

Os factores que influenciaram e conduziram à escolha da linguagem de programação principal foram os seguintes:

- Uma linguagem de programação independente da plataforma operacional que oferece um elevado grau de flexibilidade sobre eventuais alterações do sistema operativo por parte da empresa;
- Uma linguagem de programação com um conjunto de funcionalidades alargado, de forma a aproveitar ao máximo funcionalidades oferecidas;
- Uma linguagem de programação robusta no que concerne à gestão de memória e à execução de programas, oferecendo garantias elevadas de fiabilidade;
- Uma linguagem de programação "viva", ou seja, que está em permanente desenvolvimento e actualização, considerando as tecnologias de informação de ponta.

Perante estes factores, optou-se por eleger a linguagem de programação Java, como a linguagem principal no desenvolvimento de novas aplicações ou para a reconversão de aplicações

existentes. Não impede, no entanto, que determinados aspectos específicos sejam implementados noutra linguagem de programação, se necessário.

Como foi referido na secção 3, a interface com utilizador é preferencialmente efectuada via *web browser*. Assim, utiliza-se a tecnologia Java para gerar páginas *web* dinâmica, o JSP (*Java Server Pages*). As páginas JSP são interpretadas num programa que fornece o ambiente Java de execução que neste caso propôs-se o *Tomcat* (Apache Tomcat, 2006). As páginas JSP utilizam os componentes de negócio desenvolvidos segundo a convenção *JavaBeans*. Assim, permite que se implemente os componentes de *JavaBeans* com a lógica da aplicação e os de JSPs com a apresentação, sendo que os conteúdos dinâmicos (informação) e estáticos (design e estrutura) também podem ser trabalhados em separado. Os *JavaBeans* permitem a utilização, reutilização e, eventual, alteração dos componentes de negócio codificados na linguagem Java. Os dados, preferencialmente, devem estar armazenados em base de dados relacionais que são manipulados pelos componentes de negócio através da tecnologia JDBC (*Java Database Connectivity*).

Transversalmente, para tratar de aspectos de segurança na comunicação de dados entre o utilizador num *web browser* e o servidor *web* utilizou-se o protocolo de segurança HTTPS, ou seja, o dados transmitidos em HTTP são cifrados com SSL (*Secure Sockets Layer*). Pretende-se implementar de uma forma geral a cifra dos dados de comunicação para evitar ataques de escuta da comunicação que podem conter informação confidencial, como por exemplo *passwords*. O protocolo HTTP sobre SSL é utilizado especialmente em *web browsers* e em aplicações de comércio electrónico (*e-commerce*). É responsável por gerir um canal de comunicações seguro entre cliente e servidor, baseado num poderoso mecanismo de cifra de dados com chaves públicas e privadas.

6. Conclusão

A empresa Dielmar, não possuía um método adequado de controlo de acessos às aplicações, uma vez que essas aplicações foram desenvolvidas de forma independente umas das outras, ou seja, métodos de acesso a bases de dados diferentes, métodos de autenticação diferentes, o que dificulta o processo de integração das aplicações existentes, assim como das aplicações futuras. Isso acontece também devido às aplicações serem desenvolvidas por diferentes programadores e por não terem nenhum modelo comum de desenvolvimento. Então com o modelo proposto pretende-se minimizar essas dificuldades de integração, através de um controlo de acessos centralizado.

A empresa validou e adoptou este modelo para o futuro na integração das aplicações de *back-office*. De significativa relevância é a capacidade de integração destas aplicações com pacotes de software ERP, CRM e SCM, cuja aquisição é estudada consecutivamente pela gestão de topo da Dielmar, à procura de competitividade no mercado.

A arquitectura aqui descrita permite integrações posteriores de vários tipos, sejam por dados, por funcionalidades ou pela interface com o utilizador. Cada aplicação desenvolvida nesta arquitectura é fechada como um componente reutilizável, que pode fazer parte integrante de uma intranet ou configurar uma nova estrutura de integração.

Em estudo para trabalho futuro ficou também a possibilidade de desenvolver *Web Services* para integração *front-office* de algumas aplicações numa possível *Extranet*. Outra solução para o futuro é instalar um servidor de aplicações J2EE, integrando todas as aplicações desenvolvidas nesta arquitectura, convertendo os componentes *JavaBeans* para *Enterprise JavaBeans* ao nível de *back-office*.

A arquitectura referida é modular e assenta na linguagem Java, podendo o sistema operativo, o servidor Web e a base de dados serem de tecnologias diferentes, colocando de lado uma dependência excessiva de um fabricante. A opção do Java, permite explorar posteriormente um

conjunto de serviços, como por exemplo evoluir esta arquitectura para um servidor de aplicações J2EE, utilização de *Web Services*, chamada a métodos remotos (*Java RMI*), etc.

Com esta arquitectura passa a ser possível efectuar a integração de aplicações tanto a nível de dados ao migrar para novas versões de sistemas de gestão de bases de dados, como a nível de serviços recorrendo às regras de negócio concentradas nas APIs *JavaBeans*, através de chamadas remotas de procedimentos ou outra tecnologia de invocação de procedimentos.

Referências

Apache Tomcat, http://tomcat.apache.org/, Dezembro 2006

Boiling, S., Dasika, P., Krishnan, S., http://searchwebservices.techtarget.com, 10 Oct 2006

Castela, N. Recolha, análise e validação de informação para a Modelação de Processos de Negócio, 10, 2001

Cunha, P., Projecto de sistemas de informação para a realidade emergente: proposta baseada num modelo de carteira de soluções, cap. 4, 2001

Delen, D., Dalal, N., Benjamin, P., "Integrated Modeling – The key to holistic understanding of the enterprise", Communications of the ACM, Vol. 48, 04, 2005

Ferreira, H., "I@DI - Integração at Dielmar", 2005

Galguinho, H., Tavares, S., "IADI3 – Integração de Aplicações na Dielmar", 07, 2004

Linthicum, D., *Next Generation Application Integration*, Addison-Wesley, 2003

Martins, S., Marques, R., "IADI2 – Integração de Aplicações na Dielmar", 06, 2004

Mourinho, N., "Sistema de Documentação para Gestão, Análise e Exploração de Informação no Sistema de Informação da DIELMAR SA", 09, 2002

Neves, P., Pereira, O., Silva, J., Ramalho, P., "Web Service Based Multi-Channel Application", proceedings of IADIS WWW/Internet, 10, 2004

Oliveira, B., Silva, B., "IADI – Integração de Aplicações na Dielmar", 10, 2003

Themistocleos, M., "Evaluating the Adoption of Enterprise Application Integration In Multinational Organisations", 03, 2002

Plataforma Web de Acesso e Partilha de Imagens Médicas

Sérgio Lima

IEETA – Universidade de Aveiro, Aveiro, Portugal

Sérgio.lima@biocant.pt

Natércia Sousa

IEETA – Universidade de Aveiro, Aveiro, Portugal

natercia.sousa@lactogal.pt

Carlos Costa

IEETA – Universidade de Aveiro, Aveiro, Portugal
ccosta@det.ua.pt

Augusto Silva

IEETA – Universidade de Aveiro, Aveiro, Portugal
asilva@det.ua.pt

Jacek Kustra

IEETA – Universidade de Aveiro, Aveiro, Portugal
Jacek.kustra@gmail.com

Resumo

A produção de imagens médicas em formato digital tem vindo a crescer nas mais variadas instituições prestadoras de cuidados de saúde, representando neste momento um importante e imprescindível elemento de suporte à decisão médica [1]. Os formatos e processos de armazenamento e transmissão destas imagens são definidos pela norma internacional DICOM [2]. A implementação desta norma contempla, na maioria das vezes, serviços de comunicação da camada de rede TCP/IP. No entanto, a maior parte das firewalls institucionais bloqueia este tipo de comunicações com redes abertas e inseguras como, por exemplo, a Internet.

O presente documento apresenta uma arquitectura de partilha e acesso remoto a imagens médicas, baseada em WebServices e no encapsulamento de objectos DICOM em HTTP. A solução apresentada é passível de ser implementada de forma transparente em qualquer actual unidade de saúde.

Palavras chave: Imagem Médica, PACS, DICOM, WADO, WebServices

1. Introdução

Hoje em dia, é ponto assente a enorme importância que as novas tecnologias de informação e comunicação têm vindo a angariar; a elas não está imune o Serviço de Saúde. Efectivamente, um serviço clínico dito na vanguarda e devidamente equipado para prestar cuidados de saúde em quantidade e, sobretudo, em qualidade, não pode prescindir dum bom investimento a este nível, ou seja, "em sistemas de informação e em infra-estruturas telemáticas capazes de

permitirem e assegurarem de forma eficiente e escalonada no tempo todo um conjunto de actividades relacionadas com os actos médicos que se praticam".

A par da evolução exponencial das potencialidades dum sistema informático, cresce também o domínio de aplicações médicas carentes de processamento computacional, como seja a própria gestão demográfica dos pacientes, bem como todo o processo inerente aos meios de diagnóstico e/ou intervenção médica (aquisição, processamento, comunicação e arquivo de informação). Ora, é aqui que entra o recurso às mais avançadas tecnologias multimédia, sendo perfeitamente viável a co-existência de informação em formato texto com a informação em formato digital das modalidades típicas de imagem médica. A ideia será, portanto, melhorar a aquisição, processamento, transmissão e arquivo de imagens clínicas no sentido da evolução e concretização eficaz de diagnósticos fidedignos, de motivação do corpo clínico e ainda duma redução de custos associados.

2. Norma DICOM e os sistemas PACS

A crescente utilização de equipamentos médicos electrónicos e a consequente produção de informação em formato digital tornou imprescindível a criação e adopção duma série de normas e "standards". Definindo os formatos e processos de armazenamento e transmissão na área da imagem médica, a norma DICOM (*Digital Imaging and Communications in Medicine*) [2], desenvolvida pelo consórcio ACR-NEMA, nasce da necessidade de existir um conjunto de recomendações e linhas de base responsáveis pela convergência de informação de imagem digital médica entre diferentes fabricantes de equipamento electrónico. Desta forma, esta norma não só promoveu a comunicação de dados, mas ainda o desenvolvimento e expansão de sistemas de arquivo, denominados como PACS (Picture Archiving and Communications Systems) [3]. Mais ainda, define interfaces com os restantes sistemas de informação hospitalares. Por outro lado, objectivou-se ainda a criação de bases de dados de informação de diagnóstico que poderiam ser utilizadas por uma grande variedade de dispositivos distribuídos geograficamente.

Resumidamente, podemos dizer que o protocolo DICOM contempla os seguintes aspectos [4]:

- a semântica de comandos e respectivos dados associados para que os dispositivos possam interagir;
- a semântica dos serviços de ficheiros no que diz respeito ao seu formato e estrutura de directórios necessários à comunicações off-line;
- permite e facilita as operações em ambientes de rede fazendo, sempre que possível, uso de normas existentes;
- apresenta-se estruturada para suportar a introdução de novos serviços resultantes de novas aplicações de imagem médica.

Definitivamente enraizado na comunidade médica, o seu carácter perfeitamente abrangente e, de certo modo, versátil, proporciona a interoperabilidade de sistemas em ambientes multi-vendedores, desde que em conformidade com as suas recomendações e especificações.

Actualmente, a norma DICOM encontra-se na sua versão 3.0, está dividida em 18 partes e contempla 123 suplementos de trabalho sobre aspectos específicos a uma ou várias partes da norma.

O termo PACS compreende várias tecnologias que contemplam a aquisição, arquivo e distribuição ou importação de imagem em formato digital numa rede de computadores, para diagnóstico e revisão em estações dedicadas, Figura 1. Sob o ponto de vista clínico, os sistemas PACS constituem mais valias essenciais na prestação dos cuidados médicos. Várias são as vertentes que tiram vantagem deste tipo de tecnologia:

- é eliminada a necessidade de recurso à película ou ao filme analógico como mecanismo de arquivo, ganhando na capacidade e velocidade de pesquisa e consulta (visualização) de arquivos;
- é permitida a replicação das imagens, sem perdas, de forma rápida e a baixo custo;
- é proporcionado o armazenamento permanente, sem degradação da qualidade de diagnóstico ao longo dos anos, com recurso a dispositivos de baixo custo e de utilização em massa como o CDROM ou o DVD.

No que respeita à manipulação concreta da imagem, é viável o desenvolvimento de técnicas de processamento de imagem, adaptadas a cada modalidade, com a finalidade de aumentar a quantidade e qualidade de informação clínica; tarefas como manipulação do tamanho da imagem, região a observar, amplificação, brilho e contraste, entre outras, facilmente são desempenhadas. Por outro lado, a quantidade de informação complementar que se pode extrair destas imagens é bastante superior com o recurso a ferramentas de quantificação que nos permitem efectuar medições precisas e processamentos quantitativos complexos.

Desta forma, é oferecido ao corpo clínico um ambiente tecnológico que os dota da possibilidade de enviar ou consultar remotamente a informação pretendida de uma forma rápida, abrindo as portas a sessões de teleconsulta ou telemedicina.

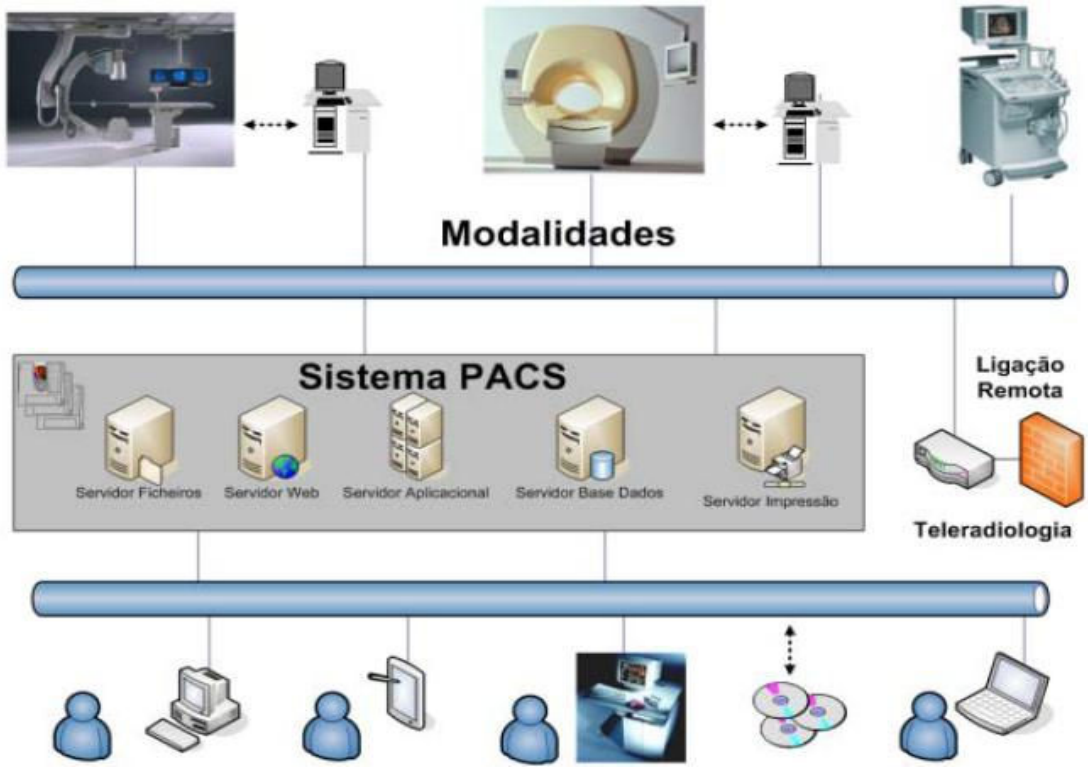


Figura 1: Ambiente DICOM/PACS

3. Problemática da arquitectura DICOM

Actualmente, uma arquitectura PACS-DICOM representa um elemento fundamental no sector da saúde pois, entre outros, elimina a necessidade do suporte físico dos exames clínicos (raio-x, filme, cd-rom,...), reduz drasticamente os tempos de acesso/distribuição dos dados (intra ou inter institucional), estando estes disponíveis logo após a execução do procedimento.

Se, ao longo da última década, não se têm verificado problemas de acessibilidades aos dados dentro da Intranet (ambiente controlado das instituições), quando se tenta promover a partilha e acesso remoto ao sistema PACS a partir do mundo exterior, nomeadamente a Internet, surgem grandes problemas.

Desde logo, as politicas de segurança implementadas pelos administradores de rede das instituições, que usualmente disponibilizam apenas o protocolo HTTP, bloqueando todas as ligações directas aos servidores utilizadas pelo protocolo DICOM (porto 104), impedindo a utilização desta arquitectura numa perspectiva inter-institucional. Desde logo, conseguimos identificar dois importantes cenários de utilização que, actualmente, estão comprometidos: - o caso das clínicas e laboratórios de exames privados que prestam serviços para os hospitais da rede pública, que não conseguem enviar electronicamente os procedimentos realizados, obrigando muitas vezes ao transporte do estudo em formato analógico (i.e. película ou filme); - os profissionais de saúde que não conseguem realizar tele-trabalho sobre a plataforma PACS da sua instituição.

Sensível a esta problemática, a norma passou recentemente a suportar encapsulamento de objectos DICOM em HTTP, esbatendo desta forma as limitações acima indicadas. Baseado nessa nova parte 18 da norma DICOM, denominada como WADO (Web Access to DICOM Persistent Objects) [5], desenvolvemos uma plataforma em ambiente Web que permite o acesso à estrutura PACS de qualquer instituição. Esta interface possibilita o registo de vários servidores PACS existentes nos vários departamentos da instituição, assim como o modo de operação de cada um deles (i.e. consulta e/ou recepção de estudos imagiológicos).

Assim, e no modo de consulta, é possível pesquisar (por nome, data do estudo, tipo de estudo, etc) e listar todos os pacientes que existem nos vários sistemas PACS de determinada instituição. Seleccionando determinado estudo, é possível visualizar as imagens num qualquer computador remoto. No modo de recepção, é possível fazer o upload remoto de um estudo em determinado sistema PACS, desde que as parametrizações do sistema e políticas de segurança o permitam.

4. Arquitectura Proposta

Em termos gerais, considere-se uma unidade hospitalar, constituída por diversos departamentos, como sejam a radiologia, a cardiologia, entre outros (Figura 2). Em cada uma destas unidades existe um ou mais servidores onde estão armazenadas as imagens clínicas produzidas pelos respectivos laboratórios, servidores estes que estão conectados através do protocolo DICOM (porto 104). A nível institucional, estes servidores estarão ligados entre si e a um servidor Web público, o qual se encontra protegido por uma *firewall* que apenas disponibiliza o protocolo HTTP (porto 80). À instituição poderão estar ligados através da Internet clínicos e/ou outras instituições espalhados por todo o mundo.

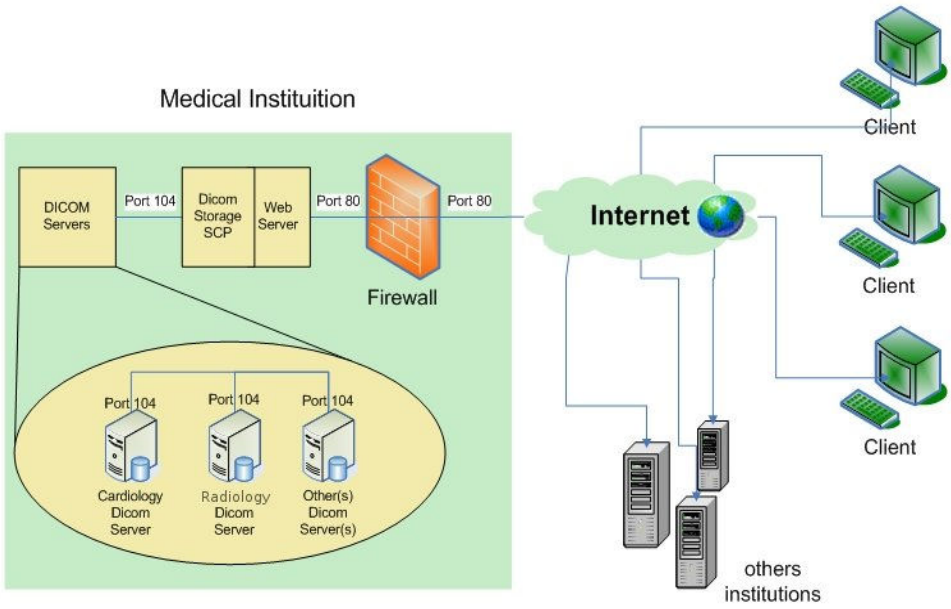


Figura 2: Arquitectura da solução implementada

Na Figura 3 apresentar-se um diagrama do fluxo de processos da plataforma desenvolvida. onde. Neste diagrama é possível visualizar sequência de acções inerentes a cada funcionalidade implementada na nossa solução. Como se pode verificar, é possível fazer uma pesquisa aos vários servidores DICOM, existentes na instituição, de forma a listar os pacientes que satisfazem essa pesquisa.

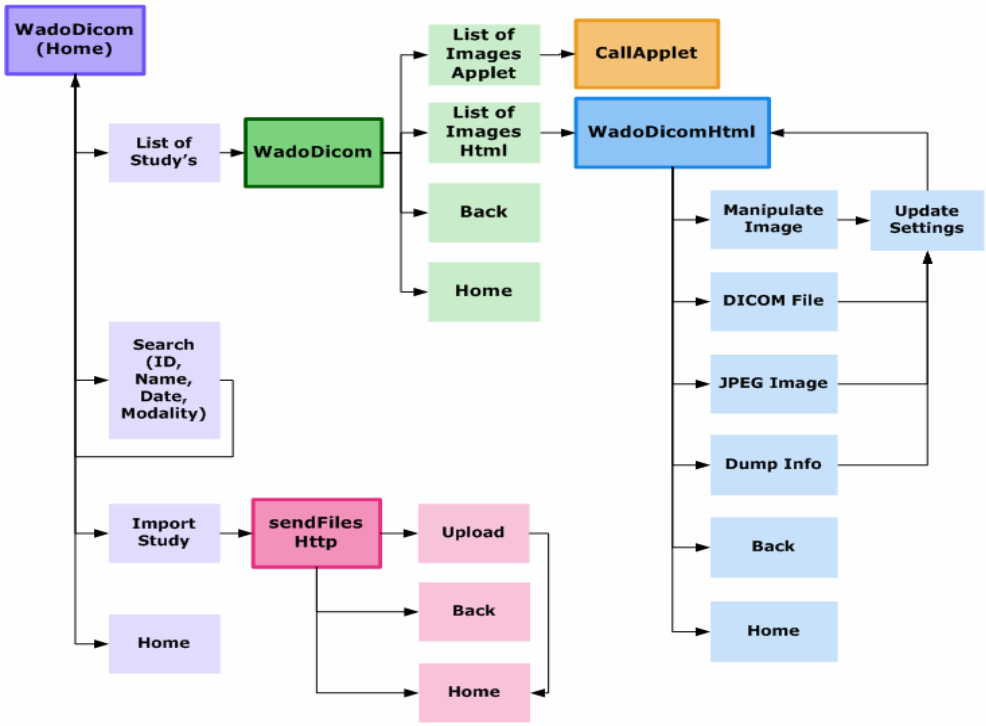


Figura 3: Diagrama de fluxo de processos da plataforma desenvolvida

5. Solução Implementada

A exploração da arquitectura proposta teve necessidade de ser suportada por um conjunto de módulos de software desenvolvidos para o efeito. Desde logo, na Figura 4 é possível observar a interface principal de acesso ao sistema WADO-DICOM. Nesta janela é possível pesquisar e listar estudos existentes nos vários servidores PACS registado no sistema. A visualização da informação imagiológica de um determinado estudo é possível, bastando para o efeito seleccionar o estudo/paciente pretendido.

Como os web browser's não têm capacidade nativa para visualização de formato de imagem DICOM, foi necessário desenvolver um visualizador específico para esse fim (Figura 5). Optou-se por desenvolver este módulo em Java na forma de uma Applet que corre do lado do cliente. Neste momento, este módulo para além das manipulações básicas da imagem (brilho, contraste, zoom, etc.), também permite visualizar sequências dinâmicas de imagem médica (i.e. filme). Na Figura 5 é possível ver um desses exemplos, a modalidade de angiografia (XA) .

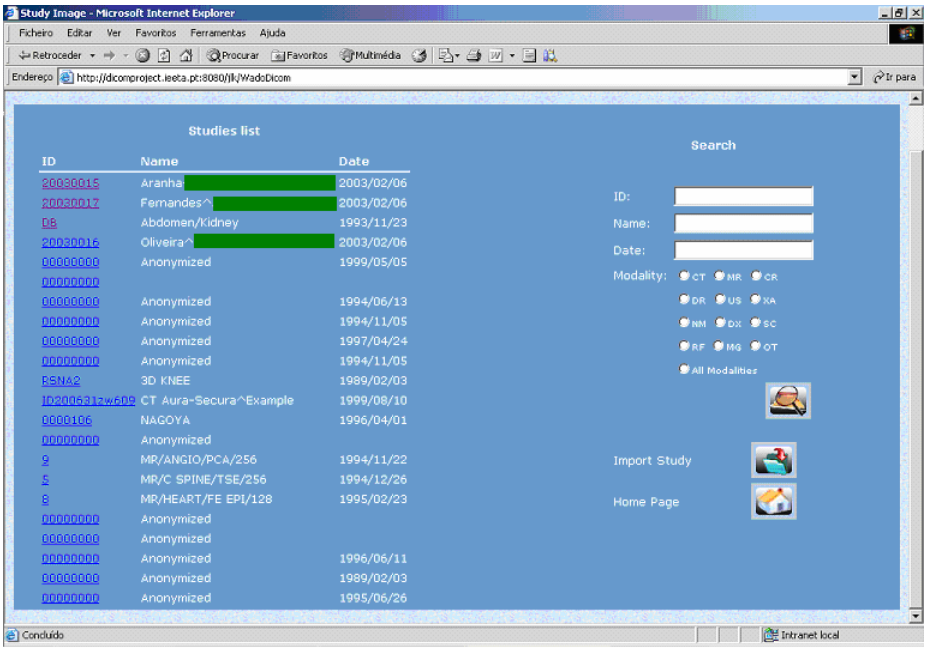


Figura 4: Página Principal – Interface Wado Dicom

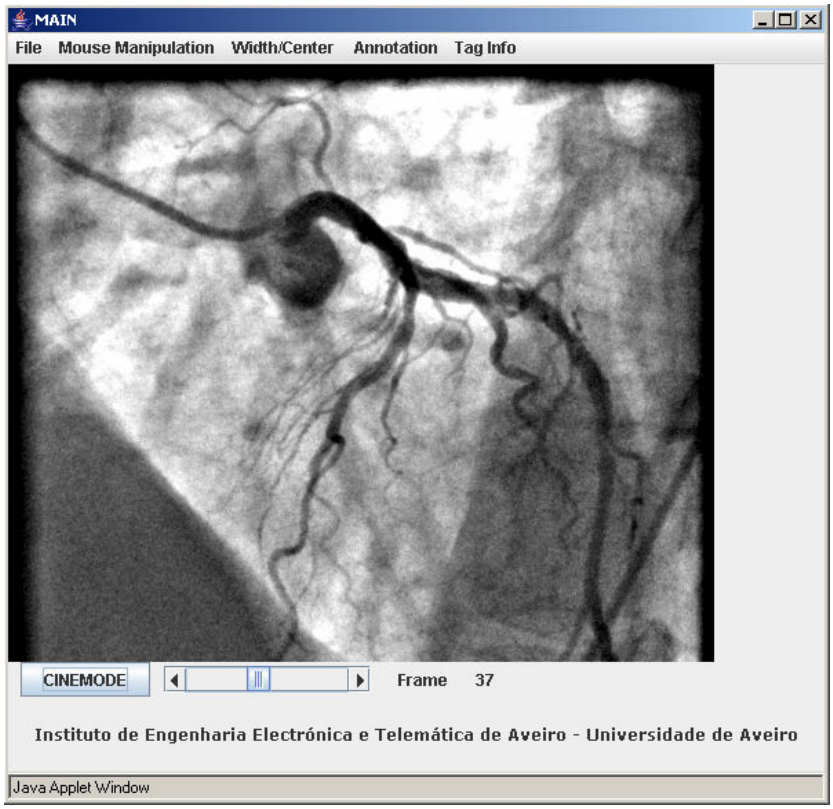


Figura 5: Visualizador de imagem DICOM (Applet)

No modo de recepção, a plataforma permite fazer o upload de estudos realizados no exterior da instituição, para um dos servidores internos de armazenamento de imagens médicas (Figura 6). Como se pode verificar, o utilizador pode escolher o servidor PACS onde pretende efectuar o upload do estudo ("List of Servers"). Do lado destes servidores PACS foi necessário criar uma nova "proveniência" para albergar/separar estes estudos dos restantes produzidos na instituição.

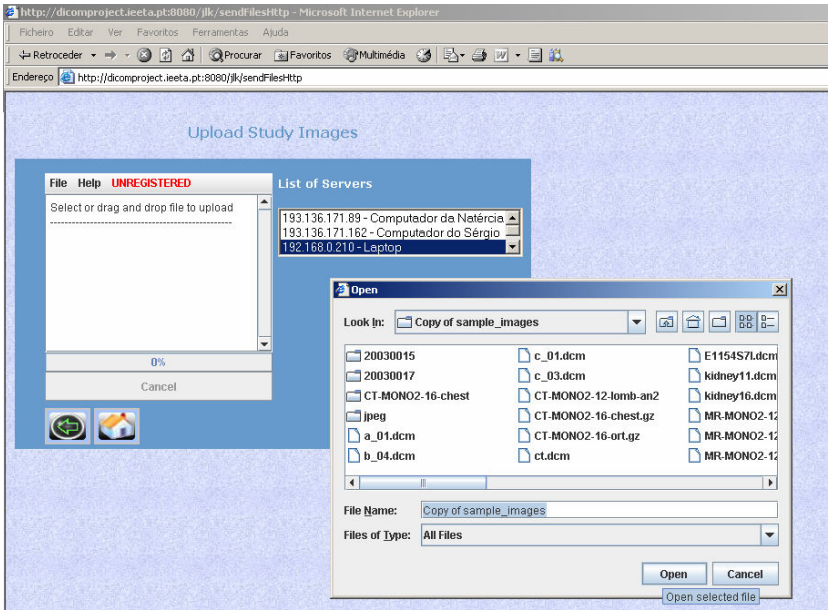


Figura 6: Interface para upload de estudos imagiológicos

Este modo de upload é de facto uma funcionalidade extra, não especificada pela norma DICOM, mas cuja implementação se entendeu pertinente, na medida em que viabiliza um cenário de aplicação de extrema importância para as unidades de saúde, i.e. o upload de exames realizados no exterior por terceiras entidades contratadas para o efeito.

Aspectos de Engenharia de Software e Redes

No que diz respeito a ferramentas de desenvolvimento utilizámos, para além de HTML e Javascript, a linguagem JAVA [6] para a construção dos vários módulos aplicacionais de servidor e cliente, i.e. as Servlet [7, 8] e a Applet [9] respectivamente.

Para implementar a parte 18 da norma DICOM, foi criada a WadoClass, a qual contém métodos e estruturas de dados necessárias à interacção com os servidores PACS-DICOM internos. Foram também criados Servlets que funcionam como interfaces entre o browser e a classe ou, a nível prático, entre o clínico (utilizador) e o arquivo de imagens médicas (Figura 7). A manipulação de objectos DICOM foi efectuada com recurso ao PixelMed Java DICOM Toolkit [10] que apresenta, como principal vantagem, o facto de ser 'open-source'.

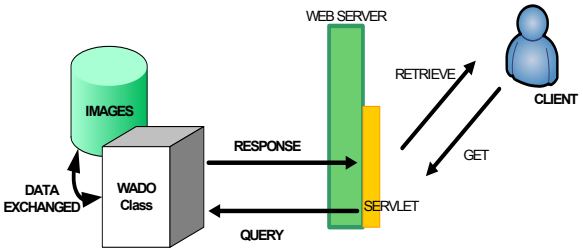


Figura 7: Interface entre o Cliente e o arquivo de imagens

Em termos de protocolo de comunicações (Figura 8), todas as interacções entre o cliente e o portal da instituição (i.e. o WebService) são WADO, i.e. objectos DICOM encapsulados no protocolo HTTP. Entre o WebService e os diversos servidores PACS da instituição é utilizado o protocolo DICOM convencional definido na Parte 8 da norma (Network Communication Support for Message Exchange) [11].

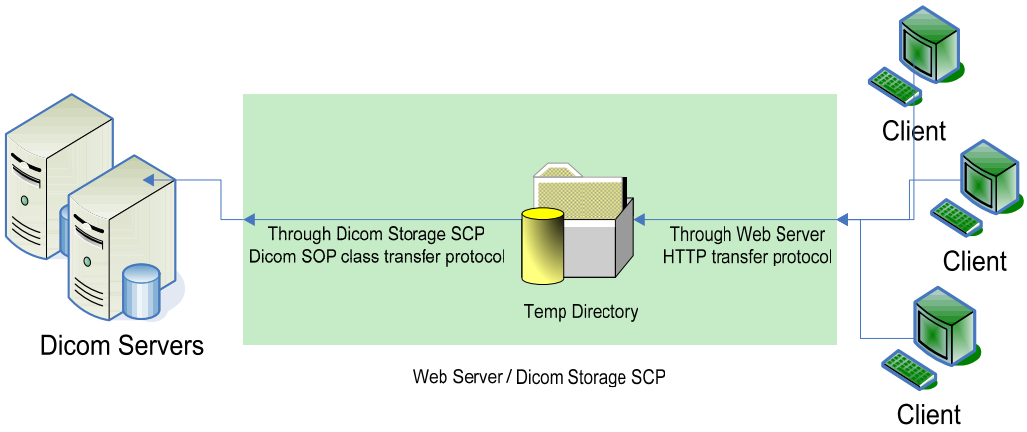


Figura 8: Protocolo de Comunicações de Dados

Em termos de parametrização, e de forma a tornar a plataforma mais escalável e facilmente reconfigurável, foi utilizado um ficheiro de configuração WEB.XML. Neste é possível incluir todas as definições dos servidores (endereços IP dos servidores DICOM, servidor WEB), bem

como informação sobre directorias temporárias, e servidor de armazenamento instalado no servidor Web, necessário para as transferências de imagens, dentro da instituição hospitalar/clínica.

6. Desenvolvimentos Futuros

Com a solução desenvolvida, é possível utilizar esta plataforma para transferência de objectos DICOM (incluindo as imagens) através do protocolo HTTP. No entanto, existem modalidades de imagem com elevados requisitos de largura de banda para transmitir a informação com custos temporais aceitáveis. Nomeadamente, estamos a falar de modalidades como o XA ou o US cardíacos [12].

Se a transmissão destes procedimentos na Intranet da instituição não levanta problemas, o mesmo não se pode garantir em acessos remotos. Neste sentido, estamos a realizar estudos para implementar estratégias de transmissão com compressão que não afecte a qualidade de diagnóstico dos estudos [13] (Figura 9).

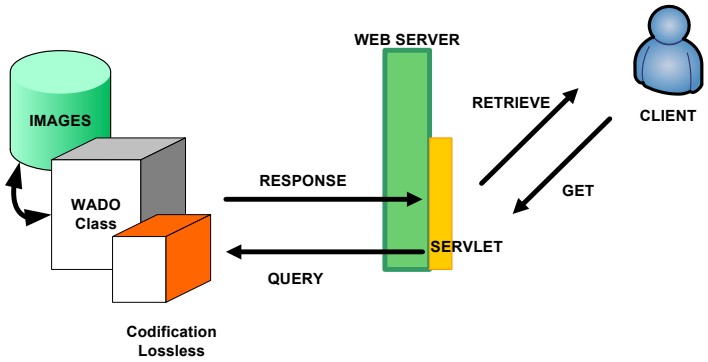


Figura 9: Interface entre o Cliente e o arquivo de imagens com codificação visual-lossless

7. Conclusões

Os sistemas PACS compreendem várias tecnologias que contemplam a aquisição, arquivo e distribuição ou importação de imagem em formato digital numa rede de computadores, para diagnóstico e revisão em estações dedicadas.

A implementação da norma DICOM na componente de transmissão de imagem contempla, na maioria das vezes, serviços de comunicação da camada de rede TCP/IP. No entanto, o carácter sensível da informação suportada pelos sistemas de informação da saúde leva a que os administradores de rede inibam este tipo de tráfego permitindo exclusivamente tráfego HTTP para o portal Web da instituição.

Neste sentido, e de forma a potenciar a partilha e acesso remoto a dados imagiológicos, desenvolve-se uma plataforma que permite encapsular objectos DICOM no protocolo HTTP. Em termos de exploração, a solução apresentada tem requisitos de recursos computacionais muito reduzidos, pelo que a sua instalação num qualquer servidor Web da instituição não compromete o seu normal funcionamento. Mais ainda, a solução apresentada pode ser transposta de forma simples e transparente para qualquer unidade de saúde que tenha um sistema PACS-DICOM, bastando para o efeito mexer exclusivamente no ficheiro de parametrização da plataforma desenvolvida.

8. Referências

[1] Hannan, T.J., *Variation in health care--the roles of the electronic medical record.* Int J Med Inf, 1999. **54**(2): p. 127-36.

[2] DICOM, *Digital Imaging and Communications in Medicine version 3.0*, ACR (the American College of Radiology) and NEMA (the National Electrical Manufacturers Association). http://medical.nema.org/dicom/2003.html.

[3] Huang, H.K., *PACS and imaging informatics: Basic Principles and Applications*. 2004: Wiley.

[4] Philips, *Philips "DICOM CookBook for Implementations in Modalities"*. 1997. ftp://ftp.philips.com/pub/pms-3003/DICOM_Information/CookBook.pdf.

[5] DICOM-P18, *Digital Imaging and Communications in Medicine (DICOM), Part 18: Web Access to DICOM Persistent Objects (WADO)*. 2004, National Electrical Manufacturers Association. http://medical.nema.org/dicom/2004/04_18PU.PDF.

[6] Sun, *Java*. http://www.java.sun.com.

[7] Bergsten, H., *JavaServer Pages, 2nd Edition*. August 2002: o'Reilly.

[8] Hall, M. and L. Brown, *Core Servelets and Javaserver Pages, 2nd Edition*: A Sun Microsystems Press/Prentice Hall PTR Book.

[9] Prokhorenko, A. and O. Prokhorenko, *Creating a Trusted Applet with Local File System Access Rights*. http://www.developer.com/java/data/article.php/3303561.

[10] Clunie, D., *PixelMed Java DICOM Toolkit*. http://www.pixelmed.com.

[11] DICOM-P8, *Digital Imaging and Communications in Medicine (DICOM), Part 8: Network Communication Support for Message Exchange*. 2004, National Electrical Manufacturers Association. http://medical.nema.org/dicom/2004/04_08PU.PDF.

[12] Soble, J.S., *MPEG Digital Video Compression.* Digital Cardiac Imaging in the 21st Century: A Primer, 1997: p. 192-197.

[13] Maede, A. and M. Deriche, *Establishing perceptual limits for medical image compression.* Proceedings of SPIE - Image Perception and Performance, 2001. **4324**: p. 204-210.

TEMA 1 – CICLO DE VIDA DOS SISTEMAS DE INFORMAÇÃO

1.3. Gestão de Sistemas e Tecnologias de Informação

Broadband Technologies and the Access Network

João Paulo Ribeiro Pereira [1], José Adriano Pires [2]

1) ESTiG - IPBragança, Bragança, Portugal

jprp@ipb.pt

2) ESTiG - IPBragança, Bragança, Portugal
adriano@ipb.pt

Abstract

Broadband technologies have view great growth over the past decade. But, the main part of this development and growth has been in the core networks, and the capacity of the access network to delivery broadband services remains as a challenge ("last mile problem"). The access network remains a bottleneck in terms of the bandwidth and service quality it affords the end user. By other side, the access network is much more spread geographically and covers larger areas. Then, this part of the network is usually the most expensive component in terms of capital investment and OAM cost. Some studies reefer that this networks required 70% of the total investment. Several access technologies can be used in this part of the network, which can be used to resolve the bandwidth bottleneck and the investment problem: xDSL, HFC, FTTx, FWA, WiMAX, PLC, Satellite, etc.

This paper presents a methodology and a tool that compare and evaluate broadband access technologies. The paper presents a techno-economic analysis of eight broadband technologies for access networks: digital subscriber line (DSL), hybrid fiber coax (HFC), power line communications (PLC), fiber to the home (FTTH), fiber to the curb (FTTC), fiber to the cabinet (FTTCab), and wireless alternatives such as WiMAX and satellite. Several actors (such operators, service providers, …) could use this tool to compare different technological solutions, forecast deployment costs, compare different scenarios, etc.

Palavras chave: Access networks, Broadband access technologies, Cost model.

1. Introduction

Full broadband coverage is highly important for bridging the digital divide. The digital divide expresses the difference in ability for people to communicate and utilize e-based business and services relative to their geographical location, their living standards and their level of education (Tardy *et al.*, 2004). It is viewed ultimately as an indicator of a country's economic and social situation. Access to ICT as a development tool for society is recognized as a political, economic and social issue of high importance. One of the factors mitigating this digital divide is broadband access, not only in urban areas, but also in rural and developing areas. However, the residential access network infrastructure remains one of the major obstacles to the deployment of a broadband telecommunications network.

An access network infrastructure must be established for delivery of a complex mix of broadband services like fast Internet access, telecommuting, and video on demand in a market of great demographic diversity (Ims *et al.*, 2004). Actually, residential broadband access is

mainly used for high-speed Internet and telephony services. The requirements for higher access capacities are primarily driven by the increase use of several broadband user terminals in the home and introduction of live TV streaming.

Until few years ago, Internet access for residential users was almost exclusively made via the public switched telephone networks (PSTN) over the twisted copper pair (Ibe, 2002). The new services requirements demand high-speed broadband access, and led to the development of several last-mile solutions (xDSL, HFC, FTTx, PLC, FWA, Satellite…) capable to support the services that internet has to offer.

The digital society increased the rapid bandwidth demand that put pressure on the network. Although national networks are largely upgraded to cope with this demand, a bottleneck remains over the 'last (or first) mile' between the customer and the first node in the network.

The methodology and tool presented in this paper has been developed to evaluate the most relevant broadband access technologies. These solutions include xDSL, HFC, PLC, FTTH, FTTC, FTTCab, and wireless alternatives such as WiMAX and satellite. The work should enable development of guidelines for broadband strategies.

2. Broadband Access Networks

Data transmission and networking technologies have view great growth over the past decade. But, the main part of this development and growth has been in the core networks where high capacity routers and ultra-high-capacity optical links have created a truly broadband infrastructure (Dravida *et al.*, 2002).

Globally it has been observed that several operators and vendors have chosen to deploy a hierarchic network model with three levels, taking into account geographical aspects and not necessarily functionality and capacity only (Ims *et al.*, 2004). A telecommunication network is commonly divided into three main parts (Figure 1): Core network; Access network; and Customer premise network.

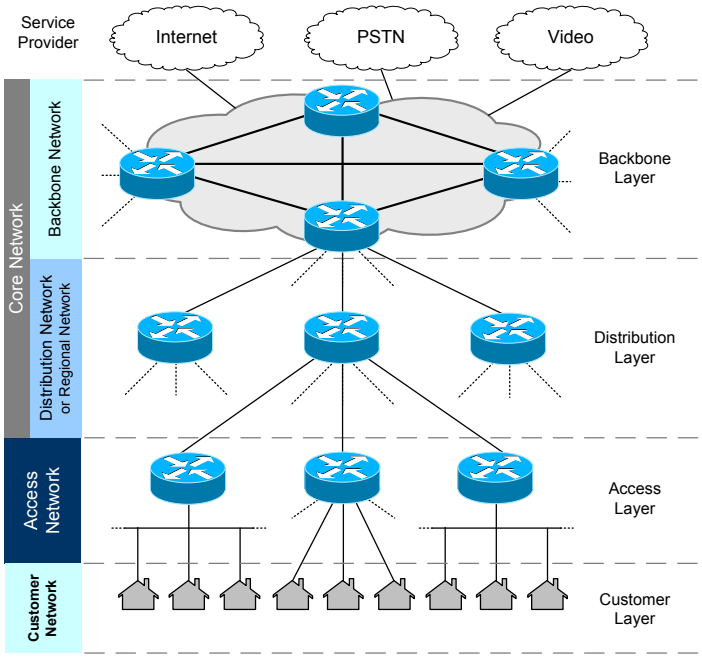


Figure. 1 - Hierarchical Telecommunication Network (Three levels).

Access Networks

Access network refers to the network between the Distribution / Regional network and the subscriber. The local access network is also often referred to as “the last mile” and “the local loop”. The access network remains a bottleneck in terms of the bandwidth and service quality it affords the end user. The last mile problem has impeded the growth of broadband services and applications.

The access network is much more spread geographically and covers larger areas. The available capacity on access lines is modest with respect to availability in aggregation/distribution and core parts of the network (Ims *et al.*, 2004). The capacity of the access network to delivery broadband services remains as a challenge ("last mile problem").

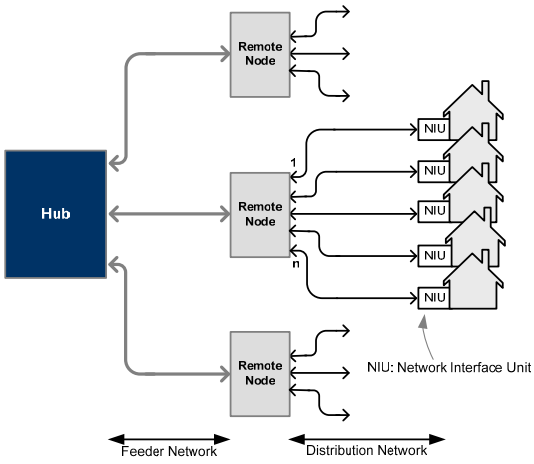


Figure. 2 - Architectures for Access Networks.

Then, this part of the network is usually the most expensive component in terms of capital investment and operation, administration and maintenance costs (OAM cost). Some studies reefer that this networks required 70% of the total investment.

The inadequacy of the access link (bottleneck problem) is particularly felt when the user require applications and services likely to become popular in the future. Interactive video applications, interactive gaming, video telephony, videoconferencing, remote storage, virtual DVD, and high-speed virtual private networks (VPNs) between geographically separated office locations or between homes and office locations for telecommuters are just a few such applications. It is likely that once the bandwidth and quality barrier of the access link is removed, new and unforeseen applications will emerge and attain widespread popularity. Many of these emerging applications cannot be supported with the asymmetric bandwidth available on today’s access technologies. Moreover, they require guaranteed quality of service (QoS) in terms of packet delays and throughput. Thus, the capability to provide symmetric high-quality high-bandwidth bearer services will be a key requirement for access media of the future (Dravida *et al.*, 2002).

As we have refer above, there is a increasing array of access technology available. The reason for this initially was the variety of types of access infrastructure available, for example twisted pair copper, cable TV co-axial networks, the electricity distribution system, point-to-point fiber, passive optical networks (PONs), and fiber rings. More recently, alternative systems that avoid the need for infrastructure have also emerged, such as radio, satellite and free-space optics (Cameron *et al.*, 2004). Several access technologies can be used in this part of the network: xDSL, HFC, FTTx, FWA, WiMAX, PLC, Satellite, etc. Each of these new access technologies has different performance, geographical and economic scaling attributes.

3. Broadband Access Technologies

Broadband access technologies are being deployed to address the bandwidth bottleneck for the "last mile," the connection of homes and small businesses to this infrastructure.

As we have see above, there are a many competing technologies which can provide the bandwidth required to deliver broadband services, but each technology has its limits in terms of bandwidth, reliability, cost or coverage. At present no single technology or network architectures seems the obvious choice.

In general, broadband access technologies can be classified by the physical medium in two major groups, namely (Ims *et al.*, 2004):

- Wired (or fixed line) technologies: The fixed line solutions communicate via a physical network that provides a direct "wired" connection from the customer to the service supplier. Some authors divide the wired technologies in Copper-based and fiber-based.
- Wireless technologies: Wireless solutions use radio or microwave frequencies to provide a connection between the customer and the operator's network.

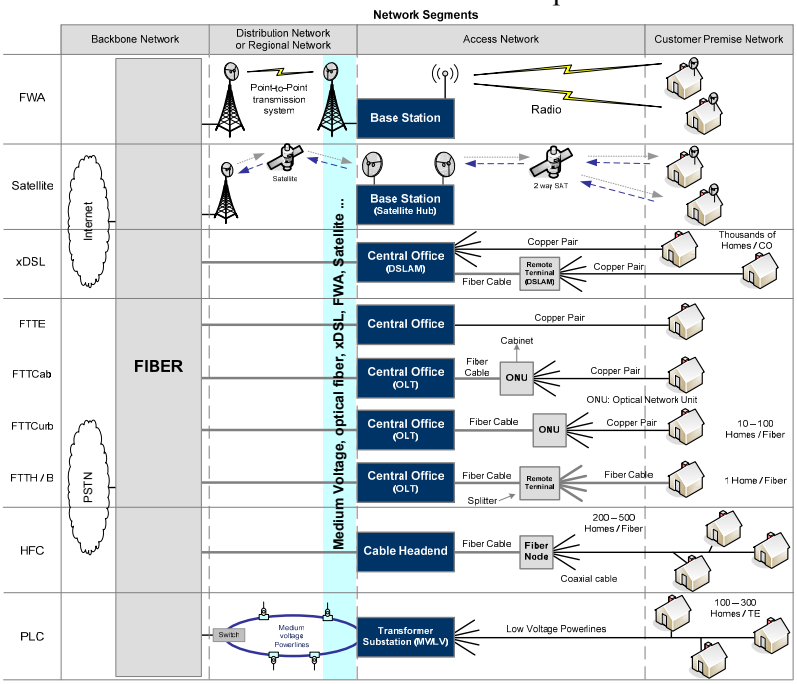


Figure. 4 - Broadband Access Network Technologies.

More recently, have emerged alternative systems that avoid the need for infrastructure. Wireless broadband is the "great equalizer" of broadband technologies. It enables even those without twisted copper, coaxial cable, fiber, or satellite to enter into competition for the broadband dollar. It therefore expands the definition of who can be a carrier (WiMAX Forum, 2005). Covering sparsely populated areas is not economically viable using conventional techniques. The newest wireless technologies are a solution that offers wide-area coverage well suited to small villages.

Today, a large variety of access network technologies and architectures are available for the operators to choose from. These include both narrowband and broadband technologies, with and without wires (Smura, 2004). The selection of the best solution requires understanding of the technical possibilities and limitations of the different alternatives, as well as understanding of the costs resulting from building and operating the networks.

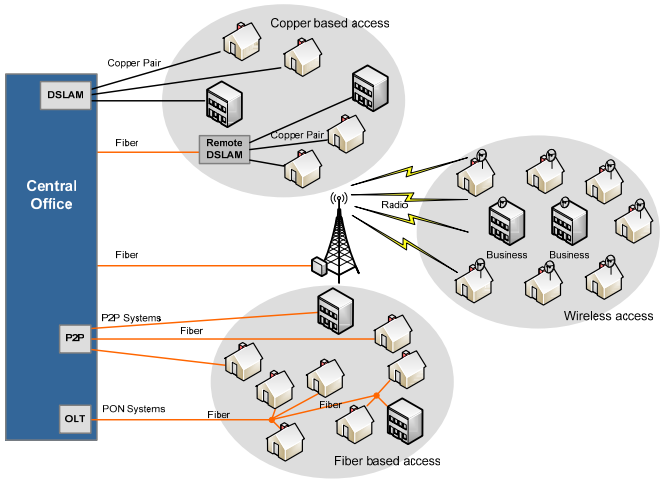


Figure. 4 - Combination of Access Technologies.

The choice of access technology depends on various variables like demography and geography. Others important variables are: Number of subscribers; Clients dispersion and service area; Existing infrastructures; Network architecture (Wireline, Wireless or hybrid); Services to support (like voice, data and video) and associated bit rate; Geographic characteristics; Infrastructure costs; Labor cost; Operation costs; and Access costs.

4. Techno-Economic Model

There are different network architectures that provide varied levels of service to the customer. Each architecture varies in complexity, network functionality, services supported and overall network costs. The network architectures and technologies will vary by cost, risk, flexibility, scalability, complexity and time to market considerations. It was argued in that the most critical parameters to include in a techno-economical model for broadband systems are subscriber density, civil works configuration, component cost evolution, and demand assessment (service penetration).

The Techno-Economic model is based on a modelling methodology for network value analysis that involves CAPEX and OPEX calculations.

Cost Model

The model analyzes several technical parameters (distances, bandwidth, equipment performance, etc.) and economic parameters (equipment costs, installation costs, service pricing, demographic distribution, etc.). Eight different technologies have been used to construct cost model. Network architectures considered include wireline technologies like xDSL, HFC, PLC, FTTH, FTTC, FTTCab, and wireless alternatives such as WiMAX and satellite.

Figure 5 show the framework architecture that supports our tool. The framework has three main components: Inputs, Processing and Outputs.

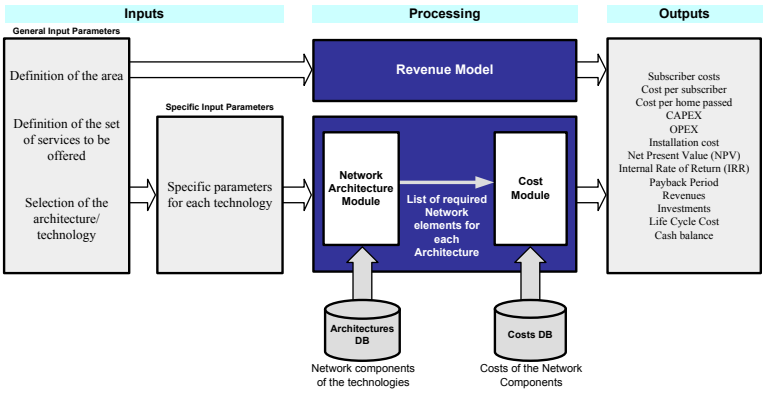


Figure. 5 - Framework Architecture.

Inputs

As we have seen above, the definition of the Input attributes is fundamental to obtain the right outputs. Then, we define three main activities:

- The area definition: Selection of the geographical area; Customer segments to be served; and Existing network infrastructure situation.
- The definition of the set of services to be offered: Services to be provided; Bandwidth; Service definitions for each user segment with adoption rates and tariffs.
- The selection of the architecture/technology to be used to provide the selected services: Network architecture and technologies; Cost of network equipment and installation; Cost of operation, administration, and maintenance procedures

The model divides the inputs into two main categories: general and specific input parameters. General parameters are those that describe the area and service characteristics and are common to all the technologies (figure 6). The specific parameters are those that characterize each solution, in technological terms (see figure 7).

Processing

Architecture Module

This part of the framework is the responsible for the main outputs production. Architectures Database contains information about the components necessary to construct network architecture. This module is used to calculate the number of network elements for the set of services and the network architectures defined. The other database used in this framework is the Cost Database, containing information about the costs of the all components of each technology in study. This Database is constantly updated with data gathered from the market. The quality of the Database has a big influence in the output results. For each technology, the following attributes are necessary input to the cost module: The equipment price in a given year; Price trends; Operation, administration, and maintenance cost parameters. The technical architectures are translated into basic cost structures.

Revenues Module

This module uses some of the general input parameters to calculate the residential and business estimated revenues. The model Revenue assumptions: Type of subscribers: Residential household; and Business customer; Tariffs / fees: Yearly trend of the tariffs/fees (Connection fee (€) and Subscription fee (€ / month)).

Output Requirements

The cost module gives a range of outputs necessary for the calculation of the necessary outputs. The outputs calculated are (Figure 8): Subscriber costs; ARPU: Average Revenue Per User; Cost per subscriber; Cost per home passed; Mbit cost; OAM costs; Installation cost; Net Present Value (NPV); Internal Rate of Return (IRR); Payback Period; Revenues; Investments; Life Cycle Cost; Cash balance

5. Tool Description

The objective of this tool is the calculation of several outputs in a given scenario, and for several BB access technologies. The tool implements a methodology for the techno-economic analysis of access networks for residential / business customers, and was being developed in MSExcel. Figure 6 show the general input form used by the tool, where the user introduce the parameters of a given scenario:

- Area Characteristics: Geographical Area to be Covered (Km2); Number of buildings in serving area; Building density per km2; Average number of floors per building; Average number of living units per floor; Average number of living units per building; Total number of Living Units in area (# of potential customers); Customers units density per Km2; % Residential/Business; Total number of potential subscribers; Technology penetration rate (expected market penetration); Number of subscribers; Subscribers density per km2.
- Service Characteristics: Required Downstream bandwidth (kbps); Required Upstream bandwidth (kbps); Expected traffic demand in the area (Mbps); Expected traffic demand in the area (Mbps) per Km2; QoS for customer: % of time the Service Level Agreement is satisfied; One-time Activation/connection fee (€); Subscription fee (€ / month).

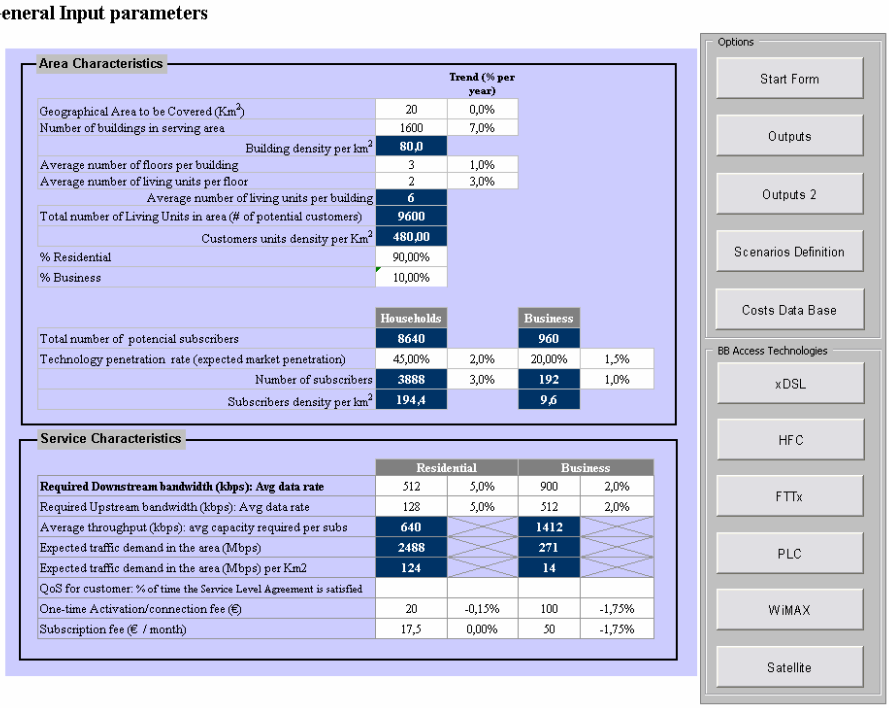


Figure. 6 - General Input Parameters

After the introduction of the general parameters, are required specific parameters for each technology. Figure 7 show the example of the specific parameters for the FTTx technology. The tool calculates the cost and shows the results in table, graphics and schemes (Figure 8 and 9).

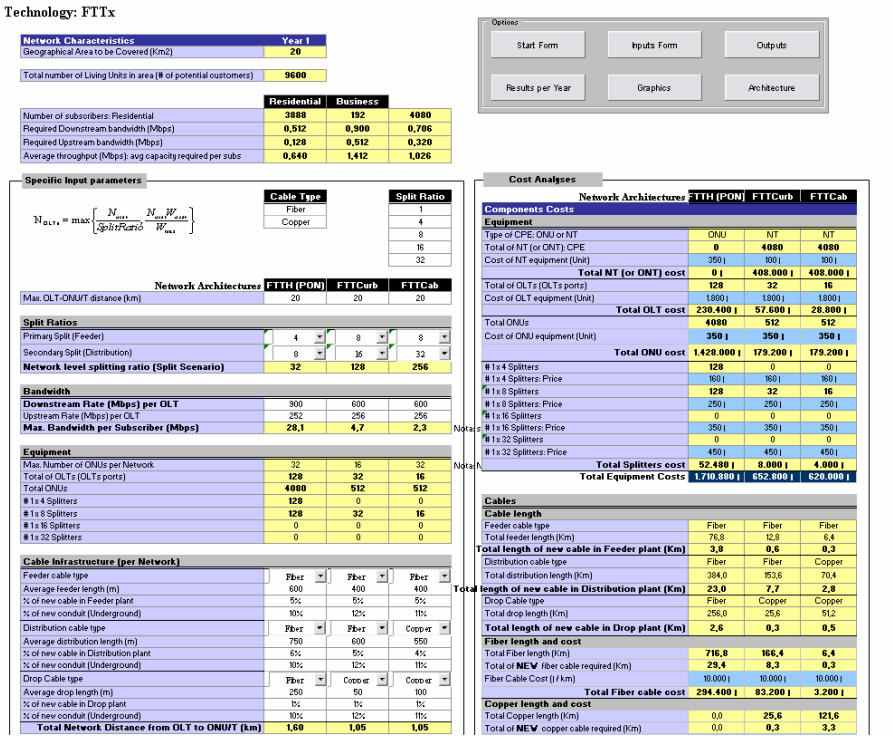


Figure. 7 - Specific Input Parameters (example of FTTx Technology)

The outputs produced by the tool are presented in many formats (table, graphics and pictures). Figure 8 show some of the outputs calculated for all the solutions used in this study.

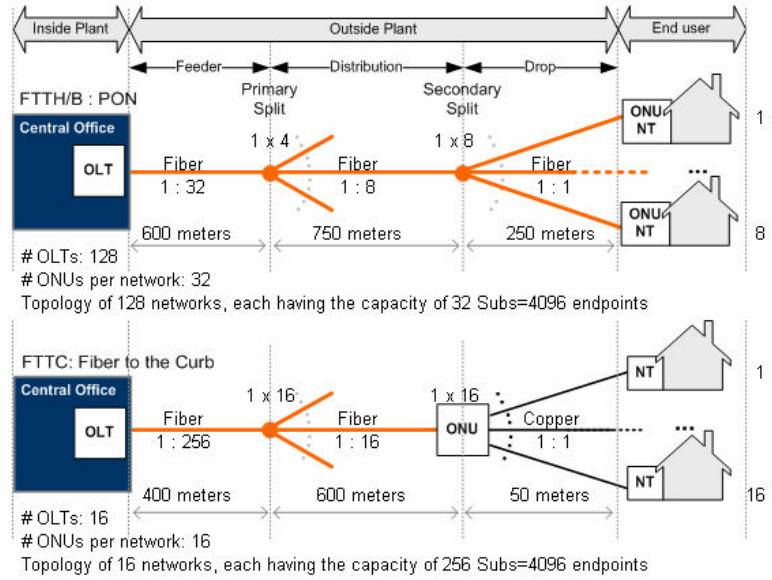
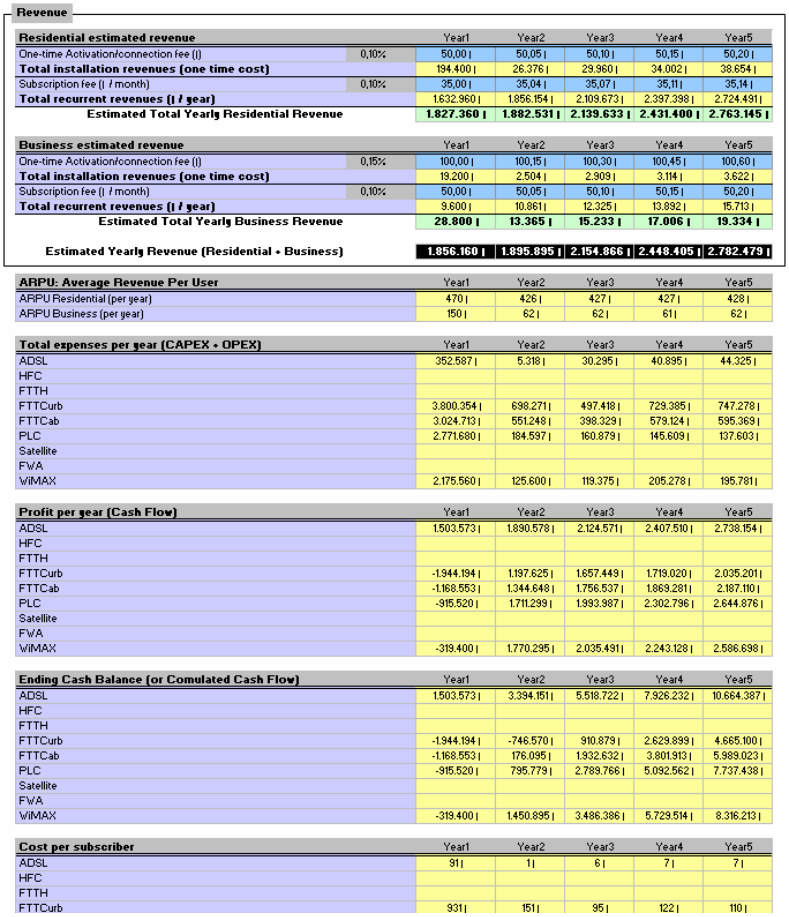


Figure. 8 - Outputs (Example of FTTx Technology)



Revenue

Residential estimated revenue		Year1	Year2	Year3	Year4	Year5
One-time Activation/connection fee (€)	0,10%	50,00 €	50,05 €	50,10 €	50,15 €	50,20 €
Total installation revenues (one time cost)		194.400 €	26.376 €	29.960 €	34.002 €	38.654 €
Subscription fee (€ / month)	0,10%	35,00 €	35,04 €	35,07 €	35,11 €	35,14 €
Total recurrent revenues (€ / year)		1.632.960 €	1.856.154 €	2.109.673 €	2.397.398 €	2.724.491 €
Estimated Total Yearly Residential Revenue		**1.827.360 €**	**1.882.531 €**	**2.139.633 €**	**2.431.400 €**	**2.763.145 €**

Business estimated revenue		Year1	Year2	Year3	Year4	Year5
One-time Activation/connection fee (€)	0,15%	100,00 €	100,15 €	100,30 €	100,45 €	100,60 €
Total installation revenues (one time cost)		19.200 €	2.504 €	2.909 €	3.114 €	3.622 €
Subscription fee (€ / month)	0,10%	50,00 €	50,05 €	50,10 €	50,15 €	50,20 €
Total recurrent revenues (€ / year)		9.600 €	10.861 €	12.325 €	13.892 €	15.713 €
Estimated Total Yearly Business Revenue		**28.800 €**	**13.365 €**	**15.233 €**	**17.006 €**	**19.334 €**
Estimated Yearly Revenue (Residential + Business)		**1.856.160 €**	**1.895.895 €**	**2.154.866 €**	**2.448.405 €**	**2.782.479 €**

ARPU: Average Revenue Per User	Year1	Year2	Year3	Year4	Year5
ARPU Residential (per year)	470 €	426 €	427 €	427 €	428 €
ARPU Business (per year)	150 €	62 €	62 €	61 €	62 €

Total expenses per year (CAPEX + OPEX)	Year1	Year2	Year3	Year4	Year5
ADSL	352.587 €	5.318 €	30.295 €	40.895 €	44.325 €
HFC					
FTTH					
FTTCurb	3.800.354 €	698.271 €	497.418 €	729.385 €	747.278 €
FTTCab	3.024.713 €	551.248 €	398.329 €	579.124 €	595.369 €
PLC	2.771.680 €	184.597 €	160.879 €	145.609 €	137.603 €
Satellite					
FWA					
WiMAX	2.175.560 €	125.600 €	119.375 €	205.278 €	195.781 €

Profit per year (Cash Flow)	Year1	Year2	Year3	Year4	Year5
ADSL	1.503.573 €	1.890.578 €	2.124.571 €	2.407.510 €	2.738.154 €
HFC					
FTTH					
FTTCurb	-1.944.194 €	1.197.625 €	1.657.449 €	1.719.020 €	2.035.201 €
FTTCab	-1.168.553 €	1.344.648 €	1.756.537 €	1.869.281 €	2.187.110 €
PLC	-915.520 €	1.711.299 €	1.993.987 €	2.302.796 €	2.644.876 €
Satellite					
FWA					
WiMAX	-319.400 €	1.770.295 €	2.035.491 €	2.243.128 €	2.586.698 €

Ending Cash Balance (or Comulated Cash Flow)	Year1	Year2	Year3	Year4	Year5
ADSL	1.503.573 €	3.394.151 €	5.518.722 €	7.926.232 €	10.664.387 €
HFC					
FTTH					
FTTCurb	-1.944.194 €	-746.570 €	910.879 €	2.629.899 €	4.665.100 €
FTTCab	-1.168.553 €	176.095 €	1.932.632 €	3.801.913 €	5.989.023 €
PLC	-915.520 €	795.779 €	2.789.766 €	5.092.562 €	7.737.438 €
Satellite					
FWA					
WiMAX	-319.400 €	1.450.895 €	3.486.386 €	5.729.514 €	8.316.213 €

Cost per subscriber	Year1	Year2	Year3	Year4	Year5
ADSL	91 €	11 €	6 €	7 €	7 €
HFC					
FTTH					
FTTCurb	931 €	151 €	95 €	122 €	110 €

Figure. 9 - Outputs (Some of the outputs produced)

6. Conclusion

The broadband access network remains a major challenge for operators due to the high cost sensitivity of this network segment. In this paper we present a model and a tool which can be used in the development of access networks. The tool calculates several outputs, and can be used to compare the economic viability of different network architectures in a given scenario. One of the objectives is to provide guidelines for broadband introduction in the access networks.

7. References

1. Cameron,A.F., Thorne,D.J., Foster,K.T. & Fisher,S.I. (2004) Fixed access network technologies. *BT Technology Journal,* **22**, 48-59.

2. Dravida,S., Gupta,D., Nanda,S., Rege,K., Strombosky,J. & Tandon,M. (2002) Broadband access over cable for next-generation services: a distributed switch architecture. *IEEE Communications Magazine,* **40**, 116-124.

3. Ibe,O.C. (2002) *Fixed Broadband Wireless Access Networks and Services*. John Wiley & Sons, Inc..

4. Ims,L.A., Bhatnagar,A., Østlyngen,E. & Bozorgebrahimi,K. (2004) Towards the next generation broadband network platform. *Telektronikk,* 107-125.

5. Smura,T. (2004) Techno-Economic Analysis of IEEE 802.16a-Based Fixed Wireless Access Networks. In: Helsinki University of Technology.

6. Tardy,I., Bråten,L., Bichot,G., Settembre,M. & Seseña,J. (2004) Hybrid architecture to achieve true broadband access in rural areas. In: *BroadBand Europe,* pp. 1-6.

7. WiMAX Forum (2005) WiMAX Deployment Considerations for Fixed Wireless Access in the 2.5 GHz and 3.5 GHz Licensed Bands. In: WiMAX Forum (ed.).

A influência da Economia de Rede na opção por *standards* de Sistemas de Informação : elementos para a construção de uma métrica

João Carlos do Rosário [1], António Palma dos Reis [2]

1) ESCS - Escola Superior de Comunicação Social, Lisboa, Portugal

jrosario@escs.ipl.pt

2) ISEG – Instituto Superior de Economia e Gestão, Lisboa, Portugal
apreis@iseg.utl.pt

Resumo

Este artigo identifica os principais factores que influenciam a evolução dos mercados dos Sistemas de Informação, como externalidades de rede, *switching costs*, e *lock-in*, que constituem especificidades dos mercados dos Sistemas de Informação quando comparados com outros mercados, levando a diferentes processos de decisão por parte de fornecedores e consumidores. Após a análise do processo de decisão sobre a opção entre *standards Open Source* e *standards* proprietários, tanto do ponto de vista dos fornecedores como dos consumidores do mercado, é proposto um "Indicador de Escolha de *Standards* nos SI" a partir da modelização do processo de decisão dos consumidores.

Palavras chave: efeito de rede, externalidades de rede, *switching costs*, *lock-in*, *standards, Open Source*

1. Introdução

A partir dos anos 80 do século XX, com o desenvolvimento dos mercados de tecnologias de informação e comunicação, começou a verificar-se que, nesses mercados, o valor para o consumidor dos produtos ou serviços aumentava com o número de consumidores do mercado, o que não podia ser explicado à luz das teorias económicas tradicionais. Tal levou a que a comunidade académica começasse a investigar estes mercados de uma forma mais sistemática. Como resultado foram-se desenvolvendo novas teorias que viriam a constituir uma nova área de investigação com a denominação de Economia de Rede, ambiente em que se insere o objecto de estudo deste artigo, o processo de decisão entre *standards* incumbentes e suas alternativas e entre *standards* proprietários e *Open Source* no mercado dos Sistemas de Informação.

2. Revisão da Literatura

A revisão de literatura apresenta os conceitos relacionados com a Economia de Rede, tais como externalidade de rede, que ocorre quando os consumidores constituem uma rede e o valor desta para o consumidor aumenta com o número de consumidores a aderirem à rede; *switching costs*, o custo de mudança entre tecnologias ou *standards*; e *lock-in*, que ocorre quando o consumidor tem demasiados custos de mudança pelo que ele próprio em particular, e os vários consumidores que constituem a rede em geral, ficam "presos" a uma dada tecnologia mesmo que esta já esteja ultrapassada.

Também será analisada a forma como estes factores influenciam a escolha dos consumidores, particularmente a escolha entre *standards* proprietários e *standards Open Source*.

2.1 . Externalidade ou efeito de rede

Em certos mercados o consumidor obtém maior utilidade dos produtos e serviços que consome quanto maior o número de consumidores que já aderiram ao produto ou serviço em questão. Por exemplo, quanto maior o número de consumidores com telemóveis 3G, maior o valor para o consumidor actual e para o consumidor potencial da rede 3G dado que obtêm maior beneficio por comunicar por videoconferência com um número de utilizadores cada vez mais alargado e de terem acesso a um conjunto de serviços de conteúdo multimédia cada vez mais alargado. Kaz e Shapiro [1985] denominaram este efeito de externalidade de rede ou efeito de rede, servindo de ponto de partida para um conjunto de conceitos nesta área de investigação. Para estes autores existem duas formas da externalidade de rede se manifestar:

- Externalidade de rede directa, em que o aumento de utilidade do consumidor advém de existirem mais consumidores a consumir o mesmo produto ou serviço, efeito existente noo mercado das telecomunicações móveis e das comunicações por fax, por exemplo.
- Externalidade de rede indirecta, em que o aumento de consumidores num dado mercado leva ao aumento da oferta de produtos ou serviços complementares [Katz e Shapiro 1985], [Economides 1996] como é o caso dos sistemas operativos (Windows, MacOS, Linux) em relação ao *software* ou periféricos.

2.2 . *Switching Costs e Lock-in*

Após a selecção e adopção de uma dada tecnologia ou *standard,* ou seja a escolha da rede onde estará presente, o consumidor terá naturalmente custos se no futuro quiser mudar para uma nova rede mesmo que esta tenha tecnologia ou *standards* superiores à rede onde ele está presente actualmente. Esses custos, denominados de ***switching costs***, serão tanto maiores quanto maior for a dimensão da rede actual, pela perda por parte do consumidor dos benefícios da externalidade que a rede actual lhe oferece**.** Adicionalmente os *switching costs* podem incluir os custos de transacção, a compra de novos equipamentos e software; os custos de aprendizagem necessários à aprendizagem para a utilização de novos equipamentos e software, e; os custos contratuais, como é o caso, por exemplo, da perda de benefícios relacionados com campanhas de fidelização de cliente [Klemperer 1987].

Por outro lado, não mudando para a nova rede, o consumidor permite a manutenção da dimensão da rede actual, mesmo que com tecnologia inferior à tecnologia da rede alternativa. A manutenção da dimensão da rede actual levará mais consumidores a aderirem a essa rede dado que quanto maior a dimensão da rede maior a utilidade para o consumidor que a ela adira. Ou seja, o adiamento da decisão de mudança do consumidor devido aos *switching costs* poderá levar a que o mercado fique em situação de ***lock-in*** ("preso") a uma tecnologia técnicamente inferior. [Farrel e Saloner 1985, 1986].

2.3 . *Standards*

Um *standard* é um conjunto de especificações que determina a compatibilidade de diferentes produtos. São estas especificações que permitem que os telefones, telemóveis e faxes comuniquem uns com os outros; que todos os leitores de DVD, CD, DAT consigam ler DVDs, CDs e cassetes DAT; que se possam trocar ficheiros de MS-Office ou correr aplicações "compatíveis Windows". Ou seja, são os *standards* que permitam a compatibilidade no mercado. Também existem *standards* fora dos mercados das Tecnologias de Informação e Comunicação (TIC) cuja análise está fora do âmbito deste artigo.

A existência de *standards* em mercados com externalidades de rede, como é o caso das TIC, leva a que nestes mercados se possa verificar uma convergência para um dado *standard* beneficiando dessa forma a empresa ou empresas que o utilizam nos seus produtos ou serviços.

2.3.1. Competição entre *standards*

Besen e Farrell [1994] distinguiram quatro diferentes formas de definição de *standards* num mercado, ou seja, as formas de um dado *standard* vir a predominar sobre os outros tornando-se o *standard* dominante, ou mesmo único, no mercado:

1. Competição entre *standards* proprietários incompatíveis;
2. Competição sobre compatibilidade, quando as empresas fornecedoras do mercado optam por não lançar novos *standards* para competir com o incumbente, mas antes lançar produtos compatíveis com o *standard* incumbente através dos chamados "adaptadores", como é o caso dos filtros para tornarem os ficheiros do OpenOffice ou Word Perfect compatíveis com o MS-Office;
3. *Standards* definidos por acordos voluntários;
4. *Standards* mandatados pelo Estado.

As duas últimas formas de definição de *standards*, por acordo e mandatados, criam os chamados *standards de jure,* como é o caso do GSM, 3G, OpenDocument Format e dos protocolos da Internet. As duas primeiras formas criam os *standards de facto,* através de processos competitivos de mercado sem coordenação entre os intervenientes no mercado ou entidades externas ao mercado, como é o caso daqueles criados por empresas como a Microsoft, Autodesk, Adobe ou Winzip que hoje dominam os mercados de *software* "Office", *software* de CAD, *software* de edição electrónica e *software* de compressão de ficheiros, respectivamente

No que diz respeito às duas primeiras formas de definição de *standards*, que criam *standards de facto*, as empresas que entram no mercado têm de decidir entre utilizar o *standard* incumbente ou avançar com um novo *standard* contra o incumbente. Segundo Liebowitz e Margolis [1996], existem várias condições para uma empresa poder vencer um *standard* incumbente. Em primeiro lugar, o novo *standard* terá de ser técnicamente superior ao incumbente e, nesse caso, é de esperar que esse novo *standard* venha a ultrapassar e substituir o incumbente. Para o novo *standard* se impor mais fácilmente poderá também ser necessário ter um preço baixo para a penetração no mercado ou uma forte campanha de comunicação.

2.3.2. Decisões dos consumidores na escolha dos *standards*

Dalle [1997] analisou a influência conjunta do efeito de rede e do grau de heterogeneidade dos consumidores na definição dos *standards*. Segundo Dalle, o efeito de rede fomenta a coordenação das decisões dos consumidores e a heterogeneidade dos consumidores fomenta a

diversidade das suas decisões. Desta forma, uma empresa que introduza um novo *standard* num mercado com algum grau de heterogeneidade dos consumidores deve, em primeiro lugar, procurar atingir nichos de mercado que permitam obter uma massa crítica de consumidores.

Dalle [1997] introduziu também o conceito de efeito de rede local. O efeito de rede referido até aqui neste artigo, denominado por este autor como efeito de rede global, é aquele em que o consumidor tem em conta o número de consumidores que já aderiram ou espera que venham a aderir a um dado *standard*. Mas, nem sempre o consumidor tem informação sobre as quotas de mercado para poder optar por um *standard* com base no número de consumidores que já aderiram a esse *standard*. Então irá utilizar aconselhamento sobre a decisão de compra em relação a um dado conjunto de *standards* a partir da sua rede de contactos: o conjunto de amigos, familiares, fornecedores, clientes ou concorrentes. A influência desse aconselhamento na decisão do consumidor denomina-se efeito de rede local. A existência de efeitos de rede globais e locais associados ao grau de heterogeneidade dos consumidores poderá levar então a que não haja convergência para um único *standard* no mercado e coexistam vários *standards*, incluindo *standards* técnicamente ultrapassados.

2.3.3. Decisões das empresas fornecedoras dos *standards*

Jongseok et al. [2003] analisaram, no quadro dos Sistemas de Informação, a opção de uma empresa entre introduzir uma nova tecnologia compatível com o *standard* dominante do mercado ou criar um novo *standard,* superior e incompatível. A análise tem como base de estudo o mercado dos microprocessadores, em que estão presentes os microprocessadores Intel e AMD com o *standard* dominante x86 da Intel, e os microprocessadores RISC incompatíveis utilizados principalmente em potentes computadores denominados estações de trabalho. Os autores concluiram que uma empresa só deveria criar um novo *standard* se a maioria dos consumidores for *power-users* (utilizadores técnicamente competentes e exigentes em termos de desempenho dos SI), ou se o *standard* incumbente ainda estiver no seu início no mercado, sem que a massificação se tenha dado e a força do efeito de rede influenciado fortemente as opções dos consumidores. No caso de uma alta taxa de penetração de mercado por parte do *standard* incumbente, e quando os *power-users* no mercado são poucos ou nenhuns, verifica-se a existência de *lock-in* no *standard* incumbente. O *lock-in* surge devido à inércia por parte dos consumidores do *standard* incumbente, motivada pelo facto de que o efeito de rede actual os beneficia mais do que a mudança para um novo *standard*, técnicamente superior, mas sem vantagens na utilização do dia-a-dia para quem não é *power-user*. Nesse caso, a empresa que entra no mercado deve procurar criar um "adaptador" de compatibilidade com o *standard*

incumbente, como fez a AMD no mercado de microprocessadores, ou evitar entrar no mercado dado que as probabilidades de sucesso são pequenas.

2.4. Concorrência com o *standard* incumbente: o caso do *Open Source*

Os conceitos de Economia de Rede enquadram a análise da concorrência entre um *standard* proprietário incumbente e um *standard* aberto *Open Source* no mercado dos Sistemas Operativos e aplicações para computadores pessoais e servidores, como se passa a explicitar.

2.4.1. O que é o *software Open Source*

O *software Open Source* é aquele pode ser modificado, desenvolvido, adaptado e incorporado noutros programas pelos programadores, sem pagamento de direitos a qualquer dos autores do mesmo. A restrição mais comum é que quaisquer futuras modificações ou programas derivados devem ser também *Open Source* [Schiff 2002]. O licenciamento do *Open Source* é gratuito.

O *software* proprietário, por outro lado, é normalmente distribuido em código binário (*Closed Source*) sendo técnicamente difícil de modificar e tendo geralmente direitos de propriedade que o impedem. O seu licenciamento envolve custos, com excepção do *software* direccionado para o consumidor individual como forma de divulgação de versões mais completas mas não gratuitas (AVG Free, por exemplo), e do *software* que tem como objectivo a definição de *standards* de mercado (Internet Explorer, MS Media Player), cujo licenciamento é gratuito.

2.4.2. O *Open Source* no mercado

Só muito recentemente começou a surgir investigação aplicada relevante sobre *Open Source*, suas características e condições para o sucesso no mercado. Um dos primeiros trabalhos nesta área é de Bonaccorsi e Rossi [2003]. Neste trabalho foram analisados três aspectos-chave do *Open Source*: ser resultado da actividade de uma comunidade sem motivações de ganhos financeiros, com uma coordenação hierárquica sem direitos de propriedade, e conseguindo difundir-se em mercados dominados por *standards* proprietários incumbentes.

Bonaccorsi e Rossi [2003] defendem, de acordo com a literatura existente, a existência nos mercados dos Sistemas de Informação de externalidades de rede, e *lock-in*, que no entanto poderá ser ultrapassado. Consideram também que se houver heterogeneidade dos consumidores e efeito de rede local, conforme também referido por Dalle [1997], e os custos de mudança para um novo *standard* não forem muitos elevados, será possível que o *standard* proprietário incumbente seja ultrapassado pelo *standard Open Source*.

Para Bonaccorsi e Rossi [2003], se existir o efeito de rede a favor do *standard* incumbente baseado no número de consumidores que já o utilizam, também existe no caso do *Open Source* um efeito de rede, neste caso relacionado com o acesso legítimo dos utilizadores de *Open Source* a uma grande variedade de *software* gratuito, variedade esta proporcional ao número de programadores que aderem ao *Open Source*. Será a relação entre estes dois efeitos, efeito de rede do *standard* incumbente e efeito de rede do *standard Open Source* devido ao acesso a uma variedade de aplicações gratuitas, que influenciará a evolução do mercado a favor de um *standard* ou outro. A análise da evolução do mercado terá também de ter em conta que o *standard* incumbente reagirá à pressão concorrencial aumentando os recursos para Investigação e Desenvolvimento, melhorando a qualidade do seu *software* proprietário. Desta forma vários resultados poderão surgir no mercado, sendo mesmo possível a coexistência entre o *standard* proprietário incumbente e o *standard Open Source*.

Lin [2004] analisou a evolução de um mercado em que estão em concorrência um *standard* proprietário incumbente e um *standard Open Source*, tomando como pressuposto a heterogeneidade dos consumidores. Os seus resultados vão no mesmo sentido da investigação já aqui apresentada, mas o autor adicionalmente analisa duas situações :

- Se existe falta de recursos humanos com capacidade técnica e experiência para a implementação do *software Open Source*, o custo de implementação do *standard* gratuito *Open Source* é relativamente elevado o que permite ao *standard* proprietário incumbente manter preços elevados e ser pouco afectado na sua rentabilidade por uma descida da sua quota de mercado, descida essa que é sempre atenuada pelo efeito de rede com que o *standard* incumbente mantém e atrai consumidores.
- Conforme aumenta o número daqueles que têm capacidade técnica e experiência de implementação do *software Open Source*, o custo global do *standard* gratuito *Open Source* desce, aumentando a pressão concorrencial sobre o *standard* proprietário incumbente e levando os preços deste a descerem.

Desta forma, o *software Open Source* poderá não vencer o *software* proprietário incumbente porque a empresa detentora deste poderá decidir estratégicamente sobre o preço de forma a capturar ou recuperar quota de mercado de forma a manter a rentabilidade e o efeito de rede do seu *standard*. Então, se o *Open Source* tiver apenas como vantagem o baixo custo, poderá não ter sucesso do mercado. Como resultado, em mercados com efeito de rede e no longo prazo o *standard Open Source* apenas terá sucesso se apresentar melhores características relevantes para o consumidor que o *standard* proprietário incumbente.

3. Modelização do processo de decisão do consumidor

Com base na revisão da literatura pode definir-se um modelo que representa o processo de tomada de decisão do consumidor na sua escolha entre o *standard* incumbente e *standards* alternativos. Os factores que influenciam a escolha entre o *standard* incumbente e o *standard* alternativo, independentemente de qual destes *standards* é o proprietário e qual é o *Open Source*, estão representados na Figura 1. Neste modelo e através dos símbolos "+" e "-" é representada a influência de cada factor na opção do consumidor pelo *standard* incumbente ou manutenção no mesmo. É essa influência que permite avaliar, para cada consumidor no mercado de Sistemas de Informação, qual a probabilidade deste optar pelo *standard* incumbente, pelo *standard* alternativo ou de mudar se já tiver aderido a um desses *standards*. A variável dependente, que define um indicador, é a "Opção pelo *standard* incumbente em detrimento do *standard* alternativo". Este indicador é calculado com base na situação actual do consumidor em termos de SI e das opiniões e percepções do consumidor, recolhidas para cada segmento do mercado tendo em conta os factores indicados.

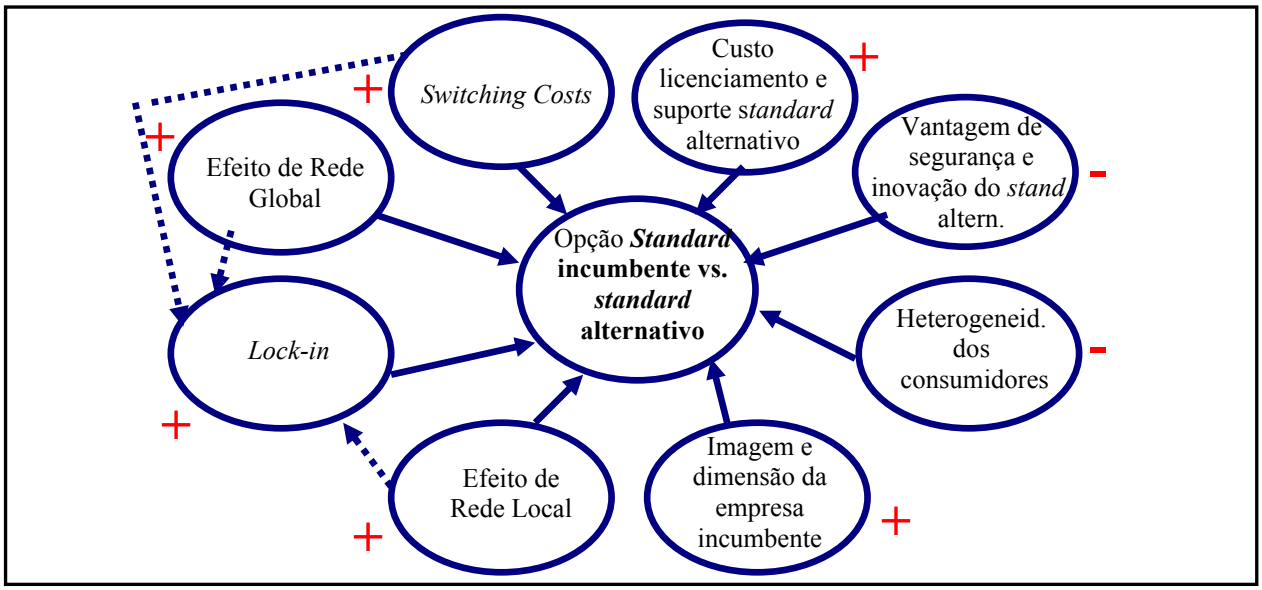


Figura 1 - Modelo de decisão sobre a opção entre *standard* incumbente e *standard* alternativo *(+)* e *(-)* Factores que influenciam a opção pelo *standard* incumbente ou alternativo respectivamente

Para o cálculo do indicador são realizados inquéritos a amostras representativas dos diferentes segmentos de mercado que se pretende analisar. O inquérito é constituido por um conjunto de questões fechadas multi-resposta e questões com escala de Liketer para cada um dos factores apresentados na Figura 1, estabelecendo desta forma o valor da métrica denominada "Opção pelo *standard* incumbente em detrimento do *standard* alternativo".

O modelo proposto será testado com base em dados recolhidos por um inquérito administrado a uma amostra aleatória estratificada pela dimensão das empresas portuguesas procurando representar a estrutura empresarial portuguesa, com base nas listagens "500 Maiores Empresas Portuguesas" e "1000 maiores PMEs" publicadas por jornais de referência. Para os consumidores finais podem ser utilizadas amostras aleatórias ou amostras de conveniência para ajudar a compreender o processo de decisão deste segmento.

4. Conclusões

Existe um conjunto de factores que influencia o processo de escolha de *standards* no mercado de Sistemas de Informação. Se por um lado alguns destes factores podem reter um mercado num dado *standard* incumbente, também se reconhece que o carácter inovador de um *standard* alternativo, ainda mais se estiver associado a um preço competitivo, criará condições para que a inovação vença e os mercados evoluam. Essa evolução poderá acontecer através do *standard* alternativo ou mesmo através do *standard* incumbente, que se verá obrigado a inovar e tornar-se competitivo em termos de preço para manter a sua posição no mercado, o que é benéfico para o mercado em geral e para os consumidores em particular.

Sendo o processo de escolha de *standard* por parte dos consumidores influenciado por um conjunto de factores como foi já referido, deve também ter-se em atenção que não só a influência dos vários factores é diferente de segmento de mercado para segmento de mercado (Particulares, Grandes Empresas e Pequenas e Médias Empresas por exemplo), como também pode ser diferente devido à heterogeneidade de cada segmento de mercado, heterogeneidade essa que pode ser avaliada pela dimensão e características sócio-económicas dos consumidores que constituem o segmento de mercado.

Torna-se importante para um melhor conhecimento do mercado dos Sistemas de Informação a criação do "Opção pelo *standard* incumbente em detrimento do *standard* alternativo". Este indicador será desta forma útil para análises previsionais de evolução do mercado e tomadas de decisão por parte dos consumidores e fornecedores. O indicador será calculado para os segmentos de mercado em que se pretende analisar o processo de decisão sobre o *standard* de Sistemas de Informação, baseando-se na recolha da situação actual, opiniões e percepções dos consumidores no mercado. O objectivo do indicador é dar a conhecer probabilidades de escolha de um dado *standard*

pelos consumidores sem que a recolha de dados junto dos mesmos os obrigue a revelar uma opção explícita entre os *standards* existentes no mercado, mas que o indicador revele a probabilidade de cada uma das opções (*standard* incumbente versus *standard* alternativo) para cada cliente do mercado analisado.

5. Referências

Besen, S. e Farrel, J., "Choosing How to Compete: Strategies and Tactics in Standardization", *Journal of Economic Perspectives*, 8 (1994), pp. 117-131.

Bonaccorsi, Andrea e Rossi, Cristina, "Why Open Source *software* can succeed", *Research Policy*, 32 (2003), pp. 1243–1258.

Dalle, Jean-Michel, "Heterogeneity vs. Externalities in technological competition: A tale of possible technological landscapes", *Journal of Evolutionary Economics*, 7 (1997), pp. 395-413.

Economides, Nicholas, "The Economics of Networks", *International Journal of Industrial Organization*, 14 (1996), 669-71.

Katz, L. Michael e Carl, Shapiro, "Network Externalities, Competition and Compatibility", *The American Economic Review*, 75, 3 (Junho 1985), 424-440.

---- e ----, "Technology Adoption in the Presence of Network Externalities", *Journal of Political Economy*, 94, 4 (1986), 822-841.

Klemperer, Paul, "The Competitiveness of Markets with Switching Costs" *Rand Journal of Economics*, 18, 1 (Primavera 1987), pp. 138-150.

Lee, Jongseok e Lee, Jeho e Lee, Habin, „Exploration and exploitation in the presence of network externalities", *Management Science*, 49, 4 (Abril 2003), pp.553-570.

Liebowitz, Stan J. e Margolis, Stephen E., "Market processes and the selection of standards", *Harvard Journal of Law and Technology*, 9 (1996), pp. 283-318.

Lin, Lihui, "Competition between Proprietary and Open Source Software with Heterogeneous Users", *Working Paper Boston University School of Management*, (Agosto 2004), pp. 1-35.

Schiff, Aaron, "The Economics of Open Source Software: A Survey of the Early Literature", versão revista da apresentação na *Review of Network Economics*, 1, 1 (2002), 66-74.

Proposta de um Sistema para Compreensão de Aplicações Web

Eva Oliveira[1], Pedro Rangel Henriques [2], Maria João Varanda [3]

1) Instituto Politécnico do Cávado e do Ave, Barcelos, Portugal

eoliveira@ipca.pt

2) Universidade do Minho, Braga, Portugal
prh@di.uminho.pt

3) Instituto Politécnico de Bragança, Bragança, Portugal
mjoao@ipb.pt

Resumo

Desenvolver software para a web é sempre um desafio dada a diversidade de tecnologias e linguagens que se interligam neste tipo de aplicações. Assim, a tarefa de compreender este tipo de software é bastante complexa. A área de Compreensão de Programas (CP) já disponibiliza técnicas e ferramentas para aplicações ditas "tradicionais", no entanto para as Aplicações Web, ainda não existem ferramentas que transfiram o mesmo nível de conhecimento, em termos estáticos, dinâmicos e comportamentais. Este artigo tem como objectivo propor um sistema de visualização que mostre as características das aplicações Web, nomeadamente a geração dinâmica de páginas e relações de dependência entre ficheiros e páginas, permitindo deste modo compreender detalhadamente a aplicação em toda a sua extensão e profundidade com vista a efectivamente ajudar em qualquer tarefa de manutenção e reengenharia. A abordagem proposta é baseado em modelos cognitivos e técnicas de visualização de Software.

Palavras chave: aplicações web, compreensão de programas.

1. Introdução

Com o crescimento exponencial de aplicações projectadas para utilização através de um *browser*, muitas tecnologias foram desenvolvidas para melhorar o desempenho e funcionalidades das Aplicações *Web* (AW). A importância das ferramentas de compreensão de AW advém da capacidade de traduzir a sua estrutura/tecnologias/paradigmas em representações visuais, organizadas e facilmente exploráveis [Storey 2005]. Em [Oliveira E., Henriques P.R., Varanda M.J. 2005], foi feito um estudo que passou por caracterizar aplicações Web, nomeadamente os componentes/relacionamentos que dela fazem parte, bem como enumerar as características que em geral uma ferramenta de compreensão de programas deve ter para atingir esse objectivo, distinguindo-se o uso de vistas, facilidade na navegação, interactividade, contextualização, apresentação e animação. Neste artigo iremos enumerar os domínios de conhecimento essenciais ao entendimento de aplicações Web, baseados em modelos cognitivos

e nas características deste tipo de programa e suas relações [Fracchia, Muller, Storey 1997]. Seguindo estes modelos, juntamente com a aplicação de técnicas de visualização de software reconhecidas como auxiliares num sistema de compreensão de programas [Storey 2005], parece haver um cruzamento de interesses que achamos fundamental ao entendimento deste tipo de aplicações com características próprias e mecanismos envolvidos. Seguidamente, iremos propor tipos de representações que satisfazem esses requisitos. Estas representações fazem parte de uma tese de mestrado [Eva Oliveira, 2006] no entanto apenas algumas irão ser apresentadas. Assim, apresentaremos as que de alguma forma, abarcam os conceitos que queremos aqui transmitir. Á laia de conclusão, queremos também destacar que as representações foram analisadas por programadores Web, embora a analise tenha sido meramente informativa para os autores, pois não foi sistemática nem foram analisados formalmente os resultados dessa análise.

Domínios de Conhecimento de uma Aplicação Web

Compreender uma AW, passa por compreender, tal como nas aplicações "Clássicas" a sua representação estática, dinâmica e comportamental [Holt and Ahmed 2001], mas também passa por compreender a funcionalidade dos componentes que a constituem, as tecnologias associadas, linguagem e paradigma de programação. Assim, e de acordo com modelos cognitivos de compreensão de programas, existem quatro domínios de conhecimento que é preciso explorar para retirar a informação que interessa representar. São eles:

- Análise do Domínio do Problema: A análise do domínio do problema tem em vista recuperar toda a informação possível acerca do problema tratado pela aplicação, quer através de documentação, quer através de entrevistas a utilizadores da aplicação, ou mesmo através de testes ao software. É portanto toda a informação relativa às regras de negócio da aplicação, as entidades que interagem e relações entre elas.
- Análise do Domínio do Programa: A análise do domínio do programa engloba, tanto o conhecimento do paradigma, da linguagem de programação e das tecnologias que fazem parte da aplicação e conhecimento do funcionamento dos componentes da aplicação.
- Análise da Representação Estática e Dinâmica: A análise da representação estática, envolve o estudo do fluxo de dados e de controlo, o conhecimento acerca das variáveis (tipos e dependências), das funções, objectos ou procedimentos, e dependências de ficheiros. Relativamente à análise dinâmica não se pode fazer a mesma abordagem que se segue tipicamente no âmbito das aplicações ditas "*Clássicas*", pois o fluxo de controlo da aplicação não depende exclusivamente da aplicação, mas também depende das acções de navegação do utilizador que podem alterar o fluxo da aplicação.
- Análise Comportamental: Esta representação relaciona a análise do domínio do problema como a análise da representação estática e dinâmica. Sabendo as soluções que existem para o problema que a aplicação se propõe resolver, queremos por um lado relacioná-la com o código que a implementa e visualizar todo o percurso de execução e, por outro lado, dada uma operação que a aplicação efectue, visualizar o seu percurso pelos componentes que a realizaram. Um exemplo de representações deste tipo de análise são os chamados *Use-Cases* de UML, que representam como a aplicação se comporta nas várias situações que podem ocorrer durante a sua execução, já usados noutras ferramentas cujo propósito é perceber o funcionamento da aplicação [Di Lucca 2004]. Durante uma operação específica teremos todos os passos até que essa operação esteja finalizada.

Achamos que através do conhecimento destes quatro domínios de conhecimento, conseguimos abranger todo um conjunto de factores que embora externos ao programa, deles dependem a compreensão efectiva de uma AW.

2. Requisitos de um Sistema de Visualização para Compreensão de Aplicações Web

Os requisitos de um sistema de visualização obtêm-se do nosso ponto de vista, através de cruzamentos entre as características essenciais numa ferramenta de compreensão, e entre os domínios de conhecimento que as AW exigem que se conheçam. Embora as questões relacionadas com a visualização de software, sejam válidas para as aplicações clássicas, neste artigo iremos contextualiza-las no dominío da Web.

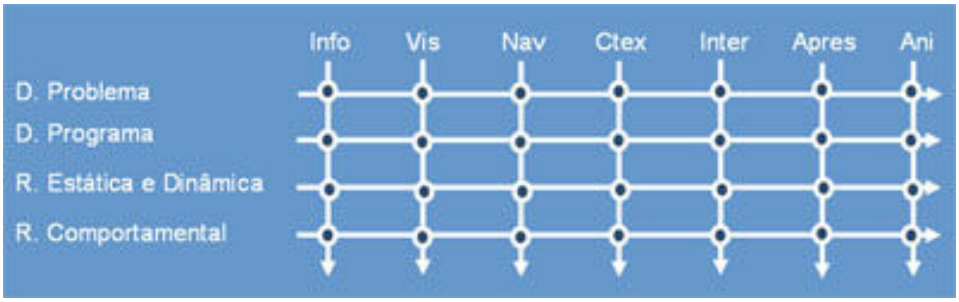


Figura 1- Requisitos de Ferramentas de Compreensão Web

A Figura 1 representa, através das suas colunas (Info - Informação, Vis - Vistas, Nav - Navegação, Ctex - Contextualização, Inter - Interactividade, Apres - Apresentação e Ani - Animação), as características que uma ferramenta de visualização deve englobar, e das linhas (Análise do Domínio do Problema, Análise do Domínio do Programa, Representação Estática e Dinâmica e Representação Comportamental) a informação que queremos analisar para compreender uma AW.

Aplicando os domínio do conhecimento à Web, iremos descrever que tipo de representações queremos ver representadas, numa ferramenta. Os parâmetros a testar o **domínio do problema**, são a representação dos conceitos/entidades, relações e regras de funcionamento (regras de negócio). A avaliação do **domínio do programa**, requer informação relativa aos paradigmas, linguagens e tecnologias associadas à aplicação. A **análise da representação estática e dinâmica** nestas aplicações deve ser efectuada em conjunto, e os critérios incidirão sobre as representações do fluxo de controlo e de dados, de páginas estáticas e dinâmicas e o seu mecanismo de geração dinâmica de páginas; Note-se que sobre uma página estática deverá ser feita uma análise estática e dinâmica, bem como sobre as páginas dinâmicas. Também se deve analisar se existem representações relativas à interacção entre as linguagens de programação usadas e entre ficheiros de código / páginas; Devem ser avaliadas as visualizações que representem relações entre componentes e de documentos gerados no Cliente com o ficheiro que o criou. Finalmente, a **avaliação da análise comportamental** será avaliar de que forma são representados comportamentos, i.e., para cada funcionalidade do sistema, determinar se podemos ver representados todos os percursos possíveis até que esta seja considerada como finalizada. Os parâmetros a avaliar nesta dimensão são relacionados com a representação dos comportamentos da aplicação relativamente às suas funcionalidades.

3. Proposta de Representações Visuais

Já se falou do que se deve mostrar numa ferramenta de CP para AW, no entanto, ainda não existem representações que ilustrem a informação que achamos ser essencial para as compreender. Nesta secção iremos propor algumas representações dos componentes de uma AW bem como do Mecanismo de Geração Dinâmica de Páginas.

Domínio do Problema

A visualização desta dimensão do conhecimento é um passo importante para a construção de um modelo mental. A representação do domínio do problema deve passar pela visualização do dicionário de dados da aplicação, quando este existe. No caso de não ser possível, propõe-se a visualização de informação extraída da base de dados da aplicação - os meta dados das tabelas. Também se propõe a apresentação de um mapa de conceitos, extraído a partir da base de dados, no entanto, o sistema deverá interagir com o utilizador [1] para que este possa introduzir manualmente os significados. O sistema deverá permitir, para cada campo de cada tabela, introduzir o respectivo significado - também pode agrupar conceitos e associar-lhes uma imagem. Para esta introdução de dados, propomos a apresentação de um ecrã como na Figura 2, em que para cada campo da tabela, o utilizador possa introduzir um texto informativo - um significado - possibilitando agrupar conceitos e associá-los com uma imagem, se assim pretender. Um mapa de conceitos mostrará o resultado desta associação.

Figura 2 - Introdução de Significados

Domínio do Programa

O domínio do programa irá passar pela representação do paradigma, tecnologias e linguagens que implementam a aplicação. A nossa proposta para representar o **paradigma** passa por mostrar um conjunto de Clientes, com ligação a um Servidor que por sua vez se encontra relacionado com uma base de dados e um sistema de ficheiros. Para saber mais detalhadamente o que se passa em cada um destes componentes deve poder-se seleccionar cada um e através de navegação descer um nível de abstracção, detalhando por exemplo no caso do servidor, os ficheiros que compõem a aplicação, e ainda sobre estes ser possível descer para outro tipo de representações - em última instância até ao código. Ainda no caso do servidor, deve existir informação acerca do tipo de requisitos que a aplicação necessita que estejam instalados para que este execute a aplicação, bem como informação acerca do relacionamento deste com os seus ficheiros e ainda a sua relação com o servidor de base de dados - no âmbito da aplicação. A nossa proposta está exibida na Figura 3 em para cada um dos elementos representados devem existir etiquetas informativas sobre as suas funções e relacionamento com o servidor.

[1] Para a introdução dos significados considera-se que o utilizador seja o programador que desenvolveu a aplicação.

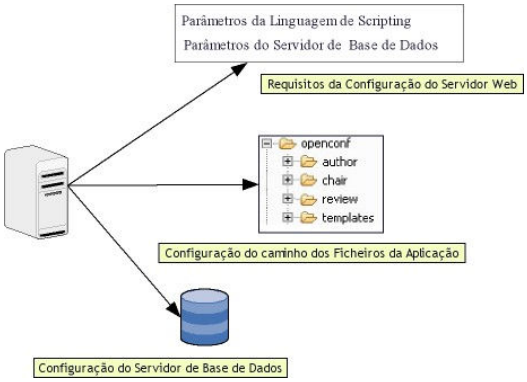


Figura 3 - Representação do Servidor

A representação da base de dados pode ser feita segundo vários tipos de diagramas, no entanto deve sempre verificar-se a possibilidade de se visualizarem apenas as tabelas que se pretende, e sempre com a possibilidade de ser descer em níveis de abstracção que permitam ir à estrutura de uma tabela, através de navegação. Relativamente aos ficheiros de uma aplicação, deverá ser possível visualizá-los numa estrutura em árvore com directorias e ficheiros e descer até ao código, se assim pretendermos. Seria útil a representação de tipos de ficheiros diferenciados com cores, distinguindo ficheiros com extensão *HTML, PHP, CSS, GIF e SQL.*

Actualmente, em que as **tecnologias** do lado do cliente estão em grande desenvolvimento, consideramos para já o uso de sessões e de *CSS*. Estas devem permitir, tal como nas outras representações, navegar para níveis mais detalhados, possibilitando seguir a visualização até aos ficheiros fonte. Parece útil detectar, no caso das *CSS*, se estas estão num ficheiro externo ou se estão inseridas na própria página Cliente, e de acordo com a situação, diferenciar a representação. Temos consciência que estas representações, apenas servem para alguém, cuja experiência em desenvolvimento *web* é nula, ou quase nula, pelo que são apenas uma ajuda na compreensão deste mecanismo, em que se apresentam os componentes envolvidos e as suas relações.

Relativamente à **linguagem**, a informação útil que propomos disponibilizar é a apresentação de um texto informativo geral, das principais características da linguagem, nomeadamente sobre o seu paradigma - lógico, funcional ou objectos.

Análise Estática e Dinâmica

A representação dos componentes de uma AW passa por mostrar as páginas estáticas, dinâmicas, e suas relações. Acreditamos que se se compreender como são obtidas as páginas e de que forma se relacionam, é possível levantar hipóteses acerca dos dados que circulam e quando circulam. Desta forma achamos que se criam pistas, que segundo os autores dos modelos cognitivos, são um auxiliar bastante importante num sistema de compreensão.

Página Estática

Uma página estática é considerada, como um ficheiro que contém apenas código interpretado pelo browser. Nas páginas estáticas, e para destacar a informação que interfere no fluxo da aplicação, estão representados os links e os campos dos formulários. Pelo que propomos cores e

formas diferentes para representar links com outras páginas, campos de formulários, links externos e links para emails.

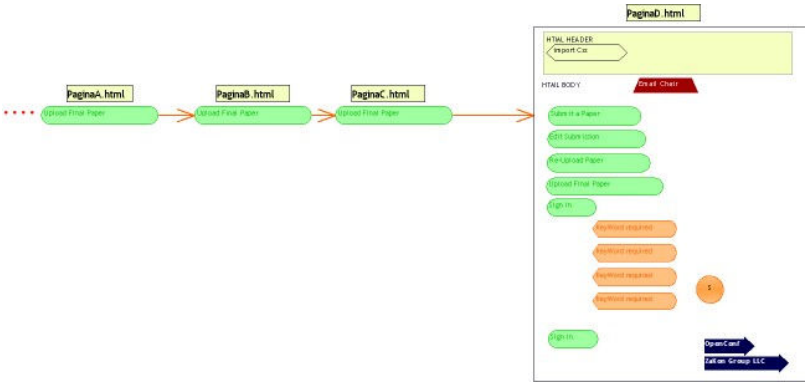


Figura 4 - Caminho entre duas Páginas Estáticas: Vista 1

Também consideramos ser útil a divisão da página em *header* e *body* para reflectir a estrutura de uma página *HTML*. Esta representação respeita a forma da página, pois sendo o *HTML* uma linguagem essencialmente de formatação de informação, importa respeitar a sua forma para um melhor entendimento. Propõe-se que seja possível navegar no sistema, seguindo os *links* apresentados. A página destino do *link* deve ser, então, mostrada; a original deve diminuir de tamanho ficando apenas visível o *link* seleccionado. Desta forma teremos uma representação natural do caminho despoletado pelo *link* em estudo. Existe porém uma condicionante espacial, pois é tecnicamente impossível representar infinitamente esta cadeia de antecessores; pelo que se propõe a representação de no máximo quatro níveis. Esta representação também é uma maneira de apresentar um dos tipos de relações que existem entre páginas Cliente, a relação "*liga a*". A representação dos antecessores permite o levantamento de hipóteses acerca do fluxo de dados e de controlo da aplicação. A sequência de passos até uma página, bem como a observação das ligações alternativas, são pistas que ajudam a compreender o fluxo e a informação que circula até à página destino [Eva Oliveira 2006].

Todas as páginas estáticas possuem relações que necessitam ser analisadas para a sua compreensão. Em resumo teremos as relações entre páginas Cliente são: *"depende de", "liga a", "é gerado por"* e *"é acedida por"* , e a nossa proposta para a sua representação está a Figura 6. É importante que a partir desta representação se possa passar directamente para a análise de cada um dos elementos nela representados, i.e., se quisermos analisar mais detalhadamente o ficheiro que a gerou, deve ser possível a selecção da página. A forma de passagem de uma representação para a outra (Navegação) não pode ser descuidada, pelo que se pode aplicar por exemplo a técnica de navegação *"pan and zoom"* aplicada noutras ferramentas, nomeadamente na ferramenta SHRimP [Lintern Michaud and Storey 2003].

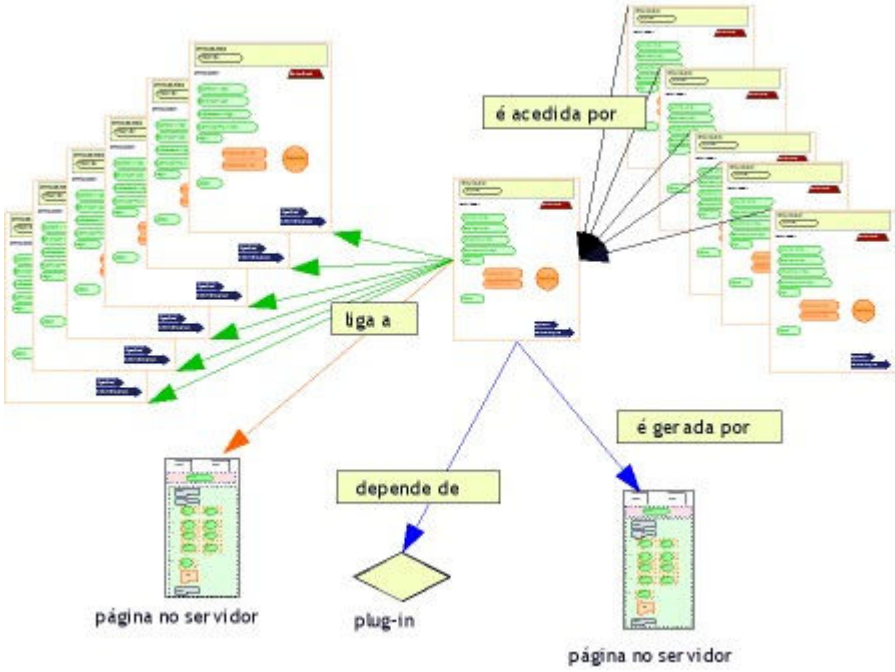


Figura 5 - Página Estática: Relações – Vista 1

Página Dinâmica

As páginas dinâmicas podem ser de dois tipos, *Tipo 2* (T2) – código interpretado pelo browser e código a ser processado no Servidor - e *Tipo 3* (T3) – código a ser processado no Servidor. A representação destas páginas obedece a critérios que permitem distinguir qualquer linha de código passível de ser enviado ao cliente, distinguindo-se também a forma como é produzido conteúdo para o cliente (se através de código processado no Servidor ou através de código interpretado no Cliente). Os blocos HTML, serão destacados dos blocos de código processado no Servidor, bem como os pedaços de código que pura e simplesmente processam código mas não enviam directamente código ao Cliente. A representação desta página está na incluída na Figura 5.

Mecanismo de Geração Dinâmica de Páginas

Para além das representações dos componentes de uma AW, existem dois conceitos que entendemos ser importantes representar. Por um lado queremos mostrar aquilo de que necessita uma página dinâmica para ser processada e onde vai buscar os conteúdos; por outro, de que forma processa a informação para gerar páginas diferentes. Comecemos por representar todas as páginas Cliente que podem ser geradas por uma página dinâmica. Uma forma de o fazer está representada na Figura 8 onde estão representadas todas as páginas geradas por uma determinada página dinâmica sendo o arco que as liga a representação de uma das relações que este tipo de página possui: a relação *"gera"*. Se se quiser ver com mais detalhe essa relação, ao seleccionar uma página é visualizada uma imagem da mesma página aumentada e as condições que precisam de ser verificadas para esta ser enviada ao Cliente.

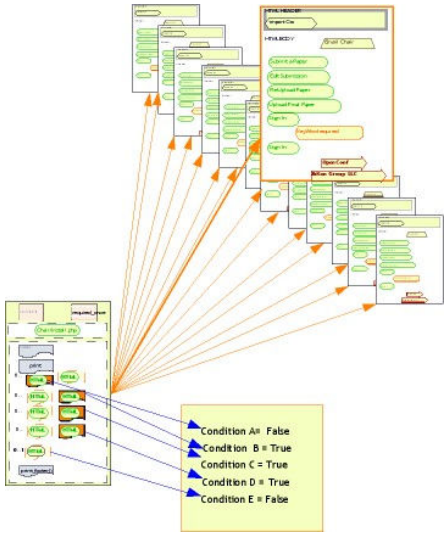


Figura 6 - Página Dinâmica: Relação

Para mostrar a forma como são gerados os conteúdos das páginas Cliente, é importante perceber o seu processamento no lado do servidor - perceber quais as condições que precisam ser satisfeitas. A Figura 6 pretende representar essa informação, mostrando os ficheiros que importa e relações com bases de dados. Ficheiros que nunca são directamente chamados pelo Cliente mas que são chamados por outros --- considerados como ficheiros do *T3* --- têm fundo avermelhado. A Figura 7 representa as dependências da página dinâmica, no Servidor, para a geração de determinada página no Cliente; é útil para detectarmos quais os ficheiros implicados na geração de uma página. Para se perceber que condições precisam ser satisfeitas para obter determinada página no Cliente, a Figura 7 é uma possível outra forma de representar a mesma informação, no entanto, foca objectivamente uma página estática e permite sabermos - através da visualização das condições que são precisas ser satisfeitas - como foi possível obter determinada página estática.

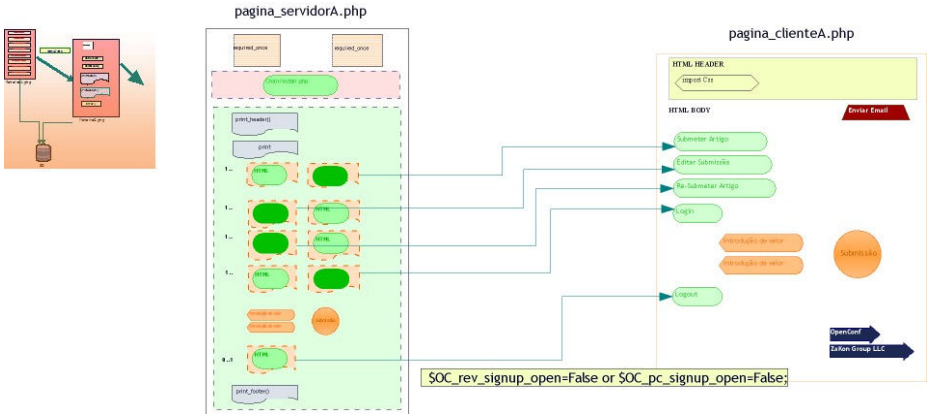


Figura 7 - Condições para Geração de Página Cliente

Análise Comportamental

As representação que iremos considerar, na análise comportamental das AW, são as adoptadas nas aplicações *WARE [Di Lucca 2004]* e *PBS[Hassan and Holt 2001]* - descritas no artigo [Oliveira E, Henriques P.R., Varanda M.J., 2005]. Pressupomos por isso, que para representar as funcionalidades da aplicação se podem usar Diagramas de Classe. Para representar o percurso da aplicação relativamente a uma dada funcionalidade podem usar-se Diagramas de Sequência em UML.

4. Conclusão

Compreender aplicações, requer por um lado o conhecimento de factores externos que podem influenciar o seu funcionamento, mas o essencial reside na análise ao código. Para isso é preciso encontrar formas de o representar logicamente, de modo a proporcionar ao utilizador abstracções que o ajudem a, mais facilmente, captar e transformar internamente em representações mentais, construindo desta forma modelos da aplicação. Neste artigo propusemos uma forma de representar logicamente algumas das características das AW, evidenciando as suas relações e lógica de funcionamento. Por um lado, representámos as páginas no Cliente, destacando os elementos que dinamizam a aplicação. Por outro, representámos as páginas no Servidor, através dos elementos responsáveis pela geração de páginas, destacando as partes do ficheiro que podem ou não ser enviadas ao Cliente. Encontrámos também uma forma de representar todos os possíveis fragmentos de código enviados ao Cliente – pela página no Servidor - e ao mesmo tempo conseguimos informar sobre as condições que precisam ser satisfeitas para que cada fragmento seja enviado. Para além destas representações, do nosso ponto de vista essenciais na compreensão de AW, completámos a análise com os tipos de características que um sistema de visualização para compreensão de AW deve proporcionar, como o uso de vistas, interactividade, navegação, contextualização, apresentação e animação.

5. Referências

Storey, Margaret-Anne , "Theories, methods and tools in program comprehension: Past, present and Future", *IWPC '05: Proceddings of the International Workshop on Program Comprehension*, , 181-191. Washington, DC, USA, 2005. IEEE Computer Society.

Richard C. Holt Ahmed E. Hassan. "Towards a better understanding of Web applications", *WSE '01 Proceddings of the 3rd Internations Workshop on Web Site Evolution*, 2001. IEEE Computer Society

Eva Oliveira, "Características de um sistema de visualização para compreensão de aplicações Web", *dissertação de mestrado, Universidade do Minho*, 2006.

Giuseppe Antonio Di Lucca, Anna Rita Fasolino, and Porfirio Tramontana, "Reverse Engineering web applications: the ware approach", *Journal of Software Maintenance and Evolution*, 16(1-2):71-101, 2004.

Rob Lintern, Jeff Michaud, Margaret-Anne Storey, and Xiaomin Wu. "Plugging-in visualization: experiences integrating a visualization tool with eclipse". *In SoftVis '03: Proceedings of the 2003 ACM symposium on Software visualization*, pages 47–ff, San Diego, California, 2003. ACM Press.

A. Hassan and R. Holt. “A visual architectural approach to maintaining web applications”, 2003.

M.-A. D. Storey, F. D. Fracchia, and H. A. Mueller. “Cognitive design elements to support the construction of a mental model during software visu-alization”. *In WPC ’97: Proceedings of the 5th International Workshop on Program Comprehension (WPC ’97),* page 17, Washington, DC, USA, 1997.IEEE Computer Society.

Oliveira E., Henriques P.R., Varanda M.J. “Compreensão de Aplicações Web: O Processo e as Ferramentas” *Procedings of the 6th CAPSI, Bragança.*, 2005.

TEMA 2 – CASOS DE APLICAÇÃO DE SISTEMAS DE INFORMAÇÃO

2.1. Casos de Estudo em Sectores de Negócio ou Mercados Específicos

Modelo de *e-Procurement* centrado no fornecedor para a Administração Pública

Ercília Rosa Correia Ribeiro Lopes [1], Leonel Duarte dos Santos [2]

1) Departamento de Sistemas de Informação, Escola de Engenharia, Universidade do Minho, Azurém, 4800-058 Guimarães, Portugal

ercilialopes@yahoo.com

2) Departamento de Sistemas de Informação, Escola de Engenharia, Universidade do Minho, Azurém, 4800-058 Guimarães, Portugal

leonel@dsi.uminho.pt

Resumo

O objectivo deste artigo é propor um modelo de *e-Procurement* centrado no fornecedor para a Administração Pública tendo em atenção as tecnologias actuais e o contexto português. Realizou-se um estudo sobre os modelos de *e-Procurement* existentes, identificando as vantagens e desvantagens de cada um deles.
A literatura refere com frequência que o modelo a usar para a Administração Pública seria o modelo centrado no comprador, no entanto propomos uma abordagem diferente, a utilização de um modelo centrado no fornecedor. Neste artigo faz-se também uma descrição do funcionamento do modelo apresentado e uma análise critica das suas vantagens e desvantagens.

Palavras chave: *e-Procurement*, modelos *e-Procurement*, compras electrónicas, *e-Procurement* centrado no fornecedor, Administração Pública.

1. Introdução

Na abordagem ao *e-Procurement* podem ser encontrados alguns modelos, ao fazer a abstracção dos projectos e iniciativas que têm vindo a ser realizados. É possível iniciar esta caracterização dividindo os modelos em privados e públicos. Os primeiros resultam da iniciativa de uma entidade, fornecedora ou cliente, que procura, respectivamente, oferecer a capacidade de comprar electronicamente aos seus clientes ou dotar-se dessa mesma capacidade. Nos modelos públicos podem-se identificar iniciativas de entidades que procuram oferecer um espaço de transacção electrónico tanto a fornecedores como a clientes.

Existem três modelos básicos de arquitecturas *e-Procurement* que evoluíram ao longo dos últimos anos. Esses modelos são conhecidos pelo tipo de relação que estabelecem entre os intervenientes (centrado no fornecedor, centrado no comprador e mercador electrónico) [AGIMO, 2003].

Estes modelos têm vantagens e desvantagens na perspectiva dos fornecedores e compradores, mas geralmente envolvem sistemas fechados, limitando a flexibilidade dos que se encontram na periferia. Existe portanto, a necessidade de se criar/propor novos modelos, baseados em padrões abertos, com o intuito de superar as limitações dos modelos conhecidos.

O sistema proposto, é um sistema aberto e distribuído, que não requer muita tecnologia nem grandes investimentos de ambas as partes do processo. Este sistema tem vantagens e desvantagens, sendo estas últimas facilmente contornáveis. É também importante referir que a comunicação do sistema *e-Procurement* é baseada em protocolos e linguagens universais (*XML – Extensible Markup Language)*, que permitem a interoperabilidade entre sistemas informáticos, não sendo necessário ao fornecedor dispor de recursos financeiros para os alterar [Amaral *et al*, 2003].

2. *E-Procurement*

O *e-Procurement* serve para adquirir bens e serviços, usando meios electrónicos, tendo como objectivo um maior controlo nos processos de aquisição e redução de custos uma vez que os produtos são licitados por diferentes fornecedores [Lopes e Santos 2006].

2.1 Processo de *Procurement* operacional

As compras MRO (Manutenção, reparação e operação) actuais, funcionam da seguinte maneira ao nível operacional [Bruins 2000]:

1. O funcionário precisa de um produto e procura-o num ou mais catálogos de papel;
2. O funcionário encontra o produto e preenche uma requisição, a qual é enviada ao comprador e ao gestor do funcionário para aprovação;
3. O comprador verifica a requisição e, se esta estiver correcta, espera pela aprovação;
4. De seguida, o comprador envia a requisição ao fornecedor via fax, carta ou telefone;
5. O funcionário do fornecedor insere a requisição no seu sistema ERP (*Enterprise Resource Planning*);
6. O sistema ERP envia uma mensagem ao armazém com as necessidades da requisição, para que sejam enviados os bens ao comprador. O sistema ERP também envia uma mensagem à administração para que esta emita uma factura ao comprador;
7. O armazém do comprador recebe os produtos e faz uma entrega interna ao empregado que os encomendou;
8. O armazém também regista o que recebeu. Baseado neste registo a administração pode comparar o que foi recebido com a factura, e se tudo estiver correcto, esta é paga.

Este processo possibilita erros e demora muito tempo. O *e-Procurement* automatiza e simplifica este processo.

3. Configurações do *e-Procurement*

Existem vários tipos de *e-Procurement.* Na tabela 1 são descritas as configurações mais comuns. A principal diferença está em quem mantém o catálogo do produto (o fornecedor, o comprador ou terceiros).

TIPO	CARACTERÍSTICAS	EXEMPLOS
Centrado no fornecedor	Um fornecedor dá acesso a uma gama pré-negociada de produtos às organizações compradoras. O cliente e o fornecedor têm um relacionamento contratual.	www.cisco.com www.aspa.nl www.dell.com
Centrado no comprador	As companhias individuais têm contratos com diversos fornecedores. O catálogo e o sistema requisição são mantidos dentro da organização de compra.	Autoeuropa.supply.net
Mercado (Marketplace)	Uma terceira entidade independente tem acordos com diversas organizações compradoras e fornecedoras. Compradores e fornecedores negociam uns com os outros através de um mercado. Ambos estão limitados por acordos com o mercado.	www.productview.net www.econstroi.pt

Tabela 1 – Modelos para compras na Internet [Bruins 2000]

3.1 Modelo de *e-Procurement* centrado no comprador

O modelo de compras centrado ou gerido pelo comprador, possui e mantém um catálogo agregado ao sistema organizacional. Um sistema de compras empresarial agrega muitos catálogos de diferentes fornecedores num único catálogo universal, permitindo que os utilizadores finais requisitem do computador local, facilitando o processo de aquisição [Haller 2004]. Os requisitantes ou compradores acedem a vários catálogos de fornecedores num formato específico do comprador, seleccionam itens, e iniciam a requisição [Kim e Shunk 2004].

Este modelo é apropriado nos casos onde existem um grande número de potenciais vendedores, pois evita catálogos fragmentados e de múltiplos formatos [Singh e Thomson 2002].

As organizações compradoras têm de montar e manter catálogos para cada fornecedor. Esta solução é dispendiosa. A organização compradora pode ter um sistema de compras interna, onde o comprador pode procurar pelos produtos através do catálogo interno, tratar de autorizações, realizar encomendas e enviar isso manualmente para o fornecedor. O sistema está

totalmente integrado no controlo financeiro empresarial e nos sistemas de relatórios. Estes sistemas podem fornecer algumas ligações rudimentares aos sistemas do fornecedor, para transmissão de encomendas, mas tipicamente, não oferecem serviços mais sofisticados que disponibilidade de stock em tempo real [Bruins 2000].

O comprador pode aumentar o seu poder de compra através do uso de ferramentas de leilão invertido. Estes leilões são comuns quando os compradores desejam pagar o preço mais baixo pelo produto. Os vendedores podem acabar por ter custos ao actualizar os seus sistemas e a integrá-los com o sistema do comprador [Singh e Thomson 2002].

Além da limitação do estabelecimento e manutenção de ligações de comunicação independentes para cada fornecedor, o maior desafio está associado à manutenção do conteúdo dos catálogos. Grandes compradores são normalmente os que tomam a iniciativa de criar um sistema de catálogos multi-fornecedor baseado nas suas especificações necessárias. Os pequenos fornecedores acabam por ser forçados a fornecer, e frequentemente actualizar os dados do catálogo em diversos formatos, de forma a satisfazer as especificações dos diferentes compradores. Apesar desta solução ser satisfatória para um pequeno número de compradores ou fornecedores, não é escalável. Numa grande empresa pode haver centenas ou milhares de fornecedores, e cada um pode ter milhares de itens de catálogos [Kim e Shunk 2004]. Haverá bens equivalentes de vários fornecedores que precisam de ser consistentemente descritos para simplificar as decisões de fornecimento. Tal volume e complexidade tornam impossíveis de gerir [Hughes 2000].

3.2 Modelo de *e-Procurement* centrado no fornecedor

Neste modelo, o fornecedor disponibiliza um sistema que está acessível através da Internet, onde compradores podem pesquisar no catálogo do fornecedor e fazer encomendas [Bruins 2000]. A gestão do conteúdo do catálogo cabe aos fornecedores. Esta é uma oportunidade eficaz para as pequenas e médias empresas comprarem *online*. Os compradores podem ser confrontados com custos de integração das suas tecnologias com as dos fornecedores [Singh e Thomson 2002]. Estes catálogos podem, por vezes, estar num formato especificado pelo comprador. A maior vantagem para os compradores está no fácil acesso ao catálogo dos fornecedores sem grandes investimentos numa solução específica para esse fim [Kim e Shunk 2004].

Cada fornecedor constrói uma loja *online* e o comprador fica com a tarefa de consultar e encomendar nas diferentes lojas [Bruins 2000]. No entanto os compradores têm que criar múltiplas contas e senhas, navegar através de diferentes interfaces e perder tempo significativo

em funções que não lhes compete. Outra contra partida é não haver um padrão para os catálogos o que torna difícil a comparação e a integração com o sistema interno de dados [Compaq Inc 2001].

3.3 *e-Marketplace* ou modelos de *e-Procurement* geridos por terceiros

Neste tipo de sistema, o fornecedor e o comprador interagem numa plataforma independente disponibilizada por terceiros. O e-*marketplace* aloja os catálogos dos fornecedores e fornece possibilidade de transacções electrónicas entre os compradores e os fornecedores [Bruins 2000].

e-Marketplaces são locais neutros e independentes entre compradores e fornecedores, onde os bens estão bem padronizados ou consistentemente especificados [Kim e Shunk 2004]. O formato e o conteúdo do catálogo pode ser normalizado para facilitar os compradores nas comparações de produtos. Os catálogos são actualizados com frequência permitindo verificações de preços e disponibilidades, o que seria impraticável num ambiente de um para muitos, onde os catálogos dos compradores eram actualizados por cada fornecedor [Hughes 2000].

B2B *Marketplace*/portal de *Procurement* estabelece uma relação de muitos para muitos entre o comprador e os fornecedores. Combina as vantagens dos diferentes modelos centrados no comprador e no vendedor, evitando custos de montagem e manutenção para os participantes [Bruins 2000]. Um catálogo alojado num sistema gerido por terceiros permite uma grande economia na sua gestão uma vez que existe apenas um único catálogo, em vez de um em cada organização compradora ou fornecedora [Hughes 2000].

4. Comparação de modelos

O modelo mais prático e cómodo para a Administração Pública seria o modelo centrado no comprador, uma vez que esta é o centro de todo o sistema [Lopes e Santos 2006]. Seria muito mais fácil ter um catálogo com dados estruturados num formato especificado pela Administração Pública. Esta solução seria vantajosa para os diversos fornecedores uma vez que estes não teriam de realizar qualquer investimento para utilizar o sistema. No entanto diversas desvantagens como o elevado custo inicial para Administração Pública, a dificuldade de manutenção do sistema e a dificuldade de actualização dos dados (por parte dos fornecedores), tornam esta solução pouco adequada.

Já o modelo centrado no fornecedor permite uma maior flexibilidade na gestão dos catálogos, no entanto o comprador ainda teria de se confrontar com a diversidade de sistemas existentes e a incompatibilidade entre eles.

Uma solução razoável passaria por descentralizar os catálogos, mantendo a consistência dos dados. Assim os diversos fornecedores teriam os catálogos integrados nos seus sistemas informáticos e actualizados em tempo real. O comprador apenas se limitaria a pesquisar e seleccionar os diversos bens. Este sistema centrado no fornecedor reduz os custos para Administração Pública e permite que os fornecedores sejam consultados por outros potenciais compradores.

5. Modelo de *e-Procurement* para a Administração Pública

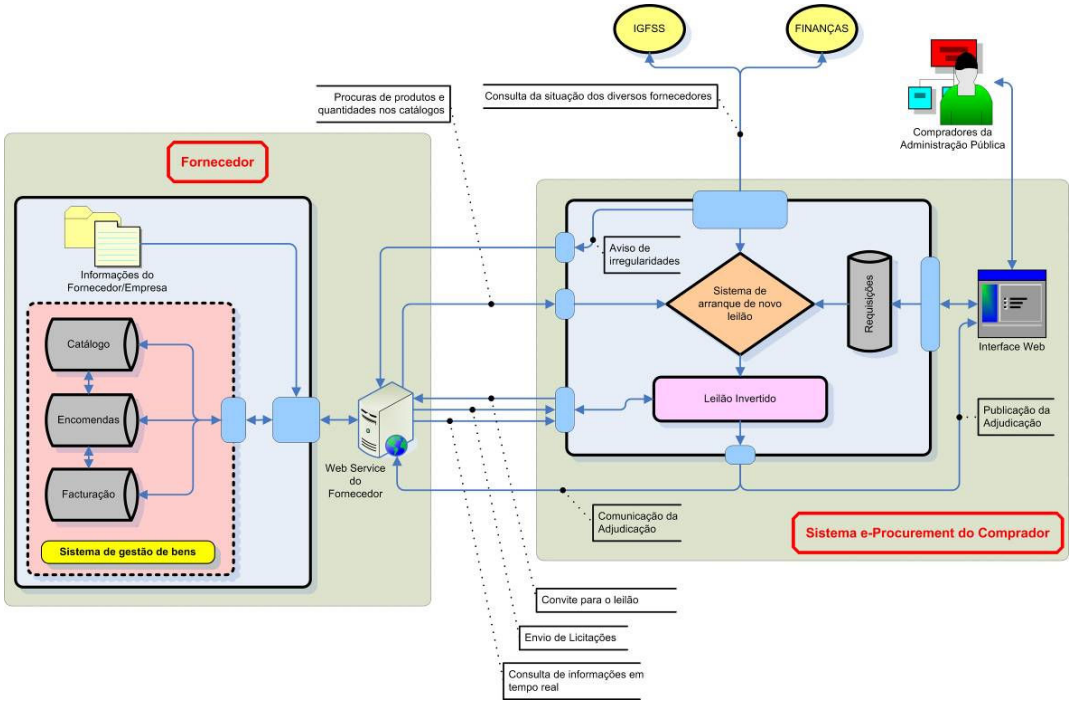


Figura 1 – Modelo de *e-Procurement* proposto para a Administração Pública

O modelo proposto na figura 1 tem como base o modelo *e-Procurement* centrado no fornecedor, herdado vantagens e funcionalidades de outros modelos e, ao qual foi acrescida a funcionalidade de leilão invertido.

Os fornecedores começarão por se registar aceitando as condições de fornecimento e pagamento impostas por lei [Portal informativo de compras]. No acto de registo os fornecedores deverão enviar uma declaração onde afirmam a sua situação financeira

regularizada. Esta situação será verificada automaticamente, utilizando mecanismos de acesso à informação dos diversos sistemas (IGFSS e Finanças).

Neste modelo, o catálogo de produtos encontrar-se-á distribuído pelos diversos fornecedores, sendo estes, responsáveis por mantê-lo e actualizá-lo. Os catálogos dos fornecedores estão armazenados localmente nos seus sistemas informáticos e serão acedidos pelo sistema *e-Procurement* através de um Serviço *Web* (*Web Service*). Cada *Web Service* deve ser capaz de se integrar nos diferentes sistemas existentes, ou deverá ser permitido ao fornecedor, transitar para um sistema de padrões abertos.

Os compradores da Administração Pública interagem com o sistema de *e-Procurement* através de uma aplicação gráfica disponibilizada na *Web*. Cada funcionário, com acesso à aplicação, pode pesquisar e requisitar bens. No entanto, para cada serviço da Administração Pública deverá existir um gestor de compras que decide se serão aceites as requisições efectuadas pelos diversos funcionários.

Ao realizarem-se pesquisas no sistema *e-procurement*, serão contactados os diversos *Web Services* dos fornecedores registados.

As requisições aprovadas ficarão registadas numa base de dados única que contém todas as necessidades da Administração Pública. Esta base de dados é consultada de tempos a tempos pelo "sistema de arranque de leilões". Caso esse sistema constate a existência de requisições aprovadas, vai verificar quais os fornecedores capazes de satisfazerem a encomenda, verificando se o produto existe nos seus catálogos. O sistema verifica também se cada fornecedor tem a sua situação financeira regularizada. Se os fornecedores estiverem em condições de participar no leilão, serão convidados através de notificação electrónica. Se, por outro lado, o fornecedor não estiver regularizado, o sistema enviará uma notificação para regularizar a sua situação no prazo de, por exemplo, 2 dias. Findo esse prazo, o sistema iniciará um leilão para os bens em causa e permitirá aos fornecedores regularizados interagir, licitando preços sucessivamente mais baixos. Os fornecedores já sabem que têm de ter a situação regularizada, e se não o fizerem no prazo indicado, perderão a oportunidade de entrar no leilão.

Durante o processo de leilão invertido, os diversos fornecedores disputam entre si o menor preço pelo qual serão capazes de fornecer os bens. No final do leilão, o sistema *e-Procurement* comunica os resultados aos compradores da Administração Pública e notifica a adjudicação ao fornecedor vencedor. Nessa altura, ser-lhe-á indicado onde deverá entregar a respectiva requisição. Uma requisição pode pertencer a um organismo ou ser um conjunto de várias requisições ao mesmo produto de vários organismos. Neste último caso, o fornecedor

adjudicado, deverá entregar cada fracção da requisição nos diversos organismos envolvidos, dentro do prazo e condições acordadas.

Todos os intervenientes (fornecedores e compradores) têm a possibilidade de acompanhar o processo desde a sua publicação até à sua conclusão, passando pelos diferentes estados de licitação.

Vantagens:

- ✓ Baixo custo e manutenção pois deixa de ser necessário armazenamento de dados dos catálogos no lado do comprador;
- ✓ Facilidade de actualização do *stock* e produtos por parte dos fornecedores;
- ✓ Maior simplicidade, menor investimento e simplificação de procedimentos para a Administração Pública;
- ✓ Possibilidade de outros potenciais compradores se ligarem aos fornecedores (possível evolução para *e-marketplace*);
- ✓ Protocolo de comunicação baseado em padrões abertos, usando *XML* na transferência de dados, permitindo interoperabilidade entre diferentes sistemas;
- ✓ Modelo distribuído e recepção assíncrona de informação, o que torna o sistema *e-Procurement* menos pesado;
- ✓ Comprador não precisa de aprender a usar diferentes sistemas como acontecia no modelo centrado no fornecedor tradicional, uma vez que todos os fornecedores usarão o mesmo protocolo de comunicação e a mesma representação de dados;
- ✓ O comprador deixa de precisar de saber onde procurar a informação, os fornecedores registam-se e passam a fazer parte do sistema.

Desvantagens:

- ✓ Sistema lento, pois para cada pesquisa o sistema terá de contactar todos os fornecedores para ver se estes têm os produtos procurados nos seus catálogos;
- ✓ Investimentos iniciais por parte dos fornecedores, caso estes não disponham de sistema informático. Caso os Sistemas Informáticos existam, há a necessidade de conversão/transporte das plataformas existentes para o novo sistema. A conversão pode limitar o acesso de fornecedores mais pequenos devido aos custos. Há a possibilidade de desenvolver um pequeno sistema de gestão de catálogo electrónico, distribuído a baixo custo ou de forma gratuita.
- ✓ Requer formação/adaptação ao novo sistema.

6. Conclusão

Este tipo de sistema, pode ser um pouco mais lento quando comparado com outros sistemas com o catálogo local. Mas o facto de ser distribuído aumentará a eficiência uma vez que a carga de processamento será partilhada pelos diversos fornecedores, não sendo necessário sistemas informáticos com grandes capacidades de ambas as partes.

A adaptação do sistema existente ou transição para um totalmente novo, implicará um custo inicial aos fornecedores, mas a médio ou longo prazo será benéfico pois resultará na abertura do sistema (anteriormente fechado). Isto porque o *Web Service* implementado nos seus sistemas pode ser utilizado por outros potenciais compradores, além da Administração Pública. Como é um sistema aberto, e como qualquer outro comprador poderá ligar-se ao sistema dos fornecedores através do *Web Service*, este sistema centrado no fornecedor pode evoluir para um Mercado Electrónico (*e-Marketplace*). Esta situação irá tornar-se benéfica para os fornecedores, pois deixam de ter de usar diferentes sistemas para cada comprador, e passarão a usar um único sistema global.

Como os pequenos fornecedores poderão não ter recursos financeiros para entrar num sistema aberto como este que foi proposto e, como se deve fornecer igualdade de direitos a todos que queiram participar nas compras electrónicas [UMIC, 2003], deverá ser implementada uma aplicação *open source* de gestão de Catálogos.

O uso de protocolos padronizados e transferências de dados em *XML* permite o desenvolvimento de diferentes aplicações para diferentes sistemas informáticos, podendo o fornecedor escolher qual o sistema que mais lhe convém, uma vez que deixa de existir incompatibilidades na comunicação entre sistemas informáticos.

Este modelo proposto torna-se mais eficaz do que os modelos tradicionais uma vez que foi efectuado um levantamento dos benefícios desses vários modelos e tentou-se explorar ao máximo as vantagens de cada um.

7. Referências

Amaral, L., Teixeira, C., Oliveira J.,"*E-Procurement*: uma reflexão sobre a situação actual em Portugal", *Associação para a Promoção e Desenvolvimento da Sociedade de Informação,* 2003.

Australian Government Information Management Office (AGIMO), Checklist for devolinp na *E-Procurement* strategy, 2000 http://www.agimo.gov.au/publications/2000/04/checklist

Bruins, A.,"Electronic *Procurement*, *Telematica Instituut,* 2000.

Compaq Inc. "E-procurement: sharpening the competitive edge throught implementation of an e-procurement solution", *White paper*, 31 August 2001

Haller, Regina, "Emerging procurement models and the effects on internal structures", *University of Viena*, 2004.

Hughues D., "Assessing the value of electronic marktpelaces for business to business trading", *Proceedings of the second Internacional Conference of Electronic Trade in the CIS and Eastern European Countries*, 2000

Kim J.-I.., Shunk D. L., "Matching indirect procurement process with different B2B e-procurement system", *Computers in Industry*, 2004

Lopes, E., Santos, L.,"Estratégias de *e-Procurement* na Administração Pública: Uma revisão de literatura", *Conferencia Ibérica de Sistemas e Tecnologias de Informação,* 2006.

Portal informativo de compras, www.compras.gov.pt

Singh M., Thomson D., "eProcurement Model for B2B Exchanges: An Australian Example", *15th Bled Electronic Commerce Conference eReality: Constructing the eEconomy*, 2002

UMIC, Programa Nacional de Compras Electrónicas, 2003 http://www.umic.pcm.gov.pt/UMIC/CentrodeRecursos/Publicacoes/apresentacao_compras.htm

Levantamento e Reengenharia de Processos nos Municípios do Distrito de Évora

Henrique O'Neill [1], Mauro Nunes [2], Judite Esteves Pinto [3], Virgílio Cruz Machado [4], Teresa Batista [5], Gisela Vitorino [6]

1) ISCTE / PARTNERTEC, Lisboa, Portugal
henrique.oneill@iscte.pt

2) ISCTE / ADETTI, Lisboa, Portugal
mauro.nunes@iscte.pt

3) PARTNERTEC, Setúbal, Portugal
jep@mail.pt

4) FCT/UNL, Lisboa, Portugal
vcm@fct.unl.pt

5) AMDE/EDD / Universidade de Évora, Évora, Portugal
tbatista@amde.pt

6) AMDE/EDD, Évora, Portugal
gisela.vitorino@evoradistritodigital.pt

Resumo

Neste artigo apresenta-se o Estudo de Levantamento e Reengenharia de Processos desenvolvido pelos autores no âmbito do projecto Évora Distrito Digital. Este projecto tem como objectivo fomentar a coesão e o desenvolvimento da região através da actuação no domínio das Tecnologias e Sistemas de Informação, promovendo um melhor interrelacionamento dos cidadãos com as suas autarquias através de serviços electrónicos.

Este levantamento envolveu cerca de 100 participantes - eleitos, dirigentes e funcionários - das diversas autarquias do Distrito e permitiu o levantamento de cerca de 340 procedimentos de gestão autárquica em diversos domínios processuais que, em particular, envolvem a prestação de um serviço ao cidadão.

Foi seleccionado um grupo de procedimentos mais representativos no conjunto dos municípios, que foram analisados detalhadamente no sentido de identificar o conjunto das actividades estruturantes, constrangimentos e oportunidades para melhoria do funcionamento.

Este conjunto, designado por Procedimentos de Referência, foi redesenhado tomando como base princípios de reengenharia, designadamente: a utilização de aplicações informáticas adequadas; a racionalização dos circuitos, em particular do circuito do expediente; a redução do número de transição entre participantes, procurando promover a simplificação e desburocratização dos circuitos de decisão. Estas iniciativas têm como objectivo a melhoria da qualidade do serviço prestado, a redução dos tempos de resposta, a eficácia e a satisfação de munícipes, funcionários, dirigentes e eleitos.

O estudo apresenta propostas de reorganização em três domínios de intervenção: Mudança Estrutural e Funcional, Tecnologia e Qualificação do Capital Humano. Estas propostas materializam-se num conjunto de 25 recomendações operacionais.

A abordagem participativa que foi adoptada, baseada na realização de sessões de formação para apoio ao trabalho realizado, envolvendo eleitos, dirigentes e funcionários, constitui um contributo importante para o reforço das competências de gestão dos participantes e um primeiro passo para a sua preparação como agentes de mudança.

É convicção dos autores que existem condições para a consolidação desta iniciativa na generalidade dos municípios do distrito de Évora, e que será possível aprofundar as propostas aqui enunciadas e materializá-las em novos projectos.

Palavras chave: Reengenharia de processos, Gestão autárquica, Modelos de organização, Administração Pública electrónica, Cidades e Regiões Digitais

1. Enquadramento do Estudo

O Estudo de Levantamento e Reengenharia de Processos, desenvolvido no âmbito do Projecto Évora Digital, pretende contribuir para fomentar a coesão e o desenvolvimento da região através da actuação no domínio das Tecnologias e Sistemas de Informação, promovendo um melhor interrelacionamento dos cidadãos com as suas autarquias através de serviços electrónicos.

O caderno de encargos identificava desde logo, os domínios processuais que iriam sustentar os serviços camarários a disponibilizar on-line aos munícipes a partir do Portal e dos sites das Autarquias do Distrito, designadamente: Urbanismo, Mercados e Feiras, Publicidade e Ocupação da Via Pública, Águas e Esgotos, Cemitérios, Higiene e Salubridade, Trânsito Actividades Económicas, Habitação, Transportes. Complementarmente, pela sua importância no domínio da gestão de informação, foram também objecto do estudo os procedimentos transversais de expediente e de arquivo.

Pretendia-se com o Estudo promover a harmonização entre as diversas autarquias dos procedimentos de back-office que suportam o atendimento ao munícipe a disponibilizar por via electrónica. Pretendia-se igualmente propor a adaptação dos procedimentos autárquicos às exigências resultantes do novo modelo de atendimento ao público através da adopção de propostas de reengenharia e de racionalização de procedimentos, reforçada pela utilização de plataformas electrónicas de gestão de processos (workflow).

Os trabalhos desenvolvidos tiveram como referencial a orientação dos serviços autárquicos para o cidadão, baseados em processos organizacionais que respondam às suas necessidades concretas. Neste contexto, o modelo de atendimento ao público suportado pelo guichet único virtual é um dos vectores fundamentais do projecto, disponibilizando serviços idênticos aos que estarão disponíveis num balcão de atendimento presencial.

2. Metodologia

A condução do Estudo de Levantamento e Reengenharia dos Processos autárquicos foi dinamizada por uma equipa multidisciplinar de consultores da Partnertec, técnicos da AMDE e funcionários autárquicos, e implicou o desenvolvimento das seguintes actividades:

- Preparação do projecto
- Levantamento e análise da situação actual de cada município
- Validação / avaliação dos processos / circuitos
- Harmonização e normalização dos procedimentos entre as várias autarquias

- Apresentação de propostas de uma nova organização de processos
- Apoio à implementação dos processos electrónicos
- Sensibilização para os novos processos de trabalho

Cada consultor assumiu a responsabilidade pelo acompanhamento de um conjunto de autarquias ao longo do projecto. Em cada autarquia foi formada uma equipa de trabalho constituída por representantes do poder político, pelo interlocutor designado para o efeito, pelos colaboradores da autarquia mais directamente ligados aos procedimentos a levantar. As reuniões efectuadas foram acompanhadas por um representante da Associação de Municípios do Distrito de Évora, agindo como facilitador.

Atendendo à natureza e ao objectivo do projecto, fortemente marcado pela necessidade de envolver todos os participantes, foi adoptada uma metodologia participativa baseada na realização de sessões de trabalho apoiadas pelo consultor (workshops), que foram complementadas por tarefas desenvolvidas pelos participantes de forma autónoma. Para o levantamento e caracterização dos procedimentos no âmbito dos domínios processuais já identificados foram previstas e realizadas 52 sessões de trabalho conjunto com a duração de 3.5h cada. Para facilitar o levantamento dos procedimentos foram construídos instrumentos de trabalho estruturados com base nos requisitos da Norma NP EN ISO 9001:2000, e suportados em folhas de cálculo (Excel): Ficha de Procedimento, Tabela Síntese e Diagrama.

Cada autarquia elencou os procedimentos a levantar que considerou mais importantes para disponibilizar através de serviços on-line, registando-se, nalguns casos, o alargamento a novos domínios processuais. Foram levantados cerca de 340 procedimentos e identificaram-se 132 categorias distintas de procedimentos. O número médio de procedimentos levantados por autarquia foi de 26.

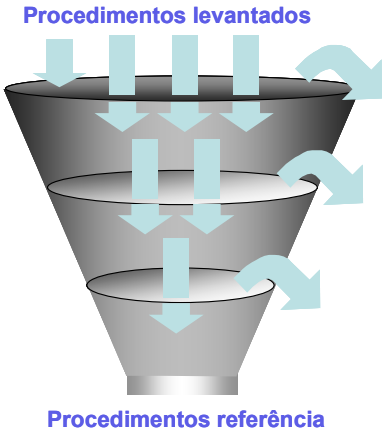
Utilizaram-se como fontes de informação os organogramas que espelham o modelo de departamentalização das actividades e as estruturas que definem o enquadramento funcional dos recursos humanos das diversas autarquias.

Após a selecção pelas autarquias dos serviços mais procurados pelos cidadãos, procedeu-se ao mapeamento de alto nível dos fluxos de informação (processos de trabalho) que serviram para a criação da base de conhecimento, elaboração de propostas de melhoria e para a elaboração de um Manual Base de Procedimentos de Gestão Autárquica.

Concluída a fase do levantamento e caracterização dos procedimentos identificou-se um conjunto de procedimentos considerados representativos e sobre os quais iria incidir as propostas de reengenharia. Como critério foram seleccionados os procedimentos que foram identificados em pelo menos 5 autarquias, tendo-se obtido o número de 30 procedimentos.

Cada consultor procedeu a uma análise transversal, por procedimento, no subconjunto de autarquias da sua responsabilidade procurando identificar exemplos de boas práticas. Num segundo momento, a análise comparativa das melhores práticas, por procedimento, alargou-se ao conjunto dos municípios. Finalmente foram desenhados os procedimentos de referência tomando como base as melhores práticas identificadas, como se ilustra na figura seguinte.

Este projecto foi desenvolvido nos 14 municípios do distrito durante os meses de Abril, Maio e Junho, envolveu mais de 100 participantes, e a realização de 52 workshops, totalizando mais de 182 horas de formação.



1. Apoio dos consultores ao levantamento de processos nas respectivas autarquias
2. Selecção e análise dos procedimentos de comparação.
3. Análise crítica dos procedimentos de comparação, modelando os procedimentos de referência.

Os procedimentos de referência foram desenhados em Visio, utilizando os diagramas de actividade UML. Complementarmente, para cada procedimento de referência foi elaborada uma Ficha de Proposta de Procedimento, com uma estrutura semelhante à utilizada na fase de levantamento.

O conjunto de procedimentos de referência encontra-se descrito detalhadamente no documento Manual de Procedimentos de Gestão Autárquica – Procedimentos de Referência.

Foi ainda desenvolvida uma aplicação informática com interface web (site) que agrega toda a documentação produzida no âmbito do projecto, e que inclui cerca de 450 ficheiros.

3. Apresentação de Resultados

Os 340 procedimentos levantados e caracterizados pelas autarquias dispersaram-se para além dos dez domínios processuais inicialmente previstos, evidenciando algum interesse do alargamento do objecto do Estudo aos processos organizacionais de suporte (recursos humanos, aprovisionamentos, actividades de acção sócio-cultural, etc.).

3.1 Distribuição dos Procedimentos Levantados

O domínio processual do Urbanismo representa 39% das preferências, seguindo-se o sector das Águas e Esgotos com 15%, Mercados e Feiras e Cemitérios próximos dos 9%. Mais modestamente as Actividades Económicas situam-se nos 8%, a Publicidade e Ocupação da Via Pública 5%, o Trânsito 4%, representando cada um dos restantes grupos menos de 2% dos processos levantados.

3.2 Procedimentos de Referência

Conforme anteriormente se referiu os Procedimentos de Referência foram identificados em pelo menos 5 autarquias, o que corresponde a mais de 60% dos municípios. No seu conjunto os Procedimentos de Referência cobrem a maioria dos procedimentos levantados no total dos municípios (64%). Estes procedimentos foram revistos de acordo com as melhores práticas identificadas, de modo a serem disponibilizados através de serviços on-line. Antecipa-se que as alterações propostas pelos procedimentos de referência poderão ser facilmente propagáveis para os restantes procedimentos.

4. Apreciação Global

Assentes numa estrutura burocrática e de pendor centralista as organizações públicas não conseguiram ainda fazer face, de uma forma rápida e eficaz, às necessidades crescentes dos cidadãos e das empresas. É contudo justo realçar o esforço de modernização administrativa que as organizações públicas, designadamente as autarquias, têm vindo a realizar e o papel importante que as associações de municípios têm vindo a desenvolver como catalizadores nesse processo.

O Estudo de Levantamento e de Reengenharia de Processos assumiu intencionalmente um carácter prático e funcional na análise do conjunto das actividades conducentes à realização dos processos, procurando alcançar a forma como o trabalho é realizado nas organizações atárquicas, prestando uma especial atenção aos constrangimentos e às sugestões apresentadas pelos participantes ao longo dos trabalhos.

De uma forma geral, constatou-se que os processos decisionais se encontram centrados na figura do Presidente. Verificaram-se algumas (poucas) delegações nos Vereadores e também nos Chefes de Divisão para decisões de carácter operacional e assinatura de correspondência e, pelo menos numa autarquia, identificou-se a subdelegação na chefia da secção. Verifica-se pois que as autarquias não retiram o alcance possível do disposto nos artºs nºos 69º (distribuição de funções) e 70º (delegação de competências no pessoal dirigente) do actual quadro de competências do regime jurídico de funcionamento das autarquias.

Os meios técnicos e tecnológicos existentes ao serviço das autarquias, numa apreciação quantitativa e sumária, apresentavam uma relação colaborador/equipamento globalmente muito boa, sem prejuízo de alguns domínios carecerem de aplicações informáticas. A maior dificuldade prende-se com o nível da sua utilização: identificaram-se aplicações sem utilização há vários anos por falta de formação dos potenciais utilizadores; outras registam níveis de utilização bastante baixos face às suas capacidades funcionais. A formação dos utilizadores das aplicações no local de trabalho poderia ajudar a ultrapassar as dificuldades encontradas. Paralelamente a esta utilização, existem também livros de controlo manual que constituem redundâncias e duplicações dos meios informáticos existentes.

O desfasamento existente entre o somatório dos tempos parciais necessários à realização das actividades processuais e o tempo real que medeia entre o pedido do cidadão e sua satisfação, que foram estimados pelos participantes, encontra uma parte significativa da sua justificação devido às transições no fluxo de actividades que se verificam entre os diversos intervenientes, resultado a uma especialização excessiva que provoca um consequente alongamento da cadeia processual. Tal facto é comum às organizações públicas que tendem a funcionar com base na especialização. No âmbito da reengenharia de processos (Hammer e Champy, 2001), o conceito de especialização tem vindo a perder terreno a favor da ideia da reunificação de tarefas em processos coerentes. Essa reunificação pode constituir uma forma interessante para o enriquecimento das funções e para agilizar procedimentos (aumento de eficiência), reduzindo o número de transições no controlo das actividades entre os diversos participantes.

A gestão da informação segue de perto o modelo de pendor centralista e uma prática baseada em processos de natureza manual. Embora com pequenas variações no que respeita ao circuito dos documentos recebidos em mão, o expediente que vem através do correio sofre um despacho inicial do presidente, é registado e datado, geralmente em aplicação informática específica, e somente depois é que se inicia a sua circulação interna. O mesmo documento pode ter mais do que um registo de entrada. Os documentos entregues nos balcões de atendimento nem sempre são enviados para o registo central de expediente. No que se refere à correspondência expedida, há autarquias onde o Presidente assina toda a correspondência e outras em que assina apenas a que é enviada a determinadas entidades previamente seleccionadas.

No domínio da gestão da informação é ainda de referir o facto os formulários produzidos por algumas autarquias não possuírem um identificador único, potenciando a utilização em simultâneo de versões (ligeiramente) distintas do mesmo formulário.

As matérias que se prendem com a gestão da correspondência (entrada e saída) e a da delegação de competências são de importância fulcral na racionalização dos processos, e devem ser repensadas.

O controlo de gestão parece ser feito de forma empírica, não se tendo confirmado a existência de aplicações informáticas para apoio ao controlo da gestão.

5. Reengenharia de processos

Um procedimento é, de acordo com o artº1º do Código do Procedimento Administrativo uma sucessão ordenada de actos e formalidades tendentes à formação da manifestação da vontade da Administração Pública ou à sua execução (CPA, 1998).

Da análise efectuada constatou-se que, em geral, num procedimento típico da administração local podemos encontrar os seguintes actos e formalidades:

> 1. Requerimento; 2. Recepção e saneamento do pedido; 3. Pagamento; 4.Registo de entrada; 5. Criação de um processo; 6. Apreciação e distribuição do processo; 7. Elaboração de um parecer técnico; 8. Consulta interna e/ou externa (se necessário); 9. Elaboração de um parecer final; 10. Despacho do processo (pré-decisão/decisão); 11. Execução; 12. Comunicação da decisão (ofício); 13. Numeração/Expedição; 14. Arquivo do processo.

A operacionalização dos procedimentos nas organizações analisadas são suportadas, na sua quase totalidade, por um conjunto de aplicações informáticas designadamente: Expediente; Tesouraria/Pocal; Águas; Feiras e Mercados; Urbanismo; Cemitérios. Com menor incidência identificou-se a existência de Gestão documental e Workflow.

Os procedimentos levantados nas várias autarquias apresentavam, contudo algumas diferenças, não obstante terem revelado uma grande semelhança na sua operacionalização mercê do seu enquadramento legal comum. O facto de diversas autarquias utilizarem as mesmas aplicações informáticas, bem como partilhem formulários idênticos que são produzidos pela mesma empresa gráfica, permite antecipar a possibilidade de uma harmonização de processos mais alargada.

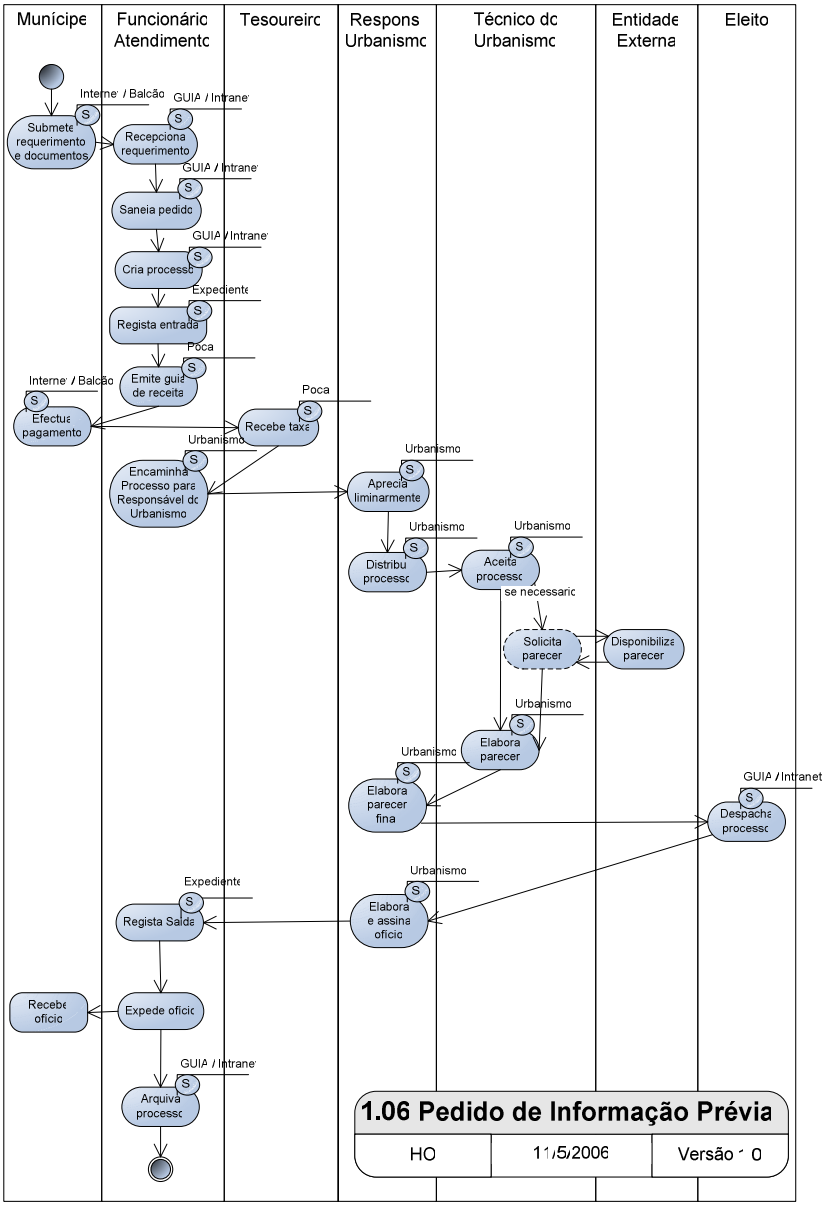
As principais diferenças, de impacto significativo na cadeia processual, residem sobretudo no circuito do expediente, nos patamares de decisão e no nível de utilização dos recursos técnicos disponíveis. Nas autarquias onde os eleitos e o pessoal dirigente têm competências delegadas e /ou subdelegadas, a cadeia processual é mais curta, com claros benefícios para os custos de funcionamento associados ao procedimento e para a satisfação para os clientes. O grau de utilização da tecnologia pode gerar também significativos ganhos de produtividade.

5.1 Desenho dos Procedimentos de Referência

A reengenharia de processos tem por objectivo a simplificação e integração sequencial das actividades, a redução de sobreposições e redundâncias, a minimização dos atrasos e dos custos, o que implica delegação de poder e daí exigir uma maior responsabilização de cada participante no processo.

O desenho dos procedimentos de referência, assenta na utilização de aplicações informáticas adequadas, na racionalização do circuito do expediente, na diminuição das situações de transição entre participantes, procurando responder à necessidade de simplificação e desburocratização dos circuitos de decisão.

Para ilustrar a aproximação adoptada, seleccionamos o procedimento de Pedido de Informação Prévia, da área do Urbanismo, que foi mencionado pela generalidade das autarquias como passível de ser disponibilizado on-line aos cidadãos. Como se pode constatar o Munícipe poderá solicitar este serviço de pedido de informação prévia através da Internet ou presencialmente no Balcão de atendimento. O contacto com o munícipe é centrado no Funcionário do Atendimento (guichet único / front office), que faz a recepção, regista, saneia de acordo com uma checklist do procedimento, numera sequencialmente (independentemente do local de inserção), classifica a documentação de acordo com o classificador existente (dispensando assim despacho inicial do presidente) e é ele que se articula com os serviços técnicos (back-office). No back-office o Responsavel de Área analisa o processo e remete-o ao Eleito que tutela o respectivo pelouro (grupo processual), para despacho definitivo e executório. Emitido o despacho o processo regressa ao Responsável de Área que comunica ao requerente (elabora e assina o ofício) o respectivo teor que recaiu sobre a sua pretensão, que seguidamente envia para o sector de Expediente.



Para a realização das tarefas os funcionários terão à sua disposição um sistema informático integrado constituído por diversas aplicações informáticas de natureza operacional. Esse sistema de informação terá por base duas aplicações de gestão de informação: uma aplicação de gestão documental e uma aplicação de workflow, que asseguram a integração organizacional ao nível dos dados de base (cidadãos, propriedade, processos, contactos, documentos, etc.) e ao nível do acompanhamento das tarefas. Na perspectiva do utilizador (funcionário, dirigente, eleito, ou mesmo cidadão), a uniformização do interface e o controlo de acessos às funcionalidades adequadas às actividades desempenhadas, será facilitada pelo uso de uma aplicação baseada princípios de Intranet e Extranet. Para efeitos ilustrativos esta aplicação surge referida como GUIA – Gestão Unificada da Informação Autárquica.

No caso específico do Pedido de Informação Prévia a aplicação informática operacional de base será a de Urbanismo, utilizada predominantemente ao nível das tarefas técnicas de back-office. Todavia o funcionário do atendimento têm à sua disposição o GUIA / INTRANET para poder

aceder aos dados pessoais do munícipe e dos processos de urbanismo com ele relacionados, à aplicação de Expediente para controlo dos documentos recebidos e expedidos, à aplicação de POCAL para emissão de Guia de Receita, bem como à aplicação de Urbanismo.

Os mesmos princípios de reengenharia – simplificação, automatização e integração – ilustrados no processo de Pedido de Informação Prévia são aplicados aos restantes procedimentos de referência.

5.2 Vectores da Reorganização Organizacional

Os vectores da reorganização deverão contemplar mudanças em diferentes domínios designadamente: Estrutural e Funcional; Tecnologia (TI/SI); Qualificação dos Funcionários

Estrutural e Funcional

- Adoptar um funcionamento por processos e não por funções
- Separar o atendimento (front-office) do apoio técnico (back-office)
- Focalizar a organização no munícipe e não no requerimento.
- Generalizar o uso do processo associado ao cidadão enquanto elemento agregador da actividade realizada pela Câmara.
- Promover uma gestão integrada da informação
- Promover a delegação de competências (artº 69 e 70 da Lei das Autarquias Locais)
- Estimular a melhoria de desempenho da organização

Tecnologia (TI/SI)

- Optimizar o uso das aplicações informáticas na tomada de decisão e para o controlo de gestão
- Adoptar a transmissão electrónica da informação
- Utilizar as aplicações informáticas (workflow) para agilizar os procedimentos operacionais
- Organizar a plataforma tecnológica (Intranet e Extranet) para disponibilizar serviços aos munícipe e aos funcionários.

Qualificação dos Funcionários

- Enriquecimento das funções
- Formação
- Reforço da motivação

6. Recomendações

A realização do Estudo de Levantamento e Reengenharia de Processos conduziu-nos a um conjunto recomendações simples e objectivas, as quais levadas à prática conduzirão a benefícios que justificarão plenamente os recursos aplicados. No contexto deste artigo destacaremos as seguintes:

1. Construir de forma abrangente e partilhada uma visão da Câmara Municipal, colocando a ênfase na satisfação das necessidades concretas dos munícipes de forma integrada e, como resultado, repensando criticamente a estrutura e o funcionamento da autarquia;

2. Instituir o munícipe (e não o requerimento) como o centro de toda a actividade autárquica e gerir a informação de forma integrada (ciclo de vida do munícipe);
3. Criar e operacionalizar o conceito do gestor do processo;
4. Criar um guichet único, com separação entre front-office e back-office.
5. Agilizar os procedimentos (eficiência) verticalizando os fluxos de tarefas e reduzindo o número de transições entre participantes, como forma de diminuir a diferença entre o tempo efectivo estimado de trabalho dedicado ao processo e o tempo de resposta ao munícipe/cliente;
6. Reforçar as práticas de gestão integrada de informação utilizando os meios informáticos, melhorando a funcionalidade das aplicações no domínio da gestão documental e acompanhamento das actividades (workflow);
7. Criar um sistema de indicadores de desempenho nas dimensões de qualidade, tempo, custo e satisfação (quadro de bordo);
8. Permitir aos dirigentes e eleitos um melhor conhecimento dos documentos entrados e acompanhamento da execução das actividades através de relatórios/listagens de controlo de gestão;
9. Melhorar a capacitação técnica e comportamental dos funcionários, promovendo o trabalho em equipa e enriquecendo o seu conteúdo funcional, em particular alargando o espectro de competências adequadas ao desempenho das actividades de atendimento (front-office);
10. Promover iniciativas de colaboração inter-municipal, centradas na harmonização de procedimentos e na melhoria do desempenho do funcionamento das autarquias.

7. Referências

Michael Hammer, James Champy, Reengineering the Corporation, Harper Collins, 2001

CPA, Código do Procedimento Administrativo, Secretariado para a Modernização Administrativa, Gráfica Jesus, 1998

Modelação e Verificação Formal de Processos de Negócio em Ambiente Hospitalar

Carlos Santos [1], Carla Ferreira [2], José Tribolet [3]

1) Instituto Superior de Contabilidade e Administração, Universidade de Aveiro, Rua Associação Humanitária dos Bombeiros de Aveiro, 3811-902, Aveiro, Portugal;

CEO, Centro de Engenharia Organizacional – INOV, INESC-Inovação, Rua Alves Redol, nº 9, 1000-029, Lisboa, Portugal

carlos.santos@isca.ua.pt

2) Departamento de Engenharia Informática, Instituto Superior Técnico, Avenida Professor Cavaco Silva, 2780-990, Porto Salvo, Portugal
carla.ferreira@dei.ist.utl.pt

3) Departamento de Engenharia Informática, Instituto Superior Técnico, Av. Rovisco Pais, 1049-001 Lisboa, Portugal;

CEO, Centro de Engenharia Organizacional – INOV, INESC-Inovação, Rua Alves Redol, nº 9, 1000-029, Lisboa, Portugal
jose.tribolet@inesc.pt

Resumo

Embeber mecanismos de controlo interno em processos de negócio em ambiente hospitalar, com o objectivo de minimizar o risco que lhes está associado, pode comprometer a consistência dos modelos de negócio. Por esta razão, a verificação formal de processos hospitalares, com mecanismos de controlo embebidos, assume um papel relevante em direcção à auditoria contínua dos sistemas de informação hospitalares.

Este trabalho descreve a aplicação de técnicas e metodologias do âmbito da Engenharia Organizacional à modelação do negócio de um serviço de urgência hospitalar. Aos modelos assim obtidos são aplicadas técnicas formais, que permitem a sua simulação e verificação de conformidade das regras fundamentais do negócio. A existência de eventuais inconsistências será despistada através da aplicação destas técnicas, o que permitirá que se proceda às correcções adequadas.

A representação de sistemas complexos deve ser feita sob duas perspectivas: uma perspectiva "micro" em que o sistema é analisado ao nível das transacções organizacionais e uma perspectiva "macro" em que o sistema é analisado ao nível dos processos de negócio. A análise ao nível transaccional, focada na troca de mensagens entre processos, proporciona um bom suporte à aplicação da técnica formal "*model checking*".

Palavras-chave: engenharia organizacional; "*model checking*"; verificação formal; PROMELA; SPIN.

1. Introdução

A complexidade dos sistemas de informação hospitalares, requer que a sua modelação seja feita com grande detalhe, relativamente aos seus componentes, estrutura, actividades e procedimentos. Nesta perspectiva, a avaliação deste tipo de sistemas pressupõe uma descrição tão precisa quanto possível de como trabalham, antes de elaborar modelos de simulação [Combes, et al. 1993] que representem formalmente as situações reais que se pretendem estudar.

A modelação de processos hospitalares coloca-nos perante dois dilemas, ambos muito importantes e que devem ser tidos em atenção durante o processo de modelação:

- a necessidade de adoptar modelos informais de alto nível que possam ser facilmente construídos, dando uma percepção aproximada do que se pretende;
- a necessidade de adoptar modelos formais e detalhados, de difícil e complexa construção [Janssen and Mateescu 2000], mas necessários se se pretender uma representação rigorosa que possa influenciar favoravelmente a confiança dos "*stakeholders*" nos processos modelados.

Face ao exposto, o objectivo do presente trabalho consiste em modelar processos hospitalares recorrendo a técnicas semi-formais e na sua transformação, através da sua especificação numa linguagem formal, em modelos que, embora não exprimam totalmente a realidade, representam a sua abstracção e descrevem as características essenciais do processo em estudo. Estes modelos tornam possível a utilização de técnicas formais, para simular a execução de processos em condições reais de funcionamento, permitindo validar o seu comportamento.

Este trabalho está estruturado em 5 secções, para além desta secção em que é feita a sua introdução, temos ainda: a secção 2 em que se faz o enquadramento teórico, referindo as metodologias fundamentais que foram usadas; a secção 3 em que é apresentado o caso de estudo desenvolvido num serviço de urgência hospitalar com foco no processo "admissão de doente" no contexto "macro" (processo, sub-processos e actividades) e no contexto "micro" (transacções); a secção 4 em que é desenvolvido o aspecto da simulação e da verificação formal de processos hospitalares temporizados com recurso à linguagem PROMELA (linguagem de entrada no verificador de modelos SPIN) e finalmente a secção 5 em que se dá nota das conclusões e proposta de trabalho futuro.

2. Enquadramento teórico

Vários exemplos[1] têm demonstrado que os erros num sistema podem causar danos económicos, físicos e humanos significativos. Com o objectivo de minimizar os potenciais erros afectos a qualquer sistema surgiu a necessidade de simular e verificar formalmente os seus modelos com o objectivo de analisar o comportamento e a conformidade das suas propriedades.

Para que os modelos desenvolvidos tenham uma perspectiva modular e sejam independentes da tecnologia que possa vir a ser utilizada na sua implementação utilizamos como linguagem de modelação a linguagem UML [Booch, et al. 1999].

No presente trabalho utilizaremos uma abordagem mista proposta por [Gospodarevskaya, et al. 2005], baseada na modelação de processos e na simulação e verificação formal de modelos utilizando a técnica *"model checking"*.

[1] Intel Pentium Chip; ARIANE; Therac-25

2.1 Modelação do negócio

Para modelar o negócio recorremos à framework do Centro de Engenharia Organizacional[2] (FCEO), figura 1. Esta estrutura de referência assenta na definição de três conceitos fundamentais: estratégia de negócio, processos de negócio e sistemas de informação. A FCEO recorre à criação de um novo perfil para a linguagem UML, que permite descrever os conceitos organizacionais referidos [Tribolet 2002].

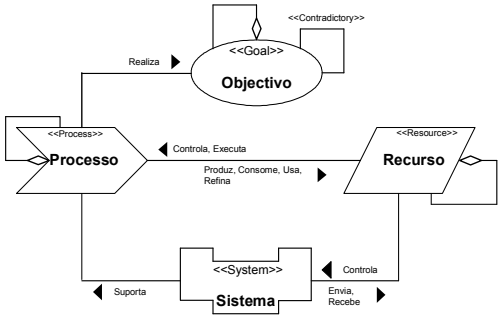


Figura 1 – Estrutura de referência CEO

A framework CEO contém um conjunto limitado de objectos de negócio com os quais se constroem modelos de negócio. Os objectos de negócio considerados nesta estrutura de referência são: os objectivos, para a modelação de estratégia; os processos, para a modelação de processos de negócio; os recursos, para a modelação dos recursos de negócio; e os blocos, para modelação dos blocos arquitecturais dos sistemas de informação. Para além dos objectos de negócio esta estrutura de referência descreve também as relações entre eles, bem como um conjunto de elementos predefinidos, construídos a partir dos conceitos base, e que representam as melhores práticas na área da modelação organizacional.

2.2 Processos temporizados

No âmbito da modelação de processos hospitalares não tem sido dado o relevo merecido à variável tempo. Ainda que grande parte das análises feitas sobre estes processos possa ser independente daquela variável, existem outras em que o tempo é determinante. Neste trabalho damos especial atenção à modelação de processos de negócio temporizados.

Embora se comecem a fazer sentir preocupações com a modelação do tempo, ainda há muito a fazer neste domínio. Por exemplo, a linguagem PROMELA que serve de entrada ao verificador de modelos SPIN, que é por nós utilizado neste trabalho não dispõe das necessárias primitivas para modelar o tempo sendo necessário estabelecer alguns pressupostos.

Na abordagem que fazemos, neste trabalho, para modelar o tempo consideramos dois mecanismos para sua medição: temporizadores e relógios. Os temporizadores têm por função controlar o tempo de execução de uma transacção incluindo o tempo que se possa verificar na comunicação entre os recursos envolvidos nessa transacção. Os relógios têm a função de contar as unidades de tempo ao longo de todo o percurso do processo organizacional. Vamos seguir de perto estudos feitos sobre a modelação do tempo por [Alur, et al. 1966,Bosnacki 1998,Guelfi and Mammar 2005,Li and Lilius 1999,Mammar and Guelfi 2004,Matousek 2003,Li, et al. 2001]

[2] Centro de investigação do Instituto de Engenharia de Sistemas e Computadores (INESC-INOV).

Para que a transposição do modelo semi-formal de uma transacção temporizada (ver figura 2) se possa fazer para linguagem PROMELA, de forma transparente e rigorosa, devem ser conhecidas a sua sintaxe e semântica que permitem fazer o enquadramento formal de todos os pressupostos associados aos modelos em análise. Assim, a definição da sintaxe e da semântica para modelação do tempo associado aos processos hospitalares temporizados é uma peça essencial para que se possa proceder à sua verificação formal.

O tempo pode ser modelado segundo uma de duas perspectivas: ou se trata o tempo de forma discreta; ou se trata o tempo de forma densa. Neste trabalho vamos modelar o tempo de forma discreta.

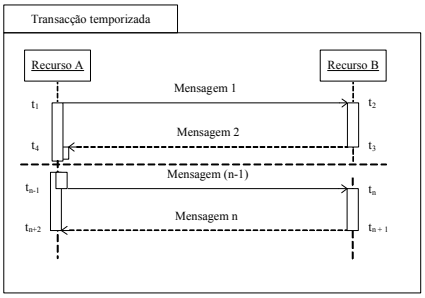


Figura 2 – Transacção temporizada

2.3 Validação formal

A validação formal de processos hospitalares está relacionada com a especificação desses processos em termos formais (com base em conceitos matemáticos, tais como lógica, teoria dos conjuntos, etc.). A especificação formal tem vantagens e desvantagens (ver tabela 1).

Existem diversas técnicas de validação de modelos, sendo de destacar as duas seguintes:

- validação por via semântica *"model checking"*. O processo a validar é representado por um conjunto de estados finito, sendo a validação automática e exaustiva;
- validação por via dedutiva "*theorem proving*". O processo a validar é representado por um conjunto de estados infinito e a validação é apenas parcialmente automática embora seja "inteligente".

Vantagens	Desvantagens
Os requisitos são definidos formalmente (contrato entre o cliente e equipa de desenvolvimento).	Não é possível provar formalmente que a especificação satisfaz os requisitos do cliente (validação).
Documentação do sistema com um nível elevado de abstracção.	Notações formais dificultam a comunicação com o cliente.
Demonstração formal das propriedades do sistema.	Nem todos os problemas podem ser formalizados de forma elegante.
Prototipagem rápida (execução do modelo formal) e geração automática de código.	Uso de métodos formais requer uma equipa de desenvolvimento qualificada.
Refinamento sucessivo do modelo abstracto até à implementação.	O desenvolvimento formal de um sistema é trabalhoso.

Tabela 1 – Vantagens e desvantagens da especificação formal de modelos.

A técnica *"model checking"* verifica as propriedades de um modelo enumerando exaustivamente todos os estados alcançáveis, devem assim, seguir-se os seguintes passos, (ver figura 3):

- construir o modelo formal do sistema – Modelo M;
- especificar as propriedades que o modelo deve ter para poder ser considerado correcto – Propriedade ↖;
- executar o verificador de modelos.

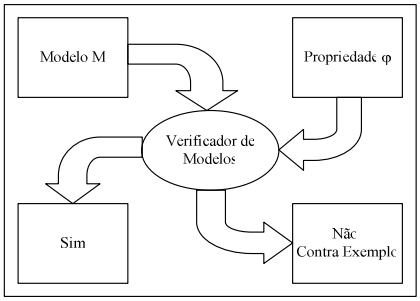


Figura 3 – Verificação formal

A utilização da técnica *"model checking"* deve-se ao seu elevado grau de automatização e à existência de ferramentas de simples utilização bastante utilizadas em vários tipos de modelos. Uma dessas ferramentas é o verificador de modelos SPIN que utiliza como linguagem de entrada a linguagem PROMELA, esta linguagem é muito semelhante à linguagem C o que torna fácil a sua interpretação para quem tenha alguns conhecimentos desta linguagem de programação.

3. Caso de estudo

O presente caso de estudo foi desenvolvido num serviço de urgência hospitalar, serviço em que a qualidade e a quantidade de informação associada a qualquer processo são factores determinantes para o seu bom funcionamento, particularmente no que respeita à qualidade que deve ser dispensada aos processos hospitalares.

Nestes processos é frequente acontecerem situações de concorrência. Um determinado número de recursos A (doentes), necessita de ter acesso concorrentemente a um determinado número de recursos B, sendo A> B. Nesta situação para além da concorrência pode ser importante conhecer o comportamento temporal da execução do processo para poder decidir sobre a afectação de recursos ao referido processo.

3.1. Modelação de processos

A cadeia de valor de um serviço de urgência hospitalar, figura 4, é constituída por três processos principais: processo "*admissão de doente*"; processo "triagem de doente"; e, processo "prestar cuidado de saúde". Por uma questão de espaço, neste trabalho, fazemos somente a modelação do processo "*admissão de doente*".

Figura 4 – Cadeia de valor de um serviço de urgência hospitalar

Caso de utilização "admissão de doente"

Qualquer doente à chegada ao serviço de urgência pode apresentar um dos seguintes estados: emergente ou urgente. No caso de ser um doente emergente é encaminhado directamente para a prestação de cuidados de saúde, podendo a sua identificação ser feita paralela ou posteriormente à prestação de cuidados de saúde. No caso de ser um doente urgente pode ainda apresentar dois estados: independente ou dependente. Se o doente é independente segue pelos seus meios para a identificação, se o doente é dependente é transferido para uma cadeira de rodas ou maca rodada, sendo a sua identificação feita simultaneamente pelo seu acompanhante.

Vista funcional

A funcionalidade controla os requisitos na maioria das aplicações. Na perspectiva do doente o sistema deve oferecer informações ou comportamentos para dar suporte ao processo ou aos objectivos de um serviço de urgência hospitalar [Pender 2004]. Com os diagramas de casos de utilização é possível modelar o que os doentes esperam encontrar quando interagem com a admissão de um serviço de urgência hospitalar, figura 5. O subsistema "admissão de doente" é responsável pela recepção e identificação dos doentes que se dirijam a um serviço de urgência hospitalar.

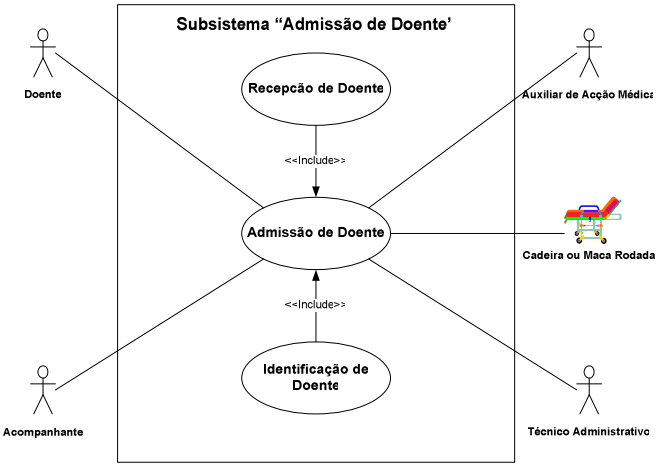


Figura 5 – Diagrama de caso de utilização do subsistema "admissão de doente"

Diagrama de sequência do subsistema "admissão de doente"

O diagrama de sequência do subsistema "admissão de doente" é representado na figura 6. Neste diagrama podem ver-se as principais interacções (transacções) que ocorrem neste subsistema:

- TID1: transacção que se verifica entre o doente e o técnico administrativo (doente urgente independente);
- TID2: transacção entre o técnico administrativo e o repositório de dados;
- TID3: transacção que se verifica entre o acompanhante e o técnico administrativo (doente urgente dependente). Esta transacção é em tudo idêntica a TID1;
- TRD1: transacção entre o doente e o auxiliar de acção médica (doente urgente dependente);
- TRD2: transacção entre o auxiliar e o repositório de cadeiras ou macas rodadas.

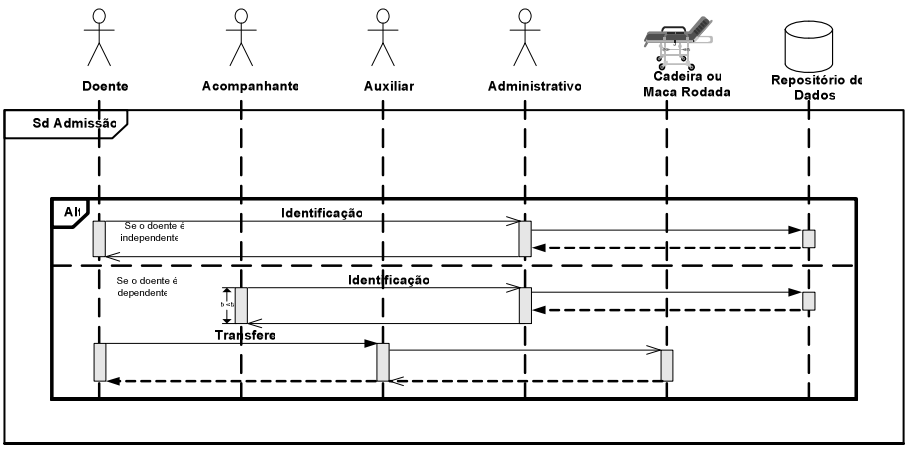


Figura 6 – Diagrama de sequência do subsistema "admissão de doente".

Temporização dos processos

A aplicação de técnicas formais a modelos de negócio temporizados em ambiente hospitalar exige a identificação de todas as transacções que deles fazem parte (análise" micro"). Numa perspectiva sócio-tecnológico, qualquer transacção (ver figura 2), deve estar sujeita aos seguintes pressupostos: (1) uma transacção é a unidade básica de um processo; (2) são trocadas mensagens, entre os intervenientes, até se verificar uma mudança de estado; (3) em cada transacção intervêm apenas dois recursos; (4) qualquer mudança de estado é permanente; (5) as mensagens são autorizadas; (6) deve ser garantida a autenticidade das mensagens; (7) deve ser garantida a confidencialidade das mensagens [Guiochet 2003].

Sintaxe

Para modelar o tempo no contexto do presente trabalho consideramos que um diagrama de sequência UML temporizado pode ser representado como um grafo no sentido matemático. Assim, de facto, o diagrama de sequência pode ser representado pelo seguinte túplo DST = (recurso, mensagem, relógio, temporizador):

- ***recurso***: representa os dois recursos envolvidos na transacção modelada através do diagrama de sequência;
- ***mensagem***: é desencadeada por cada um dos recursos quando estes necessitam de manifestar algum tipo de comportamento que exija a participação do outro recurso envolvido na transacção;
- ***relógio***: tem a finalidade de gerir a duração de uma determinada transacção independentemente dos temporizadores associados a cada um dos recursos;
- ***temporizador***: está afecto a cada um dos recursos envolvidos na transacção e refere-se à duração da execução da sua participação numa determinada transacção.

Utilizando esta sintaxe podemos representar o diagrama de sequência representado na figura 2, como sendo um túplo (R, M, Rlg, T), sendo:

- $R = \{recurso_1; recurso_2\}$;
- $M = \{mensagem_1; mensagem_2; \ldots, mensagem_{n-1}; mensagem_n\}$;
- $Rlg = \{t_1=0; t_2=0; \ldots; t_n=0\}$;
- $T = \{t_1, t_2, t_3, t_4 \ldots\}$

Semântica

Para modelação dos processos, vamos recorrer aos diagramas de sequência UML para tratar as restrições relativas ao tempo (ver figura 7). Quando um recurso envia uma mensagem deve ser activado um temporizador, como se pode facilmente depreender a acção desencadeada pelo envio de uma mensagem terá uma certa duração, que pode ser importante considerar na verificação dos modelos. Também pode acontecer existirem atrasos entre transacções que devam ser considerados.

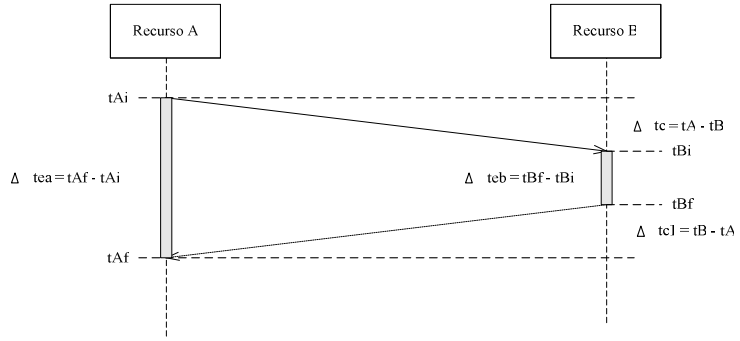


Figura 7 – Semântica do diagrama de sequência de um processo temporizado

Para modelar o tempo introduzimos restrições temporais associadas às diversas mensagens trocadas. Cada mensagem temporizada terá associado um intervalo de tempo $[t_A, t_B]$, em que t_A e t_B são relativos ao instante em que a mensagem temporizada tem início (t_A) e ao instante em que tem início a resposta à mensagem (t_B). Este intervalo de tempo é associado a um temporizador inerente a cada uma das mensagens temporizadas ($\Delta t_C = t_A - t_B$).

Também a execução das transacções tem uma determinada duração. Cada execução tem associado um intervalo de tempo $[t_{Ai}, T_{Af}]$, em que t_{Ai} é relativo ao instante em que a execução tem início e t_{Af} é relativo ao instante em que a execução termina. Também a estes tempos são associados temporizadores, à semelhança do que acontece para a temporização de mensagens ($\Delta t_{ea} = t_{Af} - t_{Bi}$).

Para além dos temporizadores utilizaremos relógios. Considerando que num processo hospitalar podem coexistir, simultaneamente, várias transacções que podem ser tratados concorrentemente, associaremos a cada um dos objectos um relógio com o objectivo de contar as unidades de tempo associadas a cada um deles em todo o percurso ao longo do processo. Considera-se o tempo dividido em fatias, o relógio permanecerá em cada fatia enquanto existir algum temporizador associado a um objecto em actividade.

Conforme se pode verificar (ver figura 8) podemos estabelecer a medida em unidades de tempo entre dois estados do mesmo temporizador, mas só poderemos estabelecer uma comparação relativa entre temporizadores diferentes na mesma unidade de tempo.

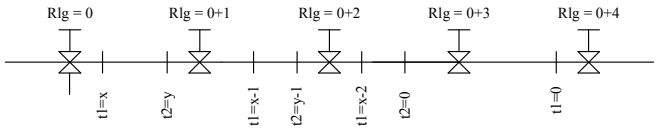


Figura 8 – O relógio é incrementado enquanto algum temporizador for positivo

4. Validação formal

A validação formal consiste na construção de um modelo utilizando uma linguagem formal que permita simular o seu funcionamento e verificar as suas propriedades. Entre as diversas técnicas existentes, para verificação formal de sistemas, a verificação de modelos *"model checking"* tem alcançado resultados bastante relevantes e tem sido cada vez mais utilizada fora do meio académico, devido à sua facilidade de uso, aplicabilidade e à existência de ferramentas para verificação automática de propriedades.

As regras associadas ao comportamento de uma transacção podem facilmente ser transpostas para LTL (Linguagem Temporal Lógica). Estas regras associadas ao modelo formal do processo permitem através da utilização de um verificador de modelos verificar a conformidade dos processos. Estas técnicas são úteis na verificação de processos hospitalares, sendo actualmente consideradas imprescindíveis em sistemas que envolvam risco.

O verificador de modelos SPIN opera em dois modos. O primeiro modo permite que o utilizador se familiarize com o comportamento do sistema modelado, simulando a sua execução. O segundo modo faz a verificação de todos os estados possíveis verificando a conformidade de regras [Bérard, et al. 2001]. O SPIN pode ser utilizado para verificar a consistência de sistemas que executem processos concorrentes [Holzmann 2004].

Inicialmente todos os relógios são colocados a "zero". Os temporizadores quando activados são colocados no valor correspondente à duração da transacção temporizada correspondente. Os temporizadores são decrementados e enquanto algum deles estiver activo o relógio correspondente vai sendo incrementado.

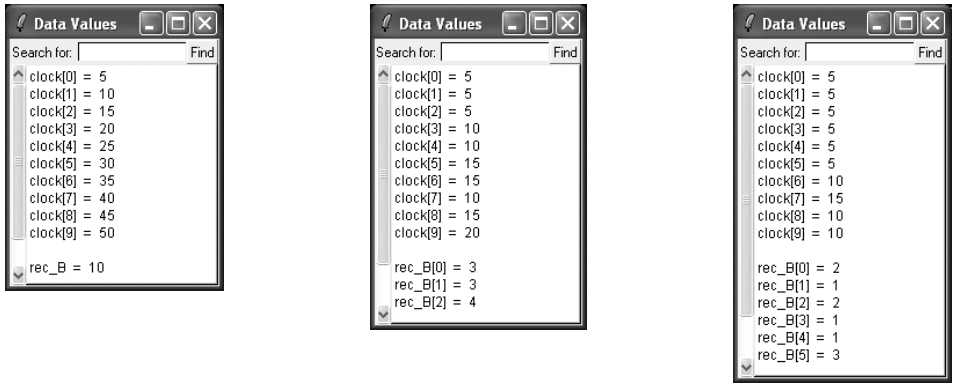
Definidos os processos, é usual testar os seus modelos quanto à sua correcção formal [Fisteus, et al. 2004]. Isto pode ser feito com recurso à verificação formal dos modelos com recursos a ferramentas adequadas. No entanto é necessário que os modelos ao serem construídos tenham em atenção a utilização de uma semântica assim como de uma sintaxe, conforme já referido. Em [Eshuis 2002] é adicionada semântica aos digramas de actividade UML para facilitar a transformação destes modelos na linguagem formal PROMELA.

4.1. Simulação

O modelo formal do diagrama de sequência UML de um processo hospitalar concorrente temporizado pode ser transposto para linguagem PROMELA. Após a sua transposição para PROMELA pode efectuar-se a sua simulação com o verificador de modelos SPIN.

Neste caso tomamos como exemplo a transacção já referida, admitindo que existem 10 doentes que necessitam de ter acesso a serviços fornecidos pelo recurso B. Vamos simular o seu comportamento para três situações: a) existe somente um recurso B; b) existem três recursos B e c) existem seis recursos B. Os resultados para cada uma das situações referidas podem ver-se na figura 9, considera-se que a resposta a cada pedido demora em média 5 unidades de tempo.

Note-se que no caso representado em a) os pedidos são satisfeitos sequencialmente demorando a resposta ao último pedido 50 unidades de tempo. Na situação mostrada em b) a resposta às solicitações é feita de forma concorrente, demorando a resposta ao último pedido 20 unidades de tempo. Similarmente na situação retratada em c) o último pedido é satisfeito em 15 unidades de tempo porque existem mais recursos a fornecerem serviços.



a) b) c)

Figura 9 – Resultado da simulação efectuada com o SPIN

4.2. Verificação

A verificação do modelo permite averiguar se determinadas regras do negócio são verificadas. No caso em análise há algumas regras básicas do negócio que podem ser modeladas e verificadas. Vamos considerar uma regra simples que consiste em garantir que todo o doente (recurso A) que solicita a participação de um recurso B é identificado.

Regra
Todo o doente deve ser identificado
#define s (len (fila_admin) == 0)
Formula As Typed: <> [] s

Quadro 1 – Descrição da regra que garante que todo o doente é identificado

A identificação de todos os doentes que chegam a um serviço de urgência hospitalar é um factor determinante para minimizar o risco de prestar cuidados de saúde a doentes não identificados, minimizando-se assim o risco de não identificar atempada e adequadamente o episódio clínico no registo clínico do doente.

Na figura 10 vemos o quadro com o resultado da verificação do modelo do processo concorrente temporizado que estamos a considerar neste trabalho.

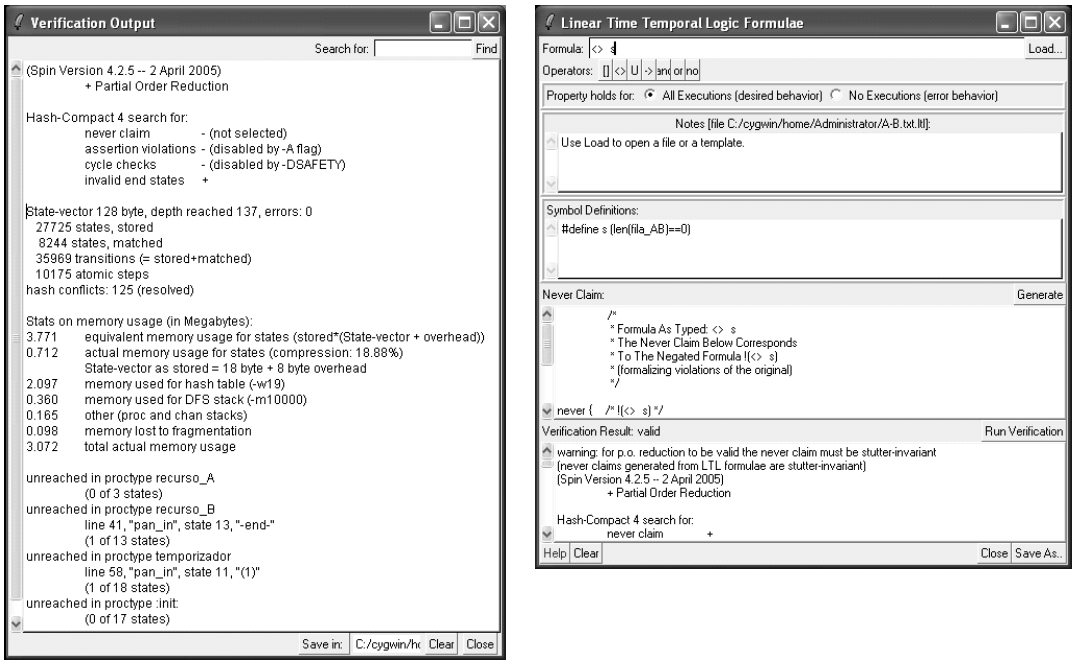


Figura 10 – Resultados da verificação do modelo e fórmula temporal para verificar se todos os doentes são identificados

5. Conclusão e proposta de trabalho futuro

A modelação de processos de negócio em ambiente hospitalar constitui um importante instrumento na definição de uma metodologia que associada ao conceito de transacção, é fundamental para que se possa modelar o processo subjacente a cada transacção e para que se possam seguidamente aplicar metodologias com base em modelos formais que permitam a sua validação formal.

A proposta de modelação, que apresentamos no presente trabalho, utilizando a estrutura de referência CEO juntamente com diagramas de sequência UML é uma proposta viável conforme se pode ver pela apreciação do exemplo apresentado.

A simulação e a verificação formal de modelos de negócio são fundamentais para que se possa garantir a sua coerência e consistência. Esta garantia permite avaliar o comportamento de negócios tão complexos como um serviço de urgência hospitalar permitindo pela análise dos modelos propor soluções que melhorem o desempenho do negócio em situação real.

O presente trabalho permite concluir que a modelação de processos de negócio em ambiente hospitalar com um nível de abstracção que permita a sua compreensão sem tornar a tarefa demasiado complexa e em que seja igualmente possível modelar os mesmos processos com recursos a lógica formal permite efectuar a sua validação formal (simulação e verificação) com recurso a um verificador de modelos.

A variável tempo, em determinados processos de negócio em ambiente hospitalar, reveste-se de extraordinária importância sendo oportuno desenvolver esforços no sentido de a representar formalmente e poder verificar regras temporais.

6. Referências

Combes C, Force C, Kellert P: Modelling Methodology for Hospital Systems: Application to an Emergency Department. In IEEE (ed): Systems, Man and Cybernetics. Le Touquet, France, 1993.

Janssen W, Mateescu R: Verifying Business Process Using SPIN. InEnschede, Telematics Institute, 2000.

Booch G, Rumbaugh J, Jacobson I: The Unified Modeling Language User Guide. Addison Wesley, 1999.

Gospodarevskaya E, Churilov L, Wallace L: Modelling the Patient Care Process of an Acute Care Ward in a Public Hospital: A Methodological Perspective. In 38th Hawaii International Conference on System Sciences. Big Island, Hawaii, 2005.

Tribolet J: Engenharia Organizacional: A Engenharia ao Serviço das Organizações. In, CEO, INESC-INOV, 2002.

Alur R, Holzmann GJ, Peled D: An Analyzer for Message Sequence Charts. Software-Concepts and Tools 1966;17:70-77.

Bosnacki D: Extending Promela and Spin with Discrete Time. In VIRES project meeting. Grenoble, 1998.

Guelfi N, Mammar A: *A Formal Semantics of Timed Activity Diagrams and its PROMELA Translation*. In Press ICS (ed): APSEC'2005: Asia Pacific Software Engineering Conference. Taipei, Taiwan, 2005.

Li X, Lilius J: Timing Analysis of UML Sequence Diagrams. In TUCS Technical Report No 281. Turku, University of Turku, 1999:16.

Mammar A, Guelfi N: A Formal Approach for the Verification of E-business Processes Using the Promela language. InLuxembourg, University of Luxembourg, 2004.

Matousek P: Verification of Business Process Models. In departament of Computer Science. Ostrava, Technical University of Ostrava, 2003:94.

Li X, Meng C, Yu P, Jianhua Z, Guoliang Z: Timing Analysis of UML Activity Diagrams. in Gogolla M, Kobryn C (eds), 2001.

Pender T: UML, a Bíblia. ed tradução, Rio de Janeiro, Elsevier, 2004.

Guiochet J: Maîtrise de la Sécurité des Systèmes de la Robotique de Service: Approche UML Basée sur une Analyse du Risque Système. In Laboratoire d'étude des systèmes Informatiques et Automatiques. Toulouse, L'Institut National des Sciences Appliquées de Toulouse, 2003.

Bérard B, Bidoit M, Finkel A, Laroussinie F, Petit A, Petrucci L, Schnoebelen P, McKenzie P: Systems and Software Verification: Model-Checking Techniques and Tools. Berlim, Springer, 2001.

Holzmann GJ: The Spin Model Checker: Primer and Reference manual. Boston, Addison-Wesley, 2004.

Fisteus JA, López AM, Kloos CD: VERBUS: A Formal Model for Business Process Verification. In Information Resources Management Association International Conference. New Orleans, Louisiana, USA, 2004.

Eshuis H: Semantics and Verification of UML Activity Diagrams for Workflow Modelling. In Centre for Telematics and Information Technology (CTIT). Enschede, Unversity of Twente, 2002:240.

Tecnologias e Sistemas de Informação para a Indústria da Panificação e Pastelaria

António Jorge Mesquita [1], João Varajão [2], Ramiro Gonçalves [3]

1) Departamento de Engenharias - UTAD, Vila Real, Portugal
jmesquita@csgnet.org

2) Departamento de Engenharias - UTAD, Vila Real, Portugal
jvarajão@utad.pt

3) Departamento de Engenharias - UTAD, Vila Real, Portugal
ramiro@utad.pt

Resumo

Procurando contribuir para a melhoria dos níveis de utilização de tecnologias de informação na indústria da panificação e pastelaria em Portugal, desenvolvemos um estudo em que se procuraram respostas para as questões "quais são as necessidades de informação actuais das empresas da indústria da panificação e pastelaria?" e "as soluções de tecnologias de informação actualmente disponíveis no mercado suportam devidamente as necessidades de informação das empresas da indústria da panificação e pastelaria?". Neste artigo são apresentados os resultados desse estudo.

Palavras chave: tecnologias de informação, panificação, pastelaria

1. Introdução

A indústria da panificação e pastelaria, de acordo com a Classificação Portuguesa das Actividades Económicas (CAE - Rev. 2.1), está classificada na secção D - Indústrias Transformadoras, subsecção DA – Indústrias Alimentares das Bebidas e do Tabaco, Divisão 15 - Indústrias Alimentares e das Bebidas, Classe 158, Fabricação de Outros Produtos Alimentares, Subclasse 1581 – Panificação e Pastelaria [INE, sd].

Em Portugal, a panificação e pastelaria é a actividade mais importante do grupo 158, representando, em 2001, 36% do volume de produção e 37% do valor total das vendas deste grupo. No período 1999 e 2001 teve a maior variação do valor das vendas, dentro das Indústrias alimentares e das Bebidas, com 163.909 mil euros. Sendo a 6ª actividade mais importante na Classe, com um valor de vendas de 424.838 mil euros em 1999, passando para 5º em 2001 com um valor das vendas de 588.001 mil euros [INE, 2003a]. O produto com maior produção é o

Pão de Trigo com 72% do volume total e 56% do valor total das vendas e apresentou um aumento de 25% (mais de 50 mil toneladas), entre 1999 e 2001 [INE, 2003b].

Nesta indústria, o controlo da produção é verdadeiramente fundamental, de modo a ser possível aferir-se o que é produzido, que matérias se utilizam para produzir, como se produz, que quantidades se utilizam, o que se perde, quanto se desperdiça e, muito importante, que custos se verificam na produção. Não obstante a importância desta área, é observável que, na prática, o fabrico é tipicamente um processo sem controlo de stocks, sendo o conhecimento das unidades produzidas tipicamente obtido de forma subjectiva [Ferrete 2004].

Por outro lado, esta é uma indústria em que verifica uma grande competitividade e uma grande exigência em termos da necessidade de cumprimento de requisitos legais, de qualidade, de certificação e de conformidade, o que torna a gestão sem suporte a Tecnologias de Informação (TI), que se verifica ainda hoje num grande número de organizações, verdadeiramente ineficiente.

Estes são problemas graves, dado trazerem profundas implicações em termos da eficácia e produtividade das empresas deste sector de actividade.

Procurando contribuir para a melhoria dos níveis de utilização de TI nesta indústria, desenvolvemos um estudo em que se procuraram respostas para as questões "quais são as necessidades de informação actuais das empresas da indústria da panificação e pastelaria?" e "as soluções de TI actualmente disponíveis no mercado suportam devidamente as necessidades de informação das empresas da indústria da panificação e pastelaria?". Neste artigo são apresentados alguns dos resultados desse estudo.

Nas secções subsequentes, após a apresentação do processo de investigação seguido, identificamos as necessidades de informação das empresas que fazem parte desta indústria, caracterizamos as soluções tecnológicas actualmente existentes no mercado e verificamos se as mesmas satisfazem as necessidades de informação identificadas.

2. Processo de investigação

A definição do processo de investigação teve subjacente a procura de respostas para as questões "quais são as necessidades de informação actuais das empresas da indústria da panificação e pastelaria?" e "as soluções de TI actualmente disponíveis no mercado suportam devidamente as necessidades de informação das empresas da indústria da panificação e pastelaria?".

Para a identificação das necessidades de informação actuais das empresas da indústria da panificação e pastelaria, primeiro, foi efectuada uma revisão de literatura específica deste sector de actividade. Desse trabalho, resultou um modelo com a identificação das principais áreas da organização interna das empresas do sector e a identificação preliminar das respectivas necessidades de informação. Seguiu-se a validação e refinamento do modelo, através da realização de um conjunto de entrevistas a entidades que têm como função apoiar as empresas deste sector: colaboradores de empresas; especialistas em engenharia alimentar, contabilistas e fornecedores. As entrevistas ocorreram de forma informal e não estruturada, tendo durado aproximadamente duas horas cada. As necessidades de informação identificadas são apresentadas na secção 3.

Após a identificação das necessidades de informação das empresas deste sector de actividade, procuraram-se diversas soluções existentes no mercado, fez-se um estudo das funcionalidades e informação suportadas por cada solução e foram contrapostas as necessidades de informação para o sector com o suporte identificado. O resultado desta avaliação é apresentado na secção 4.

3. Necessidades da informação para o sector

Nesta secção, são apresentadas e descritas as necessidades de informação das empresas da indústria da panificação e pastelaria, estruturadas de acordo com as principais áreas da sua organização interna, identificadas na figura 1:

- HACCP[1];
- Administrativa;
- Armazém de matérias-primas;
- Produção;
- Expedição;
- Distribuição/vendas.

[1] Hazard Analysis Critical Control Points.

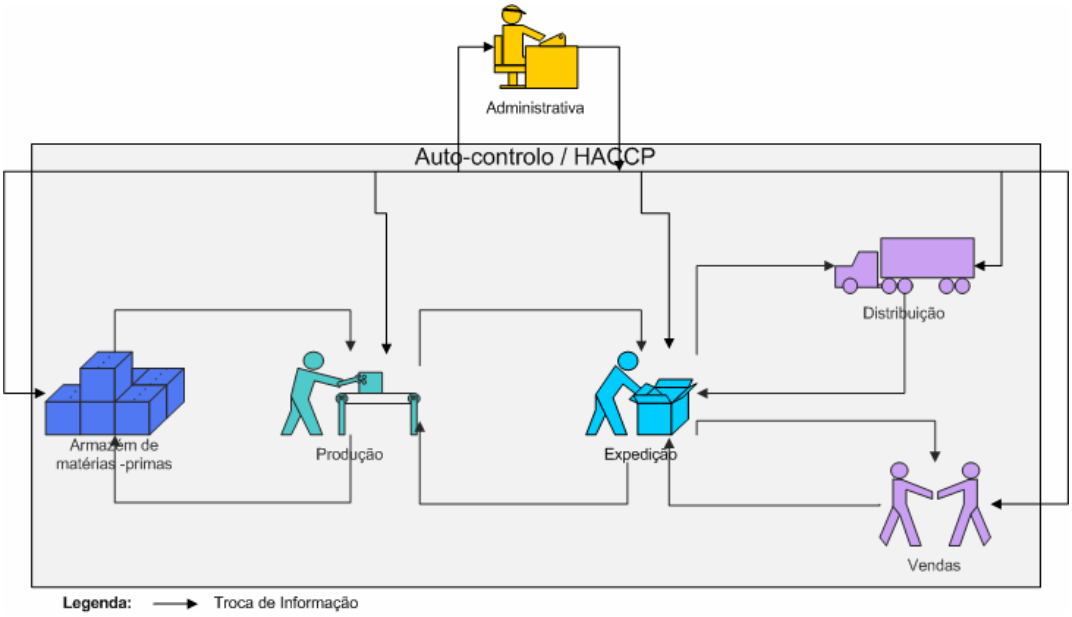


Figura 1 – Organização interna das empresas da Indústria da Panificação e Pastelaria

HACCP

O HACCP é um sistema reconhecido e internacionalmente utilizado pelas indústrias do sector alimentar. A sua implementação previne/minimiza os riscos alimentares, através da eliminação ou redução da probabilidade de ocorrência de uma eventual toxinfecção alimentar. O sistema é aplicável a todas as fases da produção, transformação e distribuição [DQA 2006].

As necessidades de informação para o sistema HACCP são as seguintes:

- Fluxograma de esquema de processos de fabrico;
- Registo de temperaturas;
- Registo do estado de conservação das instalações;
- Planos de higiene e manutenção de equipamentos;
- Fichas técnicas de produtos acabados.

Administrativa

Na área administrativa as necessidades de informação identificadas são as seguintes:

- Informação de avaliação de fornecedores;
- Informação de classificação de clientes;
- Documentos administrativos;
- Informação sobre frotas;
- Informação sobre rotas;
- Informação sobre recursos humanos.

Armazém de matérias-primas

O armazém de matérias-primas é o local onde se recebem e armazenam as matérias-primas. Aqui a informação necessária está relacionada com as operações sobre as matérias-primas, como, por exemplo, a entrada e a saída de stocks:

- Registo de entradas em stock e classificação por lotes das matérias-primas;
- Registo de saídas das matérias-primas;
- Registo de devolução/rejeição de matérias-primas;
- Registo de operações de planos de higiene e manutenção de equipamentos.

Produção

A área de produção é o local onde se dá a transformação da matéria-prima em produto acabado. A principal informação a tratar está relacionada com as operações de transformação de matéria-prima em produto acabado:

- Mapa de fabrico diário e registo de entradas de matérias-primas na produção;
- Registo dos testes aos óleos de fritura;
- Registo de operações de planos de higiene e manutenção de equipamentos.

Expedição

É nesta área que se processa a saída do produto acabado. Aqui a informação a tratar é diversificada. Nesta área recebe-se o produto acabado e classifica-se por lotes, recebem-se as encomendas e processam-se as mesmas, etc. A informação necessária a esta área é a seguinte:

- Notas de encomenda/devoluções;
- Classificação por lotes da produção;
- Guias de transporte;
- Registo de operações de planos de higiene e manutenção de equipamentos.

Distribuição/Vendas directas

A distribuição/vendas directas representam a última área da organização. É a área que tem um contacto directo com o cliente final, podendo ser cliente final o consumidor (vendas directas) ou um revendedor (distribuição). Como não se verificam diferenças de informação entre as duas áreas, estas são aqui descritas conjuntamente. A informação a tratar está essencialmente relacionada com as encomendas e as vendas:

- Notas de encomenda/devoluções;
- Facturas, vendas a dinheiro, notas de devolução, contas correntes, talões;
- Registo de operações de planos de higiene e manutenção de equipamentos.

4. Soluções tecnológicas

Após a caracterização das necessidades de informação das empresas da indústria da panificação e pastelaria procuraram-se diferentes soluções tecnológicas existentes no mercado e identificaram-se as respectivas funcionalidades. Das soluções encontradas, foram seleccionadas quatro soluções para o sector, três desenvolvidas por *software houses* nacionais e uma desenvolvida por uma *software house* internacional.

As soluções identificadas resultaram de pesquisas nas publicações do sector, pesquisas na *WWW* e contactos com agentes da indústria. Da primeira pesquisa (publicações do sector) resultou o sistema1[1] e o sistema2. De seguida, recorrendo a motores de busca *WWW* e contactando especialistas da indústria, foram procuradas mais soluções. Existem várias soluções genéricas que se podem adaptar ao sector, mas pelo facto de serem soluções genéricas e não suportarem algumas áreas específicas, não se tornaram objecto de estudo, com a excepção do sistema3. Este é um sistema que se adequa facilmente à indústria da panificação e pastelaria, visto cumprir muitos dos requisitos exigidos pelo sector como, por exemplo, a rastreabilidade, gestão de lotes, etc. No mercado internacional encontramos o sistema4.

Cada um destes sistemas foi estudado de forma a se identificarem as respectivas funcionalidades e suporte de informação. Na tabela 1, na primeira coluna, são descritas as necessidades de informação organizadas por áreas e, nas colunas seguintes, é classificado o seu suporte por parte das soluções existentes segundo uma escala discreta: suporta (S), suporta parcialmente (SP) e não suporta (NS). No primeiro caso, "suporta (S)", a solução estudada cobre de forma satisfatória um determinado requisito de informação. No segundo caso, "suporta parcialmente (SP)", a solução estudada cobre apenas parcialmente o requisito. E, finalmente, no caso "não suporta (SP)", a solução estudada não cobre o requisito.

[1] Por questões de sigilo e de concorrência não apresentamos o nome das empresas que desenvolveram os sistemas de software, sendo estes identificados como sistema1, sistema2, sistema3 e sistema4.

	Sistema1	Sistema2	Sistema3	Sistema4
HCCP				
Fluxograma de esquema de processos de fabrico	**NS**	**NS**	**NS**	**NS**
Registo de temperaturas	**NS**	**NS**	**NS**	**NS**
Registo do estado de conservação das instalações	**NS**	**NS**	**NS**	**NS**
Planos de higiene e manutenção de equipamentos	**NS**	**NS**	**NS**	**NS**
Fichas técnicas de produtos acabados	**S**	**S**	**S**	**S**
Administrativa				
Informação de avaliação de fornecedores	**S**	**NS**	**S**	**S**
Informação de classificação de clientes	**S**	**NS**	**S**	**S**
Documentos administrativos	**NS**	**S**	**S**	**S**
Informação sobre frotas	**NS**	**S**	**NS**	**NS**
Informação sobre rotas	**NS**	**NS**	**NS**	**NS**
Informação sobre recursos humanos	**S**	**NS**	**NS**	**NS**
Armazém de matérias-primas				
Registo de entradas em stock e classificação por lotes das matérias-primas	**S**	**S**	**S**	**S**
Registo de saídas das matérias-primas;	**S**	**S**	**S**	**S**
Registo de devolução/rejeição de matérias-primas	**S**	**S**	**S**	**S**
Registo de operações de planos de higiene	**NS**	**NS**	**NS**	**NS**
Produção				
Mapa de fabrico diário e registo de entradas de matérias-primas na produção	**S**	**SP**	**S**	**S**
Registo dos testes aos óleos de fritura	**NS**	**NS**	**NS**	**NS**
Registo de operações de planos de higiene e manutenção de equipamentos	**NS**	**NS**	**NS**	**NS**
Expedição				
Notas de encomenda/devoluções	**S**	**S**	**S**	**S**
Classificação por lotes da produção	**S**	**NS**	**S**	**S**
Guias de transporte	**NS**	**S**	**S**	**S**
Registo de operações de planos de higiene e manutenção de equipamentos	**NS**	**NS**	**NS**	**NS**
Distribuição / Vendas directas				
Notas de encomenda/devoluções	**S**	**S**	**S**	**S**
Facturas, vendas a dinheiro, notas de devolução, contas correntes, talões	**NS**	**NS**	**S**	**S**
Registo de operações de planos de higiene e manutenção de equipamentos	**NS**	**NS**	**NS**	**NS**

***Legenda*: NS** – Não suporta / **S** – Suporta / **SP** – Suporta parcialmente

Tabela 1 - Necessidades de informação das empresas da indústria da panificação e pastelaria vs soluções de TI existentes no mercado

Na tabela 1 verifica-se que nenhuma das soluções estudadas satisfaz integralmente as diferentes necessidades de informação identificadas para as empresas da indústria da panificação e pastelaria. As soluções que apresentam um melhor suporte das 25 necessidades de informação identificadas, são o sistema3 e o sistema4, no entanto suportando apenas 52% dessas mesmas necessidades. Dez das necessidades não são suportadas por nenhuma das diferentes soluções e a área que apresenta um suporte mais deficiente é a área HCCP (apenas 20% das necessidades são suportadas). Por outro lado, a área que tem um melhor suporte é a área administrativa, com cerca de 83% das necessidades cobertas. De notar a importância de algumas das necessidades a que nenhuma das soluções dá resposta, como é o caso, por exemplo, dos registos de temperaturas e registo de operações de planos de higiene, dado que são essenciais para as empresas em termos do cumprimento de requisitos legais pois fazem parte do HACCP.

5. Conclusão

A indústria da panificação e pastelaria revela hoje uma grande exigência em termos da necessidade de cumprir requisitos legais, de qualidade, de certificação, de conformidade, e muito particularmente de controlo, o que torna a gestão sem suporte de TI, que se verifica ainda hoje num grande número de organizações, verdadeiramente ineficiente. Este é um problema que traz profundas implicações em termos da eficácia e produtividade das empresas e, num sentido mais amplo, a nível da competitividade da própria indústria.

Procurando contribuir para a melhoria dos níveis de utilização de TI nesta indústria, desenvolvemos um estudo em que se procuraram respostas para as questões "quais são as necessidades de informação actuais das empresas da indústria da panificação e pastelaria?" e "as soluções de TI actualmente disponíveis no mercado suportam devidamente as necessidades de informação das empresas da indústria da panificação e pastelaria?".

O estudo realizado permitiu identificar diversas necessidades de informação que apresentamos neste artigo e concluir que as soluções que foram estudadas não suportam devidamente essas necessidades. Na melhor das situações, suportam apenas 50% do total das necessidades identificadas. Inclusivamente verificou-se que muitas das necessidades não são suportadas por nenhuma das soluções estudadas.

As TI são o alicerce da organização contemporânea. A cada dia que passa torna-se mais claro que sem uma utilização eficiente e eficaz das TI, as organizações não podem ser competitivas ou rentáveis e que, em muitos casos, é a sua própria sobrevivência que depende dessa capacidade [Varajão 2002].

Identifica-se assim uma oportunidade de proposta de uma nova arquitectura de informação que sirva de base para o desenvolvimento de sistemas que suportem as necessidades de informação das empresas deste sector de actividade, de forma completa e integrada.

6. Referências

DQA, *O que é o HACCP?*, http://www.dqa.pt/002.aspx?dqa=0:0:0:22:0:0:-1&ct=29, (10 de Outubro de 2006), 2006.

Ferrete, A., *A importância da Informação em Gestão na Panificação*, *a Padaria Portuguesa*, 42, 2004.

IIF, Instituto para a Inovação e Formação, *Estudo de diagnóstico e prospectiva e caracterização de perfis profissionais para levantamento das necessidades de formação no sector agro-alimentar, I e II* , 2001, 5- 36.

INE, *Classificação das Actividades Económicas - Rev. 2.1,* www.ine.pt/Prodserv/nomenclaturas/cae_rev2-1.asp, (15 de Outubro 2005), sd.

INE, *Estatísticas Agro-industriais 1999-2001*, http://www.ine.pt, (25 de Outubro 2005), 2003a.

INE, *Grupo 158 – Fabricação de Outros produtos alimentares*, http://www.ine.pt, (25 de Outubro 2005), 2003b.

Varajão, J., *Função de Sistemas de Informação - Contributos para a melhoria do sucesso da adopção de tecnologias de informação e desenvolvimento de sistemas de informação nas organizações*, Tese de Doutoramento, Universidade do Minho, 2002.

VVSHOP - O Directório de Retalhistas Electrónicos da Região dos Vinhos Verdes

José Luís Reis [1]
Comissão de Viticultura da Região dos Vinhos Verdes, Porto, Portugal
jlreis@cvrvv.pt

Miguel Calejo[2]
Declarativa, Porto Portugal
mc@declarativa.pt

Paulo Martins
Comissão de Viticultura da Região dos Vinhos Verdes, Porto, Portugal
pmartins@cvrvv.pt

Mário Araújo
Declarativa, Porto, Portugal
mfa@declarativa.pt

Resumo

A VVSHOP é um Sistema de Informação que organiza num sistema unificado os diferentes produtos com Denominação de Origem provenientes da Região dos Vinhos Verdes, permitindo associar aos produtos os seus retalhistas electrónicos, e-tailers, localizados em diferentes países, num directório disponível através da Internet, com diferentes formas de interface, nomeadamente através dos mapas disponibilizados pela Google (googlemaps). Este sistema está disponível no Portal da Região dos Vinhos Verdes e está integrado nos processos de gestão da informação dos produtores e das respectivas marcas, associando-as a outras componentes existentes no sistema de backoffice da CVRVV, que integra as aplicações que facilitam a interacção entre os produtores, os produtos, a entidade que os certifica e os retalhistas electrónicos desses mesmos produtos em todo o mundo.

Palavras chave: Retalhistas electrónicos (e-tailers), Comércio Electrónico, Portal, Produtores, Produtos, Marcas, Bases da Dados, Verde Global, Customização, WAM (Web Application Maker), Interfaces Públicos e Privados, e-marketplaces, GoogleMaps.

[1] Docente no IPAM e no ISMAI. Director de SI/TIC da CVRVV
[2] Professor Associado Convidado da UM e Director da Declarativa.

1. Introdução

O Vinho não é um produto "globalizado": tem origem numa determinada Região, com determinadas características que se reflectem na identidade dos produtos, conferindo-lhes carácter e uma história; esta conexão é fundamental para diferenciar os vinhos portugueses dos outros vinhos produzidos em todo o mundo. O produtos provenientes da Região dos Vinhos Verdes, com denominação de Origem Vinho Verde e Regional Minho são vendidos em mais de 100 países e a Comissão de Viticultura da Região dos Vinhos Verdes (CVRVV) é a entidade interprofissional responsável pela certificação, controlo de qualidade e promoção, bem como orientação técnica aos produtores e entidades comerciais. Para além das acções relacionadas com a publicidade, promoção e relações públicas em todo o mundo, a Comissão, através do seu Departamento de Sistemas de Informação, opera um portal disponível numa plataforma WWW (Internet, Intranet e Extranet), que monitoriza os principais aspectos associados à produção de vinho, proveniente das cerca de 300.000 parcelas de vinha, propriedade dos 32.000 produtores, comercializado sob diferentes marcas, por cerca de 550 entidades económicas.

2. O Universo de retalhistas electrónicos (e-tailers) do Vinho Verde

O negócio electrónico está dividido em 3 tipos: B2C: Business to Consumer (relações de negócio electrónico com o consumidor final), B2B: Business to Business (relações de negócio electrónico com empresas) e C2C Consumer to Consumer (relações de negócio electrónico entre consumidores finais). Quando a Internet atingiu a sua maturidade e as organizações passaram e investir muito mais em plataformas associadas ao comércio e ao negócio electrónico. Após uma fase em que o B2C foi alvo de maior enfoque, percebeu-se que os ganhos estão maioritariamente nas relações B2B, nomeadamente devido a economias no processo de aquisição / fornecimento, melhoria de relação com parceiros de negócio, economias no preço, acesso a novas oportunidades de negócio. No entanto as apostas no B2C continuam a ser efectuadas pelos diferentes sectores de actividade e o sector dos vinhos também tem feito as suas apostas. Um dos primeiros projectos consolidados de comércio electrónico B2C foi precisamente no sector dos vinhos, denominado VirtualVineyard.com. Em 1990 a VirtualVineyard.com adquiriu a wine.com, sendo o nome wine.com perfeito para uma empresa que pretende ter um site sobre vinho na Internet. A wine.com transformou-se na primeira empresa que integrou num portal Web, conteúdos, comércio e uma comunidade num único site sobre vinhos. A evolução e as apostas efectuadas por diferentes tipos de organizações do sector vitivinícola no comércio electrónico permitem que os vinhos sejam comercializados na Internet em 3 tipos de plataformas (Strauss, 2003) (EGP, 2003):

- eMarketplaces de produto: dedicados só ao vinho (ex.: www.beverageonline.com)
- eMarketplaces verticais: dedicados ao sector agro-alimentar (ex.: www.tfir.com)
- eMarketplaces horizontais: dedicados a vários tipos de produtos (ex.:www.furniturexnet.com).

O retalhista electrónico e-tailer, define-se como um retalhista de um determinado produto que usa o canal Internet para vender, tornado assim mais eficiente a forma como o cliente pode adquirir esse produto.

No decorrer de um novo projecto denominado Verde Global – O Sistema de Informação Global da Região dos Vinhos Verdes, foi promovido pela CVRVV e desenvolvido pelo Centro de Estudos da Escola de Gestão do Porto (EGP) da Universidade do Porto um estudo denominado "Concepção de um Modelo Logístico para a Distribuição via Internet de Vinhos Verdes nos mercados EUA e Alemanha". No decorrer do ano de 2005, foram efectuados levantamentos de informação que permitiram concluir que em 14.10.2005, para as 73 marcas que já se tinham presença na VVSHOP existiam já retalhistas electrónicos em 13 Países de 3 Continentes diferentes – ver Quadro 1, com especial ênfase para o continente Europeu, posteriormente, em 2006 verifica-se que existem mais marcas com e-tailers na América do Norte e América do Sul (Brasil). Indo ao encontro dos resultados deste estudo, a CVRVV orientou e desenvolveu durante o projecto a sua plataforma electrónica base (VVSHOP) ao nível do B2C para um conceito de e-tailers de Vinhos Verdes, e especificamente com este projecto, pretendia-se também potenciar as exportações quer nos EUA quer na Alemanha dos AE, através da disponibilização da plataforma electrónica para o apoio à distribuição nesses países.

	2005	**2006**
Total de Marcas registadas na VVSHOP	**73**	**125**
Total Marcas com retalhistas electrónicos – e-tailers	**36**	**90**
Total Marcas sem retalhistas electrónicos – e-tailers	**37**	**35**
País com retalhistas electrónicos que vendem Vinho Verde	**Marcas**	**Marcas**
Alemanha	12	31
Áustria	1	1
Brasil	3	11
Canadá	2	2
Espanha	1	2
Estados Unidos	17	35
França	2	3
Holanda		1
Noruega	1	3
Polónia	2	2
Portugal	16	18
Reino Unido	7	9
Republica Checa	2	2
Suécia	1	2

Quadro 1 – Países e e-tailers com produtos provenientes da Região Demarcada dos Vinhos Verdes.

Na data de 12.09.2006 a Região tinha mais de 1000 marcas activas, de cerca de 550 AE distintos. Na VVSHOP estão representadas 125 marcas de 54 AE com e-tailers em 14 Países – Quadro 1. Como veremos posteriormente, uma das ferramentas disponíveis no backoffice VVSHOP baseia-se num agente de software que ajuda a descobrir para os e-tailers as respectivas páginas de produtos relevantes (vinho verde), bem como a detectar ao longo do tempo quais as páginas (URLs) que ficam obsoletas (URLs quebrados).

3. A arquitectura de suporte à VVSHOP

A arquitectura da VVSHOP está integrada com o sistema de informação que suporta o portal da Região, nas suas componentes Intranet, Extranet e Intranet, que se baseia num conjunto de servidores integrados que suportam o Sistema de Gestão de Bases de Dados (SQL Server), onde está armazenada toda a informação técnica, como por exemplo: - produtores, marcas e componentes de rotulagem, produtos comercializados pelos produtores, dados dos visitantes, estatísticas, etc.

A VVSHOP providencia a apresentação de um directório em formato electrónico com os respectivos produtores, marcas e produtos, associando-os aos respectivos e-tailers que vendem os produtos através de um link, o que permite a ligação directa à respectiva página web na loja electrónica que tem o produto disponível para ser encomendado. Para além de permitir a apresentação dos produtores, a VVSHOP faz a apresentação das respectivas marcas dos produtos disponíveis, permitindo que os visitantes, seleccionem a partir do seu Pais de origem os e-tailers que vendem os produtos pretendidos.

O sistema inclui ainda dois assistentes de apoio à manutenção da Base de Dados dos e-tailers:

- Um agente Java que verifica periodicamente todos os endereços (URLs) já conhecidos, para garantir que não tenham ficado obsoletos (quebrados).

- Outro agente Java, um "Robot" de pesquisa, que em função das directivas recebidas do gestor da VVSHOP efectua pesquisas automáticas nocturnas (para não sobrecarregar a rede CVRVV), acedendo aos dois maiores motores de busca generalistas através das suas APIs: Google e Yahoo. Na manhã seguinte o gestor tem no backoffice um conjunto de propostas de páginas de produto e retalhistas, que valida com um click, após o que passam a incorporar o directório de e-tailers.

As aplicações de backoffice suportam as novas funcionalidades e possibilitam também a gestão das tabelas SQL Server (por exemplo produtores, marcas, produtos, e-tailers) e foram

desenvolvidas com base no "WAM", o gerador de aplicações Intranet, construído pela Declarativa para SQL Server, ver http://www.declarativa.com/wam.

O VVSHOP assenta assim num sistema gestor de bases de dados relacionais, onde armazena a maior parte da informação textual e onde mantém as relações necessárias à consistência dos dados disponibilizados.

4. A informação no Sistema de Gestão da VVSHOP

Foram vários os projectos implementados nos últimos anos na CVRVV, para permitir que o Sistema de Informação de Gestão tenha um controlo absoluto sobre todos os processos associados à certificação das Denominações de Origem provenientes da Região Demarcada dos Vinhos Verdes. Um dos projectos desenvolvidos, denominado "Rotulagem e Marcas no Portal", tem como principal objectivo efectuar a optimização do registo das diferentes componentes associadas à rotulagem para se criar uma *visão colectiva das marcas nacionais e regionais*, com uma imagem forte, que permita evidenciar um produto de características vinícolas únicas com base em castas que são cultivadas na Região dos Vinhos Verdes (Reis et al, 2005).

4.1. Interface privado (Intranet)

Na fase de análise das necessidades de informação (requisitos de sistema), para a definição do modelo de negócio, para posterior desenvolvimento do modelo de dados e construção do sistema de suporte à VVSHOP, foram definidos os diferentes objectos de negócio (produtores, marcas, produtos, tipos de produtos, rótulos, países, e-tailers,…). Estabeleceram-se as suas relações nos diferentes processos e actividades associados aos processos de registo de informação no sistema de backoffice que suportam as diferentes aplicações existentes na CVRVV. Este sistema denominado **VERDE GLOBAL: Gestor de Backoffice v2.1** está disponível na Intranet da CVRVV. Na Figura 2, podemos ver algumas das opções do ecrã principal da aplicação de gestão do backoffice da VVSHOP na sua versão 2.1.

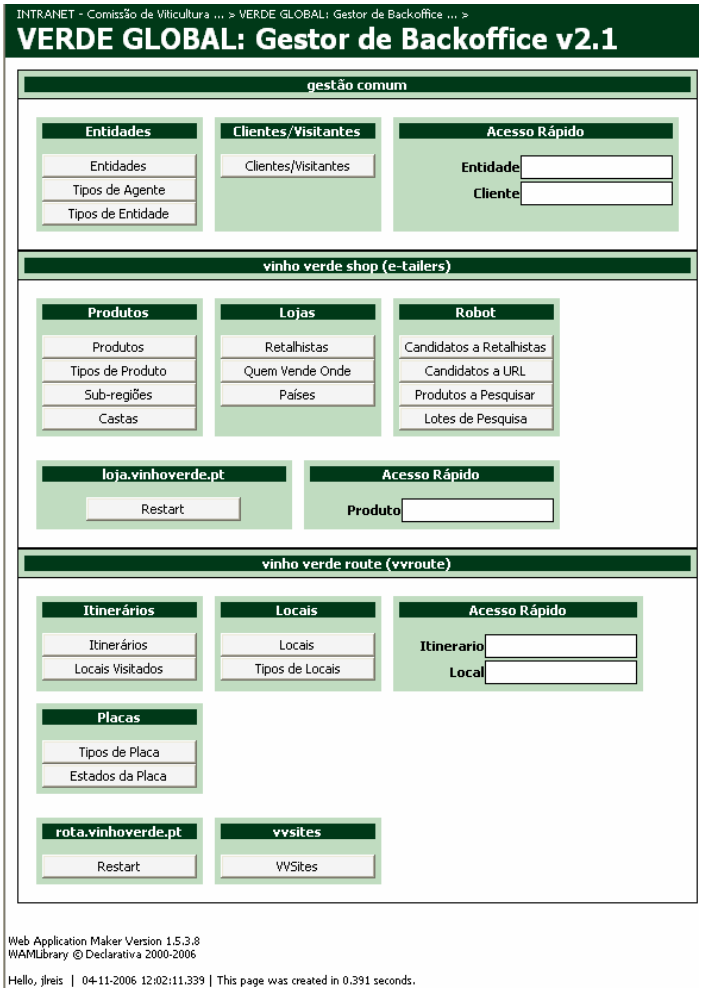


Figura 2 – Componentes do sistema de Back-office do VVSHOP

O Gestor de Backoffice permite o acesso aos produtos, tipos de produtos, sub-regiões (são 9 as sub-regiões na região dos Vinhos Verdes), castas, retalhistas (e-tailers), países e também o acesso do gestor, nas opções "Robot", aos endereços URL encontrados pelo agente de software, para poderem ser validados.

O processo de descoberta automática de e-Tailers começa pela especificação de um "lote de pesquisa", em que o gestor indica manualmente, ou através de um critério de pesquisa, quais os produtos a pesquisar (ver Figura 3):

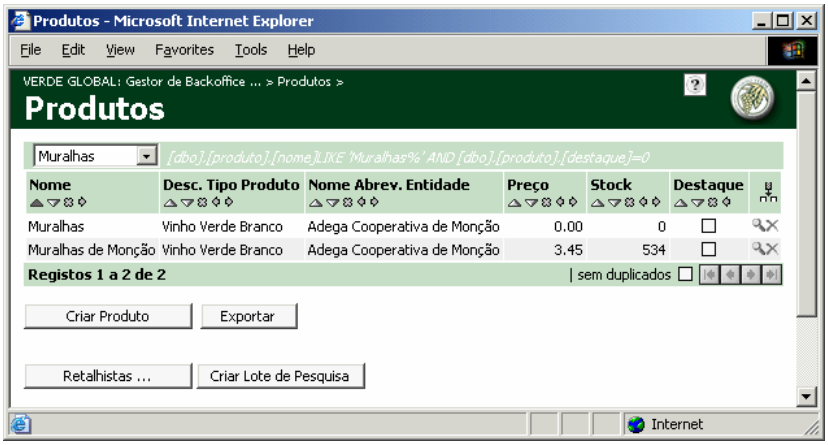


Figura 3 – Produtos a pesquisar na VVSHOP

Após premir o botão "Criar o lote de pesquisa" obtém-se a especificação de um conjunto de pesquisas, opcionalmente restrito a um domínio de topo Internet (no exemplo, .com), e indicando-se o motor de busca a usar , Google ou Yahoo, (ver figura 4) :

Figura 4 – Lote de pesquisa na VVSHOP

Na manhã seguinte o gestor dispõe de uma lista de candidatos a página de (venda de) produto (ver Figura 5):

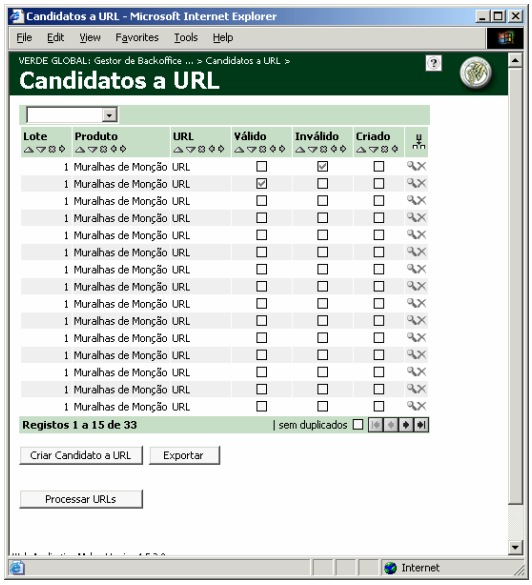


Figura 5 – Lista de candidatos a e-tailers na VVSHOP

Clickando na coluna URL o browser abre a página (supostamente) do e-tailer. No exemplo acima, o primeiro endereço encontrado é inválido porque é uma mera página com notas de prova (abaixo à esquerda), mas o segundo é válido pois permite a compra do produto em causa (ver Figura 6):

Figura 6 – Página de um e-tailer na VVSHOP

O sistema descobre automaticamente os e-Tailers a partir das suas páginas, inferindo o país a partir do endereço IP do site (ver Figura 7):

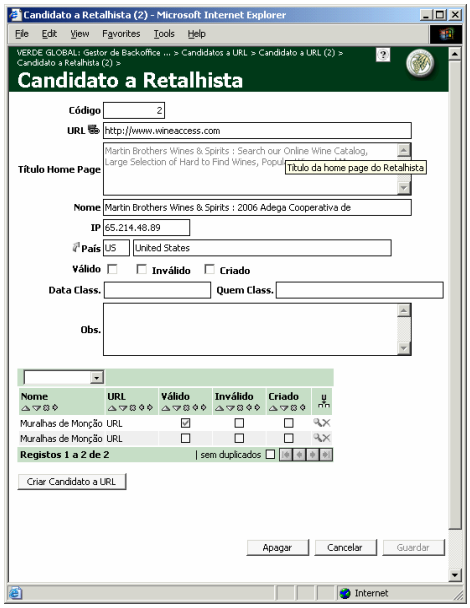


Figura 7 – Página final de um e-tailer no gestor da VVSHOP

Após classificar os URLs, premindo o botão "Processar URLs" na "Candidatos a URL" os válidos transitam automaticamente para o directório de e-Tailers. Os inválidos são retidos para serem descartados em futuras pesquisas.

4.2. Interface privado (Extranet)

Os Agentes Económicos têm acesso através da Extranet da CVRVV, denominada INETSIV, a um conjunto de opções que lhes permitem obter informação em tempo real, nomeadamente: o cadastro das suas vinhas, os seus dados de colheita e produção, os saldos das contas correntes dos seus produtos, os resultados dos relatórios das análises laboratoriais (permitem obter a certificação do produto), a visualização dos processos de registo das marcas e das componentes que podem constituir uma ou mais vestimentas que vão “vestir” a garrafa, etc.

Para além destas opções, foi criada a opção “VVSHOP Produtos” que permite ao AE visualizar e actualizar as informações sobre as características dos produtos que estão disponíveis na VVSHOP, nomeadamente a descrição, sabor, aspecto visual - cor, temperatura a que deve ser servido, ano de colheita e castas. O AE não tem acesso aos endereços URL de lojas electrónicas que os AE tenham em todo o mundo associados a cada produto, pois essa actualização é somente da responsabilidade do gestor da VVSHOP. Ver na figura 8 alguns ecrãs de acesso da Extranet que permitem aos AE’s alterar em Português e Inglês as características comerciais dos produtos.

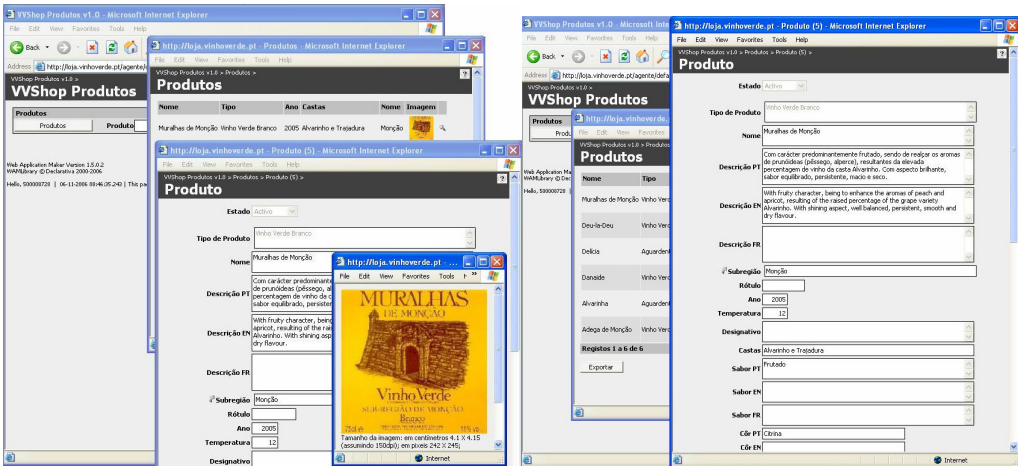


Figura 8 – Componentes Extranet da VVSHOP.

5. O Interface Público da VVSHOP

Os objectivos são claros; pretende-se associar a protecção de uma denominação de origem em questões ambientais, culturais, turísticas associadas ao produto vinho. Esta visão global do produto consegue-se integrando todos os aspectos relacionados com a certificação dos produtos num sistema unificado, daí o projecto Verde Global. Como vimos anteriormente, o sistema da VVSHOP está suportado num sistema de backoffice que contém um conjunto de elementos de informação que são transparentes para os utilizadores que dominam a gestão de todas as actividades associadas aos processos de certificação. Este sistema é customizado para o operador económico através da Extranet, bem como para o utilizador que gere as actividades de gestão através da Intranet.

Figura 9 – Interface de acesso à rotulagem de uma marca no Portal

No portal, está disponível a informação mais relevante, para o público em geral, consumidor ou não, em versão Portuguesa e Inglesa. É permitido o acesso aos dados da marcas disponíveis na VVSHOP, ou através da Home Page do portal na opção VVSHOP (Vinho Verde SHOP) ou através da pesquisa da informação sobre as marcas/rotulagem da Região, como podemos ver na Figura 9, é dada a possibilidade de acesso à VVSHOP se o visitante seleccionar a ligação "ONDE COMPRAR".

A partir do memento em que o visitante entra numa página de um produto, surgem as informações que caracterizam o produto (nome, produtor, descrição, cor, temperatura a que deve ser bebido, ano de colheita, castas). Se o visitante da página pretende saber mais sobre o produtor, pode fazer clique sobre o nome do produtor e surgem-lhe informações detalhadas sobre a empresa.

O sistema determina a localização, através do IP do visitante, o Pais onde este se encontra no momento em que entrou na página e surge no ecrã por baixo da mensagem "Onde Comprar", os endereços URL dos e-tailers que vendem o produto nesse País. Caso o visitante pretenda encontrar os e-tailers de outro País, pode seleccioná-lo no menu Popup que surge por baixo da mensagem "O seu País". Na Figura 10, encontra-se uma sequência de ecrãs que permitem visualizar a informação pública da VVSHOP.

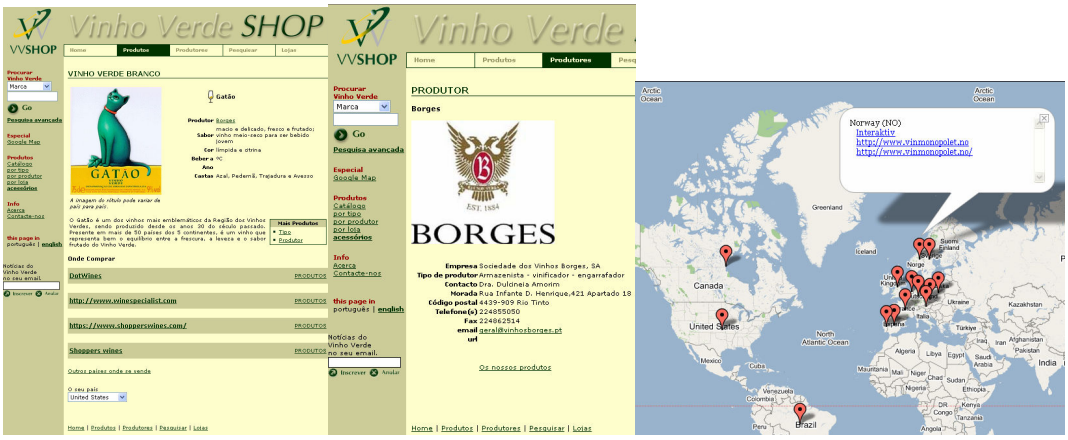


Figura 10 - Algumas componentes da informação pública da VVSHOP.

A partir do momento em que o visitante efectua clique num dos links de um dos e-tailers que surgem na lista associada a cada produto, o sistema da VVSHOP limita-se a reencaminhar para a página da empresa que vende o produto, surgindo de imediato a informação sobre esse produto nas páginas do respectivo e-tailer.

Como referimos anteriormente existe uma versão Portuguesa e Inglesa da VVSHOP, podendo ser efectuada a mudança de língua da página em qualquer momento da sua visualização. É dada

ainda a possibilidade aos visitantes receberem notícias em formato electrónico da Região dos Vinhos Verdes, ao escreverem e enviarem o seu endereço mail na caixa que surge em baixo da mensagem "Notícias do Vinho Verde no seu email".

6. Impacto do sistema e trabalhos futuros

A VVSHOP evolui ao longo do tempo, partindo de uma plataforma de comércio electrónico que envolvia o tratamento das compras, para o desenvolvimento de um directório de e-tailers. Os recursos envolvidos nas duas fases da VVSHOP são quantificados de uma forma global e envolveram a participação de um conjunto de especialistas em várias áreas; sistemas de informação (analistas, administradores de base de dados, programadores, técnicos de comunicações, operadores, …), designers, gestão de marketing, juristas, técnicos de viticultura e enologia. Para além do desenvolvimento específico da solução VVSHOP é preciso ter em consideração que muito do esforço de trabalho no sistema da VVSHOP é transversal a outros sistemas de existentes na CVRVV. O projecto VVSHOP foi desenvolvido desde 2001 até 2006 e contou com o apoio de fundos estruturais no âmbito do ON, em projectos como o e-Verde e o Verde Global, representando um esforço em termos de quantidade de trabalho (2 homens/dia) descrito e quantificado no Quadro 2.

Acções	**Dias (2 homens/dia)**
Elaboração, coordenação, desenvolvimento e acompanhamento de projecto	20
Levantamento de especificações	15
Desenvolvimento do sistema	120
Integração na arquitectura do portal	5
Design gráfico e sua integração no portal	5
Reforço da estrutura tecnológica	2
Carregamento e migração de dados para a VVSHOP	15
Acompanhamento técnico nos testes e colocação em produção	5
Formação de utilizadores e de operadores económicos	3
Total dias (2 homens trabalho/dia):	190

Quadro 2 – Acções e esforço de trabalho no Sistema VVSHOP.

A VVSHOP é um dos elementos fundamentais de um conjunto de campanhas publicitárias que decorrem na Internet, tendo como principais alvos os mercados dos EUA e da Alemanha, onde se pretende também alargar a distribuição dos Vinhos Verdes com um novo modelo logístico e de distribuição. O grupo do Centro de Estudos da EGP, fez uma selecção dos parceiros/agentes ideais e no futuro, espera-se que os Vinhos Verdes possam estar disponíveis de uma outra forma nestes mercados, procurando a máxima eficiência em toda a operação logística. Decorreram também campanhas publicitárias online em sites portugueses durante o mês de Setembro de 2006, convidando à compra na VVSHOP. Um dos principais problemas que se pretendem resolver, prende-se com a necessidade de um maior envolvimento dos AE em todos os processos relacionados com a actualização da informação imediata das características dos seus produtos na VVSHOP, bem como aumentar o número de marcas e respectivos produtos representados na VVSHOP.

Quando se iniciou o trabalho de desenvolvimento da VVSHOP com o conceito e-tailers, estabeleceram-se um conjunto de prioridades que passavam por efectuar uma análise cuidada dos dados relacionados com as visitas efectuadas, bem como o impacto que este projecto pode ter nas vendas dos retalhistas electrónicos. Em 2007, prevê-se a possibilidade de criação de um grupo de trabalho que possa extrair informação sobre os perfis dos visitantes à VVSHOP, bem como implementar um sistema que automaticamente disponibilize ao marketing da CVRVV, informação relevante para definir as acções a desenvolver em campanhas de e-marketing.

7. Conclusões

A equipa técnica que efectua a coordenação do projecto VVSHOP, está satisfeita com o caminho seguido na evolução da VVSHOP, e considera que o sistema baseado nos retalhistas electrónicos é uma solução abrangente e global, pois não está limitada nem em termos de numero de marcas e produtos representados na VVSHOP, nem limita de forma alguma a entrada de qualquer retalhista electrónico, independentemente da sua localização geográfica. É no entanto de salientar que sem o envolvimento desde da primeira hora do Departamento de Marketing da CVRVV, será impossível avançar com sucesso que se pretende. O envolvimento efectuado até agora revelou-se preponderante nas diferentes fases do projecto VVSHOP, principalmente no relacionamento com os AE, na manutenção da informação actualizada sobre as marcas, os produtos associados e respectivos e-tailers, bem como na preparação e acompanhamento das campanhas publicitárias à VVSHOP. Ao nível interno foi fácil a familiarização do Gestor da VVSHOP com o sistema, no entanto a gestão do projecto está consciente do enorme esforço que é necessário fazer para melhorar a relação entre os AE, muitos deles ainda sem capacidades informáticas, nem sensibilidade para a utilização deste tipo

de ferramentas. Por outro lado o automatismo em algumas funcionalidades da VVSHOP, bem como acções de promoção da VVSHOP, atendendo às características inovadoras do projecto, obrigam a um esforço que o grupo de trabalho está a realizar, com o objectivo de melhorar a quer a comunicação quer os serviços prestados, numa perspectiva de ir ao encontro das necessidades dos diferentes tipos de utilizadores envolvidos no sistema (produtores, técnicos de informática, gestor e visitantes). Apesar dos produtos provenientes da Região dos Vinhos Verdes serem de proveniência Regional são produtos Globais, pois estão presentes em muitos Países do Mundo. A tendência é de crescimento do B2C através da Internet, e com toda a certeza teremos cada vez mais e-tailers a vender os produtos o que significa que a VVSHOP terá também cada vez mais marcas e correspondentes e-tailers registados.

8. Agradecimentos

Salientamos o apoio das organizações Antero Ferreira Design, do Centro de Estudos da Escola de Gestão do Porto, na análise e na concepção das metodologias associadas ao desenho, construção e implementação da VVSHOP. No levantamento de informações, desenho dos modelos de dados, na programação do sistema, naquelas e noutras entidades ou a título individual, de Marta Valente, Alcibiedes Guedes, Eduardo Santos, António Fernandes e Jorge Gomes. O projecto VVSHOP teve ainda o apoio do programa ON – Operação Norte Desenvolvimento Regional Norte, no âmbito da medida 1.4. de Valorização e Promoção Regional e Local.

9. Referências Bibliográficas

Calejo, M e Araújo, M e Moreira, M., “Arquitectura do portal e-Verde”, 2001

CVRVV – Candidatura ON “e-Verde – O Portal da Região dos Vinhos Verdes”, 2000

CVRVV – Candidatura ON “Verde Global – O sistema de informação da Região dos Vinhos Verdes”, 2003

CVRVV, Lançamento da VVSHOP nos jardins da CVRVV, citado em 30/10/2006, disponível em http://www.vinhoverde.pt/pt/noticiasfrescas/lancamentoVVShop.htm

Declarativa, WAM - Web Application Maker , citado em 2009/10/30, disponível em http://www.declarativa.com/wam/

E-tailer from Wikipedia the free encyclopedia, citado em 2006/1031, disponível em

http://en.wikipedia.org/wiki/E-tailer

EGP- Centro de Estudos da Escola de Gestão do Porto, da Universidade do Porto, "Concepção de um Modelo Logístico para a Distribuição via Internet de Vinhos Verdes nos mercados EUA e Alemanha", 2006.

Laudon, J.P. e Laudon, K. "Management Information Systems", 9ª ed Prentice Hall, 2006

Lencastre, P. Brito, C.M. – Os Horizontes do Marketing. Verbo, 1º Edição, 2000

Picão, M. Calejo, M. e Mota, S., "Processos Administrativos na CVRVV", Maio de 1999

Porter, M. Estratégias de Marketing para Vinhos Portugueses na GB e EUA (PT), 2003

Reis, J.L., "Sistema de Informação Vitivinícola",1998

Reis, J.L., Cerdeira A.e Calejo, M. "The e-Verde portal - putting the Vinho Verde circuit on the web, from vineyards to consumers", Outubro de 2001

Reis, JL, M. Valente, Leite, P.P. As Marcas no Portal da Região dos Vinhos Verdes", CAPSI 2005

Reis. JL, Calejo, M. Valente, M; VVRoute - Sistema de Informação da Rota da Região do Vinho Verde, CAPSI 2003.

Sinton, P. VirtualVineyard Swallows Up Wine.com Palo Alto firm wanted Napa company's name, 1999 citado em 2006/11/03, disponível em http://www.sfgate.com/cgibin/article.cgi?file=/chronicle/archive/1999/09/14/BU39238.DTL

Strauss, J., Frost, R., *"E-Marketing",* Prentice Hall 3ª Ed, 2003

ViniPortugal – Resumo da Competitividade do Cluster e Introdução a Campanhas de Acção (PT), 2003

Wine.com, 2006, citado em 2006/10/30, disponível em http://www.wine.com/customercare/

Avaliação Remota do Desempenho da Empresa

Pedro José Guerra de Araújo[1], João Vieira Baptista[2]

1) IT – Instituto de Telecomunicações

Departamento de Informática, Universidade da Beira Interior

Av. Marquês d'Ávila e Bolama 1, 6200 Covilhã, Portugal

paraujo@di.ubi.pt

2) IESIG – Instituto de Estudos Superiores Isidoro da Graça

São Vicente, Cabo Verde

aaug_99@yahoo.com.br

Resumo

Qualquer que seja o sector de actividade, a crescente concorrência entre as empresas coloca desafios que estas têm de enfrentar e vencer para sobreviverem a longo prazo. Este problema que se observa à escala internacional, é particularmente importante para as empresas de Portugal e de Cabo Verde, dada a fragilidade dos respectivos tecidos empresariais. Afectadas por um elevado nível de desorganização, estas empresas precisam de melhorar o seu desempenho de modo a manterem-se competitivas. Porém, o nível de desorganização é um parâmetro difícil de avaliar e consequentemente de controlar. Neste trabalho apresenta-se uma proposta para avaliação do nível ou quantidade de desorganização de uma empresa, aplicado à indústria têxtil e do vestuário mas que pode também estender-se a outros sectores. Um sistema de informação de produção que utiliza este conceito está já em fase de testes e os resultados até agora obtidos apontam no sentido de confirmar a sua importância para a gestão da empresa. O sistema apresenta também a capacidade de comunicação de dados via SMS(Short Messages System) para os telemóveis dos empresários ou responsáveis da produção. Isto permite-lhes dispor de informação em tempo real sobre parâmetros de desempenho da empresa, onde quer que se encontrem.

Palavras-chave: gestão remota da produção; avaliação do desempenho; comunicações móveis

1. Avaliação do nível de (des)organização

As empresas precisam melhorar não só a sua eficácia, colocando no mercado os produtos que este procura, como também a sua eficiência, racionalizando os custos de produção. Essa racionalização passa por reduzir os desperdícios, quer em termos de tempo quer em termos de materiais, desperdícios esses que se observam num grande número de empresas, tanto do sector privado como em organismos públicos. Esta situação conduz ao elevado nível de desorganização que caracteriza o tecido empresarial de países como Portugal e Cabo Verde.

Entretanto, a desorganização é um conceito abstracto, difícil de mensurar. Esta dificuldade complica a sua utilização em diversos sentidos. Por um lado não é fácil comparar o nível de desorganização das várias empresas que leve a uma classificação de quais são mais desorganizadas do que outras. Por outro lado, não sendo a desorganização fácil de medir também não é fácil de controlar, nomeadamente pela dificuldade em determinar quais os factores que a afectam e de como actuar sobre eles.

Diversos autores (Mukhopadhyay *et al.*, 1995; Mahmood & Mann, 2000) referem a melhoria da eficiência organizacional como um dos factores em que os sistemas e tecnologias da informação (SI/TI) podem contribuir para a melhoria do desempenho das empresas, embora não apresentem um processo de avaliar essa contribuição. Ao longo deste trabalho apresenta-se uma proposta para avaliar o nível de desorganização referente à função produção em empresas industriais de mão-de-obra intensiva, como é o caso da indústria de confecção de vestuário. O primeiro objectivo é evidenciar perante os gestores que a desorganização existe, quantificando-a de modo a demonstrar as vantagens do uso das tecnologias da informação para a sua redução e consequente melhoria na produtividade. A resolução deste objectivo contribui de forma significativa para justificar o investimento em SI/TI, um assunto controverso objecto de intensa investigação (Strassmann, 1997; Brynjolfsson & Hitt, 2000; Serrano & Caldeira, 2001). O segundo objectivo é permitir aceder remotamente à informação que caracteriza o estado organizacional da empresa, mesmo quando os responsáveis estão ausentes das respectivas instalações. De facto, a importância do acesso à informação em tempo real é cada vez maior na sociedade actual e na indústria em particular, uma vez que permite responder adequadamente às diversas solicitações assim que elas se apresentam (Sapateiro *et al*, 2004).

2. Cálculo do índice de (des)organização

Propõe-se aqui um método de cálculo do nível de (des)organização de uma empresa, o qual será monitorizado pelo sistema de controlo de produção referido adiante (alínea 4). Esse valor será designado por "índice de (des)organização", representando-se por Ĩ. O método será desenvolvido considerando a aplicação a unidades industriais de mão-de-obra intensiva, como sejam empresas do ramo da confecção de vestuário, mas poderá estender-se a empresas de outras actividades. Definem-se inicialmente os conceitos de rendimento (R) e de eficiência (E) (referentes a um operador, um posto ou a uma secção). R é obtido pelo quociente entre o tempo de produção e o tempo de presença na empresa; E obtém-se pelo quociente entre o mesmo tempo de produção e o tempo útil de trabalho:

R = Tprodução / Tpresença * 100 [%] , E = Tprodução / Ttrabalho * 100 [%]

Ttrabalho = Tpresença – Tperdido ↗ R ↑ E

Tempo de produção é o tempo que o operador levou para produzir um certo número de peças. A distinção entre tempo de presença e tempo de trabalho, é que o primeiro engloba todo o tempo em que o operador esteve dentro das instalações da empresa (em que pode ter estado não só a produzir peças úteis, mas também pode ter estado parado à espera da reparação de uma máquina) enquanto o segundo se refere exclusivamente ao tempo que o operário teve à sua disposição para produzir trabalho útil. Em termos produtivos, a diferença entre estes tempos é um tempo perdido. A eficiência é um parâmetro que qualifica um operador individualmente, pois representa a capacidade que este tem de produzir peças úteis no tempo que é posto à sua disposição. O rendimento por sua vez é um parâmetro que qualifica a instalação industrial, pois representa a capacidade que esta tem de aproveitar o trabalho dos seus operadores. Por exemplo, se a linha de produção está mal desenhada ou se há avarias prolongadas nas máquinas, existirão muitos tempos perdidos para os operadores. Então, apesar de individualmente eles poderem ser muito eficientes, o rendimento que vão dar á empresa será baixo. A diferença entre E e R está assim directamente relacionada com o tempo perdido, podendo ter-se:

a) Tperdido = 0 ↗ R = E : situação óptima em que o trabalho está correctamente organizado, não existindo tempos improdutivos;

b) Tperdido ← 0 ↗ R ↑ E : existe alguma desorganização na produção, devida à existência de tempos improdutivos;

Na situação (a) toda a capacidade produtiva do operador está a ser aproveitada para efectuar trabalho útil, enquanto em (b) parte dessa capacidade produtiva está a ser desperdiçada. Este último caso conduz portanto a perdas de produtividade. A relação entre os valores de E e de R é pois um indicador do maior ou menor grau de (des)organização a nível da produção. Algumas empresas controlam esta situação calculando o valor de Tperdido (em % do tempo total), porém outros factores contribuem igualmente para a diminuição da produtividade. Segundo E. Goldratt (1988) mais importante que manter a ocupação dos vários postos de trabalho importa manter o fluxo, ou seja, importa assegurar que os produtos em curso de fabrico atravessam a linha de fabrico a ritmo constante, sem flutuações que provoquem acumulações ou faltas de materiais. Deste modo, é igualmente importante conhecer e quantificar (e também qualificar) a variabilidade dos parâmetros E e R, de modo que na sua posse o responsável pela produção possa actuar no sentido de a diminuir.

De novo, a variabilidade de E resulta de causas que podem ser atribuídas ao operador, tais como a sua boa disposição física e mental, o gosto pelo trabalho que realiza, etc. Já a variabilidade de R é provocada pelas condições de trabalho na empresa, em termos da diversidade de modelos fabricados, do tipo de máquinas utilizadas e da sua disposição no terreno, do escalonamento das ordens de fabrico, etc. A variabilidade quer de E quer de R, traduz-se em flutuações no fluxo de fabrico, que conduzem a perdas de produção e que importa diminuir.

A variabilidade de R(ïR) e de E(ïE) são medidas de dispersão que podem ser calculadas através dos respectivos coeficientes de variação(CV):

$$ïR = (1\text{-}CV_R) = (1\text{-}\ S_R/E(R)) \ , \ 0 \uparrow ïR \uparrow 1 \ ^{(a)} \quad [\ CV_R = S_R/E(R) \ , \ S_R : \text{desvio-padrão}\ ]$$

$$ïE = (1\text{-}CV_E) = (1\text{-}\ S_E/E(E)) \ , \ 0 \uparrow ïE \uparrow 1 \ ^{(a)} \quad [\ CV_E = S_E/E(E) \ , \ S_E : \text{desvio-padrão}\ ]$$

O valor ideal para estes coeficientes é a unidade, significando variabilidade nula.

Importa também considerar a relação entre o valor médio de R e de E, que idealmente deverá também ser a unidade, significando tempo perdido nulo:

$$RE = E(R)/E(E) \ , \ 0 \uparrow RE \uparrow 1 \quad [\ E(X) : \text{valor esperado de X (média)}\ ]$$

Finalmente, define-se o índice Ĩ (índice de desorganização) como sendo:

$$Ĩ = RE * ïR * ïE \ , \ 0 \uparrow Ĩ \uparrow 1$$

3. Interpretação do índice de (des)organização

O índice Ĩ será usado para classificar um operador, um posto, uma secção ou mesmo toda a linha de produção. O seu significado é:

- Ĩ = 0 : sistema desequilibrado, em que não se estão a aproveitar devidamente as capacidades dos operadores;
- 0 ↓ Ĩ ↓ 1 : indica o maior ou menor grau de organização da produção;
- Ĩ = 1 : sistema equilibrado, em que não existem tempos perdidos e o ritmo da produção se mantém constante;

[a] Nas condições de estudo, em que $S_X \leq E(X)$

O índice é uma medida de desempenho que permite detectar irregularidades no funcionamento da linha de produção. Quando num determinado posto é detectado que este índice ultrapassa um valor limite (ou seja, está abaixo desse valor), isso significa que esse posto está a funcionar deficientemente podendo pôr em causa o funcionamento de toda a linha. Como exemplo, uma situação destas pode comprometer os prazos de entrega. Nesse caso é gerado um alerta que pode ser consultado localmente ou que será enviado via telemóvel ao responsável para que tome as medidas adequadas. Valores de Ĩ abaixo de 0.8 representam já uma situação de potencial desorganização.

Importa notar que o cálculo de Ĩ não introduz complexidade adicional na recolha de dados, pois os valores de E e de R são já calculados pela maioria das empresas do ramo.

4. Ambiente experimental

Para realizar este estudo recorreu-se à utilização de um sistema de controlo de produção em tempo real desenvolvido por uma empresa de informática da Beira Interior, zona de Portugal aonde existe um significativo número de empresas do ramo têxtil e confecção. Esse sistema é composto por hardware para recolha de dados e um pacote de software para os processar e apresentar (fig.1). O sistema permite monitorizar o andamento das ordens de fabrico ao longo da linha de produção, fornecendo informações sobre a sua localização, bem como efectuar cálculos da eficiência e do rendimento de operários, de postos de trabalho e de secções.

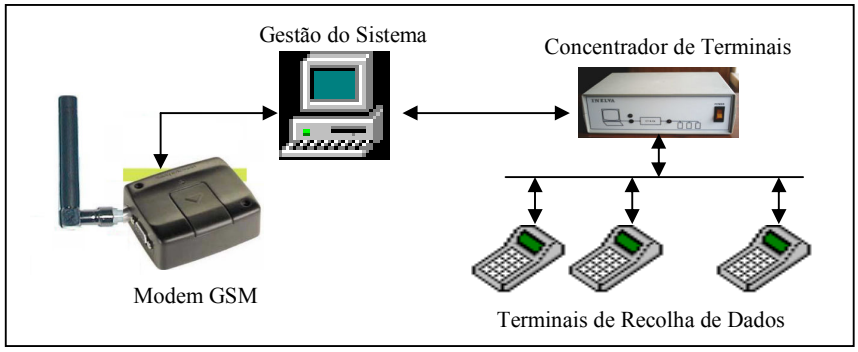


Figura 1 – Ambiente experimental

Ao sistema existente foram acrescentados posteriormente dois módulos. O primeiro para efectuar o cálculo de Ĩ conforme a definição dada nas alíneas anteriores; o segundo para permitir o seu envio remotamente para o telemóvel ou PDA dos responsáveis, conjuntamente com outros parâmetros de produção. Para além dos dados que são enviados automaticamente, este módulo é também capaz de responder a consultas às bases de dados efectuadas remotamente.

O sistema tem também a capacidade de tratar dados não directamente relacionados com a produção. Entre eles incluem-se a detecção de problemas na rede eléctrica como sejam falhas de energia, a detecção de incêndios, de intrusão e de avarias, principalmente no período nocturno. A fig.2 representa o modelo do sistema.

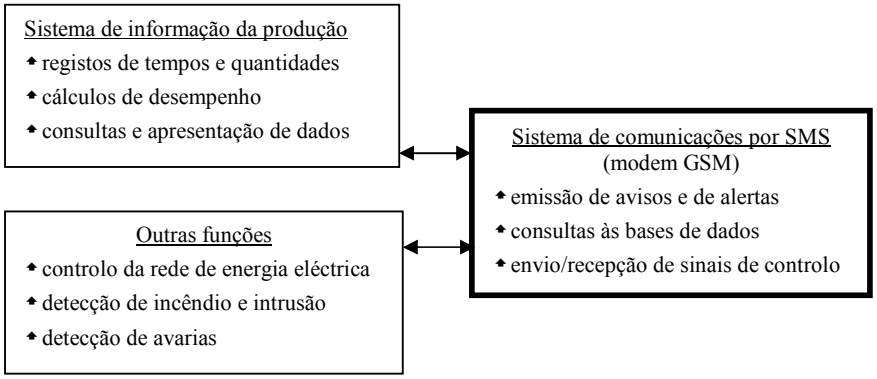


Figura 2 – Modelo do Sistema

5. Resultados experimentais

As figuras 3 e 4 ilustram duas situações opostas para valores de Ĩ. Na primeira, referente ao operador 170, observa-se uma quase coincidência da curva de R com a curva de E além de apenas uma pequena variabilidade em ambas. O resultado é um valor relativamente alto de Ĩ=0.76 (na figura representado por "id"), significando alguma desorganização. Na segunda figura, relativa ao operador 278, observa-se uma grande discrepância e variabilidade entre as duas curvas, conduzindo a um valor bastante baixo de Ĩ=0.33, o que corresponde a um elevado nível de desorganização.

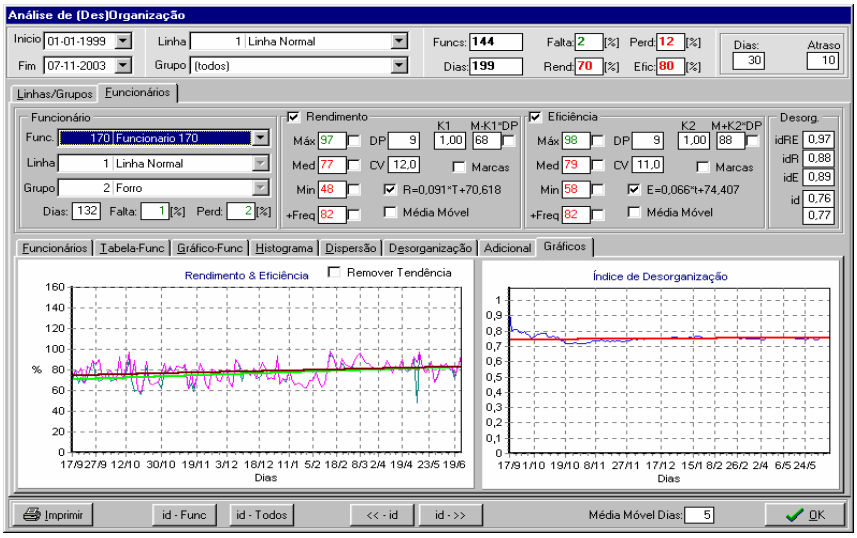


Figura 3 – Situação de pequena desorganização

Figura 4 – Situação de elevada desorganização

Nos exemplos seguintes ilustra-se a troca de mensagens entre o sistema e os telemóveis dos responsáveis da produção.

No primeiro caso, o valor muito baixo de id=0.33 (id=Ĩ), despoletou uma mensagem automática de alerta, pois no posto-chave ocupado pelo operador 278 está a ocorrer um estrangulamento à produção. Perante esta situação o responsável deverá averiguar a causa do problema, que pode ser devida ao operador, ao equipamento ou a outro factor qualquer.

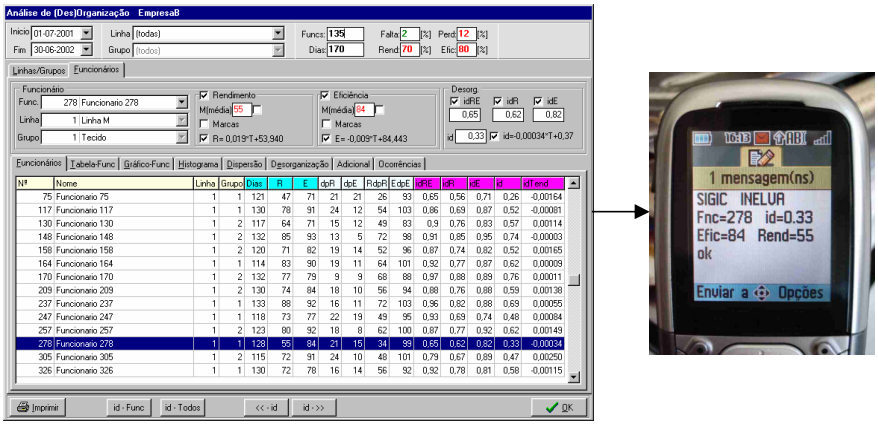


Figura 5 – Alerta automático

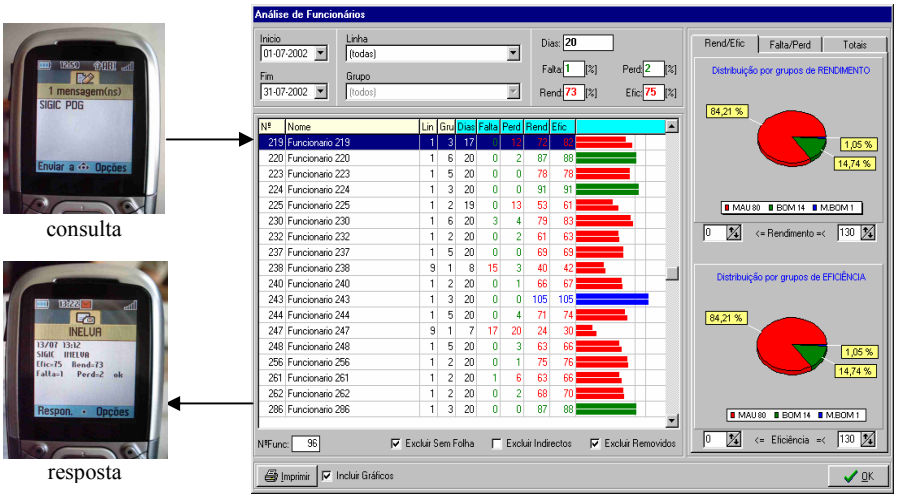


Figura 6 – Consulta à base de dados

Este exemplo mostra a resposta do sistema a uma consulta à base de dados do desempenho. Aqui pretendia-se conhecer os parâmetros Eficiência(Efic), Rendimento(Rend), Tempo de Faltas(Falta) e Tempos Perdidos(Perd). Para efectuar a consulta o responsável enviou ao sistema uma mensagem contendo o comando "PDG" (Parâmetros de Desempenho Gerais), obtendo como resposta os valores pretendidos. Neste caso, os baixos valores de Rend(73%) e de Efic(75%) indicam a existência de problemas na produção, requerendo-se a atenção do responsável.

Finalmente, as figuras seguintes representam a comparação entre níveis de (des)organização em três empresas aonde foi possível recolher dados de produção (que normalmente são considerados confidenciais pelas empresas).

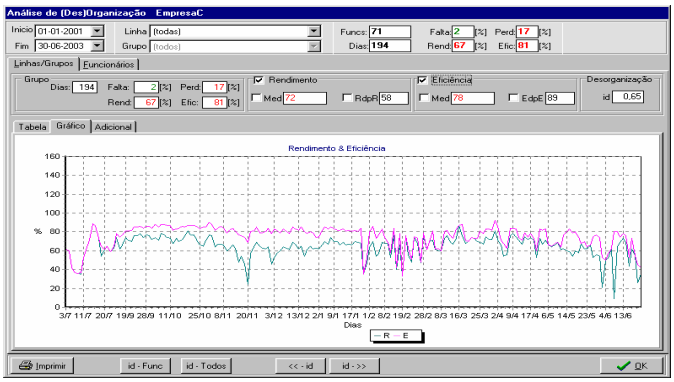


Figura 7.1 – Empresa C

A empresa C (fig. 7.1) apresenta um alto valor de desorganização, resultado do seu tipo de produção caracterizado por muitos modelos e materiais diferentes, o que conduz à ineficiência da linha de produção. A empresa A (fig 7.3), caracterizada por um tipo de produção mais constante, com pequenas variações em termos de modelos e materiais, apresenta um valor de Ĩ mais elevado e portanto menor desorganização. A empresa D (fig 7.2), apresenta um nível intermédio de Ĩ.

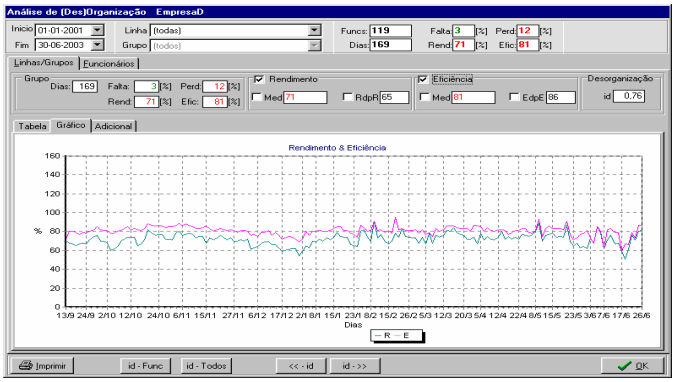


Figura 7.2 – Empresa D

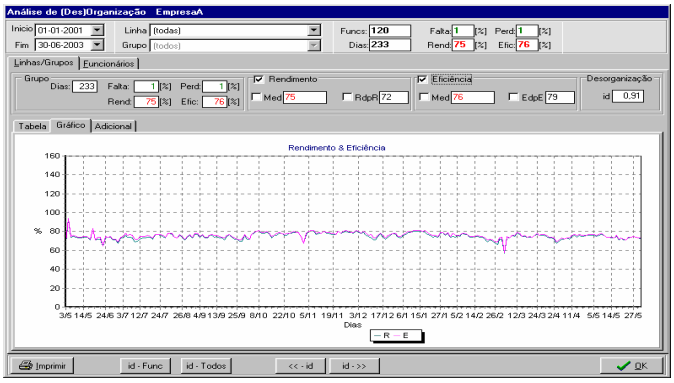


Figura 7.3 – Empresa A

6. Conclusão e trabalho futuro

Este trabalho abordou a definição de um parâmetro que permite avaliar o "estado de desorganização" de uma empresa industrial, a nível da função produção e da gestão das operações. O índice de (de)sorganização (Ĩ) tal como foi definido será usado para avaliar e melhorar o nível de organização interna das empresas, diminuindo os desperdícios e contribuindo assim para o incremento da produtividade e competitividade das mesmas.

Em termos empresariais, a implementação de um sistema de avaliação com as características do apresentado representa uma preocupação real com a problemática da produtividade e competitividade da indústria. Os sistemas e tecnologias da informação são há muito reconhecidos como uma arma estratégica por excelência para apoiar esta luta. Se os grandes desafios colocados aos SI/TI estão em grande parte já resolvidos, subsistem áreas em que podem ser melhorados ou complementados. O exemplo aqui abordado é o do acesso a dados operacionais quando os responsáveis estão ausentes da empresa (quando em viagem por exemplo). O seu desconhecimento pode levar a que não sejam tomadas decisões que seriam necessárias para manter o desempenho da empresa. De facto, a concorrência actual leva a que é cada vez mais importante a capacidade de resposta rápida assim que surgem os problemas, afim de evitar que estes se avolumem.

Em termos informáticos, o desenvolvimento do sistema representa uma oportunidade de realização de experimentação em SI/TI no terreno, tal como defendido por alguns autores (Tichy, 1998). Possibilita também a demonstração do valor dos SI/TI junto das empresas e do modo como estas as podem usar para melhorar o seu desempenho (Mahmood & Mann, 2000). Representa também uma excelente oportunidade para fomentar a ligação universidade-empresa, áreas em que Portugal e Cabo Verde são deficitários.

Determinar o valor económico para a empresa resultante do uso do sistema, é uma tarefa assaz difícil e ainda por realizar. Em parte isso deve-se à dificuldade em comparar os dados após a instalação do sistema, com os dados anteriores, por vezes inexistentes. No entanto, uma medida apenas qualitativa é dada pela taxa de utilização do sistema, quer através do número de consultas efectuadas pelos utilizadores, quer através das inúmeras sugestões que estes apresentam para que sejam incorporadas novas funcionalidades.

As orientações para trabalho futuro passam pelo aperfeiçoamento da definição do índice Ĩ, bem como pela sua aplicação a outros domínios. Em termos da definição apresentada ela não considerou a influência de aspectos como o absentismo, a moda ou a diversidade de materiais, que também podem influenciar a eficiência organizacional da empresa. A expressão obtida para o "índice de desorganização" não é pois única nem definitiva, podendo ser objecto de modificações para reflectirem influências adicionais. Por outro lado, embora até agora apenas tenham sido efectuados testes em empresas industriais do ramo do vestuário e em particular na função produção, os mesmos princípios poderão aplicar-se a outras funções da empresa e mesmo a empresas de outros ramos, uma vez que os problemas tendem a ser comuns. Essa aplicação pode estender-se, por exemplo, à função comercial, a empresas privadas de serviços ou ainda a organismos da administração pública. Neste sentido será necessário determinar que factores influenciam a qualidade organizacional dessas empresas ou de sectores dentro delas, bem como o modo como serão combinados para fornecerem o respectivo índice de (des)organização. Isto abre excelentes perspectivas de investigação em SI/TI.

7. Referências

Brynjolfsson, E., Hitt, L. (2000), Beyond Computation: Information Technology, Organizational Transformation and Business Performance, Journal of Economic Perspectives, Volume 14, Number 4, Pages 23-48.

Goldratt, E. (1988), Computerized Shop Floor Scheduling, International Journal of Production Research, 3, 443-55.

Mahmood, M., Mann, G. (2000), Special Issue: Impacts of Information Technology Investment on Organizational Performance, Journal of Management Information Systems/Spring 2000, Vol.17, No.1, pp 3-10.

Mukhopadhyay, T., Kedre, S., Kalathur, S. (1995), Business value of information technology: A study of electronic data interchange, MIS Quarterly; Jun 1995; 19,2; ABI/INFORM Global; pp. 137-156.

Sapateiro, C., Grosso, S., Cruz, J., (2004), Recolha de dados remotos sobre a rede móvel e sua publicação na Internet, actas da 5ª Conferência da APSI, Nov.2004, ISBN:972-99387-1-7, Lisboa.

Serrano, A., Caldeira, M. (2001), Gestão de Investimentos em Sistemas e Tecnologias de Informação – uma revisão crítica, actas da 2ª Conferência da Associação Portuguesa de Sistemas de Informação (APSI), Évora.

Strassmann, P. (1997), The Squandered Computer, The Information Economics Press.

Tichy, W. (1998), Should Computer Scientists Experiment More?, IEEE Computer, vol 31(5), pp.32-40.

TEMA 2 – CASOS DE APLICAÇÃO DE SISTEMAS DE INFORMAÇÃO

2.2. Casos de Estudo Relativos a Domínios de Aplicação Específicos

Building a collaboratory: Sharing knowledge and resources among the State Laboratories

Zita P. Correia[1], Maria Joaquina Barrulas[1], Luís Arriaga da Cunha[2], Aristides P. Preto Jr[3]

1) INETI, Lisboa, Portugal

zita.correia@ineti.pt; joaquina.barrulas@ineti.pt

2) LNEC, Lisboa, Portugal

lac@lnec.pt

3) INOV, Lisboa, Portugal

aristides.preto@inov.pt

Abstract

This paper introduces a research project whose main goal is to build a collaboratory, understood as a laboratory without walls, in which scientists are connected to each other, to instruments and to data and information, independently of time and location. For this purpose, two State Laboratories (SL) and a private non-profit R&D organisation engaged in a consortium with the aim of implementing in national R&D organisations, innovative models of scientific collaboration already tested and implemented successfully in other countries, with particular incidence in the USA. The research strategy to be adopted is that of the case study. Although the study is limited to just three SL, the generalizability of the results for the remaining SL is guaranteed due to the replication criteria used. The building of the collaboratory will be anchored on three main tasks: a) information audits to be carried out in each of the case study organisations, in order to map out the main information flows, information entities, repositories and systems in a subset of each audited organisation; b) information behaviour research, in order to gain insight into the researchers' information-seeking behaviour and networking patterns, and to relate them to a number of factors, such as education/degree, technical discipline, type of work performed or role played; c) implementation of the appropriate infrastructure and tools, in order to accommodate information archiving and intelligent search, collaborative software tools, multi-channel access, common feel and familiar human interface, and open source software tools-based solutions, whenever possible.

Keywords*:* Collaboratory; Collaborative working environments; Information resource sharing; Scientific collaboration; State Laboratories.

1 Introduction

The State Laboratories (SL) are public research institutions with an important role in supporting Government policies, but also in transferring knowledge to society at large, and to the productive fabric in particular, through activities of service provision, support to the industry, standardization, certification and regulation.

The SL (presently eleven, although the number has varied slightly throughout the years) gather a research community of about two thousand researchers, representing 17% of the national investment in R&D in 2003 and 67% of the investment in the government sector in that same year [Contzen, 2006, Appendix 1][1]. In its evaluation exercise of 2006, the International Working Group (IWG), classified the current state of health of the SL as "quasi sub-critical", emphasizing, among other aspects, that "the institutional framework does not provide any incentive for collaboration between the State Laboratories" [Op.cit., p.14], and noting that the use of common infrastructures should be encouraged [Op. cit., p.17].

To increase the capacity of the SL in fostering innovation within their specific sectors of intervention, a number of measures are needed, and many have been pointed out by the IWG. The contribution of the project described consists in addressing an area that remains to be dealt with: to improve the informational environment of the laboratories, in order to make it amenable and effective for scientific collaboration, without incurring in additional costs for those organizations. The fact that SL work under different ministries and cover a wide variety of scientific and technical fields has hindered their potential for resource sharing and for the establishment of synergies. These barriers may be overcome through the articulation of the SL in a network, using distributed, media-rich network connections to link people to each other, to facilities and to information, thus enabling the creation of a true collaboratory [2].

Collaboratories represent a potential transformation of the idea of laboratory. For scientists and engineers, they have the potential to revolutionize what they can do, how they do it, and who participates in what is being done. The capabilities provided by collaboratories are expected to increase the effectiveness of existing resources. Scientists and engineers become able to interact as if they were using the same physical location, sharing data, high-performance computing systems and instrumentation independently of location.

2 Conceptual framework

2.1 Collaboratories

Scientific work is eminently collaborative, and this trend has accelerated over the past decades, with the boost of the Internet and its role in erasing geographical and temporal barriers. Finholt

---

[1] Contzen, J. P., et al., *Report of the International Working Group on the Reform of the State Laboratories: Redesigning the Governance of the State Laboratories' System*. Lisbon, May 2006. http://www.portugal.gov.pt/NR/rdonlyres/A36261D2-C18F-4FF5-9A4A-DA31FA9FF7AE/0/Relatorio_Internacional_Avaliacao_Laboratorios_Estado.pdf

[2] The project discussed in this paper has been submitted as a project proposal to a funding agency, and is presently waiting for the results of the selection process. However, the two proponent State Laboratories have started a fruitful collaboration for over a year now. The team involved is multidisciplinary, as required by the nature of the project.

[2005] notes that the collaboratory is a remarkable instance of Internet-mediated science, and describes the collaboratory as "a laboratory without walls, in which scientists are connected to each other, to instruments, and to data, independent of time and location".

It has been noted that scientific collaboration has depended on physical proximity. Katz [1994] demonstrated that proximity has a direct effect on the quality and frequency of collaboration. Similarly, access to rare and costly instruments provides an additional reason for co-location in scientific premises [Hagstrom 1965; Traweek 1992]. The conventional response to these limitations, Finholt argues, "has historically been residency, either permanent or temporary, at an instrument site. Therefore, in terms of the performance of joint intellectual work, sharing a common work setting has evolved as a critical factor to support both frequent interaction among collaborators and the use of unique facilities".

Two main factors are generally identified as having contributed to alter significantly this state of affairs: a) the trend toward large scale projects led to greater size and complexity of research tasks and seem to generate higher need for collaboration; b) the developments in ICT favoured the emergence of the concept of "distributed intelligence", meaning that the research activity can be organized in a way that experience and knowledge held by scientists at one location can be easily shared and utilized by scientists elsewhere. The notion of science as "distributed intelligence" is highly dependent on ICT.

The challenge posed by scientific collaboration irrespective of time and location has inspired innovations in ICT. Research programs such as the cyber-infrastructure program in the USA [Berman & Brady 2005] and the e-science program in the UK [Hey & Trefethen 2003] are funding development in ICT to support the needs of remote scientific collaboration, including applications that support collaborative, synchronous access to remote scientific instruments.

Although the concept of "collaboratory" has been in use for more than a decade [Finholt, 1995; Kouzes, Myers & Wulf, 1996; Finholt & Olson, 1997; Agarwal, Sachs & Johnston, 1998; Sonnenwald & Whitton, 2001; Sonnenwald, Maglaughlin & Whitton, 2004; Bondarescu et al., 2005], relatively few scientific collaboratories have been implemented. Sonnenwald [2003] investigated the scientists' expectations concerning scientific collaboratories. Her results raise interesting issues for the design and evaluation of collaboratories, since meeting expectations is one factor that influences the adoption of innovations.

2.2 Information auditing

Information auditing is a process of discovery, monitoring and evaluation of an organization's information flows and resources in order to implement, maintain or improve the organization's

management of information. The information audit should be considered as a required step towards the determination of value, function and utility of information resources in order to completely explore their strategic potential [Buchanan & Gibb 1998].

The information audit can be regarded as a critically important information management tool, since it provides detailed and accurate information on the organisational information environment, as well as an understanding of the way in which the organisation functions. As an information management tool, the information audit contributes for the understanding of the information management processes, at personal, operational, organisational and strategic levels [Botha & Boon 2003]. Ultimately, the information audit aims at aligning the organisational information strategy with the organisational business strategy.

According to Buchanan & Gibb [1998] the information audit seeks to a) identify the organization's information resources; b) identify the organization's information requirements; c) identify the costs and benefits of information resources; d) identify the opportunities to use information resources for strategic competitive advantage; e) integrate IT investments with strategic business initiatives; f) identify information flows and processes; g) develop an integrated information policy; h) create awareness of the importance of IRM [Information resources Management] and define the management role; i) monitor and evaluate compliance with information-related standards, legislation and policy guidelines.

2.3 Information behaviour of scientists and engineers

Wilson [2000, p.49] defined information behaviour as "the totality of human behaviour in relation to sources and channels of information, including both active and passive information seeking and information use". The individual information needs and information seeking and use behaviours have been widely studied: Ellis & Haugan [1997], Wilson [1996, 1999, 2000] Case [2002], Pettigrew, Fidel & Bruce [2001], among many others, emphasise the need to adequate information management practices and information systems to the needs and expectations of the users.

The ways in which information behaviour has been conceptualized and studied have changed profoundly over the last three decades. The most obvious influences have been various strains of the sense-making paradigm as well as constructivist models of thought. At present, the dynamic, personal, and context-laden nature of information behaviour seems to be taken for granted.

Understanding the differences between scientists and engineers [Allen, 1997] has significant implications in planning information services for those two groups. Pinelli et al. [1993] consider

that «such an understanding is also critical to enhancing economic competitiveness, improving productivity, and maximizing the process of technological innovation". It has also been highlighted that social networks play an important role in learning and innovation in organizations: «[…] research has consistently shown that whom you know has a significant impact on what you come to know, because relationships are critical for obtaining information, solving problems, and learning how to do your work» [Cross 2004].

2.4 Free Open Source Software

Free and Open Source Software (FOSS) is freely distributed software that gives the user: the free access to the source and binary code; the authorization to modify the original program; the obligation to distribute those changes according to the terms of the type of license under which the software is available, that can not be discriminative of people, group or initiative neither can it restrict another software or technology (Open Source Initiative).

FOSS is an innovative approach to software development, with a positive impact in the economy. Investing in FOSS is investing in human capital, giving local technicians the chance to develop innovative software solutions, improving competencies, creating local communities, developing local services and local know-how.

From a financial point of view, money is saved on licenses that can be applied on services and specialized staff, depending on the organization's strategy and goals. Furthermore, if the organization decides to buy services, it will be able to change provider avoiding the undesirable "vendor-lock in". FOSS is considered as a viable strategy and an alternative to proprietary technology [Fugetta 2003].

3 Methodology

The research strategy adopted is that of the case study, obeying the logic of theoretical replication (as opposed to literal replication), meaning that the selection of the cases for replication is based on the assumption that each of them will provide different results, due to foreseeable reasons. The research methods to be used include document analysis, observation and inquiry (survey and interviewing).

One of the laboratories was selected for being a multidisciplinary SL, with an interesting mix of scientists and engineers, while another was selected for being a "vertical" SL, focused exclusively on specific engineering fields. A third laboratory will be studied, also of the vertical type, but focused on a specific scientific field.

The study is limited to just three SL, due to the characteristics of the methodology to be used, which is work-intensive and time-consuming, but we believe it is the most appropriate for the tasks that will be undertaken. On the other hand, we think that the generalizability of the results for the remaining SL (according to the case study research strategy, i.e., analytical generalization, as opposed to statistical generalization) is guaranteed due to the replication criteria used.

4 Project outline

4.1 Goal and objectives

The project goal is to build a "collaboratory". The objectives are:

1) To enable the creation of a collaborative working environment, intra and inter-laboratory, increasing the probabilities of occurrence of "cross pollination" among disciplines, through the sharing of digital institutional repositories, databases with information on individual experts and team competencies, and on project ideas.

2) To contribute to the adaptability of the information infrastructure and information systems to the researchers' needs and expectations, by studying their information behaviour and networking patterns.

3) To propitiate the sharing of informational resources and laboratory tools among the laboratories, by mapping the resources available in each of them and enabling the necessary conditions (infrastructure, software applications and other tools).

4) To foster the use of open source software, demonstrating it is a viable strategy and an alternative to proprietary technology, leading to significant savings.

5) To install a pilot plant with shareable resources made available by each of the participating laboratories.

4.2 Tasks

Task 1: Information audit

This task aims at performing a total of three independently conducted information audits, in three SL. Each of these information audits contains three main phases.

During the department level organization audit, a total of three departments, in each of the organizations will be chosen to be the target of the audit. Interviews which each department director will be conducted, in order to identify the main information flows, information entities,

repositories and systems used by the key operational processes for the department. Also a total of four key-persons inside the organization and belonging to these departments will be selected to be interviewed in the next phase.

During the key-person level organization audit, those four key-persons will be interviewed, in order to specify in detail the main information flows, information entities, repositories and systems previously identified. Finally, the information audit analysis will identify the information flows and resources used in common by the three organizational departments audited, and recommendations for improvement will be issued whenever applicable.

Task 2: Information behaviour research

This task aims at providing a framework that will allow an effective planning of the information systems and services for the State Laboratories' researchers taking into consideration: a) The context of the information seeker: professional role played, activities / tasks performed and organizational environment; b) Information sources and channels used; c) Barriers to information seeking identified; d) Time management in information seeking tasks; e) Personal information management strategies; f) Interpersonal relations / networks concerning research activities; g) Information sharing and dissemination patterns.

The success of this task depends on a careful preparation of the field work, on the quality of the data collection instruments and on the quality of the data analysis. The set of data collection instruments developed may be applied subsequently to other State Laboratories, whenever a new SL decides to join the collaboratory. The data analysis will be made using statistical analysis to deal with the questionnaire results and content analysis and network analysis to deal with the interview results.

One of the outcomes of this task will be a model or models depicting the information behaviour of the State Laboratories' researchers.

Task 3: Infrastructure and tools

The collaboratory puts into play, from an IT perspective: a) information archiving and intelligent search; b) collaborative software tools; c) multi-channel access to the services, and d) integration of information from various sources.

A first task will be the definition of a common model (from semantic to formatting) for the different types of information that the collaborative community needs to access, and the implementation of a digital document archive to store it. A second task concerns the offer of the typical collaborative/interactive tools for the researchers, based on open source software, offering not only the more common passive approach but also an active interaction.

Another area that will be developed stems from the fact that nowadays the user wants to be able to access and interact using "any device from anywhere" with a common feel interface. Finally, an IT platform, with the adequate mechanisms, will allow the information that actually resides in different institutions to be seen as "centrally" integrated.

4.3 From laboratories to collaboratory

Collaboratories represent a potential transformation of the idea of laboratory. As explained before, this project envisages the articulation of the SL in a network, using distributed, media-rich network connections to link people to each other, to facilities and to information, but also the improvement of the informational environment of the State Laboratories, in order to make it amenable for scientific collaboration, without incurring in additional costs for these organisations.

The development of the collaboratory will be based: a) on the mapping of the information resources available in the State Laboratories, through information audits (task 1); b) on the study of the researchers' information seeking behaviour and patterns of networking, through inquiry (task 2); c) on the establishment of infrastructures and on the identification of the tools needed for the working of a collaboratory, including the localization and implementation of a set of open source software applications of interest to most of the State Laboratories (task 3).

5 Final considerations

At this stage, we would like to underline what we consider to be the added value brought by this project, in a context of administrative reform and cost-reductions, to the organisations involved and to the country, apart from the advantages that derive directly from knowledge and resource sharing among the contributing networked organisations.

At organisational level, the convergence of the project tasks' outcomes will enable the project team to work out with each of the participating laboratories information policy guidelines, in accordance with each laboratory's mission. These will, in turn, work as a driver for the drafting of the organizational information strategy in a subsequent stage.

At country level, the collaboratory has the potential to "regionalise" the use of costly resources. Building a collaboratory linking three SL (in a first stage, possibly all in the future) located in Lisbon, does no longer mean that only those laboratories have access to their pool of resources. These may become available also to other organisations located elsewhere in the country, from higher education institutions and R&D companies, to technological centres.

6 Bibliographic references

Agarwal, D. A., S. R. Sachs & W. E. Johnston, "The reality of collaboratories", *Computer Physics Communications*, 110, 1998, 134-141.

Allen, T.J., "Managing Technical Communications and Technology Transfer. Distinguishing Science from Technology", in Katz, R. (ed.), *The Human side of managing technology innovation: a collection of readings*, Oxford University Press US, 1997, 307-319.

Berman, F & H. Brady, "Final Report: NSF SBE-CISE Workshop on Cyberinfrastructure and the Social Sciences". www.sdsc.edu/sbe/ [12-10-2006].

Bondarescu, R. et al., "The Astrophysics Simulation Collaboratory Portal: a framework for effective distributed research", *Future Generation Computer Systems*, 21, 2005, 259-270.

Botha, Hannerí & J.A. Boon, "The Information Audit: Principles and Guidelines", *Libri*, 53, 2003, 23-38.

Buchanan, S. & F. Gibb, "The Information Audit: An Integrated Strategic Approach", *International Journal of Information Management*, 18, 1, 1998, 29-47.

Case, D.O., *Looking for Information. A Survey of Research on Information Seeking, Needs, and Behavior*, Academic Press, Elsevier Science, USA, 2002.

Cross, R., & A. Parker, *The Hidden Power of Social Networks: Understanding How Work Really Gets Done in Organizations*, Harvard Business School Press, 2004.

Ellis, D., & M. Haugan, "Modelling the Information seeking Patterns of Engineers and Research Scientists in an Industrial Environment". *Journal of Documentation*, 53, 4, 1997, 384-403.

Finholt, Thomas A., "Collaboratories", *Annual Review of Information Science and Technology*, 36, 1, 2005, 73-107.

Finholt, Thomas A., "Evaluation of electronic work: research on collaboratories at the University of Michigan", *SIGOIS Bulletin*, 16, 2, 1995, 49-51.

Finholt, Thomas A. & Gary M. Olson, "From laboratories to collaboratories: a new form of scientific collaboration", *Psychological Science*, 8, 1, 1997, 28-36.

Fugetta, A., "Open source software – an evaluation", *The Journal of Systems and Software*, 66, 2003, 77-90.

Hagstrom, W. O., *Scientific community*, New York: Basic Books, 1965.

Hey, T & A. Trefethen, "e-Science and its Implications", *Phil. Trans. R. Soc. Lond.*, A, 361, 2003, 1809–1825.

Katz, J. S., "Geographic proximity and scientific collaboration", *Scientometrics*, 31, 1, 1994, 31-43.

Kouzes, Richard T., James D. Myers & William A. Wulf, "Collaboratories: doing Science on the internet", *IEEE Computer*, 8, 1996.

https://collaboratory.emsl.pnl.gov/presentations/papers//IEEECollaboratories.html [28-09-2006]

Pettigrew, K.E., R. Fidel, & H. Bruce, "Conceptual frameworks in Information Behavior", *Annual Review of Information Science and Technology* (ARIST), 35, 2001, 43-78.

Pinelli, T.E., et alli, "The Information-Seeking Behavior of Engineers", NASA/DoD Aerospace Knowledge Diffusion Research Project, paper 31, reprinted from the *Encyclopedia of Library and Information Science*, 52, 15, 1993, 167-201.

http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19930020848_1993020848.pdf [21-06-2005].

Sonnenwald, Diana, Kelly L. Maglaughlin & Mary C. Whitton, "Designing to support situation awareness across distances: an example from a scientific collaboratory", *Information Processing and Management*, 40, 2004, 989-1011.

Sonnenwald, Diana, Mary C. Whitton & Kelly L. Maglaughlin, "Evaluating a Scientific Collaboratory: Results of a Controlled Experiment", *ACM Transactions on Computer-Human Interaction*, 10, 2, June 2003, 150–176.

Traweek, S., *Beamtimes and lifetimes: The world of high energy physics*, Cambridge, MA: Harvard University Press, 1992.

Wilson, T.D., "Human Information Behaviour", *Informing Science, Special Issue on Information Science Research*, 3, 2, 2000, 49-55.

http://inform.nu/Articles/Vol3/v3n2p49-56.pdf [13-09-2004].

Wilson, T.D., "Models in Information Behaviour Research", *Journal of Documentation*, 55, 3, 1999, 249-270.

Wilson, T.D. & C. Walsh, *Information Behaviour: an interdisciplinary perspective*, The British Library Board, British Library Research and Innovation Report, 10, 1996.

http://www.shef.ac.uk/~is/publications/infbehav/prelims.html [10-09-2004].

Avaliação do impacto da Televisão Digital Terrestre (TDT) no desempenho individual

Jacinto Ramos [1] e António Palma-dos-Reis [2]

1) IPQ – Instituto Português da Qualidade, Monte de Caparica, Portugal

jramos@mail.ipq.pt

2) ISEG – Instituto Superior de Economia e Gestão, Lisboa, Portugal
apreis@iseg.utl.pt

Resumo

Apesar da Televisão Digital Terrestre (TDT) se apresentar como uma novidade tecnológica promissora, várias iniciativas empresariais que ousaram a sua utilização não conseguiram afirmar-se no mercado. Em Portugal continua-se à espera da decisão do Governo para que um novo concurso de atribuição da exploração da TDT tenha lugar, depois do consórcio concessionado em 2001 ter falhado os seus compromissos. Tendo em conta esta contradição, o presente trabalho proporciona recomendações para a estratégia dos operadores da TDT pela análise das expectativas sobre o impacto potencial da nova tecnologia no desempenho dos seus adoptantes. Para avaliar estas expectativas, foi adaptada uma métrica de impacto no desempenho a partir de Davis (1989), e foram revistos os fundamentos da Theory of Reasoned Action (TRA), do Technology Acceptance Model (TAM), da Theory of Planned Behavior (TPB) e do Model of Adoption of Technology in Household (MATH), o que permitiu construir um modelo de investigação que inclui as hipóteses colocadas acerca da relação entre a "atitude", os "impactos esperados da TDT" e os vários aspectos que caracterizam a oferta da nova tecnologia.

Identificaram-se os seguintes aspectos como potencialmente relevantes quanto à oferta de serviços da TDT: a oferta de mais canais com melhor som e imagem, o correio electrónico, o comércio electrónico (t-shopping), a formação à distância (t-learning), a assistência médica à distância (t-medicine), o vídeo a pedido (VoD), a utilização do gravador pessoal de vídeo (PVR), o acesso à "internet" e a interactividade. Tendo em conta o MATH, estes aspectos foram analisados considerando três dimensões para medir o impacto esperado pelos indivíduos: a utilitária, a hedónica ou de entretenimento e a social.

Construiu-se um inquérito, administrado a 71 indivíduos, acerca da relevância que a oferta potencial da TDT (contemplada no conjunto de serviços identificados) poderia vir a assumir no dia a dia dos inquiridos. Com o ajustamento do modelo estrutural às respostas obtidas, verificou-se que o impacto da TDT no desempenho individual assume maior importância nos casos da oferta de "t-medicine", "t-learning" e "t-shopping", importância intermédia nos casos da oferta de VoD, utilização do PVR e correio electrónico, e importância muito moderada no caso da oferta de mais canais.

Palavras chave: Televisão Digital Terrestre; TDT; Desempenho individual; Adopção da tecnologia; TRA; TPB decomposta; MATH

1. Introdução

Segundo dados da EU (Comissão das Comunidades Europeias, 2003), 58% das casas portuguesas recebe o sinal aberto dos canais de televisão que operam em Portugal por via hertziana através da rede terrestre de difusão de TV. Assim, no conjunto dos países Europeus, Portugal, a par da França, da Grécia, da Itália e da Espanha, pertence ao grupo dos países em que predomina a *difusão terrestre* do sinal de TV, em contraponto com o grupo dos países que pode ser designado de forma genérica como de *difusão não-terrestre*, onde o sinal de TV é difundido por satélite e por cabo. É o caso do Benelux e da generalidade dos países nórdicos (Adda e Ottaviani, 2005).

Na televisão analógica, a uma frequência corresponde um só canal e um só programa. Na plataforma digital, o número de programas por canal pode ser aumentado significativamente graças à norma ISO MPEG-2 para codificação de televisão de média e alta definição (Bruin e Smits, 1999). A televisão digital terrestre permite difundir por via hertziana imagem, som de qualidade CD-audio, texto e dados digitalizados e pode ser captada com antenas clássicas de hastes colocadas no telhado, podendo o sinal ser enviado directamente a um televisor digital ou então convertido para um televisor analógico por meio de uma "set-top box". O sinal digital pode ser recepcionado em vários tipos de equipamento, como o computador pessoal, o aparelho de televisão e o telefone móvel, dando forma material ao conceito de *multimedia* (Ariño e Ahlert, 2004).

Possibilitando a interacção do espectador com o serviço que lhe está a ser oferecido, a televisão digital é um desafio permanente à sua passividade habitual. Citando Thomas et al. (2001), Brown (2003) considera que as aplicações de iTV que irão proporcionar maiores rendimentos serão: "a publicidade interactiva", "a telefonia premium", "o t-comércio", "os jogos" e "as apostas", atribuindo ainda grande importância ao PVR (gravador de vídeo pessoal).

A televisão digital terrestre apresentou-se pela primeira vez ao mercado europeu na Inglaterra, em Novembro de 1998, com a ITVDigital a apresentar um modelo de negócio assente na subscrição de um serviço pago. Na primavera de 2002 a ITVDigital declarou a sua falência. A 5 de Maio de 2000, em Espanha, o jornal espanhol *ABC* noticiou o lançamento comercial da cadeia de televisão digital terrestre Quiero TV. O mesmo jornal, a 29 de Maio de 2002, noticiou a decisão de encerramento da plataforma por acumulação sucessiva de avultados prejuízos anuais. Em Portugal, em 2002, o operador PDTP (Plataforma de Televisão Digital Portuguesa, SA) é obrigado, por falta de cumprimento dos compromissos assumidos perante o Estado, a entregar a licença de exploração da TDT que lhe tinha sido concedida em 2001.

Parecendo ser várias as vantagens da nova tecnologia para as empresas, para os governos, e para os consumidores, porque é que estes parecem não ter muita propensão a adoptá-la, como é demonstrado pela falência das duas empresas pioneiras e pelo receio do seu lançamento em Portugal? A avaliação do impacto da TDT no desempenho individual, objectivo deste trabalho, poderá contribuir para a resposta a esta questão.

2. Enquadramento Teórico

A TRA (Theory of Reasoned Action) é uma teoria que pretende compreender as principais razões do comportamento humano em face de um problema de dupla ou múltipla escolha. Para a TRA, as "intenções" são determinadas pelas "atitudes" e pelas "normas sociais" e, segundo Ajzen e Fishbein (1980), as "crenças" individuais influenciam tanto as "atitudes" como as "normas sociais". Afirmam estes autores que é inapropriado usar as "crenças" numa relação directa de causalidade com as "intenções" ou com o "comportamento". Consideram ainda que um indivíduo, nas decisões que tem de tomar, apenas consegue usar algumas das crenças que

possui. Ajzen e Fishbein (1980) designam este pequeno conjunto como "crenças com significado".

Segundo a TPB (Theory of Planned Behavior), a "atitude em relação a um comportamento" resulta da associação de uma "crença" a um determinado "resultado", na linha do que Fishbein e Ajzen (1975), inspirados em Vroom (1969), definem como "Modelo Expectativa-Valor", modelo que assume que a maior ou menor inclinação para agir de determinada maneira depende da maior ou menor expectativa de que a acção a empreender produzirá uma determinada consequência (ou resultado) e do valor ou atractividade dessa consequência (ou resultado) para o indivíduo em causa. Podemos assim identificar uma clara ligação das "consequências esperadas" às "atitudes em relação ao comportamento" individual, sendo aqui que encontramos de forma explicita um "constructo" – "as crenças acerca das consequências ou resultados" – que se identifica com o conceito de "impacto esperado". Nesta relação encontra-se também a importância que é dada a cada um dos "impactos esperados", podendo estabelecer-se uma métrica da sua valorização pelos indivíduos, o que permite responder ao aspecto da "avaliação" que existe na questão de investigação. Também para a TPB as "crenças com significado" desempenham um papel determinante na previsão das intenções dos indivíduos.

Baseados nos princípios do enquadramento teórico da TPB, Venkatesh e Brown (2001) realizaram um estudo empírico extensivo nos EUA para identificar os factores que orientavam as famílias americanas na adopção do computador pessoal (PC) em casa. Concluíram que o que mais influenciou as famílias que adoptaram o PC em sua casa foi o que designaram por expectativas em relação aos "resultados utilitários", "resultados hedónicos" e "resultados sociais". Este estudo deu origem ao enquadramento teórico MATH (Model of Adoption of Technology in Household), ao qual Brown e Venkatesh (2005) acrescentaram algumas características demográficas que desempenham um papel moderador importante na intenção de adopção do PC em casa.

Taylor e Todd (1995) estudaram os factores que determinam o uso das tecnologias de informação no seio das organizações recorrendo a uma abordagem integrada da TAM e da TPB e de alguns dos conceitos relativos às teorias da inovação de Rogers (1995). Esta nova abordagem integradora designada "TPB decomposta", apresenta, de forma explicita, uma relação de causalidade entre as "crenças" e as três variáveis que determinam as "intenções": as "atitudes", as "normas sociais" e a "sensação de controlo", conduzindo à construção de um modelo que contempla uma estrutura multi-dimensional de "crenças". Parece de fácil aceitação que o comportamento individual seja multi-dimensional por natureza, tal como o "constructo" que possa permitir a avaliação das "crenças" de um indivíduo relativamente às consequências esperadas por adoptar e usar a TDT.

3. Modelo de investigação

Este trabalho pretende identificar as "consequências esperadas" da adopção da TDT, isto é, quais são os aspectos da oferta potencial de serviços da TDT que os indivíduos crêem poder ser mais influentes na sua vida quotidiana. Vamos por isso recorrer à "elicitação das crenças com significado" a que se referem a TRA e a TPB, para identificarmos os "factores relevantes" que permitem compreender a intenção comportamental dos indivíduos.

No âmbito da TRA, Ajzen e Fishbein (1980) identificam de forma inequívoca a mediação da "atitude" na causalidade das "crenças" sobre as "intenções".

Tendo em conta os trabalhos de Tan e Teo (2000), Venkatesh e Brown (2001), Choi et al. (2003), Hsu e Chiu (2004), Choudrie e Dwived (2004), Lim e Dubinsky (2005) e Brown e Venkatesh (2005), produzidos tendo em consideração o modelo TPB decomposto, pode afirmar-se que a compreensão das especificidades das origens do comportamento podem ser

melhor evidenciadas tendo em conta os modelos que fazem recurso à decomposição das crenças individuais numa perspectiva multi-dimensional.

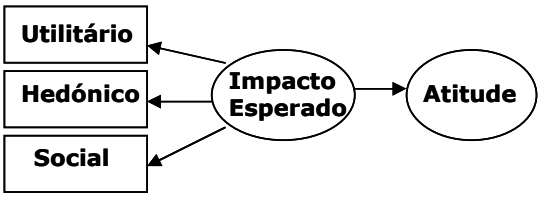


Figura 1 – Modelo de investigação

Na sequência dos trabalhos desenvolvidos no âmbito do MATH, o modelo de investigação que adoptamos, e apresentamos na Figura 1, considera que os impactos da TDT nos indivíduos podem ser de três tipos: 1) impactos que se identificam com a "utilidade" da nova tecnologia; 2) impactos relacionados com o hedonismo, isto é, com a vertente de "entretenimento e lazer" dos indivíduos; e, 3) impactos de cariz social.

4. Hipóteses de estudo

Quase metade dos programas mais vistos na TV são os "Noticiários" (Cardoso e Santos, 2004). Isto evidencia a dimensão "utilitária" da vertente informativa da televisão. A TDT pode, no entanto, aumentar significativamente a "utilidade" da televisão através da oferta da capacidade de interacção em linha que, uma vez colocada à disposição do utilizador, aproxima o uso da TV das funcionalidades do computador pessoal no que respeita à sua capacidade de comunicação e interacção com a "internet", à utilização de correio electrónico, ou de outro tipo de serviços. É o caso da telemedicina (Niiranen et al., 2002), da educação (Damásio, 2004), do t-shopping (Wade, 1999) e do comércio electrónico em geral, e ainda da vertente utilitária do PVR ("personal video recorder") e do VoD ("video on demand") (Brown, 2003).

Com base na literatura mencionada anteriormente, as hipóteses de carácter utilitário que colocamos são as seguintes: (H1) o aumento da oferta informativa através do aumento de canais disponíveis permitido pela TDT é um dos impactos úteis esperados pelos indivíduos; (H2) a possibilidade de usar um correio electrónico através da TDT é um dos impactos úteis esperados pelos indivíduos; (H3) a possibilidade de fazer compras através da TDT (t-comércio) é um dos impactos úteis esperados pelos indivíduos; (H4) a possibilidade de ser assistido em casa por um médico à distância (telemedicina) é um dos impactos úteis esperados pelos indivíduos; (H5) a possibilidade de utilizar a TDT como meio de aprendizagem formal (t-learning) é um dos impactos úteis esperados pelos indivíduos; (H6) a possibilidade de armazenamento local de conteúdos permitindo visionar os programas no horário que mais lhes interesse através da utilização do PVR é um dos impactos úteis esperados que pelos indivíduos, e, (H7) a possibilidade de descarregar vídeo à medida das necessidades permitindo utilizar um clube de vídeo em linha (VoD) é um dos impactos úteis esperados pelos indivíduos.

De acordo com Davis et al. (1992) a "sensação de prazer" em usar uma tecnologia corresponde ao maior ou menor grau de deleite que alguém usufrui pelo facto de a usar independentemente da melhoria de desempenho que possa resultar do seu uso. Os conceitos de "sensação de divertimento" (Igbaria et al., 1994) e de "sensação lúdica" (Moon e Kim, 2001) estão relativamente próximos do de "sensação de prazer". Yu et al. (2005) conclui que a "sensação de prazer" tem uma influência sobre a atitude individual em relação ao "t-comércio" muito superior à "sensação de utilidade". Van der Heijden (2004) comprova a hipótese de que para sistemas hedónicos, a "sensação de prazer" explica melhor a intenção de um indivíduo em relação a um comportamento do que a "sensação de utilidade". Também Choi et al. (2003) afirma que a "sensação de prazer" é um dos factores que mais influencia a atitude dos indivíduos em relação à adopção de tecnologia em casa.

Como "a televisão é antes de mais um meio de difusão de ficções e de diversão" (Guillaume, 1989), e a TDT pode potenciar esta característica, é plausível que um dos impactos esperados

que pode ter mais significado para os indivíduos seja o do acréscimo da quantidade e da qualidade das alternativas de entretenimento.

Desde a possibilidade de interacção do telespectador com os conteúdos, passando pela webTV, o VoD e o uso do PVR, até à convergência entre a TV digital e os brinquedos (Quico, 2004), existe uma panóplia incomensurável de oportunidades de diversão e entretenimento. Assim, as hipóteses de carácter hedónico que colocamos são as seguintes: (H8) o aumento da oferta de possibilidades de entretenimento é um dos impactos esperados pelos indivíduos; (H9) o aumento das possibilidades de entretenimento é mais valorizado pelos indivíduos do que os outros impactos; (H10) a oferta de canais com melhor som e imagem é um dos impactos de entretenimento esperados pelos indivíduos; (H11) a melhoria da qualidade do som e da imagem é um dos aspectos de entretenimento mais valorizado pelos indivíduos; (H12) a possibilidade de interacção com os conteúdos visionados é um dos impactos de entretenimento esperados; (H13) a possibilidade de aceder à "internet" através da TDT (webTV) é um dos impactos lúdicos esperados pelos indivíduos; (H14) a possibilidade de utilizar um clube de vídeo em linha (VoD) é um dos impactos de entretenimento esperado pelos indivíduos; (H15) a possibilidade de visionar os programas no horário que mais lhes interesse utilizando o PVR, é um dos impactos de entretenimento esperados pelos os indivíduos.

Regra geral, os primeiros indivíduos a usar uma inovação são os que acham que esse comportamento lhes aufere perante os outros um estatuto de superioridade, de liderança e de exemplo perante o grupo – são os utilizadores precoces (Rogers, 1995). Assim, colocamos a seguinte hipótese: (H16) o aumento da imagem e do estatuto social é um dos impactos esperados pelos indivíduos.

A televisão é um veículo que proporciona a interacção social entre as pessoas que a vêem (Morley, 1993). No entanto, a capacidade de interactividade da TDT permite colocar à disposição dos indivíduos uma tecnologia com as funcionalidades de um computador pessoal, que, em casa, é usado habitualmente de forma individual, inibindo com facilidade a interacção social (Morrison e Krugman, 2001). Assim, devemos colocar a seguinte hipótese: (H17) o aumento do uso individual da televisão é um dos impactos esperados pelos indivíduos.

A TDT, ao permitir usar a televisão como se fosse um computador pessoal pelo menos no que respeita ao entretenimento com jogos electrónicos, tenderá também a aumentar as preocupações dos pais relativamente ao conteúdo dos referidos jogos e ao seu poder aditivo em relação ao comportamento das crianças, sentimento que é expresso pelas famílias entrevistadas no trabalho de Frohlich et al. (2001), quando referem a sua grande preocupação em controlar a utilização que os seus filhos fazem do computador pessoal. Assim, devemos colocar a seguinte hipótese: (H18) a adopção da TDT em casa irá aumentar a preocupação dos pais em relação à necessidade de controlo do uso que os seus filhos possam fazer dela.

5. Metodologia

Para testar as hipóteses apresentadas, foi elaborado um inquérito com vista a recolher os dados para a análise estatística, tendo sido feita a sua validação com uma "Confirmatory Factor Analysis" e ajustado um modelo estrutural (SEM) para obtenção dos resultados (Bollen, 1989).

O inquérito foi elaborado tendo em conta Davis (1989), que considera um conjunto de seis itens para testar a variável "utilidade" duma estação de trabalho pessoal, usando uma escala de Likert de sete pontos. Para medir a variável "entretenimento e lazer", considerámos a operacionalização feita por Van der Heijden (2004) e Igbaria et al. (1994), que propõem um conjunto de 4 itens bipolares para testar a variável "hedónica" usando uma escala de Likert de sete pontos. Para medir os "impactos sociais esperados", considerámos três vertentes: uma que respeita à imagem e estatuto pessoais, tendo em conta Choi et al. (2003); outra relativa ao poder paternal

no seio da família, operacionalizada a partir de McDonald (1982) e, por fim, uma relativa à interacção da família com a televisão, com operacionalização desenvolvida no âmbito deste trabalho. Este inquérito foi administrado a 71 indivíduos residentes em Lisboa que possuem, pelo menos, uma licenciatura, e proporcionou os resultados apresentados na secção seguinte.

6. Resultados e Discussão

Dos 71 indivíduos inquiridos, 60% são do sexo masculino e 40% do sexo feminino, apresentam idades compreendidas entre os 23 e os 58 anos, verificando-se que 17% têm menos de 29 anos, 18% têm mais de 42 anos e os restantes 64% têm entre 29 e 42 anos de idade. Quanto ao seu estado civil, 58% são casados, 35% são solteiros e 7% têm outra situação.

Observando as respostas às questões dicotómicas (sim ou não), constatamos que: 96% possui computador pessoal em casa, 93% possui ligação à "internet" em casa, 96% reside num local em que existe TV por cabo, 76% tem assinatura de TV por cabo, 96% considera que a TDT tem vantagem em relação à TV actual e 27% manifesta a intenção de aderir de imediato à TDT.

As expectativas dos indivíduos apresentam-se bastante mais positivas em relação aos impactos de tipo lúdico (ou de entretenimento) do que em relação aos outros dois tipos de impacto esperado – utilitário e social. A grande relevância dos impactos de tipo hedónico é claramente confirmada pelo número de indivíduos cujas respostas se situam acima do ponto médio da escala em relação a todos os respectivos itens, sem excepção. Para todos estes itens, as respostas acima do ponto médio da escala variam entre o mínimo de 59% e o máximo de 87%, facto que se traduz nos 5,17 da média global de resposta e no valor negativo de -0,66 do enviesamento global, evidenciando uma grande incidência de respostas à direita da escala.

Ao contrário do impacto de tipo lúdico, os indivíduos inquiridos não valorizam de forma significativa os potenciais impactos sociais. As expectativas dos indivíduos em relação a este tipo de impacto são mesmo as mais baixas dos três tipos de impacto considerados. De facto, neste caso, a resposta maioritária dos indivíduos a todos os itens, sem excepção, é na banda abaixo do ponto médio da escala. Isto é comprovado pelo baixo valor de 2,58 da média global de respostas e pelo valor positivo de 0,83 do enviesamento global que corresponde à existência duma esmagadora quantidade de respostas à esquerda da escala. Para estes itens, a quantidade de respostas abaixo do ponto médio oscila entre os 58% e os 86%.

Já a distribuição das respostas relativas ao impacto utilitário apresenta-se muito menos enviesada, revelando o valor global de 3,90 do Teste de Bowman-Shelton estarmos em presença de um conjunto com as características de uma distribuição normal (Newbold et al., 2002), embora com uma média global de 3,35, um pouco abaixo do ponto médio (PmedU) da escala, parecendo existir uma tendência genérica para que as respostas considerem que este tipo de impactos tem alguma importância mas não demasiada.

Por último, saliente-se o facto de mais de 90% das respostas em relação aos itens que medem a "atitude" serem iguais ou superiores ao ponto médio da escala, parecendo revelar um potencial de aceitação da tecnologia muito elevado.

O ajustamento do modelo da Figura 1 aos dados recolhidos resultou nos coeficientes (pesos) apresentados na Figura 2.

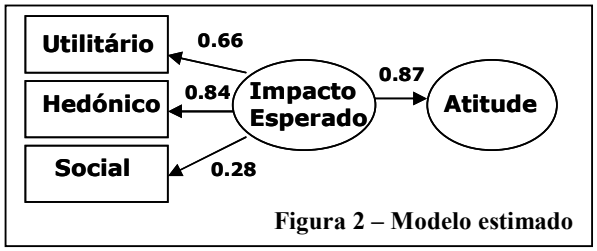


Figura 2 – Modelo estimado

Salientando que todas as variáveis observadas têm significância estatística ($p<0,05$), verificamos que a que tem mais peso nesta estrutura é o "impacto esperado de entretenimento" (0,84) e a que tem menor importância é o "impacto social

esperado" (0,28) pelos indivíduos, o que valida a hipótese H9 de que "o aumento das possibilidades de entretenimento é mais valorizado pelos indivíduos do que os outros impactos". Assim, não basta que um serviço seja útil. A componente de entretenimento tem também grande relevância para a aceitação dos serviços prestados pela nova tecnologia.

Embora com pesos específicos diferenciados em todos os modelos estimados relativos a cada um dos sete aspectos utilitários avaliados, a probabilidade do respectivo peso ser nulo é sempre inferior a 0,001. Podemos assim afirmar que todas as hipóteses colocadas acerca dos aspectos utilitários são válidas (H1 a H7). A Tabela 1 apresenta os pesos relativos de cada uma das variáveis utilitárias nos respectivos modelos estruturais.

Var.Utilitárias	Peso (λ)
mais canais	0,39 *
e-mail	0,48 *
t-shopping	0,46 *
t-medicine	0,60 *
t-learning	0,48 *
PVR	0,40 *
VoD	0,56 *

* p<0,001

Tabela 1

Var.Entretenimento	Peso (λ)
som e imagem	0,51 *
interactividade	0,58 *
internet	0,58 *
VoD	0,71 *
PVR	0,67 *

* p<0,001

Tabela 2

Var.Sociais	Peso (λ)
imagem e estatuto	0,51 *
TV em família	0,58 **
poder paternal	0,67 **

* p<0,01; ** p>0,15

Tabela 3

Observe-se que o impacto utilitário mais valorizado pelos indivíduos é o relativo ao serviço de teleMedicina logo seguido pelo impacto utilitário do vídeo a pedido (VoD). O peso específico do primeiro é de 0,60 e o do segundo é de 0,56. O impacto utilitário menos valorizado é o da oferta de mais canais, cujo peso é 0,39.

O facto de a teleMedicina aparecer como o mais saliente dos impactos percebidos pelos indivíduos apresenta-se como um enorme desafio para as empresas e para o Governo. Assim, a procura para o exercício da medicina à distância existe, esperando-se que, como diz Cádima (2004), "os decisores dêem maior importância aos factores que influenciam as preferências de consumo e a procura, antes de avançarem com estratégias de inovação e objectivos industriais".

A Tabela 2 apresenta os pesos relativos de cada uma das variáveis de entretenimento nos respectivos modelos estruturais. Como se vê, todas as variáveis observadas consideradas são estatisticamente significativas e todas têm pesos específicos de sinal positivo. Deste modo, estão verificadas as hipóteses H8 a H10 e H12 a H15, não se verificando H11. De facto, a melhoria de som e imagem não é dos aspectos mais valorizados pelos indivíduos, apresentando mesmo o seu peso específico (0,51) o menor valor de todos os aspectos de entretenimento.

Realce-se a semelhança dos pesos do par "interactividade" e "internet". Considerar estas duas vertentes como um único aspecto de marketing para uma televisão digital que ofereça interactividade em linha, pode ser uma questão a ter em conta pelos profissionais ligados à publicidade. Por exemplo, é plausível que um anúncio comercial interactivo remeta de imediato para a questão do comércio electrónico e, consequentemente, para a necessidade de acesso à "internet".

A Tabela 3 mostra os pesos específicos de cada uma das variáveis relativas aos aspectos sociais. Como pode ser visto, só a "imagem e estatuto" pode ser considerada estatisticamente significativa, tendo um peso específico de sinal positivo. Qualquer das outras duas variáveis consideradas não são estatisticamente significativas (p>0,15). Assim, quer os aspectos da "relação da família com a TV" quer as questões relacionadas com o "poder paternal" parecem não ser valorizados pelos indivíduos.

Para estimar as perspectivas dos inquiridos sobre o impacto da TDT no seu desempenho, utilizaram-se, como aproximação ao conceito de desempenho individual, cinco dos seis itens

propostos por Davis (1989) para medir a utilidade: a realização mais rápida das tarefas, a melhoria da eficácia pessoal, a maior facilidade de realização de tarefas, a melhoria do desempenho, e o aumento da produtividade. Não foi incluído o item "utilidade para o dia a dia" por não se apresentar adequado ao conceito de desempenho, facto que comprovámos com um teste de contingência.

Assim, apresentamos por último uma medida das expectativas acerca do impacto da TDT no desempenho individual. Designamos esta medida por "taxa de impacto esperado da TDT no desempenho individual" que definimos como "a probabilidade de encontrar um indivíduo que afirme que a TDT pode ter algum impacto no seu desempenho – P(z>PmedU), ponderada pela importância estrutural relativa – $\lambda_j/(\lambda_1+\lambda_2+\lambda_3)$ – que cada um dos aspectos da oferta de serviços da TDT tem no respectivo modelo estrutural. A sua formulação matemática é a seguinte:

$$\mathbf{Tx.Impacto_j = P(z>PmedU_j) \ . \ \lambda_j/(\lambda_1+\lambda_2+\lambda_3)}$$

Como podemos verificar na Tabela 4, a Taxa Global de impacto esperado da TDT no desempenho individual é de 7,9%, sendo a "oferta de mais canais", com 1,5%, o aspecto que menos expectativas cria nos indivíduos quanto ao impacto que possa ter no seu desempenho. Deste ponto de vista, os aspectos mais relevantes da oferta de serviços da TDT parecem ser a "teleMedicina", o "t-Learning" e o "t-Shopping". De facto, são claramente identificáveis três conjuntos distintos quanto ao impacto no desempenho individual: 1) o conjunto de maior impacto esperado constituído pela teleMedicina", com 12,6%, pelo "t-Learning", com 11,6% e pelo "t-Shopping", com 10,4%; 2) o conjunto de impactos esperados mais moderados constituído pelo e-Mail, com 8,5%, pelo PVR, com 8,0% e pelo VoD, com 9,1%; e, 3) a oferta de mais canais com apenas 1,5% de impacto esperado no desempenho individual.

Tabela 4

	+canais	e-mail	t-shop	t-medic	t-learn	PVR	VoD	Global
P(z>PMedU)	0,06	0,28	0,35	0,35	0,39	0,31	0,28	0,21
$\lambda_j/(\lambda_1+\lambda_2+\lambda_3)$	0,28	0,30	0,29	0,36	0,30	0,26	0,33	0,37
Tx.Impacto	1,5%	8,5%	10,4%	12,6%	11,6%	8,0%	9,1%	7,9%

7. Conclusões

Primeiro, e de acordo com os trabalhos de Van der Heijden (2004) e Yu e tal. (2005), os impactos hedónicos esperados pelos indivíduos são os que têm maior importância, revelando que a televisão continua a estar muito associada às questões do entretenimento e do lazer. Os impactos hedónicos são mais importantes que os impactos de tipo utilitário e muito mais importantes que os impactos de tipo social, que têm muito pouca importância para os indivíduos.

Segundo, em relação aos aspectos de tipo utilitário, apesar de podermos considerar que existe alguma expectativa por parte dos indivíduos, não nos parece contudo que esta seja demasiado relevante para que possamos dizer que existirá uma corrida em massa à adopção da nova tecnologia por sua causa.

Terceiro, os aspectos mais tecnológicos da oferta potencial da TDT são menos valorizados do que os serviços propriamente ditos que a mesma tecnologia permite colocar à disposição dos indivíduos. São assim quase irrelevantes os impactos esperados pelo facto de poderem existir mais canais, bem como se constata uma menor valorização dos impactos perspectivados com a oferta de canais com melhor som e imagem do ponto de vista do entretenimento. Portanto, não será pelo facto de existirem mais canais com melhor som e imagem que os indivíduos vão aderir à nova tecnologia.

Quarto, a adesão dos indivíduos à TDT parece poder ser mais motivada pela possibilidade de ter acesso a serviços como os de um clube de vídeo onde terá disponível a qualquer hora uma videoteca com a quantidade e qualidade de filmes, documentários e outros tipos de vídeo que permitam preencher as horas de lazer. Este sim, parece ser um motivo que poderá facilitar a adesão dos indivíduos à nova tecnologia, quer do ponto de vista utilitário quer do ponto de vista do entretenimento e do lazer.

Quinto, a interactividade, ou por falta de experiência concreta, e portanto sem grande capacidade de perspectivar uma forma de aproveitamento lúdico dos aspectos que lhe estão subjacentes, ou por outra qualquer razão, é um aspecto da oferta de serviços da TDT relativamente menos valorizado em termos de impacto percebido pelos indivíduos, frustrando um pouco as nossas expectativas iniciais. Este aspecto está, no entanto, ao nível da percepção do impacto esperado da "internet" como factor de entretenimento, o que indicia de alguma maneira que ambos os aspectos podem constituir um conjunto com um significado e peso específicos no que respeita às expectativas dos indivíduos em relação à TDT.

Sexto, a teleMedicina é o aspecto utilitário a que é atribuída maior importância, quer em termos de impacto utilitário esperado, quer no que respeita ao aspecto específico do impacto esperado no desempenho individual.

Sétimo, as questões sociais não são reconhecidas como relevantes pelos indivíduos. Quer no que respeita à dinâmica da família em relação à televisão quer no que diz respeito às preocupações paternas em relação aos filhos, as expectativas de impacto da TDT não têm significado relevante para os inquiridos. No entanto, de acordo com Rogers (1995), constata-se alguma tendência para se considerar que a adopção da TDT pode melhorar a imagem pessoal junto do respectivo grupo social, o que pode ser aproveitado pelo meio publicitário para lançar e promover a nova tecnologia.

Por fim, quanto ao impacto no desempenho individual são de salientar os três aspectos que geram mais expectativas nos indivíduos: em primeiro lugar a teleMedicina, em segundo lugar o t-Learning e em terceiro lugar o t-Shopping. Este trio apresenta-se como mais importante do que um segundo trio composto pelo e-Mail, VoD e PVR que, de qualquer modo, tem uma relevância mais significativa do que a oferta de mais canais. Esta constatação revela uma clara tendência para se considerarem os serviços prestados por terceiros como aqueles que mais impacto podem ter no desempenho dos indivíduos, funcionando aqui a TDT como um "media" facilitador da troca, ou seja, como um intermediário que serve a conveniência dos indivíduos.

8. Referências

Adda, Jérôme e Ottaviani, Marco, "The transition to digital television", *Economic Policy*, (2005), 159-209.

Ajzen, I. e Fishbein, M., *Understanding attitudes and predicting Social Behavior*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1980.

Ariño, Mónica e Ahlert, Christian, "Beyond Broadcasting: The Digital Future of Public Service Broadcasting", *Prometheus*, 22, 4 (2004).

Bollen, K. A., *Structural Equations with Latent Variables*, New York: Wiley, 1989.

Brown, Allan, "The Digital Future of Terrestrial Advertiser-supported Television", *Prometheus*, 21, 1 (2003).

Brown, Susan A. e Venkatesh, Viswanath, "Model of Adoption of Technology in Households: A baseline Model Test and Extension Incorporating Household Life Cycle", *MIS Quarterly*, 29, 3 (2005), 399-426.

Bruin, Ronald de e Smits, Jan, *Digital Video Broadcasting: Technology, Standards and Regulations*, Artech House, Norwood, 1999.

Cádima, Francisco R., "Desafios da Televisão Digital: da TDT à iTV", *Obercom – Revista do Observatório da Comunicação*, 10, (2004), 161-170.

Cardoso, Gustavo e Santos, Susana, "Tendências e Contradições no Sistema Televisivo: da televisão interactiva à televisão em rede", *Obercom – Revista do Observatório da Comunicação*, 10, (2004), 129-141.

Choi, H., Choi, M., Kim, J. e Yu, H., "An empirical study on the adoption of information appliances with focus on interactive TV", *Telematics and Informatics*, 20, (2003), 161-183.

Choudrie, Jyotie e Dwived, Yogesh K., "Towards a Conceptual Model of Broadband Diffusion", *Journal of Computing and Information Technology* – CIT 12, 4 (2004), 323–338.

Comissão das Comunidades Europeias, "9th Report on Implementation of the Telecommunications Regulatory Package. Annex 1: Market overview", COM (2003) 715 final, SEC (2003) 1342.

Damásio, Manuel J., "Televisão interactiva: aplicações educativas", *Obercom – Revista do Observatório da Comunicação*, 10, (2004), 77-97.

Davis, Fred D., "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology", *MIS Quarterly*, 13, 3 (1989).

Davis, F.D., Bagozzi, R.P. e Warshaw, P.R., "Extrinsic and intrinsic motivation to use computers in the workplace", *Journal of Applied Social Psychology*, 22, 14 (1992), 1111-1132.

Fishbein, M., e Ajzen, I., *Belief, attitude, intention, and behavior: An introduction to theory and research*, Reading, MA: Addison Wesley, 1975.

Frohlich, David M., Dray, Susan M. e Silverman, A., "Breaking up is hard to do: family perspectives on the future of the home PC", *International Journal of Human-Computer Studies,* 54, 5 (2001), 701-724.

Guillaume, Marc, "Novas técnicas e novas culturas da comunicação", *Revista de Comunicação e Linguagens*, 9, (1989), 117-125.

Hsu, Meng-Hsiang e Chiu, Chao-Min, "Predicting electronic service continuance with a decomposed theory of planned behaviour", *Behaviour & Information Technology*, 23, 5 (2004), 359-373.

Igbaria, M., Schiffman, S.J. e Wieckowski, T.J., "The respective roles of perceived usefulness and perceived fun in the acceptance of microcomputer technology", *Behaviour & Information Technology*, 13, 6 (1994), 349-361.

Lim, Heejin e Dubinsky, Alan J., "The Thoery of Planned Behavior in E-Commerce: Making a case for Interdependencies between Salient Beliefs", *Psychology & Marketing*, 22, 10 (2005), 833-855.

McDonald, Gerald W., "Parental Power Perceptions in the Family", *Youth & Society*, 14, 1 (1982), 3-31.

Moon, J.-W. e Kim, Y.G., "Extending the TAM for a World-Wide-Web context", *Information & Management*, 38, (2001), 217-30.

Morley, David, *Family television: cultural power and domestic leisure*, London: Routledge, 1993.

Morrison, Margaret e Krugman, Dean M., "A Look At Mass and Computer Mediated Technologies: Understanding the Roles of Television and Computers in the Home", *Journal of Broadcasting & Electronic Media*, 1, 45 (2001), 135-161.

Newbold, P., Carlson, W., Thorne, Betty M., *Statistics for Business and Economics*, Prentice Hall, Fifth Edition, 2002.

Niiranen , S., Lamminen, H., Mattila, H., Niemi, K. e Kalli, S. "Personal health care services through digital television", *Computer Methods and Programs in Biomedicine*, 68, (2002), 249–259.

Quico, Célia, "Televisão Digital e Interactiva: o desafio de adequar a oferta às necessidades e preferências dos utilizadores", *Obercom – Revista do Observatório da Comunicação*, 10, (2004), 115-128.

Rogers, Everett M., *Diffusion of Innovations*, Fourth Edition, New York, The Free Press, 1995.

Tan, Margaret e Teo, Thompson S. H., "Factors Influencing the Adoption of Internet Banking", *Journal of the Association for Information Systems*, 1, (2000), 1-42.

Taylor, S. e Todd, P. A., "Understanding Information Technology Usage: A Test of Competing Models", *Information Systems Research*, 6, 2 (1995), 144-176.

Thomas, Adam, Dyson, Simon e O'Brien, Stuart, *The Digital Broadcasting Revolution*, 4th Ed., Informa Media Group, London, 2001.

Van der Heijden, H., "User Acceptance of Hedonic Information Systems", *MIS Quarterly*, 28, 4 (2004), 695-704.

Venkatesh, Viswanath e Brown, Susan A., "A Longitudinal Investigation of Personal Computers in Homes: Adopting Determinants and Emerging Challenges", *MIS Quarterly*, 29, 1 (2001), 71-102.

Vroom, V., "Industrial Social Psychology", *The Handbook of Social Psychology*, G. Lindzey and E. Aronson Eds., 5, (1969), 200-208.

Wade, Nicola e Mckechnie, Sally A., "The impact of digital television: will it change our shopping habits?" *Journal of Marketing Communications*, 5, (1999), 71–84.

Yu, J., Ha, I., Choi, M. e Rho, J., "Extending the TAM for a t-commerce", *Information & Management*, 42, (2005), 965-976.

Analyzing competitive strategy drivers to define e-commerce implementation stage

Winnie Picoto[1], António Palma dos Reis[2]

1) Instituto Superior de Economia e Gestão – Universidade Técnica de Lisboa, Lisbon, Portugal
w.picoto@iseg.utl.pt

2) Instituto Superior de Economia e Gestão – Universidade Técnica de Lisboa, Lisbon, Portugal
apreia@iseg.utl.pt

Abstract

E-commerce plays an increasingly important role in providing companies with the required competitiveness, thus making it critical to analyze and evaluate the key aspects to have in mind when designing the e-commerce strategy. Today companies face the challenge of deciding the level of investment and commitment dedicated to the Internet channel and its degree of sophistication as well as its scope of functionalities.

This paper examines the relationship between the main drivers for a successful e-commerce strategy and the e-commerce implementation stages that should be adopted to support that strategy addressing issues such as how deep should an organization's Internet strategy go and which factors determine the adequate level of resources to commit to the Internet presence.

Key words: e-commerce, electronic commerce, strategy, website; business model, B2C

1. Introduction

E-commerce has assumed a growing role in organizational strategy over the past two decades. Although e-commerce successful strategies have been implemented, most organizations have faced difficulties and even failure in developing an effective internet channel positioning [Zeng and Reinartz 2003]. E-commerce affects the way businesses are conducted across different industries and the correct usage of Internet technology can be a major determinant of a company's competitive advantage and financial performance. Thus, for many companies, e-commerce is not just an alternative anymore, it has become an imperative to survive in today's marketplace. The question is no longer whether the company should invest in e-commerce. The real question is how can a company's investment in e-commerce become as profitable as possible.

The Internet is an extremely important technology [Porter 2001], and its usage has increased even after the bubble burst in April 2000 [Zeng and Reinartz 2003]. Consumer Internet expenditures will double from $37,552 million dollars in 2004 to $74,038 million dollars in 2009, according to Vernis Suhder Stevenson [2005]. Additionally, the number of Internet users reached 200 million in Europe in 2004 [Datamonitor 2005] and e-commerce sales in the second quarter of 2004 accounted for 1.7% of total sales while in the second quarter of 2003 e-commerce sales were 1.5% of total sales in the United States [Schaupp 2004].

According to Varadarajan and Yadav [2002], the emergence of E-Commerce has been related with major developments, including: higher richness in transactional and relational information; lower information search costs for buyers; reduction of information asymmetry between buyers and sellers; electronic proximity between buyers and sellers, and reduced time span between the moment of purchase and the moment of delivery.

However, the Internet also weakens the industry's profitability through increased competition, making it critical for companies to differentiate themselves through strategic positioning [Porter 2001]. Adequate usage of Internet technology is one of the most important contributors to leverage a company's competitive advantage, market penetration, and innovation [Torkzadeh and Dhillon 2002].

The continuing evolution of technology makes the Internet a more appealing place for doing business. Indeed, innovation and networking capabilities drive usage growth. These capabilities include broadband technology, improved user experience tools, increasingly accurate and convenient search options, personalization with more effective targeting, multi-media functionalities, payment options, wireless technology, mobile devices, improved security, and the most recent hype of local and user-generated content (movies, IM, social networking, video searches, blogs) [Morgan Stanley 2005].

Zott and Donlevy [2000] stated that companies may adopt a set of e-strategies to create value:

1. Companies can use the Internet to increase transaction efficiency by:
    a. Strengthening the supply chain by reducing supplier costs and integrating vertically;
    b. Providing a wide array of products and services;
    c. Making the transaction convenient to save time;
    d. Reducing the asymmetry of information amongst parties.
2. Companies can increase the ability to draw and retain customers to their web site by:
    a. Rewarding customers' loyalty;
    b. Personalizing products or customizing services;

c. Building virtual communities;

d. Establishing reputation enabling trust in the transaction.

Although the Internet can bring many advantages, most companies still face failure, evidenced by low customer loyalty, reduced usage and, specially, failure to reach profitability. Adequate and successful strategies have been hard to find: back in 2000, more than 50% of e-commerce ventures had gone out of business. Still today, most Internet usage is informational rather than transactional and the level of expectations regarding e-commerce has significantly decreased since 2000 [PWC 2001]. Most business models employed by e-commerce companies (other than those providing infrastructure) are not profitable [Stewart and Zhao 2000].

Despite many failures, there is no doubt that a new e-market has arisen, backed by several successful Internet ventures and, today, several Internet businesses are considered world wide successes, such as Amazon, Dell's customized PC's and E-Bay's auctions.

Many studies have depicted the drivers of a competitive strategy, while other studies analyzed a vast array of e-commerce implementation stages. However, it has not been clear on what links these two concepts. Towards avoiding the failure cases mentioned above, this paper proposes the existence of a relationship between the main drivers of a competitive strategy and the functionalities adequate for the Internet channel, based on a conceptual framework for integrating the Internet into the Competitive Strategy [Varadarajan and Yadav 2002]. Finding the organization's key drivers will help managers decide which stage of electronic commerce better suits their competitive strategy, thus reaching and sustaining profitability.

This paper is organized as follows: Section 1 presents the introduction. Section 2 deals with the e-commerce drivers in the conceptual framework for competitive strategy. In Section 3 we discuss the e-commerce implementation stages. Section 4 shows the relationship between drivers and implementation stages. Section 5 presents the conclusions and, finally, Section 6 suggests further research.

2. E-commerce drivers: conceptual framework for competitive strategy definition

The success of a company's e-commerce strategy is a consequence of the fit between its e-commerce business model and the value added by the Internet to the customers [Zeng and Reinarts 2003]. A company's ability to create value in e-commerce is a function of both how efficient its Internet business model is and how well it is able to draw customers to its web site and retain them [Zott and Donlevy 2000].

Varadarajan and Yadav [2002] developed a conceptual framework that determines the main drivers and outcomes of a marketing strategy considering that the company is competing in a broader environment represented by the Internet. This paper considers four main driver categories that will influence a company's Internet positioning: product characteristics, buyer and buying environment characteristics, firm characteristics, skills and resources and industry structure. Industry concentration, market thinness and market tippeness were not considered within industry structure because they are indirectly present on the other drivers. Since this paper focuses on the firm and its B2C strategies, macro-environment issues and individual vs. organizational issues are disregarded to focus the analysis on the core of the B2C strategy issues.

This paper extands Varadarajan and Yadav's framework [2002], adding some factors that help achieving a more complete and accurate assessment of the main drivers for a competitive strategy. Grounding on Zeng and Reinartz [2003], we introduced a set of factors, such as (i) purchase frequency, (ii) purchase risk and (iii) purchase as a process vs. purchase as an end. The first three factors refer to the buyer and buying environment characteristics, while the last one refers to product characteristics. Based on Gunasekaran and Ngai [2005], we introduced the level of Internet usage by potential consumers driver, given its relevance to the Internet implementation definition, and as well as the competition on the net driver which substituted the number of competitors and size of competitors driver since it is more related with e-strategy definition. This factor was considered as a buyer and buying environment characteristic. Figure 1 represents the resultant conceptual framework:

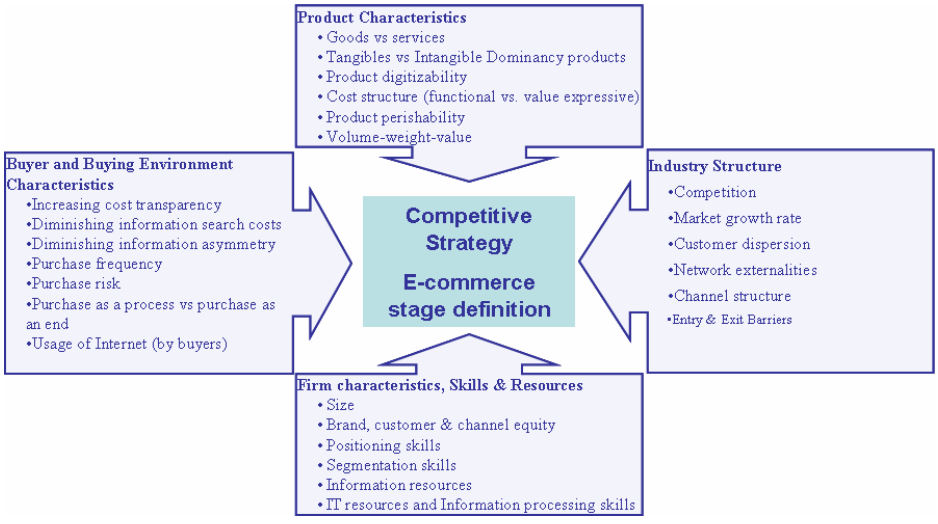


Figure 1: Drivers for integrating the Internet into Competitive Strategy: a Conceptual Framework

3. Stages of e-commerce implementation

Currently, companies face the challenge of deciding the level of investment and commitment dedicated to the Internet channel and its degree of sophistication and scope of functionalities. Executives and managers ask themselves *"How further should our Internet strategy go? What should be our positioning? What factors are key to determine the level of resources we should commit?"*

Rao et al. [2003] have defined stage as a set of descriptors that characterize the evolutionary nature of e-commerce. Belanger [2003] proposes five stages of e-commerce implementation: informational, two-way communication, transaction, integration and participation. Since interpretation and participation constitute additional functions that transaction sites may develop, we included the last two stages in the transaction stage, leaving us with three stages: Information, Two-way-communication and Transaction.

- **Stage 1 - Information:** consists in providing users with information in a content based web site. It represents the simplest type of Internet channel, which main challenge is providing information on time and accurately. Even today, the majority of Internet sites are purely informational.
- **Stage 2 - Two-way communication:** in this stage, the informational site can be coupled with additional functionalities allowing the exchange of information between the customer and the company through email, filling a pre-defined form, placing an order (without financial transactions) or setting a profile. This allows the attraction of new clients, and the possibility of engaging and retaining visitors and relating them to their individual preferences for customization purposes [Rao et al. 2003].Some examples of two-way-communication sites are: nestle.com, mercedes-benz.com, edp.pt.
- **Stage 3 - Transaction:** this stage differs from the previous one mainly by the presence of financial transaction between partners [Rao et al. 2003]. The transaction phase includes the online payment, and can also include the delivery (only possible for digitalizable products or services). Some examples of transaction sites are: fnac.com, itunes.com, barnesandoble.com, axaonline.com, tap.pt.

The stages are in ascending order in terms of complexity. The first stage is the least complex in terms of e-commerce implementation, and the last one represents the most complex stage, supporting a business that is online.

4. Relating Drivers and Stages of Internet implementation

Many researches have been made concerning the e-commerce business models classification. According to Zott and Donlevy [2000] a Business Model is the architectural configuration of the components of transaction designed to exploit business opportunities. Steward and Zhao [2000] have defined that a business model is a statement of how a firm will make money and sustain its profit stream over time. For the purpose of this study, we consider that a Business Model is part of a company's competitive strategy and determines the online and off-line functional integration extent. The relationship between e-commerce drivers, competitive strategy, e-commerce stage definition, and e-business model is graphed in Figure 2, where e-commerce stage definition is part of competitive strategy.

Figure 2: Relating Drivers with e-commerce stage definition and e-business model

Figure 3 summarizes how the drivers categories, product characteristics, buyer and buying environment characteristics, industry structure and firm characteristics, skills and resources, may influence the e-commerce stage definition. The driver's categories are likely to have different weights in influencing the selection of each e-commerce implementation stage. Besides, the relative importance of each driver may vary according to factors such as company characteristics or competitive framework.

Product characteristics and buyer environment characteristics are the most relevant categories that influence channel complexity and sophistication [Zeng and Reinartz 2003]. Of course, industry structure and firm characteristics should also be relevant, but they have less weigh when compared to the other two. However, even within a certain driver category, the factors may have different importance (see Figure 3) and different impacts on the three stages as suggested in Figure 4. The impact of each driver on e-commerce implementation stage is proposed based on the literature review. Note that the drivers that do not have an important impact (positive or negative) were considered neutral and are not weighted.

Main Drivers	Stage 1: Information	Stage 2: Two-way communication	Stage 3: Transaction
Product Characteristics	♠ Search is a phase of consumer decision making process ♠ Important for high cost products ♠ Important for high value goods/ services ♠ Not very important for "high touch" goods ♠ Important for products with high information content	♠ Interaction is key ♠ Important at evaluation stage of decision making process ♠ Important for products with high information content ♠ Important when expertise is needed to make a decision	♠ If product is digital - delivery can be online ♠ If product is not perishable - logistics is not a problem ♠ Payment may be digital ♠ Services can be delivered online ♠ Buying as an end ♠ Cost inefficient for high volume or weight products
Buyer and Buying Environment Characteristics	♠ Low information search cost ♠ Low information asymmetry ♠ Not relevant for frequent purchases ♠ Relevant for low purchase risks	♠ Communication model with consumers ♠ Very relevant for high purchase risks	♠ Low cost transparency ♠ Internet usage by target customers ♠ Important that consumers have high Internet usage ♠ If purchase is an enjoyable process in traditional commerce it is less possible to be done online
Industry Structure	♠ Internet presence as a distinction factor or a survival factor	♠ Build a more close relationship with consumers	♠ Customer dispersion has less impact for business ♠ Market is growing ♠ Channel structure is key
Firm characteristics, Skills & Resources	♠ Less IT resources needed ♠ Brand positioning is key	♠ Medium IT resources needed ♠ Company has good information resources - expertise provider	♠ High IT resources needed ♠ Medium-high size companies ♠ Strong brand can be an advantage ♠ Important to have a correct market segmentation - possible niche opportunity

Figure 3: Impact of drivers on e-commerce stage

In Figure 4, some of the drivers have varying influence along the three e-commerce stages, while others have the same impact on the three stages. One driver may be a facilitator (it promotes a positive impact) of a particular stage, while the same driver could also be a barrier (it promotes negative impact) on another stage. On Figure 4, the full circle represents a very positive influence on a particular stage, the half full circle represents a positive impact on a stage, an empty circle represents a negative impact on the stage and NA (Not Available) represents that no information was found to support the driver's impact on the stage. The most adequate implementation stage of an organization will also depend on the relative weight that drivers have when comparing with each other.

		Stage 1	Stage 2	Stage 3
Product Characteristic				
Typology[1]	Goods	-	-	-
	Services	-	-	-
Digitizability[1,2]	Digitalizable	◐	◐	●
	Non-digitalizable	◐	◐	○
Cost structure[2]	High cost	◐	●	○
	Low cost	○	◐	◐
Tangibility[2]	Tangible (low information content)	-	-	-
	Non-tangible (high info. content)	●	●	●
Perishability[2]	Perishable	○	○	○
	Non-perishable	-	-	-
Low value/weight or value/volume ratio[5]		-	-	○
Buyer and Buying Environment characteristics				
Cost transparency[1]	High	◐	◐	●
	Low	○	○	○
Diminishing information search cost[3]		●	◐	◐
Diminishing information asymmetry[1]		●	◐	◐
Purchase frequency[2]	High	○	○	●
	Low	◐	◐	-
Purchase Risk[2]	High	●	●	○
	Low	○	-	●
Purchase engagement[2,6]	Purchase as a process	○	○	○
	Purchase as an end	◐	◐	●
Internet usage (by buyers) [3,4]	High	●	●	●
	Low	○	○	○
Industry structure				
Competition on the Internet[4]	Is on the Internet	●	●	●
	Is not on the Internet	-	-	-
Market growth rate[5]	Positive	●	●	●
	Negative	-	-	-
Customer dispersion[4]	High	●	●	●
	Low	-	-	-
Network externalities[3]		◐	◐	●
Channel structure – high concentration at traditional intermediary level[5]		○	○	○
Entry barriers[1]	High	◐	◐	●
	Low	-	-	-

Figure 4: Drivers' impact on e-commerce stages

Firm characteristics, skills and resources				
Size[1]	Large	◐	◐	●
	Small	◐	◐	○
Brand, customer and channel equity[1]		◐	◐	●
Positioning skills[3]	Strong	●	●	●
	Weak	○	○	○
Segmentation skills[3]	Strong	●	●	●
	Weak	○	○	○
Information resources[3]	Strong	◐	●	●
	Weak	○	○	○
IT resources and information processing skills[3,4]	Strong	◐	◐	●
	Weak	○	○	○

5. Conclusions

Based on the literature review, an enhanced conceptual framework for integrating Internet into the competitive strategy and the e-commerce implementation stages were defined. The relationship between these two concepts was then represented on Figure 2, showing that drivers influence competitive strategy and stage definitions, and the competitive strategy and stage influence e-business model determination. The explanation of drivers' impact on e-commerce stage is graphed on Figure 3 and Figure 4, illustrating some examples of drivers' influence and the impact's type (if it is positive or negative) on a particular stage.

This study has evaluated the relative importance of those drivers when comparing to each other, and helped determine the most relevant factors. It contributes to research on E-commerce, since it links two concepts that are very important in this context: drivers for competitive strategy and stages for e-commerce implementation. For managers, it proposes a way to validate Internet competitive strategy by finding the most adequate stage for the e-commerce implementation. For academics, this paper proposes a relationship between drivers for a competitive Internet strategy and the complexity and sophistication stages of e-commerce adoption and implementation as shown in Figure 2.

[1] [Stewart and Zhao 2000]
[2] [Zeng and Reinartz 2003]
[3] [Rao et al. 2003]
[4] [Gunasekaran and Ngai 2005]
[5] [Varadarajan and Yadav 2002]
[6] [Dennis et al. 2002]

6. Further Research

Further research will verify the assumptions in Figure 3 and 4 through a survey aimed at marketing professionals. The survey will ask what are the two main driver categories that impact the most e-commerce stage definition. Accordingly, we will then present to the respondent the drivers for each of the two categories selected and respondents will be asked to pick 3 of each and to distribute them in a 100 points scale. Based on those 6 drivers (3 drivers for 2 categories), we will build hypothetical scenarios (combinations of drivers), and respondents will be asked to pinpoint which development stage would be more appropriate.

7. References

Belanger, F. (2004). "A Framework for e-government: privacy implications".

Chuang, M. and Shaw, W. (2005). "A Roadmap for E-Business Implementation", Engineering Management Journal; 17, 2

Datamonitor (2005).

Dennis, C., Harris, L. and Sandhu, B. (2002). "From bricks to clicks: understanding the e-consumer", Qualitative Market Research, 5, 4, pg 281-290.

Gunasekaran, A. and Ngai, E. (2005). "E-commerce in Hong Kong: an empirical perspective and analysis", Internet Research, 15, 2

Meeker, M. (2005). "Global technology and Internet trends", Morgan Stanley.

Porter, M. (2001). "Strategy and the Internet", Harvard Business Review 79(3), Mar., pp.63-78.

Rao, S., Metts, G. and Monge C. [2003]. "Electronic commerce development in small and medium sized enterprises", Business Process Management Journal, 9, 1.

Schaupp, L. and Belanger, F. (2004). "A conjoint analysis of online consumer satisfaction".

Stewart, D. and Zhao, Q. (2000). "Internet Marketing, Business Models, and Public Policy", Journal of Public Policy & Marketing, 19 (2), pp 287-296.

Varadarajan, P. e Yadav, M. (2002). "Marketing Strategy and the Internet: An Organizing Framework" Journal of the Academy of Marketing Science, 30(4), pp. 296-312.

Vernis Suhder Stevenson, "Consumer Internet" [2005]

Zeng, M. e Reinartz, W. (2003). “Beyond Online Search: The Road to Profitability” California Management Review, 45(2), pp. 107-130.

Zott, C. and Donlevy, J. (2000). “Strategies for value creation in e-commerce: Best Practice in Europe”, INSEAD R&D, 2000/61/ENT/3i-5

TEMA 3 - A SOCIEDADE E OS SISTEMAS DE INFORMAÇÃO

3.1. Ensino e Aprendizagem de Sistemas de Informação

Sistemas de Informação nas Pequenas e Médias Empresas

Cristiana de Freitas

Universidade do Minho, Guimarães, Portugal

cristianafreitas@iol.pt

Resumo

A partir de uma breve discussão sobre o perfil de um profissional em Sistemas de Informação e da relevância da construção de uma visão global dos processos organizacionais, é descrito um caso de estudo da intervenção de um profissional em Sistemas de Informação numa pequena e média empresa. Esta discussão serve de base para perceber a diversidade de papéis que um profissional em sistemas de Informação pode exercer. O artigo inclui um caso de estudo que descreve a intervenção, deste profissional, num pequena e média empresa. O artigo conclui argumentando a importância das pequenas e médias empresas contratarem para os quadros da empresa profissionais desta área.

Palavras chave: processos

1. Introdução

Pretende-se com este artigo descrever o perfil de um profissional em Sistemas de Informação. A necessidade deste esclarecimento deve-se à falta de conhecimento, por parte das pequenas e médias empresas, das vantagens competitivas que podem obter pela utilização das tecnologias de informação. Desta forma, este profissional tem o perfil adequado para combater o insucesso dos sistemas de informação nestas organizações.

Pretende-se também demonstrar que é relevante a construção de uma visão global dos processos organizacionais. Um processo pode percorrer vários departamentos dentro de uma organização, é necessário perceber quais os processos críticos da empresa, de forma a melhorar a sua eficiência.

Assim, neste artigo será apresentado um caso de estudo da intervenção de um profissional em Sistemas de Informação numa pequena e média empresa.

2. Perfil de um profissional em Sistemas de Informação

Este capítulo procura esclarecer quais são as funções que um profissional em Sistemas de Informação poderá desempenhar dentro de uma organização.

"O principal objectivo de um licenciado em Informática de Gestão é a capacidade em ajudar as organizações a tirar partido das Tecnologias e dos Sistemas de Informação para responder às exigências dos mercados globais e aos desafios relacionados com a adopção de novos modelos de gestão. São profissionais aptos a assumir a liderança de processos de melhoria organizacional ancorada nos Sistemas de Informação e nas Tecnologias da Informação.

Para tal, o mesmo deverá estar apto a utilizar um conjunto de métodos, técnicas e ferramentas, que lhe permitam optimizar as suas funções, com um elevado know-how tecnológico, e um conjunto de conhecimentos básicos da gestão corrente das organizações.

Segundo Peter Keen, os cursos de Informática de Gestão são descritos como híbridos, resultantes da gestão e da informática, o que possibilita ao formando uma trajectória profissional em cada um destes domínios." [DSI, 2006]

Fig. 1 - Híbridos [Keen, citado por Coelho]

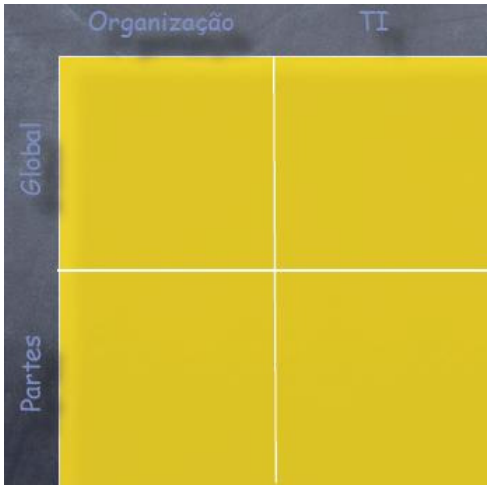


Fig. 2 - Híbridos [Carvalho e Amaral]

Um profissional de Sistemas de Informação poderá desempenhar cargos como: auditor informático, consultor de sistemas e tecnologias de informação, gestor de sistemas de informação, director de informática e analista de sistemas.

Muito aproximado ao perfil de um profissional de Sistemas de Informação encontra-se o CIO (*chief information officer*).

O CIO é um executivo que gere os recursos de TI de forma a implementar a estratégia empresarial: fornece uma visão da tecnologia e liderança para desenvolver e implementar iniciativas de TI para criar e permitir que a empresa mantenha a liderança num mercado intensamente competitivo que está em constante mudança.

Os CIO têm uma forte noção de negócio e da tecnologia. Em muitas organizações desempenham papéis que abarcam ambas as áreas. Frequentemente é solicitado aos CIO desempenhar papéis estratégicos numa parte do dia e papéis operacionais noutras, invés de ocupar o seu tempo apenas a um só papel.

Peter e Saunders identificaram doze responsabilidades principais definem frequentemente o papel do CIO [Pearlson e Saunders, 2004]:

1. *Defender a organização*. Defender as TI na empresa como uma ferramenta estratégica para o seu crescimento.
2. *Gestão da arquitectura*. Estabelecer prioridades e direcção organizacionais.
3. *Consultor de estratégia comercial*. Participar em tomadas de decisão ao nível executivo.
4. *Planeamento de tecnologia comercial*. Estabelecer grupos de negócio e de tecnologia para colaborarem no planeamento e na execução.

5. *Desenvolvimento de aplicações*. Vigiar as iniciativas das empresas legadas e das que emergem, bem como as iniciativas divisionárias e da unidade de negócio estratégico (SBU – strategic business unit)
6. *Gestão da infastrutura das TI*. Por exemplo computadores, impressoras e redes. Manter as tecnologias actuais e investir em tecnologias futuras.
7. *Estabelecer a fonte*. Desenvolver e implementar uma estratégia de pagar a alguém para fornecer a fonte (por oposição à própria empresa tê-la) dos serviços e pessoal das TI.
8. *Desenvolver parcerias*. Negociar relações com fornecedores-chave de conhecimentos técnicos e de serviços de TI.
9. *Agente de transferência da tecnologia*. Fornecer tecnologias que permitem à empresa trabalhar melhor com os fornecedores e com os clientes – tanto internos como externos – e consequentemente aumentar o valor das acções.
10. *Gestão da satisfação do cliente*. Entender e comunicar tanto com os clientes internos como os externos de modo a assegurar a satisfação do cliente.
11. *Formação*. Fornecer formação aos utilizadores das TI, bem como aos executivos superiores que têm de entender como é que as TI se encaixam na estratégia empresarial.
12. *Planeamento de descontinuidade comercial/recuperação de tragédias*. Planear e implementar estratégias para limitar o impacto de tragédias naturais e provocadas pelo homem nas TI e, consequentemente, a condução do negócio.

Um CIO deve trabalhar eficazmente não apenas dentro da área técnica, mas também no que diz respeito à gestão do negócio em geral. Este conjunto de competências único exige um background especializado. As competências que se seguem são consideradas essenciais para o sucesso como CIO [Pearlson e Saunders, 2004]:

1. Uma forte orientação em relação ao negócio numa empresa, indústria, ou através de actividades relacionadas, tais como consultoria.
2. Capacidade de se aperceber dos benefícios e de gerir o custo e riscos associados com as TI.
3. Capacidade de aproximar quaisquer falhas entre as tecnologias disponíveis e as necessidades comerciais.
4. Familiaridade com as necessidades dos clientes internos não-técnicos.
5. Capacidades organizacionais fortes para gerir recursos dos SI e aplicações localizados bem como SBU mais latos e iniciativas e recursos divisionários.
6. Capacidade de conceber, construir e implementar projectos de TI múltiplos a tempo e dentro do orçamento.
7. Capacidade de articular e defender uma visão de gestão das TI.
8. Capacidade de lidar bem com a estrutura de gestão existente.
9. Uma visão estratégica para a empresa que vai para além das TI.

3. Construção de uma Visão Global dos Processos Organizacionais

É razoável que um CIO tenha uma visão global dos processos organizacionais com vista ao sucesso dos sistemas de informação.

"Muitas vezes adopta-se o organograma como referencial orientador para delimitar o âmbito das aplicações informáticas e para o planeamento dos projectos. Acontece que os organogramas dizem quem manda em quem e em que área. São os modelos de processos que dizem o que se tem de fazer e como se faz. O cruzamento das duas representações diz-nos quem é responsável

por quais actividades. (...), o conhecimento dos fluxos horizontais do trabalho e da informação é essencial para a construção de qualquer tipo de sistema de apoio à gestão, como por exemplo, o processo de desenvolvimento de um novo produto que poderá envolver os departamentos de marketing, de produção e financeiro no caso de uma empresa industrial." [Coelho, 2005]

A especificação dos sistemas de informação deverá então decorrer das necessidades dos processos de negócio, na procura constante de concretizar a estratégia da organização.

É necessário que os sistemas de informação contribuam adequadamente para o desempenho dos processos organizacionais e na sua interacção com outras organizações, para que sejam um instrumento potenciador do desempenho e diferenciação das organizações e assim obter vantagens competitivas.

Não só a dependência na estrutura orgânica, mas também o levantamento dos requisitos por entrevistas ao utilizador e a atitude dos informáticos são algumas das atitudes que potenciam o insucesso dos sistemas de informação nas organizações.

Um processo pode percorrer vários departamentos dentro de uma organização. Ao fazer o levantamento de requisitos por entrevistas ao utilizador, o informático deparar-se-á com diversas opiniões ao longo do mesmo processo, pois cada utilizador tem uma visão restrita do problema e desconhece as potencialidades das tecnologias de informação. São por estas razões que os requisitos devem ser identificados para o processo pelo seu desempenho como um todo e não pelo desempenho de cada utilizador.

Um informático deve ter uma atitude mais pró-activa em relação aos pedidos do cliente. Deve questionar o pedido e tentar responder da melhor forma, para se evitar frases tais como: "Ele nunca resolve o meu problema!" e "Eu fiz exactamente o que me pediu!".

"Para que o informático possa questionar o utilizador terá de ter credibilidade em termos do seu conhecimento da organização. Isto implica que não se pode misturar dois perfis (...): o profissional de sistemas de informação e o profissional de tecnologias de informação, num perfil único designado informático. O primeiro terá de ser conhecedor da organização e seus processos e das potencialidades das tecnologias de informação. O segundo terá de conhecer bem as tecnologias de informação." [Coelho, 2005]

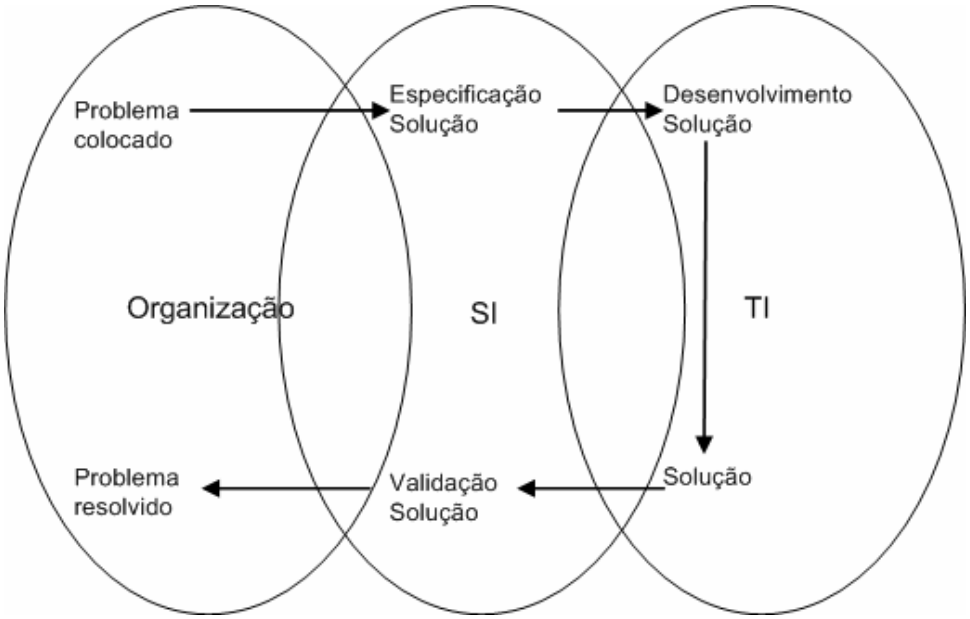


Fig. 3 - O importante papel do profissional SI [Coelho, 2005]

4. Intervenção numa PME

A PME escolhida para este caso de estudo é uma empresa nacional que fabrica e comercializa material ortopédico e hospitalar.

Ao terem-se apercebido das dificuldades em garantir o sucesso do seu sistema de informação, a gestão de topo desta empresa resolveu convidar pela primeira vez, para estágio curricular, alguém da área de sistemas de informação.

Melhorar o desempenho do sistema de informação da empresa com vista a tornar mais eficaz o seu sistema organizacional através da utilização de sistemas informáticos foi o objectivo proposto.

Mediante uma análise geral da empresa, mas sobretudo pela construção global dos seus processos organizacionais, foram identificadas algumas limitações no desempenho do seu sistema de informação. Para cada uma destas limitações identificadas, foi realizada uma análise cuidada e foram realizadas propostas para a sua resolução.

L.	Descrição	Solução	Custos	Benefícios
1. Segurança informática	Entrada acessível por qualquer indivíduo, via rede wireless, a toda a informação da empresa, permitindo também o acesso ao servidor. Esta ficava sujeita a leitura e modificação por terceiros. O backup da informação não era realizado regularmente. Má utilização do sistema informático por parte dos utilizadores. Pouco controlo sobre o software instalado em cada posto de trabalho. Software instalado sem utilização por parte do colaborador do respectivo posto (software não produtivo), provocando sobrecarga no disco. Necessidade de estabelecer uma política de conduta do sistema informático para que o seu uso seja estritamente profissional e daí tirar o máximo partido das suas funcionalidades.	Instalação de firewalls e anti-virus, protecção da rede wireless com segurança WEP. Levantamento das necessidades de cada posto de trabalho (com computador), instalação apenas dos software necessários a cada posto, contas limitadas e uma conta administrador protegidas com passwords (conta administrador acessível somente pelo administrador do sistema informático); Sistema de backup diário; Redigir um plano de políticas e normas do sistema informático, sensibilizar e formar os utilizadores para as boas práticas.	Tempo de análise e implementação das soluções propostas.	Maior segurança na rede e por consequência os dados que circulam na rede estão também mais protegidos. Só tem acesso à rede aquele indivíduo que obtenha autorização para tal. Haverá menor impacto negativo com a ocorrência de incidentes no sistema, pois os sistemas antivírus são preparados para detectar e reagir e os backups são actualizados diariamente. Facilidade na identificação de eventuais problemas de rede. Cada posto tem instalado apenas o software de que necessita, retirando sobrecarga ao disco. Maior respeito e melhor utilização do sistema informático.

2. Gestão de stocks	Compra-se material pensando que já não existe em stock, mas que por vezes não corresponde à verdade, aumentando desta forma custos com as compras. Demasiado tempo despendido para confirmar ao cliente se têm em stock determinado produto.	Criação de documentos no software Primavera mais adaptados à situação concreta da empresa. Formação aos colaboradores acerca dos novos métodos a proceder. Criação de perfis no software Primavera para que os colaboradores trabalhem apenas com os documentos de que necessitam, evitando desta forma a utilização de documentos de forma errada. Utilização do código de barras e de um dispositivo de leitura de código de barras de forma a reduzir erros humanos.	Custo baixo, se optar pelo leitor de código de barras, que no contexto da empresa consegue responder ás suas necessidades (aproximadamente 300€).	Gestão dos stocks mais eficaz e diminuição do tempo de resposta ao cliente.
3. Parametrização do sw ERP	Necessidade de certas informações que só não obtêm por desconhecimento de determinadas funcionalidades das aplicações existentes. Por exemplo, o Rappel, que é um relatório mensal que indica as vendas a determinado cliente agrupado por famílias de produto, é realizado da seguinte forma: exportação de dados da Primavera em cinco ficheiros Excel, construção de um ficheiro Excel de um mix dos outros cinco ficheiros e através de algumas fórmulas é construída a folha Rappel. Muito tempo despendido para executar esta operação. O risco de dados distorcidos é elevado. Muitas operações são feitas paralelamente ao software ERP.	Formação aos colaboradores da forma como devem adquirir determinados relatórios, sejam estes de vendas, compras, etc. Parametrização do software ERP ao contexto da empresa. Para o exemplo em concreto, o software Primavera permite a impressão de relatórios Crystal Reports. Desta forma, é possível a criação de relatórios, em Crystal Reports, que geram automaticamente a folha Rappel e outros relatórios de gestão.	Tempo de análise; Compra do software Crystal Reports para permitir a edição/construção de relatórios (aproximadamente 179€); elaboração de relatórios.	Menor tempo despendido e maior fiabilidade dos dados, concentração dos dados a um só software.
4. Terminal Expedição	As facturas devem ser separadas mediante o terminal de expedição que seguem. Muito tempo dispendido na procura do terminal correspondente a determinada expedição. Risco elevado de erro humano.	Na base de dados do ERP, existe uma ligação da tabela “Clientes” com uma tabela “Zonas”. Como cada zona corresponde um terminal, a solução passa pela criação de uma tabela designada “Terminal” com ligação à tabela “Zonas”. Desta forma adiciona-se o campo “Terminal” na tabela “Zona” que fará a ligação à tabela “Terminal”. Nas facturas, indicar qual o terminal referente à expedição, apresentando simplesmente o campo “Terminal” da tabela “Zonas”.	Tempo de análise e implementação das soluções propostas, tendo também em conta a compra anterior do Crystal Reports, para edição das facturas.	Diminuir o tempo despendido na procura do terminal e o risco de erro humano.

5. Módulo Expedição	A Ortomaia subcontrata os serviços de expedição a uma empresa especializada que disponibiliza um software para a impressão de etiquetas. Esse software para impressão de etiquetas não está integrado com o software ERP da empresa, inclusivé utiliza uma base de dados diferente, o que obriga a executar dois processos separadamente. Primeiro processo é a facturação da mercadoria, o outro é o registo da mesma informação no software de expedição para que assim se possa proceder à impressão da etiqueta. Esta situação aumento o tempo de seguimento da expedição.	Cruzamento da informação das duas base de dados e integração no software ERP. Desenvolvimento de um módulo para a Expedição, na linguagem de programação Visual Basic 6. Isto permitirá também diminuir a sobrecarga de trabalho a um colaborador da empresa, pois a impressão das etiquetas poderá ser feita pelo responsável do armazém.	Implementação de um módulo, para a impressão das etiquetas de expedição que cruza os dados das duas aplicações.	Redução do tempo para dar seguimento à expedição. Diminuição do custo com papel de etiquetas (+/- em 90%).
6. Portal web	Site meramente institucional, com uma pequena apresentação da empresa e de alguns artigos que produzem/comercializam desactualizados. Os vendedores têm de se deslocar à empresa para actualizar a informação.	Desenvolver/actualizar o site. Internet, Intranet e Extranet.	Implementação de módulos, Cliente e Colaborador, em PHP. Aquisição de terminais móveis com acesso a rede para suporte aos vendedores (aproximadamente 40€/mês por vendedor).	Promoção da imagem da empresa; apoio aos colaboradores que pelo facto de não estarem constantemente no interior da empresa estão privados ao acesso de certas informações; apoio aos clientes disponibilizando-lhes informação necessária sobre os produtos (características, inovações técnicas, etc) a possibilidade de fazer encomendas, enviar questões/sugestões/reclamações, etc.
7. Sistema VOIP	Utilização constante do telefone. É gasto um elevado valor em chamadas diariamente.	Adesão a um sistema de VOIP, por exemplo o Skype.	Tempo de análise e configuração do sistema de comunicação; Aquisição de Gateway VOIP (aproximadamente 600€).	Redução bastante alta dos gastos em chamadas (+/- 60%).

Tabela 1 – Análise Custo-Benefícios

Como se pode verificar, as soluções propostas são bastante simples, algumas passam apenas pela indicação de que a informação que a empresa pretende adquirir já é abrangido pelo seu sistema ERP. Desta forma, muitas das tarefas realizadas consistiram na realização de formação aos utilizadores das TI.

Esta empresa, adquiriu um software ERP standard, recebeu uma formação básica e não foi realizada uma parametrização do software ao seu negócio e necessidades. Consequentemente, tentam colmatar a falta de conhecimento, das funcionalidades do seu software, através da utilização de ferramentas como MS Excel e MS Access o que provoca perca de tempo,

duplicação de informação e aumenta o risco de erros. Diminuem desta forma a possibilidade de adquirirem vantagem competitiva pela utilização das TI.

No desenvolvimento destes projectos, verifica-se que foram executados alguns perfis de um profissional de Sistemas de Informação, tais como, gestor da infra-estrutura das TI (projecto 1), planeador de descontinuidade comercial/recuperação de tragédias (projecto 1), gestor da satisfação do cliente (projecto 2), formador (projectos 2, 3, 4, 5 e 6), desenvolvimento de aplicações (projecto 5 e 6), análise de requisitos (projecto 6), agente de transferência de tecnologia (projecto 7) e estabelecedor da fonte (projecto 7).

Neste caso de estudo, a empresa não tinha nenhum informático, por essa razão, o profissional em Sistemas de Informação por vezes alarga o seu papel para desempenhar tarefas de um informático.

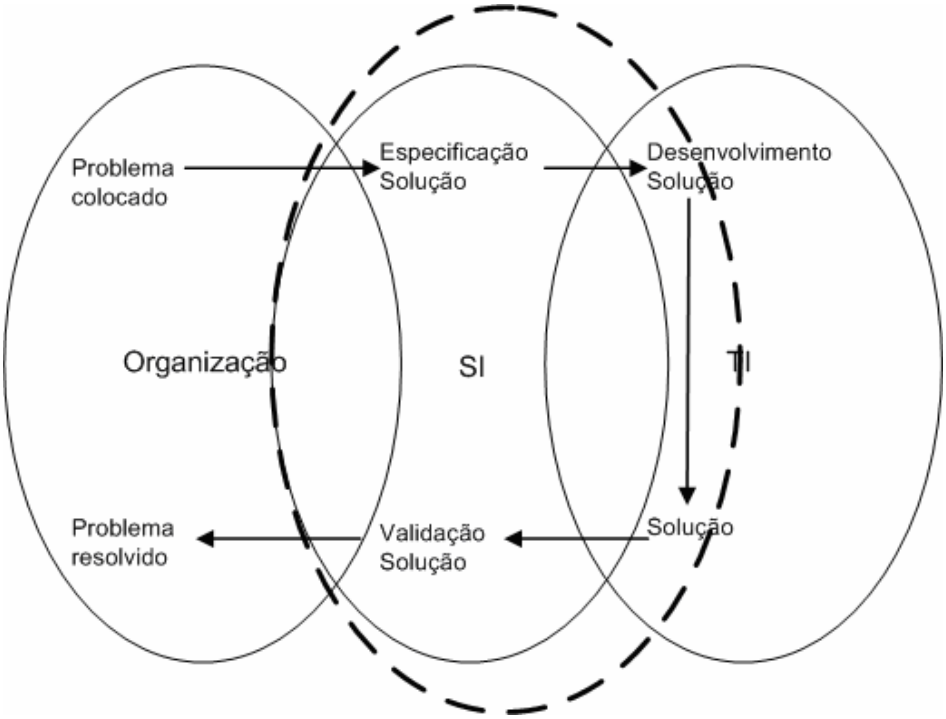


Fig. 4 - O importante papel do profissional SI Adaptado [Coelho, 2005]

5. Relevância de um profissional em Sistemas de Informação numa PME

Apesar da forte evolução das tecnologias de informação, continua-se a assistir a grandes dificuldades por parte das pequenas e médias empresas em garantir o sucesso dos Sistemas de Informação. Em vários casos, isto pode-se dever ao receio que as PME ainda têm em investir na criação de um novo departamento, o departamento de sistemas de informação.

A empresa deste caso de estudo, subcontrata os serviços de consultoria informática. Porque razão decidiu agora contratar um profissional em Sistemas de Informação? Fazer consultoria informática implica o envolvimento dos quadros da empresa e ter uma visão global dos seus processos organizacionais. Esta tarefa é muito árdua tanto para os consultores como para os quadros da empresa. Para os consultores porque num período reduzido de tempo devem perceber o "core business" da empresa e os seus processos organizacionais, e para os quadros da empresa porque têm de dispensar algum do seu tempo.

A solução encontrada neste caso de estudo foi a contratação de um profissional em Sistemas de Informação, que tem uma visão global dos processos organizacionais da empresa e percebe o seu "core business". Este profissional está lá, ouve, experimenta, e sente as dificuldades.

Com o saber deste profissional, e eventualmente em conjunto com as consultorias informáticas, é possível melhorar o desempenho do sistema de informação da empresa.

Segundo a gestão de topo, deste caso de estudo: "Durante anos estivemos dependentes das consultorias de informática, gastando imensos recursos financeiros e com resultados que não respondiam ás nossas necessidades e não solucionavam todas as nossas limitações. Com um profissional de sistemas de informação dentro da empresa, é possível tirar melhor partido do nosso sistema de informação. Temos pena não termos tomado esta decisão mais cedo!".

Um profissional em Sistemas de Informação tem um bom perfil para solucionar este insucesso das PME's.

6. Referências

Amaral, L.A.M.d. "PRAXIS - Um Referencial para o Planeamento de Sistemas de Informação", in: Escola de Engenharia, Universidade do Minho, Braga, 1994.

Carvalho, Prof. Doutor João Álvaro, apontamentos de Sistemas de Informação III, 3º ano, Universidade do Minho, 2004/2005: Modelo do processo de desenvolvimento de sistemas de informação com indicação dos resultados de cada uma das actividades.

Coelho, J. S., "Arquitectura da empresa centrada nos processos: o factor determinante para o alinhamento estratégico dos SI", in: Sistemas de Informação Organizacionais, Lisboa: Edições Silabo, 2005 - p.141-197

DSI, Objectivos, 2006, Disponível em: http://lig.dsi.uminho.pt/, Acedido em: Agosto de 2006

Pearlson, K.E., and Saunders, "C.S. Managing and Using Information Systems", (2nd ed.) John Wiley & Sons, 2004.

Uma Aplicação Web para Gestão de Estilos de Aprendizagem

Rosa Pereira [1], António Andrade [2]

1) Escola Secundária Ferreira de Castro, Oliveira de Azeméis, Portugal

rosapereira@esfcastro.pt

2) Universidade Católica Portuguesa, Porto, Portugal
aandrade@porto.ucp.pt

Resumo

O caminho da globalização no acesso e na utilização das Tecnologias da Informação e Comunicação, em particular a Internet, é um factor de mudança na Escola, com implicações a vários níveis, mas com especial e desejado impacto no processo de aprendizagem. Na escolha dos recursos educativos ou construção de conteúdos, na selecção das metodologias e das estratégias de ensino, os docentes devem ter em conta os estilos de aprendizagem dos alunos, e, estes, por sua vez, devem reflectir sobre esse conhecimento com vista a saber como aprender melhor. No entanto, a superabundância de teorias, modelos e instrumentos de diagnóstico de estilos de aprendizagem é um dos entraves à sua aplicação prática. Neste artigo é apresentado o projecto de uma aplicação, desenhada para a Web, cujo objectivo é permitir aos utilizadores a gestão de estilos de aprendizagem, a partir da informação automaticamente obtida dos questionários informatizados: o Index of Learning Styles Questionnaire (ILS) e o Questionário Honey-Mumford de Estilo de Aprendizagem – Learning Styles Questionnaire (LSQ).

Palavras chave: estilos de aprendizagem, aplicação Web

1. Introdução

No âmbito da Sociedade da Informação e do Conhecimento, várias iniciativas visam a modernização da escola, promovendo a generalização do acesso e utilização das Tecnologias da Informação e Comunicação (TIC), para que os alunos adquiram mais e melhores qualificações e competências.

Neste mesmo contexto, é incentivada a criação e utilização de conteúdos educativos digitais e de plataformas de *e-learning*, não só em situações de ensino/formação a distância, mas sobretudo como ambientes de apoio a metodologias de *b-learning*, baseadas em experiências de aprendizagem activas, centradas nos alunos e na construção do seu conhecimento.

Apesar de se reconhecer o potencial educativo da Internet, especialmente da World Wide Web (ou simplesmente, Web) como fonte de informação, e o potencial produtivo e criativo das ferramentas TIC, em geral, deve agir-se com precaução pois, só por si, essas ferramentas não garantem uma melhoria nem no processo, nem nos resultados da aprendizagem.

Não obstante, a aprendizagem electrónica permite a utilização de ferramentas que dificilmente podem ser aplicadas numa sala de aula, a partir dos métodos tradicionais, como é o caso dos estilos de aprendizagem, com resultados comprovados na melhoria da qualidade da aprendizagem, de forma mais rápida e mais eficaz. A personalização dos conteúdos, em função

dos estilos de aprendizagem dos alunos, resulta numa melhoria do rendimento escolar, bem como, num aumento do grau de satisfação dos alunos [Rodríguez & Caro 2003].

Os docentes têm a responsabilidade de compreender a diversidade cognitiva, emocional e cultural dos seus alunos e de adaptar as estratégias de ensino a essa diversidade [Miller 2001]. Por sua vez, os alunos devem reflectir sobre as suas qualidades e fraquezas e usar essa informação para obter melhores resultados escolares [Fleming & Mills 1992].

A utilização das TIC na aprendizagem permite adequar a forma de apresentar os conteúdos e conduzir as experiências de aprendizagem, às características individuais dos alunos, bastando, para isso, disponibilizar a informação, em diferentes formatos de media e desenhar percursos de aprendizagem diferenciados, tirando partido de plataformas LMS (Learning Management System) e conteúdos SCORM (Sharable Content Object Reference Model).

Assim, o processo de aprendizagem pode ser mais ajustado às características individuais dos alunos e da turma, e é em torno desta ideia que ganham novo impulso as teorias de estilos de aprendizagem [Gordon & Bull 2004].

O objectivo deste artigo é apresentar o projecto de desenvolvimento de uma aplicação, desenhada e implementada para a Web, que permite determinar automaticamente os estilos de aprendizagem dos utilizadores, segundo dois dos questionários mais utilizados e testados, mas não sem antes justificar a escolha desses questionários, numa abordagem resumida às teorias e modelos que lhe dão enquadramento e das indicações dos seus autores de como gerir essa informação para conseguir melhores aprendizagens.

2. Teorias, Modelos e Instrumentos de Estilos de Aprendizagem

Há mais de quarenta anos que se realiza investigação na área dos estilos de aprendizagem. Nos últimos anos, a pesquisa tem sido mais intensa e verifica-se uma emersão do conceito em diversas disciplinas, justificado pela centralidade que a aprendizagem têm em praticamente todos os aspectos ao longo da vida [Cassidy 2004].

3. Categorias de Modelos de Estilos de Aprendizagem

O enquadramento teórico sobre os estilos de aprendizagem é fragmentado, extenso e não consensual, o que constitui o maior entrave à sua aplicação prática. É possível encontrar, na literatura sobre o tema, uma variedade de definições de "estilo de aprendizagem", a par da utilização dos termos "estilo cognitivo", "estratégias de aprendizagem" ou "preferência de aprendizagem", com contornos pouco precisos.

O Modelo de Curry (1983, 1987), também conhecido por "The Onion Model", é um metamodelo que, a partir da análise das várias teorias e modelos, entende que cada um deles procura enquadrar diferentes dimensões de características que interferem no processo de aprendizagem.

Curry organizou os modelos de aprendizagem em quatro camadas: Modelos de Personalidade, Modelos de Processamento de Informação, Modelos de Interacção Social e Modelos de Preferências Educativas, representadas, por esta ordem, do interior para o exterior de uma cebola. Com esta metáfora, fica explicita a ideia de que as características posicionadas no exterior da cebola estão mais sujeitas às interacções com o exterior, enquanto as características da camada mais interna são muito menos sujeitas a alterações.

Uma forma alternativa de categorizar os modelos de estilos de aprendizagem é o Metamodelo dos Quatro Quadrantes de Gordon-Bull. Com base na comparação entre vários modelos, Gordan e Bull [2004] conseguiram identificar fortes equivalências para formar um "núcleo" de dimensões fundamentais: Estilo α – Sequenciais e Práticos, Estilo β – Intrapessoais e

Argumentadores, Estilo γ – Imaginativos e Globais e Estilo δ – Lógicos e Analistas, circunscritos por uma categoria que designam por Resistentes, desde que o modelo reconheça que os estilos de aprendizagem mudam ao longo do tempo e com respeito à tarefa que está a ser realizada.

Em ambientes de aprendizagem *online*, as razões que colocam um aluno na categoria dos Resistentes são: a falta de competências TIC para interagir efectivamente com os materiais *e-learning*, ou a falta de autoconfiança nas suas competências TIC, ou ainda, por se encontrarem frustrados, apáticos, desencorajados ou desobedientes [Martinez 2000, *cit in* Gordon & Bull 2004].

Estas e outras taxionomias de modelos de estilos de aprendizagem tentaram simplificar e organizar o acervo de modelos e instrumentos. No entanto, como refere Cassidy [2004:440], no final de uma extensa, mas ainda assim não exaustiva, revisão de literatura:

> *"For those working within an educational setting wishing to utilise learning style to promote more effective learning, whether through individual or group profiling, design of instructional methods, or identifying learner preferences, operationalising learning style is a necessary but highly problematic endeavour."*

Na escolha de um modelo ou de um instrumento, De Bello [1990, *cit in* Cassidy 2004:440] sugere que se avalie, não só a validade e fiabilidade, mas também, a extensão na qual o modelo foi previamente aplicado.

4. Instrumentos de Estilos de Aprendizagem

Este projecto disponibiliza dois instrumentos sujeitos a numerosos testes de validade e fiabilidade, tendo em conta o alcance dos modelos a que estão associados: o Index of Learning Styles Questionnaire (ILS) e o Questionário Honey-Mumford de Estilo de Aprendizagem – Learning Styles Questionnaire (LSQ).

O ILS [Felder & Soloman 1991] é um instrumento para avaliar estilos de aprendizagem nas quatro escalas do Modelo de Felder-Silverman: Activo/Reflexivo, Sensorial/Intuitivo, Visual/Verbal e Sequencial/Global. Em 1994, foi criada a segunda versão, e, em 1997, foi disponibilizado gratuitamente na Web, desde que usado com fins educativos.

Tendo em conta o Metamodelo de Curry, o ILS pode ser enquadrado ao nível dos Modelos de Processamento de Informação e dos Modelos de Preferências Educativas, conforme referido anteriormente, as camadas exteriores do Modelo de Curry, pelo que mais facilmente observáveis e modificáveis, por forma a melhorar os processos e os resultados das aprendizagens.

O LSQ é um questionário de estilos de aprendizagem desenvolvido por Honey e Mumford (1992) com base no Modelo da Aprendizagem Experimental de Kolb e no seu Modelo de Estilos de Aprendizagem. Este questionário permite identificar quatro estilos de aprendizagem: Activo, Reflexivo, Teórico e Pragmático, que apresentam uma relação estreita com as quatro fases do ciclo de Kolb, respectivamente: Experimentação Activa, Observação Reflexiva, Conceptualização Abstracta e Experiência Concreta.

O LSQ é um modelo de processamento de informação, tendo em conta a classificação de Curry, e perfeitamente enquadrado no Modelo de Gordon-Bull.

No âmbito deste projecto, em relação a outros questionários, o LSQ apresenta a vantagem de ter sido adaptado à população portuguesa, em 2002 [Fernandes 2004].

5. Implicações da Utilização de Estilos de Aprendizagem

A maior parte dos professores concorda que os estilos de aprendizagem existem e reconhece o efeito significativo que os estilos de aprendizagem têm no processo de aprendizagem [Vicent & Ross 2001, *cit in* Gordon & Bull 2004].

Miller [2001] refere que as investigações comprovam que a motivação e o desempenho aumentam quando as estratégias de ensino são adaptadas às preferências e estilos de aprendizagem dos alunos. No mesmo sentido, Fernandes [2004] aponta que os estilos de aprendizagem, apesar do problema do diagnóstico preciso e rigoroso, permitem a concepção de planos de estudo, didácticas e de estratégias em consonância com o estilo de aprendizagem dominante.

Vários estudos têm mostrado que a aprendizagem é mais efectiva quando o estilo de ensinar corresponde com os estilos de aprendizagem dos alunos (*matching hipotheses*), do que quando tal não acontece [Schmeck 1988; Felder & Silverman 1998; Hayes & Allison 1993, 1996, *cit in* Felder & Brend 2005].

Zywno e Waalen [2001, *cit in* Felder & Brent 2005:60] relataram o desenvolvimento e a implementação bem sucedida de actividades de aprendizagem com recursos hipermédia dirigidas a alunos com estilos de aprendizagem menos favorecidos pelos métodos tradicionais.

Por outro lado, têm sido apontadas várias críticas à utilização dos estilos de aprendizagem. Segundo Gordon e Bull [2004], as mais fortes são: que a classificação em estilos de aprendizagem pode, perigosamente, estereotipar ou rotular os alunos; que os estilos de aprendizagem não são um factor cognitivo estável e que os estilos de aprendizagem também variam com a tarefa ou situação.

Na verdade, existe uma família de modelos de estilos de aprendizagem que reconhecem o carácter instável dos estilos de aprendizagem e que os descrevem como um processo evolutivo, tais como, Lewin's Cycle (Kolb 1984), Conceptions of Learning de Saljo (Marton 1976), e Bloom's Taxonomy (Bloom 1956) [Gordon & Bull 2004].

A utilização das TIC permite esbater estes argumentos. Os ambientes de aprendizagem virtuais podem ser programados para reconhecer o tipo de aluno, sem que os docentes ou os alunos conheçam o "estereótipo" em causa, e para reavaliar periodicamente os estilos de aprendizagem, em tarefas ou situações de aprendizagem diferentes.

Quando um aluno é colocado perante situações em que escolhe frequentemente determinado formato (texto, áudio, imagens, animações) ou, quando sistematicamente escolhe começar por visualizar exemplos de aplicação dos conceitos, essas interacções têm informação sobre as preferências que melhor se adequam ao seu estilo de aprendizagem.

A maior parte dos LMS, inclusivamente, armazena nas suas bases de dados informações que permitem estabelecer padrões de comportamento, que revelam as preferências de cada aluno e que podem permitem detectar o seu estilo de aprendizagem. [Graf & Kinshuk 2006].

Isto não significa, contudo, que os docentes devam apenas usar estratégias de acordo com o estilo de aprendizagem do aluno, porque isso impede que o aluno desenvolva capacidades nos restantes estilos, o que será muito importante ao longo da sua vida [Felder & Brent 2005].

Além disso, nem o professor nem a tecnologia devem forçar o aluno a criar um percurso, mesmo que pareça muito ajustado ao seu estilo de aprendizagem, apenas recomendar, dando-lhe flexibilidade para que siga caminhos de aprendizagem alternativos, se assim o entender [Gordon & Bull 2004].

Em suma, enquanto os docentes devem ter em conta a diversidade de estilos de aprendizagem, os alunos devem conhecer os estilos de aprendizagem, tendo presente o seu carácter mutável, no

sentido de desenvolver estratégias que lhe permitam melhorar os seus resultados e a capacidade de aprender melhor.

6. Aplicação Web para Gestão de Estilos de Aprendizagem

A aplicação para Gestão de Estilos de Aprendizagem (GEA) foi desenhada e implementada para funcionar na Web, com os seguintes objectivos:

- Informatizar o trabalho de distribuição, recolha e tratamento de dados dos questionários, uma tarefa monótona e morosa se realizada pelos métodos tradicionais, já não justificáveis no séc. XXI;
- Incentivar os docentes a conhecer os estilos de aprendizagem dos seus alunos e a usar esses resultados para definir estratégias, metodologias e experiências de aprendizagem;
- Motivar a reflexão dos alunos sobre as suas características individuais, que lhe permitam encontrar estratégias para aprender melhor.

De acordo com estes objectivos, o desenho da interface da aplicação (Figura 1) é o mais simples possível.

A interacção com o utilizador é conseguida a partir de um *browser* (sem necessidade de instalação de qualquer componente adicional) que, através da Internet e do protocolo HTTP (HyperText Transfer Protocol), comunica com o servidor Web, onde estão instalados e configurados o PHP (acrónimo recursivo de PHP: Hypertext Preprocessor), o SGBD (Sistema de Gestão de Bases de Dados) e o GEA.

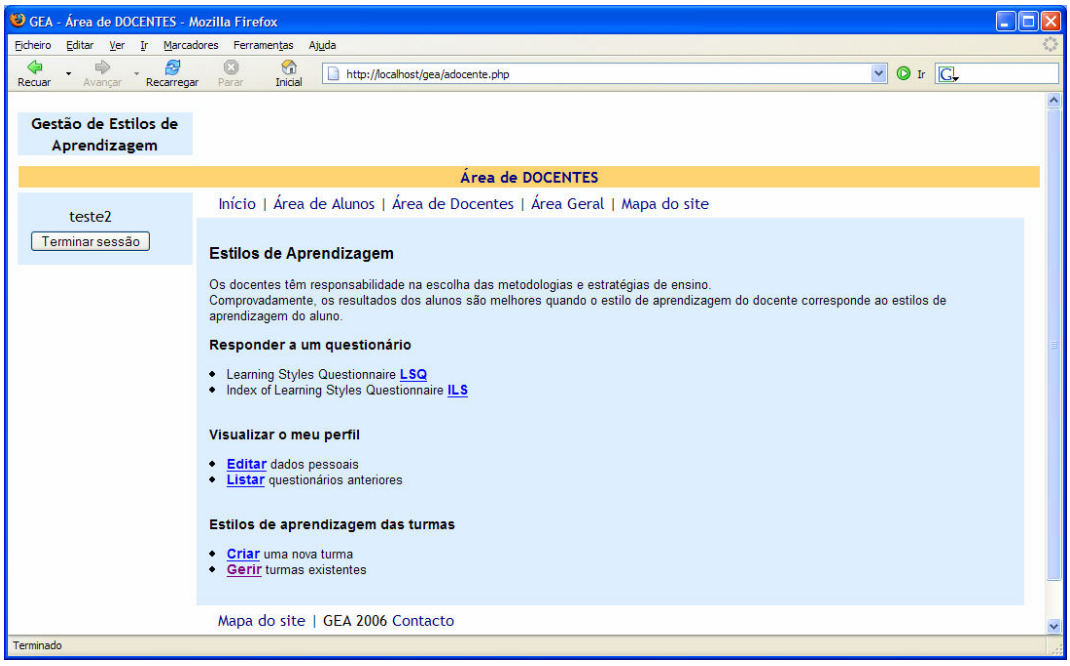


Figura 1: Interface de entrada da «Área do Docente» do GEA

A implementação da aplicação é baseada numa arquitectura cliente-servidor e utiliza no seu desenvolvimento apenas *software* livre: uma linguagem de *script* embutida no HTML (HyperText Markup Language) – PHP, um servidor Web – Apache e um SGBD – PostgreSQL.

A aplicação é suportada por uma base de dados que armazena, em várias tabelas, alguns dados pessoais dos alunos e dos docentes, os questionários disponibilizados (LSQ e ILS) e os resultados obtidos da realização dos questionários pelos utilizadores.

Ao solicitar uma página da aplicação, por exemplo, para realização de um questionário, um utilizador registado recebe os dados processados relativos às questões do questionário que escolheu. Depois de preenchido o questionário, este utilizador submete as suas respostas, para obter os resultados do seu perfil de aprendizagem, que são processados e registados na base de dados (ver Figura 2 e Figura 3Figura 2: Interface do GEA para o questionário LSQ).

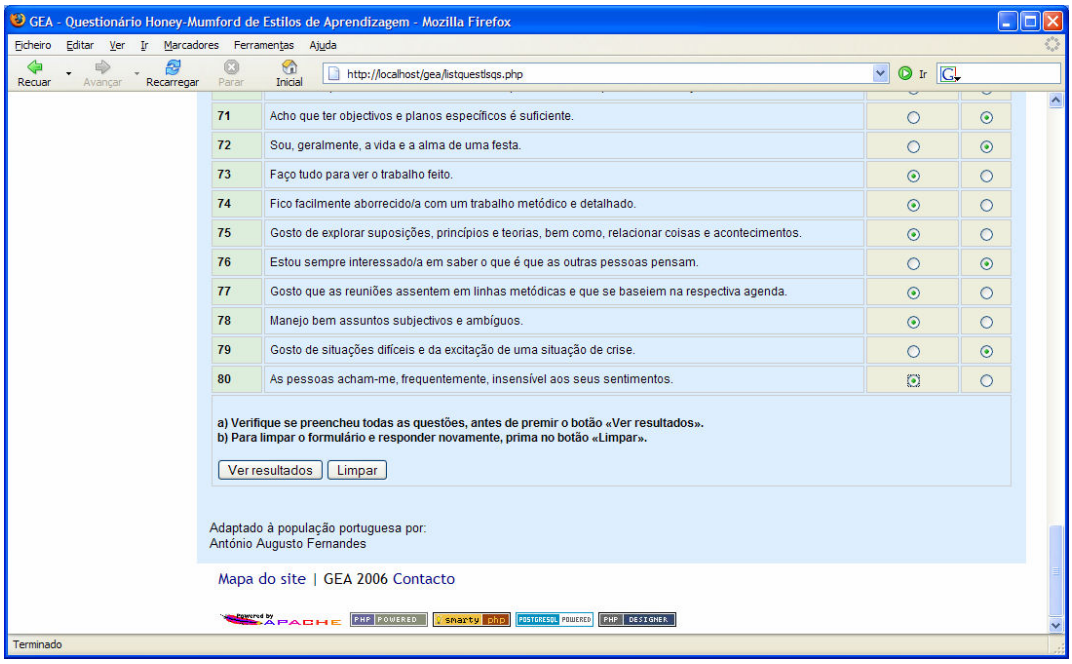


Figura 2: Interface do GEA para o questionário LSQ

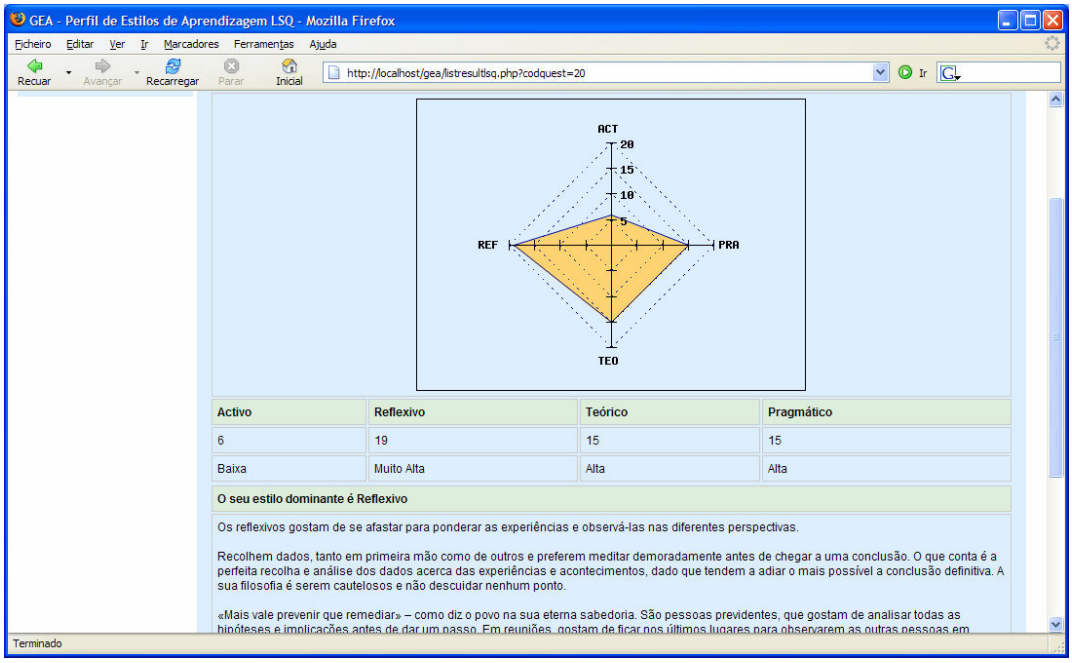


Figura 3: Interface do GEA para visualizar os resultados individuais no LSQ

O utilizador pode escolher qual o questionário que pretende realizar, podendo, inclusivamente, realizar ambos e repetidas vezes.

Tendo em conta os modelos em que se enquadra cada um dos instrumentos, é fornecido ao utilizador uma descrição do seu perfil de estilo ou estilos de aprendizagem dominantes e sugerido um conjunto de estratégias para melhorar o desempenho escolar e a capacidade de aprender a aprender de forma mais eficaz.

Qualquer utilizador poderá efectuar o seu registo, como aluno ou como docente, o que permite que seja encaminhado para a área específica do seu interesse.

Uma área de acesso geral permite que qualquer utilizador possa obter informação sobre os estilos de aprendizagem, bem como, visualizar os valores médios obtidos em cada estilo de aprendizagem, pelos alunos e pelos docentes (tendo em conta a data mais recente), relativamente a cada questionário.

Após os registos, um aluno poderá realizar qualquer um dos questionários e obter um relatório individualizado do seu perfil, enquanto que um docente poderá, além disso, obter um relatório do perfil de estilos de aprendizagem de cada uma das suas turmas. Os alunos são associados às turmas, a partir de um código, que é fornecido ao docente no momento da criação da turma e que deve indicar aos seus alunos.

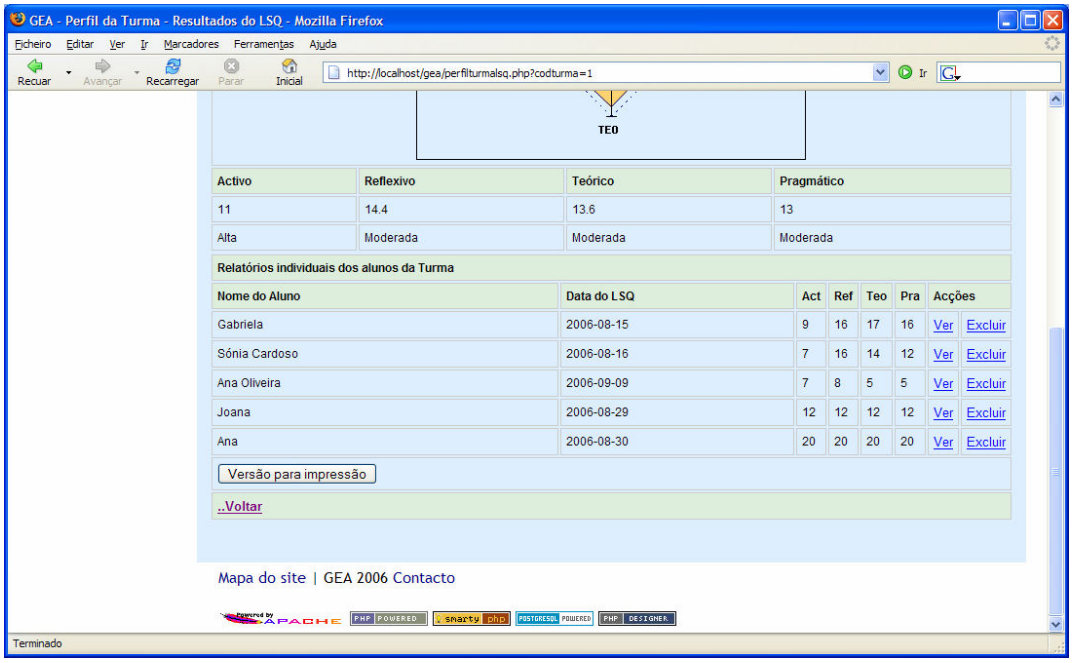


Figura 4: Interface do GEA para visualizar os resultados de uma turma, relativamente ao LSQ

Os relatórios do perfil de estilos de aprendizagem dos alunos de uma turma (ver Figura 4) indicam ao professor as pontuações que cada aluno obteve, relativamente a cada estilo de aprendizagem, no questionário mais recente; informação que poderá completar com as estratégias que mais se adequam a cada grupo, segundo os modelos em que cada questionário se aplica.

Sustentados na informação sobre os estilos de aprendizagem de uma turma, os professores podem desenhar actividades mais personalizadas e preparar/seleccionar contextos e objectos de aprendizagem diversificados, de acordo com os objectivos e as competências a desenvolver.

7. Conclusões

A informação obtida a partir da aplicação de GEA pretende ser um estímulo à reflexão em torno dos estilos de aprendizagem e das suas potencialidades, especialmente quando associados às TIC, para melhorar o autoconhecimento do aluno e aumentar a eficácia da função docente, ao conhecer melhor o perfil de aprendizagem dos alunos e a sua evolução.

Assim, ao informatizar os questionários, espera-se facilitar a recolha e tratamento dos dados relativos aos estilos de aprendizagem dos alunos e dos docentes, que sustentem e conduzam as estratégias mais adequadas à interacção com os conteúdos programáticos a abordar, bem como, motivar os alunos a reflectir sobre as suas características individuais, que lhe permitam encontrar estratégias para aprender melhor, criando ambientes de aprendizagem mais activos, flexíveis e geradores de conhecimento.

8. Referências

Cassidy, S. (2004). Learning Styles: An overview of theories, models, and measures, in *Educational Psychology,* Vol. 24, n.º 4, pp. 419-444.

Felder, R. M., & Brent, R. (2005). Understanding Student Differences, in *Journal of Engineering Education,* Vol. 94, n.º 1, pp. 57-72.

Felder, R. M., & Soloman, B. A. (1991, 1994). *Index of Learning Styles Questionnaire*. Acedido em 19/07/2006, 2006, em http://www.engr.ncsu.edu/learningstyles/ilsweb.html

Fernandes, A. A. (2004). Estilos de Aprendizagem. In *Didáctica das TIC* (1.ª ed., pp. 6-41): Instituto de Educação - Universidade Católica Portuguesa.

Fleming, N. D., & Mills, C. (1992). Not Another Inventory, Rather a Catalyst for Reflection. In *To Improve the Academy* (Vol. 11, pp. 137).

Gordon, D., & Bull, G. (2004). *The Nexus Explored: A Generalised Model of Learning Styles.* Paper presented at the 15th International Conference of Society of Information Technology & Teacher Education, Atlanta, Georgia, USA.

Graf, S., & Kinshuk. (2006). *An Approach for Detecting Learning Styles in Learning Management Systems.* Paper presented at the The 6th IEEE International Conference on Advanced Learning Technologies, Kerkrade, The Netherlands.

Miller, P. (2001). Learning Styles: The Multimedia of the Mind. Research Report., in, Vol., n.º.

Rodríguez, A. G., & Caro, E. M. (2003). *Estilos de aprendizaje y e-learning. Hacia un mayor rendimiento académico*. Acedido em 18-07-06, em http://www.um.es/ead/red/7/estilos.pdf

Ambientes de Aprendizagem na Concepção e Desenvolvimento das Instituições de Ensino Superior

Eduardo Luís Cardoso [1], Pedro Pimenta [2], Duarte da Costa Pereira [3]

1) Universidade Católica Portuguesa – CRPorto–ESB, Porto, Portugal

elcardoso@esb.ucp..pt

2) Universidade do Minho – Departamento de Sistemas de Informação, Guimarães, Portugal
pimenta@dsi.uminho.pt

3) Universidade do Porto – Faculdade de Ciências, Porto, Portugal
dcpereira@fe.up.pt

Resumo

As tecnologias de informação e comunicação (TIC), entendidas como tecnologias educativas, são defendidas como podendo suportar de forma mais efectiva a construção de conhecimento no ensino superior e a promoção de aprendizagens relevantes, com sentido, em particular se utilizadas como ferramentas cognitivas de amplificação das capacidades mentais e não, como tradicionalmente acontece, apenas como meios de envio de informação, como comunicadores de conhecimento ou tutores de alunos. Os ambientes de aprendizagem (AA) baseados em plataformas de *e-learning* (PeL) e integrando aproximações construtivistas aos processos de ensino e aprendizagem são considerados particularmente adequados para o ensino superior. A proposta de trabalho de investigação consistiu num estudo de caso visando aprofundar a compreensão do fenómeno da adopção de PeL no suporte a AA nas IES. Foi definido como objectivo a identificação do quadro de condicionantes do sucesso deste processo de incorporação de TIC na prática lectiva, ao nível da pré-graduação. Pretendeu-se caracterizar o papel desses factores condicionantes no processo de inovação, as suas formas de acção e as relações com os resultados. Foi desenvolvido o que pode ser apresentado como um modelo de referência. O modelo valoriza as influências e interacções entre o contexto institucional e as fases do processo de adopção de tecnologia e contempla a influência deste processo na evolução do próprio contexto institucional. O modelo proposto constitui um contributo para a compreensão do fenómeno e um referencial para a análise e avaliação de iniciativas institucionais na área e para a própria definição e planeamento de políticas institucionais e sectoriais.

Palavras chave: adopção de tecnologia, plataformas de *e-learning*, ensino superior

1. Enquadramento

As necessidades de desenvolvimento da sociedade assentam crescentemente, na criação, difusão e utilização de conhecimento, impondo desafios ao desenvolvimento da sociedade em geral, da educação, em particular e, nomeadamente, das instituições de ensino superior (IES) no quadro de um sistema de ensino superior em que se assiste a uma evolução de uma educação

para poucos e para objectivos limitados, para uma educação para muitos e para objectivos expandidos.

A organização do ensino superior e as IES tradicionais podem ser consideradas como portadoras de uma concepção essencialmente mecanicista dos sistemas de ensino, com origens na sociedade industrial e nas perspectivas de base tayloristas da organização da produção, sendo que estamos actualmente num contexto de reconhecimento e valorização de outras concepções do processo de ensino e aprendizagem, em que em oposição à tradicional ênfase na transmissão, na memorização, nos conteúdos e numa aprendizagem isolada, se contrapõe uma aprendizagem construída, situada num contexto, em interacção e em comunidade [Hills e Tedford 2003].

Estamos pois num contexto de mudança de paradigma e de referências, com implicações na organização do sistema de ensino e na prática dos processos de ensino e aprendizagem, também extensíveis ao subsistema de ensino superior.

As tecnologias de informação e comunicação (TIC), entendidas como tecnologias educativas, são defendidas como podendo suportar de forma mais efectiva a construção de conhecimento no ensino superior e a promoção de aprendizagens relevantes, com sentido, em particular se utilizadas como ferramentas cognitivas de amplificação das capacidades mentais e não, como tradicionalmente acontece, apenas como meios de envio de informação, como comunicadores de conhecimento ou tutores de alunos [Jonassen *et al.* 1998; Jonassen 1994]. Assim, as TIC funcionariam como *mindtools*, apoiando a interpretação e a organização do conhecimento pessoal do aluno.

Os ambientes de aprendizagem baseados nestas tecnologias e integrando aproximações construtivistas aos processos de ensino e aprendizagem são considerados particularmente adequados para o ensino superior onde o objectivo é a aquisição de conhecimento avançado, em particular nas fases intermédias e finais dos programas de pré-graduação, como defendido por Jonassen *et al.* [1993] no "Manifesto para uma aproximação construtivista ao uso de tecnologia no ensino superior".

Os ambientes de aprendizagem (AA) baseados em plataformas de *e-learning* (PeL), ofereçem uma implementação e utilização facilitada, constituem oportunidades de inovação e desenvolvimento para as instituições de ensino superior, suportando processos de ensino e aprendizagem ao nível da pré-graduação que podem ser avaliados como mais flexíveis em múltiplas dimensões [Grabinger 1998]. Desde logo em termos pedagógicos, pela diversidade de aproximações valorizadas, designadamente em função das necessidades dos alunos. Em termos de uma noção alargada da distância que, para além de considerar respostas aos obstáculos tradicionais associados à distância espacial e temporal, considera também os associados a distâncias psicológicas nomeadamente pela valorização da personalização da aprendizagem e da interacção social contextualizada. Flexibilidade, ainda em termos das múltiplas tecnologias e suportes que podem estar associadas a um AA.

Às TIC está, então, associado um grande potencial de aplicação em educação e uma expectativa de incorporação nos sistemas educativos que contribua para estes responderem adequadamente aos desafios que lhes estão colocados pela sociedade.

Embora possa ser identificada, em definições de política educativa, a emergência de um papel central para as TIC na educação [CEU/CEC 2000], a capacidade tecnológica disponível, associada à rápida evolução destas tecnologias, excede claramente a capacidade da comunidade educativa responder adequadamente às oportunidades e desafios que lhes são colocadas actualmente [Bates 2001].

A tradicional resistência à mudança das estruturas das IES, que pode ser uma vantagem em ambientes mais estáveis, atrasa os processos de adaptação às exigências de uma sociedade

baseada no conhecimento. As IES tem dificuldade em dar resposta a uma envolvente mais instável e complexa, como é exemplo esta incapacidade de explorar todo o potencial das tecnologias de comunicação e da economia digital. Em ordem a preservar o seu papel e integridade na sociedade do conhecimento as IES devem pois evoluir no seu modelo organizacional [Santos *et al.* 1998]. Embora as IES, sendo organizações dedicadas à criação de novo conhecimento, possam ser vistas como geradoras de mudança a vários níveis da sociedade, são muitas vezes tidas como organizações «conservadoras» no que respeita a alterações da sua própria forma de funcionar.

2. Objectivos da Investigação

Foi privilegiado, no trabalho de investigação que aqui se apresenta, o estudo sobre a adopção de tecnologias de informação e comunicação nos processos de ensino e aprendizagem, numa perspectiva organizacional, a nível das IES, o que pressupunha uma consideração alargada de temáticas. A área de aplicação foi delimitada ao ensino de pré-graduação nas instituições tradicionais do sistema nacional de ensino superior, actividade central das IES e fulcral à sua missão na sociedade.

A proposta de trabalho de investigação consistiu num estudo de caso visando aprofundar a compreensão do fenómeno da adopção de PeL no suporte a AA nas IES. Foi proposto ser desenvolvida uma compreensão do processo de adopção da tecnologia e da forma como este é influenciado e influencia o contexto em que se desenvolve. Foi definido como objectivo a identificação do quadro de condicionantes do sucesso deste processo de incorporação de TIC na prática lectiva, ao nível da pré-graduação. Pretendeu-se caracterizar o papel desses factores condicionantes no processo de inovação, as suas formas de acção e as relações com os resultados. O estudo deveria poder contribuir para o desenvolvimento de modelos de inovação tecnológica nas IES que considerem possíveis especificidades da realidade nacional e ser uma forma de colaborar na reflexão sobre o papel que as TIC, como tecnologias de ensino e aprendizagem, poderão desempenhar na evolução do ensino superior e das IES.

Esta definição do trabalho de investigação procura criar condições para compreensão da inovação baseada em tecnologias de *e-learning* no ensino superior que permitam uma análise sistematizada do quadro de condicionantes associado a um dado contexto de desenvolvimento de um processo de adopção e utilização destas tecnologias nas IES.

O ponto de partida e a questão orientadora do trabalho que aqui se apresenta, é sobre o processo de adopção de tecnologia:

Q - Como se desenvolve nas instituições de ensino superior o processo de adopção e utilização de plataformas de *e-learning* para implementar ambientes de ensino distribuído, ao nível da pré-graduação?

Com uma sub-questão mais específica de caracterização do processo:

Q1- Qual o quadro de condicionantes fundamentais para o sucesso deste processo de inovação?

Perspectivou-se um trabalho na área dos sistemas de informação preocupado com as possibilidades de uma utilização adequada de tecnologias de informação, com uma melhor gestão dos sistemas de informação e com os impactos e as implicações nas organizações e na sociedade, neste caso particular as IES e o sistema de ensino superior.

Pretende-se contribuir para o desenvolvimento de conhecimento sobre a inovação de base tecnológica nas IES, envolvendo a adopção de plataformas de *e-learning* para o desenvolvimento de ambientes de ensino distribuído e sobre os factores que condicionam estes processos de inovação, focando nas implicações associadas em termos de mudança pedagógica

e de mudança organizacional. Espera-se, assim, contribuir para a definição de políticas de desenvolvimento das IES e do ensino superior.

3. Metodologia

O estudo de caso, como método de investigação, tem as características para se ajustar à tipologia do trabalho e aos objectivos definidos e para estruturar um desenho da investigação em que se procura explicitar interacções entre uma multiplicidade de factores considerando vários aspectos relevantes para um processo de inovação educacional, nomeadamente, aspectos organizacionais, socioculturais, tecnológicos, pedagógicos e metodológicos, com a ênfase colocada no processo pelo qual o contexto social da adopção e utilização, influencia e é influenciado pela tecnologia [Walsham 1993]. O estudo de caso, como método de investigação qualitativa em sistemas de informação, permite uma abordagem holística ao fenómeno em estudo, permite considerar as condições contextuais e contemplar múltiplas fontes de evidências, criando condições para uma análise sistemática do processo de mudança nas organizações. É um método empírico de inquérito para investigar um fenómeno contemporâneo no seu contexto de inserção real [Yin 1994].

O trabalho de investigação foi baseado na estruturação de um estudo de caso sobre uma iniciativa do Departamento de Sistemas de Informação (DSI) da Universiade do Minho (UM) de promoção do ensino distribuído na prática lectiva, a nível do ensino de pré-graduação da responsabilidade do Departamento que foi designada DSI XXI. O início desta iniciativa foi sinalizado no início do ano lectivo de 1998/1999.

Numa perspectiva de paradigma interpretativista-construtivista, a investigação deverá conduzir a uma compreensão da realidade, do fenómeno em estudo, através de um conhecimento partilhado, construído nomeadamente a partir do significado atribuído por diferentes intervenientes que tem sido apontada como apropriada para avaliar e compreender a complexidade dos contextos de ensino e aprendizagem baseados em TIC.

A entrevista semi-estruturada constitui o instrumento central de interrogação da realidade neste trabalho de investigação. A entrevista é uma das técnicas mais usadas em investigação qualitativa em ambientes organizacionais. As entrevistas apresentam-se como adequadas para conhecer o entendimento, o significado e o posicionamento de indivíduos que vivem em dada realidade [King 1994].

Um instrumento de recolha de dados particularmente útil quando, como é o caso, o ponto de partida da investigação não é um conjunto de hipóteses a testar mas um conjunto de temáticas associadas a um modelo de interpretação do fenómeno da adopção de tecnologia nas organizações. Temáticas que se pretende aprofundar e detalhar a partir da diversidade de opiniões, experiências pessoais, pensamento e reflexões que venha a ser coligida, em particular, a partir das entrevistas aos grupos de actores no processo em estudo.

A análise qualitativa das entrevistas embora possa ser concretizada de diferentes formas é, em geral, considerada como adequada para o tipo de trabalho de investigação que aqui se pretende desenvolver. "A operação intelectual básica de uma análise qualitativa de materiais de entrevistas consiste essencialmente em descobrir «categorias», quer dizer, classes pertinentes de objectos, de acções, de pessoas ou de acontecimentos. Seguidamente, trata-se de definir as suas propriedades específicas e de conseguir construir um sistema ou um conjunto de relações entre essas classes" [Maroy 1995: 118-119].

4. Resultados e Análise

O trabalho de análise de conteúdo realizado com base nas entrevistas realizadas, a líderes, docentes, alunos e técnicos da IES e com base na comunicação intradepartamental, através de uma lista distribuição, permitiram a identificação e categorização de conceitos que emergiram como elementos chave caracterizadores do contexto institucional composto pelos contextos envolventes, organizacional e tecnológico, bem como do processo de adopção de tecnologia composto pelas condições para a adopção e utilização de PeL, da própria adopção e utilização e das suas consequências que, no conjunto, constituem os elementos essenciais do modelo de mudança que nos propusemos desenvolver, associado à adopção e utilização de PeL e de AA, na actividade de ensino nas IES, a nível dos cursos de pré-graduação.

O processo iterativo de identificação e categorização de conceitos com base nos dados empíricos permitiu desenvolver o modelo que se apresenta na Figura 1.

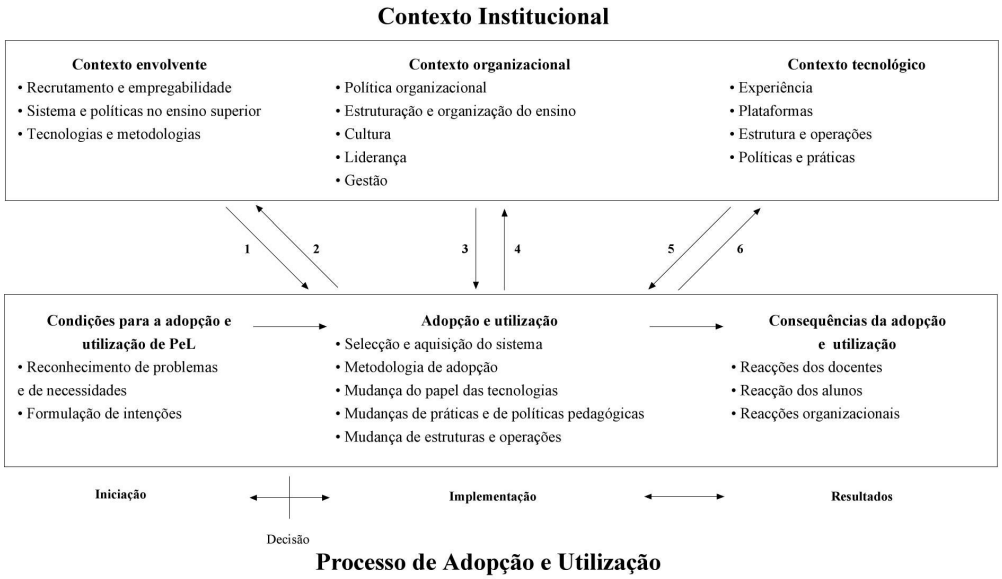


Figura 1 – Modelo de representação da mudança organizacional associada ao processo de adopção e utilização de PeL nas IES para a concepção e exploração de AA

Foi desenvolvido o que pode ser apresentado como um modelo de referência, inspirado na proposta de Orlikowski [1993], para a adopção de tecnologias nas organizações, integrando também aspectos dos modelos de fases apresentados por Rogers [1995] para os processos de mudança nas organizações e, em particular, por Fullan e Stiegelbauer [1991] na área da educação.

Este modelo foi, com base no estudo de caso realizado, detalhado para representar conceitos relevantes que constituem o quadro de condicionantes fundamentais para o sucesso destes processos de inovação, respondendo à primeira das sub-questões orientadoras do estudo. Os factores condicionantes facilitam ou dificultam a concepção e exploração de AA, baseados em PeL, na prática pedagógica nas IES. O modelo valoriza as influências e interacções entre o contexto institucional e as fases do processo de adopção de tecnologia e contempla a influência deste processo na evolução do próprio contexto institucional.

Parece ser pertinente esta perspectiva mais organizacional na análise ao fenómeno de inovação baseado na adopção de PeL no ensino superior, em oposição ou complemento a estudos mais centrados na adopção individual (como, por exemplo, nos trabalhos de Collis e Pals [2000]).

A fase de iniciação do processo de adopção e utilização, mais aprofundadamente analisada neste trabalho, que antecede a decisão de inovar e a decorrente implementação, envolveu a emergência da relevância do objectivo na organização e a concepção e preparação da proposta de mudança, características das etapas de formulação de agenda e de ajustamento, com eventos numa etapa a alterarem decisões tomadas em etapas anteriores.

Numa organização, como uma IES, em que cada indivíduo tem a percepção, identificada neste trabalho, de dispor de uma grande autonomia nas formas de desenvolver a actividade de ensino de que é responsável (e de que os sistemas de avaliação e progressão individual pouco relevam a dimensão do ensino e aprendizagem) a adesão a uma proposta de inovação educacional é, no essencial, opcional. Uma decisão independente das decisões de outros elementos da organização. A actuação de cada indivíduo pode, contudo, ser condicionada pelas decisões dos órgãos de gestão da organização e pelo posicionamento dos indivíduos de posição superior na carreira profissional docente. Parece, pois, ser importante a fase de iniciação envolver uma formulação da agenda da organização de forma partilhada e aprofundada, geradora de consensos, quando se pretende promover uma inovação com um âmbito alargado. Esta fase pode corresponder a um processo moroso como aconteceu no caso estudado.

O modelo proposto representa um enquadramento para compreender o processo de construção de condições para adoptar e utilizar ambientes de ensino em instituições de ensino superior. O sucesso deste processo exigirá um comprometimento individual numa perspectiva de melhoria contínua e, como foi evidenciado neste trabalho, um comprometimento organizacional e uma evolução do contexto envolvente, factores interrelacionados com o próprio processo de mudança. Um posicionamento das IES que pode ser avaliado como o de uma escola aprendente, em que o processo de mudança de aspectos essenciais à organização seja perspectivado, como propõe Senge *et al.* [2000], como um processo de aprendizagem, envolvendo os vários actores do sistema. A fase de iniciação permitiu a decisão de adoptar uma PeL, com a instalação do TWT (em Dezembro de 2000), a que se seguiu uma fase de implementação que apenas permitiu uma exploração da PeL no ano lectivo de 2001/2002, já que vários aspectos do planeamento foram, entretanto, modificados. A implementação foi conduzindo a medidas de avaliação dos resultados com objectivos de dinamização e melhoria da iniciativa em curso.

5. Conclusões

A análise da adopção e utilização de uma PeL na concepção e desenvolvimento de AA, numa IES, permitiu extrair conclusões sobre aspectos que caracterizam as possibilidades de concretizar o potencial dos ambientes de aprendizagem suportados em TIC no ensino superior, decorrentes da forma como estes ambientes são compreendidos e aceites.

Motivações e Objectivos

As IES podem identificar várias motivações relevantes e específicas para justificar o investimento em TIC em educação, que, no caso estudado, foram categorizadas em, necessidades pedagógicas, administrativas, de competição, de ética profissional e de aprendizagem organizacional. Parecendo inevitável a necessidade de um posicionamento claro das IES perante os desafios e as oportunidades das TIC na organização e desenvolvimento da actividade de ensino superior.

Decorrente da percepção de necessidades, registamos neste trabalho uma diversidade de objectivos formulados para a adopção e utilização de TIC no ensino superior que, em termos gerais, foram categorizados em objectivos de melhoria da qualidade e produtividade da actividade de ensino mas também de desenvolvimento da IES a vários níveis. No concreto, podiam ser identificados objectivos de curto prazo, como sejam, tornar mais expedito o processo de auto-avaliação do ensino ou informatizar a gestão de sumários, quer objectivos necessariamente de longo prazo, como os respeitantes ao fomento da auto-aprendizagem, melhoria da aprendizagem ou desenvolvimento nos alunos de competências de autonomia e iniciativa.

A adopção de TIC e em particular de PeL, como sistemas estruturantes de apoio ao ensino e aprendizagem, através designadamente da criação de AA, poderá assim representar uma estratégia de desenvolvimento do ensino superior, associada a um objectivo de melhoria contínua do processo de ensino e aprendizagem. Mas, com a ampla gama de motivações e objectivos, mostra-se necessário que, a diferentes níveis da IES, sejam clarificadas orientações para a mudança pedagógica, nomeadamente, a nível da Reitoria, da Escola ou Faculdade e do Departamento, não deixando de ficar à disposição de cada docente um espaço de autonomia significativo e determinante para a concepção de uma utilização de TIC que se adeqúe, faça sentido e se justifique no contexto específico de uma disciplina, dos conteúdos programáticos ou do grupo em aprendizagem.

Parece crescente a percepção, pelo corpo docente, de que o modo essencialmente transmissivo de informação codificada, explícita, facilmente representada, armazenada e distribuída por meios electrónicos e digitais, designadamente através da Internet, perde relevância no contexto do ensino superior, assumindo-se indispensável a focalização num outro modo de ensino e aprendizagem, para o qual assumimos como referência o Modo 2 proposto e defendido por Hills e Tedford [2003]. Passando, assim, a ser mais valorizada a construção de conhecimento implícito, tácito, associado a competências transversais, identificadas neste trabalho como centrais à missão das IES, tais como as inerentes à resolução de problemas, à construção de projectos ou ao trabalho em grupo, abordagens que podem, também, ser facilitadas e potenciadas por TIC, como foi parcialmente ilustrado no caso estudado.

Apesar de uma visão pedagógica ter sido repetidamente identificada neste estudo, muitas das motivações explicitamente associadas à iniciativa de adopção de PeL e de AA revelaram-se pouco alinhadas com esta preocupação essencial e correspondiam, muitas delas, a visões contraditórias com as da iniciativa proposta.

Processo de ensino e aprendizagem combinado

No contexto do ensino superior pré-graduado, baseado num *campus* de uma IES, os AA começarão por ser compreendidos como uma solução suplementar dos processos de ensino presenciais, muito associados a funções de publicação de informação e de distribuição de conteúdos didácticos.

Se a PeL oferecer serviços académicos integrados, os AA poderão facilmente assumir um carácter obrigatório e uma utilização massificada, ao assegurar funções de organização e gestão do processo de ensino e aprendizagem.

Os AA, compreendidos e valorizados como componentes intrínsecos da actividade de ensino superior, como os concebemos neste estudo (parte integrante e central de um processo de ensino e aprendizagem combinado, também designado *b-learning*), suportando actividades de aprendizagem, envolvendo, eventualmente, experimentação virtual e certamente relevantes em termos de avaliação, em que as ferramentas tecnológicas disponibilizadas sejam adoptadas como ferramentas cognitivas ao serviço de um pensamento pedagógico, no contexto de um

dado percurso de aprendizagem proporcionado aos alunos, terão, numa fase inicial de implementação, uma adesão limitada de um número reduzido, minoritário, de docentes de uma IES tradicional.

Interesse alargado mas concretização limitada

Foi identificado ao longo do estudo um grande interesse por parte de líderes, docentes e alunos na exploração do potencial das TIC e, em particular, de PeL, para a criação e disponibilização de ambientes de ensino e aprendizagem baseados na Internet. Este interesse, registado ao longo das entrevistas mas, também, em diversas fontes documentais, foi analisado num quadro geral de grandes, ou mesmo exageradas, expectativas em relação ao papel das TIC nos processos de ensino e aprendizagem, a par de concepções pouco alinhadas (e por vezes contraditórias) em relação às características destes ambientes que aqui designamos como ambientes de aprendizagem.

A observação das implementações de *sites* disciplinares permite constatar que, embora o número de disciplinas com um *site* de suporte que chegam a existir implementados na PeL seja significativo (25 num total de 36 disciplinas), é bastante limitado o número de implementações que incorporam algumas dimensões pedagógicas relevantes, designadamente no quadro de referência dos ambientes construtivistas de aprendizagem (ACA) [Jonassen 1993], (8 disciplinas com *sites* caracterizados como inovadores no total dos 25). Resultado que pode ser considerado agravado pelo facto de ser expectável que pelo menos em 5 destas disciplinas os docentes implementassem um AA, mesmo que não pudessem contar com o apoio de uma iniciativa departamental de promoção da adopção, já que detinham experiências anteriores em *sites* disciplinares com várias características inovadoras em relação à prática lectiva corrente.

Diversidade de implementações e exigências de concepção

A partir da explicitação, pela Direcção do Departamento, de uma visão geral sobre os ambientes que a comunidade de aprendizagem deveria vivenciar, dos objectivos pretendidos com a adopção de TIC e de modelos de ensino e aprendizagem de referência, a concepção e a decorrente implementação dos ambientes de ensino e aprendizagem baseados em PeL deveriam ser muito diversas para satisfazerem as exigências e especificidades associadas a cada contexto de aprendizagem, isto é a cada disciplina. Esta diversidade foi em alguma extensão ilustrada nas implementações observadas e analisadas, com funcionalidades que propusemos serem classificadas e caracterizadas como sendo associadas à organização do processo de aprendizagem, à exploração de conteúdos, a actividades propostas aos alunos, a componentes experimentais suportadas em TIC e a formas de avaliação implementadas. Esta diversidade de implementações, que certamente poderia ser bastante mais complexa do que a observada neste estudo, integrando componentes, estruturando actividades e utilizando ferramentas características dos AA, permite, ainda assim, justificar o interesse e o grande potencial de aplicação percepcionado. No entanto, este tipo de implementação envolve, como observado, uma grande exigência de concepção pedagógica na criação de ambientes baseados numa PeL que constituam uma aplicação específica de suporte a uma dada prática lectiva. A falta de profundidade observada em muitas das implementações, patente, por exemplo, em soluções muito focadas na exploração de conteúdos e em aspectos da organização do processo de ensino, pode ser resultado das dificuldades impostas pela concepção pedagógica. Outros estudos confirmam a manutenção de uma enorme dificuldade na concepção de soluções de TIC, enquanto ferramentas sistemáticas de auxílio ao trabalho pedagógico e didáctico em IES [Matos 2004]. Concepção esta que pode beneficiar das aproximações metodológicas da concepção e desenvolvimento de sistemas de informação, mas que implica, concomitantemente, um esforço

significativo e qualificado, especialmente na área pedagógica e, também, na área tecnológica específica.

Papel central dos docentes

Uma iniciativa de adopção de AA no ensino superior é certamente muito dependente do comprometimento dos docentes. Ao longo do trabalho realizado, foi possível constatar que os resultados observados associados à implementação de AA numa PeL se deveram, fundamentalmente, à resposta individual de cada docente, aos desafios enumerados para a iniciativa. As condições criadas, bem descritas e em si relevantes, não foram suficientes para alterar, de forma significativa, o contexto institucional caracterizado por vários aspectos inibidores das mudanças pressupostas. A iniciativa não incluía soluções para factores percepcionados por docentes como decisivos como sejam, desde logo, a valorização da actividade de ensino através de uma avaliação consequente que garantisse uma visão abrangente e contemplasse métricas adequadas à utilização de TIC e, portanto, adequadas à consideração do trabalho associado à adesão à iniciativa. Mas, também, não incluía soluções para o apoio à produção de conteúdos, nem a disponibilização de uma equipa especializada na concepção e implementação de AA, o reforço das equipas docentes para compensar o acréscimo de esforço inicial ou a integração da PeL nos sistemas de informação da IES.

Deveria ser claro e relevante um quadro de incentivos ao corpo docente que oferecesse condições sustentadas para a inovação e desenvolvimento dos processos de ensino e aprendizagem.

Impactos e Perspectivas de Desenvolvimento para as IES

Apesar da amplitude limitada dos resultados decorrentes da iniciativa departamental concreta acompanhada e estudada, foi possível constatar uma paradoxal dissonância entre um consenso sobre a necessidade de mudanças na prática pedagógica e a falta de políticas adequadas e nomeadamente de políticas de tecnologias em educação, tendo sido destacada a necessidade de existir, nas IES, uma estratégia de mudança pedagógica prévia à mudança associada à adopção e utilização de ambientes de aprendizagem. Neste quadro é maior o risco, associado ao uso de TIC, de serem automatizadas as práticas de ensino mais tradicionais.

Foram caracterizadas implicações do processo de adopção e uso desenvolvido em termos da necessidade de uma compreensão e visão partilhadas, assim como de exigências de capacidades de liderança pedagógica do processo e de exigência de um ciclo de aprendizagem a ser percorrido, para poderem ser atingidos objectivos significativos à escala institucional. Ainda que possam ser conseguidos resultados significativos a nível individual ou de pequenos grupos de docentes, para serem atingidos objectivos a nível organizacional, a mudança proposta é ambiciosa.

O estudo permitiu reforçar a necessidade de mudanças no papel do docente para poderem ser atingidos objectivos pedagógicos relevantes.

Foi ainda identificada a necessidade das IES disporem de um sistema integrado de apoio à actividade de ensino em função da realidade tecnológica.

A efectiva adopção e utilização de PeL que permita a concepção e exploração de AA permite perspectivar, de acordo com o estudo realizado, a introdução de vários factores de flexibilização que podem suportar estratégias de desenvolvimento das IES tradicionais. Sendo particularmente destacado o papel que os AA podem ter no suporte a uma estratégia de centrar

a intervenção das IES mais no aluno e no processo de aprendizagem, reclamada pela sociedade e projectada pela própria governação das universidades, como no caso da Reforma de Bolonha em curso [MEE 1999].

Também relevante para um papel de crescente importância das IES numa sociedade determinada pelo conhecimento é a possibilidade dos ambientes de aprendizagem, conseguindo acrescentar valor pedagógico significativo e implementando processos combinados, potenciarem a abertura a novos públicos e sustentarem uma necessária diferenciação do sistema de ensino superior nacional.

Serão no entanto necessários novos e renovados serviços ao nível de uma IES tradicional no âmbito de um processo de reorganização e restruturação que lhe permita uma capacidade reforçada de gestão de conhecimento e de aprendizagem de forma a qualificar a sua intervenção ao nível da actividade de ensino.

O modelo proposto foi desenvolvido tendo por base um modelo oriundo da área dos sistemas de informação, proposto por Orlikowski [1993], integrando contributos de teorias clássicas sobre inovação organizacional e inovação educacional. Desenvolvido e detalhado para a adopção de tecnologia no ensino superior, o modelo proposto constitui um contributo para a compreensão do fenómeno e um referencial para a análise e avaliação de iniciativas institucionais na área e para a própria definição e planeamento de políticas institucionais e sectoriais.

6. Referências

Bates, A., "The Continuing Evolution of ICT Capacity: The Implications for Education," *The Changing Faces of Virtual Education*, Farrel, G. (Ed.), The Commonwealth of Learning, Vancouver, 2001, 29-46.

CEU/CEC, *e-Europe 2002. An Information Society For All. Action Plan prepared by the Council and the European Commission for the Feira European Council, 19-20 June 2000*, Brussels, 2000, 29.

Collis, B. e Pals, N., "A Model for Predicting an Individual's Use of a Telematics Application for a Learning-Related Purpose", *International Journal of Educational Telecommunications*, 2000, 63-103.

Grabinger, S., "REALS for Distributed Learning", *A paper delivered at the EuroConference98*, Aveiro, 1998.

Fullan, M. e Stiegelbauer, S., *The New Meaning of Educational Change*, Continuum, London, 1991.

Hills, G. e Tedford, D., "The Education of Engineers: The Uneasy Relationship Between Engineering, Science and Technology", *Global Journal of Engineering Education (UICEE)* 7, 1 (2003), 17-28.

Jonassen, D., Mayes, T. e McAleese, R., "A Manifesto for a Construtivist Approach to Technology in Higher Education", in Mayes, T., Jonassen, D., Duffi, T. e Lowyck, J. (Eds.), *Designing Constructivist Learning Environments*, Springer-Verlag, Heidelberg, 1993, 231 - 247.

Jonassen, D., "Technology as Cognitive Tools: Learners as Designers", [Disponível em 21.05.2001 em: http://it.coe.uga.edu/itforum/paper1/paper1.html], 1994.

Jonassen, D., Carr, C. e Yueh, H., "Computers as Mindtools for Engaging Learners in Critical Thinking", *TechTrends* 43, 2 (1998), 24-32.

King, N., "The Qualitative Research Interview", *Qualitative Methods in Organizational Research. A Practical Guide*, in C. e Symon Cassel, G. (Eds.), Sage Publications, Londres, 1994, 14-36.

Maroy, C., "A Análise Qualitativa das Entrevistas", *Práticas e métodos de investigação em Ciências Sociais*, in Albarello, L., Digneffe, F., Hiernaux, J.-P., Maroy, C., RuQuoy, D., e Saint-Gcorges, P. (Eds.), Gradiva, Lisboa, 1995, 117-155.

Matos, J., *As Tecnologias de Informação e Comunicação e a Formação Inicial de Professores em Portugal: Radiografia da Situação em 2003*, Centro de Competência Nónio Século XXI da Faculdade de Ciências da Universidade de Lisboa, Lisboa, 2004.

MEE, *Declaração de Bolonha - Declaração dos Ministros da Educação Europeus*, Ministros da Educação Europeus, Bolonha, 1999.

Orlikowski, W., "CASE Tools as Organizational Change: Investigating Incremental and Radical Changes in Systems Development", *Management Information Systems Quarterly,* 1993, 309-340.

Rogers, E., *Diffusion of Innovation*, The free press, New York, 1995.

Santos, F., Heitor, M., e Caraça, J., "Organisational Challenges for the University", *Higher Education Management,* 1998, 87-107.

Senge, P., Cambron-McCabe, N., Lucas, T., Smith, B., Dutton, J. e Kleiner, A., *Schools That Learn. A Fifth Discipline Fieldbook fo Educators, Parents, and Everyone Who Cares About Education*, Nicholas Brealey, Londres, 2000.

Walsham, G., *Interpreting Information Systems in Organizations*, John Wiley & Sons, Chichester, 1993.

Yin, R., *Case Study Research*, Sage, California, 1994.

TEMA 3 - A SOCIEDADE E OS SISTEMAS DE INFORMAÇÃO

3.2. Aspectos Sociais dos Sistemas de Informação

Social Network Framework

Ricardo Duarte, Alberto Silva

INESC-ID
ricardomduarte@gmail.com

INESC-ID
alberto.silva@acm.org

Abstract

Social networks are an emergent phenomenon on the Internet. It becomes increasingly necessary to be able to create models that can help us represent these structures, with appropriate focus on the underlying main concepts. A clear and formal definition of these concepts will allow both the explicitation of social networks, which already exist implicitly in several Information Systems today, and also the ability to integrate networks from different systems and from different domains. Moreover, an uniform representation of concepts allows the development and reutilization of algorithms specific to the social networking domain. In this text a Social Network Framework is presented to answer these needs. Focus is given on the conception of a modeling strategy to represent social networks. A discussion of model transformations and algorithms specific for social networks is also presented.

Palavras chave: social networks, modeling, integration, transformations, link mining.

1. Introduction

Social networks are emerging on the Internet as one of the most important recent phenomena in Information Society [Hempel 2005]. This kind of networks which have the main goal of capturing information about social relationships between people in digital support, enclose an intrinsic value, that can and should be explored in several different ways, ranging from direct marketing to the retrieval of people and extraction of information. These abilities are related to a more general potential that is inherent to social networks, which is the capability to express a context based on information about the neighborhood around a given individual in the network. Social networks also have the power to disclose the true identity of Internet users. In fact, when a link is established between two users, both are acknowledging and certifying the existence and identity of one another. This kind of identification will bring a greater degree of responsibility to the usage of information systems, because abusive behavior can be observed by the social network neighbors of each individual, and this will strongly relate digital responsibility to real life responsibility. Although the popularization of social networks is recent, these networks already exist more or less implicitly in all Information Systems which somehow support an online collaboration between people or organizational structures [SocialNetwork]. In this context, it becomes necessary to extract the information present in these systems, and to represent it in an explicit way, unraveling true social networks in a format which may be recognized by algorithms able to extract new information that was previously unknown from these networks. A new requirement that emerges from the growing proliferation

of social networks throughout the Internet is the ability to integrate information captured from different social networks. Most social networks are specific to given domains or otherwise too general to contain all the relevant information needed to solve specific problems. The integration of networks from different domains may reveal new information which can be explored and analyzed. It is thus necessary to create integration mechanisms that allow us to infer new networks, with different relations and topologies, from existing social networks.

To answer the need for social networks analysis and integration, a Social Network Framework (SNF) is proposed in this paper. This framework encloses (1) the definition of a set of models to allow the representation of social networks, (2) importing mechanisms that will retrieve social networks from existing systems and represent them according to the defined models, (3) a set of transformations, allowing conversions between models, and (4) algorithms which may analyze the social networks to retrieve relevant information.

Figure 1 shows the package hierarchy of the Social Network Framework, and the existing dependencies between packages.

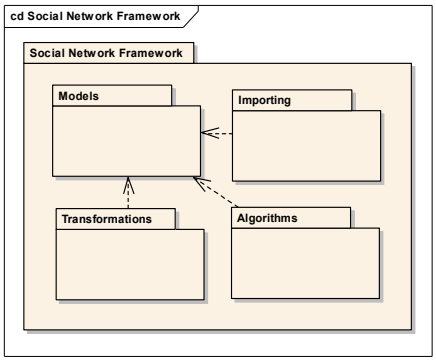


Figure 1 – Social Network Framework

In this Section we explained the motivation behind this work and proposed a Social Network Framework to deal with the identified problem. The remainder of this paper is divided into five additional sections: In Section 2, social network history is briefly introduced and an overview on previous work done on this area is given. In Section 3 a three layer modeling strategy to represent social networks is presented. We discuss different kinds of social network models intended for specific purposes: Specific Domain Models and a Canonical Model. In Section 4, some features that should be implemented to take advantage of the defined models are enunciated. Importing mechanisms to extract implicit social network from existing systems are exemplified, possible transformations between social network models are discussed, and some algorithms able to analyze social networks are listed. Finally, in Section 5 we present the main conclusions taken and discuss future work that can be done beyond the initial approach presented in this text.

2. State of the Art

Although the concept of "social networks" has been referenced for the first time in 1954 by J. A. Barnes [SocialNetwork], only in the mid-90s has this phenomenon found its firmed place in Information Society, through the large scale dissemination of the Internet. Before, the establishment of large social networks was difficult since it was hard to keep people constantly in contact with each others. It was also difficult to keep track of such networks. With the emancipation of the Internet, the natural human ability to build social relationships through

communication has been bestowed, and some networks started to emerge based on websites, forums and mailing lists.

The multiplication of social networks may have started in 1995 with the website classmates.com [Classmates]. Since then, other networks have emerged, with a greater or lesser number of members: as an example, mySpace [MySpace] counts with more than 68.000.000 users, as of 2005. The list of websites supporting social networks continues to grow, counting with more general or more specific networks [SNList]. But social networks are not only present in internet websites: the implicit network in MSN Instant Messenger [IM] contact lists, in Google's Gmail [Gmail] invite-only e-mailing service, or in phone companies' call logs are possibly much larger than those contained in any particular website, and are already strongly related to real user identities.

In the scientific community, research in the area of social networks divides itself between systems for finding experts, such as ReferralWeb [Kautz 1997], whose objective is to explore the links between people in an organization, to quickly find experts on a given topic that can be reached through personal contact lists of people in the organization; agent-based computing paradigms [Wooldrige 2002]; or Link Mining [Getoor 2003], which cares about development of algorithms which can explore the relationships in linked structures, mostly focused on the hyperlinked structure of the World Wide Web.

3. Social Network Framework – Models

There are several ways in which the domain of social networks can be modeled. Although this domain is not inherently complex, some choices must be made regarding its representation. These choices will vary according to the purpose intended by the representation.

As a guide for taking such decisions, three general considerations have been taken into account: (1) our representation should be broad enough to allow the representation of most social networks implicit in existent applications; (2) the representation should be flexible to allow integration and transformation between different social networks; (3) the representation should allow a simple implementation, such that generic graph algorithms may be applied on it, with an acceptable performance, even within large scale sets of data.

3.1. Basic Concepts

A social network, as a representation of social relationships between people, encompasses two basic concepts: (1) Person, which represents the basic nodes of the network; (2) Relationship, which denotes the existence of some social relationship between person nodes.

The Person concept represents individuals, human beings, which will potentially display social behavior between them. Care must be taken when attempting to generalize the Person concept. Considering, for example, that an organization is a (collective) person, may be sometimes adequate and others not, but as a rule of thumb, such extensions should be avoided, or dealed with special care, because they may invalidate the rationale behind some decisions taken during the construction of the model here presented.

A social relationship, henceforth referred to only as relationship, represents some interaction between two or more people. Care must also be taken when considering which relationships may represent social interaction between people. A social relationship implies that all participants in the relationship are active in that relationship. For example, "I know George" will only be considered a social relationship, if George also knows me, and thus participates in the relationship. This does not invalidate that relationships may be asymmetrical, such as "Mary is John's mother".

3.2. Modeling Strategy

Social networks, in the real world, are complex, as there are many different kinds of relationships which may exist between people. Therefore, a social network model, which attempts to simultaneously represent multiple kinds of relationships between different sorts of people, must be complex to some extent. This complexity goes against the simplicity intended by our previously stated guidelines, yet, simply ignoring this complexity would result in the conception of a shallow model, with a very limited expressive power. The proposed solution to accommodate both simplicity and expressiveness is the creation of a set of models instead of one, each intended for a specific purpose, and then providing suitable mechanisms to transform ones into others.

Figure 2 shows the three layers architecture proposed for social network modeling. A Social Network Meta Model (SNMM) is defined to model the representation of Social Network Models (SNM), which model Social Network Instances (SNI). A SNI represents a set of relationships between actual people. The respective SNM represents all kinds of people and relationships allowed in that SNI. Finally the SNMM defines the possible relations between People Types and Relationship Types that can be represented in SNM's.

At the SNM layer we define the different models that will allow us to level the tradeoff between simplicity and expressiveness. To allow a simple representation we define a Canonic Model (CM). Instantiations of the CM are simple enough to be readily analyzed by generic graph algorithms, which is the CM's main purpose. Specific Domain Models (SDM) allow us to represent social networks from specific domains in our Social Network Framework. For each existing social network which we intend to integrate or analyze, a SDM should be defined. With appropriate transformations, the instances of different SDM's can be integrated into a more general SDM instantiation or converted into instances of the CM that can be analyzed by specific algorithms. These transformations are defined at the models level and then applied on actual instances.

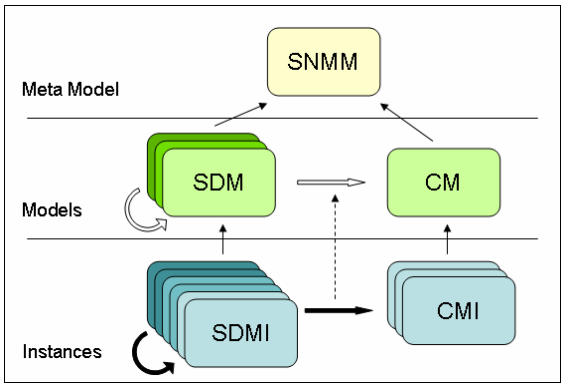


Figure 2 – Social Network Meta Model

Figure 2 shows the various models at their respective levels. Upward arrows crossing layers represent instantiations. Thick arrows represent transformation definitions (unfilled) and their respective instantiations (filled).

The SNMM and the CM are part of the core SNF models package. All SDM's should be developed as extensions to the framework, according to the particular social network domains we wish to use within the framework.

3.3. Social Network Meta Model

The Social Network Meta Model defines the possible elements and relations that can be expressed in Social Network Models. Figure 3 presents the SNMM. Instances of this meta model define social network types composed by a set of relationship types between given types of people, through specific role types.

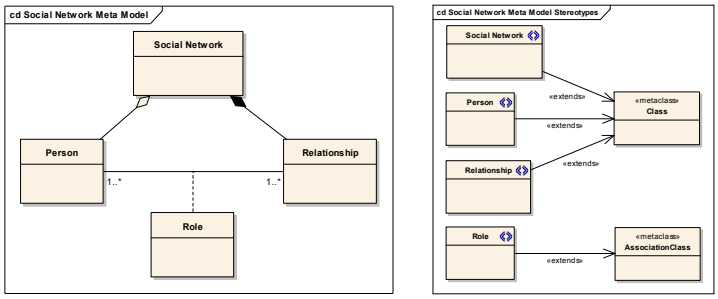


Figure 3 – Social Network Meta Model

3.4. Social Network Models

The SNMM allows us to represent Social Network Models, which define social networks as aggregations of relationships between people. These social networks also enclose the set of people who share relationships contained in the social network. A person not related to any other by a relationship, is not social by nature, and thus cannot be part of a social network. A person represents the identity of a real person, who can be part on several relationships which may pertain to different social networks. Roles allow people to participate differently in a relationship.

3.4.1 Specific Domain Models

Figure 4 shows, as example, an instantiation of the SNMM defining a SDM which represents a social network in a school domain. In this network there are two kinds of people, students and teachers, which relate between themselves through school class relationships.

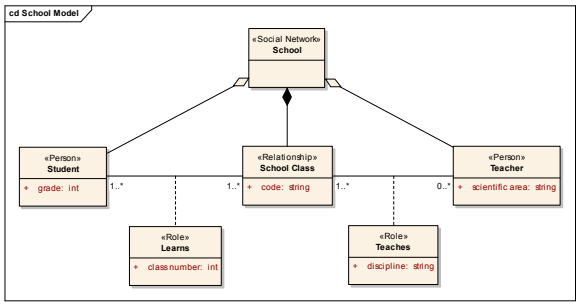


Figure 4 – School Specific Domain Model

An instantiation of the School SDM is shown on Figure 5, representing a simplified social network containing only one school class relationship between Teacher John Parker, and students, Paul, Susan and Mary.

3.4.2 Canonical Model

The Canonical Model represents a single view of a social network. This model allows only one specific kind of relationship to be represented, and is therefore a projection of the social network representing all existing relationships on a single dimension, which is the one regarding the relationship type of interest. The main purpose of the CM is to allow a simple representation that will be used to generically run graph algorithms, which may extract relevant information from the social network. As though, the representation for this model becomes very similar to the modeling of generic graph structures.

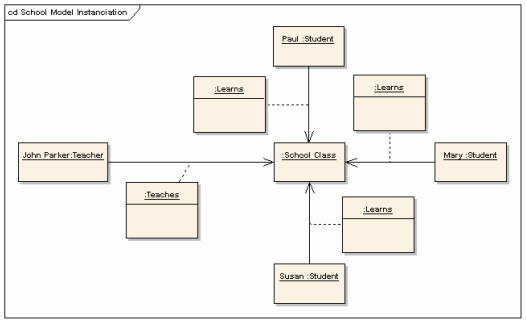


Figure 5 – School Model Instantiation

Figure 6 shows the proposed model for canonical social networks. As stated before, there is only one kind of relationships and each relationship associates only two people. A constraint should be enforced preventing that person who participates in the relationship through the source role is the same as the one assuming the target role.

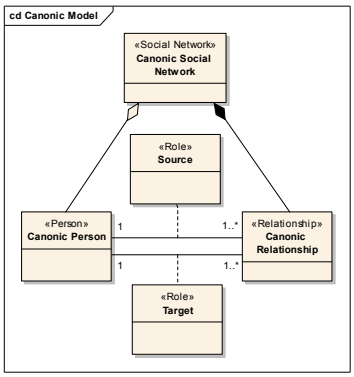


Figure 6 – Canonical Model for Social Networks

Each SDM may generate different CM's, one for each analytic purpose. Domain specific information is actually lost in the transformation process from SDM's to CM's, and so if we wish to analyze a network in some aspect of the problem domain, it is the transformation's responsibility to filter the generation of the canonic instantiation according to that semantic. For example, if we are interested in analyzing only classmate relationships in the 5th grade, the transformation should discard all teacher instances and filter students only to consider those from the desired grade. Some of the possible transformations are discussed in Section 4.2.

4. Social Network Framework – Features

To take advantage from our previously defined models, mechanisms should be implemented to (1) import social networks from existing systems, (2) specify transformations to integrate and

convert models suited for different purposes, and (3) develop algorithms to extract previously unknown information from the represented social networks.

4.1. Importing

One of the main purposes of this work on social network modeling is that social networks that implicitly exist in actual systems may be expressed as true social networks. Developing a proper model to represent social networks, and allowing existing networks to be fit in this model, will allow us the take advantage of specific graph algorithms which can extract valuable information from the existing social networks. The SDM's allows importing data from these systems into our social network framework. Once a given social network is represented as an instance of a SDM, it can be easily transformed into another type of model such as the CM, to take advantage of the SNF's capabilities.

Figure 7 shows an example of a domain model of an existing learning management system [EscolaNaNet].

Instances of this domain model can be mapped into instances of the SDM defined in Figure 3. Importing teachers, students and school class relationships is straightforward from the entities with the same names in the domain model. Relations involving disciplines, school classes and teachers are mapped into teach roles, and registration entities are transformed into learn roles.

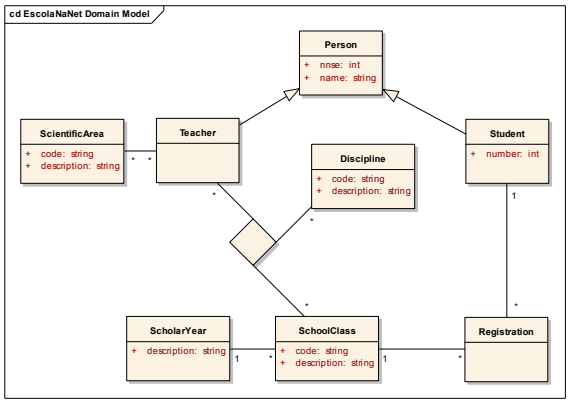


Figure 7 – Domain model of a learning management system in respect to the school class relationship.

4.2. Transformations

The representation we created of our school class SDM enables us to easily import data from an existing application, but in order to apply specific graph algorithms to extract information from the network, we need to transform our model into a canonic model, so that the algorithms can run effectively and efficiently. There are several possible transformations that we can apply to our models. We present the ones which we may use to convert general models into a canonic model, which tend to produce simpler models, but other transformations can also be built to produce more general models from simpler ones. Our proposed transformations are:

1. Projection - a social network model can be projected into a new one, by simply removing certain types of relationships, roles or persons. This way we can discard information which is not relevant to the solution of a specific problem, thus simplifying our model.
2. Selection - a transformation that allows us to obtain more specific networks from existing ones by filtering instances in the network according to a set of conditions enforced on attributes associated with the elements that may be defined in the model.

3. Equivalence Matching - It is possible to infer relationships based on the matching of topological units containing different relationships, roles and person types. In practice, this involves solving the problem of graph structure matching, whose solutions are computationally complex and generally don't scale well.
4. Canonicalization - the transformation of social networks into instances of the canonic model, entails two transformations: (1) simplification, which drops all the information expressed in the source model not contemplated in the canonic model, such as person types, roles and relationships' attributes, and (2) spanning, a transformation that combinatorially spans n-ary relationships into binary relationships.

The first three transformations are applied to convert SDM's into other more restrictive SDM's. The fourth converts a SDM into a CM. As our primary intent is to convert an existing SDM into a CM representation, we can first apply SDM to SDM transformations to produce intermediate models representing the subset of our domain that is pertinent for the intended analysis, and then apply the SDM to CM transformation that allows the algorithms to be applied.

4.3. Algorithms

Several generic graph algorithms exist which can be directly applied to retrieve information from the final canonic representation of our social networks. Some other algorithms specific to the domain of link analysis are especially appropriate to extract information from social networks.

Some generic graph algorithms can be applied being of practical interest in the context of social networks [Cormen 1997]:

1. Searching, which is one of the most common operations that we may want to apply on a social network. This may be breadth-first search or depths-first, although breadth-first search will be often more useful, since it looks first on individuals closer to the search root, which are often more important from a social network point of view.
2. Minimal Spanning Trees, which may be used to discover which individual should be chosen to more rapidly deliver a word-of-mouth message to an entire network. This can be interesting for marketing purposes, since a company could offer a product to clients with high network value, so that they will spread the word about it to others.
3. Shortest Paths, can reveal which is the more efficient way in which an individual may reach another. This algorithm may be useful to choose from the set of people retrieved by a search on a given criteria, the one which can be more easily contacted.
4. Strongly Connected Components, may be of some interest in finding communities. These communities are not strong, since we only know that people in the community are accessible to one another through their contact network, but may not be known to one another. Moreover, if relationship directionality is irrelevant in the social network, then the entire network will be a strongly connected component and in such cases the algorithm has no practical usefulness.

Algorithms from Link Analysis, an area pertaining to the more general research field of data-mining with the purpose of extracting new information on linked structures such as graphs, can be also applied on social networks. These include classification algorithms, which predict the value of nodes based on their links, and clustering algorithms, which attempt to group nodes of a same kind and are of special interest in finding communities on linked structures. Some examples follow:

1. Clustering in Directed Graphs – If our social network relationships are directional and if we believe that the nodes can be classified into two distinct categories such that nodes in one category typically have links to nodes in the other, we can apply algorithms to find these two distinct communities by partitioning the network in its two bipartite components. An example is the HITS algorithm [Kleinberg 1998], proposed by Kleinberg to find authority and hub pages on the hyperlinked structure of the Internet, which can be adapted to solve this problem, also retrieving the most important members in each discovered partition.
2. Clustering in Undirected Graphs – If our social network relationships' direction is irrelevant, maximum-flow algorithms may be used to identify communities conforming to the definition that a community member has more links to members inside the community than to outside members. Flake, Lawrence and Giles [Flake 2000] propose an algorithm to solve this problem, performing efficiently in large scale environments. This approach was proposed to find communities of websites, in order to identify communities of pages related to a topic defined by a starting set of chosen seed pages, but it has an even greater potential on the social network analysis field, since the proposed definition of a community fits better a real community of people than an artificial community of hyperlinked documents.
3. Clustering in Concentric Circles – Another approach on identifying communities in linked structure has been proposed by Zhou et al. [Zhou 2002]. This approach tries to identify communities at different scopes, defined by concentric circles around the some seed set of individuals, thus identifying communities with variable granularity.
4. Emergent Community Retrieval – Kumar et al. [Kumar 1999] discuss an algorithm to find communities in linked structures as soon as they start emerging on the network. This algorithm is based on crawling abilities to search for such emergent communities.
5. Viral Marketing Algorithms – One of the greatest potential of social networks is in identifying the network value of clients for marketing purposes, this activity is known as Viral Marketing [Domingos 2004]. Domingos et al. [Domingos 2001] propose an algorithm to assess this network value, based on markov random fields and bayesian theory.

5. Conclusions and Future Work

From the investigation discussed in this paper we conclude that much work on the area of social network is still open, and that this area has a growing potential which is now starting to be recognized and explored.

A Social Network Framework has been proposed to integrate and analyze existing social networks. Throughout this work, a set of models has been developed which are able to represent social networks extracted from real application scenarios, and mechanisms have been discussed to transform these models into a representation that is suitable to be understood by generic algorithms. We also conclude that several algorithms exist, both in graph literature and in the Link Mining research field, which may be found useful and applicable in the context of social networking, to allow the extraction of the valuable information that these networks enclose.

The discussion presented in this paper opens some research paths that can be taken beyond the current investigation: (1) An interchange format should be discussed and defined in order to allow the representation and exchange of these models, so that integration from different kinds of systems can be made with a reduced effort. (2) The transformations discussed can be explored, and others may be found necessary. These transformations can be concretely defined and implemented, eventually evolving as a social network querying language or another mechanism that will implement the defined transformations. (3) Algorithms that perform over

these networks should be implemented, refined and optimized to be of practical use in the context of social networks. (4) Moreover, real applications should be used to validate the presented models in real case scenarios. Integration can be obtained directly from application databases from specific systems, or by the use of crawling agents over the World Wide Web. (5) Privacy concerns must also be thoroughly evaluated, since they are essential in the assessment of what information can ethically and legally be extracted from social networks. (6) Finally, it would also be of interest the development of new systems, taking advantage of the defined model, transformations and algorithms, to support social network related tasks.

6. References

[Classmates] *Classmates.com*, classmate searching portal, http://www.classmates.com, Classmates Online, Inc. , (October 9, 2006) 1995-2006

[Cormen 1997] T. H. Cormen, C. E. Leiserson, and R. L. Rivest, "Introduction to algorithms", MIT Press and McGraw-Hill Book Company, 18th edition, 1997.

[Domingos 2001] Pedro Domingos, et al., "Mining the network value of customers", In 7th ACM SIGKDD International Conference on Knowledge Discovery and Data mining, 57{66, ACM Press, 2001.

[Domingos 2004] Pedro Domingos, "Mining Social Networks for Viral Marketing", Department of Computer Science and Engineering, University of Washington, 2004

[EscolaNaNet] *EscolaNaNet*, http://www.escolananet.org, INESC-ID, (October 9, 2006) 2006

[Flake 2000] Gary William Flake, Steve Lawrence, C. Lee Giles, "Efficient Identification of Web Communities", in Proceedings of the 6th International Conference on Knowledge Discovery and Data Mining (ACM SIGKDD-2000), 2000

[Getoor 2003] Lise Getoor, "Link Mining: A New Data Mining Challenge", ACM SIGKDD Explorations Newsletter archive, Volume 5, Issue 1 column: Position papers on MRDM, 2003

[Gmail] *Gmail*, e-mailing service, http://www.gmail.com, Google, (October 9, 2006) 2006.

[Hempel 2005] Hempel, J., Lehman, P. "The MySpaceGeneration", cover story in Business Week Magazine, December 12, 2005

[IM] *MSN Instant Messenger*, messaging service, http://messenger.msn.com, Microsoft, (October 9, 2006) 2006

[Kautz 1997] Henry Kautz, Bart Selman, Mehul Shah, "Referral Web: Combining Social Networks and Collaborative Filtering", AT&T Laboratories, in Communications of ACM, vol.40, nº3, March 1997.

[Kleinberg 1998] Jon M. Kleinberg, "Authoritative Sources in a Hyperlinked Environment", in Proceedings of the ACM-SIAM Symposium on Discrete Algorithms, 1998

[Kumar 1999] Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, Andrew Tomkins, "Trawling the web for emerging cybercommunities", in Proceedings of The 8th International World Wide Web Conference, 1999

[MySpace] *MySpace*, personal webspaces portal, http://www.myspace.com, MySpace.com, (October 9, 2006) 2003-2006

[SocialNetwork] *Social Network*, http://en.wikipedia.org/wiki/Social_network, Wikipedia, (May 3, 2006) 2006

[SNList] *List of Social Network Websites*, list of the largest social network websites, http://en.wikipedia.org/wiki/List_of_social_networking_websites, Wikipedia, (May 3, 2006) 2005

[Wooldrige 2002] Michael Wooldridge, "An Introduction to MultiAgent Systems", John Wiley and Sons, ISBN 047149691, 2002

[Zhou 2002] Wei-Jun Zhou, Ji-Rong Wen, Wei-Ying Ma, Hong-Jiang Zhang, "A Concentric Model for Community Mining in Graph Structures", 2002

Proposta de um Modelo para o Processo de Desenvolvimento de um Código de Ética/Conduta/Práticas

Helena Campos

Universidade do Minho (DSI), Guimarães, Portugal

helendoc@sapo.pt

Luís Amaral

Universidade do Minho (DSI), Guimarães, Portugal

amaral@dsi.uminho.pt

Resumo

Um código é a expressão escrita dos princípios de conduta ética. Funciona como um instrumento de gestão para estabelecer e articular valores corporativos, responsabilidades e obrigações e as ambições éticas de uma organização. Tem como função promover um comportamento ético. O código deve ser um documento dinâmico e não estático. E, para se conseguir tal dinamismo, os procedimentos necessários a ter em conta durante o processo de preparação e construção de um código, são relevantes. É um processo faseado que exige a contribuição activa quer da gestão de topo quer das partes interessadas (*stakeholders*). Para que uma organização consiga ter uma gestão da sua ética é necessário ter um programa ético. Muitos autores consideram que o coração de tal programa é precisamente o código. Neste trabalho é apresentada uma proposta de um modelo para o processo de desenvolvimento de um código, que oferece passo a passo, as directrizes necessárias para a sua construção.

Palavras chave: Valores, ética, código de ética, código de conduta, código de práticas, gestor de ética, comité de ética, processo de desenvolvimento de um código

1. Introdução

O mundo corporativo está em contínua mutação. Assiste-se a uma mudança que se deve a uma ampliação significativa dos parâmetros de impacto nos negócios**.** A começar pela quantidade e diversidade de relacionamentos: clientes cada vez mais exigentes e participativos, fornecedores integrados na cadeia de produção influenciando nas decisões, funcionários mais preocupados com qualidade de vida e outras formas de remuneração, entre outros. A ampliação está a acontecer em diversas dimensões: económica, ambiental, social e até humana, considerando-se que estamos em última instância a falar de relacionamento entre pessoas. Mas, não podemos

esquecer de que na base de tais dimensões se encontra a dimensão tecnológica e de que esta só foi possível com a intervenção humana [Campos 2006a]. Paralelamente a esta dinâmica, também não nos podemos esquecer, do perigo eminente da Sociedade de Informação globalizada, de que se a dimensão tecnológica não for bem dirigida e planeada – planeamento e gestão eficientes dos Sistemas de Informação (SI) e a correcta utilização das Tecnologias de Informação (TI) - todas as outras dimensões correm o risco de serem seriamente afectadas, uma vez que, a sua mobilização depende da mobilização tecnológica [Campos 2006a].

A gestão de qualquer área de negócio, quer seja, produção, marketing, financeira, recursos humanos, sistemas de informação, ou outra, requer uma série de actividades que cubram o planeamento, a implementação e o controle através de meios formais e informais. Paralelamente também, todo o negócio deveria integrar na sua gestão, técnicas e instrumentos típicos comuns de gestão da ética (declarações de valores, códigos de ética, comités ou gestores de ética, programas de responsabilidade social, entre outros) cujo o papel principal será a gestão do comportamento ético dos empregados, a gestão das relações com os *stakeholders* e a avaliação do desempenho ético. Dentro dos instrumentos citados, destacam-se os códigos de ética por serem declarações que se preocupam em delinear explicitamente qual o tipo de conduta desejado e esperado dos profissionais do ponto de vista ético dentro de uma organização, profissão ou indústria [Campos 2006a].

De salientar que a organização continua a ser o local onde as pessoas passam grande parte do seu tempo, correspondendo a um terço da sua vida produtiva [Gonçalves ?]. Assim sendo, motivar os profissionais para que desempenhem com ética, competência, responsabilidade e com satisfação a sua actividade profissional, representa hoje um grande desafio para qualquer liderança organizacional. Sendo vital, o reconhecimento por parte desta, de que: (1) a ética não é uma opção mas uma necessidade; (2) mais do que condenar, promove; (3) é um compromisso de valores duradouros; e (4) de que a existência de um Código de Ética/Conduta/Práticas Profissional, é uma componente essencial indispensável para o exercício livre e responsável de qualquer profissão digna de "confiança pública" [Moreira 2004].

Dada a importância dos códigos, o presente trabalho tem por finalidade a proposta de um modelo para o processo de desenvolvimento de um código, tendo como visão, a construção de um código para os profissionais de Tecnologias e Sistemas de Informação (TSI) em Portugal, inexistente até ao presente.

2. Código de Ética vs Conduta Vs Práticas

São inúmeras as definições e as funções de um código (Código de Ética (CE), Código de

Conduta (CC), Código Deontológico (CD), Standards de Conduta (SC), Código de Práticas (CP), entre outros. Existe uma variedade de títulos para os códigos, e uma definição definitiva não é providenciada pela literatura [Campos 2006b]. Na sua essência, podem ser todos descritos como linhas de orientação com o objectivo comum de promover um comportamento ético.

Alguns são declarações resumidas ou versões condensadas que reúnem (descrevem) um conjunto de princípios éticos fundamentais e de aspirações ou ideais ao mais alto nível de abstracção – Códigos de Ética (CE). Outros, são versões com um maior nível de detalhe das declarações dos valores e princípios que definem o propósito da organização – Códigos de Conduta (CC). Daqui poderemos considerar que o CE e o CC se complementam, dado que o CC é a versão mais completa do CE, destinando-se a ser mais rigoroso na prática tendo um carácter de maior obrigatoriedade. E ainda outros, como guias de orientação mais específica (CP), visando a aplicação de "boas práticas" no ambiente profissional, preocupando-se com a regulamentação de como o profissional deve aplicar os domínios específicos das competências e respectivos *skills* profissionais, na sua prática diária. São mais dinâmicos do que os anteriores, devido ao rápido desenvolvimento do contexto técnico, implicando uma revisão mais frequente [Campos 2006b].

Não é intenção deste trabalho indicar qual a melhor denominação para o título específico do código. Cada grupo profissional que desenvolva o seu próprio código deve tomar a decisão de como o intitular [Berleur et al. 2004].

Partindo do princípio de que os códigos são declarações voluntárias que comprometem organizações, indústrias, ou profissões a especificar crenças, valores e acções e/ou demonstram o comportamento ético apropriado para os seus profissionais há que reflectir sobre o tipo de código que se pretende construir. Crane e Matten [2004] definiram 4 tipos principais de códigos: corporativos ou organizacionais, profissionais, industriais e os universais. Contudo, verificou-se através da literatura que as propostas para os passos do processo de desenvolvimento de um código, são independentes do tipo de código ou do título. Assim sendo, e, por uma questão de simplificação será utilizado o termo "código" de uma forma genérica.

É importante notar que a primeira tarefa a realizar para a construção de um código, é a definição exaustiva de todos os passos necessários a ter em conta antes, durante e depois da sua construção – **processo de desenvolvimento do código**. Este processo consiste na articulação de mecanismos possíveis para a construção e sustentabilidade do código e para a renovação das energias que levaram ao seu desenvolvimento. De realçar ainda que uma vez construído o

código, é necessário existir uma dinâmica de suporte ao código. Se tal não acontecer, a pressão que inicialmente conduziu à sua criação pode dissipar-se.

3. Processos de Desenvolvimento de um Código / Análise comparativa

Desenvolvimento de um código

Da análise conjunta das diferentes propostas para o processo de desenvolvimento de um código, não se encontrou um processo "padrão" para os passos a seguir.

A tabela 1 sintetiza todos os passos propostos pelas diferentes aproximações analisadas e a respectiva ordem de execução. Na referida tabela, pode verificar-se uma multiplicidade de passos contemplados quer em número quer em conteúdo.

Da análise da tabela há a destacar:

- o passo "Implementação", contemplado em metade dos estudos, não aparece no estudo de *Simone Webley* [2001a] porque o considerou como um processo a ser tratado separadamente devido ao seu grau de importância [Webley 2001b]. Também a *Joint Initiative* [1998] subscreveu que o passo "Implementação" merece ter um tratamento à parte, considerando no entanto dever ser um dos passos obrigatórios no processo de desenvolvimento de um código.
- o passo "Decidir estratégias de comunicação e educação" é intitulado por "Publicação e disseminação do código" nos 4 estudos restantes, tratando-se do mesmo passo com títulos diferentes;
- a identificação de 9 passos coincidentes nos processos para o desenvolvimento de um código: os 6 primeiros em maioria e os últimos 3 em metade dos estudos (ver Figura 1):

(1) Criação de um grupo de trabalho
(2) Reunir informação
(3) Esboço preliminar do código
(4) Testes e revisões finais do código
(5) Publicação e disseminação do código
(6) Revisão e actualização do código
(7) Começar por conseguir apoio da liderança
(8) Consultas/Pareceres sobre o esboço preliminar
(9) Implementação

Figura 1- Passos coincidentes nas várias aproximações

Passos para o Processo de Desenvolvimento de um Código	Berleur et al. (2004)	Driscoll et al. (2000)	Joint Initiative (1998)	Murray (2000)	SHRC (2001)	Webley (2001a)
▪ **Encontrar um "campeão"**						1º

▪ **Determinar os objectivos e significado do sucesso**		1º			1º	
▪ **Começar por conseguir apoio da liderança**		2º			2º	2º
▪ **Perceber quais as preocupações das pessoas**						3º
▪ **Contexto**				1º		
▪ **Criar um grupo de trabalho (*task force*)**	3º	3º	3º		3º	
▪ **Reunir informação**	1º	4º	1º		4º	
▪ **Discussões preliminares com os principais *stakeholders***	2º		2º			
▪ **Seleccionar ou não um modelo bem testado**				2º		4º
▪ **Manter a liderança informada**		5º			5º	
▪ **Fazer um esboço preliminar do código**	4º	6º	4º	3º	6º	
▪ **Submeter primeiro o código para revisão à liderança**		7º			7º	
▪ **Consultas/Pareceres sobre o esboço preliminar**	5º		5º	4º		
▪ **Apresentar o código internamente (à organização)**						5º
▪ **Testar o código e fazer revisões finais**		8º		6º	8º	6º
▪ **Publicação e disseminação do código**	6º		6º	8º		7º
▪ **Ter o código revisto por um parecer legal**		9º			9º	
▪ **Obter aprovação da versão final pela alta direcção**		10º		5º	10º	
▪ **Conformidade na adopção do código**				7º		
▪ **Decidir estratégias de comunicação e educação**		11º			11º	
▪ **Fazer o código trabalhar (exemplos práticos do código)**						8º
▪ **Implementação**	7º		7º	9º		
▪ **Revisão e actualização do código**	8º	12º	8º		12º	

O modelo da Join Initiative [1998] é baseado numa série de casos de estudo e foi adoptado pela IFIP (*International Federation for Information Processing*) como guia para o desenvolvimento e implementação de um código [Berleur et al. 2004].

Tabela 1 – Processos de desenvolvimento de um código – Análise comparativa

4. Proposta de um modelo para o processo de desenvolvimento de um código eficaz

O processo seguido no desenvolvimento de um código é tão importante como o próprio código [Berleur 2004].

Um grupo de interesses especiais (SIG9.2.2) recomenda fortemente um processo de consulta e *feedeback* no desenvolvimento de um código – adepto da ideia de permitir "espaços para discussão" [Berleur et al. 1996].

A definição do processo de desenvolvimento de um código é uma tarefa significativa em si mesmo. É a plataforma de base que define exaustivamente todo o processo de encontrar os

requisitos para a comunicação eficiente de padrões de organização da ética. Um código por si só não chega. A participação e conformidade de todos os *stakeholders,* para que mais tarde o código não seja questionado; revisões e actualizações contínuas; formação e aconselhamento na sua utilização, são cruciais para garantir a continuidade e o funcionamento do código.

Como vimos, existem diferentes procedimentos na preparação de um código. Os códigos tratam de assuntos variados e foram criados por diversas razões. Tal como em qualquer processo de desenvolvimento, não há a garantia de que contemple todos os passos completos e perfeitos para que resulte num código de sucesso [Join iniciative 1998].

O modelo proposto (ver esquema I), resulta da análise comparativa de propostas de aproximações validadas para o processo de desenvolvimento de um código, e do estudo de literatura adicional na área. De uma forma exaustiva e abstracta do tipo de código, indica quais os passos a ter em conta em todo o processo de preparação para a construção de um código. Caso se trate da construção de um código voluntário (não desenvolvido pela organização) os passos 2, 6, 9, e 11 serão excluídos, uma vez que não há a participação directa da liderança na construção do código.

O processo de desenvolvimento de um código inicia-se pelo exercício de uma reflexão acerca de um conjunto de questões críticas a considerar (passo 0). Obtidas as respostas a estas questões, a tarefa seguinte será encontrar alguém preparado para a elaboração do código – um *expert* ou "campeão" [Webley 2001a], ou alguém da gestão sénior com o perfil adequado (o ideal seria um CEO caso não se trate de um código voluntário) para dirigir a introdução da política ética do negócio.

Passo 1. Decidir objectivos e significado de sucesso

Decida os objectivos e o significado de sucesso para um programa ético, o mais cedo possível. Um código como um documento isolado não é eficaz. Os seus benefícos resultam do seu uso estratégico como uma parte das principais iniciativas éticas do negócio. O desenvolvimento de um código deve basear-se num conjunto de objectivos específicos que vão além do próprio código. Devem permitir que os empregados usem o código dirigido às suas próprias preocupações éticas e promovam os objectivos éticos do negócio.

Passo 2. Começar por conseguir apoio por parte da liderança

Ética e valores corporativos são matéria de política governativa. Sendo assim necessário o suporte e o compromisso por parte dos líderes para o sucesso no desenvolvimento do código. Como tal, é crucial, conseguir antecipadamente o compromisso da liderança e comunicá-lo a toda a organização.

Passo 3. Reunir informação

Para um código ser eficaz, os profissionais devem sentir que o código é relevante. A informação recolhida (respostas às questões básicas iniciais) através de entrevistas, *focus groups* e discussões informais pode ser importante para direccionar questões chave que vão formar a plataforma para diálogos acerca da ética. A análise da informação reunida permitirá tomar conhecimento das preocupações e questões éticas chave por parte dos profissionais; perceber quais os tópicos em que estes requerem orientação; obter sugestões que indicarão o modo de criar um código mais útil e ajudar no desenvolvimento do esboço e conteúdo do código.

Passo 4. Discussões preliminares com os principais stakeholders

Testar preliminarmente impressões e opiniões com os principais *stakeholders* das conclusões obtidas na fase da recolha da informação, e identificar os parceiros dispostos a ajudar no desenvolvimento do código. Pode ser útil obter outras opiniões além das dos colegas, dos colaboradores e da(s) organização(ões) - incluir de um modo mais alargado interesses éticos dos consumidores, organizações ambientais e sindicais, associações profissionais e governamentais e ONG´s [Joint Initiative 1998]. Tais opiniões poderão ajudar na confirmação de percepções de interesses e preocupações éticas iniciais, bem como adicionalmente levar a uma maior participação de pessoas e organizações no desenvolvimento do código.

Passo 5. Criar um grupo de trabalho (Task Force)

É essencial ter as pessoas certas no grupo de trabalho. Devem ser profissionais seguros, credíveis, conhecedores do assunto, representantes dos diferentes *stakeholders* que vão ser afectados pelo código, e com tempo disponível para se entregarem exclusivamente a este trabalho.

A criação de uma força de trabalho ou comité atribuindo-lhe responsabilidades pessoais e cívis é crucial para desenvolver um código [Silva 2006]. Poderão surgir discussões abertas acerca dos objectivos do grupo, responsabilidades de cada participante, cargas de trabalho e decisões de como o grupo vai funcionar. Recomendando-se actualizações regulares entre o grupo de trabalho.

Passo 6. Manter a liderança informada

Revelar finalmente à liderança, as impressões acerca da compreensão e da visão que os *stakeholders* chave têm do ambiente corrente. Na sua divulgação, não deve ser feita qualquer indicação da(s) fonte(s) nem de quem especificou quaisquer comentários ou impressões.

Passo 7. Esboço preliminar do código

Na preparação preliminar do esboço, o grupo deve identificar quem será responsável pelos aspectos específicos da implementação. Por exemplo, determinados profissionais podem executar as funções chave, e uma nova estrutura de gestão pode ser criada para assistir na implementação. Na medida do possível, todas as pessoas que tenham sido identificadas como profissionais chave no desempenho de funções chave, devem fazer parte das conversas do grupo de trabalho, por forma a que as suas idéias, potencialidades e apoio, ou resistência, possam ser contempladas nos termos do código em esboço.

O grupo de trabalho necessita de decidir os valores, provisões, estruturas, etc, para o conteúdo do código. A importância da escrita e da definição da estrutura do código, terá de ser objecto de um estudo posterior.

Desenhar um código é como que "costumizar" a organização e servir as suas necessidades específicas. É então fundamental, fazer bom uso da informação recolhida dos colaboradores, voluntários e outros *stakeholders* chave.

Passo 8. Consultas sobre o esboço preliminar

Consultas completas no começo podem prevenir problemas mais tarde.

Uma boa aproximação será começar pelos que provavelmente irão ser afectados directamente pelo código e que já tenham conhecimento dele. Para os que ainda não têm conhecimento do código, deverá proceder-se formalmente à sua divulgação de um modo individual ou em grupo. Um plano de consulta pode ser útil - deve incluir funções para profissionais de elevado perfil e com bons *skills* de comunicação, para explicarem o código e recolherem opiniões e/ou sugestões. Este *feedback* permitirá ao grupo de trabalho discutir e reflectir a melhor forma de como tais comentários deverão ser incluídos na revisão da versão final do código, fortalecendo não somente o código, como também, permitindo que todos os que forneceram *feedback*, sintam em parte o código como seu.

Passo 9. Submeter o código à liderança para revisão

Assegurar o suporte da liderança através da sua presença na revisão e comentários da versão final do código. Este passo permite que os líderes mantenham um papel participativo no processo de desenvolvimento do código, mesmo que diariamente não estejam activamente envolvidos. Tal participação reforça o comprometimento total na iniciativa da construção do código.

Passo 10. Testar o código e fazer quaisquer revisões finais

O código deve ser validado através de uma amostra que integre profissionais de todos os níveis e de diferentes posições, e/ou por uma fonte externa à organização. Dá-lo novamente a ler para quaisquer sugestões e revisões finais. Há sempre detalhes que são esquecidos no esboço inicial.

Passo 11. Obter aprovação da versão final por parte da liderança

Antes de ser publicado e disseminado, a versão final do código deve ser formalmente aprovada pelo conselho executivo (pela direcção).

Passo 12. Públicação e disseminação do código

Decidir a estratégia de comunicação - tornar o código conhecido e pô-lo a circular. Esta estratégia deverá ter no mínimo os 4 objectivos que se seguem [ERC´s 2004]: (1) assegurar que todas as partes foram informadas e receberam o código; (2) que todas as partes compreenderam a sua responsabilidade pessoal para aceitar e executar as normas e prescrições residentes no código; (3) que o compromisso da organização é claro e não ambíguo para todas as partes e (4) assegurar que todas as partes estão expostas a exemplos abundantes da utilidade do código.

É importante um bom plano de comunicação para dar a conhecer o código e informar como deve funcionar eficazmente [Joint Initiative 1998]. Não esquecer que um código tem um enorme impacto na expansão de futuras comunicações éticas (formação ética, artigos, *newsletter*, *posters*, *e-mails*, discussões formais ou informais acerca da ética) para formular ou enfatizar questões chave contempladas no código.

Recomenda-se que o código seja anunciado num evento especial, envolvendo a liderança da organização e outras partes que contribuiram para a sua criação. No evento convém realçar que a distribuição do código não é um fim em si mesmo, mas mais um passo chave para o progresso do processo [SHRM 2001].

Deve-se divulgar que a organização tem um código e um programa de implementação. Bem como, colocar o código na *web,* no site da organização, e enviá-lo a parceiros do negócio [Webley 2001a].

Passo 13. Implementação

É importante notar que a implementação - fazer o código trabalhar - faz parte do processo de desenvolvimento de um código. Os códigos são acontecimentos, avanços e documentos flexíveis que devem ser actualizados de modo a reflectirem novas circunstâncias e percepções. É crucial assegurar a efectiva concordância do código pelo envolvimento de todos os seus participantes.

Exemplos práticos da implementação do código e respectivos programas de formação devem ser introduzidos internamente em toda a organização. Deve ser nomeado um "gestor" ou "coordenador" do código para garantir a sua sustentabilidade.

Este passo devido ao seu grau de importância merece um tratamento mais promenorizado, exigindo um maior detalhe para que a implementação do código seja eficaz. Como tal, será objecto de um estudo posterior.

14. Revisão e actualização regulares

O código não deve ser visto como um documento estático. Deve ser revisto regularmente tendo em conta mudanças que possam ocorrer no ambiente organizacional ou operacional. Para medir o impacto do código são necessários dados de situações em que tenha sido aplicado e com que frequência. Dados estes que podem ser utilizados não só para determinar fraquezas na elaboração e funcionamento do código, mas também, para demonstrar quaisquer resultados atingidos e o progresso conseguido através da sua utilização.

Pode ser útil usar os passos 3, 4 e 5 do modelo proposto (ver esquema I) como um processo de revisão [Joint Initiative 1998]:

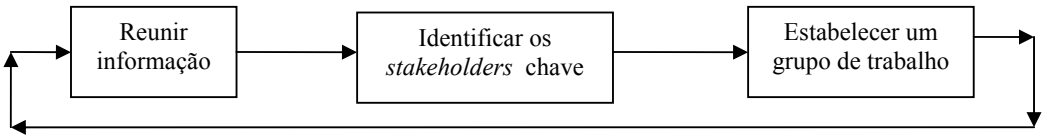
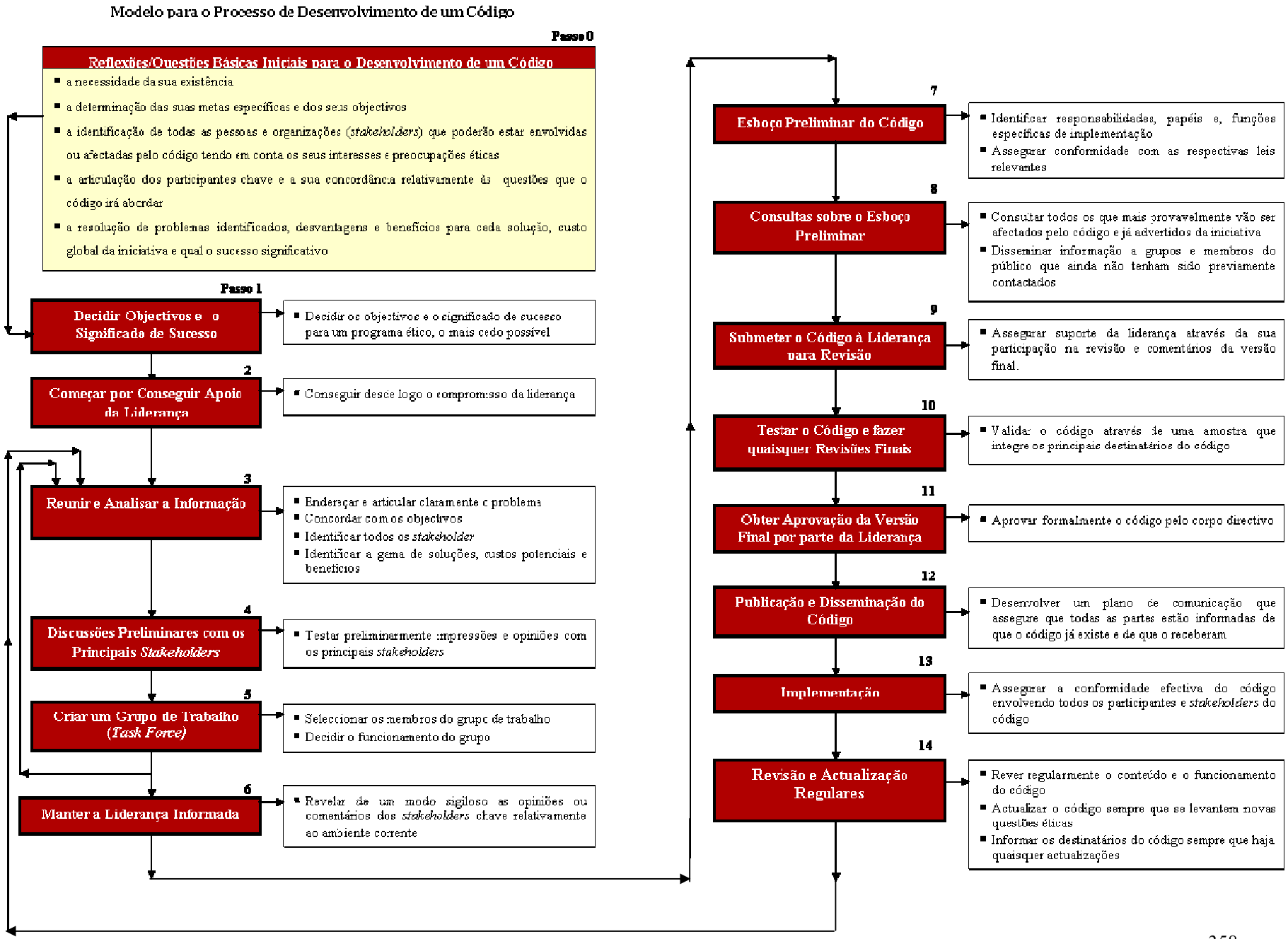


Figura 2 – Processo de revisão

As revisões periódicas do código podem enviar uma mensagem simbólica para todos os colaboradores da organização: de que a ética é uma área que requer uma atenção contínua, sentirem o empenho da organização em atingir um estado de excelência ética e permitir-lhes um ambiente saúdavel no desempenho das suas competências



Esquema I – Modelo para o processo de desenvolvimento de um código

5. Conclusões

A construção de um código não é um processo ad-hoc, deve obedecer a um conjunto de passos/ procedimentos. O modelo proposto para o processo de desenvolvimento de um código contempla a hipótese de o código ser construído ou não pela organização, i.e., uma organização que pretenda construir o seu código, deve seguir os 15 passos do modelo. Caso se trate da construção de um código voluntário (não desenvolvido pela organização), devem seguir-se somente os 11 passos que se seguem – 0, 1, 3, 4, 5, 7, 8, 10, 12, 13 e 14 (ver esquema I). A exclusão dos passos 2, 6, 9, e 11, deve-se ao facto de não haver a participação da liderança na construção do código.

A selecção dos elementos que vão fazer parte do grupo de trabalho deve ser cuidadosa e rigorosa devido à sua responsabilidade durante todo o processo: construir o código, divulgar os padrões de conduta profissional fora e dentro da organização, avaliar e esclarecer dúvidas sobre questões éticas e julgar casos de violação do código.

O código deverá ser construído com base nas informações recolhidas dos vários *stakeholders*.

Numa organização que deseje construir o seu próprio código, o compromisso e o envolvimento da liderança no processo de desenvolvimento do código é essencial [Berleur et al. 2004], [Birkett 1999], [Brown 2003], [Discrcoll et al. 2000], [Joint Initiative 1998], [Murray 2000], [Navran 2003], [SHRM 2001], [Webley 2001a].

O desenvolvimento de um código é essencialmente um processo de aprendizagem [Joint Initiative 1998]. Uma vez implementado o código, a confiança, a comunicação e a experiência podem levar à criação de obrigações mais rigorosas, estruturas de responsabilidades podem ser corrigidas, revisões regulares podem melhorar o conteúdo do código, conseguindo-se aumentar a sua credibilidade e eficácia.

O processo de desenvolvimento de um código deve ser flexível. As concepções originais do problema e soluções possíveis poderão mudar devido a novas e emergentes informações e à mudança de circunstâncias. A flexibilidade e o alerta para a necessidade de mudança podem ajudar a modificar expectativas e resultados.

É importante que todos os que são afectados pelo código sejam advertidos da sua existência e participem na sua elaboração.

Depois de construído o código, é vital que sejam asseguradas estratégias de comunicação, de formação e de actualização para que o código continue a reflectir acerca das prioridades éticas

da organização [ERC´s 2004]. Tarefas que podem ser asseguradas pelo comité ou gestor de ética criado no processo de desenvolvimento.

O sucesso da escrita e do conteúdo do código irá depender da excelência conseguida no seu processo de desenvolvimento.

6. Trabalhos Futuros

A **implementação** do código (pôr o código a trabalhar); a **escrita** (o código ser de fácil leitura e compreensão) e a **estrutura do conteúdo** (parte substancial do código), devido ao seu grau de importância, serão objecto de estudos posteriores.

Bibliografia

Berleur, J. e M. d´Udekem-Gevers, *Codes of Ethics or of Conduct Within IFIP and Others Computer Societies*, em *Ethics of Computing: Codes, Spaces for Discussion and Law*, Berleur, J. e Brunnstein, K., (Eds.), London: Chapman & Hall (now Kluwer), 1996, 3-41.

Berleur, J. Duquenoy P. Holvast J. Jones M. Kimppa K. Sizer R. e Whitehouse D., *IFIP Framework for Ethics of Computing. Criteria and Prodecures for Developing Codes of Ethics or of Conduct*, IFIP-SIG9.2.2., 2004.

Birkett, W. e Barbera, M., *Codifying Power and Control – Ethical Codes, Codes in Action Contents*. Financial and Management Committee (FMAC) of the International Federation of Accounts (IFAC), The Australian Centre for Management Accounting Development, University of New South Wales, Sydney, Australia, 1999.

Brown, J., *Ten Writing Tips for Creating an Effective Code of Conduct* http://www.ethics.org/resources/code_writing.html (26 de Setembro de 2006), 2003.

Campos, H. D. O., Gestão Ética nos Negócios: Não Há Responsabilidade Social sem Ética nos Negócios, *Recursos Humanos Magazine*, 44, 2006a.

Campos, H. D. O., *Codes of Ethics in IST - Comparative Analisys*, Softwars 2006, Universidade do Minho, Portugal, 2006b.

Crane, A. e Matten, D., *Business Ethics: A European Perspective – Managing Citizenship and Sustainability in Age of Globalization*, Oxford University Press, 2004.

Driscroll, D. e Hoffman, W.M., *Ethics Matters: How to Implement Values-driven management,* http://www.ethics.org/resources/code_construction.html (25 de Setembro de 2006), 2000, 77.

ERC´s, Ethics Resource Center, *Creating a Workable Company Code of Conduct - What to Do After Your Code of Conduct is Written*, http://www.erc.org (6 de Outubro de 2006), Ethics Today Online, volume 2: September 2003 to July 2004.

Gonçalves, E. L., *A Empresa No Mundo Moderno*, http://www.fides.org.br (6 de Outubro de 2006), ?.

Joint Initiative of the Office of Consumer Affairs, Industry Canada, and the Regulatory Affairs Division, Treasury Board Secretariat, *Voluntary Codes: A Guide for Their Development and Use*, http://strategis.ic.gc.ca/volcodes (26 de Setembro de 2006), 1998.

Moreira, J. M., *A Ética no Contexto da Sociedade do Conhecimento e da Informação*, APDSI (Associação para a Promoção e Desenvolvimento da Sociedade da Informação), Ordem dos Engenheiros, 2004.

Murray, H. and Aitken, R., *Creating a Code of Ethics: Developing Ethical Standards for a Sector*, Canadian Council for International Co-operation (CCIC), 2000.

Navran, F., *Seven Steps For Changing the Ethical Culture of an Organization,* http://www.erc.org (3 de Outubro de 2006), Ethics Today Online, volume 1: September 2002 to July 2003.

SHRM, Society For Human Resource Management, *Code of Ethics Toolkit: A Guide To Developing Your Organizations´s Code of Ethics,* http://www.shrm.org/ethics/chapter-coe.pdf (26 de Setembro de 2006), 2001.

Silva, A. M. C., *Crie Uma Comissão de Ética*, http://www.ethos.org.br (5 de Outubro de 2006), 2006.

Vargas, R. *Os Meios Justificam os Fins – Gestão Baseada em Valores: da Ética Individual à Ética Empresarial*, Gradiva, 2005.

Webley, S., *Eight Steps for a Company Wishing To Develop Its Own Corporate Ethics Programme,* http://www.ibe.org.uk/developing.html (2 de Outubro de 2006), 2001a.

Webley, S., *Twelve Steps for Implementing a Code,* http://www.ibe.org.uk/implementing.html (2 de Outubro de 2006), 2001b.

O jogo e a alternância de papéis formando/formador como facilitadores do processo de aprendizagem nas organizações

Vítor Santos[1], Luís Amaral[2]

1) Microsoft

vitors@microsoft.com

2) Universidade do Minho

amaral@dsi.uminho.pt

Resumo

Neste artigo é proposta uma nova estratégia para formação on-line de membros de organizações de média e grande dimensão, baseada na partilha e troca de conhecimentos entre esses membros. A estratégia seguida suporta-se na criação de comunidades on-line intra-empresarias, onde todos os membros assumem, simultaneamente, o papel de formando e de formador. A interacção na comunidade é obtida através de um jogo, em que cada participante desafia os outros a aprenderem o que ele tem para ensinar na área de sua especialidade profissional, partilhando para isso informação e recursos sobre as matérias que domina e é, ao mesmo tempo, desafiado a aprender matérias de outras áreas profissionais. A adopção deste modelo permite melhorar, de forma lúdica e desafiante, o nível de conhecimento dos membros de uma organização contribuindo para que os seus membros cooperem melhor e ,em consequência, que esta se torne mais competitiva e resilente.

Palavras-chave: organizações aprendentes; comunidades de aprendizagem; *e-learning*; jogo;

1. Introdução

O papel das empresas no desenvolvimento das competências individuais é, hoje, unanimemente considerado vantajoso, tanto para os trabalhadores como para as empresas.

Para desempenho deste papel, uma alternativa importante a considerar é a utlização de tecnologias de informação e comunicação, baseadas em Internet vocacionadas para o ensino [1].

A utilização destas tecnologias justifica-se pelas vantagens que as mesmas trazem [2] em termos da flexibilidade no ensino, bem como pelos baixos custos associados, permitindo, na maior parte dos casos, uma rápida ampliação da oferta formativa que os colaboradores podem realizar no seu posto de trabalho.

Tradicionalmente, a formação on-line assenta na lógica formador/formando na qual os papéis do formador e dos formandos se encontra totalmente distinguido. Tal estratégia de formação on-line, a que podemos chamar tradicional, não se mostra ser adequada para um ambiente de partilha e troca de conhecimentos entre diferentes colaboradores, dificultando por isso que a formação on-

line seja usada para ajudar a que o conhecimento de colaboradores mais experientes e especialistas seja passado a outros.

É, contudo, possível recorrendo à tecnologia, implementar outras lógicas diferentes da convencional como por exemplo a de construção de comunidades de aprendizagem, em que os participantes desenvolvam de forma lúdica as competências da alternância de papéis entre formando e formador e com isso permitir a partilha e troca de conhecimentos [3].

Neste artigo propomos uma estratégia para formação on-line de membros de organizações de média e grande dimensão baseada na partilha e troca de conhecimentos entre esses membros.

Esta estratégia possibilitará que cada colaborador desempenhe alternadamente o papel de formador e de formando, participando numa comunidade on-line, intra-empresarial onde, através de um jogo, cada participante desafie os outros a aprenderem o que ele tem para ensinar na área da sua especialidade profissional, partilhando para isso informação e recursos sobre as matérias que domina e é, ao mesmo tempo, desafiado a aprender matérias de outras áreas profissionais.

2. A aprendizagem nas organizações

O "conhecimento" é um dos principais activos das organizações e a capacidade de o manter e aumentar constitui um desafio.

Este desafio assenta, essencialmente, na qualificação dos recursos humanos e na capacidade de reter e fazer circular o conhecimento na organização criando condições que favoreçam o desenvolvimento das competências dos respectivos colaboradores.

Uma condição necessária é o envolvimento de todas as pessoas da organização, na edificação de uma cultura de aprendizagem colectiva, que estruture uma organização aprendente, onde os conceitos de colaboração, partilha, aprendizagem e conhecimento são chave.

As organizações que se propõem a construir estruturas e estratégias visando a dinamização e aumento do potencial de aprendizagem têm sido denominadas por organizações aprendentes "Learning Organizations".

Senge [4], define Learning Organizations como organizações onde as pessoas estão focadas na aprendizagem colectiva e comprometidas com resultados motivadores. Segundo este autor, as organizações que aprendem são "organizações nas quais as pessoas expandem continuamente a sua capacidade de criar os resultados que realmente desejam, onde se estimulam padrões de pensamentos novos e abrangentes, a aspiração colectiva ganha liberdade e onde as pessoas aprendem continuamente a aprender em conjunto"

Para Friedman [5], por organização aprendente entende-se uma organização que "está num processo permanente de mudança, porque os trabalhadores são estimulados a realizar continuamente modificações e adaptações. Uma organização aprendente [...] concentra-se na criação, na aquisição e na transmissão de conhecimentos e na adaptação dos comportamentos em função desses conhecimentos"

É de importância fundamental para uma organização, ao assumir o modelo "Learning Organization", que esta possa contar com um ambiente que incentive a aprendizagem, com informações fluidas e claras, tranversais a toda a estrutura da organização, que os membros partilhem do mesmo objectivo e que a empresa conte com uma visão de gestão alinhada.

As organizações aprendentes são particularmente vantajosas quando se encontram inseridas em ambientes imprevisíveis e turbulentos, como são os casos dos mercados das tecnologias. Um ambiente que propicie a aprendizagem, facilita a troca e a aprendizagem com o ambiente externo, reforça comportamentos receptivos e não-defensivos, e facilita a recuperação e retenção intencional do conhecimento.

A procura de estratégias que ajudem aumentar, nomeadamente com recurso às tecnologias da informação, o nível de conhecimento dos membros de uma organização, contribuindo para que os seus membros cooperem melhor e, em consequência, que esta se torne mais competitiva e resilente tem, portanto, um crescente interesse.

3. Modelo Conceptual

O modelo que propomos tem por principal objectivo possibilitar a criação de comunidades on-line intra-empresarias, onde todos os membros, de forma lúdica, assumem, alternadamente, o papel de formando e de formador partilhando dessa forma os seus conhecimentos com todos e conseguindo aprender com todos.

A interacção nesta comunidade é obtida através de um jogo, em que cada participante desafia os outros a aprenderem o que ele tem para ensinar na áreas da sua especialidade profissional, ou eventualmente, pessoal, partilhando para isso informação e recursos sobre as matérias que domina e é, ao mesmo tempo, desafiado a aprender matérias fornecidas por outros membros.

A partilha dos recursos é feita on-line através de uma submissão onde é indicada a área ou áreas de especialidade que o preponente sugere para classificar os recursos.

Entendem-se, neste contexto, por recursos de aprendizagem textos de apoio, imagens, vídeos, *webcasts*, *links*, referências bibliográficas, dicas, etc. Juntamente com os recursos são submetidos questionários de avaliação coerentes com as matérias de especialidade dos recursos fornecidos que irão validar as aprendizagens construídas pelos "formandos".

A validação das matérias sobre o ponto de vista da adequabilidade do contéudos e dos respectivos questionários é feita, após submissão, por um ou mais índividuos denominados "Gurus" que podem ser externos ou internos à organização.

A validação visa garantir que os conteúdos se encontram correctos sobre o ponto de vista técnico e que são suficientes para responder aos questionários. Na figura 1 apresentamos um exemplo de uma grelha de validação de recursos de um "Guru".

O "Guru" tem também a tarefa de classificar os conteúdos e respectivos questionários relativamente à sua pertinência e posicionamento na área de especialidade e a aspectos didáticos e pedagógicos. Esta avalição atribui uma pontuação ao conteúdo. Na figura 2 apresentamos um exemplo de uma grelha de avaliação de recursos de um "Guru".