

BASES DE DADOS EM CIÊNCIA E TECNOLOGIA

Cecília Carmen Cunha PONTES

Deptº de Pós-graduação em Biblioteconomia
PUCCAMP

RESUMO

PONTES, Cecília Carmen Cunha. Bases de dados em ciência e tecnologia. *Trans-in-formação*, 2(2/3): 33-42, mai/dez. 1990

O impacto do desenvolvimento da ciência e da tecnologia na sociedade moderna e o surgimento de uma indústria de informação é analisado face o desenvolvimento da tecnologia de informática e telecomunicações. Também apresenta dados sobre a indústria de informação no Brasil, com ênfase nos bancos de dados apoiados pelo Programa de Apoio ao Desenvolvimento Científico e Tecnológico PADCT/CNPQ/SCT.

INTRODUÇÃO

Informação refere-se a todos os fatos e opiniões emitidos ou recebidos durante o curso da vida diária. A informação é obtida de outros seres humanos, através dos meios de comunicação de massas, dos bancos de dados eletrônicos e de outros tipos de fenômenos observáveis no ambiente. O interesse pelo fenômeno informação cresceu no século XX com o desenvolvimento da ciência e da tecnologia e hoje é objeto de estudo em numerosas disciplinas tais como filosofia, física, biologia, lingüística, engenharia, administração, biblioteconomia, ciências sociais e, mais recentemente, a informática(1).

Informação em ciência e tecnologia é uma expressão que designa os conhecimentos obtidos a partir das atividades de pesquisa e desenvolvimento. É insumo essencial e produto final da

pesquisa, e veicula os novos conhecimentos que geram o progresso técnico e o avanço das ciências.

A Informação em ciência e tecnologia assume no mundo moderno uma importância estratégica no desenvolvimento da sociedade. As nações desenvolvidas têm investido grandes somas no processo de geração de informação, propiciando o surgimento de uma indústria com grande rentabilidade.

Para dimensionar corretamente as dificuldades que envolvem o desenvolvimento de uma indústria da informação, é necessário reportar-se ao próprio desenvolvimento da atividade de pesquisa na atualidade.

A pesquisa científica e tecnológica se apresenta, no mundo moderno, com um alto grau de complexidade, desaparecendo aparentemente as fronteiras entre a ciência e a técnica. A ciência, denominada por Sola Price de "Big Science", caracteriza-se pelo advento dos grandes projetos de desenvolvimento tecnológico(2).

Essa ciência, que chamaríamos de ciência industrial porque a pesquisa passa a ser encarada como atividade empresarial e o seu produto uma "mercadoria", determina uma mudança profunda tanto a nível da estrutura e organização das instituições de pesquisa como, especificamente, no tipo de informação que estas instituições passam a gerar(3).

A produção de tecnologia, que durante a história passada foi uma atividade comum a todas as sociedades, converteu-se cada vez mais em privilégio de um pequeno grupo de países, e, dentro destes, de organizações e empresas que podem financiar seus custos sempre crescentes. Para uma grande parte da humanidade, a tecnologia, que havia sido sempre parte integrante da cultura individual e coletiva, tornou-se um fator exógeno. "O que fazer" e "como fazer" começam a ser determinados por valores sociais e culturais de outras sociedades, implícitos nas tecnologias importadas(4).

Após a Segunda Guerra Mundial, a atividade científica sofreu mudanças importantes e radicais. De uma atividade marginal e complementar, transformou-se num campo essencial da produção. O número de cientistas atualmente vivos supera de longe o número daqueles que existiram durante milênios de história. E esta quantidade aumenta num ritmo ainda mais rápido,

duplicando a cada 10 anos. Por outro lado os gastos com as atividades científicas cresceram, nos últimos anos, num ritmo idêntico(5).

Quadro 1. Dados Sobre Pesquisa e Desenvolvimento por Regiões.

1980			
Regiões	Número de Cientistas	Cientistas por milhões de pessoas	Gastos em R&D - %
USA	674.725	2.679	2,33
EUROPA	839.473	1.725	1,79
OCEANIA	33.889	1.472	1,11
ÁSIA	693.659	273	1,18
AMÉRICA LATINA	90.936	251	0,49
ÁFRICA	40.812	91	0,36

STATISTICAL YEARBOOK, 1986 - UNESCO

Segundo dados da UNESCO apresentados na tabela anterior, nos países desenvolvidos, existem aproximadamente 2.000 cientistas para cada um milhão de pessoas. Em países em desenvolvimento esse número se reduz a aproximadamente 250. Quanto aos gastos em Pesquisa e Desenvolvimento os países ricos gastam em média 2% do seu PND (produto nacional bruto), ao passo que os países pobres gastam por volta de 0,5%.

Hoje em dia, a ciência passa a atuar universalmente como força diretamente produtiva e a indústria por inteiro converteu-se em aplicações do processo de trabalho simples e sua conversão em processo científico. Cada vez mais a tecnologia torna-se um campo de aplicação do conhecimento científico, eliminando progressivamente os fatores aleatórios e subjetivos da produção, para substituí-los por atividades objetivas e planejadas apoiadas na máquina, isto é, no computador.

Em decorrência dessas transformações surge uma nova divisão internacional, cabendo às nações economicamente

desenvolvidas a responsabilidade das novas descobertas científicas e tecnológicas, atribuindo-se aos países em desenvolvimento a função de adaptadores de tecnologia importada.

A INDÚSTRIA DE BANCO DE DADOS

A ciência da atualidade está produzindo uma quantidade de dados que vai atingir as proporções de um dilúvio, milhões de vezes maior que a capacidade humana de manipulá-los ou estudá-los. Os pesquisadores produziram e armazenaram em máquina, no últimos 20 anos, um volume de informações equivalentes a tudo que foi descoberto nos seus 20 mil anos de história da civilização do homem(6).

Neste contexto, o fortalecimento de uma indústria de informação passa a ser prioridade nacional e considerado um importante indicador do desenvolvimento sócio-econômico do país. Alguns índices como número de publicações, registros de normas e patentes, micro-fichas editadas, número de bibliotecas e centros de informação e principalmente bancos de dados disponíveis à população, mostram a dimensão do desenvolvimento técnico-científico e sócio-econômico do país.

As características estruturais e mercadológicas da indústria da informação, identificada como predominantemente monopolista, estão provocando a concentração das empresas que atuam nesta área em meia dúzia de países localizados no hemisfério norte. Por outro lado, a predominância da língua inglesa, nos canais de comunicação, tem reforçado a liderança dos Estados Unidos na área, levando outros países a esforços semelhantes.

A organização e disseminação de informação tornou-se, no mundo moderno, assunto de políticas públicas e muitos países fundaram bibliotecas, e arquivos nacionais para proceder à aquisição ordenada de registros e documentos. Posto que essas instituições não estão capacitadas a executar por si a disseminação desses documentos, novas atividades cooperativas e compartilhamento de material gravado propiciaram o surgimento de redes de bibliotecas e banco de dados públicos e privados, nacionais e internacionais.

Na área de documentação, pela primeira vez na história, está se tornando econômica e tecnicamente viável a automação

das redes de bibliotecas e a criação de grandes distribuidores de informação.

TECNOLOGIA DE INFORMAÇÃO

Nas últimas décadas, o volume de documentos produzidos pelos países industrializados ocasionou a busca de tecnologias que permitam o processamento mais rápido e eficaz dessas informações, provocando uma substancial mudança quantitativa e qualitativa no panorama internacional e propiciando o surgimento de uma estrutura de prestação de serviços de dimensões intercontinentais, com utilização de novas técnicas nas áreas de computação e tele-processamento e a progressiva hegemonia dos países industrializados nesse setor.

Após a década de 70, com a utilização do computador na recuperação automatizada de informação em linha (on-line) e a proliferação das bases de dados, essa técnica tornou-se essencial permitindo a triagem de milhões de referências bibliográficas. Este material, armazenado em memórias magnéticas e mais recentemente memórias óticas, tornou a atividade de informação uma das mais complexas e ao mesmo tempo mais lucrativas nos países desenvolvidos.

Avalia-se que estão disponíveis ao público, em todo mundo, por volta de 4.300 bases de dados distribuídas pelas grandes empresas.

Quadro II. Disponibilidade de Bases de Dados

1972	6 bases de acesso público
1976	75 bases de acesso público
1985	2800 bases de acesso público
1989	4300 bases de acesso público

- Dados do DIALOG - Curso de acesso às Bases de Dados do Sistema - 1991

Hoje com uma produção mundial estimada de 5 milhões de artigos de periódicos anualmente, quase que na sua totalidade

registrada em bases de dados automatizadas, a atividade de recuperação de informação de forma precisa e exaustiva passou a constituir-se numa tarefa que exige sofisticada tecnologia de computação e telecomunicações. Em decorrência, pesquisas estão sendo realizadas pela Ciência da Informação, utilizando metodologias emprestadas das áreas de Lingüística e Ciência da Computação, visando a identificação de algoritmos que permitam a otimização dos processos de recuperação das milhares de bases de dados hoje disponíveis no mercado.

Essas pesquisas, hoje identificadas na área de desenvolvimento de "sistemas especialistas", possuem como objetivo reproduzir as etapas do raciocínio de um ser humano para identificar e selecionar informações armazenadas em computador. Esses sistemas possuem três componentes básicos: um **software** de interface através do qual o usuário pode formular perguntas para as quais o sistema solicita mais informações e explica ao usuário o processo empregado para chegar à resposta; um banco de dados composto de axiomas, chamado base do conhecimento e regras para se fazer inferências; programa de computador que executa as inferências.

Esses sistemas que têm como suporte processadores digitais com velocidade de acesso e confiabilidade de dados são capazes de emitir milhões de sinais por segundo a longa distância, permitindo a utilização dos recursos computacionais por uma grande quantidade de usuários, de forma simultânea. Por outro lado, o desenvolvimento de memórias de armazenamento de dados, através de discos óticos, está propiciando o surgimento de bancos de dados de tamanhos incalculáveis e permitindo a representação em um mesmo meio de imagem, vídeo e som.

A implantação de sistemas sofisticados de recuperação automática de informação e processamento de grandes volumes de informação pressupunha, até o final da década de 70, a utilização de processadores digitais de grandes dimensões e requeria investimentos consideráveis. Entretanto, hoje com o desenvolvimento da tecnologia de microinformática e a pesquisa de sistemas especialistas, a recuperação de documentos passou a ser uma tarefa viável para centros de informação/documentação de pequeno porte ou usuários individuais.

O advento da micro-informática e da tecnologia de armazenamento de dados por processos óticos (CD-Rom), está propiciando o surgimento de uma nova forma de acesso às bases de dados, tornando viável a adoção de uma rede descentralizada de informação com centros operando localmente sistemas automatizados. Este modelo poderá representar uma opção mais adequada à realidade dos países em desenvolvimento.

BANCOS DE DADOS EM CIÊNCIA E TECNOLOGIA NO BRASIL

No Brasil, o fortalecimento de uma indústria de informação nacional passou a ser um dos objetivos dos órgãos de planejamento de Ciência e Tecnologia.

Estima-se que, atualmente, a rede de bibliotecas universitárias do país é composta por mais de 1000 centros, possuindo um acervo de 15.000.000 de livros e 500.000 coleções de títulos de revistas.

O IBICT - Instituto Brasileiro de Informação em Ciência e Tecnologia, criado em 1954, com o nome de Instituto Brasileiro de Bibliografia e Documentação, está vinculado a Secretaria Especial de Ciência e Tecnologia e tem como objetivo estimular a implantação de uma indústria de informação nacional.

Os serviços automatizados de informação mais importantes estão nas áreas da saúde (BIREME - Centro Latino Americano e do Caribe de Informação em Ciência da Saúde), Energia (CIN-CNEN-Comissão de Nacional de Energia Nuclear) e Agricultura (EMBRAPA e Ministério da Agricultura). Essas áreas possuem serviços organizados, conectados a sistemas internacionais, disseminando informações em ciência e tecnologia. (7)

A coordenação de sistemas automatizados para catalogação cooperativa de documentos é de responsabilidade de duas instituições brasileiras: a Fundação Getúlio Vargas, que opera um sistema de catalogação de monografias, Bibliodata-Calco, seguindo os padrões da Biblioteca do Congresso Americano e interligando 70 bibliotecas do país; o IBICT, que opera um Catálogo

Coletivo Nacional de Publicações Seriadadas - CCN, com a participação da rede de bibliotecas especializadas do país.

Em 1984, com o intuito de implementar os serviços de informação em C&T, foi instituído o Sub-programa de Informação em Ciência e Tecnologia no âmbito do Programa PADCT/CNPQ. Este sub-programa, em sua fase I, reforçou principalmente os serviços já existentes sem enfatizar os aspectos de inovação. Entretanto, um dos projetos iniciados nesta fase, o "Sistema Público de Acesso a Bases de Dados", propiciou a definição e o teste, em fase experimental, de um modelo de sistema cooperativo para acesso à bases de dados nacionais e internacionais. Este projeto contou com a participação de três instituições brasileiras com tradição na operação de serviços On-Line: BIREME, CIN-CNEN e IBICT. Os resultados alcançados pelo PADCT, na fase I, levaram o grupo técnico do Sub-programa de ICT a reformular sua estratégia no PADCT II, concentrando esforços na implementação de um sistema integrado de informação em C&T.

Hoje, o IBICT, apoiado pelo Programa de Apoio ao Desenvolvimento Científico e Tecnológico PADCT/CNPQ/SCT, está coordenando dois projetos que são iniciativas importantes para a implantação de uma indústria de informação nacional.

O primeiro projeto no âmbito do sub-programa "Informação em Ciência e Tecnologia" do PADCT, que tem como finalidade a ampliação e melhoria da qualidade dos serviços de informação em ciência e tecnologia do país, está implantando o "Sistema Público de Acesso a Bases de Dados" que se propõe a administrar um serviço automatizada de informação interligando via rede de comunicações de dados da Embratel (RENPAQ), centros nacionais que operam serviços de informação automatizados, colocando à disposição da comunidade de ciência e tecnologia do país, acesso on-line a bases de dados nacionais e estrangeiras. Também pretende instalar sessenta postos de serviços (PS), a serem localizados nos principais centros de documentação/informação do país(8).

O segundo projeto, no âmbito do sub-programa Tecnologia Industrial Básica do PADCT, através do projeto, desenvolve um serviço de informação e extensão industrial, operando uma "Rede de Informação Tecnológica", composta de vinte e cinco núcleos

setoriais e regionais, transferindo tecnologias disponíveis para a pequena e média empresa.

Concluindo, indagaríamos: hoje, a indústria de informação está contribuindo para a democratização da sociedade brasileira na sua globalidade? Ou está reforçando uma divisão das nações: de um lado, aquelas detentoras dos bancos de dados disseminados em redes nacionais e internacionais, de outro, aquelas somente usuárias e portanto dependentes desses bancos de dados para implementar seu progresso técnico-científico?

Acreditamos que a operação de uma rede **On-Line** nacional de informação em Ciência e Tecnologia, com dimensões similares às empresas existentes nos países mais avançados, atuando de forma cooperativa, tornará disponível para a comunidade de Ciência e Tecnologia as principais bases de dados nacionais e internacionais e poderá contribuir para a democratização do acesso à informação da rede de bibliotecas especializadas do país.

SUMMARY

PONTES, Cecília Carmen Cunha. Science and technology data base. *Trans-in-formação*, 2(2/3): 33-42, may/dec. 1990.

The impact of science and technology on modern society is analysed in the light of computer and telecommunication technology. Data of brazilian information industry is presented with emphasis on the implementation of data bases supported by the National Development of Science and Technology Program.

REFERÊNCIAS BIBLIOGRÁFICAS

- (1) SLAMECKA, W. - Information Processing and Information Systems-in Encyclopedia Britanica, vol 1989, p. 552.
- (2) SOLLA PRICE, D. SPIEGEL R. - Science Technology and Society - Londres Sage Publication, 1977
- (3) PONTES, C. C. - Ciência e Tecnologia em São Paulo: Política e Realidade - Tese de doutorado defendida pela F. F. L. C. H. da USP, 1982
- (4) HERRERA, A. - El nuevo papel de la Tecnologia - in Anais do Seminário sobre Teoria do Desenvolvimento da Ciência e Tecnologia - São Paulo, Secretaria de Ciência e Tecnologia do Estado de São Paulo - 1979

- (5) SANTOS T. - Revolução Científica e Técnica e Capitalismo Contemporâneo - Petrópolis, Ed. Vozes, 1983.
- (6) CARVALHO, F. - Ciência aguarda dilúvio de informação - Jornal Estado de São Paulo, 25-3-1990, p. 24
- (7) CUNHA, M. B. - Bases de dados e bibliotecas brasileiras - Brasília, Ed. ABDF, 1984
- (8) PADCT/MCT/CNPQ - Doc. Sub-programa de Informação em Ciência e Tecnologia - Brasília, versão 7, junho de 1990.