

PÁGINAS

ARQUIVOS & BIBLIOTECAS

a&b

Informação

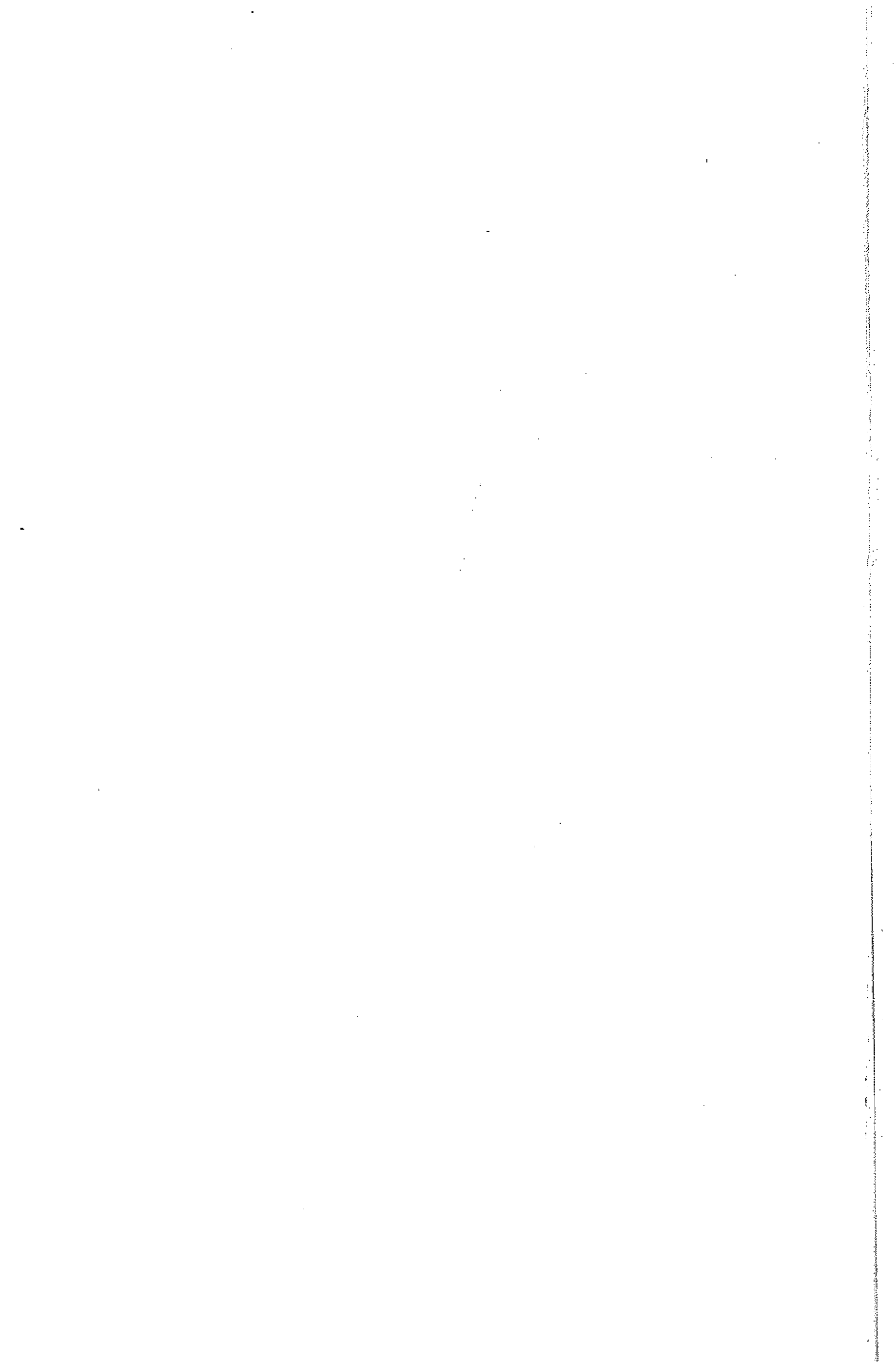
Electrónica

6

20
01

GABINETE DE ESTUDOS

a&b



Gabinete de Estudos A & B

Rua Almirante Barroso, 54 - 5.º

1000-013 LISBOA

PÁGINAS
ARQUIVOS & BIBLIOTECAS

a&b

Director Maria Luísa Cabral **Redacção** Maria Luísa Cabral • Fernanda Ribeiro • Armando Malheiro • Carlos Abreu **Direcção Gráfica** Carlos Abreu **Edição e Propriedade** Gabinete de Estudos a&b **Periodicidade** Semestral **ISSN** 0873-5670 **Depósito Legal** 118 811/97 **Tiragem** 750 exemplares **Execução Gráfica** Textype - Artes Gráficas, Lda. **Distribuidora** HT-Distribuidores - Rua Rodrigues Sampaio, 77 - 1150 Lisboa - **Tel.:** 21 352 90 06/8 • **Fax:** 21 315 92 59

Os artigos assinados são da exclusiva responsabilidade do(s) seu(s) autor(es)

Normas para publicação de artigos

PÁGINAS a&b aceitam artigos para publicação escritos em Português, Inglês, Francês ou Espanhol, de preferência originais, e desde que os mesmos não tenham sido submetidos simultaneamente para apreciação junto de outro Editor.

Qualquer que seja a decisão do Concelho Redactorial de *PÁGINAS a&b* sobre o artigo submetido a apreciação, o(s) autor(es) será(ão) devidamente informado(s).

Os direitos de reprodução dos artigos publicados passarão a ser propriedade de *PÁGINAS a&b*.

Os textos submetidos a apreciação deverão ser acompanhados por disquete e, se o texto contiver gráficos, tabelas, quadros, etc., estes deverão ser fornecidos à parte.

Os artigos deverão vir acompanhados de sumário, em Inglês qualquer que seja a língua original do texto e a sua extensão não deverá ultrapassar as 150 palavras.

As referências bibliográficas deverão seguir a NP 405-1 e NP 405-2 e, no caso de não estarem conforme esta, *PÁGINAS a&b* reservam-se o direito de introduzir as alterações necessárias.

A revisão de provas é da responsabilidade de *PÁGINAS a&b* e só em caso de solicitação específica aquelas serão entregues ao(s) autor(es).

A colaboração com *PÁGINAS a&b* é sempre a título gratuito mas o(s) autor(es) receberá(ão) 4 exemplares do número da revista na qual o artigo for publicado.

Sumário

Editorial

5

GINA GUEDES RAFAEL
Documentos electrónicos
da biblioteca de papel à biblioteca digital

7

ANABELA PRISTA SARAIVA SERRANO
A biblioteca digital ou o acesso global

21

OLÍVIA PESTANA
Elementos para uma avaliação
de fontes de informação na Internet

41

LICÍNIA SANTOS
Metadata
algumas codificações

51

ABBY SMITH
Authenticity in a Digital Environment

59

CHARLES T. CULLEN
Authentication of Digital Objects
Lessons from a Historian's Research
63

PETER B. HIRTLE
Archival Authenticity in a Digital Age
73

GEORGE STEINER
Um mestre de leitura
91

Editorial

Foi complicado arranjar um título suficientemente lato ao mesmo tempo específico que pudesse sintetizar o conteúdo deste número das *PÁGINAS*. Os artigos das Colegas debruçam-se sobre questões diferentes que enchem as bibliotecas em fase de transição, questões que nos obrigam a pensar e sobretudo que nos levam a perguntar que biblioteca é esta onde trabalho. Para onde quero ir? Onde quero chegar? Que tipo de aprovisionamento informativo vou eu providenciar para responder aos meus leitores ou, ainda mais difícil, que expectativas devo alimentar? Não são tempos fáceis mas o desafio que se nos coloca, a nós profissionais, é muito interessante e pena é que nem sempre as instituições estejam à altura de compreender de forma clara a sua missão, proporcionando aos profissionais os meios e as estruturas que lhes permitam quer a progressão devida nas carreiras quer o fornecimento de serviços à medida das necessidades dos leitores.

Acabou, assim, por nos parecer que o lema INFORMAÇÃO ELECTRÓNICA representava da melhor maneira possível o ambiente que neste número perpassa. *Da biblioteca de papel à biblioteca digital*, ao discutir documentos electrónicos, sua origem, situação e tratamento, abre o número rasgando um espaço adequado às outras contribuições, mais ou menos virtuais, que se vão completando entre si. A análise sobre *Metadata* constitui mais uma achega para a compreensão da nova informação digital. Aprender a navegar é preciso e o artigo sobre a *Internet* é disso testemunho: não basta dispor, é indispensável tirar proveito duma ferramenta muito rica e interveniente. Tudo para culminar no *Acesso global*, problema dos problemas e nosso objectivo último. Independentemente do ambiente, acesso, a questão fulcral. Depois artigos de origem norte-americana a revelarem um enorme amadurecimento e um grande à vontade no *Digital*

Environment, artigos que envolvem duma forma magistral a problemática e aparente contradição entre preservação e difusão. Finalmente em *Ler, Muito Prazer* uma entrevista magnífica a George Steiner, filósofo e pensador, sobre a leitura, o seu papel os seus prazeres.

Com o número anterior começámos um período novo da vida das *PÁGINAS*. Em cada número procuraremos acertar aspectos que vão desde a forma ao melhoramento do conteúdo sempre com a preocupação de lançarmos à discussão as chamadas «questões quentes» que nos rodeiam. Ao editarmos esse número, procuramos também fazer alguma publicidade sobre as *PÁGINAS* e sobre o GABINETE DE ESTUDOS a & b e apraz-nos registar, hoje, que o apelo feito tem vindo a ser correspondido. As assinaturas lá vão aparecendo, as colaborações também. Mas precisamos de mais apoios. Prosseguir sem depender de ninguém mas das nossas próprias preocupações é um luxo a que só poderemos atender se os Colegas nos continuarem a procurar. Daqui, pois, um apelo bem simples. Como assinante que é, tente conseguir outro assinante; se já é leitor, porque não fazer-se assinante?

Como sempre, e para além deste desafio, deixamos mais um: as *PÁGINAS* querem espelhar a realidade portuguesa. Faça-nos sugestões de temas que lhe pareçam pertinentes, ou colabore directamente com o seu próprio artigo. Se quiser discutir as suas ideias antes de as pôr sobre uma folha de papel, contacte-nos e, juntos, tentaremos construir cada número. A estrutura do GABINETE DE ESTUDOS a & b está aberta às suas ideias e, quem sabe, não as porá em marcha?

Como sempre, os meus votos de boa leitura.

MARIA LUÍSA CABRAL

Documentos electrónicos da biblioteca de papel à biblioteca digital

GINA GUEDES RAFAEL

RESUMO A biblioteca e o bibliotecário estão a sofrer mudanças muito significativas, devido não só às novas tecnologias, como ao novo tipo de documento. Estas mudanças conduzem a alterações na cultura profissional, levando-nos não só a repensar nas funções do bibliotecário, como obrigam as bibliotecas a enfrentar um público diferente do seu público habitual. Definir documentos electrónicos, e desenvolvimento de colecções, propor novas formas de catalogar e armazenar, considerando as novas descontinuidades tecnológicas assim como as consequentes mudanças de atitudes, são sugestões deste artigo.

ABSTRACT Libraries and librarians are going through very deep changes due either to new technologies or to different sort of documents. These changes cause alterations within the professional culture making us to rethink the librarians' functions as at the same time they push libraries to face a public different from the traditional one. This article aims therefore to define electronic documents, and collections development, to propose new ways of cataloguing and archiving bearing in mind the new technological capabilities as well as the changes in attitude.

Introdução

Na era das redes e da digitalização, deparamo-nos com o pensamento em movimento permanente, numa difusão constante, num fluxo contínuo de informação entre os leitores.

O pensamento materializado já não assenta unicamente numa unidade fixa, como era o livro, o jornal ou a própria revista, mas cada vez mais nas edições electrónicas, sejam edições com suporte físico, disquetes, CD-ROM's, DVD-ROM's, sejam as publicações editadas na Internet. Somos, obrigados também a reflectir para determinar,

quais de entre as publicações na Internet, devem ser tratadas do ponto de vista bibliográfico, uma vez que relativamente às bibliotecas, esses documentos são objecto de depósito legal e/ou compra. Depois, deixámos de ser um simples *site* ou portal, para nos tornarmos também uma plataforma de informação na rede... e assim todos são uma realidade que urge tratar e definir.

Estas novas tecnologias e este novo tipo de documento estão também a conduzir-nos a alterações na cultura profissional, levando-nos não só a repensar nas funções do bibliotecário, como obrigam as bibliotecas a virarem-se para um público diferente do seu público habitual.

Afinal, documentos electrónicos, o que são?

Com o aparecimento dos recursos electrónicos de representação e de transmissão de informação, a história da comunicação acelerou-se ao longo dos últimos 40 anos. Números, imagens, textos, som, tudo pode ser representado em formato electrónico [1].

Por outro lado, a publicação em suporte electrónico pode ser concebida, como o é, para suporte tradicional em papel, quer dizer oferecendo a dupla característica de formato lógico e de aspecto estético e físico. Pois, existe muitas vezes o desejo de projecção no ecrã da estética do papel e que vai até à semelhança, entre o que o ecrã nos mostra e o resultado de uma impressão a pedido, efectuada a partir do suporte electrónico.

Agora, para o presente propósito e para evitar ambiguidades, convém definir publicação electrónica, não especificando ao pormenor as definições de suportes, por demais conhecidos de todos nós, como sejam CD-ROM, bases de dados, etc.

Uma definição proposta para publicação electrónica é apresentada pelo British Library Working Party on Electronic Publishing [2]:

Publication which requires user to employ an electronic device at some stage for its reception and/or its reading. Electronic publications may most conveniently be classified in terms of two characteristics these are whether the content is in bit-mapped or character-coded form, and whether it is delivered to the user on a material medium or as a transmitted signal.

Ou, também segundo a *ISBD(ER)* [3], podem-se considerar documentos electrónicos como materiais controlados por computador, ou que requerem o uso de um periférico (ex. leitor de CD-ROM associados a um computador e que podem ser utilizados de modo interactivo ou não). Incluem-se dois tipos de documentos:

- Dados – informação em forma de número, letras, grafismo, imagens, som e possíveis combinações;
- Programas – instruções ou rotinas para o desempenho de certas tarefas incluindo o processamento de dados.

Esta definição aplica-se à maior parte dos documentos incluindo trabalhos multimédia interactivos, que estão geralmente disponíveis e incluem também aqueles que são acessíveis por rede ou por telecomunicações.

Ainda, segundo «Standards for electronic publishing: an overview»[4] podem-se categorizar as publicações e até sistematizá-las:

- publicações electrónicas sem conexão à rede, são todas as que estão armazenadas em suportes físicos digitais, tais como disquetes, CD-ROM ou discos ópticos;
- publicações electrónicas híbridas, significam todas as publicações sem conexão à rede que contêm ligações a documentos em linha;
- publicações electrónicas em linha, significam aqueles documentos publicados na Internet ou em redes proprietárias.

Depois, ainda se podem caracterizar as publicações em linha, através da fixação do seu texto:

- Documentos estáticos, que são aqueles cuja forma e/ou conteúdo é reconhecido como fixado substancialmente, através do seu ciclo de vida;
- Documentos cumulativos, todos aqueles conteúdos a que se vão adicionar novos elementos, permanecendo contudo texto fixo;
- Documentos dinâmicos, aqueles cuja forma e/ou conteúdos mudam continua ou dinamicamente o seu ciclo de vida.

Em termos globais, pode-se definir publicação ou edição electrónica, como sendo uma publicação que utiliza meios electrónicos para colocar a informação, à disposição do público.

Em suma, estas publicações estão disponibilizadas, quer sobre a forma de disquete, CD-ROM ou outro suporte autónomo, quer por ligação a um servidor central ou directamente a uma rede informática. Podem-se apresentar unicamente sob forma electrónica, ou em suporte electrónico e também em suporte papel. As publicações electrónicas são memorizadas nos computadores, que por sua vez permitem quer a visualização quer a impressão, assim, podem existir várias publicações electrónicas:

- publicações electrónicas equivalente a publicações impressas: livros, revistas, brochuras, etc.;
- bases de dados interactivas contendo por exemplo: bibliografias, estatísticas, imagens digitalizadas ou textuais;
- produções multimédia interactivas;
- programas e sistemas especiais;
- novas formas de publicação de informação: serviços telemáticos, grupos de discussão, etc.

A publicação electrónica pode ter também um carácter retrospectivo, logo que se converta para formato electrónico o fundo documental existente, não só para o tornar mais acessível como para preservar o seu conteúdo ou utilizá-lo na produção de novos documentos.

Tipologia dos documentos

Publicações sem conexão à rede

Na categoria das publicações sem conexão à rede ou *offline* [4], consideramos o CD-ROM, a disquete. Podemos ainda constatar que grande parte do mercado está inundado de bases de dados com material bibliográfico ou de referência, sendo que nos anos mais recentes o desenvolvimento se centrou em bases de texto completo.

Os CD-ROM's contêm uma enorme variedade de informação e é um tipo de suporte que apresenta uma tecnologia relativamente barata, armazena uma enorme quantidade de informação e por outro lado, os CD-ROM's não são tão caros se comparados com o equivalente em suporte papel e são bem aceites pelos utilizadores. Contudo podem apresentar, alguns problemas aos bibliotecários, a nível de: acesso, preços e licenças, actualizações, catalogação, etc.

- A nível do acesso, coloca-se-nos a dúvida entre providenciar o acesso livre contra o armazenamento do disco e a prevenção da perda/roubo do disco;
- Quanto a preços e licenças, a política de licenças está em desenvolvimento e constatamos que alguns produtores proíbem a cópia ou transferência da informação armazenada, enquanto outros cobram taxas/encargos de maneiras diversas; relativamente às actualizações, apesar da grande capacidade de armazenamento do disco, pesquisar numa base de dados em CD-ROM pode ser complicado e um processo lento, caso a base de dados seja actua-

lizada mensal ou quinzenalmente, ou um conjunto de disco seja substituído;

- que se refere à catalogação, deve ou não, por exemplo o bibliotecário catalogar um item de cada disco ou só catalogar o próprio disco, no caso em que existam centenas títulos? Ou como definir em termos de categoria de documento, uma base de dados bibliográfica cumulativa? Obra em vários volumes por simbologia com o suporte papel? Publicação em série?

Obviamente, também podemos apresentar uma porção de facilidades: relativamente ao uso, são bastantes amigáveis na sua utilização e podem ser pesquisáveis facilmente; quanto ao espaço, a capacidade de armazenamento é bastante grande (600Mb por disco) e sem sombra de dúvidas, sabemos que algumas prateleiras de CD-ROM's podem substituir uma biblioteca inteira em papel, e quanto ao custo, uma vez comprado, um CD-ROM pode ser pesquisado sem custos adicionais, quantas vezes forem necessárias.

Quanto à disquete, digamos que é um tipo de documento que está em grande parte a ser substituído pelo CD-ROM. As disquetes entram numa biblioteca de variadíssimas maneiras: inseridas num livro, como parte de um pacote, como parte principal de um manual, e mesmo como documento único. Por vezes apresentam alguns vírus, ficam inutilizados, os editores têm dificuldade em substituir.

Contudo, prevê-se que num futuro mais próximo, qualquer um destes suportes seja rapidamente substituído pelas publicações em linha.

Publicações em linha e híbridas

Bases de dados em linha, bases de dados bibliográficas e bases de dados em texto completo com facilidades de manipulação e processamento de dados, são documentos que existem há alguns anos disponíveis nas bibliotecas e com a evolução da tecnologia passaram a ser de acesso directo para os utilizadores, o que de imediato facilita as pesquisas, prestando serviços que revelam rapidez e eficácia. As publicações em linha também podem ser transferidas e pesquisadas usando motores de pesquisa da Web, requerendo quando muito, algum programa adicional para que se possa aceder ao seu conteúdo.

Quanto ao livro electrónico é um tipo de documento que não se enquadra com facilidade em nenhuma destas categorias mencionadas, pelo simples facto de ser produzido em linha e acedido sem

conexão à rede, diferenciando-se e enquadrando-se em cada uma delas. Um elemento essencial na comercialização deste tipo de documento é a segurança. Considerando o modo de distribuição, é diferente do modelo em linha onde o conteúdo é limitado e controlado, neste caso, os conteúdos dos livros electrónicos são ou devem ser encriptados de modo a que não possam ser transferidos, o que agrada aos editor. Contudo, este tipo de solução cria alguns problemas aos bibliotecários nomeadamente no que respeita a gestão da preservação a longo prazo. No que respeita ao processo de depósito, é um processo complicado quando as versões não estão encriptadas, uma vez que não se tem controlo sobre o conteúdo.

O crescimento da informação digital

O aparecimento da rede electrónica marca o início de nova era. O acesso à informação é o distintivo do final do século xx e princípios do século XXI, permitindo melhorar a troca de documentos entre as bibliotecas e desenvolvendo a autonomia do utilizador.

Entre 1970 e 1990, a informatização dos ficheiros e as principais funções internas das bibliotecas foram o ponto de viragem para as bibliotecas. Os instrumentos criados na época para esta primeira vaga de mudanças foram complementadas pelo acesso às bases de dados, em diferido ou em linha e aos serviços partilhados, como a PORBASE, em Portugal.

Graças a essas bases de dados, os leitores puderam consultar não somente os catálogos da sua biblioteca, como também os das outras.

Com o crescimento da informação digital, dos documentos electrónicos, as bibliotecas tradicionais estão a mudar gradualmente para bibliotecas virtuais, digitais e os bibliotecários vêem-se, agora, confrontados com um sem número de alterações e com um grande desafio – as mudanças a nível do acesso e da disponibilidade da informação. A Internet e a Web cresceram de tal modo que aceleraram o problema de explosão da informação e hoje para melhor identificar as bibliotecas, podemos classificá-las em três tipos [5]:

- bibliotecas de papel/análogicas – biblioteca clássica com o seu ficheiro manual;
- biblioteca híbrida/automatizada – biblioteca analógica com catálogo automatizado;
- biblioteca digital – biblioteca automatizada em que maior parte da informação é digital.

Para ultrapassar esta descontinuidade tecnológica, há a considerar dez grandes pontos, segundo Sólveig Thorsteinsdóttir [6]:

1. Descobrir e recuperar informação digital e disponibilizar quer material nacional quer estrangeiro;
2. Disponibilizar de forma eficaz mecanismos de pesquisa;
3. Disponibilizar a leitura de informação digital salvaguardando a propriedade intelectual;
4. Arquivar e preservar toda a informação digital que se considere significativa culturalmente como fazendo parte do património cultural;
5. Investir em tecnologias;
6. Estabelecer relações de colaboração entre bibliotecas e vendedores de tecnologias;
7. Providenciar diferentes espécies de bibliotecas para criar documentos electrónicos que sejam úteis para a educação e pesquisa;
8. Ensinar os utilizadores a aceder e a utilizar informação digital;
9. Oferecer cursos de formação para bibliotecários de forma a que estes possam ir ao encontro dos desafios da mudança;
10. Alterar a formação na bibliotecas de forma a encontrar as necessidades da biblioteca virtual.

Mudança de atitude

A Web e os novos métodos de pesquisa não substituem as bibliotecas clássicas, o que se pretende é que as novas bibliotecas venham a proporcionar todos os serviços das bibliotecas tradicionais, acrescidas de novos serviços.

O progresso tecnológico veio mudar a forma como as bibliotecas fazem o seu trabalho. O mais profundo desenvolvimento tecnológico – a ligação de computador a computador numa inquebrantável cadeia à volta do mundo – veio alterar o conceito fundamental da biblioteca no século XXI, mais concretamente ao estabelecer a ligação directa das pessoas à informação. Assim, devido ao avanço da informática, das redes e das tecnologias para o armazenamento, como devemos nós bibliotecários, assimilar e utilizar estas tecnologias? A qualidade de acesso depende da infra-estrutura, instrumentos e experiência do pessoal da biblioteca, identificação das necessidades dos utilizadores, localização dos documentos e providência para organizar as suas próprias fontes de informação.

Significa que esta particular mudança na gestão da biblioteca, vai desenvolver as facilidades do serviço facultado ao público, o que obriga os bibliotecários a adquirirem novos conhecimentos no que se refere a sistemas de informação em suporte electrónico, incluindo a habilitação para pesquisar na variada gama de bases de dados que se nos oferecem e para a resolução de problemas técnicos.

Depois, outras questões se nos colocam. Sabendo nós que a comunidade dos editores de livros e publicações em série, é uma comunidade heterogénea e que adopta normas, que são benéficas do ponto vista interno, mas pouco aplicáveis e funcionais para o exterior, assim, como definir uma política de desenvolvimento de colecções? Como gerir o depósito das publicações electrónicas? Como catalogar? Como arquivar e preservar?

Desenvolvimento das colecções

A aquisição de publicações electrónicas pelas bibliotecas está repleta de dificuldades. A primeira delas relaciona-se com a variedade de formatos em que as publicações estão disponíveis, pois cada formato exige o seu próprio equipamento para ser adquirida e mantida. Por outro lado, para além dos custos das próprias publicações há a considerar, a variedade de fontes de informação que envolve frequentemente substituições de equipamento obsoleto.

Outros problemas surgem, como a distinção entre aquisição e acesso. A disponibilidade do acesso do documento electrónico vai encorajar a tendência para direcções várias, na política de aquisições. A aquisição de publicações em série electrónicas coloca diferentes problemas. Para resolver, estas questões em alguns países, criaram-se comités de bases de dados, os quais optaram por estabelecer determinados critérios.

Na Dinamarca, por exemplo foi criado um Comité de Bases de Dados [7], Março de 1998, cujo objectivo era:

- descobrir que bases de dados estavam disponíveis em formato digital;
- descobrir que tipo de tecnologia necessitavam;
- recomendar como estabelecer licenças de utilização;
- recomendar um modelo económico para armazenar informação digital.

Após identificação destes parâmetros criou-se um grupo de pesquisa/observatório para a selecção de bases de dados. Deviam aceder tanto a bases de dados nacionais como a estrangeiras, formando-se consórcios para áreas temáticas semelhantes, beneficiando quer em simplicidade da gestão, quer na organização da tecnologia. Facilitar a acessibilidade às

bases de dados o mais rapidamente possível, considerando a variedade dos temas e benefício dos vários grupos etários era o objectivo último.

A negociação das licenças para o acesso, ainda segundo o Comité de Bases de Dados, deve considerar os direitos de autor e propriedade dos documentos electrónicos. A propriedade dos direitos deve ser negociado com os editores, devendo para tal o bibliotecário ter assessoria jurídica em matéria de licenças.

Os documentos da Internet devem ser seleccionados tão cuidadosamente como qualquer outro documento a ser integrado na colecção da biblioteca. Devem estar envolvidos os gestores das colecções, especialistas de assuntos e outro pessoal da biblioteca que de algum modo esteja relacionado com esta tarefa.

Catálogo

Com o advento e difusão da Internet e de World Wide Web, as bibliotecas instruíram, exploraram e em alguns casos influenciaram o impacto da Internet na descrição e acesso dos documentos electrónicos.

Para efeitos de pesquisa e catalogação, os documentos electrónicos são tratados de duas maneiras, consoante o acesso seja local ou remoto.

Assim, quando existe:

- Um «suporte» físico que pode ser descrito (ex. disco, cassete, etc.) e que geralmente tem de ser inserido pelo utilizador num computador ou num periférico associado a um computador, dizemos que se trata de acesso local;
- Acesso remoto, quando não existe um «suporte» físico que possa ser manuseado pelo utilizador. O documento existe num disco rígido ou noutro dispositivo de armazenamento de um computador ou de uma rede e só é possível aceder-lhe utilizando um dispositivo saída-entrada (ex. terminal, cliente, PC, etc...).

Usar técnicas e procedimentos biblioteconómicos já existentes e criar registos para recuperar informação através de catálogos em linha, é o método mais eficiente de aceder a estes documentos, segundo a OCLC – Online Computer Library Center, que desenvolveu um guia para ajudar os participantes nesta tarefa designado *Cataloging Internet Resources: a Manual and Practical Guide* [8] e integrado no projecto «Building a Catalog of Internet Resources».

A IFLA também desenvolveu e redefiniu em 1997 a (*ISBD(ER)* – *International Standard Bibliographic Description for Electronic Resources*) [4], assim como estudou novos campos do UNIMARC [9] aplicados aos documentos electrónicos¹.

Relativamente à *ISBD(ER)*, para além das zonas comuns a outro tipo de documentos temos a considerar:

- Zona 3 – Zona do Tipo e extensão do documento, esta zona descreve as características básicas de um documento electrónico e inclui duas partes: 1. Designação do documento que é obrigatória, identifica e particulariza o tipo de documento, exemplo como o recurso é constituído por dados e programas, bases de dados em linha, jornal electrónico, etc. a 2. Extensão do documento consiste no número de ficheiros que compõem o documento, é opcional e regista-se quando estiver a informação disponível;
- Zona 4 – Zona da Publicação, produção e distribuição, no contexto da *ISBD(ER)* determinou-se que todos os documentos electrónicos, de acesso remoto ou local fossem considerados «publicados»;
- Zona 5 – Zona da Descrição física, esta zona só deve ser preenchida para documentos em que exista um suporte físico, ou seja para documentos de acesso local, sendo especialmente específica para recurso que existam em diferentes tipos de suporte, ou em diferente tamanhos;
- Zona 7 – Zona das Notas, temos Nota sobre a fonte do título, que é obrigatória; Nota sobre a história bibliográfica do item, inclui informação sobre frequência de alteração dos conteúdos de um documento de acesso remoto dinâmico; Notas sobre requisitos do sistema informático, para documento de acesso local e Nota sobre o modo de acesso, para documentos de acesso remoto, ambas são obrigatórias.

Quanto ao formato UNIMARC e aos novos campos a implementar e que correspondem às zonas da *ISBD(ER)* atrás mencionadas, definimos cinco:

- 135 – Campos de dados codificados: Documentos electrónicos;

¹ No âmbito do Comité Permanente do UNIMARC e da Secção de Catalogação da IFLA e sob orientação técnica da BN, estes dois documentos estão em fase de estudo e tradução, prevendo-se a sua conclusão para finais de 2001.

- 230 – Zona específica de alguns tipos de materiais: documentos electrónicos, campo equivalente à zona 3 da *ISBD(ER)*;
- 336 – Nota relativa ao tipo de documento electrónico, relacionada com o tipo e extensão do documento;
- 337 – Nota relativa a requisitos do sistema (documentos electrónicos), equivalente à mesma nota da *ISBD(ER)*;
- 856 – Acesso electrónico, equivalente à nota sobre modo de acesso, ou seja permite localizar um item electrónico.

Grandes esforços têm sido feitos para catalogar documentos electrónicos, principalmente os designados de acesso remoto ou via Internet, uma vez que, existe um grande conjunto de informação válida disponível na Internet e que estes documentos precisam de ser organizados, para se tornarem acessíveis.

Pesquisa e consulta

Na tradição documental clássica, a pesquisa de informação efectua-se quer sobre um registo estruturado, interrogando campos específicos, quer em texto integral do documento completo sem uma estrutura particular. As bases de dados correspondentes estão quer em formato tipo bibliográfico, quer em ficheiro de texto tipo ASCII.

Considera-se que num ambiente de trabalho, como será o da biblioteca, o leitor poderá ter acesso a numerosos documentos da cadeia da publicação electrónica e pode-se começar a imaginar funções de consulta utilizando as novas possibilidades de pesquisa. Pode-se também, apoiar sobre o código da estrutura lógica para folhear o documento, não mais na sua lógica sequencial, mas construindo uma nova lógica de consulta fundada sobre um percurso personalizado da estrutura, criado e modificado pelo leitor à vontade (figuras, fórmulas, tabelas, bibliografias, resumos...).

Há que desenvolver métodos para facilitar o acesso na recuperação da informação e serviços que fornecem documentos – um leitor pode obter um artigo em formato electrónico via servidor do editor, ou uma referência de uma base de dados para uma determinada área temática do seu interesse.

Um dos problemas maiores que se tem com os documentos electrónicos, é que os suportes em que estão armazenados mudam ou desenvolvem-se com grande rapidez.

Arquivar e preservar

Os suportes electrónicos são geralmente efémeros, se não se tiver em consideração que possam ser reformatados. Importa então que as recomendações relativas ao depósito sejam formuladas de tal maneira, que as instituições depositárias sejam autorizadas a copiar, reformatar, ou a transferir as publicações depositadas afim de assegurar a sua preservação. Sem esta autorização os documentos não podem ser conservados para a posteridade.

Segundo o grupo de trabalho da CDNL – Conférence des Directeurs de Bibliothèques Nationales [1], a preservação dos documentos electrónicos está ainda em fase de estudo e de desenvolvimento, sendo que os problemas que coloca devam ser tratados e/ou estudados em separado.

Prevêem que toda a preservação tenha assegurado o direito de fazer cópias de salvaguarda. Este direito deve figurar já nas disposições que regem o depósito legal das publicações impressas e é essencial que se estenda às publicações electrónicas, de modo a que toda a informação que esteja armazenada possa posteriormente ser lida, escutada, mostrada ou transmitida.

A preservação da informação digital culturalmente mais significativa é muito importante. Para este tipo de informação e documentos, deve ser elaborada legislação apropriada e rapidamente, pois este tipo de documento sofre alterações contínuas e por outro lado, num pequeno período de tempo pode-se perder muita informação.

Conclusão

Perceber que estamos a lidar com um mundo de informação diferente e desenvolver um plano de acção é o nosso novo desafio.

De facto, temos de considerar a construção de novos instrumentos de trabalho, uma vez que a informação digital é quase toda acessível por Internet e Web, constatamos que o volume das publicações aumenta extraordinariamente (multiplicação dos formatos, dos suportes e dos meios tecnológicos), os preços são cada vez mais elevados e verificamos que é um problema a forma como se acede à informação, que está em mudança contínua e é protegida pelos direitos de autor.

Para os editores, a colocação no mercado de informação sob a forma electrónica necessita de estudos de *marketing* aprofundados que permitirão o retorno do investimento e que são consideráveis. Não se

trata da simples transferência de produtos em suporte papel para suporte electrónico, mas sobretudo de criar produtos que apresentem funcionalidades suplementares oferecidas pelas novas técnicas e em que se torna necessário precisar o estatuto jurídico quanto à propriedade intelectual.

As bibliotecas são evidentemente os mediadores para medir a reacção do utilizador quanto ao documento electrónico. São as bibliotecas, quer sejam virtuais, híbridas ou clássicas que adquirem e gerem o arquivo dos documentos electrónicos, dominam as funcionalidades da pesquisa não somente a nível da identificação mas também a nível do conteúdo do documento electrónico, quem oferece postos de leitura e possibilita a consulta de documentos.

Quando se discute a organização das bibliotecas que contenham documentos digitais devem ser considerados vários tópicos para discussão e na perspectiva de Kristiina Hormia-Poutanen [10], devem-se ter em conta : cooperação no campo da licenças, preços, métodos de trabalho, condições de trabalho e empréstimo interbibliotecas, projectos nos campos das edições electrónicas, da digitalização de materiais com imagem e texto, catalogação e arquivo dos documentos electrónicos, etc.

Em conclusão, provavelmente a solução mais evidente reside na partilha de recursos, como forma de melhorar o acesso à informação, no fornecimento dos documentos e na coordenação do desenvolvimento das colecções. Por outro lado, é claro, também para todos nós que este problema requer a atenção, não só dos bibliotecários mas também dos políticos, especialistas em direito, informáticos e editores.

A importância da informação para a economia, cultura, indústria, governo, serviços públicos em geral deve ser reconhecida.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Le depot legal des publications electroniques [Em linha]. [Consult. 22 Nov. 2000]. Disponível na WWW: <http://www.unesco.org/webworld/memory/depot1.html>
- [2] INGLATERRA. British Library. Working Party on Electronic Publishing *The Impact of electronic publishing on library services and resources in the UK: report of the BLWPEP* [Em linha]. [Consult. 21Out 2000]. Disponível na WWW: <http://www.ifla.org/documents/libraries/net/ukrtp.txt>
- [3] IFLA *ISBD(ER) – International standard bibliographica description for electronic resources*. [Em linha].Muchen: KG Saur, 1997 (UBCIM Publications, New Series, vol. 17). [Consult. 22 Nov. 2000]. Disponível na WWW: <http://www.ifla.org/VII/s13/pubs/isbd.html>

- [4] MARK BIDE & ASSOCIATES
Standards for Electronic Publishing : an overview. Amsterdam: Nedlib, 2000.
(Nedlib Report Series 3)
- [5] SHARON, Taly; FRANK, Ariel J.
Bibliotecas digitais en internet [Em linha]. [Consult. 22 Nov. 2000]. Disponível na WWW: <http://www.ifla.org/VI/ifla66/paper/029-142s.html>
- [6] THORSTEINSDÓTTIR, Sólveig
Meeting the Challenge of virtual libraries in Iceland, Nordinfo – NYTT. Helsingfors ISSN 0356-9624. 2 (1999) 25-30
- [7] THORHAUGE, Jens
Developing a new cooperative culture: Denmark's Electronic Research Library. Nordinfo – NYTT. Helsingfors ISSN 0356-9624. 2 (1999) 10-15
- [8] OLSON, Nancy B., ed. lit
Cataloging Internet Resources: a manual and practical guide [Em linha]. Dublin, Ohio: OCLC On-line Computer Library Center, 1995 [Consult. 9 Nov. 2000]. Disponível na WWW: <http://www.oclc.org/oclc/man/9256cat/chap1.html>
- [9] IFLA
UNIMARC *Guidelines n.º 6: Electronic Resources* [Em linha]. [Consult. 22 Nov. 2000]. Disponível na WWW: <http://www.ifla.org/VI/3/p1996-1/Sec-Uni.html>
- [10] HORMIA-POUTANEN, Kristiina
Developing the National Electronic Library in Finland. Nordinfo – NYTT. Helsingfors ISSN 0356-9624. 2 (1999) 17-22

A biblioteca digital ou o acesso global

ANABELA PRISTA SARAIVA SERRANO

RESUMO Este documento constitui um estudo de caso de gestão estratégica aplicada a uma grande biblioteca universitária com enormes carências do ponto de vista financeiro, de equipamento e, sobretudo, de recursos humanos. A solução encontrada, no seio do estudo de gestão estratégica, é a de criar uma biblioteca digital como forma de proporcionar um «volte-face» estimulante que dinamize o referido serviço, fazendo-o sair do mutismo em que se encontra para o abrir ao exterior de uma forma directa e transparente. São descritas as várias etapas do estudo, assim como os aspectos mais relevantes da construção da referida biblioteca digital. Conclui-se que a biblioteca digital permite uma maior visibilidade e que, nos moldes em que foi planeada, simplifica a gestão corrente do serviço, desconhecendo-se ainda as implicações a médio / longo prazo desta decisão estratégica embora os indicadores apontem para uma dinamização do serviço, não obstante o factor de risco implícito.

ABSTRACT This paper aims to describe a case study in strategic management applied to a large academic library in great need of human, financial and equipment resources. The strategic management study suggests that creating a digital library could result in a stimulating turn-around strategy which could lead the library into a more dynamic approach, getting it out to the world throughout an opened and cleared way. All stages of the study are described, as well as the main steps to create the digital library. Finally, the conclusion achieved is that a digital library could bring more exposure and make current managerial tasks easier, although future developments of such a strategic decision are still unveiled. However, in spite of necessary risk-taking, that library could well turn into a more dynamic workplace.

Introdução

O presente artigo pretende equacionar a criação de uma biblioteca digital como uma medida estratégica para a dinamização de uma grande

biblioteca universitária. Tomar-se-á como exemplo uma biblioteca universitária com área em m² e um fundo documental consideráveis inserida numa grande e dinâmica comunidade académica e científica, mas esvaziada, desde sempre, de recursos humanos e financeiros, o que permitiu que permanecesse desactualizada em termos de gestão biblioteconómica, apesar de alguns esforços mais recentes que, apenas, tiveram resultados temporários e pouco consistentes, pois, não se alicerçaram num plano prévio nem em medidas sólidas a nível de gestão de recursos humanos e de recursos financeiros.

Para introduzir uma alteração na gestão de um serviço é desejável que se proceda de acordo com o convencionado nas escolas de gestão na disciplina de «política de gestão» ou «gestão estratégica»¹. A existência e manutenção de qualquer organização só tem razão de ser enquanto a mesma «produz» algo para alguém. O mesmo se aplica aos serviços. Um serviço surge porque é pressuposto que o mesmo venha a satisfazer uma ou várias necessidades de potenciais utilizadores. Mas, para que um serviço ou organização consubstancie a sua actividade num «produto» (qualquer que ele seja) é necessário que se defina, previamente, qual a missão e quais os objectivos daquele, e disso dar conhecimento a todos os colaboradores do mesmo. Há que equacionar os factores externos e internos que interferem na definição/estruturação do plano estratégico, e depois de este ser elaborado há que implantá-lo, o que será, decerto, a fase mais complexa e de prognóstico mais incerto.

Este artigo está, portanto, estruturado da seguinte forma: definição da missão do serviço e dos seus objectivos; identificação dos factores externos e internos; elaboração do plano estratégico; elaboração de um plano operacional a aplicar de imediato (plano de actividades para 2001); implementação de uma das medidas do plano estratégico – a criação da biblioteca digital; tópicos para avaliação da qualidade das bibliotecas digitais.

Definição da missão

O que é que se entende por missão de uma organização ou de um serviço? Segundo Luís Cardoso, a «missão identifica uma visão, cenário desejável do futuro para a organização, que possa consti-

¹ Designação preferida, ultimamente, dada a rapidez com que muda o meio envolvente o que obriga os gestores (de topo e intermédios) das organizações (públicas ou privadas) a tomarem decisões com maior rapidez e agilidade.

tuir-se numa referência, ideal comum, orientando os comportamentos e decisões e motivando a acção. É expressa de forma qualitativa...»².

«*Informar com precisão e rapidez o maior número possível de utilizadores*» é o lema que aqui se propõe para o serviço em causa. A uns parecerá redundante e vulgar, a outros (os que conhecem a realidade do referido serviço) parecerá utópico e pretensioso. Todavia, talvez seja necessária uma operação de choque e uma mudança radical de mentalidade para revitalizar a biblioteca universitária que é objecto deste artigo. O que é perfeitamente possível de levar a cabo, dado que o maior obstáculo a este tipo de tática – um quadro de pessoal arregaçado a concepções antiquadas de serviço público – praticamente não existe. De entre os funcionários do quadro de pessoal da faculdade à qual pertence a biblioteca em causa, apenas, estão em funções na biblioteca uma (1) Técnica Profissional de Biblioteca e Documentação (BD), uma (1) Auxiliar Técnica de Biblioteca e Documentação e um (1) Auxiliar Administrativo recentemente transferido para o serviço; há, ainda, seis (6) lugares vagos de Técnico Profissional BD de 2.^a classe e um lugar vago de Técnico Superior BD de 2.^a classe. Isto significa que a renovação total desta biblioteca é perfeitamente exequível – dotando a mesma de um *staff* totalmente novo, assegurando formação profissional a todos os funcionários, e motivando os actuais funcionários (progressivamente) para as mudanças a concretizar. Importa, igualmente, ter a preocupação de integrar os diferentes elementos da equipa de trabalho a criar, de modo a que todas as funções exercidas por cada elemento conduzam a um resultado prático tangível. Cada elemento da equipa deve ter a noção de qual a missão do serviço e de quais os objectivos a atingir mediante o trabalho realizado no decurso da sua actividade profissional.

Definição dos objectivos

Os objectivos a serem atingidos podem ser de âmbito mais geral ou mais específico (operacionais e quantificáveis). Mas, nesta fase, convém começar por definir os objectivos gerais que, neste caso, se podem considerar objectivos permanentes do serviço. Voltando a citar Luís Cardoso, «os objectivos constituem aquilo que a organização pretende atingir. Concretizam a missão e englobam os valores que se

² Luís Cardoso – *Gestão estratégica das organizações : ao encontro do 3.º milénio*. p. 46.

pretendem conseguir obter no futuro como resultado da actividade da organização.»³.

Para uma biblioteca do ensino superior universitário na área da ciência e da tecnologia poder-se-ão considerar os seguintes objectivos gerais:

1. Recolha, tratamento, e difusão de informação científica e técnica com interesse para o ensino e a investigação e desenvolvimento em ciência e tecnologia;
2. Permitir o acesso à informação disponibilizando documentos e informações com celeridade;
3. Disponibilizar aos utilizadores os meios adequados à pesquisa e à consulta de informação;
4. Conservação, restauro e, eventualmente, reprodução noutra tipo de suporte dos documentos de forma a que não se perca a informação neles contida.

É evidente que estes objectivos gerais são comuns, na sua essência, à totalidade dos serviços deste tipo.

Identificação dos factores externos

Após a definição da missão e dos objectivos do serviço, devemos proceder a uma análise do meio em que aquele se insere identificando os factores externos que poderão vir a condicionar a aplicação do plano estratégico.

A. Sendo o serviço em causa uma grande biblioteca universitária, haverá que equacionar a *política de informação* definida a diferentes níveis, isto é, *na União Europeia, em Portugal, e na Universidade do Porto* (instituição em que a biblioteca se insere).

No Programa Impact 2 lançado pela Comissão Europeia em 7 de Junho de 1991, preconizava-se, na sua quinta área relativa às bibliotecas, o cumprimento dos seguintes objectivos:

- a) a disponibilidade e a acessibilidade de serviços modernos de bibliotecas em toda a Comunidade (leia-se, actualmente, União Europeia);
- b) a penetração acelerada, mas rentável, de novas tecnologias neste sector;

³ IDEM, *ibidem*.

- c) a normalização necessária à cooperação;
- d) a harmonização e a convergência das políticas nacionais.

A nível nacional, foi criada uma estrutura de missão (Missão para a Sociedade da Informação) em Março de 1996 para planeamento e aplicação de uma política de informação nacional e, em Abril de 1997, o Ministério da Ciência e Tecnologia deu um importante passo estratégico ao lançar o *Livro Verde para a Sociedade da Informação*, em cujo capítulo três, com o título «O saber disponível», se estabelecem as medidas 3.2 e 3.3 relativas, respectivamente, ao «Desenvolvimento de Bibliotecas Digitais» e à «Digitalização do Património Cultural». Algumas das ideias expressas neste documento são reafirmadas, mais tarde, pelo MCT no *Programa Operacional Sociedade de Informação* do actual quadro comunitário de apoio (2000-2006):

O Programa Operacional da Cultura elegeu como instrumento privilegiado de promoção do acesso à cultura a utilização de novas tecnologias na divulgação do nosso património histórico, arquitectónico, arqueológico, museológico e documental.

Contém, por isso, diversas acções relativas ao tratamento, digitalização e informatização dos elementos essenciais do património nacional, com vista à sua disponibilização – por meios telemáticos ou em suporte digital – como forma de afirmação da identidade nacional e europeia mas igualmente como factor de desenvolvimento económico.»⁴

A mesma ideia tem seguimento, já a nível operacional, no *Livro Branco do Desenvolvimento Científico e Tecnológico Português : 1999-2006*, do MCT, na parte que respeita à «Intervenção Operacional Sociedade da Informação», sob o título «Desenvolver e adaptar sistemas de armazenamento, relacionamento, procura e transformação de informação»:

A integração de grandes quantidades de informação suportada nos vários modos de apresentação – dados, texto, imagens ou som – e a sua acessibilidade e manuseamento rápido e fácil, requer avanços consideráveis em relação às tecnologias disponíveis. São necessárias tecnologias de interacção multimodal homem-computador de diversos tipos – em qualquer linguagem, clicando um rato, tocando um ecrã, falando com o computador ou por interacção gestual – de forma a reduzir a dependência da facilidade de acesso da cultura, educação ou destreza física do utilizador. São também necessários novos métodos de localização, relacionamento e integração de informação em bases de dados de muito grande dimensão. Estas novas metodologias devem ser desenvolvidas, aprendidas ou adaptadas em torno de situações concretas: *bibliotecas digitais*, gestão de crises e mitigação de riscos

⁴ Acessível na Internet: <http://www.fct.mct.pt/qca/posi/posi.htm>

públicos, monitorização e gestão do mar e do ambiente, gestão de transportes, investigação de mercado em sectores críticos, *modernização da administração pública, apoio ao ensino*, informação médica.»⁵

Das citações acima transcritas se infere que a criação de uma biblioteca digital se insere, perfeitamente, em programas de apoio ao desenvolvimento das instituições no seio da política de informação vigente em Portugal e na Europa. Mas, a biblioteca universitária em causa enferma de ausência quase total de recursos humanos qualificados de nível superior que permitam a formação de uma equipa que assegure o planeamento de um projecto, o seu lançamento em concurso e que, se e uma vez aprovado, garanta a prossecução do mesmo na calendarização prevista. E será muito problemático, para não dizer impossível, conseguir constituir uma equipa – com elementos em número suficiente e com perfil adequado – para concretizar uma candidatura a um programa no actual quadro comunitário de apoio que cessa em 2006. Possivelmente, o próximo quadro comunitário de apoio a decorrer a partir de 2007 já não visará as bibliotecas digitais, o que significa que se se aguardar uma nova oportunidade para dar um passo em direcção à era digital ela, hipoteticamente, não surgirá.

A nível institucional, na Universidade do Porto (UP), o *Guia de Desenvolvimento Estratégico: 2000-2004*, refere como componente da sua definição de missão (alínea d do ponto dois) «a conservação e divulgação dos conhecimentos, das obras de cultura e das técnicas que configuram, em cada momento, o património disponível para utilização criativa dos especialistas e do público»⁶. Em termos de planeamento plurianual, nas propostas de «Grandes Opções de Desenvolvimento 2000-2004», no ponto dois do ponto oito (8.2) relativo a redes de informação e comunicação prevê-se a «qualificação e articulação dos serviços de recolha, tratamento, conservação e difusão de informação, tais como bibliotecas, arquivos e centros de documentação; e desenvolvimento de novos serviços de informação e lazer.»⁷. A inclusão dos objectivos dos serviços de informação na missão da Universidade do Porto é muito significativa, mas carece da definição de uma política de informação no seio daquela instituição. A essa mesma conclusão se chegou na reunião de 16 de Junho de 2000 realizada na Reitoria da UP – é necessário

⁵ Acessível na Internet: <http://www.fct.mct.pt/Livro-BrancoCT/>

⁶ Acessível na Internet: <http://www.up.pt/conhecaup/orghist/docinstitui/guiadeup/htm>

⁷ *Ibidem*.

começar por definir uma política de informação para toda a UP para que se possa concretizar o catálogo comum das suas bibliotecas e outras ideias como consórcios de assinatura de publicações periódicas, de bases de dados, etc. Entretanto, a Reitoria avançou com a assinatura da *Current Contents* via Internet, o que, pelo menos para a biblioteca analisada, constituiu uma boa novidade. É que enquanto se discutem políticas em reuniões anuais, há licenciaturas, doutoramentos, mestrados e projectos de investigação que se vão concluindo sem o apoio desejável do ponto de vista documental. Defendendo o planeamento para uma melhor prossecução dos nossos objectivos, não podemos deixar de observar que, por vezes, os directamente interessados (os utilizadores) são prejudicados – mesmo que disso não tenham consciência – pela ausência de iniciativa de cada serviço em particular. E se algumas carências financeiras e de recursos humanos são, ou parecem, inultrapassáveis, outras há que se podem (com muito esforço!) contornar. E é esse o raciocínio subjacente à elaboração do plano estratégico para a biblioteca estudada, adiante delineado, e do primeiro plano operacional a que este último deu origem (o plano de actividades para 2001) – a criação de uma biblioteca digital como face visível de toda uma série de outras medidas que serão tomadas, simultânea ou subseqüentemente, a partir do ano 2001.

B. Como se trata de uma biblioteca universitária há que equacionar, igualmente, a *política de educação* subjacente para o ensino superior universitário e em termos gerais. A tendência actual é a da educação permanente, pelo que a necessidade de informação actualizada é, cada vez mais, uma constante. O ensino à distância através da Internet (*e-learning*) encontra-se em franca expansão, o que estimula o aumento do número de utilizadores da Internet.

C. Em termos de *política científica*, também se incentivam as actividades de investigação e desenvolvimento (I&D), pelo que se faz sentir uma maior necessidade de actualização por parte dos utilizadores no meio académico e científico.

D. Paralelamente, o *desenvolvimento tecnológico* conduz a uma progressiva aceleração na procura de informação por parte das pessoas, dado que computadores e telemóveis têm cada vez mais funcionalidades e são cada vez mais rápidos e interconectáveis. Encontramo-nos num mundo, tendencialmente, digital e sem fios, pulverizado de redes de comunicação onde tudo parece acessível, mas onde há

ainda muito mais para disponibilizar do que aquilo que se possa supôr. A *globalização* preconizada por McLuhan, na sua imagem de «aldeia global», é já uma realidade.

E. E até mesmo os *editores* começaram a publicar noutro suporte, que não o papel ou os discos ópticos, – isto é, na Internet –, pois, a edição torna-se muito mais rápida e mais barata. Recentemente, pudemos verificar isso mesmo aquando da publicação do dossier sobre o genoma humano pela *Science* e pela *Nature*: o dossier foi lançado na Internet numa segunda-feira, a versão impressa saiu na quinta-feira seguinte e, apenas, chegou à biblioteca analisada (que é assinante) na segunda-feira seguinte. Se já tivéssemos a biblioteca digital construída, teríamos uma ligação para aquelas páginas que disponibilizasse de imediato o documento a todos os nossos utilizadores virtuais em simultâneo.

F. Para além de todos os factores atrás expostos, e também por influência daqueles mesmos (entre outros que aqui não foram referidos), o *tipo de utilizador* dos serviços de informação do ensino superior está a mudar. O utilizador actual procura, sobretudo, informação acerca de um determinado assunto, ou seja, pretende localizar todos os documentos relevantes para esse mesmo assunto. Pretende, igualmente, aceder ao conteúdo dos documentos que julgue pertinentes para a sua pesquisa, o mais rapidamente possível, onde quer que aqueles se encontrem, da forma mais cómoda possível. O actual utilizador do ensino superior está habituado a utilizar computadores e a aceder à Internet, embora possa desconhecer os operadores de pesquisa a utilizar em bases de dados bibliográficos. A faixa etária dos utilizadores é cada vez mais ampla, dado que se pretendem formar as pessoas ao longo de toda a vida (conceito de educação/formação permanente), e os utilizadores são, frequentemente, pessoas muito ocupadas que distribuem o seu tempo entre a carreira, a família e a sua educação/formação. Têm pressa de obter a informação de que necessitam e, desde que possam, não se importam de pagar uma série de dólares para obterem uma reprodução via fax ou *email* de um documento que só exista, por exemplo, numa biblioteca americana. A pesquisa directa do documento primário por autor/título deixou de ser predominante, sendo mais frequente nos primeiros anos dos cursos de licenciatura.

Este quadro conjuntural obriga-nos a repensar os serviços que disponibilizamos e a forma como os disponibilizamos. Que tipo de acesso à informação estamos a facultar aos nossos utilizadores? Será

o mais adequado? Um inquérito periódico aos utilizadores facilita a tomada de decisão em gestão estratégica, mas os resultados apurados vão depender muito da própria elaboração do questionário – quanto mais flexível fôr o mesmo, mais interessantes poderão ser os resultados.

Identificação dos factores internos

Temos, ainda, que considerar os factores internos que agilizam ou condicionam, consoante o caso, o serviço perante a mudança estratégica a que se pretende proceder.

Na biblioteca estudada os recursos humanos são escassos e pouco qualificados, os recursos financeiros parcos (cerca de 4900 contos de orçamento anual para todo o tipo de despesas), e o catálogo foi informatizado mediante um programa informático de produção caseira cujas verdadeiras funcionalidades não vão além da pesquisa por autor, por título e por termo aleatório. Com um quadro tão negro, o maior obstáculo para a recuperação do serviço poderá ser o aparente oásis daquele – o catálogo informatizado! Sem recursos humanos qualificados, a conversão dos dados inseridos no programa criado no Lotus Approach para o Porbase 5.0 ou para o Aleph redundará numa tarefa impraticável. Mas, impraticável é, também, a manutenção da base de dados existente! E quanto mais se adiar a decisão mais difícil será a conversão de todos os registos. Sem pesquisa por assuntos, o cenário quotidiano de atendimento ao utilizador neste serviço é o seguinte: a sala de leitura é utilizada para leitura de apontamentos ou de documentos não solicitados à biblioteca; a taxa de requisição anual é baixíssima; os utilizadores que procuram documentos sobre um dado assunto, quando se apercebem de que os documentos não estão indexados e de que a base de dados não permite a recuperação por palavra-chave, ou vão embora de imediato ou tentam uma pesquisa por título com a taxa de sucesso que se adivinha (!) e, depois de examinarem rapidamente os documentos solicitados, acabam por abandonar a biblioteca desiludidos. Mas, este serviço usufrui de uma vantagem que é comum a toda a comunidade dos seus utilizadores mais directos (discentes, docentes, investigadores e funcionários) – a rede informática da faculdade. A Intranet e a ligação diária à Internet poderão vir a ser a tábua de salvação da Biblioteca Geral se a soubermos utilizar – a criação e manutenção de uma biblioteca digital ficam largamente facilitadas.

Diagnóstico do serviço

Tipologia do serviço: biblioteca universitária.

Fundo documental: cerca de 100 000 documentos, incluindo um fundo de Livro Antigo de cerca de 10 000 volumes e uma colecção de gravuras de cerca de 950 peças.

Área de especialização do fundo: ciência e tecnologia.

Recursos financeiros:

- orçamento anual de cerca de 4900 contos,
- receitas no ano 2000 de cerca de 170 contos,
- receitas anuais até 1999 não quantificadas.

Recursos humanos:

- 1 Director de Serviços que não pertence à carreira de Biblioteca e Documentação (BD),
- 1 Técnico Superior BD contratado a termo certo,
- 1 Técnico Profissional BD,
- 1 Auxiliar Técnico BD,
- 1 Auxiliar Administrativo.

Nota: há oito (8) Técnicos Profissionais BD distribuídos pelas bibliotecas departamentais que não têm sido orientados nem auxiliados pela Biblioteca Geral.

Área em m²: cerca de 430 m².

Sala(s) de leitura: uma com 130 m², 44 mesas individuais, e um posto de consulta da base de dados local (a base de dados está, também, acessível na Internet).

Serviço de empréstimo: o programa informático utilizado inclui um módulo de circulação e empréstimo integrado.

Serviço reprográfico: em regime de *self-service* mediante cartão magnético.

Informatização do fundo documental:

- 22 000 monografias e teses;
- publicações periódicas entradas por compra.

Publicações da Biblioteca Geral: não foram encontrados registos nos últimos anos.

Exposições realizadas, exclusivamente, com documentos da BG: não foram encontrados registos; está prevista uma em 2001.

Exposições em parceria: duas nos últimos vinte e cinco anos (1979, 1987).

Cedência de documentos para exposições: várias; quatro (4) cedência a exposições de vários documentos desde Setembro de 2000 até Setembro de 2001.

Plano estratégico

O diagnóstico do serviço atrás apresentado poderia ser mais exaustivo e mais preciso, mas, mesmo assim, a sua leitura permite-nos retirar algumas conclusões. Os recursos humanos e financeiros são escassos, mas a situação é reversível, até porque a qualidade do fundo documental desta biblioteca permite que se consigam criar algumas receitas. Para além do mais, o quadro de pessoal não se encontra totalmente preenchido e deveria ser alargado.

A estratégia de mudança irá, basicamente, assentar em:

1. Motivação do pessoal existente para a «mudança» no serviço;
2. Formação profissional do pessoal existente;
3. Preenchimento das vagas existentes no quadro de pessoal da instituição e eventual alargamento do quadro;
4. Contratação de prestadores de serviços até o quadro estar preenchido;
5. Estabelecimento de uma rede de comunicação entre as diferentes bibliotecas (geral e departamental) via *email*, por exemplo;
6. Realização de reuniões periódicas de coordenação e informação e de troca de impressões com todo o pessoal das diferentes bibliotecas;
7. Criação de receitas próprias do serviço;
8. Uniformização do tratamento técnico nas várias bibliotecas;
9. Coordenação da selecção, aquisição e armazenamento dos documentos pela biblioteca geral;
10. Criação de um grupo consultivo de apoio à gestão das bibliotecas constituído por docentes e/ou investigadores dos diferentes departamentos;
11. Uniformização do sistema de informação em todas as bibliotecas e constituição de um catálogo comum acessível por *http* na Internet;
12. Criação de uma biblioteca digital acessível por *http* na Internet.

Este é um plano a implantar plurianualmente, cuja aplicação vai ser demorada.

Plano operacional 2001

Para o ano de 2001 foi proposto, no plano de actividades da biblioteca, um plano operacional que aplica, parcialmente, algumas das medidas previstas alicerçando-se em:

1. «Reorganização dos fundos documentais uniformizando os sistemas de cotação;
2. Reconversão do catálogo para um sistema informático compatível com os das outras bibliotecas universitárias;
3. Estabelecimento de uma política de indexação, e realização da indexação retrospectiva dos diferentes fundos documentais de acordo com as prioridades estabelecidas no plano delineado;
4. Continuação da informatização dos fundos documentais;
5. Criação de uma biblioteca digital;
6. Estabelecimento de uma actividade editorial de difusão/promoção do serviço;
7. Realização de actividades de divulgação dos fundos documentais do serviço;
8. Preservação dos documentos de colecções especiais mediante a sua protecção com material *acid-free*;
9. Desenvolvimento de acções conducentes à criação de verbas próprias do serviço.»⁸

A biblioteca digital

Constante do ponto cinco dos objectivos gerais do referido Plano de Actividades para 2001, a criação de uma biblioteca digital pode parecer utópica e descabida a muitos, sobretudo após a leitura do diagnóstico do serviço. Porém, dado o dramatismo da situação actual desta biblioteca, a criação de uma biblioteca digital poderá constituir uma medida estratégica de mudança daquela. Apenas, teremos de ter o cuidado de a utilizar de forma a que não aumente drasticamente o volume de serviço da biblioteca, e que, se possível, agilize alguns dos serviços prestados habitualmente pela biblioteca.

Antes de mais, teçamos algumas considerações em torno deste conceito e realidade emergente em biblioteconomia. O que é que se entende

⁸ Biblioteca Geral da Faculdade de Ciências da Universidade do Porto – *Plano de actividades : 2001*, p. 7.

por «biblioteca digital»? Após a análise de registos recuperados na ERIC e na Dissertations Abstracts, e de múltiplos artigos recentemente publicados em publicações periódicas da especialidade, chega-se à conclusão de que o conceito de biblioteca digital é ainda muito vago porque é encarado em múltiplas acepções. Todavia, há uma definição básica que julgo que não escapa a ninguém: biblioteca digital é um serviço de informação virtual de acesso global porque disponível numa rede informática ligada a múltiplas outras (World Wide Web). Ora, isto tem várias implicações e uma delas poderá ser mesmo o facto de o ciclo de Gutenberg ou da imprensa se estar a completar, tendo início algo que se poderá designar de «ciclo digital» ou «era digital». As bibliotecas digitais são, neste momento, objecto de muitos projectos de investigação – uns para obtenção de graus académicos de mestrado ou de doutoramento, outros à escala nacional como nos Estados Unidos da América, no Reino Unido, ou na Alemanha – em várias áreas científicas das ciências da informação. Os modelos de construção são variados podendo concentrar-se, segundo Tefko Saravecic da Rutgers University⁹, em três (3) tipologias de acordo com o conceito de biblioteca digital subjacente:

- A – o modelo de digitalização;
- B – o modelo de consórcio;
- C – o modelo de integração de serviços.

O modelo A baseia-se na digitalização dos fundos documentais, de que é exemplo o American Memory Project da Library of Congress¹⁰. O modelo B baseia-se na criação de uma biblioteca digital nacional, de que é exemplo a National Digital Library of the Library of Congress. O modelo C baseia-se na disponibilização de serviços de biblioteca na Internet para além de permitir o acesso a documentos digitalizados, de que é exemplo a California Digital Library¹¹.

De entre os três modelos atrás descritos, afigura-se-nos com evidência que o modelo de consórcio não é aplicável à nossa situação e que o modelo de digitalização não satisfaz as necessidades dos nossos utilizadores. Logo, o modelo de integração de serviços afigura-se-nos como sendo o mais adequado à biblioteca. É claro que o modelo C é bastante mais complexo que o modelo A, e terá de ser

⁹ Acessível na Internet: <http://www.scils.rutgers.edu/people/faculty/tefko.html>

¹⁰ Acessível na Internet: <http://memory.loc.gov>

¹¹ Acessível na Internet: <http://www.cdlib.org/directory/>

muito bem pensado para que a biblioteca digital daí resultante não origine uma avalanche de solicitações a que o serviço não consiga responder. Porém, actualmente, o problema da capacidade de resposta já existe, dado que se trata de uma grande biblioteca sem recursos humanos suficientes. O que não existe ainda é a possibilidade de preenchimento de formulários *online* nem um catálogo cujas funcionalidades de pesquisa permitam que os utilizadores recuperem informação relevante, o que poderia agilizar muito a gestão corrente do serviço.

A biblioteca digital a criar irá integrar no mesmo espaço áreas relativas a fornecedores, a utilizadores, a OPAC's e a associações científicas. Será uma biblioteca de acesso bilingue – em Português e em Inglês – com total duplicação de todas as ligações, menus e textos para a língua inglesa.

A estrutura do menu principal irá integrar as seguintes ligações:

ÁREA DE REGISTO DO UTILIZADOR

1. Registo de utilizador/edição de perfil;

ÁREA DE AQUISIÇÕES

2. Ligação para um submenu com ligações para editores e livreros;
3. Formulário para preenchimento *online* de propostas de aquisição;
4. Lista das aquisições mais recentes;

ÁREA DE REFERÊNCIA

5. Catálogo comum da faculdade (integrando os das bibliotecas departamentais);
6. Ligação para um submenu com ligações para as bibliotecas departamentais;
7. Ligação para um submenu com ligações para catálogos de bibliotecas de ciência e tecnologia portuguesas e estrangeiras;
8. Ligação para um submenu com ligações para bibliotecas universitárias portuguesas e estrangeiras;
9. Ligação para um submenu com ligações para bibliotecas nacionais;
10. Ligação para um submenu com ligações para bases de dados *online* (incluindo a *Current Contents*);
11. Ligação para um submenu com ligações para obras de referência (dicionários, enciclopédias, *thésauri*, etc.);

ÁREA DE PUBLICAÇÕES PERIÓDICAS

12. Ligação para um submenu com ligações para publicações periódicas *online* que poderá ser subdividido por área temática;

ÁREA DE LEITURA

13. Reserva de documentos (formulário);
14. Empréstimo Inter-Bibliotecas (EIB);
15. Pedido de reprodução de documentos (formulário);
16. Pedido de emissão/recarregamento de cartão magnético (formulário);

ÁREA DE RESERVADOS

17. Coleção de gravuras digitalizada;
18. Fundo de Livro Antigo digitalizado;

ÁREA DE PUBLICAÇÕES DA BIBLIOTECA GERAL

19. Boletim de sumários;
20. Guia do utilizador;
21. Catálogos de exposições;

ÁREA DE EXPOSIÇÕES

22. Agenda de exposições;
23. Exposições da Biblioteca Geral – galeria virtual;
24. Pedido de cedência de documentos para exposições (formulário);

VÁRIA

25. Ligação para um submenu com ligações para associações científicas;
26. Ligação para um submenu com ligações para operadores de pesquisa;
27. Outras ligações;

ÁREA DE SERVIÇO INTERNO

28. Informações;
29. Organograma do serviço;
30. Contactos.

Tópicos para avaliação da qualidade das bibliotecas digitais

Uma biblioteca digital pode ser avaliada qualitativamente em várias vertentes. Cada uma delas oferece um campo aberto de discussão que

merece um tratamento mais profundo e individualizado que, como é evidente, não poderá ter lugar neste artigo dada a extensão que o mesmo já leva. Evocar-se-ão aqui aquelas cuja inobservância possa prejudicar o sucesso da criação da biblioteca digital atrás descrita:

- o interface;
- o tratamento dado às imagens digitalizadas;
- as funcionalidades de pesquisa (a existência de *thésauri*/índices como uma mais-valia);
- a actualização dos dados.

Perante as características da estrutura acima descrita, torna-se necessário pensar num *interface* que permita controlar o acesso – a ligações a bases de dados e a publicações periódicas assinadas – autenticando os utilizadores, que permita realizar pesquisas simultâneas em diferentes bases de dados, que permita o maior número de funcionalidades de pesquisa. Todos estes requisitos são assegurados pelo protocolo Z39.50 que constitui um interface comum a muitos catálogos de bibliotecas acessíveis na Internet, pelo que a sua utilização constituirá, decerto, um valor acrescentado.

A constituição de dois núcleos de *imagens digitalizadas* de tamanho considerável, relativos à Colecção de Gravuras e ao Fundo de Livro Antigo, coloca outra questão muito importante ao nível da pesquisa das próprias imagens. É nossa preocupação oferecer mais do que um simples rastreio sequencial das imagens. À medida que os documentos forem sendo digitalizados, atribuir-se-á uma legenda a cada imagem de modo a facultar a pesquisa textual das mesmas. Paralelamente, criar-se-ão os respectivos catálogos de documentos reservados, a cujos registos irão sendo acopladas todas as imagens relacionadas. Deste modo, a pesquisa das imagens ficará largamente facilitada.

Quanto às *funcionalidades de pesquisa* a disponibilizar na biblioteca digital, o ideal seria poder facultar aos utilizadores o maior número possível de funcionalidades. Aqui o lema a adoptar é o «quanto mais melhor». No entanto, algumas delas são mais imprescindíveis do que outras. A melhor forma de recuperar registos relevantes em bases de dados bibliográficos (não se obtêm os mesmos resultados em texto integral) consiste na utilização de *thésauri* e de índices *online* que, programados para funcionarem como controladores de qualidade dos catálogos, orientem os utilizadores direccionando-os para os termos preferenciais equivalentes utilizados nos catálogos da faculdade sempre que aqueles pesquisem por termos não-preferenciais. Os operadores booleanos (os três, e não apenas um ou dois como tem vindo

a tornar-se moda) são imprescindíveis para a combinação de termos de qualquer tipo. E os operadores de adjacência, de proximidade, e a truncatura constituem auxiliares preciosos da pesquisa em texto integral. O estabelecimento de um menu com diferentes níveis de pesquisa é muito importante, uma vez que também existem vários níveis de utilizadores – desde o discente recém-chegado à universidade até ao investigador familiarizado com todos os operadores de pesquisa e que necessita de informação mais específica e mais precisa. A pesquisa por um dado campo revela-se útil para restringir pesquisas em bases de dados, e a pesquisa lateral permite orientar o utilizador nos casos menos bem sucedidos. A pesquisa lateral consiste na possibilidade de pesquisar uma palavra-chave não solicitada desde que a mesma se encontre num registo recuperado de outra forma, e esteja activada permitindo a ligação para uma instrução de pesquisa que recupera todos os registos que contêm o mesmo termo no mesmo campo de dados.

A actualização dos dados que compõem a biblioteca digital deve ser uma preocupação constante, pelo que há que ter muito cuidado com os aspectos de administração/manutenção da mesma. A biblioteca digital a criar irá constituir uma medida estratégica para implantar a mudança no serviço, mas constituirá também uma operação de *marketing* à escala global sendo as suas páginas na Internet a imagem exterior do serviço. E uma boa imagem constituirá sempre um factor de sucesso... não esquecendo que ela será tanto melhor quanto melhor fôr o conteúdo da biblioteca digital!

Conclusão

O objectivo proposto na introdução deste artigo – equacionar a criação de uma biblioteca digital como uma medida estratégica para a dinamização de uma grande biblioteca universitária – é passível de uma conclusão em aberto. O percurso feito na análise do objecto de estudo seleccionado demonstrou a necessidade de se tomarem medidas urgentes que se enquadrem na conjuntura criada pelos condicionantes externos e internos, mas que, de igual modo, chamem a atenção dos potenciais utilizadores para aquele serviço adormecido de que já poucos sentem a falta. E, teóricamente, demonstrou-se que a biblioteca digital poderá ser uma aposta adequada e que, seguramente, trará maior visibilidade à biblioteca que foi objecto deste estudo de caso agilizando a sua gestão corrente. Se, na prática, se irão alcançar os resultados pretendidos não se correndo um risco demasiado ele-

vado, isso só o futuro próximo o dirá. Mas, se aquela ocasionar um maior volume de solicitações, essa é a prova bastante de que tem cabimento preencher os lugares vagos no quadro, e de que o argumento segundo o qual não é necessário mais pessoal porque a biblioteca geral não tem utilizadores não tem razão de ser. E a taxa de sucesso irá depender do próprio empenhamento do serviço na prossecução dos objectivos aqui expostos. Daí que esta seja uma conclusão em aberto!

REFERÊNCIAS BIBLIOGRÁFICAS

- CARDOSO, Luís
Gestão estratégica das organizações : ao encontro do 3.º milénio. 2.ª ed. revista. Lisboa : Verbo, 1997.
- HOLMBERG, Melissa
 «Using publishers' web sites for reference collection development». *Issues in Science and Technology Librarianship* [online]. Winter 2000, 25 [citado em 2000.12.02]. Acessível na Internet: <http://www.library.ucsb.edu/istl/00-winter/article3.html>. ISSN 1092-1206.
- LAVAGNINO, Merri Beth (ed.)
 «System security in the networked library». *Library Hi Tech* [online]. 1997, 15, 1-2 [citado em 2000.10.06]. Acessível na Internet: <http://www.lib.msu.edu/hi-tech/ht15.12.html>. ISSN 0737-8831.
- MCGEACHIN, Robert B.
 «Selection criteria for web-based resources in a science and technology library collection». *Issues in Science and Technology Librarianship* [online]. Spring 1998, 18 [citado em 2000.10.06]. Acessível na Internet: <http://www.library.ucsb.edu/istl/98-spring/article2.html>. ISSN 1092-1206.
- MERCER, Linda S.
 «Measuring the use and value of electronic journals and books». *Issues in Science and Technology Librarianship* [online]. Winter 2000, 25 [citado em 2000.12.02]. Acessível na Internet: <http://www.library.ucsb.edu/istl/00-winter/article1.html>. ISSN 1092-1206.
- PORTUGAL. Ministério da Ciência e Tecnologia
Livro Branco do Desenvolvimento Científico e Tecnológico Português : 1999-2006. [online]. [Citado em 2000.12.28]. Acessível na Internet: : <http://www.fct.mct.pt/Livro-BrancoCT/>
Livro Verde para a Sociedade da Informação. [online]. Lisboa : MCT, 1997. [Citado em 2000.12.28]. Acessível na Internet: : <http://www.fct.mct.pt/>
Programa Operacional Sociedade de Informação. [online]. [Citado em 2000.12.28]. Acessível na Internet: <http://www.fct.mct.pt/qca/posi/posi.htm>
- SARACEVIC, Tefko
Digital libraries : variety of perspectives and models. [online]. [Citado em 2000.11.23]. Acessível na Internet: <http://www.scils.rutgers.edu/people/faculty/tefko.html>
- SMITH, Alastair G.
 «Search features of digital libraries». *Information Research*. [online]. April 2000, 5, 3 [citado em 2000.10.10]. Acessível na Internet: <http://www.shef.ac.uk/~is/publications/infres/paper73.html>

STACKPOLE, Laurie E.; KING, Richard James

«Electronic journals as a component of the digital library. *Issues in Science and Technology Librarianship* [online]. Spring 1999, 22 [citado em 2000.12.02]. Acessível na Internet: <http://www.library.ucsb.edu/istl/99-spring/article1.html>. ISSN 1092-1206.

STARR, Susan S.

«Building the collections of the California Digital Library». *Issues in Science and Technology Librarianship* [online]. Winter 1998, 17 [citado em 2000.10.06]. Acessível na INTERNET: <HTTP://WWW.LIBRARY.UCSB.EDU/ISTL/98-WINTER /ARTICLE2.HTML>. ISSN 1092-1206.

TURNER, Fay

«Selecting a Z39.50 client or web gateway». *Library Hi Tech* [online]. 1998, 16, 2 [citado em 2000.10.06]. Acessível na Internet: <http://www.lib.msu.edu/hi-tech/lht16.2.html>. ISSN 0737-8831.

UNIVERSIDADE DO PORTO

Guia de Desenvolvimento Estratégico : 2000-2004. [online]. [Citado em 2000.12.28]. Acessível na Internet: <http://www.up.pt/conhecaup/orghist/docinstitui/guiadeup/htm>

UNIVERSIDADE DO PORTO. Faculdade de Ciências. Biblioteca Geral

Plano de actividades : 2001. porto : [bgfcup], 2001.

XIAO, Daniel; MOSLEY, Pixey Anne; CORNISH, Alan

«Library services through the world wide web». *The Public-Access Computer Systems Review* [online]. 1997, 8, 4 [citado em 2000.11.23]. Acessível na Internet: <http://info.lib.uh.edu/pacsrev.html>.

WELLBORN, Victoria

«Building websites for science literacy». *Issues in Science and Technology Librarianship* [online]. Winter 2000, 25 [citado em 2000.12.02]. Acessível na Internet: <http://www.library.ucsb.edu/istl/00-winter/article2.html>. ISSN 1092-1206.

WISHARD, Lisa

«Precision among Internet search engines : an Earth sciences case study». *Issues in Science and Technology Librarianship* [online]. Spring 1998, 18 [citado em 2000.12.02]. Acessível na Internet: <https://www.library.ucsb.edu/istl/98-spring/article5.html>. ISSN 1092-1206.

Elementos para uma avaliação de fontes de informação na Internet

OLÍVIA PESTANA

RESUMO O objectivo deste artigo é definir as linhas essenciais para um processo de avaliação das fontes resultantes da pesquisa na Internet. São, assim, apresentados os elementos fundamentais para o estabelecimento duma análise que permita determinar a adequabilidade da fonte ao pretendido tendo em conta a informação e a navegabilidade.

ABSTRACT The aim of the current paper is to define the essential structure to allow an evaluation process of resources identified by searching in the Internet. Therefore, fundamental elements to establish an analysis are introduced in order to determine adequacy between resource and expectations bearing in mind both existing data and navigation capabilities.

Porquê avaliar

No contexto da explosão de informação ocorrida nas últimas décadas, a Internet é, sem dúvida, o recurso mais revolucionário de sempre pela facilidade e liberdade permitidas no acesso a milhares de fontes de informação. Vista como uma espécie de «poção mágica», o entusiasmo daí decorrente tem ocasionado a ausência de avaliação criteriosa das fontes e da respectiva origem por parte da generalidade dos utilizadores.

O aparecimento, cada vez mais crescente, de páginas pessoais, a desactualização da própria informação, dificuldades no acesso e alguma falta de exactidão na apresentação dos factos contribuem para que seja necessário criar mecanismos de avaliação das fontes.

Pode-se partir para a pesquisa estabelecendo uma boa estratégia de utilização das formas disponíveis, seja em motores de pesquisa, em catá-

logos de assuntos e directórios ou em bibliotecas virtuais. No entanto, será sempre aconselhável uma avaliação por parte do utilizador final.

Não se pretende, aqui, fazer uma análise da melhor orientação a tomar no acto de pesquisa, mas sim traçar as linhas básicas de um processo de avaliação das fontes resultantes da pesquisa. Para tal, apresenta-se, seguidamente, os elementos fundamentais para o estabelecimento de uma análise a executar pós ou em simultâneo com a pesquisa de modo a determinar a adequabilidade da fonte ao pretendido e tendo como base dois campos fundamentais: a informação e a navegabilidade.

Avaliação da informação

1. Finalidade e audiência

Numa primeira abordagem importa determinar se estão mencionadas a finalidade e a audiência a que se dirigem. Estes aspectos podem, de imediato, clarificar se se vai ou não encontrar a informação procurada ao estar mencionado o nível de conhecimentos dos utilizadores.

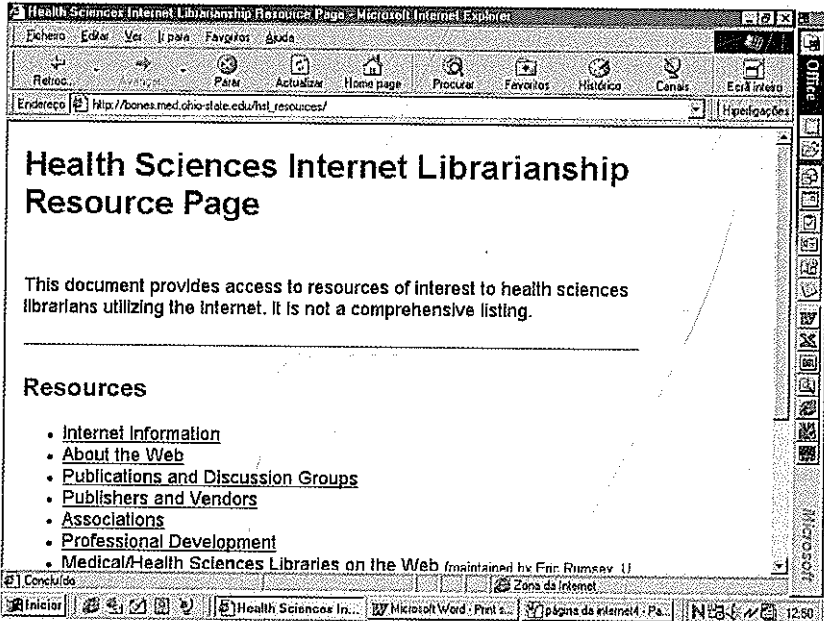


Figura 1 – Exemplo de identificação da audiência-alvo

Na figura 1 pode ver-se um exemplo que identifica a audiência-alvo, mencionando que o documento apresentado fornece acesso a recursos de interesse a bibliotecários da área das ciências da saúde (Health Sciences Internet Librarianship Resource Page – http://bones.med.ohio-state.edu/hsl_resources – acesso efectuado em 28.11.2000).

Na figura 2 pode visualizar-se um exemplo de esclarecimento da finalidade da fonte, sendo referido que não é uma listagem de fontes primárias, mas sim indicação e aconselhamento de *sites* que permitem chegar a esses recursos (Librarian and Information Science, Librarianship. Finding Resources – <http://www.lub.lu.se/netlab/documents/lisres.html> – acesso efectuado em 28.11.2000).

2. Autoridade

Deve-se, obrigatoriamente, determinar a origem da fonte que se encontra: quem é o autor e se apresenta contacto, se é especialista na área, se tem por detrás uma organização ou instituição ou se a autoria se deve a uma organização reconhecida. No último caso importa, ainda, verificar se é o *site* oficial.

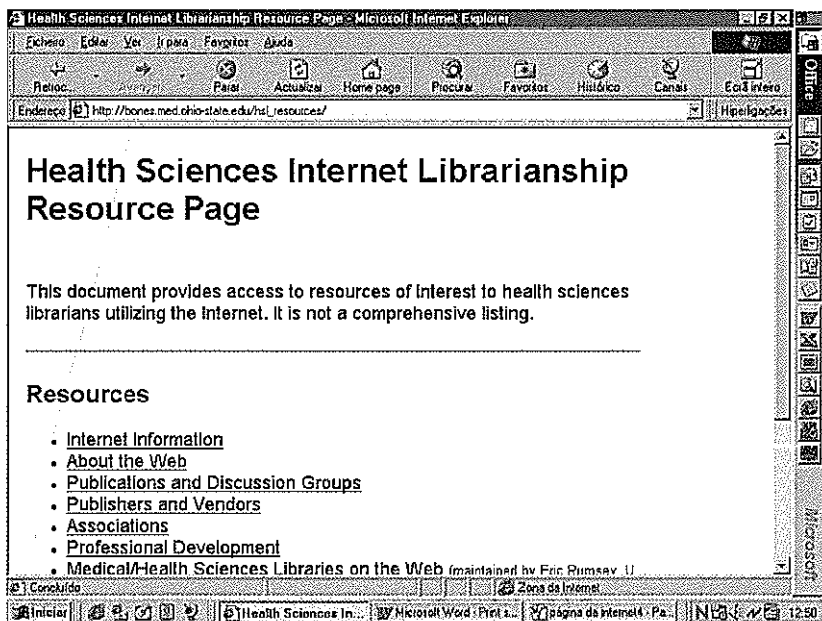


Figura 2 – Exemplo de esclarecimento da finalidade da fonte

A figura 3 é ilustrativa de uma fonte cuja autoridade é inquestionável: Organização Mundial da Saúde (<http://www.who.int> – acesso efectuado a 28.11.2000).

3. Cobertura

É sempre necessário confirmar se a informação diz respeito ao texto integral ou se é um resumo de uma outra fonte. Sendo a original deve ser avaliada a profundidade tomada no tratamento do assunto e se está de acordo com a audiência. Por outro lado, é importante determinar as áreas de assuntos e o tipo de materiais abrangidos bem como a existência de *links* para outras fontes.

4. Exactidão

É fundamental verificar se a informação é factualmente exacta ou se pode ser alvo de tratamento tendencioso. A motivação dos indiví-

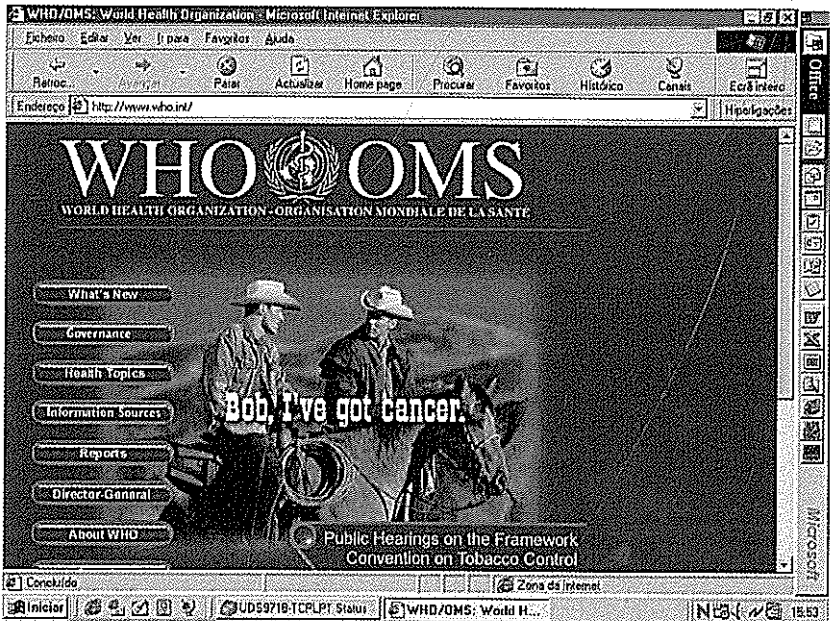


Figura 3 – Exemplo de uma fonte segura relativamente à autoridade

duos ou organizações envolvidos na produção da fonte é uma questão que merece alguma atenção.

5. Actualização

A generalidade das fontes apresenta a data de criação e/ou revisão. É importante saber se é um recurso estático ou se é actualizado com frequência. Havendo existência de *links*, deve-se procurar saber se esses estão realmente acessíveis ou se aparece a frustrante mensagem «*File not found*». Pese embora o dinamismo de actualização de certas páginas invalide os *links* criados noutras recentemente, é verdade que o aparecimento desta expressão pode pôr em causa a veracidade da data de revisão da fonte encontrada.

Na figura 4 pode identificar-se uma fonte com menção da data de criação, da data de revisão parcial e, ainda, a data de registo de direitos de autor (Librarian and Information Science, Libranship. Finding Resources – <http://www.lib.lu.se/netlab/documents/lisres.html> – acesso efectuado em 28.11.2000).

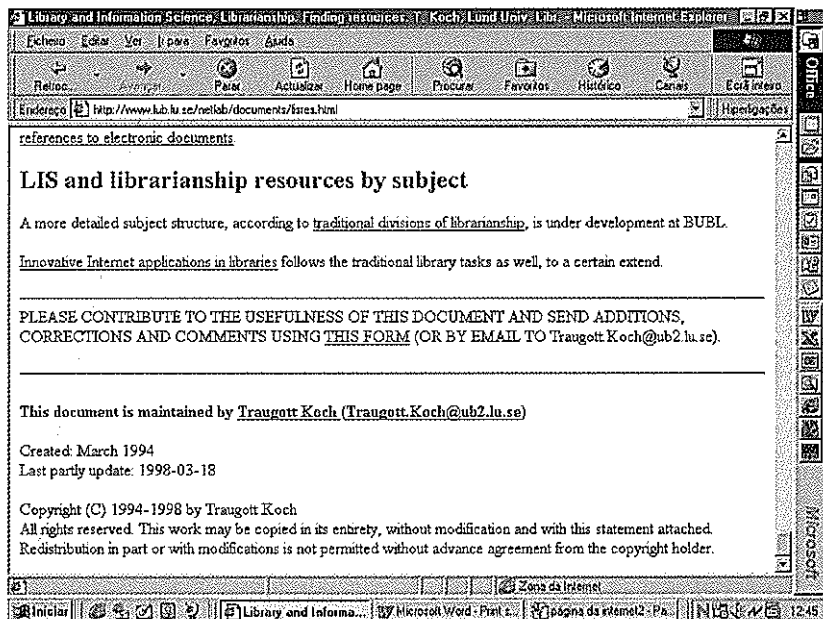


Figura 4 – Exemplo de apresentação das datas de criação e revisão

6. *Comparação com outras fontes*

Numa procura de informação é sempre relevante estabelecer pontos de comparação com outras fontes existentes, a fim de determinar a originalidade da informação encontrada e o grau de cobertura do assunto. É, também, importante verificar a vantagem de aceder à informação via Internet comparando com outros eventuais formatos da mesma fonte.

7. *Qualidade da escrita*

À semelhança da informação noutros suportes, o nível de escrita é um dos aspectos mais relevantes na avaliação. A ausência de erros gramaticais ou tipográficos, a aplicabilidade do uso de calão, a clareza da terminologia e da argumentação constituem aspectos que pontuam positivamente uma fonte de informação.

Avaliação da navegabilidade

1. *Apresentação da informação*

A apresentação da informação tem importância não só pelo lado estético, mas também, pelo facto de transmitir a organização do conteúdo e, até, a cobertura. A existência de um mapa do *site*, de um sumário e, paralelamente, a possibilidade de utilizar uma ferramenta de pesquisa valorizam a fonte de informação, dado que permitem, de várias formas, o acesso ao conteúdo.

Muito embora a aquisição de bons equipamentos se tenha generalizado, é sempre importante verificar a existência de uma versão só de texto, pois nem todos os equipamentos recuperam a totalidade de gráficos ou imagens muitas vezes com animação.

Por outro lado, a utilização de anúncios deverá ser apropriada quer ao tema quer à disposição da informação, uma vez que o seu efeito pode ser o de distrair a atenção do que é fundamental.

Outro aspecto a apreciar diz respeito à presença de informação normalmente intitulada de «*Frequently asked questions*» ou «Perguntas Frequentes». É uma forma de, rapidamente, se encontrar muita da informação que necessitamos evitando a dispersão.

Por último, é necessário verificar a existência de uma característica que pode ser determinante na compreensão do conteúdo: a presença de uma versão traduzida para uma das línguas mais faladas mundialmente.

A figuras 5, 6, 7 e 8 ilustram algumas formas de apresentação da informação. As figuras 5 e 6 mostram exemplos de páginas com a resposta a «Perguntas frequentes» (Comissão de Acesso aos Documentos Administrativos – <http://www.cada.pt> – acesso efectuado em 28.11.2000 e National Library of Medicine – <http://www.nlm.nih.gov/services/faq.html> – acesso efectuado em 28.11.2000, respectivamente). A figura 7 ilustra um exemplo de arrumação chamado «Mapa do site» (Ministério da Cultura – <http://www.min-cultura.pt/mapa.html> – acesso efectuado em 28.11.2000). Uma possibilidade de chamar à atenção da existência de uma versão traduzida pode ver-se na figura 8, a qual mostra uma página com a bandeira inglesa dando indicação da tradução nessa língua (Nederlands Institut voor Wetenschappelijke Informatiediensten ou Netherlands Institute for Scientific Information services – <http://www.niwi.knaw.nl> – acesso efectuado em 28.11.2000).

2. Utilização

É importante determinar se a fonte é, em geral, funcional, amigável e intuitiva e se é compreensível a mobilidade entre as diversas

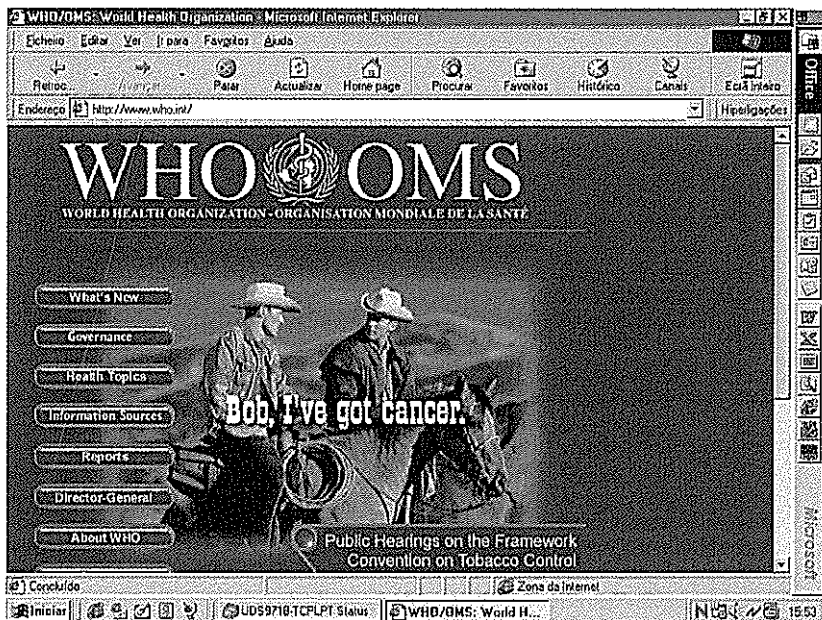


Figura 5 – Exemplo de uma página com «Perguntas frequentes»

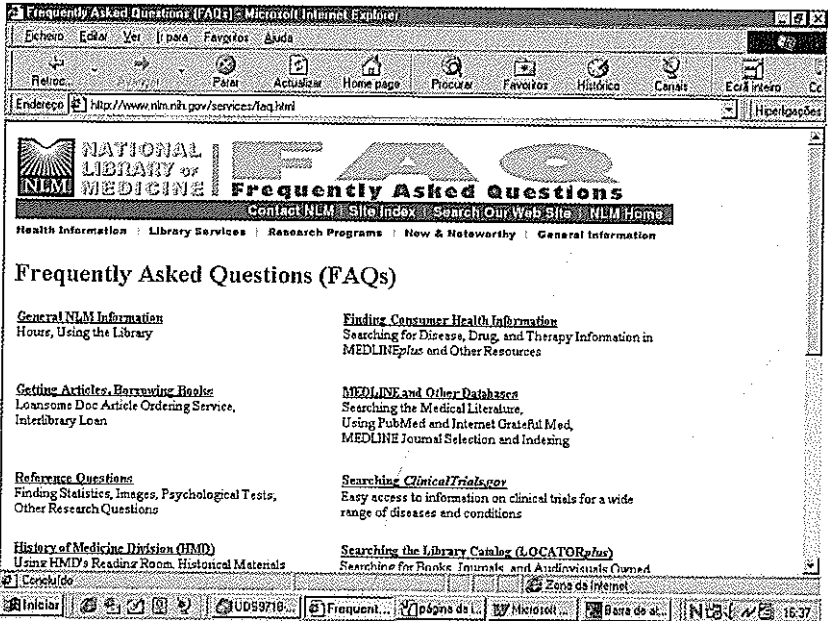


Figura 6 – Exemplo de uma página com «Frequently asked questions»

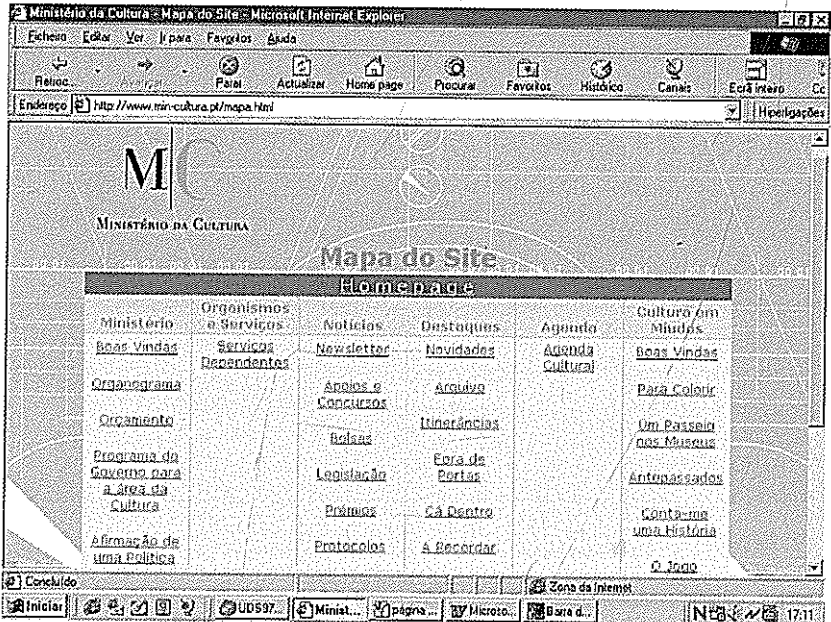


Figura 7 – Exemplo de um «Mapa do site»

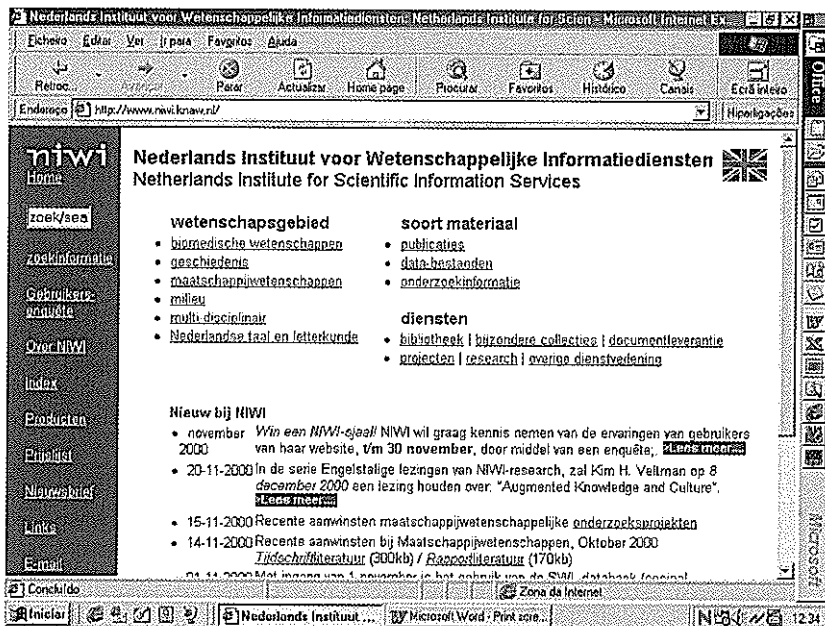


Figura 8 – Exemplo de identificação de existência de tradução através da bandeira inglesa

páginas constituintes da fonte. Mensagens como «voltar ao topo» ou «regresso à página principal» ajudam a deslocar dentro das páginas.

3. Acessibilidade

Um aspecto que, de imediato, valoriza uma fonte de informação na Internet diz respeito à sua localização. Frequentemente, aparece uma mensagem indicativa de que não há informação no endereço procurado ou de que há uma baixa no servidor onde a fonte está localizada. A disponibilidade não permanente causa desânimo quer num primeiro contacto quer numa posterior tentativa de recuperar informação mais actualizada.

O uso de imagens de elevada capacidade contribui, por outro lado, para a lentidão no acesso, pelo que a demora na abertura de uma página chega, em muitos casos, a invalidar o acesso à totalidade da informação disponível.

Paralelamente, importa verificar se a fonte é acessível na sua totalidade ou se exige registo e, até, custos adicionais para os utilizadores.

BIBLIOGRAFIA

BRANDT, D.

«Evaluating information on the internet». *Computers in Libraries*. ISSN 1041-7915. 16: 5 (1996) 44-46.

CHOWDHURY, G.

«The internet and information retrieval research: a brief review». *Journal of Documentation*. ISSN 0022-0418.55: 2 (1999) 209-225.

DAWSON, A.

The internet for library and information service professionals. 2nd ed. London: ASLIB, 1997. ISBN 0-85142-400-7.

HAWKINS, D.

«What is credible information?» *Online*. ISSN 0146-5422. 23: 5 (1999) 86-89.

HOCK, R.

The extreme searcher's guide to web search engines: a handbook for the serious searcher. New Jersey: CyberAge Books, 1999. ISBN 0-910965-26-9.

TSENG, G.; POULTER, A.; HIOM, D.

The library and information professional's guide to the internet. 2nd ed. London: Library Association Publishing, 1997. ISBN 1-85604-221-9.

Metadata

algumas codificações

LICÍNIA SANTOS

RESUMO Fala-se muito de Metadata, aparentemente como sendo uma nova questão, resultante da aplicação das novas tecnologias às bibliotecas, nomeadamente com o aparecimento das chamadas bibliotecas digitais e/ou bibliotecas digitalizadas. Pretende-se com este artigo desmistificar o fenómeno Metadata como realidade recente, fazendo um paralelismo entre a sua utilização pelas bibliotecas tradicionais e a sua utilidade nas bibliotecas digitais. Refere-se ainda o Dublin Core como estrutura *standard* de Metadata, para identificar documentos em suporte digital e/ou na Internet.

ABSTRACT We hear a lot about Metadata as being a new issue in result of the application of new technologies to libraries, namely to digital libraries. The aim of the current paper is to dismount the Metadata issue taken as a new issue, making a parallelism between its application by traditional libraries and its usefulness in digital libraries. The Dublin Core is also referred as a Metadata standard structure to identify documents either available on digital media or in Internet.

Introdução

Fala-se muito de Metadata, aparentemente como sendo uma nova questão, resultante da aplicação das novas tecnologias às bibliotecas, nomeadamente com o aparecimento das chamadas bibliotecas digitais e/ou bibliotecas digitalizadas. Pretende-se com este artigo desmistificar o fenómeno Metadata como realidade recente, fazendo um paralelismo entre a sua utilização pelas bibliotecas tradicionais e a sua utilidade nas bibliotecas digitais. Refere-se ainda o Dublin Core como estrutura *standard* de Metadata, para identificar documentos em suporte digital e/ou na Internet.

1. A utilização de Metadata nas bibliotecas tradicionais

Em sentido mais lato, Metadata significa dados sobre os dados, isto é, um conjunto de atributos ou elementos necessários para identificar cada documento de uma determinada colecção ou fundo de uma biblioteca. Também num sentido mais abrangente, a função primeira de qualquer biblioteca é organizar a sua colecção e estruturar os dados de forma a permitir o acesso aos documentos por parte dos utilizadores. Nesta medida, podemos dizer que os catálogos tradicionais já encerram um conjunto de Metadata ou Meta-informação importante sobre os documentos. Neste tipo de Metadata incluem-se os dados resultantes da catalogação do documento segundo as ISBD's e RPC, no caso português. Representa-se pela catalogação quando se descreve um título, um complemento de título, um autor, um local de publicação, um editor, uma data de publicação, um assunto ou outros tantos elementos descritivos, que identifiquem um dado documento.

Com a transferência dos catálogos manuais para sistemas automatizados esta questão da Metadata, está cada vez mais implícita, quer pela necessidade de normalização necessária para poder disponibilizar informaticamente a informação, quer pelo facto de se poder trocar essa mesma informação entre as várias entidades que a produzem. Em Portugal, as questões normativas, no que respeita à descrição catalográfica dos documentos em sistemas automatizados passam pela utilização do UNIMARC. Assim, podemos dizer que, também o UNIMARC é um sistema de codificação de Metadata, tal como todos os restantes formatos MARC ou outros formatos para sistemas automatizados. Relativamente a esta questão e, já que estamos a referir a automatização, apenas temos de ter em atenção que a normalização é importante, sob pena de não comunicarmos uns com os outros.

A ligação de Metadata a um determinado documento pode fazer-se de duas formas distintas:

- Elementos ou atributos existentes num registo separado do item bibliográfico;
- Elementos ou atributos integrados no próprio documento.

Embora possa parecer que a primeira situação diz respeito à forma de organização da informação nas bibliotecas tradicionais e que a segunda situação diz respeito à organização do conhecimento nas bibliotecas digitais, tal não é verdade. De facto, nas bibliotecas tradicionais já se convive com estes dois tipos de organização da informação, senão vejamos: quer nos catálogos manuais, quer nos catálo-

gos automatizados, temos uma situação de organização da informação em que os Metadados estão dissociados do item. Mas, também temos situações em que essa Meta-informação se encontra integrada no próprio documento, como é exemplo da catalogação CIP (Cataloging In Publication), da ficha técnica que por vezes se apresenta nas publicações, dos elementos que se encontram na página de rosto, muitas vezes com uma disposição própria que melhor permite a sua identificação.

2. A Metadata nas bibliotecas digitais

No que respeita às bibliotecas digitais é comum conceber que este tipo de bibliotecas tem como característica principal, a gestão das suas colecções apenas em formato digital ou que, o acesso à informação se faz essencialmente recorrendo à tecnologia digital. O acesso neste tipo de bibliotecas, embora assente na estrutura de Metadata traz, uma mais valia significativa no que respeita ao acesso imediato ao documento e a possíveis conexões com outros documentos similares ou correspondentes ao assunto seleccionado. Não há, portanto, dúvidas que o acesso imediato ao texto integral e a vários documentos em simultâneo, marca a grande diferença relativamente às bibliotecas tradicionais.

Este conceito surge, intimamente, associado à Internet e, intrinsecamente ligado à «explosão» da produção de documentos electrónicos e ao aparecimento de bibliotecas digitais. Neste contexto, a Metadata é muito importante para identificar, descrever os documentos e criar a sua identidade de forma a evitar duplicações e tornar único cada documento. Estes atributos ou elementos, normalmente, aparecem integrados no próprio documento, numa estrutura própria que serve os princípios da disponibilização da informação na Internet, contexto mais comum em que se inserem as bibliotecas digitais. Desde o início da produção deste tipo de documentos que, os seus próprios autores/produtores, identificam os seus documentos com informações respeitantes à identificação do documento (título, nome do autor, data de criação, etc.). Foi este tipo de identificação, isto é, Metadata que, permitiu até então a pesquisa na Web por parte dos motores de busca. No entanto, a recuperação da informação tem sido dificultada pela forma como se apresenta a Metadata, dada a falta de uniformização na apresentação dos dados. Começamos por um conjunto de dados integrados na linguagem HTML que orienta o modo como a informação é disponibilizada na Internet, por vezes de difícil identi-

ficação por parte do bibliotecário que quer representar a informação, ou por parte do utilizador que pretende identificar o documento que tem em mãos.

Numa pesquisa de um qualquer assunto na Internet, a pertinência dos resultados faz-se, muitas vezes, tendo em consideração a entidade que produz a informação, a data de produção, data de actualização, a relação com outros documentos congéneres, etc. pelo que, a não identificação rápida deste tipo de Metadata traz alguns problemas na recuperação dos documentos. Uma outra desvantagem, da integração da Metadata na linguagem HTML (HyperText Markup Language) tem a ver com a necessidade de se entrar na programação sempre que se verifiquem alterações ao conteúdo do documento, a fim de as poder representar. Esta situação está a ser ultrapassada com a utilização do XML (eXtensible Markup Language) que, há quem designe por linguagem do futuro, por possibilitar a codificação destes elementos, visíveis no conteúdo do documento. Apesar de tudo isto, sente-se ainda uma falta de organização no conjunto dos dados a codificar ou que identifiquem o documento, daí o surgimento do Dublin Core.

3. Estrutura de Metadata de acordo com as regras do Dublin Core

O Dublin Core não é mais do que um *standard* desenvolvido no seio da OCLC (Dublin Core Metadata Initiative), com tendência a ser adoptado pelo resto do mundo, a fim de designar uma semântica comum na descrição e identificação da informação em formato digital, ganhando importância à medida que se vai desenvolvendo a Internet. Reconhece-se que a consistência na criação de Metadata é «a chave do sucesso» para uma boa recuperação da informação.

O aparecimento desta nova forma de codificar dados que, já se codificavam em formatos automatizados, surge devido à proliferação dos formatos MARC, e à adopção de regras diferentes de catalogação nos diferentes países, mesmo usando um formato igual e, também, pela dificuldade de integrar os diferentes formatos MARC na linguagem SGML (Standard Generalized Markup Language).

É uma estrutura que assenta em três objectivos bem delimitados:

- Simplicidade na criação e na manutenção dos dados (pretende-se que o produtor possa preencher os vários elementos que compõem esta estrutura, de forma simples, sem necessidade de

quaisquer conhecimentos técnico-documentais, daí o não se ter escolhido nenhum *standard* já existente);

- Semântica comum (pretende-se seguir uma mesma terminologia para representar conceitos idênticos, de forma a orientar o utilizador leigo na obtenção de resultados mais específicos);
- Âmbito internacional (inicialmente foi desenvolvido em inglês mas, já foram criadas versões em muitas outras línguas).

O Dublin Core Metadata é uma estrutura *standard* que organiza os elementos descritivos que identificam um documento de acordo com uma estrutura de 15 elementos. Esta estrutura é, ainda alvo de discussão concertada a nível internacional, e coordenada por um grupo interdisciplinar de profissionais que tem por objectivo definir o *standard* tendo em atenção uma semântica única e, não esquecendo que se destina a identificar documentos World Wide Web. Cada elemento é opcional e repetível.

3.1 Estrutura do Dublin Core

Título – nome atribuído ao documento, nome pelo qual o documento é formalmente conhecido.

Produtor – entidade primária responsável pelo conteúdo do documento (pode ser uma pessoa física, uma organização ou um serviço).

Assunto – termo de indexação que designa o conteúdo do documento (normalmente pode ser expresso por palavras-chave, frases ou através de uma classificação categorial).

Descrição – *Abstract*, índice, referência a representações gráficas, etc.

Editor – entidade responsável por disponibilizar o documento.

Colaborador – entidade responsável por prestar contribuição ao conteúdo do documento.

Data – data que especifica a vida do documento, pode ser a data de criação, de validação, data de disponibilização e/ou modificação. Deve ter um formato *standard* de oito dígitos (YYYYMMDD), conforme o recomendado pela ISO 8601.

Tipo – especifica a natureza dos dados. Normalmente inclui termos que descrevem funções ou agregações relativas ao conteúdo do documento. Também deve ser descrita aqui a forma como vão ser apresentados os dados, nomeadamente, no que respeita ao formato (de acordo com as normas emanadas por DCMI (Dublin Core Metadata Initiative) *Type Vocabulary*).

Formato – identifica o formato físico ou digital do documento. Normalmente, deve indicar-se a extensão (dimensão) do ficheiro e o tipo dos dados. Pode ainda usar-se para especificar o tipo de *software*, *hardware* ou qualquer outro equipamento necessário para visualizar o documento.

Identificador – indica-se uma referência que identifique de forma inequívoca o documento. Identificações atribuídas por diferentes sistemas, conforme o tipo de documento, como por exemplo: URI (Uniform Resource Identifier), URL (Uniform Resource Locator), DOI (Digital Object Identifier), ISBN (International Standard Book Number), etc.

Fonte – referência à fonte do documento.

Língua – língua do texto do documento (deve ser usada uma codificação, de acordo com o definido na ISO 639)

Relação – referência a documentos relacionados. Recomenda-se que a referência seja feita de acordo com os identificadores dos documentos referenciados, como por exemplo URL. Podem ser também indicações do tipo: versão de, substituído por, parte de, referência de, etc.

Cobertura – localização espacial (cidade, país ou quaisquer coordenadas geográficas), período temporal (data) ou jurisdição (entidade administrativa). Recomenda-se a utilização de vocabulário controlado (Por exemplo *Thesaurus of Geographic Names* (TGN) ou as normas referenciadas na ISO 3166)

Direitos – informação acerca de direitos de autor (*copyright*)

3.2 *A correlação do Dublin Core com outras codificações de Metadata*

Analisando os 15 elementos atrás definidos, podemos verificar que não são muito diferentes, na sua essência, daqueles que já usamos para identificar os documentos que descrevemos e dos quais automatizamos as suas referências bibliográficas. O objectivo do Dublin Core é criar uma estrutura organizada da informação de modo a permitir aos motores de busca uma maior sistematização e ao utilizador, uma pesquisa mais fiável. É importante estruturar a informação de modo a que se, se pretender pesquisar artigos sobre braços (*arms*), não nos venha como resultado a obra técnico-científica de William Arms, um especialista na área das bibliotecas digitais. Utilizando-se esta estrutura de Metadata proposta, os motores de busca podem não ter em linha de conta as palavras do atributo Autor, em pesquisas

como a referida anteriormente, fazendo apenas referência às palavras no autor se o utilizador assim o indicar.

No mundo das bibliotecas digitais, a orientação do utilizador é um problema fulcral, mais do que nas bibliotecas tradicionais, uma vez que o contacto com o bibliotecário se perde, deixa de existir a interacção que possibilitava, muitas vezes, novas orientações no trabalho desenvolvido pelo bibliotecário. Assim, como em qualquer trabalho que se desenvolva, é necessário avaliar constantemente os resultados alcançados de modo a proceder aos ajustes necessários.

Relativamente ao Dublin Core, há quem defenda uma maior semelhança no preenchimento destes elementos de Metadata e aqueles que se utilizam numa qualquer codificação do tipo MARC. Tendo em atenção que o Dublin Core foi criado para ser preenchido pelo produtor do documento, há também opiniões que defendem que um preenchimento mais minucioso não é funcional.

A título conclusivo, é importante não esquecermos como a normalização é definitiva, tanto mais quando se trata de informatização de catálogos, de bibliotecas digitais e da disponibilização de informação na Internet. Sendo assim, e, não sendo esta matéria uma completa novidade como ficou demonstrado, aqui fica um desafio – porque não aproveitar os mesmos dados, para as diferentes codificações ou estruturas de Metadata, utilizando-os conforme as necessidades e evitando ao máximo as duplicações de trabalho.

BIBLIOGRAFIA

BECKETT, Dave; MILLER, Eric; BRICKLEY, Dan

Using Dublin Core in XML <http://purl.org/dc/documents/wd/demes-xml-20000714.htm>

Dublin Core Metadata Initiative : Making it easier to find information

<http://dublincore.org>

Dublin Core Qualifiers

<http://purl.org/dc/documents/rec/demes-qualifiers-20000711.htm>

FOX, Edward A.; SORNIL, Ohm

Digital Libraries. Draft prepared to appear as *Chapter 11 in Modern Information Retrieval, AWL England, 1999*

HILLMANN, Diane

Using Dublin Core

<http://purl.org/dc/documents/wd/usageguide-20000716.htm>



Authenticity in a Digital Environment*

ABBY SMITH

Introduction

What is an authentic digital object? On January 24, 2000, the Council on Library and Information Resources (CLIR) convened a group of experts from different domains of the information resources community to address this question. To prepare for a fruitful discussion, we asked five individuals to write position papers that identify the attributes that define authentic digital data over time. These papers, together with a brief reflection on the major outcomes of the workshop, are presented here.

Our goal for this project was modest: to begin a discussion among different communities that have a stake in the authenticity of digital information. Less modestly, we also hoped to create a common understanding of key concepts surrounding authenticity and of the terms various communities use to articulate them.

«Authenticity» in recorded information connotes precise, yet disparate, things in different contexts and communities. It can mean being original but also being faithful to an original; it can mean uncorrupted but also of clear and known provenance, «corrupt» or not. The word has specific meaning to an archivist and equally specific but different meaning to a rare book librarian, just as there are dif-

* Copyright The Council on Library and Information Resources, Washington, DC (May 2000). Republished with permission

Neste número publicam-se duas das comunicações apresentadas ao Workshop *Authenticity in a Digital Environment* sendo as restantes comunicações publicadas nas *PÁGINAS* n.º 7

ferent criteria for assessing authenticity for published and unpublished materials. In each context, however, the concept of authenticity has profound implications for the task of cataloging and describing an item. It has equally profound ramifications for preservation by setting the parameters of what is preserved and, consequently, by what technique or series of techniques.

Behind any definition of authenticity lie assumptions about the meaning and significance of content, fixity, consistency of reference, provenance, and context. The complexities of these concepts and their consequences for digital objects were explored in *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information*, published by the Commission on Preservation and Access in 1996. There is no universally agreed-upon mandate about what must be preserved and for what purpose. For example, an archivist will emphasize the specifications of a record that bears evidence; a librarian will focus on the content, knowing that it could serve multiple purposes over time. That being the case, there may be many ways to describe an item being preserved and what aspects of that item must be documented to ensure its authenticity and its ability to serve its intended use over time. For certain purposes, some argue, migration may suit the preservation needs of a digital object. For those objects most valued as executable programs, others argue, emulation is preferable. Beyond the technical options undergirding metadata and preservation decisions, numerous nontechnical questions beg to be asked. The issue of authenticity must be resolved before humanists and scientists can feel confident in creating and relying upon digital information.

Creating a common understanding about the multiple meanings and significance of authenticity is critical in the digital environment, in which information resources exist in many formats yet are interactive. From peer-reviewed journal articles to unpublished e-mail correspondence, these resources are integrated; they can interact and be modified in a networked environment. We wanted to know whether the distinctions that have proved to be helpful heuristic devices in the analog world, such as edition or version, document or record, could help us define a discrete piece of digital information. Can we define the distinct attributes of an information resource that would set the parameters for preservation and mandate specific metadata elements, among other important criteria?

We charged the five writers—an archivist, a digital library expert, a documentary editor and special collections librarian, an expert on document theory, and a computer scientist—to address one essential

question: What is an authentic digital object and what are the core attributes that, if missing, would render the object something other than what it purports to be? We asked each to address this question from the perspective he found most congenial. We emphasized our interest in the essential elements that define a digital object and guarantee its integrity, but left the writers free to grapple with that question as they saw fit.

In considering this central issue, we asked that they think about the following:

- If all information-textual, numeric, audio, and visual-exists as a bit stream, what does that imply for the concept of format and its role as an attribute essential to the object?
- Does the concept of an original have meaning in the digital environment?
- What role does provenance play in establishing the authenticity of a digital object?
- What implications for authenticity, if any, are there in the fact that digital objects are contingent on software, hardware, network, and other dependencies?

These are some of the issues that we anticipated would arise in the course of the workshop.

In thinking of which communities to include in the workshop discussion, CLIR sought expertise from the major stakeholders in these issues: librarians, archivists, publishers, document historians, technologists, humanists, and social scientists. Because so many concepts of authenticity derive directly from experience with analog information, we called upon experts in the traditional technologies, such as printing and film, to elucidate key concepts and techniques for defining and securing authenticity of information bound to a physical medium.

The authors were given time to revise their papers in light of the discussion and any comments they received from the participants. Some chose to revise their papers, and others did not. The task of writing a position paper on this complex subject (a paper that we limited in size but not scope) was quite difficult. Each writer took a different approach to the subject, and the papers differ greatly one from another. This seeming disparity proved a boon to the discussions. During that time, each writer had a chance to «unpack» the various nuances of thought that the papers held in short form only, and participants were confronted with the diverse ways that such com-

mon words as *copy*, *original*, *reliable*, or *object* are used. Much of the substance of the discussion is included in the concluding essay.

As one participant remarked, authenticity is a subject we have avoided talking about, primarily because the issues it raises appear so intractable. We are deeply grateful to Messrs. Cullen, Hirtle, Levy, Lynch, and Rothenberg for agreeing to form the advance party as we ventured into *terra incognita*. They were willing not only to think deeply about a vexing issue but also to commit their thoughts to writing and to careful scrutiny by others. Their papers, together with the oral summaries they delivered at the meeting, marked out several different trails to follow, each of which opened onto ever-larger vistas—some breathtaking, some daunting. We are also grateful to the participants, many of whom came from very distant places. Their thoughtful preparation and frank discussion confirmed our sense that authenticity is important for many communities, and that they are ready to engage the issue.

ABBY SMITH, Director of Programs, Council on Library and Information Resources

Authentication of Digital Objects Lessons from a Historian's Research

CHARLES T. CULLEN

The issues stemming from authenticating digital objects are quite similar, and in some cases identical, to those relating to holographs or printed books. Everyone dealing with important material in any form should approach it with a bit of skepticism, but scholars especially need to question what it is they are using. In other words, they need to authenticate all documentation they use in the processes of learning and of creating new scholarship. An authentic object is one whose integrity is intact—one that is and can be proven or accepted to be what its owners say it is. It matters little whether the object is handwritten, printed, or in digital form.

Over time, we have established various measures of authenticity for analog forms that we trust almost without question. Our trust is, however, much greater for printed books than for handwritten objects. In fact, handwritten objects raise many of the same questions of authenticity as digital objects do. The difference is that in the case of the former, the answers may be more easily found. Take Thomas Jefferson's manuscript «Report on the Navigation of the Mississippi», for example. Could he have written it? Is it his handwriting? Is the paper watermarked, and from the appropriate time period? Is the ink contemporary? Do other copies of the manuscript exist? Has its recipient or any other contemporary endorsed it? Is there other internal evidence? Who has described it for us? Has it been identified by a trusted third party?

Is a book authentic? Who published it, and who wrote it? Can they be trusted (are they worthy of one's research time?) Is the rare book what it purports to be? Is the manuscript correspondence actu-

ally by the person to whom it is attributed, and is its date accurate? These questions are now being asked more openly of objects that originate in digital form because we have not yet adopted practices or standards for providing ready answers to them. When objects are presented digitally, deciding what is required to authenticate them may be informed from past practices with non-digital objects.

Two experiences with paper objects inform my views of this subject. The first is a multi-page autograph document that lies in the John Marshall Papers at the Virginia State Library. It is labeled in the hand that wrote the entire piece, «John Marshall's Notes on Evidence in Commonwealth v. Randolph, 1796.» Although the title itself might raise some question about who penned it (How often does an author—even an eighteenth-century author—use his own name in a title of one of his documents?), this document has been used for decades for the source of historical articles and at least one full-length book on the investigation of Richard Randolph for murder. Randolph, a member of the famed Randolph family of Virginia, was related to Marshall and to Thomas Jefferson and to many other members of Virginia's «first families.»

Examination of the writing by those familiar with John Marshall's hand, however, quickly reveals that he did not pen this document. Knowing who did write it is important, but does not help make it more authentic as a Marshall document. The possibility that someone in possession of Marshall's holograph could have copied the document raises new questions, not the least of which assigns significant importance to the value of the original of a document, regardless of its form. Internal evidence, obtained by a close reading of this document, reveals that it might be a partial transcript of a hearing in Cumberland County Court where witnesses are questioned by attorneys, and it has been used by historians as a partial record of Randolph's «trial.» But Marshall's name never appears as one of the questioners, and a knowledge of Virginia law at the time would reveal that whatever was taking place could not be an actual trial, because white men could not be tried for felonies at the county court level during that period. In short, efforts to authenticate this document raise more questions than they answer. At the least, such efforts reveal that the document may not be what many had long thought it to be, and that it may not be even what its title says it is.

This example is somewhat esoteric, to be sure, because it is unlikely that one would use a similar digital document without asking the questions that eventually were asked of the document attributed to Marshall. But the questions asked of it suggest attributes that

must be held by a holograph as well as by a digital object that is to be regarded as authentic. Is it the author's work or a copy? Is it what its title purports it to be? What tests can be applied to answer these questions convincingly?

A second example is more to the point. In the collection of Thomas Jefferson's papers at the Library of Congress is a document that appears to be a list of letters written and received between 1791 and 1793, a period of time during which Jefferson was Secretary of State. An examination of the handwriting reveals that it is most likely Jefferson's, but the list is unlike his other journals of letters sent and received. A close look at the original document suggests that it was written in only one or two sittings (the ink changes only once or twice), rather than over three years. The most significant evidence relating to this document's authenticity lies in the paper itself. Holding it before a light source reveals a watermark that indicates the paper was manufactured in 1804. The document, therefore, could not be an authentic 1791-1793 document.

Almost all these tests can be applied to digital objects, and they need to be. But because digital objects bear less evidence of authorship, provenance, originality, and other commonly accepted attributes than do analog objects, the former are subject to additional suspicion. Tests must be devised and administered to authenticate them.

In many cases, problems of authentication arising from objects that originate as digits are obvious. In trying to find solutions to those problems, however, we must carefully test all suggestions to ensure that they do not themselves open new issues that may be inherent in this medium. The problems of preserving digital objects have received more attention than have questions of authentication (people, I suppose, are less worried about authenticity than about preservation). But why preserve what is not authentic? Might the preservation of a digital object imply an endorsement of authenticity, even if nothing else is done to it? More than one archivist has stated that the only sure means of preserving a digital object is to save a printed copy. Concerns with format codes, migrations from version to version, dependence on hardware-would all be solved by printing a copy (or many copies) and putting it (them) in a safe place. Do that to a digital object before confronting the questions of authenticity, and all that is valuable may be lost. Converting a digital object from one program to another, or migrating it from version to version, could present problems of authenticity that may or may not be solved by careful attention to provenance.

A digital object must be authenticated at the time of its creation by a means that will convey a high degree of confidence to all users, including subsequent use by the originator. Clifford Lynch wrote an interesting and convincing article on the integrity of digital information, published in the December 1994 issue of the *Journal of the American Society of Information Science*. He seems to assume, from traditional experience perhaps, that readers will be responsible for authenticating copies being used on the basis of cataloging data to which they must be alert. Retrieving electronic files by title, for example, might lead one to a revised work, different from the original. The reader must exercise caution, Lynch writes, and be ready to detect signs of alteration. «The expectation should be that violations of integrity cannot be trivially accomplished», he says. Accepting this in the world of printed objects is relatively easy. It is much more difficult in the realm of electronic digital information.

Andy Hopper of Cambridge University suggests an authentication strategy that is worthy of consideration, if not adoption. In his system, the concept of a trusted third party is borrowed from the print world. According to this concept, trusted librarians help authenticate their print holdings through recognized acquisition processes, accepted cataloging procedures, and careful stewardship of their collections, especially those in manuscript form. If a special collection librarian tells us, either directly or by means of a catalog card, that the book in hand is one of two extant copies of Ariosto's *Orlando Furioso* printed on vellum in Venice in 1542, and that it was prepared for the dauphin of France, the library's and the librarian's reputation go a long way toward instilling some degree of confidence that the document is indeed authentic. Moreover, all of this information may be checked. If another librarian delivers to a reader a box of letters cataloged as Ernest Hemingway's, authentication is assumed until internal or physical evidence suggests someone has made a mistake. Knowing that the materials-hard-copy objects-have gone through a process of description and identification, if not authentication, conveys a sense of trust that they are authentic, at least until proved otherwise. Some of the problems of description that help authenticate printed special collection objects have similar, if not identical, examples in the digital world. Take one final example as evidence: in the Newberry Library's special collections is a printed copy of the classic book on rhetoric in Renaissance England, *Arte of Rhetorique* by Thomas Wilson (1525-1581). This particular copy is identified as having belonged to Elizabeth I as part of her royal library, and it is authenticated as such by its original binding, which bears the mark of the royal arms.

Book historians know that until the time of James I, the royal arms were put only on the books within the monarch's own library. (After 1603, King James allowed them to be placed on books bound for other members of court.) Elizabeth's coat of arms on the binding of this copy of Wilson's *Rhetorique* therefore marks it as authentic, as long as external evidence does not dispute it. (If it could be shown that the binding was not sixteenth century, for example, or that it resembled the work of a sixteenth-century forger, the authenticity might be questioned).

Some accepted system of similar assumption of authentication needs to exist in the digital world, but it is more difficult to achieve because digital material is more changeable, accidentally or deliberately. Andy Hopper and others suggest that some means of marking digital objects could help solve many of the problems of authentication. Hopper argues that libraries might serve as authenticators by marking digital objects by some means that would remove doubt as to their characteristics at time of origin. A method must be developed whereby a trusted third party, ideally a trusted librarian, would put a marker on a digital document—a marker that could not be predicted or devised (guessed)—that would mark the document's time and date. The marker might be a number based on sonic rays at various times during the day, a number large enough to prevent guessing (Hopper suggests 100 digits). A professor writing a paper could send the document to the librarian to be marked, and it could then be returned to and held by the author. In the future, the object could be authenticated by its marker, regardless of who held it. Any change in the document would remove the marker. This procedure would be used by librarians who receive digital objects from donors. The marker would ensure that digital objects are as authentic as analog objects at time of cataloging.

Despite its science-fiction flavor, such a method seems to meet accepted tests of authenticity. A trusted third party can claim nothing more about an object, analog or digital, than what can be cataloged, and that information derives largely from physical evidence. Identifying an object in a catalog record or a collection description puts a marker on it that most of us use as the first step in the process of authentication. Is the document what it purports to be or what its owner claims it to be? Scholars often require means to test the cataloger, and the physical attributes of analog objects offer more opportunities to do such testing than do those of digital objects. Handwriting, publishing history, bindings, watermarks, inks, and various forms of internal evidence provide answers to questions of authen-

ticity in analog objects that are lacking in digital objects. Digital objects have attributes that can be used to help with authentication, but none is sufficiently trustworthy or stable to be acceptable unless a workable system of certain marking can be devised.

Certifying that a digital object is the product of its author is difficult when the object originates in electronic form. Without a deliberate and distinctive marking caused by the author that could not be guessed by another or altered by anyone, it seems impossible to authenticate an electronic document beyond doubt. Authors of files or images must take steps to establish authorship of their work; if not, our only option is to accept the assertions of others. Electronic files left behind by someone who has not taken action to establish authorship are subject to suspicion if authorship is asserted by anyone else at the time of «cataloging.» This leaves us where we have been all along—at the mercy of catalogers. But, in the case of a digital object, we are actually worse off than we would be if we were dealing with an analog object. This is because we lack the physical evidence provided by analog objects—evidence that offers the means to test the cataloger. This ability to test both reassures the user and helps keep the cataloger honest. I find no corollary in the digital object realm.

The concern over authenticating digital copies of analog objects is almost as important as that relating to objects that originate in digital form. Scholars are keenly interested in having access to documentary evidence in digital form, and librarians have begun to consider digitization a desirable means of preservation, in spite of the recognized problems inherent in it. Those who hope to use this material, once it has been digitized, must be able to rely on its authenticity, just as they have become accustomed to do in all the forms currently available. Documentary editors, as well as librarians, have new responsibilities as they publish and provide access to their materials in digital form with all the value they have added intact. The work of documentary editors offers some insight into the questions raised over authenticity of digital objects, especially those that derive from analog or holographic objects.

The first task of a documentary editor who is working on an edition of a subject's papers is to locate all the objects that have ever existed as part of the *corpus*, incoming as well as outgoing. This sometimes requires reliance on copies of papers that evidence suggests once existed in original form but which have not been found. Once the collection is organized, each item must be dealt with separately. That is the first stage for authentication tests, starting with the question of whether the item is what it appears to be. All the available phys-

ical attributes assist in answering these questions, but sometimes only internal evidence leads to a final answer (as in the Marshall and Jefferson examples described earlier). The editor is obliged to share these findings with readers and to describe the item in such a way that few, if any, questions remain about the document as object. Not unimportant in this description is all available information about other copies of an item, be they photocopies, carbons, letterpress, polygraph, drafts, or additional holographs. Knowing as much as the editor about all copies is the only sure way for other readers to test the «cataloger's» description, and only by having this information available can a reader have full confidence that all questions of authenticity have been asked. In preparing digital files of historic documents, editors begin their publication by attaching a full document description to a transcript. This is the scholar's seal of authenticity, as it were, or at least as much of a seal as a scholarly editor can provide.

Preparing a digital transcript of a historic object introduces new problems to the issue of authentication. How do we know the transcription is accurate and that it is exactly what the editor prepared originally? The method of providing access to journal articles adopted by JSTOR may offer the best answer for authenticating modern digital transcripts of manuscripts or printed material that originated in analog form. They provide the user with a digital transcription of the text, which is fully searchable and otherwise subject to all the vagaries of digital files. They also provide an image of the original text. If both copies could carry some form of marking that could not be manipulated, the problem of authentication would be solved. This system should work quite well for documentary editors and the readers of their digital publications. Providing an image of the document that is transcribed would be an important improvement over present forms of presentation, because it would permit easy verification of transcriptions. Inaccurate transcriptions are the downfall of documentary editors (as they should be), and mistakes often go undetected. The reader, who may have a high level of confidence in the scholarly work of the editor, is left to assume that the transcription is accurate and authentic. Having a means of testing this assumption would be a great improvement.

Related problems that arise from considerations of authenticity seem to offer little to assist us in answering the primary question. Creating a digital file, and even marking it in such a way that will ensure authenticating it as my own, will mean little if the file itself cannot be read at any point in the future. If the file cannot be read, it cannot be authenticated as mine. (It would be even more mad-

dening if the file could be authenticated but not read.) The same can be said for provenance. If a file can be marked in such a way that its authenticity is assured, issues of its subsequent provenance might not matter in questioning its authenticity. But if a file cannot be read, its provenance will mean little, even if it can be tracked over a long period. Without a marker of authenticity, provenance of a digital object would be of limited use in establishing authenticity. It would help test the cataloger, but the current technology would render uncertain any assertions of authenticity. The instability of software alone would introduce questions that would challenge any claims of authenticity suggested by a trusted provenance.

Paul Conway (1999) says the existence of digital objects moves challenges of preservation from guaranteeing the physical integrity of objects to assuring their intellectual integrity, including their authenticity. He adds that librarians can control this by «authenticating access procedures and documenting successive modifications» to digital files. Authenticating access procedures may affect provenance more than the integrity of the digital object itself, but it would be difficult to guarantee authentication with only this control. It seems that, in this argument, the alteration of an original record is acceptable as long as it is documented. Acceptance of changes with documentation is unreasonable over time and places unnecessary burdens on users. In this case, as in others, preservation without authentication results in a loss of intellectual integrity.

We are not close to having a means of marking digital documents that cannot be challenged – a means that would establish authenticity. Absent such a technique, we are left to consider what other attributes, if any, might approach the establishment of authenticity. Few suggest any high degree of confidence that would come close to what we have for analog materials, but consideration of the problem raises some issues that relate to other concepts that bear on the problem. How confident can one be when an object whose authentication is crucial depends on electricity for its existence? Surely there are higher degrees of confidence in some cases than in others, but something more than provenance or traditional testing methods established for analog objects is needed. I believe it is easier to describe the characteristics of an authentic digital object than to support the authentication beyond a reasonable doubt. My definition is conditional; it depends on an object's capability of being proved to be authentic. Establishing a method of authentication of digital objects that would be unconditional may be possible. At the least, we must agree on some means of testing the authentication of digital objects. The consequences of not doing so are dire.

REFERENCE

CONWAY, Paul.

1999. The Relevance of Preservation in a Digital World. Technical Leaflet, Section 5, Leaflet 5, p. 8, Andover Mass.: Northeast Document Center. Available from <http://www.nedec.org/plam3/tleaf55.htm>.

CHARLES T. CULLEN is president and librarian of the Newberry Library in Chicago. A legal historian, Mr. Cullen earned his Ph. D. degree from the University of Virginia. He has taught at Princeton University, the College of William and Mary, and Averett College. While at the College of William and Mary, he worked on the Papers of John Marshall, and became editor of that project in 1976. At Princeton, he was editor of the Papers of Thomas Jefferson.

A charter member of the Association for Documentary Editing, Mr. Cullen served as its president in 1982-1983. He currently represents that organization on the National Historical Publications and Records Commission. He has served on the New Jersey Historical Commission, and was vice-chairman of the board of the Founding Fathers Papers, Inc., from 1986 until 1992. He is currently vice-president of the board of the Modern Poetry Association, and chairs the board of the Heartland Literary Society.

Mr. Cullen has published or contributed to 30 books and articles. He has lectured widely on subjects relating to the Age of Jefferson, the scholarly use of computers, and the role of humanities research libraries. His scholarly computing expertise was recognized by the Association for Documentary Editing, which in 1987 awarded him its Distinguished Service Award «for outstanding contributions to the field of documentary editing through the use of computers.



Archival Authenticity in a Digital Age

PETER B. HIRTLE

Archival Authenticity: An Example

Downtown Baltimore is a vibrant, dynamic place filled with new office towers and hotels that rise above shops, plazas, and museums. At the heart of Baltimore is the Inner Harbor, an area that is crowded year-round with residents and tourists who are sightseeing, dining, shopping, or watching baseball at nearby Camden Yards. Over the past two decades, the Inner Harbor has become the living center of a revitalized downtown.

The defining feature of Baltimore's Inner Harbor, unlike that of so many American cities, is not a glass structure, a shining space needle, or a distinctive sculpture. The Harbor is marked instead by the sturdy masts and graceful spars of the USS *Constellation*, a historic wooden-hulled naval vessel permanently moored there.

The famous ship arrived in Baltimore in 1955, and for the next 35 years, the city celebrated its frigate, taking pride in the illustrious history of a ship that had been built in Baltimore in 1797 as a sister ship to the equally famous USS *Constitution* anchored in Boston. The story of the *Constellation* took a different turn in 1991, however, with the publication of *Fouled Anchors: The Constellation Question Answered*, a report by Dana Wegner, the chief of ship models at the U.S. Navy's David W. Taylor Research Center in Carderock, Maryland. Rumors had circulated for half a century that the *Constellation* was not what its promoters claimed it to be, and Wegner's report confirmed them. Investigators from the Navy discovered that the supposed Revolutionary War-era frigate in Baltimore Harbor was actu-

ally a Civil War era sloop that had been built in Norfolk, Virginia, in 1854. All it shared with the frigate built in Baltimore in the eighteenth century was its name. It resembled a Revolutionary War-era frigate because during early renovations, some of the ship's admirers had «restored» the *Constellation* to appear to be almost 60 years older than it was; for example, they added a second gun deck and made other alterations. For most of its tenure in Baltimore, the *Constellation* was living a lie (Wegner 1991; LeDuc 1999).

Many themes are at work in the story of the true identity of the *Constellation*. Early citizens of Baltimore, for example, seemed to have a stronger need to connect to the Revolutionary War than to the Civil War. They may have felt that «older is better», and that the ship would be of greatest interest if it was thought to have a Baltimore connection (i.e., if it had been built there). Nonetheless, their distortion of history came at the expense of the *Constellation's* very interesting own history. It was, for example, the last and largest all sail-powered sloop commissioned by the U.S. Navy, and while it did not engage in a famous sea battle, as did its predecessor, it did work to interdict the slave trade during the mid-1800s.

The most interesting themes in the *Constellation* story, however, revolve around the issue of authenticity—not the authenticity of the ship itself, but rather the authenticity of the *documentation* about the ship. For it was not just the appearance of the ship that was «forged», but also the written record concerning the ship.

Some of the changes to the written record may not have been an intentional effort at deceit. Between 1854 and 1908, for example, the annual reports of the Navy listed the ship as having been built in Norfolk in 1854; however, from 1909 onward, the reports listed the ship as having been built in Baltimore in 1797. Was this an intentional effort to deceive or an honest effort to correct what naval officers may have thought was a past mistake? Wegner could not determine the answer.

In the 1950s however, documents began to appear that Federal Bureau of Investigation (FBI) investigators later determined were forged. One document, allegedly written in 1918, was found to have been written with a typewriter made after 1946. Some of the forged documents in the possession of researchers bore forged stamps indicating that they were copies of records found in the National Archives. Other forged documents were inserted into historical files at the National Archives and at the Franklin Roosevelt Presidential Library, where they were subsequently «found» by researchers.

The need to alter the archival written record to conform to a particular historical interpretation speaks to the power of archives to

authenticate. At rest in Baltimore Harbor was a physical artifact, a wooden ship, measuring over 180 feet long and weighing several hundred tons. The existence of the artifact per se, however, was not enough to establish its authenticity. To confirm beyond doubt the nature and history of the *Constellation*, both supporters and critics of the «*Constellation* as frigate» theory turned to a few sheets of paper housed in a few archives.

What characteristics of traditional analog archives give them the power to authenticate? And how can this power be maintained in the digital world, both for archives and for other cultural heritage repositories in general?

The Nature of Archives

To understand why users turn to and trust information found in analog archives, it is necessary to understand the nature of archives. In the vernacular, the word *archives* has come to mean anything that is old or established, be it collections of old movies (such as the Pacific Film Archive), a journal that publishes what the editors hope will be papers of enduring value (for example *Virchows Archiv*, the official journal of the European Society of Pathology), or even rock-and-roll oldies on cable television (in the V Archives) (Maher 1997). Even information professionals have not been loath to extend the definition of archives beyond that found in the American Library Association (ALA) Glossary or other official lexicons when they speak of «digital archiving», a generic term for the preservation of electronic information.

While archivists often inherit responsibility for old things, a collection of historic documents or artifacts, in and of itself, does not make an archives. A true archives is a contextually based organic body of evidence, not a collection of miscellaneous information. A manual written by Dutch archivists almost a century ago codified existing German and French archival theory and developed a modern basis for archives. According to these authors, archives are «the whole of the written documents, drawings and printed matter, officially received or produced by an administrative body or one of its officials...» (Muller, Feith, and Fruin 1968). This definition has been adopted in one form or another by most of Western society.

Found within this definition are the essential elements that define an archives and are the source of much of its power to authenticate. First, archives consist of documents. For the Dutch, these documents

had to be written or printed; modern archivists extended the definition to include multimedia records, including sound recordings and motion pictures. More recently still, archivists (and the courts) have added electronic records to the definition of documents. A recent court case even argued (unsuccessfully) that «cookies», the small transactional files created by many Web browsers when surfing the Internet, were government records when found on a computer used by a government official; others have argued that voice-mail messages are documents (Welch 1998). In short, archives consist of documents, regardless of their form¹.

The documents constituting a formal archives are further distinguished by the fact that they have to have been officially produced or received by an administrative body. Such documents become records. According to the most recent glossary of archival terms, published by the Society of American Archivists, a record is a «document created or received and maintained by an agency, organization, or individual in pursuance of legal obligations or in the transaction of business» (Bellardo and Bellardo 1992). When someone requests a Social Security card, when a business reports its revenues for tax purposes, or when President Clinton issues a proclamation, documents are created. These documents are records because the agencies or officials involved in each transaction are fulfilling legal obligations as they conduct their business. Similarly, when a faculty committee approves tenure for an assistant professor, or when an organization issues an invitation to a meeting, a record is created.

Note that under this definition, the archivist is not concerned about the value, accuracy, or utility of the content of the record. A document may contain lies, errors, falsehoods, or oversights-but still be evidence of action by an agency. Nor does a record have to be particularly interesting or important, or even something that anyone would ever want to consult again. Pure archival interest in records depends not on their informational content, but on the *evidence* they provide of government or business activity. As the Australian archivist Glenda Acland has noted, the «pivot of archival science is evidence, not information» (Acland 1992).

For a time, the essence of records as evidence slipped from center of the archival vision. Ironically, the challenges inherent in dealing with the most modern of records-electronic records-forced creative

¹ Of course, the question of what constitutes a «document» can be problematic (Buckland 1997).

archivists to reinvestigate basic archival principles. Perhaps the most notable of these individuals is David Bearman, author of many publications on electronic records. His collection of essays on *Electronic Evidence: Strategies for Managing Records in Contemporary Organizations* is particularly noteworthy (Bearman 1994). Similar analysis has been conducted by the Australians Sue McKemmish, Frank Upward (McKemmish and Upward 1993), and Glenda Acland, and by the archival educators Luciana Duranti in Canada (Duranti 1998) and Margaret Hedstrom in the United States (Hedstrom 1995). All these authors have concluded to some extent that one can deal effectively with electronic records only if one returns to the first principles of archival theory, including the importance of records as evidence.

Records as evidence provide internal accountability for an agency and make it possible for the agency to determine what it has done in the past. More important, archives-when they contain records that can serve as evidence-can force leaders and institutions to be accountable for their actions. Government archives that contain evidence of the actions of the government can ensure that the rights of individual citizens are protected². They can also provide evidence of when, where, and why the Navy might build and name a new ship.

Records preserved as evidence may also be interesting because of their informational content. For example, census records retained in an archives because of the evidence they provide about the activity of the Census Bureau, may be of great interest to genealogists. To many archivists, however, the fact that the Census Bureau creates census returns in the course of conducting its legally mandated business-not the information contained in the record-is of paramount importance³.

At the heart of an archives, therefore, are records that are created by an agency or organization in the course of its business and that serve as evidence of the actions of that agency or organization. The agency or organization maintains those records for its business pur-

² These two themes-the ability of archives to hold public officials accountable and to protect the rights of individual citizens-form the basis of the new mission statement of the National Archives and Records Administration, i.e., «to ensure ready access to essential evidence [and note the emphasis on *evidence*] ... that documents the rights of American citizens, [and] the actions of federal officials ...

³ While most archivists would agree with the definition of a record as presented in this paper, there are strong differences about what criteria should be used in the appraisal of records for retention or possible destruction. Some archivists argue that only the evidentiary value of the records should be taken into account, others argue that sociocultural requirements, including the need to establish memory) should be considered (Cook 1997; Cox 1994; Cox 1996).

poses. At the point when the records are no longer of immediate value to the organization, it may elect to transfer its records to an archives. The archives become responsible for maintaining the evidentiary nature of the materials after the records have left the control of the agency that created them.

One way in which archivists working with analog records have sought to ensure the enduring value of archives as evidence is through the maintenance of an unbroken provenance for the records. Archivists need to be able to assert, often in court, that the records in their custody were actually created by the agency specified. Furthermore, the archivist must be able to assert that the records have been in the custody only of the agency or the archives. In an analog environment, the legal and physical transfer of the documents from the agency to the archives ensures an unbroken chain of custody.

Archives truly exist only when there is an unbroken chain of custody from the creating agency to the archives. For a government archives, the transfer of custody is best accomplished as a matter of law. As Margaret Cross Norton, a pioneer theorist of American archives, noted:

We must disabuse ourselves of the concept that the acquisition by the state historical society of a few historical records ... automatically transforms the curator of manuscripts into an archivist ... An archives department is the government agency charged with the duty of planning and supervising the preservation of all those records of the business transactions of its government required by law or other legal implication to be preserved indefinitely (Mitchell 1975).

In a nongovernmental agency, policy can take the place of law if the policy identifies what records of business transactions need to be preserved indefinitely. Either law or policy, however, should govern the transfer of records to an archives.

Why is the authorized transfer of a complete set of records to an archives with an unbroken chain of custody important? First, it helps maintain the evidentiary value of the records. An archivist can be called upon to testify in court about the nature of the records in his or her custody. That archivist would not be expected to testify as to the accuracy of the contents of the records. However, he or she should be able to assert that on the day when the records left the custody of the originating agency or organization, a particular document was included as part of the records.

Equally important as unbroken custody in establishing the integrity of records is the completeness of the documents. Only records that are complete can ensure accountability and protect personal rights.

As soon as records become incomplete, their authority is called into question. For example, when information is missing in a record, we do not know if it is because the information was never created or because it has been discarded. Individual records must be complete; they must contain all the information they had when they were created. They must also maintain their original structure and context.

In addition to each individual record being complete, it is also necessary that the record series in which the record is created be complete. Because records gain meaning from their context, it is important to know the nature of other records. Take the example of a case file. A case file is a record relating to one person as he or she interacts with a government agency. It might be an application for food stamps, an assessment of eligibility for veterans' benefits, or a request for a reproduction of a photograph in an archives. By itself, a case file can tell the user a great deal, but it does not reveal whether the individual in question was treated differently from other people in the same situation. To understand a single record in context, one needs the whole series. There may be references from the case file to other records in the same series. Whenever possible, therefore, archivists seek to preserve entire series.

This does not mean that archivists never throw anything away. The normal archival principle is to save only 2 to 4 percent of an organization's records. What archivists try to avoid, however, is assessing individual records or parts of records. One either keeps the entire record or discards the entire record. Similarly, the normal presumption is that one either keeps or discards an entire series of similar records (though there may be times when the bulk of the records makes this impossible).

Hilary Jenkinson, a leading archival theoretician, neatly summed up the importance of both the legal basis for the transfer of records to an archives and the need for completeness within the record series and the individual records. He noted the importance of authenticity to archives and defined it as the principle that archives are «preserved in official custody ... and free from suspicion of having been tampered with» (Jenkinson 1965). According to Jenkinson, the archivist's primary task is «to hand on the documents as nearly as possible in the state in which he received them, without adding or taking away, physically or morally, anything: to preserve unviolated, without the possibility of suspicion, every element in them, every quality they possessed when they came to him» (Jenkinson 1984).

Archivists have a responsibility to ensure the integrity of the documents even after they are legally transferred to a repository. In an

analog environment, this is done by a number of mechanisms. Users of archives, for example, normally must work under the supervision of an archival staff member. The users are instructed to maintain the order of records as they are found and are cautioned against adding material to or removing it from the file. In some cases, especially when documents are known to be of great economic value, an archival staff member may count the documents delivered to and then returned by a researcher. (Normally, however, the volume of material in an archives works against any sort of item control.)

The example of the *Constellation* illustrates both the promise and the dangers associated with the evidentiary power of traditional archives. Some of the forged documents that seemingly proved that the ship in the Baltimore harbor had been built in 1797 were found among the records of the U.S. Navy located in the National Archives and Records Administration. Transfer of the records presumably took place under the legal authority of the Federal Records Act, and an unbroken chain of custody had been established. Users of the records, therefore, could assume that any documents found in the record series had been created and maintained by the Navy until they were transferred to the National Archives. The National Archives then maintained the records as they were received from the Navy. The powerful presumption must be that documents found in the Navy files in the Archives are an accurate reflection of the Navy's files at the time of the transfer. Regardless of the content of the records, the organizational context alone would be enough to argue for their authenticity.

We now know that in the case of the *Constellation*, it was wrong to presume that all of the documents in the Navy files, as they were found in archives, were authentic. Archivists had sought to preserve the records in the context of the office that had created them and they had accessioned a complete series into the archives. Normally, this would be enough to ensure the authenticity of the records. In this case, however, it was also necessary to turn away from the context of creation of the record and to examine the individual record itself.

When Wegner, assisted by forensic document examiners at the FBI, examined the problematic documents, he found a number of elements within the documents that led him to question their authenticity. Since most of the documents were copies, it was not possible to test inks and papers. On the basis of the typeface on some of the documents, however, the FBI could determine that the documents had been typed on typewriters that did not come into existence until 30 years after the documents had supposedly been created. Other documents were

undated and unsigned, raising questions about their authenticity. In yet another instance, the investigators noticed 14 spelling and typographical errors in a simple document. The investigators knew that the office from which this document supposedly originated had strict requirements for accuracy; the suspect document could not have originated in an office that enforced those requirements.

Without realizing it, the investigators had used one of the oldest archival sciences to test the authenticity of the documents: the science of diplomatics. Diplomatics is a body of concepts and methods, originally developed in the seventeenth and eighteenth centuries, «for the purpose of proving the reliability and authenticity of documents.» Over time it has evolved into «a very sophisticated system of ideas about the nature of records, their genesis and composition, their relationships with the actions and persons connected to them, and with their organizational, social, and legal context» (Duranti and Eastwood 1995, quoted in Duranti and MacNeil 1996). Perhaps because diplomatics emerged from the need to understand and authenticate medieval charters, patents, and other legal documents, American archivists knew little about the field until quite recently. In addition, the primary problem facing American archivists for most of this century has not been to understand individual documents but rather to deal with the flood of documents on paper and in other formats generated by a bureaucratic, paper-intensive society.

Fortunately, in 1989 an Italian archivist teaching in Canada introduced North American archivists to the primary concepts of diplomatics through a series of six articles published in the Canadian journal *Archivaria* (Duranti 1998). In these articles and in her later work on reliability and integrity, Duranti expands on the interrelationship between the form, structure, and authorship of documents. The form of a record and the procedure for its creation, she asserts, determine the reliability of the record. A record is more likely to be reliable when its form is complete than when it is incomplete. While documents can require many elements, the two most commonly required elements of form are the date and an element, usually a signature, that assigns responsibility to a person for the content of the record (Duranti 1995).

Diplomatics also provides a mechanism for evaluating the authenticity of copies. Why is an original more reliable as evidence than a copy? It is because the original has the maximum degree of completeness and a higher degree of control in the procedure of creation of the document. Creating a copy always introduces the possibility for variation or change from the original.

On the other hand, there are times when a copy may be more reliable than an original. For example, a contract for the sale of the house that is copied into the deed books of a village government may be more reliable than the original, because a third, impartial, authority can attest to the agreement of the parties represented in the contract. Archives have a long tradition of producing authentic copies, i.e., copies that have not been subject to manipulation, substitution, or falsification after the completion of the process that created the original record. Such copies often entail a change in format (for example, from paper to microfilm) and require that procedures be in place to ensure the authenticity of the resultant copies. If the latter condition is met, archivists willingly discard the originals.

An archivist could use the principles of diplomatics to judge the reliability and the authenticity of the individual documents in the *Constellation* case. For example, questioned documents that lacked a date or a signature would fail the fundamental test for reliability. The document filled with misspellings and typographical errors would also fail. The form of a document that does not follow the documentary conventions of the creating office is suspect; the document itself may be unreliable.

In summary, traditional archival theory has developed two approaches for ensuring the authenticity of the document. The first approach, the basis for most American archives, seeks to understand and control the context in which records are created. Records that are generated in an agency, transferred by law or policy to an archival agency through an unbroken change of custody, and maintained complete and inviolate by that archival agency are presumed to be authentic. The second approach, as exemplified in the works of Duranti, focuses on the individual record: its form and the circumstances of its creation. Together, these two approaches are used to ensure the authenticity of records in the analog world.

Archival Authenticity in a Digital World

The archival profession has established a theoretical base to justify the assertion of authenticity when dealing with analog records. But will the principles that have worked so well in the analog environment transfer to the new digital world? Wendy Duff has noted, 'As records migrate from a stable paper reality to an intangible electronic existence, their physical attributes, vital for establishing the authenticity and reliability of the evidence they contain, are threatened' (Duff 1996). The ease

with which records in electronic form can be created, transferred, and modified only heightens the importance of maintaining their integrity. The central question facing all archivists, therefore, is how to ensure the authenticity of records in digital form. Can the traditional archival methodologies developed for analog records be used for digital records? Or must new methodologies and techniques be developed to ensure that the archival records remain authentic over time?

A number of important initiatives are under way to explore how the integrity of records can be preserved in a digital environment. None of the strategies has yet become widely accepted, primarily because they have not been tested in the field. As Philip Bantin has concluded, «In short, there are no clear-cut answers available yet, but there are plenty of very good ideas and emerging strategies out there» (Bantin 1999). Two of the more promising approaches can be summarized here.

The University of Pittsburgh Functional Requirements for Evidence in Recordkeeping Project

The University of Pittsburgh conducted one of the first and most extensive research projects that sought to identify the functional requirements for the preservation of electronic evidence. Its project, the «Functional Requirements for Evidence in Recordkeeping», consisted of three main components. First, the project identified the functional requirements for recordkeeping in a variety of communities. The project recognized that groups other than archivists (e.g., the legal, medical, and business communities) also had need for authentic, reliable records. Laws, standards, customs, and the best practices of each community contain the justifications for record keeping. To ensure that electronic records meet the needs of those communities (i.e., that they become what the project identified as «business acceptable communications»), one must identify the requirements for recordkeeping in each community and then establish metadata that meet those requirements. The project did this by establishing the recordkeeping requirements and practices of organizations—the literary warrant (Duff 1996; Bearman 1996).

Using the requirements necessary for literary warrant, the project then produced a general specification of the attributes of evidentiality. The specification consists of 13 properties that are categorized into three groups. The first group requires a *conscientious organization* that complies with legal and administrative requirements for recordkeeping. The second group specifies the requirements for *accountable recordkeeping systems*, including policies, assigned responsibility, and for-

mal methodologies for their management and accurate and complete documentation. The Pittsburgh system presupposes that accountable recordkeeping systems are used at all times in the normal course of business. The third group defines the requirements that relate to the *record itself*, specifically how the record is created or captured, how it is maintained, and what is necessary for the record to be used.

In addition to developing the general specification of the requirements for evidentiality, the Pittsburgh project developed a set of production rules to express formally each functional requirement. David Bearman, a consultant on the project, has turned the production rules and general analysis into a set of metadata requirements. The goal is to be able to create records that are encapsulated metadata objects: content in an envelope of metadata that ensures the authenticity, integrity, reliability, and usability of the content.

Implicit in the Pittsburgh approach is the assumption that «recordness» and «evidentiality» (the elements that determine the trustworthiness of records in business and legal settings) can be maintained in an electronic system only if the requisite functionality is built into the record system from the start. Several efforts have been made to implement the Pittsburgh model, most notably in projects under way at Indiana University, a Swedish pharmaceutical company, and the City of Philadelphia, but there is no consensus whether the Pittsburgh project has identified the true functional requirements for authenticity. Some worry that the Pittsburgh model may be too complex, and hence too costly, to implement. Furthermore, it presupposes radical changes in how documents are generated. For example, if one wishes to write a report, one currently opens a word processing package and begins writing. The Pittsburgh system seems to propose that in the future one would open instead a report-writing module. The module would «know» who you are, what your authority for writing the report is, and in what format you are writing the report. The software would automatically encapsulate each draft of the report with this management information. While highly desirable or even mandatory, to ensure the authenticity of the electronic file, such an approach does not reflect how people currently use software.

University of British Columbia Preservation of the Integrity of Electronic Records and InterPARES Projects

Two projects at the University of British Columbia (UBC) are investigating the integrity of digital information over time. The first proj-

ect, «Preservation of the Integrity of Electronic Records», sought to identify the best methods for preserving the reliability and authenticity of electronic records over time. The UBC analysis determined that generic information systems designed to collect, process, store, and disseminate information lack some of the functionality needed to produce, maintain, and preserve reliable electronic records. For example, most current systems do not adequately relate the content of records to business transactions. They also lack sufficient metadata to monitor the creation and maintenance of records in a way that ensures they will be both reliable and understandable when retrieved in the future. The project concluded that reliability and authenticity of electronic records are best ensured when procedural rules for record-keeping are embedded into the overall records system. This finding is similar to that of the Pittsburgh project, which expressed an interest in building into systems the automatic capture of the metadata it has determined are needed to ensure the recordness of the data (Duranti and MacNeil 1996; Hedstrom 1996).

In other ways, however, the UBC project was fundamentally different from the Pittsburgh project (Duranti and MacNeil 1996; Bantini 1999; Marsden 1997). For example, the analysis of the requirements for recordkeeping in the two projects differed greatly. The Pittsburgh project based its analysis on literary warrant, whereas the UBC project's analysis was based on diplomatics and archival theory.

In part because of the difference in starting points, the two projects reached fundamentally different conclusions in some areas. One of the most striking differences relates to the role of the archives in ensuring authenticity. The Pittsburgh project did not assume that an archives is needed to ensure the preservation and authentication of records. In the Pittsburgh system, it is the metadata, not the custodial agency, that determine the authenticity of records. Records can, and in most cases should, remain in the custody of the agency that created them. As one of the Pittsburgh project members has argued, «Archivists cannot afford-politically, professionally, economically, or culturally-to acquire records except as a last resort ... Indeed, the evidence indicates that acquisition of records and the maintenance of the archives as a repository gets in the way of achieving archival objectives and that this dysfunction will increase dramatically with the spread of electronic communications» (Bearman 1991). The UBC project, in contrast, placed archives at the heart of the authentication system for electronic records, in a fashion similar to the role played by archives in protecting and authenticating paper records. This project concluded that «the routine transfer of records to a neutral

third party, that is, to a competent archival body, invested with the exclusive authority and capacity for the indefinite preservation of inactive records, is an essential requirement for ensuring their authenticity over time» (Duranti and MacNeil 1996).

The «Preservation of the Integrity of Electronic Records» project at UBC sought to establish a theoretical framework based in traditional archival principles for the authentication of digital information. A follow-on project is now seeking to put some of these principles into action. The InterPARES (for «International Research on Permanent Authentic Records in Electronic Systems») project is an international collaboration spearheaded by UBC. Its goal is to use the tools of archival science and diplomatics to develop the theoretical and methodological knowledge essential to the permanent preservation of inactive electronically generated records. It will then formulate model strategies, policies, and standards capable of ensuring the preservation of those records. The InterPARES project has generated great interest in the archival community, in part because it is based on familiar principles and practices. The community eagerly awaits reports of its findings.

Conclusion

It is not possible at this early stage to say whether Pittsburgh or UBC has the better approach for ensuring the authenticity of records. Both approaches need to be tested in the field (Bantin 1999). As Margaret Hedstrom has noted, «What we lack is an evaluation of the usefulness of these findings from the perspective of organizations that are responsible in some way for preserving and providing access to electronic records. We need assessments from the administrators of archival and records management programs about the feasibility of putting the proposed policies, and models into practice. We need reactions from people outside the archival community especially where related research and projects are being conducted» (Hedstrom 1996).

In the interim, however, it is easy to speculate that some combination of the Pittsburgh and UBC approaches will come to dominate. The Pittsburgh project's basis in the actual documentary requirements of different communities is very appealing, and the project's desire to include administrative metadata from the very moment of creation is highly desirable.

On the other hand, it is unlikely that all information of interest to future users of records systems will be found in records creation

management systems fully compliant with the Pittsburgh metadata. Scholars will be willing to access, use, and evaluate the information found in the electronic files, regardless of whether the actual data convey the true quality of «recordness». An archival purist might insist that if information is not stored in a record keeping system, then the information cannot be a record and therefore should not be part of the archival record. In reality, however, our repositories are filled with interesting information that may not meet the formal definition of «record» or may not have been created with a record keeping system in mind.

A good example of how material that is not formally a record can be valuable to the researcher is the famed PROFS case (Bearman 1993). PROFS refers to a proprietary IBM communication system used in the White House under Presidents Ford and Reagan. Because they were system back-up tapes, the PROFS tapes lacked even the rudiments of record keeping functionality. Nevertheless, a consortium of historical groups sued for the release of the tapes. In the absence of controlled records, the information on the back-up tapes was the best the researchers could find. For researchers, the value of the tapes was great because they were still held by the agency and were surprisingly complete. However, even if only selections of the e-mail messages had survived and were located only in nongovernmental repositories, researchers would still try to use them, even though their authenticity was more questionable.

In short, social mechanisms of control promise to be the fundamental basis for the establishment of digital authenticity. It would be desirable if all digital information consisted of true records created in a system that encapsulates with the record the information needed to maintain the evidential value of the records. For most digital information, however, the fact that it is in an archives, an unbiased third party, will have to suffice. As with the paper records used in the *Constellation* example, the fact that digital information is found within a trusted repository may become the base upon which all further assessments of authenticity build.

Even if the physical presence of digital data in a trusted repository is the basis for future assessments of authenticity, archivists will still need to associate with those digital documents metadata that researchers can use to understand and assess digital information. We need self-conscious documentation by the creators and preservers of digital representations that details the methods employed in making and maintaining the representations. We also need to know what researchers need to know about the transformation from analog to digital format, as

well as about any transformations that may occur as digital data are preserved. To determine the latter, we need to understand the «digital literacy» that future researchers will need «to assess digital information, identify known artifacts introduced by particular processes, and correctly identify as yet unknown sources of distortion» (Bearman and Trant 1998). Only by understanding the interactions between researcher and document and records and repositories will we be able to convey into the future the trust mechanisms of the paper world.

REFERENCES

- ACLAND, Glenda
1992. Managing the Record Rather than the Relic. *Archives and Manuscripts* 20(1):57-63.
- BANTIN, Philip
1999. Strategies for Managing Electronic Records: A New Archival Paradigm? An Affirmation of our Archival Traditions? *Archival Issues*.
- BEARMAN, David
1991. An Indefensible Bastion: Archives Repositories in the Electronic Age. In *Archival Management of Electronic Records*, edited by David Bearman. Archives and Museum Informatics Technical Report #13. Pittsburgh: Archives and Museum Informatics.
1993. The Implications of *Armstrong v. the Executive Office of the President* for the Archival Management of Electronic Records. *American Archivist* 56(4):674-89.
1994. *Electronic Evidence: Strategies for Managing Records in Contemporary Organizations*. Pittsburgh: Archives and Museum Informatics.
1996. Virtual Archives. Paper presented at the International Congress of Archives meeting in Beijing, China, September 1996. Available from <http://www.lis.pitt.edu/~nhpre/prog6.html>.
- BEARMAN, David and Jennifer Trant
1998. Authenticity of Digital Resources: Towards a Statement of Requirements in the Research Process. *D-Lib Magazine*. Available from www.dlib.org/dlib/june98/06/bearman.html.
- BELLARDO, Lewis J., and Lynn Lady Bellardo, compilers.
1992. *A Glossary for Archivists, Manuscript Curators, and Records Managers*. Chicago: Society of American Archivists.
- BUCKLAND, Michael
1997. What is a «DocumenC»? *Journal of the American Society for Information Science* 48(9):804-9.

- COOK, Terry
 1997. What Is Past Is Prologue: A History of Archival Ideas Since 1898, and the Future Paradigm Shift. *Archivaria* 43:17-63.
- COX, Richard
 1994. The Record: Is It Evolving? *Records & Retrieval Report* 10(3)A-16.
 1996. The Record in the Information Age: A Progress Report on Research. *Records & Retrieval Report* 12(1):1-16.
- DUFF, Wendy
 1996. Ensuring the Preservation of Reliable Evidence: A Research Project Funded By the NHPRC. *Archivaria* 42:28-45.
- DURANTI, Luciana
 1995. Reliability and Authenticity: The Concepts and Their Implications. *Archivaria* 39:5-10.
 1998. *Diplomatics: New Uses for an Old Science*. Lanham, Md.: Scarecrow Press.
- DURANTI, Luciana and Terry Eastwood
 1995. Protecting Electronic Evidence: A Progress Report. *Archivi & Computer* 5(3)213-50.
- DURANTI, Luciana and Heather MacNeil
 1996. The Protection of the Integrity of Electronic Records: An Overview of the UBC-MAS Research Project. *Archivaria* 42:46-67.
- HEDSTROM, Margaret
 1996. Electronic Records Research Issues: A Summary of Recent Research. Prepared for the Invitational Conference on Electronic Records, June 28 and 29, 1996. Available from <http://www.si.umich.edu/e-recs/Research/NHPRC-Sum.html>.
 1995. Electronic Archives: Integrity and Access in the Network Environment. *American Archivist* 58(3):312-324.
- JENKINSON, Hilary
 1965. *A Manual of Archive Administration*. London: Percy Lund, Humphries & Co. Ltd.
 1984. Reflections of an Archivist. In *A Modern Archives Reader*, edited by Maygene F. Daniels and Timothy Walch. Washington, D.C.: National Archives and Records Service.
- LEDUC, Daniel
 3 July 1999. Shored Up and Shipshape: Refurbished *USS Constellation* Returns to Inner Harbor, Minus a Myth. *The Washington Post*.
- MAHER, William
 Society and Archives. 1997. Incoming presidential address to the Society of American Archivists. Available from <http://www.archivists.org/governance/presidential/maher-l.html>.
- MARSDEN, Paul
 1997. When is the Future? Comparative Notes on the Electronic Record-Keeping Projects of the University of Pittsburgh and the University of British Columbia. *Archivaria* 43:158-73.
- MCKEMMISH, Sue and Frank Upward, eds.
 1993. *Archival Documents: Providing Accountability through Recordkeeping*. Melbourne: Ancora Press.
- MITCHELL, Thornton W
 1975. *Norton on Archives: The Writings of Margaret Cross Norton on Archival and Records Management*. Carbondale, Ill.: Southern Illinois University Press.

MULLER, S., J. A. Feith, and R. Fruin

1968. *Manual for the Arrangement and Description of Archives*, translated by Arthur H. Leavitt. New York: H.W. Wilson.

WEGNER, Dana M.

1991. *Fouled Anchors: The Constellation Question Answered*. Bethesda, Md.: David Taylor Research Center.

WELCH, Matt

15 October 1998. *The Cookie Monster of Putnam Pit*. Salon Magazine. Available from <http://www.salon.com/21st/feature/1998/10/15feature.html>.

PETER B. HIRTLE is co-director of the Cornell Institute for Digital Collections (CIDC). CIDC is responsible for developing digital resources, supporting their use campus-wide, and conducting applied research that advances the production and utility of such resources. Mr. Hirtle is associate editor of *D-Lib Magazine*, a monthly magazine about innovation and research in digital libraries. Prior to his arrival at Cornell, Mr. Hirtle worked at the National Archives and Records Administration (NARA), first for the technology research staff (where he helped complete its most recent digital imaging report), and then as coordinator of electronic public access for the agency. He has also served as curator of modern manuscripts at the National Library of Medicine. Hirtle has a master's degree in History and an M.L.S. with a concentration in archival science. He has served on several of the units sponsored by the Society of American Archivists, most recently as a member of the executive committee of its governing council, and on the Commission on Preservation and Access/ Research Library Group's Task Force on Digital Archiving. He currently serves on the Research Libraries Group/Digital Library Federation Task Force on Long-Term Retention, and is a member of the National Initiative for a Networked Cultural Heritage's Working Group on Best Practices in Networking Cultural Heritage.



LER MUITO PRAZER

Um Mestre de Leitura

Ramin Jahanbegloo – Ontem falámos da sua vida. Hoje gostaria que falássemos do seu trabalho. Como o define? É um filósofo, um crítico literário ou um leitor do nosso mundo?

George Steiner – Gostaria de ser recordado – por pouco que perdue nas memórias – como um mestre de leitura, como alguém que passou a vida a ler com os outros. Para conhecermos bem o acto de leitura, devemos servir-nos das análises muito finas de Charles Péguy que deu na sua obra o exemplo de uma definição em filigrana e densa, intensa e cheia, daquilo que uma leitura bem feita implica. É uma leitura que implica uma responsabilidade, e neste termo contém-se o de resposta. Trata-se portanto de responder a um texto, à presença e à voz de outrem. E isso tornou-se difícil senão impossível numa cultura onde o ruído é constante, que não tem de reserva uma praia de silêncio ou sequer de paciência. Entendo paciência na sua acepção do século XVII, quando a etimologia prevalecia em certas fórmulas dando a «paciência» ou a «sofrer» um sentido que hoje se desvanece. Ler não é sofrer mas, falando com propriedade, estarmos prontos a receber em nossa casa um convidado, ao cair da noite. A imagem que os grandes poetas sugerem, e também emerge em Heidegger ou nos pensadores pré-socráticos, é a de um acolhimento aberto ao pensamento, ao amor e ao desejo dos outros, através da prática da leitura, da audição da música e do conhecimento da arte. Trata-se de aprender com outros a escutar melhor. É por isso que o ensino me foi sempre indispensável quando em várias ocasiões teria podido, materialmente falando, abandoná-lo. Mas na organização da minha existência, procurei-o sempre como um meio de reunir leitores à minha volta, para poder conservar a esperança de que depois da minha morte alguns outros continuem a amar os poetas e os filósofos que tanto amei. A sombra da americanização do mundo inteiro obceca-me. Depois de ter viajado por diversos países, parecia-me ver a América

que é, sem dúvida, uma libertação material do homem, sendo porém, e ao mesmo tempo, a própria adversidade, o silêncio e a solidão. O fenómeno a que me refiro contraria a resistência a um acesso demasiado imediato a um texto, a uma obra de arte ou ao pensamento. Falo de um mundo que nos vai aproximar uns dos outros, como aconteceu com a queda do templo de Jerusalém, quando deixou de existir ensino formal, tendo ele sido substituído por casas e mestres de leitura que liam e tentavam aprender a ler e a reler com alguns outros companheiros e confrades – não falemos de alunos, recuso aqui esse termo. Mas que quer tudo isto dizer? Uma leitura ideal pode ser a de um parágrafo de Montaigne, de Pascal ou de Kierkegaard, um poema de René Char, um soneto de Shakespeare, ou ainda alguns versos de Sófocles. Começo sempre por um exercício que se chama «amar o logos», quer dizer *logos philein* ou filologia. Trata-se de descobrir, com o auxílio de todas as ferramentas que os eruditos nos propõem, a saber os diversos dicionários, o sentido primeiro, ingénuo, quase inocente de cada palavra. Este questionamento é de início uma interrogação puramente filológica. Depois, é preciso passar à gramática, música do pensamento porque através das diversas formas de gramática, que formam um mundo inesgotável, é dito que o pensamento humano se faz música. Não conheço um grande poeta que não seja um mestre gramático ou um virtuoso da sintaxe, do mesmo modo que não há sintaxe que não encerre uma visão do mundo, uma metafísica, e também uma filosofia da morte. Dizer que em certas línguas o pretérito não existe, dizer que em hebraico não existem verbos no futuro, é falar de uma visão global do universo, do homem e da identidade de cada um de nós. Tais são as razões porque nos demoramos em aspectos minuciosos daquilo a que costuma chamar-se a gramática e a sintaxe. Vem depois o contexto histórico. Recuso totalmente a ideia de uma ficção que rejeite a biografia, a história e o contextual. Muito pelo contrário, não há, em meu entender, uma única frase de *Madame Bovary* que não reflecta a história do Segundo Império, da vida de Flaubert, da língua francesa, da crise da burguesia. Nenhum texto pode pretender situar-se fora de um contexto que poderíamos comparar à infinita Torre de Babel, da biblioteca imaginária de Borges. A um nível limitado, porque é uma tarefa que excede um pouco os meus meios, abordamos a semântica, quer dizer o sentido do sentido, a compreensão da intencionalidade que todos os meus livros visam, sob um ou outro prisma. Então, regresso ao método medieval, que contém quatro etapas que percorrem a leitura, essa leitura tão pregnante e tão presente que devemos confessar que não compreendemos um poema ou um parágrafo e que precisamos

de o aprender de cor. O que não decorre de uma técnica mas de uma metafísica que se faz amor, que se faz Eros. Porque aquilo que sabemos de cor é inalienável; não se pode despojar ninguém daquilo que de conhecimento traz dentro de si num mundo onde reinam a censura e a opressão, o ruído, o exílio numa condição humana reduzida à segurança material esvaziada de toda a interioridade. Houve grandes almas que foram capazes de sobreviver à opressão por saberem certos textos de cor. Saber de cor uma página de prosa não é um exercício porque esse logos entra em nós – talvez demasiado difícil ou demasiado violento, inaceitável para nós –, e saber de cor a página significa que convidamos o seu logos a habitar a casa do nosso ser e que ele e nós aceitamos viver junto! Trata-se de corrermos o risco de que uma noite um texto, um quadro, uma sonata batam à porta da nossa morada – o meu livro *Presenças Reais* foi todo ele construído em torno desta imagem – quando é possível que o convidado destrua e incendeie a casa inteira. Pode ser também que nos roube e a deixe vazia! Mas temos de aceitar tomar o texto dentro de nós, e não sei como dizer a riqueza dessa experiência que vivi mil vezes, nomeadamente ao ler a *Ética*, de Espinosa, que é para mim uma referência última. Leio todos os dias Heraclito e certos poetas modernos como Paul Celan, e ainda quando talvez não compreenda bem os textos, aprendo-os de cor para que façam parte integrante do meu ser. A obra de súbito acolhe-me, sem se explicar e eu acedo finalmente ao poema. Nem por isso posso voltar para os meus seminários proclamando que enfim compreendi a obra, o que seria ao mesmo tempo arrogante e pretensioso. Todavia, é verdade que a incompreensão se transformou em amor, em fertilidade, em acto de confiança perante aquilo que me escapa. Gostaria de ilustrar o que estou a dizer com uma experiência que realizei nos Estados Unidos e que não foi uma experiência conseguida. Fui introduzido num grupo de terapia gestual onde me propuseram que acesse ao nível mais elementar da meditação deixando-me cair para trás sem medo uma vez que alguém estaria atrás das minhas costas para me amparar. Falhei o exercício o que me perturba intensamente. Tentei a valer, via outras pessoas que se deixavam, com uma confiança absoluta, cair lentamente para trás com os olhos fechados, mas não pude chegar ao mesmo resultado porque, para que a experiência seja bem conseguida, é preciso estar descomprimido no plano espiritual, *at homeless*, quer dizer estar como em casa na vida, estar na paz da própria alma. Também experimento a mesma impressão, mas ao ler os grandes textos de filosofia ou de metafísica, ou enriquecendo a minha cultura artística. Então, sim, deixo-me cair e às vezes esbarro no chão, mas vou aprendendo como

confiar no absoluto e no inacessível. O meu voto mais querido seria ter passado a minha vida a ler, a ler no sentido mais amplo do termo, como em inglês se pode dizer *I read a painting, I read a symphony*, incluindo no ler as belas-artes e a música. Toda a minha obra assenta na apreensão das vozes que se aproximam de mim. É por isso que escrevo na primeira linha de *Tolstoi ou Dostoiewski* que toda a verdadeira crítica é um acto de amor. É deste modo que me oponho às disciplinas modernas, sejam críticas, académicas, desconstrucionistas ou semióticas. Aos meus olhos, toda a boa leitura retribui uma dívida de amor.

[*Quatro entrevistas com George Steiner*. Trad. Miguel Serras Pereira. Lisboa: Fenda, 2000]

[Transcrição gentilmente autorizada pelas Edições Fenda.]

BIBLIO soft

Informática, Formação e Serviços, Lda

Desenvolvimento de aplicações em CDS/ISIS – Conversão retrospectiva de catálogos ma

Sede	Formação e Apoio Técnico
Rua Fernando Lopes Graça, n.º 4, 8.º A	Rua Ramalho Ortigão, n.º 35 A
2675-549 – Odivelas - Portugal	1070-228 Lisboa – Portugal
Tel./Fax: 219337388	Tel.: 213712157/8 Fax: 213712159

António Manuel Freire (Tel. 96 5871550) – Luis Damas (Tel. 96 5698910)
Email: bibliosoft@esoterica.pt

A **BIBLIOsoft**, em conjunto com a **Irisdata** criou um nova linha de produtos para gestão de bibliotecas (desenvolvidos em CDS/ISIS) para correr em ambiente Windows e que designou **BIBLIObase**. Relativamente à base bibliográfica, pretende ser um produto normalizado em formato **UNIMARC** de acordo com o que está definido pela parametrização Porbase 4 (da Biblioteca Nacional). Apesar de estar estruturada em **UNIMARC** não requer qualquer introdução de codificação pelo utilizador.

Produtos – CDS/ISIS

- ✓ Interface WWW para bases de dados CDS/ISIS (*MS/DOS e BIBLIObase*)
- ✓ Módulo para detecção de registos duplicados (*MS/DOS*)
- ✓ Impressão de códigos de barras (*MS/DOS e BIBLIObase*)
- ✓ Módulo de Circulação e Empréstimo (*MS/DOS*)
- ✓ Módulo de Gestão Kardex de Publicações em Série (*MS/DOS*)
- ✓ Utilitários (*MS/DOS do CDS/ISIS*)