

Avaliando a Veracidade das Histórias sobre Documentos

Daniel Gonçalves Joaquim A Jorge

Dep. Eng^a. Informática, IST

Av. Rovisco Pais, 1000 Lisboa

daniel.goncalves@inesc-id.pt, jaj@inesc-id.pt

Resumo

Um problema cada vez mais encontrado pelos utilizadores de computadores é localizar um determinado documento electrónico. As formas tradicionais para organizar os documentos são cada vez menos eficazes ao lidar com a quantidade de informação que até mesmo um utilizador comum tem ao seu dispor hoje em dia. Adicionalmente, são soluções artificiais, obrigando à classificação dos documentos em categorias arbitrárias no sistema de ficheiros, com todos os problemas que isso implica. É urgente o desenvolvimento de novos mecanismos de recuperação de documentos que reflectam a forma de pensar dos utilizadores quando se recordam dos seus documentos, recorrendo a informação autobiográfica com eles relacionada.

Na nossa investigação, verificámos que as histórias sobre os documentos são uma forma simples e natural dos utilizadores se lhes referirem. Um conjunto de entrevistas permitiu-nos conhecer qual a forma e conteúdos típicos dessas histórias. Através da avaliação de protótipos de baixa fidelidade, construídos com base nas características identificadas, identificámos a forma que uma interface para a recuperação de documentos baseada em narrativas pode tomar.

Restava, no entanto, responder a uma questão crucial: são as histórias suficientemente precisas na descrição dos documentos? Neste artigo descrevemos um estudo em que uma versão funcional de um protótipo de interface para capturar histórias sobre documentos foi usada para recolher trinta histórias sobre documentos reais. A informação nessas histórias foi então comparada com os documentos propriamente ditos. Verificámos que, no geral, as histórias descrevem os documentos de forma fidedigna (81% a 91% de toda a informação é correcta), validando as histórias como forma de recuperação de documentos. Confirmámos, também, que os utilizadores conseguem contar histórias usando a interface: as histórias contadas ao computador são semelhantes às contadas a humanos.

Palavras-Chave

Narrativas, recuperação de documentos, gestão de informação pessoal, interfaces pessoa-máquina

1. INTRODUÇÃO

Duas tarefas que nenhum utilizador de computadores consegue evitar são organizar e posteriormente encontrar os seus documentos electrónicos. Infelizmente, os mecanismos existentes para as realizar não sofreram alterações significativas nas últimas décadas. Baseiam-se, primordialmente, na classificação de *todos* os documentos criados numa hierarquia definida pelo utilizador, no sistema de ficheiros, e na atribuição de um nome ao ficheiro que representa o documento dentro dessa hierarquia. Isto causa inúmeros problemas, uma vez que não é invulgar um documento poder ser classificado em mais do que uma categoria, ou não parecer enquadrar-se em nenhuma das existentes. As decisões a que esse tipo de classificação obriga aumentam desnecessariamente a carga cognitiva a que os utilizadores estão sujeitos e dificulta a recuperação dos documentos num momento posterior.

Thomas Malone efectuou um dos primeiros estudos sobre as formas de organizar documentos [Malone83]. Neste, a

dificuldade dos utilizadores em classificar todos os documentos ficou bem patente. Era preferível para a maioria dos utilizadores organizar os documentos em pilhas e recorrer à sua memória visual e espacial do que ter que os classificar. Hoje em dia, o problema subsiste, agravado pelo aumento significativo do número de documentos com que os utilizadores lidam.

A causa da ineficiência encontrada ao usar sistemas de ficheiros hierárquicos é estes não reflectirem a forma mais natural dos utilizadores se referirem aos seus documentos. Quando um documento é procurado, não nos lembramos do um nome ou classificação arbitrariamente atribuídos, mas sim do documento em si, de porque o necessitamos nesse momento, o que continha, etc. Em suma, recorremos a informação autobiográfica quando nos recordamos de um documento. Isto foi confirmado por estudos das caixas de correio electrónico dos utilizadores, em que o mesmo problema de classificação existe ao guardar mensagens nas pastas hierarquicamente organizadas. Verificou-se, no entanto,

que muitos utilizadores recorriam a essas mensagens como forma de encontrar documentos! [Whittaker96] Isto deve-se ao facto das mensagens, ao contrário dos documentos no sistema de ficheiros, aparecerem inseridas num contexto que permite mais facilmente encontrá-las, associadas a informação adicional (data e hora, remetente, assunto, etc.) que está mais perto da forma natural de recordar dos utilizadores.

Numa tentativa de permitir aos utilizadores fazer uso dessa informação para recuperar os seus documentos, vários trabalhos foram desenvolvidos. Alguns seguiram uma abordagem mais limitada dando um papel preponderante ao tempo. É o caso do Lifestreams [Freeman96], em que todos os documentos são apresentados sequencialmente, de forma ordenada, sendo possível obter *substreams* de documentos que cumprem um certo critério filtrado a *stream* principal. Já o Timescape [Rekimoto99] apresenta o *desktop* do ambiente de trabalho como uma janela sobre um determinado instante de tempo, que pode ser movido quer para o passado quer para o futuro.

Mais abrangentes, algumas abordagens têm em conta um leque mais alargado de propriedades. O primeiro desses trabalhos foi o Semantic File System [Gifford91], em que os documentos estão organizados em directorios virtuais, resultado de pesquisas segundo critérios como o seu autor ou a data em que foi criado. Mais recentemente, sistemas semelhantes mas mais sofisticados foram desenvolvidos. É o caso dos Placeless Documents de Paul Dourish [Dourish00] e do PACO, de Ricardo Baeza-Yates [Baeza-Yates96], em que os documentos são organizados em colecções que os agrupam por algum critério.

Alguns sistemas tentam especificamente fazer uso de propriedades cujos valores não conseguem ser apreendidos directamente dos documentos propriamente ditos. É o caso do Stuff-I've-Seen [Dumais03] que integra informação sobre os documentos recolhida de várias aplicações que o utilizador corre. O sistema de Ariel Shamir [Shamir04] tem um comportamento semelhante. Os recentes sistemas de pesquisa no *desktop*, como o Google Desktop, também partilham desta característica.

Estes sistemas, embora promissores, não resolvem totalmente o problema de recuperar documentos. Recordar valores atribuídos a propriedades arbitrárias não é muito melhor do que recordar onde, no sistema de ficheiros, está um documento guardado. Falta uma interface que permita ao utilizador, de forma natural, referir toda a informação relevante. Propomos as histórias sobre os documentos como base para essa interface.

Os seres humanos são contadores de histórias por natureza. Todos contamos histórias da infância à velhice. Em particular, contamos histórias sobre os documentos que procuramos quando temos alguém ao nosso lado durante essa pesquisa. Uma interface capaz de fazer uso dessas histórias para recuperar os documentos seria, ao mesmo tempo, natural para os utilizadores, e faria uso da

informação realmente associada pelos utilizadores aos seus documentos.

Na tentativa de criar interfaces baseadas em histórias para recuperação de documentos, conseguimos, recorrendo a entrevistas, recolher sessenta histórias. A sua análise permitiu-nos conhecer a forma e conteúdos que podemos esperar. Um conjunto de *guidelines* para a criação das interfaces foi, também, desenvolvido [Gonçalves04]. Com base nessas *guidelines*, e com a ajuda da avaliação de protótipos de baixa fidelidade, foi possível identificar um formato promissor para a interface propriamente dita [Gonçalves04a].

No entanto, duas questões cruciais para a validação das histórias como base de uma interface para recuperação de documentos ainda estavam por responder. Em primeiro lugar, era necessário confirmar que um utilizador consegue contar histórias a um computador da mesma forma que o faz a um interlocutor humano. Em segundo lugar, era preciso saber até que ponto as histórias contadas são fidedignas. Relembrar correcta e precisamente alguma informação pode não ser fácil, principalmente quando tentando descrever documentos manipulados há já algum tempo. Se a informação contida nas histórias estiver muito desfasada da realidade, não poderá ser usada para encontrar documentos.

Para dar resposta a estas questões, implementámos a interface que se mostrou mais adequada nos protótipos de baixa fidelidade. Trinta novas histórias sobre documentos foram recolhidas, tendo sido comparadas com a informação real associada aos documentos descritos. Isto permitiu-nos avaliar até que ponto a informação estava correcta. Adicionalmente, as histórias foram comparadas com as previamente recolhidas com o auxílio de um entrevistador humano, de modo a verificar se as histórias contadas a uma interface no computador eram ou não semelhantes às restantes. Verificámos que é esse o caso. Constatámos, também, que entre 81% e 91% de toda a informação contida nas histórias está correcta.

Na secção seguinte descreveremos, sucintamente, a interface implementada. Em seguida, referiremos a metodologia usada no estudo descrito neste artigo, após o qual os principais resultados serão enumerados. Seguir-se-á uma discussão dos mesmos. Concluiremos então, apresentando possíveis caminhos para o trabalho futuro.

2. O PROTÓTIPO

A interface usada no estudo descrito neste artigo foi concebida com base nos resultados de testes com protótipos de baixa fidelidade [Gonçalves04a]. Esses protótipos foram criados com base nas linhas-mestras inferidas a partir da análise de histórias contadas a ouvintes humanos [Gonçalves04]. Comparando-as com as recolhidas usando os protótipos, foi possível identificar qual deles melhor permitia aos utilizadores contarem as suas histórias. A satisfação dos utilizadores foi também medida com a ajuda de um questionário.

Uma vez escolhida a forma que a interface deveria tomar esta foi implementada usando a linguagem de

programação Python, escolhida tanto pelo seu poder e flexibilidade como pela rapidez de construção de um protótipo que permite. A aplicação resultante foi baptizada com o nome de Quill.

2.1 A Interface

A Figura 1 mostra o aspecto final da interface Quill na altura da realização do estudo. A janela está dividida em três áreas. A que mais se destaca, no canto superior esquerdo, é a área da história propriamente dita. Nela, a história que descreve os documentos procurados vai sendo incrementalmente construída. Cada elemento possível vai sendo sucessivamente apresentado ao utilizador sob a forma de uma frase incompleta. Após a introdução da informação relevante pelo utilizador, essa frase será terminada. A parte da frase que contém a informação específica sobre o documento funcionará como um elo hipertextual permitindo a posterior alteração dessa informação. Os vários elementos possíveis foram inferidos a partir da análise de conteúdos efectuada às histórias contadas a humanos num estudo anterior.

A informação é introduzida com a ajuda de diálogos especializados que aparecem no lado direito da janela. A informação a pedida em cada um desses diálogos reflecte directamente o que os conteúdos encontrados nas histórias contadas a entrevistadores humanos. Os diálogos são flexíveis. Por exemplo, se o utilizador mencionar que um documento é do tipo Microsoft Word, então no

diálogo referente ao Conteúdo do documento (em que o aspecto visual do mesmo pode ser especificado) apenas serão dadas opções referentes a possíveis aspectos de documentos desse tipo. Se, caso contrário, se tratasse de uma folha de cálculo, seriam esses os aspectos mostrados.

A história é construída como um todo coerente, legível pelos utilizadores. Assim, quando nova informação é introduzida e as frases completadas, estas podem sofrer pequenas alterações de modo a reflectir essa informação. Pretende-se manter o texto o mais correcto possível (dentro de limites razoáveis) para que a história como um todo continue a ser apercebida como tal pelos utilizadores. Dito isto, tentámos que as alterações fossem tão pequenas quanto possível evitando desconcertar e confundir os utilizadores.

A ordem na qual os vários elementos da história são sugeridos pelo programa é a inferida a partir das histórias contadas a humanos, que considerámos ser a mais natural. Apesar disso, o utilizador tem ao seu dispor um conjunto de botões que lhe permitem alterar o fluxo da história. Estes estão localizados imediatamente abaixo da área da história. O mais à esquerda, “*I Want Another*” (“Quero Outro”) permite aos utilizadores escolherem um qualquer elemento à sua escolha para referir em seguida. Para tal, apresenta uma lista ao utilizador com os elementos possíveis. O botão seguinte, “*It Didn't Happen*” (“Não Aconteceu”), pode ser pressionado pelos utilizadores

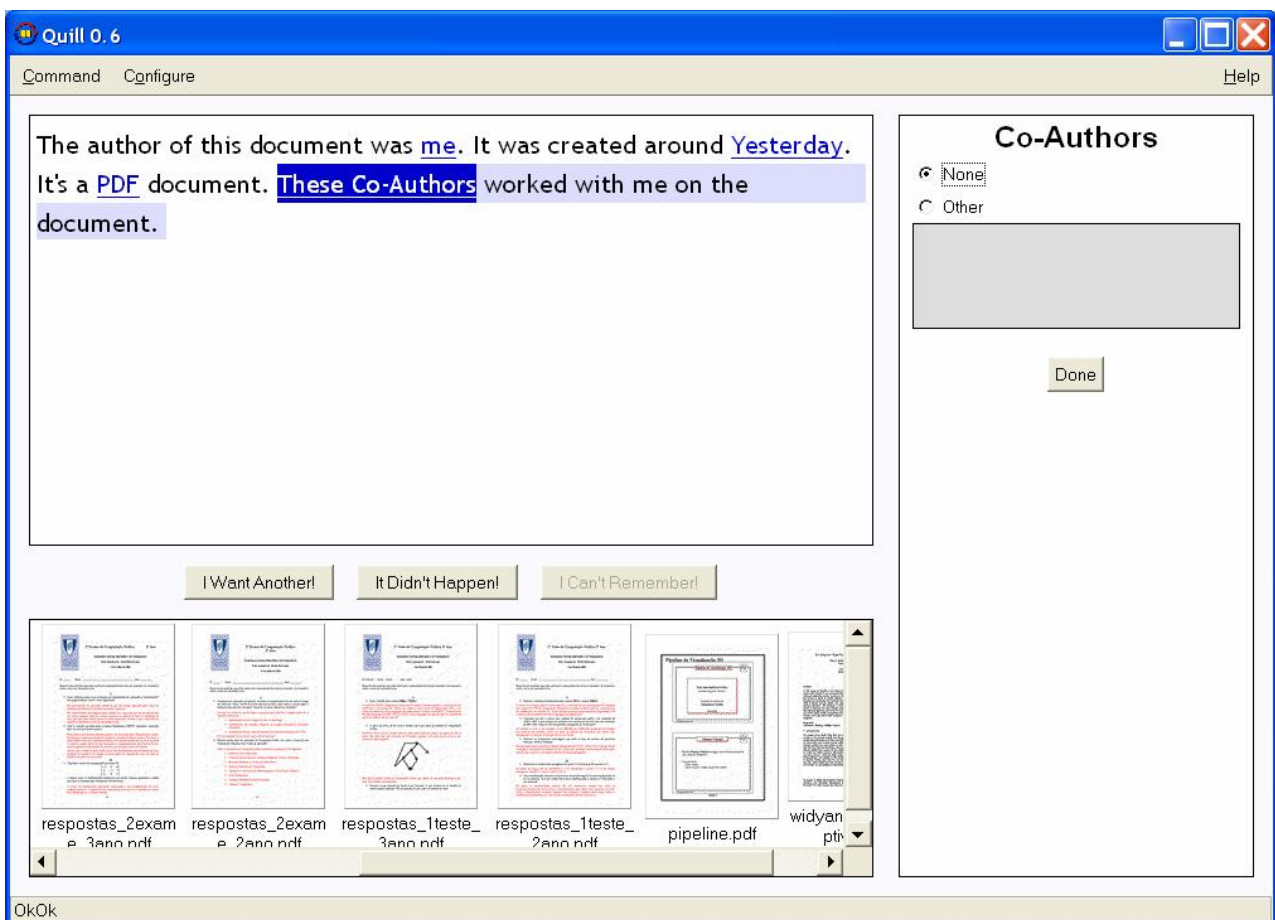


Figura 1 – Aspecto geral da interface

quando o elemento sobre o qual vão falar num dado momento se reporta a algo que não teve lugar. Por exemplo, se questionados quanto ao envio do documento por correio electrónico para outro utilizador, este botão pode ser pressionado caso isso não tenha tido lugar. Finalmente, encontramos o botão “*I Can’t Remember*” (“Não me Lembro”). Este poderá ser usado quando os utilizadores são inquiridos sobre um elemento que não recordam bem. A existência destes dois últimos botões permite ao utilizador estabelecer a diferença entre a informação que não sabe e aquela que sabe não ter acontecido. Isto acontecia recorrentemente nas histórias contadas a humanos pelo que se tentou preservar essa informação na interface.

A terceira área na janela da aplicação é a área de sugestão de documentos. À medida que a história se vai tornando mais completa, o Quill procura continuamente os documentos que mais se adequam à descrição nela contida. O seu nome é mostrado junto a um pequeno *thumbnail* da primeira página do documento (ou o seu aspecto geral, se for uma imagem, por exemplo). Efectivamente, outro dos factores mais vezes referidos nas histórias era o aspecto visual dos documentos, pelo que esperamos, apresentando este *thumbnail*, tomar partido da memória visual dos utilizadores e ajudá-los a reconhecer o documento que procuram de forma natural, sem perturbar indevidamente o processo de contar a história distraíndo-os da mesma.

2.2 A Infraestrutura

Subjacente à interface que acabámos de apresentar está uma infraestrutura que permite ao sistema compreender as histórias contadas pelos utilizadores e encontrar os documentos procurados. Cedo foi verificado que a informação contida nas histórias se reporta a um contexto alargado e não apenas ao documento em si. É informação autobiográfica, sendo necessário para a compreender conhecer o utilizador e o mundo que o rodeia. Recorrer apenas a palavras-chave ou outro tipo de meta-informação preenchida pelos utilizadores não é suficiente, e seria ineficaz dada a relutância dos mesmos em fornecê-la. Decidimos, assim, recolher automaticamente a maior quantidade de informação possível sobre os utilizadores e as suas acções. Para tal, foi criado um sistema que monitoriza continuamente o que se passa no computador e guarda essa informação.

O sistema de monitorização, ainda em desenvolvimento, começa por recolher toda a informação já existente no computador na primeira vez em que é corrido, mantendo-a, a partir daí, actualizada. No momento em que o estudo foi realizado, apenas era recolhida a informação sobre os ficheiros do utilizador. As directorias de sistema são automaticamente filtradas, e um mecanismo de *blacklist/whitelist* permite o refinar dos critérios usados para a realidade de cada utilizador. Toda a informação relevante é extraída dos ficheiros, como o seu nome e data de criação e modificação. Para ficheiros de determinados tipos outra informação disponível é extraída. Por exemplo, dos ficheiros MP3 ou OGG é

extraída a meta-informação contida no seu cabeçalho ID3. Dos ficheiros textuais são extraídas palavras-chave, recorrendo ao algoritmo tfidf [Salton88], após terem sido divididos em *tokens* e sujeitos a um algoritmo de *stemming* para obter apenas as raízes das palavras tornando a aplicação mais robusta em termos de tempos verbais, género, número, etc. O algoritmo usado foi o algoritmo de Porter [Porter80], mas tanto o *stemming* como a divisão em *tokens* está feita de modo a que outros algoritmos possam ser aplicados sem grande esforço (bastando alterar o ficheiro de configuração), para lidar com documentos noutras línguas, por exemplo.

Toda a informação recolhida pelo sistema de monitorização é armazenada numa base de conhecimento. Desde o início que se tornou patente que para compreender as histórias seria necessário ter algum conhecimento de senso comum. Até uma coisa tão simples como referir que um documento foi escrito “por volta do Ano Novo” obriga a conhecimento adicional para ser interpretada: saber que o Ano Novo ocorre a 1 de Janeiro de todos os anos, e que provavelmente como o utilizador não referiu nenhum ano, está a falar do último. Assim, foi decidido usar uma base de conhecimento em que a informação estará devidamente estruturada de modo a permitir as inferências que serão necessárias para a correcta interpretação das histórias. Com a ajuda de ontologias, será possível enriquecer a base de conhecimento com toda a informação necessária para tal. Após um estudo dos formalismos usados para construir bases de conhecimento [Gonçalves04b], decidimos usar o RDF, parte da iniciativa da *Semantic Web*. Para além de ser um standard, espera-se que, com o passar do tempo, informação semântica representada nesse formalismo passe a permear a WWW. O acesso imediato a essa informação pode ser de uma ajuda inestimável para compreender o que está a interessar ou preocupar o utilizador a cada momento. Verificámos que o RDF tem um poder expressivo suficientemente grande para a nossa aplicação e que, caso se venha a mostrar necessário aumentá-lo, basta recorrer a outro nível de complexidade de representação sobre RDF (RDFSchem ou OWL) sem invalidar tudo o que já foi criado anteriormente.

Em termos de implementação, foi criada uma biblioteca em Python, a que chamámos Scroll, que esconde os detalhes de implementação referentes ao RDF e permite a preocupação apenas com questões de mais alto nível. Isto inclui a realização de inferência de nó e de caminho, e também a representação de conhecimento de forma semelhante a uma rede semântica, cujos *case-frames* estão definidos num *schema* designado por iQUILL, com uma expressividade perto da lógica de primeira ordem (exceptuando a quantificação existencial e negação). Outro *schema*, QUILL, define todos conceitos necessários para o armazenamento da informação recolhida sobre os documentos, páginas web, etc.

Sempre que é contada uma nova história, são criadas regras de inferência referentes à informação introduzida. As inferências correspondentes são então efectuadas no

Scroll, e aos documentos resultantes atribuído um valor numérico. Da combinação dos valores resultantes das diversas regras de inferência presentes na base de conhecimento num dado momento resulta uma pontuação global para os documentos. Os que estiverem melhor pontuados serão sugeridos ao utilizador na área de sugestão de documentos da interface.

3. METODOLOGIA

O objectivo do estudo era avaliar o grau de confiança que se pode ter na informação contida nas histórias contadas a um protótipo, e comparar essas histórias com as contadas a humanos de modo a verificar se partilham das mesmas propriedades. Para o realizar, foi necessário aceder aos computadores dos utilizadores. Só assim poderiam as histórias contadas ser comparadas com os documentos propriamente ditos, de forma a verificar quanta da informação nelas contida é fidedigna. Consequentemente, foi necessário entrevistar os utilizadores na sua casa ou no seu local de trabalho. Isto colocou algumas limitações ao número de utilizadores entrevistados. Preocupações de privacidade resultantes da necessidade de aceder a documentos reais foram, também, uma barreira. No total, foram entrevistados dez utilizadores, seis dos quais no seu local de trabalho. Houve um esforço consciente de que a amostra fosse a mais diversificada possível. Foram entrevistados utilizadores com profissões como consultor na área de informática ou advogado, entre outras. As suas idades estão compreendidas entre os 26 e os 56 anos, sendo seis homens e quatro mulheres.

Cada entrevista começou com uma pequena introdução em que foi explicado aos entrevistados o objectivo do estudo e como a entrevista iria decorrer. O protótipo foi então instalado nas suas máquinas, após lhes ser garantido que este estaria contido num único directório bem definido, pelo que seria possível apagá-lo sem deixar rasto uma vez terminada a entrevista. Enquanto o programa recolhia informação sobre os documentos do utilizador, o entrevistador dava um pequeno tutorial sobre o modo de funcionamento da interface, e preenchia um pequeno formulário com informação demográfica sobre o utilizador. Foi garantida a confidencialidade total aos entrevistados. Uma vez terminado este período introdutório, o processo de recolha de informação por parte do protótipo era interrompido, caso não tivesse ainda terminado. Isto deu origem a uma base de conhecimento incompleta, mas em nada comprometeu o estudo, uma vez que a validação das histórias foi feita manualmente. Permitir que o processo de recolha de informação chegasse ao fim seria demasiado moroso, dadas as restrições temporais colocadas para as entrevistas. Tentou-se que estas não ultrapassassem a uma hora de duração. Evitou-se assim interferir no trabalho dos entrevistados mais do que o estritamente necessário.

Aos utilizadores foi então pedido que contassem três histórias sobre três documentos diferentes: um documento Recente, criado por eles até há duas semanas atrás, um documento Antigo, criado entre um a dois anos atrás, e um documento de Outro Autor. Esses três tipos de

documentos tinham já sido considerados em estudos anteriores, e mantê-los permitiu a comparação directa das histórias entre os estudos. Para evitar que existisse alguma deturpação dos resultados devido à crescente familiaridade dos utilizadores com a interface ao longo da entrevista, as histórias sobre os documentos de diferentes tipos foram pedidas em ordens distintas de utilizador para utilizador (Tabela 1).

	1º lugar	2º lugar	3º lugar
Recente	5	3	2
Antigo	3	7	0
Outro	2	0	8

Tabela 1 – Posição das histórias na entrevista

O tempo gasto com cada história foi registado. Após o término de cada uma, foi pedido aos utilizadores que encontrassem manualmente o documento procurado (caso o Quill não tivesse sido capaz de o fazer). A informação com ele relacionada foi então anotada junto da informação referida na história, tendo sido avaliado se esta era correcta ou não. Estes dados foram armazenados para uma posterior análise, mais cuidada.

3.1 Avaliando a Correção da Informação

Confirmar a correcção dos vários elementos que constituem as histórias não é trivial. Nem todos os elementos são passíveis do mesmo grau de verificação. Por exemplo, se o nome do ficheiro onde o documento está guardado pode ser verificado inequivocamente, o mesmo não pode ser dito sobre as trocas desse documento com outros utilizadores. Isso implicaria, em última análise, verificar todas as mensagens de correio electrónico do utilizador, todos os discos amovíveis, etc. e mesmo assim seria impossível ter a certeza absoluta quanto à ocorrência ou não de uma certa troca. Para os elementos cuja verificação mais inequívoca foi impossível, foi estabelecido um diálogo com os utilizadores em que estes tinham que defender a informação dada. Várias perguntas foram feitas pelo entrevistador, com base na informação concreta sobre o documento que foi possível encontrar. Para que o elemento fosse considerado como válido a explicação dada devia ser consistente com essa informação e convencer o entrevistador. Este tentou sê-lo exaustivo quanto possível nas questões colocadas.

Mesmo assim, e para garantir a correcção dos resultados apresentados, foram distinguidos dois níveis diferentes de validade dos elementos. Por um lado, encontramos aqueles que cuja veracidade foi inequivocamente verificada (comparando o nome de ficheiro do documento com o dado na história pelos utilizadores, por exemplo). Por outro, temos os que foram considerados como correctos após questionar os utilizadores. A análise dos resultados será efectuada tendo sempre presente esta distinção.

Para alguns dos elementos mais problemáticos, estratégias específicas de validação foram usadas. As mais importantes estão referidas em seguida. Uma lista mais completa pode ser encontrada no relatório técnico que descreve o estudo [Gonçalves05]:

- **Razão:** nalguns casos era possível verificar a razão de ser do documento a partir do seu conteúdo. Caso contrário, os utilizadores foram questionados.
- **Outros Documentos:** sempre que possível, pedimos para consultar também esses documentos (frequentemente guardados na mesma pasta), confrontando os utilizadores com eles, fazendo-os verificar se se os tinham recordado correctamente. Em casos em que os documentos relacionados eram citados no documento principal, a verificação foi imediata e inequívoca.
- **Vida Pessoal:** se a informação estivesse nas agendas dos utilizadores, ou fosse conhecida pessoalmente pelo entrevistador era considerada como correcta..
- **Acontecimentos no Mundo:** estes acontecimentos foram verificados a partir da memória dos utilizadores e do próprio entrevistador, recorrendo a fontes noticiosas *online* para confirmar que ocorreram na altura referida pelo utilizador.
- **Trocas:** tentámos perceber se a informação estava correcta a partir do conteúdo do documento e aparente fim a que se destinava.
- **Tarefas:** se as tarefas mencionadas se reflectissem no conteúdo do documento (inserir imagens ou preparar gráficos, por exemplo), seriam consideradas como válidas
- **Acontecimentos** (ocorridos durante a interacção com o documento): a única forma de confirmar se esta informação estava correcta ou não foi dialogar com os utilizadores

4. RESULTADOS

Nesta secção iremos descrever os principais resultados do estudo. Começaremos por verificar se as histórias contadas usando a interface são ou não similares às contadas a entrevistadores humanos. Em seguida avaliaremos a veracidade da informação contida nas histórias, tanto em termos globais com de cada um dos elementos que as constituem.

4.1 Contando as Histórias...

Nenhum utilizador necessitou de explicações adicionais quanto à utilização da interface após o tutorial inicial. Com algumas excepções, todos os elementos da interface foram correctamente compreendidos e usados. Em seguida apresentamos uma das histórias contadas por um dos utilizadores, copiada directamente da interface (após alterações para garantir o anonimato do entrevistado):

O autor deste documento fui eu. Foi criado por volta de há um mês. A razão

pela qual o criei foi para um relatório de trabalho. Quando trabalhei no documento estava no local de trabalho. Foi escrito apenas por mim. O documento é sobre Data Warehouse. este documento lembra-me outro chamado "adenda" (areas) porque foi manuseado na mesma altura e foi escrito pela mesma razão e era sobre o mesmo assunto e está guardado na mesma pasta (ou lá perto...). Dei-o a [colega1] e [colega2] usando email e LAN (pastas partilhadas, etc.). É um documento do tipo Microsoft Word. O documento contém as palavras ou expressão "Adenda Relatório Tardes DW Projectos" e tem o aspecto de um texto a uma coluna. O documento está no Computador no Local de Trabalho. Para o escrever, tive que recorrer a outros documentos. Teve várias versões. O nome do ficheiro era algo como "tardes DW". Nessa altura não aconteceu nada de especial. Na altura em que o criei, na minha vida pessoal (amigos, família, trabalho) não aconteceu nada de especial. Quando estava a trabalhar no documento, nada de especial aconteceu.

A interface podia ser configurada para funcionar em Inglês ou Português, de acordo com as preferências dos utilizadores. A historia apresentada pertence a um dos dois utilizadores que o fez em Português. O Português usado não é perfeito, mas é suficientemente bom para que a história seja satisfatoriamente legível.

Foi observada uma curva de aprendizagem da interface, reflectida no tempo levado pelos utilizadores a contar as histórias (Figura 2). Em média, o tempo levado para contar a terceira história era apenas 60% do gasto com a primeira, independentemente dos tipos de documentos descritos (de 427 a 269 segundos).

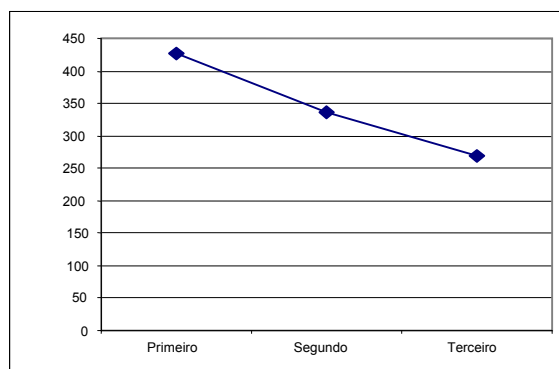


Figura 2 – Tempo gasto a contar as histórias (em seg.)

4.2 Comparando as Histórias

É importante verificar até que ponto os utilizadores continuam a sentir que estão a contar histórias e não apenas a dar valores a propriedades usando uma interface

no computador. Para tal, as histórias contadas usando o Quill foram comparadas às contadas a um entrevistador humano. Se partilhassem da mesma estrutura e conteúdos, estaria efectuada a nossa validação. Mais concretamente, usámos como base de comparação as histórias encontradas durante a fase de avaliação de protótipos de baixa fidelidade, comprovadamente semelhantes às contadas a um entrevistador [Gonçalves04]. Comparámos tanto a estrutura quer o conteúdo das histórias.

4.2.1 Estrutura das Histórias

As histórias agora recolhidas têm comprimentos semelhantes aos das anteriormente analisadas: cerca de 14 elementos (Tabela 2). A razão entre os comprimentos das histórias actuais e das anteriores são de 98%, 101,5% e 100,7%, para os documentos Recentes, Antigos e Outros, respectivamente. Testes de t-student confirmaram que esses valores são iguais, com 95% de confiança.

	Actual		Anterior	
	Média	Desv. Pad	Média	Desv. Pad
Recente	14	1.05	14.3	2.06
Antigo	13.5	1.08	13.3	1.25
Outro	13.4	1.43	13.3	2.06

Tabela 2 – Comprimento das Histórias (em elementos)

Avaliámos a ordem na qual os diversos elementos ocorrem nas histórias tendo em conta que a ordem em que eram sugeridos aos utilizadores era idêntica à inferida a partir das histórias contadas a humanos. Assim, qualquer desvio dessa ordem indicaria que esta não era considerada natural neste contexto. Foi contabilizado o número de vezes que o utilizador pediu um elemento diferente do sugerido, usando o botão para esse fim disponível na interface.

Apenas um utilizador pediu um elemento diferente do sugerido, uma vez em cada uma das suas histórias, resultando numa média global de 0.1 utilizações por história. Podemos, assim, concluir que a ordem na qual os elementos eram sugeridos aos utilizadores era considerada natural por estes.

Em suma, a estrutura das histórias contadas usando o protótipo é semelhante à das histórias contadas a humanos.

4.2.2 Conteúdo das Histórias

Uma vez que as histórias possuem estruturas semelhantes, é espectável que os seus conteúdos também o sejam. Uma verificação mais cuidada deste facto será descrita em seguida.

Um dos aspectos a ter em conta é a importância dos vários elementos nas histórias. Para verificar se esta se mantinha inalterada, verificámos as frequências com que esses elementos aparecem nas histórias recolhidas em ambos os estudos. O gráfico na Figura 3 mostra os valores dessas frequências.

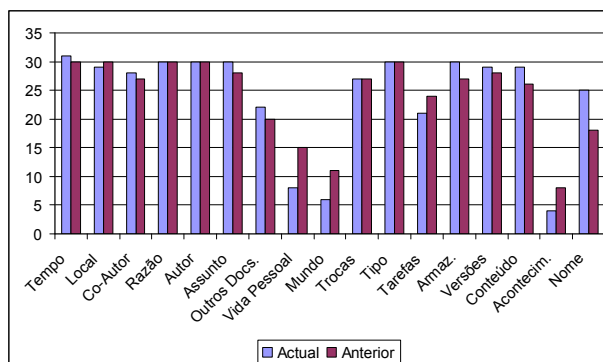


Figura 3 – Comparação das requências dos elementos

É imediatamente aparente que, com poucas excepções, as frequências são bastante semelhantes. Diferenças relevantes podem ser observadas apenas para quatro elementos. O Nome dos documentos é mencionado 39% mais frequentemente nas histórias contadas usando o protótipo. Isto pode ser explicado por duas causas diferentes. Em primeiro lugar, os utilizadores estavam, na sua maioria, a trabalhar com o computador imediatamente antes da entrevista, o que os pode ter tornado mais atentos às convenções de nomes que costumam usar, refrescando a sua memória. Em segundo lugar, o simples facto de estarem a lidar um com computador pode ter influenciado a sua propensão para mencionar informação relacionada com este, como é o caso dos nomes dos ficheiros. Seja como for, a diferença encontrada é positiva, uma vez que mais informação é sempre útil para encontrar um documento.

Já os elementos sobre a Vida Pessoal, Acontecimentos no Mundo e Acontecimentos ocorrem menos frequentemente (47, 45 e 50 por cento, respectivamente). Estes são os três elementos que têm mostrado ser menos confiáveis e difíceis de recordar ao longo de todos os estudos efectuados. A maioria dos utilizadores pura e simplesmente não os associa aos documentos que manipula. Foi encontrada uma grande variabilidade individual no que diz respeito a estes elementos. Alguns utilizadores nunca os usam, enquanto que outros recorrem a eles esporadicamente. Dada a baixa frequência absoluta com que ocorrem, não é possível tirar conclusões fiáveis a partir das diferenças observadas.

Passando à importância relativa dos vários elementos, independentemente do valor absoluto de ocorrências encontrado, comparar simplesmente a ordem na qual os elementos ocorrem não seria correcto, uma vez que pequenas variações em valor absoluto podem resultar em alterações nessa ordem. Em vez disso, os elementos foram divididos em duas classes: os que são mencionados em quase todas as histórias e os que apenas o são raramente. O valor limite que os separa é terem sido mencionados em 70% das histórias. Em redor a esse valor não há elementos, tendo todos frequências claramente maiores ou menores. Esta divisão foi feita para as distribuições de elementos nos dois conjuntos de histórias. Foi, assim, possível encontrar os elementos cuja importância relativa se alterou (i.e., mudaram de classe).

Para os documentos Recentes, não encontramos quaisquer diferenças. Para os documentos Antigos e Outros, e no geral, apenas o elemento Nome mudou de importância. Como já foi anteriormente mencionado, este elemento foi referido mais frequentemente nas histórias contadas usando a interface. Passou assim de elemento raro a comum.

Para concluir, as histórias contadas ao computador usando o protótipo mantiveram-se, no geral, idênticas às contadas a interlocutores humanos.

4.3 Validade das Histórias

Iremos agora estudar a validade das histórias, verificando até que ponto podemos confiar na informação nelas contida. Atentaremos, em primeiro lugar, à correção geral das histórias, e focaremos em seguida a de cada um dos seus elementos.

4.3.1 Validade Geral

O gráfico na Figura 4 mostra as percentagens de elementos válidos encontrados nas histórias de cada um dos diferentes tipos. Como já foi mencionado acima, considerámos duas classes diferentes de elementos: os cuja validade foi verificada inequivocamente e os que fora considerados correctos sem que tenha sido possível prová-lo cabalmente.

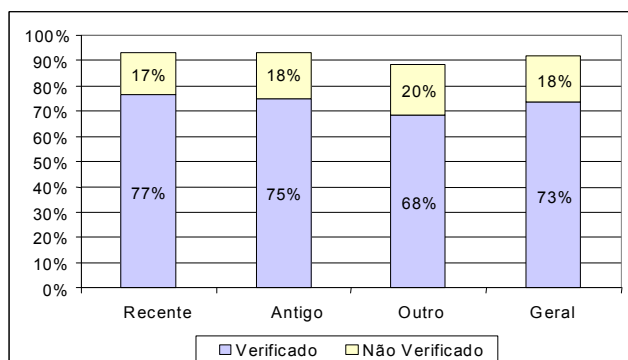


Figura 4 – Veracidade das Histórias

Não parecem existir diferenças relevantes entre as histórias para os documentos Recentes e Antigos, revelando que os utilizadores se conseguem lembrar igualmente bem da informação para documentos desses tipos. As histórias sobre documentos de Outros autores, por outro lado, parecem ligeiramente menos precisas do que as dos documentos do próprio utilizador. Testes de t-student confirmam isto, com um nível de confiança de 95%.

Já no respeitante à validade das histórias no caso geral, verificámos que, embora não perfeita, é razoavelmente boa. Mais de 91% dos seus elementos constituintes são verdadeiros. Mesmo considerando apenas os inequivocamente verificados, o valor é de cerca de 70%. Em média, entre 73% e 92% do que os utilizadores mencionam nas histórias corresponde à verdade.

Olhando mais atentamente para os dados, podemos verificar que a percentagem de elementos não verificados (17, 18 e 20 por cento para os três tipos de documentos, respectivamente) é devida, principalmente, a três

elementos: Vida Pessoal, Acontecimentos no Mundo e Acontecimentos. Esses três elementos explicam 59, 41 e 47 por cento de todos os elementos não verificados para os três tipos de documentos. Isto deve-se ao método de verificação usado: dado que a maior parte dos utilizadores se limitou a dizer que “nada aconteceu” ou que “não se lembram” desses elementos, não tínhamos forma de o verificar. O elevado número de vezes em que isto ocorreu reflecte o facto desses elementos não estarem a ser recordados em relação aos documentos. O utilizador não os mencionar é tão válido como se realmente nada tiver acontecido. Adicionalmente, mesmo que algo tenha de facto ocorrido não estamos na presença de informação incorrecta tanto como estamos na ausência de informação. Posto isto, e recalculando as percentagens de excluindo esses elementos, obtemos os resultados mostrados no gráfico da Figura 5. Podemos aí constatar que entre 81% e 91% das histórias pode ser considerada correcta. Isto corresponde apenas a de 1 a 3 elementos errados por história, mais do que suficiente para as usar como base para a recuperação de documentos.

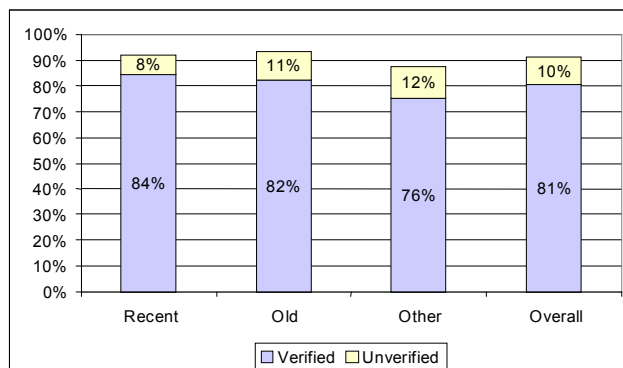


Figura 5 – Veracidade das Histórias (corrigida)

4.3.2 Correção dos Elementos

É importante saber quais os elementos mais frequentemente errados, para melhor lidar com as suas eventuais imprecisões. O gráfico na Figura 6 representa a veracidade dos vários elementos separadamente, para todos os tipos de documentos.

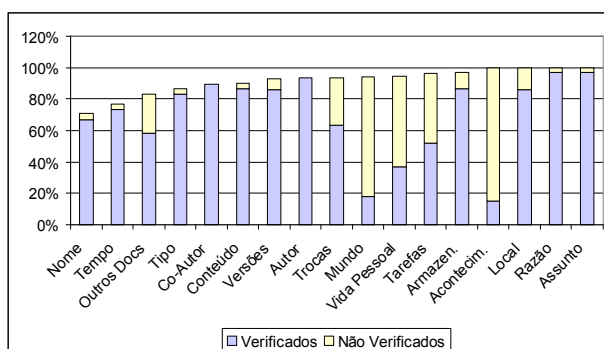


Figura 6 – Veracidade dos Elementos

Alguns elementos foram notoriamente difíceis de verificar. É o caso dos três menos precisos, Vida Pessoal, Acontecimentos no Mundo e Acontecimentos, já discutidos na secção anterior. Também nesta situação

estão as Trocas e Tarefas. Na maior parte dos casos estávamos confiantes na sua correcção, mas foi impossível confirmá-lo.

O Nome parece ser o elemento mais infrequentemente lembrado com precisão. Nalguns casos o utilizador estava errado, mas na maior deles os utilizadores tinham alguma ideia do nome correcto (“janeiro2005” em vez de “2005janeiro”, Utilizador 7), o nome real tinha sido abreviado, ou parte do nome sugerido pelos utilizadores pertencia ao nome real (“estudo produtividade bolseiros” em vez de “produtividade bolseiros – calculos”, Utilizador 4).

Quanto ao Tempo, na maior parte dos casos os elementos errados estavam perto da data correcta, mas mesmo assim fora dos intervalos de tolerância considerados. Talvez estes necessitem de ser ajustados à luz destes resultados.

O terceiro elemento menos preciso são os Outros Documentos. Frequentemente, os utilizadores lembravam-se correctamente de alguma informação, mas outra estava errada. Por exemplo, o Utilizador 9 mencionou correctamente que o documento versava sobre mesmo assunto que o procurado, mas enganou-se no nome referido. O Utilizador 4 referiu os autores correctos mas enganou-se, também, no nome. De um modo geral, os mesmos problemas encontrados na descrição do documento principal ocorreram também na descrição dos documentos relacionados.

Confusões quanto ao Tipo deveram-se, principalmente, a problemas a distinguir entre tipos semelhantes: texto simples vs Microsoft Word, Powerpoint em vez de imagem (Utilizador 1), ou PDF em vez de Word (Utilizador 7). Os documentos colocados *online* também trouxeram problemas. Por exemplo, o Utilizador 6 mostrou-se confundido sobre se uma base de dados em Microsoft Access que tinha disponibilizado no seu *web site* deveria ser considerada uma base de dados ou um documento web.

Todos os elementos restantes tiveram precisões acima dos 90%. Não foi identificado nenhum outro padrão de erro como os que acabámos de descrever.

Finalmente, no respeitante a eventuais diferenças de veracidade em função do tipo de documento, o único aspecto a referir é que o Autor é recordado correctamente muito menos frequentemente para os documentos de Outros Autores, do que para os do utilizador (80% a 100%), resultado pouco surpreendente.

5. DISCUSSÃO

No que diz respeito à comparação entre as histórias contadas a humanos e as contadas usando a interface, *não foram encontradas diferenças significativas*. A estrutura das histórias mantém-se idêntica, tal como o fazem os seus conteúdos. A única excepção é o elemento Nome, que aparece mais frequentemente usando o computador. Isto confirma que é possível manter a sensação de “contar uma história” usando o computador, sem o qual a construção de interfaces baseadas em narrativas para a localização de documentos seria impossível.

Quanto à veracidade das histórias, verificámos que, *de um modo geral, podemos acreditar nos utilizadores*. Foram envidados esforços para verificar para além de qualquer dúvida razoável a validade dos vários elementos que constituem as histórias. Conseguimos fazê-lo para 81% dos elementos e constatámos que entre 73% e 91% da informação nas histórias está correcta. Se alguns ajustamentos forem feitos tendo em conta a natureza de alguns dos elementos, este intervalo sobe para de entre 81% e 91% (1 a 3 elementos por história). Isto implica que a informação nas histórias não pode determinar taxativamente quais os documentos sugeridos aos utilizadores. Em vez disso, um mecanismo que tenha em conta de forma limitada a contribuição dos vários elementos é necessário, minorando assim os efeitos adversos de alguma informação errónea. Vimos ainda que nalguns casos é possível saber que espécie de imprecisões esperar e agir em conformidade (melhores margens de tolerância para o Tempo, por exemplo).

Os elementos menos precisos foram identificados: Nome, Tempo, Outros Documentos e Tipo. Igualmente, verificámos que nos documentos de outros autores, o elemento Autor é recordado com bastante menor precisão do que no caso dos restantes.

Outro resultado importante é que as histórias sobre os documentos do utilizador, quer Recentes quer Antigos, partilham as mesmas propriedades. Alguns resultados de estudos anteriores apontavam já nesta direcção. Este é, no entanto, um resultado surpreendente, uma vez que se poderia esperar que a informação sobre os documentos mais antigos fosse recordada com muito menor precisão. No entanto, tal não foi observado. Em termos do desenho de interfaces baseadas em narrativas, isto indica que estas podem evitar fazer a distinção entre estes dois tipos de documentos, tratando de modo distinto apenas os documentos do utilizador e os de outros autores.

Sobre a interface usada, podemos afirmar que esta é fácil de aprender e usar. Um pequeno tutorial informal foi suficiente para a sua utilização, e após apenas três histórias terem sido contadas, o tempo gasto para o fazer decresce para 60% do tempo usado com a primeira interacção. Isto mostra que as escolhas feitas com base na avaliação de protótipos de baixa fidelidade estavam correctas.

6. CONCLUSÕES

Encontrar documentos nos sistemas disponíveis hoje em dia é uma tarefa morosa e difícil. Isto deve-se não só aos números crescentes de documentos com que temos que lidar, como à diversificação dos tipos destes. As formas tradicionais de organização e localização não avançaram de modo a fazer frente a esta nova realidade.

As narrativas sobre documentos parecem uma boa base para a criação de interfaces que permitam aos utilizadores fornecer ao computador toda a informação autobiográfica de que se recordam quando procuram um documento, de forma natural e eficiente. Para validar a sua aplicabilidade, era necessário responder a duas questões:

podem histórias semelhantes às contadas a outros humanos ser contadas ao computador? É a informação contida nas histórias suficiente precisa para ser usada como base na localização de documentos?

Os resultados do estudo descrito neste artigo permitem-nos responder afirmativamente a ambas questões. As histórias contadas ao computador são semelhantes às contadas a humanos, e a informação é, de um modo geral (cerca de 90% dos casos), correcta. A adequação das narrativas para a localização de documentos foi, pois, confirmada.

No futuro, e tendo em conta as lições aprendidas neste estudo, o protótipo será melhorado de modo a mais se adequar às reais necessidades dos utilizadores. Iremos integrar o sistema de monitorização com a interface, recolhendo mais informação que nos permitirá melhor compreender as histórias contadas. Actualmente, recolhe já informação referente ao correio electrónico do utilizador, as páginas web a que este acede e as aplicações corridas.

Assim que a interface tiver evoluído o suficiente, efectuaremos testes mais alargados com os utilizadores. Nesses testes, ao longo de maiores períodos de tempo, conseguiremos ter uma melhor ideia da curva de aprendizagem da interface, e responder a uma terceira e premente questão: contém as histórias informação suficiente para distinguir entre documentos semelhantes mas diferentes?

7. AGRADECIMENTOS

Os autores agradecem a todos os utilizadores que participaram no estudo descrito neste artigo. O trabalho aqui apresentado foi financiado pelo projecto FCT POSI/EIA/59022/2004.

8. BIBLIOGRAFIA

- [Baeza-Yates96] Baeza-Yates, R., Jones, T. and Rawlins, G. A New Data Model: Persistent Attribute-Centric Objects, Technical Report, University of Chile, 1996
- [Dourish00] Dourish, P. *et al.* Extending Document Management Systems with User-Specific Active Properties. *ACM Transactions on Information Syst.*,18(2), pp 140-170,ACM Press 2000.
- [Dumais03] Dumais, S. T. et al. Stuff I've Seen: A system for personal information retrieval and re-use. In *Proceedings of SIGIR 2003*.
- [Freeman96] Freeman, E. and Gelernter, D. Lifestreams: A Storage Model for Personal Data, *ACM SIGMOD Record*,25(1), pp 80-86, ACM Press 1996
- [Gifford91] Gifford, D., Jouvelot, P., Sheldon, M. and O'Toole, J. Semantic File Systems. *13th ACM Symposium on Principles of Programming Languages*, October 1991.
- [Gonçalves04] Gonçalves, D. and Jorge, J., "Tell Me a Story": Issues on the Design of Document Retrieval Systems. In *Proceedings DSV-IS'04*, Lecture Notes on Computer Science, Springer-Verlag, July 2004, Hamburg, Germany.
- [Gonçalves04a] Gonçalves, D. and Jorge, J. Telling Stories to Computers. In *Proceedings CHI2004*, ACM Press, 27-29 April 2004, Vienna, Austria.
- [Gonçalves04b] Gonçalves, D. and Jorge, J. Why RDF? Considerations on a Language for the Representation of Document-Describing Stories. *Technical Report, IST/UTL*. June 2004, (disponível online em http://immi.inesc.pt/~djvg/phd/files/why_rdf.pdf)
- [Gonçalves05] Gonçalves, D. Real Stories about Real Documents: Evaluating the Trustworthiness of Document-Describing Stories. *Technical Report, IST/UTL*. March 2005, (disponível online em http://immi.inesc.pt/~djvg/phd/files/real_stories.zip)
- [Malone83] Malone, T. How do People Organize their Desks? Implications for the Design of Office Information Systems, *ACM Transactions on Office Information Systems*, 1(1), pp 99-112, ACM Press 1983.
- [Porter80] Porter, M. F. An algorithm for suffix stripping. *Program* 14, 130-137, 1980.
- [Rekimoto99] Rekimoto, J. Time-machine computing: a time-centric approach for the information environment. In *Proceedings of the 12th annual ACM symposium on User interface software and technology*, pages 45-54, ACM Press, 1999.
- [Salton88] Salton, G. Automatic text processing, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 1988.
- [Shamir04] Shamir, A.. A View on Views. In *Proceedings SmartGraphics04*, Banff Center, Canada, May 2004.
- [Whittaker96] Whittaker, S., Sidner, C. Email overload exploring personal information management of email. In *Conference proceedings on Human factors in computing systems*, pages 276-283, ACM Press, 1996.