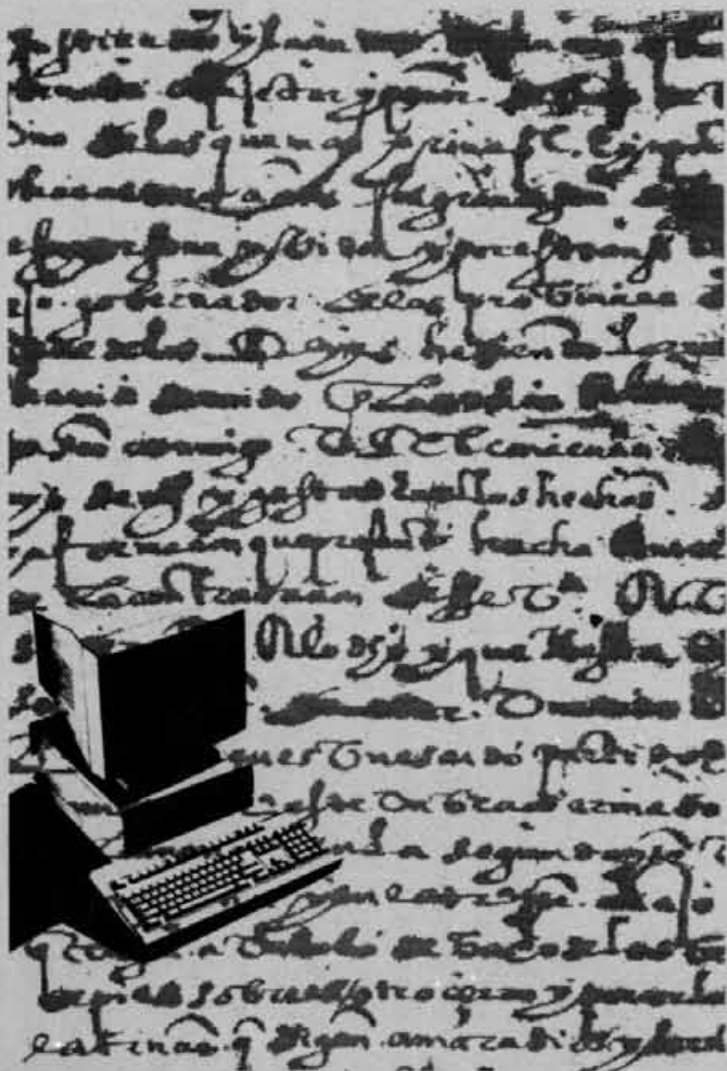


# ACERVO

REVISTA DO ARQUIVO NACIONAL

VOLUME 7 • NÚMERO 01/02 • JAN/DEZ • 1994



NOVAS TECNOLOGIAS  
EM ARQUIVOS

MINISTÉRIO DA JUSTIÇA



ARQUIVO NACIONAL

## Sistema BNDES

### A integração de tecnologias no gerenciamento dos documentos de arquivo

**Léa de Aquino**

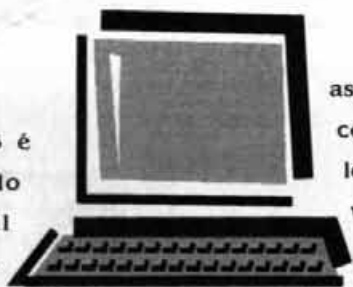
Bibliotecária e Arquivista.

Gerente de Administração de Documentos-BNDES.

#### HISTÓRICO

O Sistema BNDES é constituído pelo Banco Nacional de Desenvolvimento Econômico e Social - BNDES e por suas duas subsidiárias: Agência Especial de Financiamento Industrial - FINAME e BNDES S/A Participações - BNDESPAR.

Sua atribuição básica é apoiar os empreendimentos prioritários ao desenvolvimento da economia brasileira, com ênfase no estímulo à iniciativa privada nacional. De forma integrada, BNDES, FINAME e BNDESPAR dispõem de um conjunto de modalidades operacionais permanentemente atualizadas, visando



assegurar o atendimento às necessidades de investimento de longo prazo de empresas privadas e entidades públicas.

A ação financiadora do Sistema BNDES objetiva alocar os recursos à sua disposição de forma a garantir o maior e melhor impacto possível sobre o desenvolvimento nacional, promovendo crescimento de produção de bens e serviços, modernização e capacitação tecnológicas, geração de empregos e ampliação da gama de produtos competitivos no mercado externo.

Criado em 20 de junho de 1952, o então denominado Banco Nacional de Desenvolvimento Econômico - BNDE,

hoje subordinado à Secretaria de Planejamento, Orçamento e Coordenação da Presidência da República - SEPLAN-PR, ao longo de sua trajetória, acumulou documentos/informações de importância fundamental para a história do processo de desenvolvimento econômico e social do País.

A problemática de 'arquivos' do Sistema BNDES, até 1980, não diferia muito daquela existente na maioria das empresas destituídas de programas de gestão de documentos de arquivo.

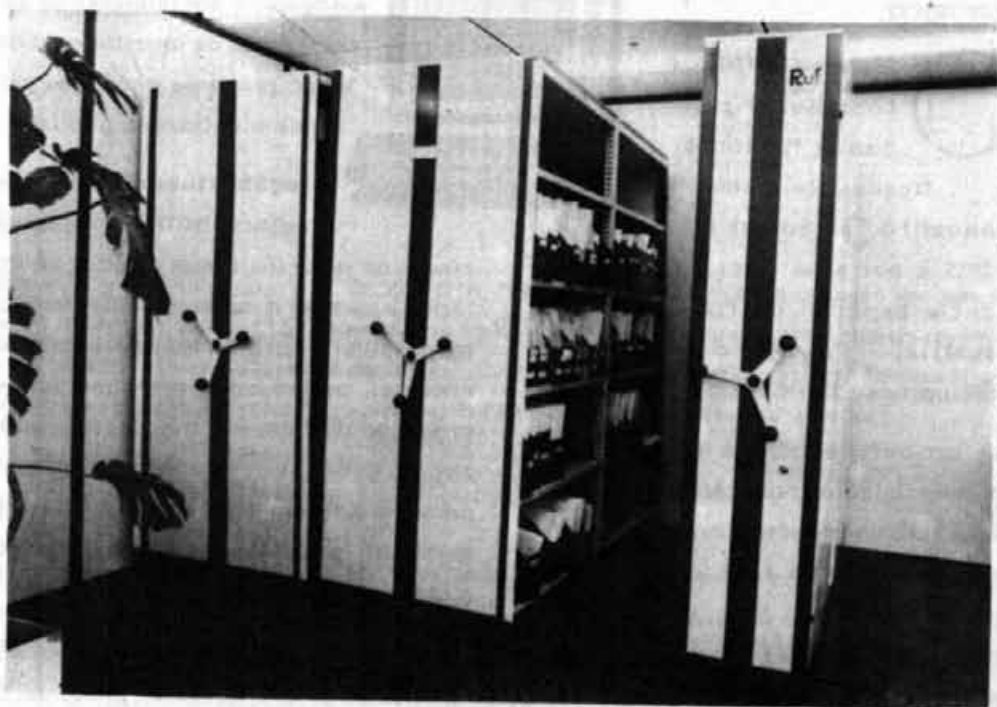
Foi detectado, no então arquivo central, um volume significativo de arquivos não-avaliados, simplesmente acondicionados em móveis-arquivo, com poucas possibilidades de localização e recupe-

ração de informações, e ocupando espaços cujos altos custos de manutenção estavam comprometendo a sua sobrevivência.

Contrastando com essa imagem do arquivo central, chamado por seus poucos clientes de 'arquivo morto', existiam, nas diferentes unidades integrantes da sua estrutura organizacional, diversos arquivos mantidos por seus gestores, sem procedimentos técnicos arquivísticos, mas com muito zelo, uma vez que continham informações indispensáveis e vitais aos diferentes processos de tomada de decisão.

#### ABORDAGEM CONCEITUAL

Diante desse cenário, estava evidenciada:



Reprodução: Agostinho Santos e Flávio Lopes. Arquivo Nacional

Arquivo intermediário - acervo no suporte papel

• A necessidade de se projetar um modelo de sistema de arquivo que intensifique o fluxo de informações da organização. Em outras palavras, o sistema de arquivo de documentos deve, especialmente, atender às necessidades de seus clientes internos e externos e integrar o programa corporativo geral de administração de informações da organização.

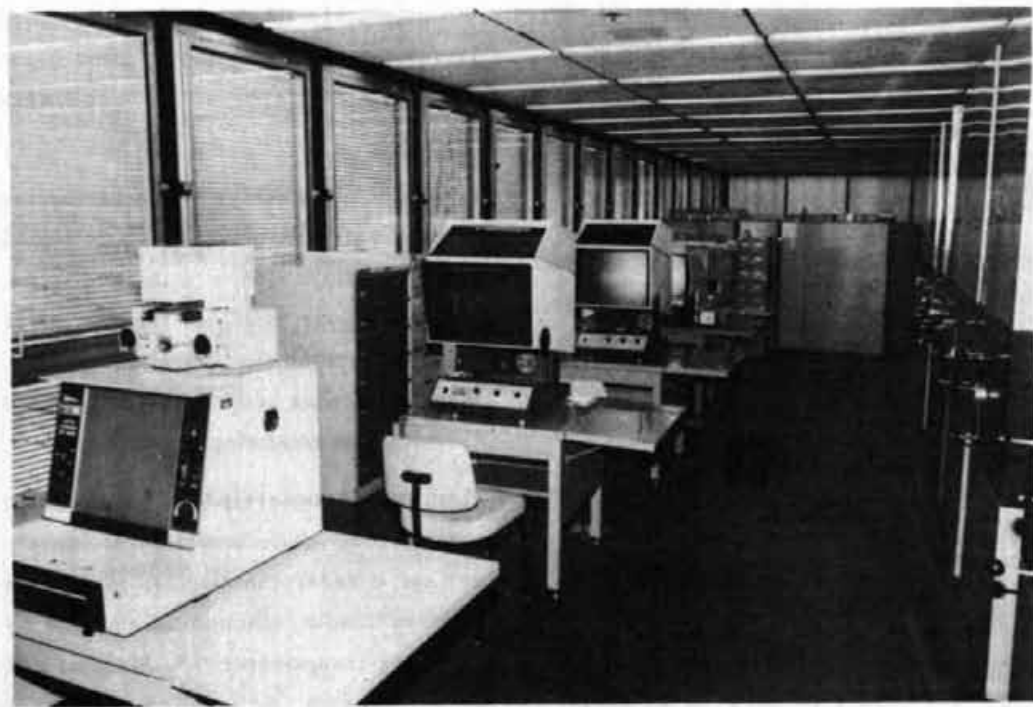
• A necessidade de que o sistema de arquivo possa acomodar o crescimento e as tecnologias futuras. Um sistema, portanto, flexível e preparado para a convivência com a evolução dos recursos da tecnologia da informação.

A necessidade de que o sistema

garanta uma redução de custos significativa, sem prejuízo da qualidade da informação requerida. Um ponto efetivo de 'venda' dos programas de gestão dos documentos de arquivo junto às administrações superiores é exatamente a economia de recursos, além dos fatores de segurança e qualidade.

#### O SISTEMA DE ADMINISTRAÇÃO DE DOCUMENTOS DO BNDES

Embasaada nos aspectos conceituais anteriormente abordados, a meta fixada em junho de 1980 - de projetar um sistema de administração de documentos com o objetivo de reduzir a massa documental do BNDES e de garantir, ao



Reprodução: Agostinho Santos e Filipe Lopes, Arquivo Nacional

Arquivo permanente - acervo em mídias diversas



informação é pelo menos de 500 milhões de caracteres por disco. O CD-ROM fornece uma capacidade de armazenagem lucrativa para computadores pessoais que excede aquela dos meios magnéticos e indubitavelmente incentivará posteriores desenvolvimentos e intensificações em armazenagem da informação e em técnicas e aplicações de recuperação.

### Recuperação

Duas questões críticas na maior parte dos meios de armazenagem da informação e dos sistemas de recuperação têm sido a estrutura do arquivo de índice (*index file*) e como as questões da pesquisa são formuladas. A estrutura do arquivo de índice implica comumente em arquivos convertidos

derivados de texto permutado, termos de assuntos marcados manualmente ou campos bibliográficos determinados (p. ex., título, resumo). Questões baseadas na lógica de pesquisa booleana ('e', 'ou', e 'não' isolados ou em combinação) junto com a proximidade da palavra são em geral desenvolvidas para identificar e recuperar documentos que reúnem critérios específicos. Arquivos de índice convertidos, texto formatado, vocabulário controlado e campos bibliográficos têm uma variedade de problemas<sup>15</sup>.

Pesquisa e recuperação de texto livre baseadas na linguagem natural e se utilizando *feedback* de relevância evitam muitos dos problemas de arquivos de índice convertidos, vocabulários controlados e campos bibliográficos. O *feedback* de relevância

utiliza as palavras num documento ou publicação que o usuário identifica como relevante. Comumente, um artigo conterà sinônimos e palavras-chaves que são utilizadas para procurar mais documentos e publicações de uma natureza semelhante. Diversas interações desse processo geralmente identificarão aqueles documentos e publicações que se casem com os resultados do *feedback*.

Um grande impedimento para o *feedback* de relevância tem sido o fato de que ele exige um computador extremamente rápido e poderoso para executar essas operações sobre volumes maciços de texto no tempo adequado (isto é, alguns segundos). O poder de processamento computacional paralelo maciço, particularmente o que é usado na *Connection Machine*, permite pesquisas instantâneas de centenas de milhares de artigos. Uma instalação da *Connection Machine* completa uma única questão em inglês natural de uma base de dados que consista em 112 milhões de caracteres em cerca de 74 milésimos de segundos, menos do que o tempo que se leva para pressionar a tecla que inicia a pesquisa.

A ausência de comandos complicados para memorizar, ou técnicas de pesquisa para aprender, e a velocidade da *Connection Machine* darão suporte ao 'folhear' eletrônico que mimetiza o processo que a maior parte das pessoas usam quando folheiam um livro ou

passam a vista num artigo de jornal. Conseqüentemente, a pesquisa de texto-livre de enormes quantidades de informação em forma eletrônica e a identificação e recuperação de documentos que preencham necessidades de informação muito específicas serão uma realidade na década que está diante de nós.

### **Compartilhamento da informação e funcionalidade integrada**

A terceira grande trilha da tecnologia da informação é a expansão da informação compartilhada e da funcionalidade integrada. O compartilhamento da informação implica na troca de informação eletrônica (imagens, textos, diagramas e coisas semelhantes) entre usuários que podem estar no mesmo escritório ou edifício ou a grandes distâncias um do outro, e cujos computadores podem ser incompatíveis. A funcionalidade integrada, por outro lado, significa aplicações disjuntas ligadas através de PCs, microcomputadores e *mainframes* de modo que conexões e traduções de dados sejam automáticas e transparentes para os usuários.

### **Compartilhamento de informação e redes digitais**

O compartilhamento de informação, particularmente de grandes quantidades de informação eletrônica, como no caso das imagens digitais, exige uma capacidade de transmissão substancial.

Redes de áreas locais (LAN - *Local Area Networks*) que geralmente incluem uma sucessão de escritórios ou um edifício único, permitem às pessoas compartilhar *hardware e software* que estejam localizados aproximadamente em um raio de 2.000 pés. Redes de áreas amplas (WAN - *Wide Area Networks*) ligam LANs que geralmente estão amplamente dispersas. Apesar da implementação em larga escala da Rede Digital de Serviços Integrados (ISDN - *Integrated Services Digital Network*), na década de 1990, permitir que a taxa de transferência de dados alcance desde 64.000 *bits* a 1,54 milhões de *bits* por segundo nas linhas telefônicas convencionais - o que facilitará a telecomputação - isso ainda é inadequado para grandes volumes de dados. A esse respeito, o uso crescente de fibras ópticas para transmissão de dados suprirá as redes de comunicações que dão apoio ao compartilhamento da informação em ampla escala<sup>15</sup>.

O crescimento em larga escala dos ISDNs e das redes de fibras ópticas na década de 1990 removerá um dos dois maiores impedimentos para a total permuta eletrônica de informação. O outro impedimento é a incompatibilidade entre sistemas computacionais, o que torna difícil, senão impossível em algumas instâncias, compartilhar a informação eletrônica.

O desenvolvimento de um contexto de 'sistemas abertos' ou de 'interopera-

bilidade' no qual aquelas partes do processamento computacional que necessitam serem compartilhadas se adequem a padrões públicos não-proprietários reduzirão grandemente essa barreira<sup>17</sup>. Uma sucessão de padrões para fomentar a interoperabilidade está sendo desenvolvida pela Organização de Padrões Internacionais sob a rubrica de Modelo de Referência de Interconexão de Sistemas Abertos<sup>18</sup>. Quando totalmente implantados no decorrer da próxima década, esses padrões permitirão aos usuários comunicar e compartilhar dados sobre sistemas desenvolvidos por diferentes distribuidores. Os padrões OSI (*Open Systems Interconnection*) não resolverão todas as questões do compartilhamento de informação, mas certamente removerão um grande impedimento. Na medida em que se expandam as redes digitais e a interoperabilidade, o escritório eletrônico, no qual o *software* computacional integra e automatiza as tarefas, se tornará uma rotina. Os maiores fabricantes de computadores já estão desenvolvendo o que se chama de aplicações 'descosidas' as quais ligam de modo visível as utilizações através de computadores pessoais, microcomputadores e *mainframes*. Uma interface gráfica que use imagens de tela ao invés de interfaces baseadas em caracteres (ou guiadas por comandos) facilitará o uso e também reduzirá a quantidade de tempo exigida para realizar uma tarefa. Por exemplo,

quando um arquivo é recuperado de um disco rígido ele aparecerá na tela como uma imagem que pode ser manipulada através de um movimento para uma aplicação visual específica, tal como o processamento verbal. Os filtros de tradução converterão automaticamente, conforme o necessitado, a informação para o formato que se deseje usar. Na

década de 1990 programas eletrônicos integrados de escritório suprirão um contexto computacional no qual os computadores pessoais se tornarão núcleos de comunicação de onde se derivará e se compartilhará a informação eletrônica.

**Tradução de Sebastião Uchoa Leite**

## N O T A S

1. Conferência realizada na Universidade de Macerata, Itália, em cinco de setembro de 1990.
2. No decorrer da última década, mais ou menos, vários arquivistas levantaram essa questão por escrito. Os arquivistas que sugeriram que os documentos eletrônicos exigem modificações na teoria e prática do arquivo são DOLLAR, Charles. "Avaliando os documentos legíveis em máquina". In: *American Archivist*; HAM, F. Gerald. "Estratégias arquivistas para a era pós-custódia". In: *American Archivist*, verão de 1981, p. 297-216; KESSNER, Richard. "Administração da informação automatizada: há uma função para o arquivista no escritório do futuro?". In: *Arquivaria 19*, inverno de 1984-85, p. 162-172; MANNE-HARITZ, Angelica. "O impacto da convergência sobre o ciclo de vida dos registros". In: *Management of Record Information: converging disciplines*. New York: 1990, p. 121-130; e BEARMAN, David. In: *Management of Electronic Records: issues and guidelines*. United Nations: 1990, p. 103-105. Os arquivistas que geralmente argumentam que a teoria arquivística satisfaz as necessidades dos documentos eletrônicos são COOK, Terry. "Da informação ao conhecimento: um paradigma intelectual para os arquivos". In: *Arquivaria*, 19, inverno 1984-85, p. 28-49; e BAILLEY,

Provavelmente a única contribuição de maior importância que os arquivistas podem dar para ajudar a estabelecer essa agenda é a defesa da incorporação da proveniência no projeto e na implementação dos sistemas complexos de informação.

No decorrer da próxima década os arquivistas têm que fazer a escolha entre continuar as coisas do modo como elas são usualmente ou redefinir e adaptar os princípios e práticas arquivísticos básicos. Aqueles que escolherem a primeira hipótese 'boiarão' dentro da estagnação em que, em esplêndido isolamento, tanto eles quanto os documentos que preservam se tornarão produtos históricos, meras curiosidades. Os arquivistas que escolherem redefinir e adaptar

princípios e práticas arquivísticos básicos, a fim de torná-los úteis para lidar com sistemas complexos de informação, transformar-se-ão em participantes fundamentais na comunidade de manejo da informação na medida em que se unirem a outros especialistas de informação que pensem igual no trabalho de assegurar que as novas tecnologias de informação atendam de modo adequado aos interesses organizacionais e societários em relação à responsabilidade documental.

**Os pontos de vista e opiniões expressos neste escrito são do autor e não devem ser considerados como algo que reflita a política oficial da National Archives and Records Administration.**

A

N

E

X

O

**E**ste exame da tecnologia da informação enfoca os desenvolvimentos que já estão dentro do mercado ou em breve estarão. Pelo fim deste século, esses desenvolvimentos e seus acréscimos caracterizarão os sistemas de informação eletrônica que apóiam as atividades de rotina - em casa, na escola, nos negócios, na ciência militar, na medicina, na educação e no governo, para citar apenas alguns poucos setores. Organizei esse exame em três amplas

categorias: apreensão da informação; processamento, armazenagem e recuperação da informação, e compartilhamento da informação.

#### **Apreensão da informação**

O veículo principal para a apreensão de informação para processamento computacional é comumente através da chave de entrada manual. Apesar dos desenvolvimentos no reconhecimento do carácter óptico (*Optical Character Recognition - OCR*), é pouco provável



mudanças significativas nesse campo na década de 1990. Contudo, um novo veículo para a apreensão da informação, chamado imagem digitalizada (*digital imaging*), será cada vez mais usado na década de 1990. A tecnologia da imagem digitalizada pega uma fotografia eletrônica de um documento de papel (textos, mapas, desenhos de engenharia, e coisas semelhantes) e a armazena digitalmente num sistema computacional<sup>1</sup>. Documentos, mapas, fotografias, diagramas, desenhos de engenharia e similares serão convertidos rotineiramente em imagens eletrônicas, indexados, comprimidos e armazenados em discos digitais ópticos<sup>2</sup>.

A expansão do uso da tecnologia de imagem digital óptica se tornou possível pelos sistemas sofisticados desenvolvidos para intensificar e manipular as

imagens de satélite dos anos de 1980. O uso difundido dessa tecnologia a um custo razoável ao lado dos novos desenvolvimentos em marcha asseguram que a apreensão das imagens digitais se expandirá rapidamente na década de 1990.

Uma das aplicações desenvolvidas mais rapidamente que se estende tanto sobre a chave de entrada tradicional de dados como sobre a tecnologia da imagem digitalizada é o *Geographic Information Systems (GIS)*<sup>3</sup>. Um GIS liga geograficamente dados espaciais e dados descritivos de mapas, gráficos, dados tabulares, fotografias, dados percebidos remotamente, e similares que auxiliam questões de bases de dados interativas complexas<sup>4</sup>. Os resultados dessas questões podem ser mostrados de modos variados, dependendo das necessidades do usuário. A maior parte das aplicações do GIS auxilia a análise de problemas relacionados com a produção agrícola, a extração de recursos naturais, a administração das florestas, a conservação de ambientes frágeis e o planejamento das cidades, para citar só algumas coisas que exigem enormes quantidades de dados e computadores velozes e poderosos.

### **Processamento, armazenagem e recuperação da informação**

A emergência de processamentos paralelos maciços, novos meios de armazenagem e de inteligência artificial têm importância enorme para o proces-

- A
- C
- E
18. Nos anos recentes, a documentação tem sido convertida para a forma eletrônica e armazenada em fitas de computador em estreita proximidade com os próprios documentos. A documentação está sendo substituída por sistemas de dicionários de fontes de informação contendo todas as relações lógicas (*metadata*) que funcionam num contexto interativo. Um proveito maior dos emergentes sistemas de diretórios de fontes de informação interativa é o fato de que provavelmente serão transparentes e convenientes para o usuário.
  19. Como foi observado anteriormente, na seção de retrospecto tecnológico, um padrão internacional para sistemas de diretórios de fontes de informação (*Information resource directory systems - IRDS*) está em desenvolvimento.
  20. DUCHEIN, Michel. Op. cit., p. 67.
  21. DUCHEIN, Michel. Op. cit., p. 75.
  22. Em geral, um quadro de designação de arquivos, que poderia assinalar nomes de arquivos de origem humana ( nos sistemas DOS isso é limitado a oito caracteres), controla o acesso a documentos ou registros individuais.
  23. Isso presume, é claro, que o usuário tenha permissão autorizada de acesso à informação.
  24. Regra geral, essas fronteiras são definidas em termos de escritório de origem.
  25. Kenneth Thibodeau, diretor do Centro de Documentos Eletrônicos da *National Archives and Records Administration* (Administração dos Arquivos e Documentos Nacionais) propôs essa abordagem, a qual, com efeito, transfere para algum tempo futuro o custo de tornar inteligível ou utilizável a informação eletrônica. Essa abordagem transfere também para os futuros usuários o prazo em que ocorrerão as conversões de *software*.
  26. Essa idéia de reavaliação periódica do valor de longo prazo é discutida em maiores detalhes na parte sobre preservação, mais adiante neste trabalho.
  27. Para mais informação sobre essas linhas de orientação, ver YOUNG, Julia Marks. "Bibliografia anotada sobre avaliação". In: *American Archivist*, v. 48, primavera de 1985, p. 190-216.
  28. Se tanto as séries de documentos de papel como os arquivos legíveis em máquina forem considerados como sistema de documentos, então essa distinção pode ser de pouca consequência. Nos Arquivos Nacionais dos Estados Unidos essa distinção se expandiu porque o que define uma série de documentos é o programa de ação e não o formato. Segundo essa visão, os documentos criados pelo mesmo escritório e que cobre a mesma atividade funcional geral são compreendidos como uma série, sem consideração para com o seu formato. Isso significou, na maior parte dos exemplos, que os documentos não-textuais - fotografias, mapas e fitas de computador - foram incorporados às séries de documentos de papel existentes.
  29. A capacidade de manuseio computacional indicou o fato de que os documentos podiam ser

processados por um computador. O potencial de ligação do computador significou que dois arquivos separados poderiam ser ligados por alguma característica comum. Embora sejam possíveis com documentos de papel, praticamente falando a quantidade de tempo exigida para realizar essas operações manualmente revela que elas não são possíveis de serem realizadas.

30. A portabilidade mostra se um arquivo legível em máquina pode ser usado num computador com um sistema de operação e um *software* diferentes daqueles nos quais ele foi originalmente operado.
31. Quando os arquivistas começaram a usar a expressão 'projeto de sistema', em geral eles tinham em mente discretas aplicações de base de dados as quais, de uma perspectiva de sistema, não eram muito complexas.
32. Ver, por exemplo, KOWOLWITZ, Alan. "Avaliação arquivística dos sistemas de informação *on line*". In: *Archival Informatics Technical Reports*, v. 3, n. 3, outono de 1988; McDONALD, John. "Um comentário". In: *Archival Information Technical Reports*, v. 3, n. 3, outono de 1988 e HEDSTROM, Margaret. *Novas técnicas de avaliação: o efeito da teoria sobre a prática*. Trabalho apresentado na Conferência de Arquivos Regionais do médio Atlântico, Albany, New York, maio de 1988.
33. Essa discussão se estende sobre o convincente argumento que David Bearman, entre outros, formulou a esse respeito.
34. Os estágios da administração do ciclo de vida dos registros de informação eletrônica são criação, avaliação, uso e disposição.
35. Os Arquivos Nacionais dos Estados Unidos organizaram recentemente um grupo de trabalho para identificar as necessidades funcionais da administração do ciclo de vida dos registros de informação eletrônica para o Governo Federal. Um projeto semelhante para as Nações Unidas está sendo conduzido pelo Comitê Consultivo para a Coordenação dos Sistemas de Informação (*Advisory Committee for the Coordination of Information Systems - ACCIS*).
36. Resultados preliminares de um estudo dos usuários dos arquivos nacionais sugerem que os pesquisadores não abordam os registros arquivísticos com assuntos amplos na cabeça, mas antes com múltiplas facetas de um assunto, incluindo-se datas, lugares geográficos, nomes próprios, acontecimentos históricos específicos. Eles estão de fato interessados em informações específicas como essas.
37. Ver PETERSON, Trudy Huskamp. "Princípios e documentos arquivísticos da nova tecnologia". In: *The American Archivist*, v. 47, n. 4, outono de 1984, p. 390.
38. Os roteiros descritivos existentes para arquivos de dados legíveis em máquina, tais como o MARC AMC e o *Cataloguing machine readable data files: an interpretative manual* (Chicago:

Mil bytes	= kilobyte (KByte) ou KB = $10^3$ = $2^{10}$ = 1.000 bytes
Milhão de bytes	= megabytes (MB) = $10^6$ = $2^{20}$ = 1.000.000 bytes
Bilhão de bytes	= gigabytes (GB) = $10^9$ = $2^{30}$ = 1.000.000.000 bytes
Trilhão de bytes	= terabytes (TB) = $10^{12}$ = $2^{40}$ = 1.000.000.000.000
Quatrilhão de bytes	= petabytes (PB) = $2^{50}$ = 1.000.000.000.000.000 bytes

Correlação entre bytes, mega, giga, tera e petabytes

magnéticos, por que comportam muito mais informações a serem pesquisadas. Vamos, então, compensar esse retardamento com a agilização de manuseio dos próprios discos. Buscou-se, então, a antiga 'geringonça' das máquinas *juke-boxes*, providas agora de braços robóticos que retiram, colocam e viram os discos em tempos mínimos.

São, então, centenas de gigabytes ou terabytes de informações prontas a serem acessadas e colocadas à disposição em uma tela super VGA (*Video*

*Graphics Array*), com resolução (precisão de imagem) de milhares de pontos por unidade de área e com centenas de nuances de cores.

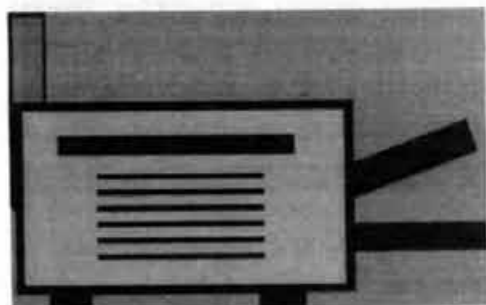
#### 'VENDO' O RESULTADO

Vamos ver agora como foi a evolução na outra extremidade dos sistemas de processamento da informação, isto é, as impressoras, os equipamentos que permitem materializar, na forma de papel ou outro meio qualquer, o resultado de um trabalho eletrônico. São os famosos dispositivos de *output*. As primeiras impressoras estavam ligadas aos grandes sistemas e funcionavam com a tecnologia dos grampos - as chamadas impressoras matriciais. Compunham cada caracter (letra, acento, espaço) por meio de grampos semelhantes a aranhas. A impressão era feita em papel que, para grandes trabalhos de computador, consistiam em formulários contínuos picotados. A qualidade de impressão não era boa, mas também não precisava ser de primeira qualidade, uma vez que o resultado eram enormes quantidades de relatórios numéricos e listagens destinadas a consultas técnicas.

Quando surgiram os processadores de gráficos e de textos, as impressoras matriciais passaram a ter melhor qualidade, aparecendo, também, as primeiras impressoras gráficas. Depois as gráficas em cores e, então, as primeiras impressoras do tipo máquina

de escrever, que imprimiam por meio de esferas e discos de tipos. Com essas, a qualidade melhorou, uma vez que as aplicações estavam voltadas também para escritório e trabalhos administrativos, como cartas, gráficos de apresentações, convites etc.

Acompanhando a evolução dos terminais e micros, as impressoras também utilizaram novas tecnologias que lhes proporcionaram maior velocidade, maior flexibilidade e melhor qualidade. Surgiram, então, as impressoras de linha, isto é, que imprimiam uma linha de cada vez, assim como apareceram,



mais rápidas ainda, as impressoras de linha bidirecionais, isto é, que imprimiam uma linha da esquerda para direita e a seguinte da direita para a esquerda, num vai-e-vem incansável. Eram as impressoras de espaço fixo, isto é, não importava que caracter estivesse sendo impresso, o espaço por ele ocupado era sempre o mesmo, definido pela 'fonte de tipos' característica da impressora. Esses tipos eram fixos e imutáveis.

Mais tarde, apareceram as impressoras

com tecnologias modernas e com grande variedade de tipos de impressão, como as impressoras de espaço proporcional. Seus tipos eram mais bem desenhados e uma letra 'i' ocupava menos espaço no papel do que a letra 'm', por exemplo. Em seguida, a mudança de tipos, de acordo com a necessidade do documento, ficou mais fácil. Bastava trocar a fonte de tipos antes da impressão e se obtinha um texto impresso em **universe, bodoni book, light italic** etc...

Ao mesmo tempo, vieram as impressoras coloridas e gráficas, com as quais era possível inserir gráficos coloridos para ilustrar um relatório ou uma pesquisa, por exemplo.

Essas impressoras podiam ser independentes de sistemas, ou seja, estarem conectadas a micros ou a sistemas semi-eletrônicos de processamento de textos. As impressoras de sistemas tornaram-se cada vez mais versáteis, pois podiam sofrer variações de fontes de tipos via programação e não mais fontes 'físicas' de tipos, ou cartões internos da máquina.

Maravilha das maravilhas: as impressoras de não impacto! Essas não tinham mais grampos ou martelos ou matrizes metálicas que precisavam ser 'borradas' com tinta para depois serem comprimidas com força no papel, deixando o caracter alfanumérico impresso em baixo-relevo. Eram as impressoras a jato de tinta que utili-

autorizadas a informações restritas. Hoje, alguns fac-símiles convertem documentos para a forma digital, possibilitando a captura da imagem em arquivo-texto para referência futura. Os fac-símiles apresentam diversas vantagens sobre os métodos de distribuição de documentos.

Quando não existe um método inteiramente eficiente para arquivar, recuperar e distribuir documentos, o usuário é forçado a utilizar uma combinação deles. Nesse caso ele sofre com a proliferação de equipamentos de escritório - incompatíveis - que requerem diferentes experiências, diferente gerenciamento, diferentes fornecedores e contatos de vendas dentro de sua própria organização. Organizações com necessidades regulares para distribuição de documentos externos usam, atualmente, fac-símiles em redes locais, associados a sistemas de correio eletrônico.

Os fac-símiles existem desde a virada do século. No passado, tiveram aplicações muito limitadas, mas seu uso cresceu significativamente nos últimos dez anos. Esse crescimento se deve a diversos fatores, como a tendência de descentralização, necessidade de redução de custos e desenvolvimento da tecnologia disponível.

Um dos fatores que levaram ao crescimento do uso de fac-símile foi o desenvolvimento de sistemas de transmissão para jornais na década de 1920.

Foram usados, primeiro, para aplicação de leis, divulgação de notícias, informações meteorológicas e transmissão de informações médicas. A decisão Carterfone, dos Estados Unidos, em 1969, possibilitou as companhias de produtos não-telefônicos a terem acesso às linhas públicas de transmissão. Esta decisão criou competição e crescimento na indústria de fac-símiles norte-americana e, conseqüentemente, mundial.

Os fac-símiles podem transmitir documentos contendo gráficos e material manuscrito, não estando limitados a informações codificadas ou em caracteres como em outros sistemas. Requer um investimento muito menor do que sistemas baseados na palavra, e o usuário não precisa de um volume grande de material a transmitir para justificar os custos de seu investimento. Apresenta um nível mais baixo de erros do que outros sistemas de distribuição porque a transmissão não envolve datilografia.

Os protocolos de transmissão têm sido padronizados a partir de quatro grupos de fac-símiles. Os modelos mais novos, normalmente, são compatíveis uns com os outros.

Em 1981 o International Telegraph and Telephone Consultative (CITT) desenvolveu padrões de compatibilidade para estes equipamentos, tendo estabelecido padrões para quatro

grupos de equipamentos. O grupo I tem como característica a capacidade de operar com modulação analógica FM, transmitir uma página de quatro a seis minutos, com uma resolução de 60 pel (*picture elements per inch*). Estes sistemas são lentos e produzem cópias de má qualidade. São usados tipicamente por organizações que têm um volume de necessidades muito pequeno (até cinco páginas por dia).

Já o grupo II opera com modulação analógica AM, transmite uma página de dois a três minutos com resolução de 90 pel. Estes sistemas são mais populares. Em geral, são usados por organizações que transmitem por volta de dez páginas por dia.

O grupo III opera com técnicas digitais, transmite uma página em um minuto ou menos, usando o método de compressão por bandas. Operam com velocidade duas vezes maior do que os do grupo II, e podem economizar tanto o tempo do operador como os custos de linha.

Já o grupo IV se caracteriza por operar com técnica digital de alta velocidade (tempo de transmissão de 56.000 bits por minuto), resolução de 200-400 pel, capaz de ser conectado diretamente a circuitos digitais, sendo compatível com satélite e usando protocolos de alto nível.

Os equipamentos do grupo IV têm dispositivos avançados, como capacidade de armazenamento e transmissão para correio eletrônico quando ope-

rando em ambiente conectado com computadores, e capacidade de comunicação aperfeiçoada para integração com outros sistemas sofisticados de escritório.

Existem dois tipos de sistemas de fac-símiles: analógicos e digitais. Estes termos se referem à maneira pela qual o equipamento processa a informação resultante do *scanning* de um documento.

Um fac-símile analógico decompõe todas as partes de uma página original, incluindo os espaços em branco. Uma célula fotoelétrica produz um sinal elétrico para cada elemento da figura decomposta. Após a transmissão, esses sinais elétricos dirigem uma impressora no equipamento de recepção, produzindo uma nova página de elementos da figura que se assemelha à original. A maioria dos fac-símiles de mesa, pequenos e baratos, são analógicos.

Esses equipamentos são caracterizados por transmissão lenta, pois transmitem todos os elementos e espaços em branco de um documento e levam de dois a seis minutos para transmitir uma única página. A qualidade de imagem é ruim porque o papel térmico, o *scanning* e a tecnologia de impressão são rudimentares, condições que se combinam para a formação da imagem de pouca qualidade que tende a se desvanecer com o tempo. Não há detecção de erro, pois os protocolos de

Um sistema de processamento de documento-imagem é um sistema que facilita a conversão de um sistema de arquivamento e recuperação de **papel físico** em um sistema computadorizado de arquivamento e recuperação de **imagem**. Suas funções possibilitam digitalizar um documento-papel, armazenar uma imagem do documento, registrar o documento, recuperar/exibir em tela e imprimir o documento-imagem, como também definir sua rota e distribuir documentos para processamento. Ou seja, fazer eletronicamente tudo o que hoje se faz com os papéis e malotes, mensageiros e arquivistas.

Esses sistemas têm a capacidade de indexar e arquivar tanto dados-objeto alfanuméricos como documentos-imagem num armazenamento *on line*. O sistema utiliza disco magnético para documentos de acesso freqüente, e meio ótico para recuperação menos freqüente de documentos que não deverão nunca ser eliminados do sistema. O movimento de dados entre arquivo magnético e ótico é controlado automaticamente de acordo com os atributos de processamento e retenção de cada documento individual.

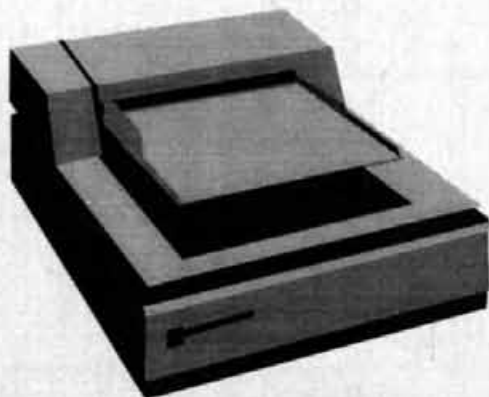
Um objeto é representado por uma coleção de dados e controles arquitetados, que podem ser intercambiados e processados por aplicações ou por componentes de um equipamento.

O que é um Sistema Ótico de Arquivamento, o SOA?

O SOA fornece o armazenamento para grande número de objetos com uma taxa relativamente baixa de recuperação. Esse sistema é formado por bibliotecas óticas e *drives*, e baseia-se na tecnologia de arquivo ótico de *write once read many* (escrever uma vez, ler muitas vezes), e na *write many read many* (escrever muitas vezes, ler muitas vezes). O SOA é controlado pelo componente de gerência de arquivos. Em uma unidade da biblioteca podem ser armazenadas até 2.500.000 imagens digitalizadas com 50 kbytes por página.

#### ESTAÇÕES DE IMAGEM.

**E**m que uma dessas estações se difere de um micro comum? As estações de imagem são definidas para trabalharem, não com dados somente, mas com imagens



geradas por aglutinação de pontos e até com voz. E pode apresentar configurações de monitor que permitam a visão simultânea de dados-imagem e dados alfanuméricos. O monitor pode ser de alta resolução e de tela grande para

apresentar todos os detalhes de uma foto, um diagrama, uma assinatura ou uma radiografia, por exemplo. Essas estações permitem, ainda, a conexão direta entre a estação e *scanners* de documento e/ou impressoras. As funções de gerenciamento de imagem dentro das estações de imagem, trabalhando em conjunto com o programa do computador, proporcionam a manipulação de uma página, modificação de documento e movimentação entre páginas dentro de um documento de múltiplas páginas.

No ambiente de rede local de comunicação, redes de estações de imagem permitem o acesso simultâneo a aplicações de transações e seus dados-imagem associados.

A indexação do documento é feita durante o processo de captura ou automaticamente pelo sistema ou pelo operador. À medida que as páginas de um documento são digitalizadas, o operador da estação de trabalho garante sua legibilidade por meio de ajustes de qualidade da imagem. Uma vez capturada a imagem, as atividades da aplicação apropriada podem ser planejadas. O documento migra automaticamente para o arquivo ótico durante o ciclo de gerenciamento de armazenamento.

A recuperação, a visualização e a impressão de documento são iniciadas pelo operador da estação de trabalho. O documento pode então ser visto, manipulado e impresso por meio de um

'diálogo' interativo, disponível ao operador da estação de trabalho, através de programas típicos, permitindo as mesmas operações executadas com as pastas de papéis.

Um sistema de processamento de documento-imagem, portanto, faz o gerenciamento da captura e processamento da informação, a qual pode estar em forma de imagem, texto, dados ou gráficos, e até associada a trechos de filmes e a voz. Esses objetos são representados como um documento que, por sua vez, pode ser composto por uma ou mais páginas que tenham sido digitalizados por um *scanner* conectado ao sistema.

A capacidade de *input batch* transfere grandes volumes de documentos, tanto em forma de dados codificados como de imagem, para o sistema, uma vez que esses documentos tenham sido formatados apropriadamente, utilizando-se padrões e arquiteturas desenhadas para processar essa nova modalidade de informação. Uma página neste contexto é toda e qualquer imagem digitalizada de uma folha física de papel, e um documento é uma coleção dessas páginas.

Quando um documento precisa ser processado, ele é direcionado para uma fila de trabalho apropriada. A aplicação possibilita ao usuário definir sua própria fila de trabalho com base no tipo de documento, e a cada documento está associada uma prioridade que o coloca em uma fila de trabalho. A aplicação

específicas e complicadas até as mais populares e prosaicas, como controlar a temperatura de um forno de micro-ondas assando uma 'pizza' - teremos diante de nós uma tarefa tão grande como têm os sistemas de gerenciamento de bibliotecas, que permitem a um estudante, na sala de estudos da biblioteca da Universidade de São Paulo, consultar um verbete de um volume da biblioteca do Congresso americano, ou a última avaliação de um estudioso, disponível na biblioteca de Jerusalém! Mas, sem sombra de dúvida, esse leque de recursos foi enormemente enriquecido com a disponibilização do processamento de informação em forma de imagem. E, diariamente, novas tecnologias, novas aplicações, novos produtos estão sendo apresentados, nos dando a certeza de que muitas maravilhas estão ainda por vir, enriquecendo

o conhecimento humano, tornando a vida na Terra mais sadia, mais segura, mais alegre e, por mais paradoxal que pareça, mais simples.

O que pode vir no futuro? A imaginação humana nos tem presenteado com especulações fantasiosas em filmes, livros e artigos de revistas. Mas o que nos vem sendo apresentado na realidade tecnológica tem sido tão fantástico que nada mais podemos classificar, com certeza, como fantasia alucinada.

Visitar outras galáxias, sem sair de nossas salas - aí está a realidade virtual!!! Poderemos nós, algum dia, projetar nossos problemas e necessidades em vídeo para nos preparar para resolvê-los antes que aconteçam? Tudo será possível, pois é ilimitada a capacidade humana em criar e aperfeiçoar. Até os definitivos limites das ciências exatas já estão deixando de ser tão definitivos assim!

## A B S T R A C T

The article is a brief survey of the development of information since the Egyptians till nowadays when it is evident the importance of computer science and of image processing systems. It analyses the evolution of each type of technology (printers, telephony, telex, facsimile, computers, etc.) and their implications for the field of archivalology, focusing on the resources for image processing and on the systems to integrate such tools. The article also speculates on the future development of computer science.

## R É S U M É

Brève rétrospective du développement de l'information, des égyptiens jusqu'à l'époque actuelle, au cours de laquelle l'importance de l'informatique et des systèmes de traitement d'images est mise en évidence. Cet article analyse l'évolution de chacune des technologies (imprimantes, téléphonie, télécopie, fac-similés, ordinateurs, etc...) et de leurs implications aux archives, en soulignant, en particulier, les ressources du traitement d'images et des systèmes d'intégration de ces techniques; il explore en outre le développement futur de l'informatique.