

A Comercialização da Informação e do Conhecimento

Rodolfo Coutinho Xavier¹
Fernando Augusto Mansor de Mattos²

Introdução

O objetivo deste artigo é discutir o processo de comercialização da informação e do conhecimento nas sociedades capitalistas atuais, destacando o papel exercido pelas mesmas na acumulação de capital.

Em primeiro lugar, pretendemos estabelecer alguns parâmetros teóricos para compreender o mercado das produções científicas legitimadas, um mercado que não segue a lógica dos bens comuns, tangíveis e tradicionais. Em seguida, investiga-se o caso concreto das bases de dados privadas, detentoras do poder de comunicação e atribuição de reconhecimento à pesquisa científica.

Quando se usa a expressão Sociedade da informação, supõe-se que se quer dizer que se está mergulhado numa sociedade em que se produz e se vende muita informação, principalmente nas economias mais promissoras. Pode-se afirmar que as máquinas nada mais são que um conhecimento entranhado na matéria, e que o próprio conhecimento pode ser produzido, distribuído e comercializado, gerando preços, estoques e custos. Ademais, é inegável a expansão da bibliografia referente ao consumo, gestão, administração da informação.

Retoma-se à teoria de Marx. É claro que aqui se percebe com toda a clareza a invasão do capitalismo sobre a esfera da produção, disseminação e uso da informação geradora de conhecimento, ou seja, sobre a superestrutura, entendida como o conjunto articulado das idéias jurídicas, políticas, econômicas, religiosas, filosóficas de uma dada classe dominante que predominam na sociedade.

Contudo, é errado interpretar que o conhecimento só tem função ideológica dentro do sistema, pois seu papel é o de promover maior acumulação tanto no seu próprio mercado quanto

¹ Mestre em Ciência da Informação pela PUC de Campinas.

² Professor e pesquisador em Carreira Docente (regime de 40 horas) na PUC de Campinas. Atua no programa de pós-graduação em Ciência da Informação e no Centro de Economia e Administração da PUC Campinas. Mestre e Doutor em Economia pela UNICAMP.

em outros de produtos e serviços não diretamente vinculados ao conhecimento, mas necessitados de informações em decorrência da concorrência.

Também não se pode repetir o erro de alguns teóricos, – Sweezy e Baran (1966) – que, segundo Bolão (2000), teriam colocado o serviço da comunicação e da informação como trabalhos improdutivos, ou seja, não geradores de acumulação de capital. Não se deve reduzir o mercado do conhecimento a ideologia, a superestrutura ou a um setor de escape da acumulação de capital, já excedente em outros ramos como a indústria bélica.

Na verdade, Marx demonstra na sua clássica obra “O Capital” , apesar de não se ater diretamente ao mercado do conhecimento, que as comunicações são indispensáveis para o desenvolvimento da acumulação do capital, da venda da mercadoria, para vencer a esfera da circulação e realizar o lucro, pois não é possível que a produção consiga ser consumida sem a comunicação.

É inegável também que a comunicação científica é parte imprescindível de todos os elos da cadeia produtiva do conhecimento, a produção, a transferência da informação e o uso por cientistas e pesquisadores:

Garvey (1979) considera que a comunicação científica está envolvida tanto nas atividades de produção quanto de disseminação e uso da informação, em um processo que se inicia na concepção da idéia a ser pesquisada até o momento em que os resultados dessa pesquisa sejam aceitos pela comunidade científica (GARVEY, apud OLIVEIRA e NORONHA, 2006, p.2).

Ademais, esse pensador percebeu que o desenvolvimento técnico relacionado às comunicações era essencial à expansão e aceleração da acumulação de capital; pela redução do tempo de transportes e comunicação, os espaços são suprimidos e os mercados se ampliam: “o principal meio de abreviar o tempo de circulação é o progresso dos transportes e comunicações” (MARX, apud BOLÃO, 2000, p.25). “Economia de tempo, a isto se reduz finalmente toda a Economia”. (MARX, apud DANTAS, 2002, p.7).

Aplica-se a teoria marxista como faz Bolão (2000), e chega-se às mesmas conclusões: os transportes e comunicações são fundamentais no processo de reprodução do capital, que expande e cria mercados de consumo e fatores de produção para além das proximidades das indústrias – como faz a indústria do conhecimento -, que produz um mercado que tem uma lógica econômica própria e especial, que pelo trabalho concreto e não apenas abstrato gera produto, mercadoria, valor de uso e de troca. (BOLÃO, 2000, p.28).

Nesse sentido, pode-se afirmar que apesar da imensa contribuição das novas tecnologias da informação e comunicação na era digital e virtual, com respeito à facilitação da comunicação científica e maior acesso ao uso dos pesquisadores do conhecimento, a rede mundial de computadores não produz conhecimento independente de uma cadeia produtiva real, com relações sociais entre produtores, distribuidores e usuários, os quais necessitam do tempo e do espaço reais para problematizar, interpretar e analisar situações e dados concretos, um tempo de reflexão:

“O produto do conhecimento, fruto de idéias, deve ser documentado na rede e, assim, facilitar o seu acesso a outros sujeitos, mas não devemos esquecer-nos que a rede não permite a produção do conhecimento em si. O conhecimento surge de perguntas sobre as mais variadas situações, o que exige tempo de reflexão por parte do sujeito. O conhecimento é produzido a partir de análises, de interpretações de dados, o que pressupõe a reflexão. A rede pode ajudar numa fase da produção do conhecimento, em especial a coleta de dados, pois sua velocidade agiliza essa etapa. Porém, a indagação inicial requerida para a problematização do conhecimento, bem como a busca do significado para as respostas encontradas, só podem ser realizadas pelo sujeito produtor de idéias.”(BOSI apud BARRETO, 2005, p.121).

Isso não significa que os serviços prestados pelas bases de dados privadas e os provedores de acesso em linha sejam improdutivos, a saber, não geram riqueza. Atualmente, não se discute que qualquer serviço que satisfaça as necessidades da sociedade é uma atividade produtiva no âmbito Macroeconômico, como argumenta Hicks:

“a razão pela qual as pessoas podiam persuadir-se de que somente as atividades primárias eram produtivas baseava-se em que a palavra produção, quando empregada em sentido restrito, sugere a obtenção de um bem material, extraído pelo homem da natureza. De fato, uma grande parte da produção econômica é composta de bens materiais diretamente originários da terra; mas, a produção de todas as utilidades necessárias à vida humana em sociedade não se completa apenas com esse tipo de bens. Os comerciantes não são originariamente responsáveis pela produção dos bens com que transacionam; todavia, o comércio desempenha a útil atividade de reunir e oferecer os bens em locais que melhor satisfaçam às necessidades dos consumidores.[...]. Além disso, há numerosos trabalhadores que não estão diretamente ligados à produção de bens materiais. Os médicos, os professores e os atores são, todos eles, exemplos de produtores de serviços que satisfazem necessidades tão importantes quanto as proporcionadas por certos tipos de bens materiais. Eles contribuem com trabalho socialmente útil, são remunerados por suas atividades e devem, portanto, ser considerados como produtores, na ampla e correta aceção do termo.”(HICKS, 1960, p.p.43-44)

Não se pode aqui simplesmente aplicar o modelo marxista sem nenhuma ressalva ou reformulação. Deve-se orientar para aplicação adaptada dos conceitos fundamentais como mais valia e geração do valor, os quais não estão restritos à produção dos bens comuns. No momento,

não existe modelo que possa mensurar quantitativamente o valor e a mais valia extraída dos bens do conhecimento, mas isso não significa que eles não existam, trata-se da maior mais valia e valor produzidos pelos trabalhadores do conhecimento, e apropriados pelo capital.

Vários são os problemas de tornar a informação uma mercadoria: seu consumo não exclui ninguém pelo fato de alguém já estar utilizando-a, não pode ser dividida e nem especificada por consumidor, não se pode mensurar o seu valor antes mesmo de adquiri-la. Isso implica dizer que o conhecimento e a informação não excluem a aquisição de ninguém porque alguém está consumindo-os, pode-se conquistar alguma exclusividade de acesso à informação e conhecimento, mas ela será temporária e sempre estarão em domínio público, não se pode medir sua utilidade sem antes tê-los adquiridos.

Outros pontos podem ser mencionados, como o fato de que ambos funcionam como custo fixo na empresa, pois depois de adquiridos podem ser usados freqüentemente, além da sua utilização ampliá-los, mas paralelamente eliminá-los. Todos esses fatores fazem do conhecimento e da informação algo de difícil comercialização, mas que incide decisivamente na produção capitalista.

A informação é comunicável e de certo modo externa ao seu possuidor, mas o conhecimento é pessoal, singular e não transmissível. Cada nova informação é interpretada de modo diferente por diferentes pessoas, pois elas têm distintos graus e tipos de conhecimento. A aquisição do conhecimento depende da capacidade cognitiva do sujeito que o recebe.

Além disso, produzir conhecimento exige intuição e criatividade do indivíduo, o que torna o conhecimento algo mais subjetivo e temperamental.

A teoria do equilíbrio geral não é capaz de incluir, em si, os processos de inserção e assimilação da informação. A ausência do controle de sua distribuição, a atuação não-rival como se fosse um bem público, a falta de avaliação possível antes mesmo da aquisição, dificultam sua apreensão pelo sujeito.

Mas a informação por si só não é útil, faz-se necessário um conhecimento a partir do qual a informação se torna útil. Assim, quem efetivamente pode aproveitar-se de uma informação é quem já possui um conhecimento prévio. Isso coloca os agentes econômicos em pé de desigualdade, uns com relação aos outros, além dessa defasagem levar muito tempo e dinheiro para ser suprimida (POSSAS, 1998, p. 03).

Outra característica do conhecimento, intratável pela teoria ortodoxa econômica, é a cumulatividade. Quanto mais se conhece mais se está apto para adquirir novas informações, e produzir mais conhecimento. Quando se aprende coisas novas isso abre possibilidades de novos conhecimentos. Escolher conhecer alguns assuntos é não escolher outros, o que faz com que só possam ser retomados a partir de um gasto adicional. A dificuldade de transmissão e de aquisição do conhecimento torna essa mercadoria complicada de ser comercializada.

Nesse sentido, o conhecimento tende a se concentrar num lugar específico, quem o detém consegue transmiti-lo em parte, especializações e pesquisas são promovidas localmente em instituições de ensino, hospitais, firmas, etc. Grande parte desse conhecimento não é traduzido em palavras.

Tecnologia, mercados, fornecedores, consumidores, todos esses elementos fazem parte da pesquisa das empresas, o que as torna um imenso repositório de conhecimento.

O conhecimento está nas organizações, as quais, com menor ou maior facilidade, transmitem e acumulam conhecimento no seu cotidiano.

Crescimento, desenvolvimento e distribuição de renda dependem necessariamente da distribuição social do conhecimento, pois é esse processo disseminador que aumenta o poder aquisitivo da população.

No que se refere à comercialização do conhecimento, do ponto de vista econômico ele é efêmero e passível de rápida obsolescência. Se não o utiliza ele se perde, além de quase na sua totalidade rapidamente se tornar obsoleto. Mercados e tecnologia mudam, o que exige mudança também no conhecimento, pois se alteram as situações em que é utilizado.

A mercadoria informacional passa por uma cadeia produtiva assim como qualquer outra mercadoria. São variáveis independentes da sua função de produção: recursos humanos e materiais, de capital e financeiro. Todos esses insumos combinados devidamente e nas proporções adequadas produzirão a mercadoria informação. Cada um desses recursos gera um determinado gasto ou dispêndio na produção, o que sugere que a combinação que otimiza a redução de custos e desperdícios desses recursos é a que reduz o custo unitário da mercadoria e do estoque, por conseguinte, são bens que devem estar mais em circulação do que apenas armazenados para viabilizarem o lucro.

Oferta, demanda, custo, preço e valor são temas inevitáveis quando se aborda o mercado do conhecimento. A informação é uma mercadoria simbólica, não se esgota com o seu consumo,

sua transferência e aquisição sempre é compartilhada entre emissor e receptor, a duração e intensidade do seu consumo será determinada por sua qualidade e valor cultural, assim como a eficiência de seu suporte material.

Como o produtor ou distribuidor nunca perde seu bem quando o dissemina, ele pode disponibilizar e vender a mesma informação quantas vezes quiser, mas a aquisição pela demanda depende de suas competências cognitivas, mesmo que a informação seja abundante, além de precisar que o consumidor a veja com relevância e utilidade para ser adquirida.

Não há uma unidade de medida exata para quantificar uma informação, como existe para um bem material comum, tão pouco seu custo tem relação com o seu preço e com o seu valor. A produção da informação científica não segue a regra da dependência da demanda para crescer ou diminuir, o desenvolvimento da produção do conhecimento se dá independentemente de sua venda ou lucratividade, o que torna esse mercado mais complexo do que os demais, que dependem dele na criação de inovações tecnológicas.

A disseminação da informação eletrônica pelas grandes bases de dados

A Internet ainda continua sendo um campo de novas oportunidades de negócios, além de conter nela várias possibilidades novas de mercados, ajuda a baratear a busca por empregos e serviços, com maior facilidade e rapidez. A ausência de informação é um problema em todos os mercados. Ademais, a Internet estabelece novas bases para uma nova relação entre informações e empresas, conhecimento e bases de dados privados, tendendo a ampliar a produtividade.

Os negócios da Internet, como as bases de dados, precisam de planos de negócios diferentes do que eram feitos no passado para obter sucesso. O montante de capital empreendido, para a produção de um grande acervo eletrônico, é monumental, o que leva os prejuízos elevados por vários anos, além de recrudescer as barreiras à saída, ou seja, não se recuperara fácil o capital investido em um negócio fracassado. Uma rede mundial de computadores interconectados, pode colocar várias e possíveis fontes de informação simultaneamente para toda demanda, várias pessoas numa mesma fonte de informação ao mesmo tempo.

É claro que o acesso às redes estimula o trabalho em equipe. A possibilidade de todos terem acesso aos mesmos dados e interagirem facilmente em sua utilização favorecem os esforços coletivos. Ao mesmo tempo, a comunicação por meio de redes pode ajudar a integrar o grupo. De fato, serve para ampliar a influência do grupo tanto em termos quantitativos quanto em extensão geográfica (MEADOWS, apud ARELLANO; MORENO, 2005, p. 80).

Teoricamente, a Internet elimina as limitações de um acervo único, disponibilizando toda a informação desejada de uma vez só. Serviços prestados no seu interior, como os dos fornecedores de acesso on-line aliados as grandes bases de dados, poderiam gerar um salto de competitividade e amenizar as desigualdades entre cientistas e pesquisadores do primeiro e do terceiro mundo.

Mas o que faz de útil e necessário às bases de dados? As bases de dados resolvem problemas com relação à recuperação do conhecimento num ambiente informacional cada vez mais complexo. Devido à explosão bibliográfica, os vários tipos de documento, a redução das barreiras geográficas e outras que impediam o fluxo da informação, a velocidade das transformações do conhecimento e seu estado atual de interdisciplinaridade, dificuldades de representação e recuperação surgem para todos os documentos.

Atualmente é impossível conter todo o conhecimento em um único lugar para consulta, mais fácil será se comunicar com todas as bases de dados que podem ter essa informação, o acesso gerado pelas tecnologias de comunicação e informação deram uma solução ao problema, ainda que não a ideal.

Muitas definições levam em consideração o aspecto computacional das bases de dados, entretanto, há definições que negam essa diferença entre as bibliotecas tradicionais e as bases de dados on-line, frisando que se trata de um conjunto de informações, organizadas de uma maneira que possam ser recuperados:

Uma coleção de itens sobre os quais podem ser realizadas buscas com a finalidade de revelar aqueles que tratam de um determinado assunto. A base de dados consiste em artefatos, como livros (o acervo de uma biblioteca é uma base de dados com certeza), ou registros que representam os artefatos, como, por exemplo, registros bibliográficos constantes de páginas impressas, de fichas ou de meios eletrônicos (LANCASTER,1993, p.305).

Chama-se base de dados um tipo de mecanismo institucional, que garante a perpetuação e a transmissão do conhecimento registrado do passado.

As bases de dados têm elevados custos fixos – custos que não variam com a produção - e baixos custos marginais – custos gerados pela produção de uma unidade a mais de produto -. Tanto a produção de informações científicas, quanto sua representação, estruturação, linguagem e recuperação produzida por softwares avançados, utilizados pelas bases de dados são dispendiosas, pois exige elevada aptidão humana em pesquisa, um trabalho que envolve muitos

especialistas. Contudo, depois disso a reprodução de um artigo, dissertação ou resumo é extremamente barata, o que traz uma economia de escala considerável do lado da oferta:

“Uma característica é o fato de que um negócio de informações ou de Internet geralmente tem altos custos fixos e baixos custos marginais (ou o custo de produzir uma unidade extra). É caro produzir a primeira unidade, seja um *software*, um serviço como apostas on-line ou uma mercadoria de informação mais amplamente definida, como uma droga de pesquisa intensiva ou um avião de projeto intensivo. Isso porque todos os custos são feitos de antemão e exigem muita habilidade humana. Entretanto, cópias subsequentes são muito baratas ou mesmo livres de custos no caso de cópias de um programa ou de um livro *on-line*. A reprodução pode ser feita de maneira barata por computadores ou outras máquinas. Em outras palavras: os negócios de informações são inerentemente caracterizados por rendimentos ascendentes a avaliar no lado da oferta.” (COYLE, 2003, p.140).

Pode-se obter uma vantagem de liderança competitiva, ao ofertar para o consumidor um padrão tecnológico de busca e recuperação avançadas, mas isso se tornará caro caso os consumidores migrem para um concorrente – os novos repositórios institucionais e temáticos gratuitos por exemplo -.

Em decorrência dos custos fixos elevados que as bases de dados têm, tanto mais é vantajoso e desejável reproduzir e vender em larga escala, o que torna a reputação crucial nesse negócio.

Todas essas peculiaridades alteram a política de preços das bases de dados; seus usuários não estão interessados no quanto essas empresas gastaram para conseguir organizar, armazenar e disponibilizar tanta informação, nem quanto de tempo será necessário para recuperar o investimento feito, ou seja, igualar pelo menos a receita total com o custo total. Contudo, a reprodução de artigos, textos, dissertações, referências e resumos em meio eletrônico é muito barata, o que significa que há sempre a possibilidade de se ofertar e vender cada vez mais.

A prática do mark-up, ou seja, estimar uma margem de lucro percentual em cima do custo unitário não é adequado às bases de dados, pois esses custos são muito baixos. Seria mais conveniente cobrar de acordo com o valor que cada cliente atribui à informação, o que permitiria que aqueles que têm melhores condições aumentassem a margem de lucro, enquanto aqueles que só esporadicamente utilizam o serviço e pagam pouco também pudessem ser incluídos. As vantagens para as bases de dados seriam o aumento das vendas, a elevação da economia de escala na oferta e a redução da capacidade ociosa do repositório on-line.

Além disso, novos modelos de negócios são exigidos das editoras científicas tradicionais, na medida em que os repositórios digitais de acesso gratuito à informação, temáticos e

institucionais, prestam serviços equiparados em qualidade aos serviços pagos dos provedores de fornecimento de acesso on-line e de suas bases de dados privadas "os novos serviços de indexação disponíveis gratuitamente têm um desempenho semelhante a seus concorrentes pagos em termos de precisão e conveniência" (MORENO; ARELLANO, 2005, p. 83).

Descontos promocionais para clientes assíduos, customizar produtos e serviços, rastrear o comportamento dos consumidores para atender mais adequadamente suas necessidades, são recursos interessantes que podem ser usados pelas bases de dados.

O rendimento econômico de uma base de dados depende da eficiência proporcionada na procura de seus documentos. Esta, por sua vez, está condicionada ao período de tempo gasto pelo usuário nessa aquisição, a taxa de precisão, a taxa de revocação, ao refinamento dos sistemas de interrogação e de indexação. Tanto a linguagem de descrição dos dados quanto de comando ou busca, contém aspectos morfológicos e sintáticos, palavras-chave. Ambas devem estar harmonicamente conciliadas, pois enquanto a descrição dos dados é importante na entrada de uma base de dados, a linguagem de busca é imprescindível para uma boa saída ou recuperação daquilo que se deseja.

Uma pesquisa mostra que se quiser procurar até a exaustividade um determinado assunto em 40 bases de dados, nas 13 primeiras se teria encontrado cerca de 80,6% de toda informação que é oferecida nessas 40 bases de dados sobre aquele assunto, ou seja, em apenas 32,5% delas. Ao passo que só 19,5% das informações relevantes estarão dispersas em 67,5% das bases de dados, o que demonstra que a exaustividade, ou seja, procurar um assunto em todas as bases é perda de tempo (CHAUMIER, 1986, p. 26 e 27).

Essa característica da mercadoria informação faz com que a reputação e o marketing cresçam em necessidade para as bases de dados, pois os consumidores somente arriscarão comprar uma informação valiosa com um vendedor que seja seguro, idôneo. A reputação inibe a redução de intermediários, freqüentemente se está querendo que alguém garanta a compra, ou algum órgão ou alguém que se possa reclamar; além disso, muitas vezes se consultam pessoas físicas para saber sobre preços, do que programas na Internet com essa função.

O excesso de informações disponibilizadas hoje na rede exige não só uma seleção rigorosa e sistemática do acervo que se quer construir, mas também o caminho mais rápido e adequado para se chegar exatamente na informação que se quer, uma busca eficiente. São esses

pontos que as bases de dados de conhecimento têm a seu favor, permitindo que os usuários não se percam no volume de informação da Internet.

OCDE – Organização de Cooperação e de Desenvolvimento Econômico -, OMS – Organização Mundial de Saúde - *World Bank* – Banco Mundial -, universidades e centros de pesquisa são organismos públicos ou privados, nacionais ou internacionais, produtores de enormes acervos de dados, os quais são processados e distribuídos por grandes bases de dados privados on-line. Os governos gastam em média 10 milhões de dólares para recuperar nessas bases de dados às informações financiadas pelo próprio governo, à custa do dinheiro público e da sociedade. De referências bibliográficas a artigos, dissertações e teses, todo conteúdo indexado nessas bases de dados foi financiado com dinheiro público. Contudo, a espanhola PIC e a ECHO – base de dados da Comunidade européia – são exemplos de bases de dados públicas, com custo reduzido ao seu acesso – apenas o pagamento da comunicação -. Contudo, quase a totalidade das bases de dados on-line é privada, e estão nas mãos dos norte-americanos. E cerca de 77% das bases de dados está em inglês. (FERIGLE, 1992, p.240).

São exemplos de fornecedores de acesso on-line a essas bases de dados: o *Dialog*, STN internacional, OCLC, *DataStar*, *LEXIS/ NEXIS*, *Questel-Orbit*, *Ovid Technologies* e *H.W.Wilson*. (CENDÓN, 2003, p.227). Algumas vezes as próprias bases de dados são fornecedoras do acesso on-line, mas o que ocorre geralmente é o predomínio do monopólio da informação dos grandes distribuidores.

Nesse contexto, a liderança monopolista norte-americana pode ser aferida com alguns dados: A Comunidade européia tem menos cinco vezes bases de dados com textos na íntegra que os EUA, a Comunidade européia tem seis vezes menos bases de dados estatísticos que os EUA, o EUA produz mais de dez vezes bases de dados que a Comunidade Européia, em 1990. (FERIGLE, 1992, p.240).

Nos anos 80, somente 30% dos vendedores de bases de dados não eram norte-americanos, sendo que os demais 70% do mercado era dominado pelas grandes transnacionais estadunidenses como a *Chase Econometrics*, sucursal do *Chase Manhattan Bank*, e o *Data Resources Inc.* o *Dialog*, hoje da *Thonsom*, tido como o maior distribuidor de informação científica do mundo. Além disso, existem as companhias que alugam linhas de telecomunicação, para vender serviços on-line de acesso às bases de dados, são elas: A *Tymnet* com 32% do mercado e a *SprinNet* com 39%, pelos menos até 1990. (FERIGLE, 1992, p. 240-241).

Nos anos 70, para tentar minimizar a dificuldade de acesso dos países subdesenvolvidos, pois não tinham a tecnologia mais eficiente, o *Dialog* começou a disponibilizar seus conteúdos em tecnologias ópticas – CD-Rom -, evitando assim os custos das telecomunicações, nessa época ainda elevados para países emergentes.

As bases de dados eletrônicas via on-line e as ópticas off-line foram criadas para facilitar o acesso à informação científica; contudo, isso nem sempre acontece.

Usuários reclamam que o índice de revocação de suas buscas dentro de um grande provedor em suas bases de dados é muito baixo, não se chega a utilizar nem metade das informações recuperadas, o que onera financeiramente as buscas e causa grande perda de tempo.

As causas desses mal-entendidos nas buscas são as complexas e inter-relacionadas variáveis que as bases de dados tem que levar em consideração, são elas: objetivo do usuário, vocabulários controlados, paradigmas eficazes de metadados, escopo da base de dados, diferentes tipos de indexação dependendo da base de dados e outras.

Uma busca com uma palavra-chave específica nessas bases certamente recuperará algum documento pertinente à necessidade do usuário. Contudo, apesar do índice de precisão ser elevado nesse caso, ou seja, a relação entre a quantidade de documentos pertinentes recuperados e a quantidade total de documentos recuperados, o índice de revocação é baixo, a saber, a quantidade de documentos recuperados dividido pela quantidade de documentos pertinentes nessas bases de dados. (MEADOWS, 1999, p. 232-233).

Utilizar uma palavra-chave menos precisa fatalmente pode aumentar as chances de se recuperar outros documentos pertinentes. Contudo, o que se observa é a recuperação considerável de documentos não pertinentes, uma redução vertiginosa no índice de precisão, sem a melhora desejada no índice de revocação.

Existem várias maneiras de se buscar informação nas bases de dados do provedor *Dialog*, mas as duas mais simples seriam: a busca registro por registro a partir de uma palavra definida pelo usuário, pelo procedimento do arquivo-invertido, o qual se assemelha a uma procura num índice com palavras-chaves, data, autor, título do periódico, para se ter acesso ao conteúdo do documento.

As bases de dados inseridas no provedor *Dialog* deveriam ser avaliadas pelas seguintes critérios: abrangência de cobertura sobre o assunto, assiduidade das atualizações, praticidade do uso e predileção pelo usuário, qualidade da saída das informações, performance da linguagem de

indexação nas recuperações, gastos de acesso e saída e quem paga, qualidade da documentação e dos sistemas de classificação, possibilidades de acesso e suportes de armazenamento, estrutura da base de dados e formato dos registros. (ROWLEY, apud VALENTIM, 2001, p. 133-134).

A empresa *Dialog* contém 520 bases de dados privadas, classificadas como: bibliográficas, diretórios, financeiras, numéricas e de texto completo.

As bases de dados bibliográficas são aquelas que indexam abstracts e referências bibliográficas dos artigos, muito utilizadas pelas cientistas. Podem trabalhar com vários tipos de documentos – *papers*, simpósios, seminários, patentes - ou com um único tipo.

A tendência é que essas bases bibliográficas não indexem apenas os resumos e as referências bibliográficas, mas os artigos na íntegra. Em 2001 a *Dialog* contava com 4 mil periódicos de texto completo, difusos em suas bases de dados.

Os critérios para a seleção e publicação de artigos nas bases de dados agrupadas sob a *Dialog* variam de uma para outra, mas há um núcleo comum que deve ser satisfeito: respeito à periodicidade, o trabalho tem que passar pela análise e aceitação de uma comissão editorial, palavras-chaves e resumos tem que ser feitos. A última exigência se torna fundamental para uma recuperação eficaz.

Ilustrativamente, um exemplo de bases de dados bibliográfica inserida no provedor *Dialog* é a *Medline*, com um modelo de metadados de 27 campos e contendo vários tipos de documentos. Por outro lado há a *Dissertation Abstracts Online*, que indexa somente dissertações.

As bases de dados de diretórios são bastante variadas, porque satisfazem uma gama muito grande de tipologias documentais, absorvendo todas as informações referentes a cada tipo de documento. Ademais, seu conteúdo pode variar enfaticamente, é o caso da *Marquis Who's Who*, que indexa informações sobre pessoas e faz parte das bases de dados do *Dialog*. Já contém mais de 790 mil registros de pessoas, e possui uma estrutura de metadados complexa com 37 campos diferentes, para dar informações detalhadas sobre todas as pessoas.

Um exemplo de bases de dados de textos originais integrais é a *U.S. Patents Fulltext*, inserida no provedor *Dialog*. São bases que detém para a recuperação textos completos, patentes, notícias de jornal, normas técnicas, artigos de revistas, etc. Os metadados dessas bases mudam muito, porque contemplam textos integrais muito diferentes entre si.

No caso da base de dados de patentes citada, ela atua desde 1971 nos EUA, mas apenas em 1974 começou a indexar o texto completo das patentes. Já têm mais de 2 milhões e 245 mil

documentos à disposição, além de mostrar o desenho da patente, seu desenvolvimento na criação, o procedimento legal para reservar o uso e a certificação pelo poder público. Essas duas últimas informações não estão contidas no documento original, derivam dos metadados da base, o que mostra a importância de agregação de valor que essas bases produzem sobre os documentos, gerando novas informações.

Nessa base, as patentes são exaustivamente cobertas por 43 campos de metadados, permitindo uma busca mais precisa e minuciosa, caso seja do interesse do usuário.

As bases de dados numéricas oferecem dados quantitativos sobre vários assuntos, além de textos completos. Diferentes textos na íntegra e dados exigem diferentes formas de tratamento da informação, no que se refere à representação e a descrição.

Para ilustrar, há uma base numérica no *Dialog* chamada *Population Demographics*, com mais de 50.457 registros de informação. Vários dados sobre a população americana, coletados em 1990, são disponibilizados: poder de compra, empregos, informações industriais, ocupação, educação, renda, família, etc.

Essa base de dados ainda disponibiliza vários estudos sobre projeções de dados como: econômicos, territoriais, planejamento estratégico, estatísticos, de situação local, de objetivo de mercado e outros. Há um modelo de metadados com 36 campos informacionais, o que permite recuperar informações sobre população, região, ocupação, etc. Bastante útil e relevante para as empresas, para que sua estratégia de mercado seja bem sucedida.

Dados de organizações sobre investimentos, balanços, bolsa de valores, indicadores econômicos, cotações na bolsa de valores e outros, são indexados nas bases de dados financeiras. Mostram desde dados referenciais até textos na íntegra. Seu modelo de metadados é específico e adequado ao conteúdo disponibilizado, por exemplo, os balanços das empresas.

Por exemplo, a base de dados *Extel International Financial Cards*, incluída no *Dialog*, detém informações empresariais financeiras, de cotações e textuais de várias partes do mundo como: Malásia, Singapura, Japão, Reino Unido, Austrália, América do Norte, Europa e outras. Já possui mais de 7.675 registros, e procura manter um histórico sobre os dados de 5 anos.

Aplicação em fundos, informação das diretorias das empresas, dados das empresas, organogramas das organizações, lucros e perdas, balanços, relatórios de gerência e financeiros, inversões em ações, informações sobre bancos e auditoria, são analisados e colocados à disposição nessa base.

O modelo de metadados com 54 campos informacionais usado nessa base exaure as possibilidades de busca, sendo bastante completo e satisfazendo as necessidades dos usuários. A descrição e representação dos dados são bastante minuciosas, além da completude das informações, isso dá alta credibilidade a base.

As bases de dados que, por meio de seus metadados, acrescentam informações não disponíveis no conteúdo original, geralmente cobram mais caro o acesso, o que mostra uma tendência no mercado da informação científica.

As grandes empresas transnacionais são as que mais se beneficiam das bases de dados privados on-line, pois detêm redes internas de comunicação que permitem uma boa velocidade de acesso às informações científicas. Informações importantes sobre o ambiente externo como patentes, recursos naturais, indústrias, inovações tecnológicas não são monopolizadas nem por governos e nem por empresas privadas, mas colocadas à disposição nessas grandes bases de dados on-line. Esses grandes distribuidores de informação, tanto mais vendem e fazem da informação uma mercadoria, mais cobram pelo acesso aos seus serviços, um processo em franca expansão, basta verificar que no Dialogue, em 1990, as bases de dados comerciais cresceram 30% nesse provedor sobre o total.(FERIGLE, 1992, p.241).

A tendência a se passar de publicações impressas para publicações eletrônicas promoveu um processo de oligopolização dentro da Internet, o qual já existia e vinha se fortalecendo no campo da impressão. De 1997 a 2000, houve uma grande fusão, na qual as cinco maiores editoras científicas foram incorporadas pelas suas concorrentes. Isso gerou um aumento significativo nos preços, os quais aumentam vertiginosamente a cada ano.

Tabela 1 – Aumento dos preços de algumas revistas
(em US\$), 1989-2001

Título da revista	Preço em 1989	Preço em 2001	Aumento percentual
American Journal of Physical Anthropology	540	1.490	176%
Chemical Engineering Science	650	4.306	562%
Clinical Psychology Review	75	984	1.212%
Communications on Pure and Applied Mathematics	280	2.190	682%
Developmental Biology	936	4.715	404%
Energy Conversion and Management	305	3.019	889%
IEEE Proceedings: All-Society Package	5,335	29,995	462%
Journal of Physics A	1,124	6,030	436%
Educational Research Quarterly	275 (2000)	505	184% (em um ano apenas)

Fonte: Revista Brasileira de Pós – Graduação.

Tais empresas praticamente proprietárias do conhecimento impunham seu poder na venda para as universidades, sem dar nenhuma chance de renegociação de preços para seus clientes. Ironicamente, os próprios produtores de conhecimento são os que pagam mais caro para adquiri-los depois, em face da apropriação dessas grandes bases de dados privadas, que exponencialmente aumentaram o preço de suas publicações. Por conseguinte, o número de revistas científicas nas universidades diminuiu, mas seu custo aumentou drasticamente.

Também confirma Tenopir e King essa absurda elevação no preço

É praticamente indiscutível que os preços das revistas aumentaram de modo considerável nos últimos 20anos. Estima-se que os preços das revistas acadêmicas e científicas norte-americanas aumentaram de 39 dólares no ano de 1975 para 284 dólares em 1995. Portanto, os preços médios aumentaram em um fator de 7,3, ou 2,6 vezes em uma taxa monetária constante. (TENOPIR; KING, 1998, p. 5).

Na área de medicina também o preço dos periódicos cresceu celeremente. A revista Blackwell de medicina aumentou em 184%, entre 1990 e 2000, enquanto na França o custo de vida teria aumentado 33%, apenas. (ANDRÉ, apud KURAMOTO, 2006, p. 3).

A situação chegou a um ponto limite para as universidades e os Estados, que já não conseguiam atualizar suas publicações e nem aos governos comprarem essas coleções com dinheiro público, o mesmo que financia a produção científica. Chegada à crise até nas universidades norte-americanas, principalmente as particulares, iniciou-se um movimento de Estados e universidades para outras soluções viáveis com preços justos.

Um problema antes restrito aos bibliotecários e administradores de universidades passou a ser tema de pesquisa central para as universidades. Mas isso não ocorreu só no Brasil, Harvard deixou de assinar cem revistas do provedor Reed Elsevier, porque não teve êxito em renegociar seus preços abusivos (SOARES, 2004, p.21).

As conseqüências desse processo de privatização da informação são diversas: a informação deixa de ser um bem acessível às classes sociais com níveis de renda diferenciados, para ser submetida a leis do mercado; somente os países, as empresas e as pessoas mais ricas terão acesso às informações de qualidade, o que tanto aumentará a distancia social entre ricos e pobres quanto ampliará o poder dos economicamente incluídos; os países pobres serão dependentes da tecnologia dos países ricos, enviando dados brutos e importando dados elaborados; empresas que não tiverem acesso a essas grandes bases de dados, ou serão dependentes da lógica econômica daquelas que têm, ou não perpetuarão; pesquisadores que não tiveram acesso a esses acervos privados terão mais problemas em levar adiante suas pesquisas do que aqueles que têm; somente as informações de preços elevados serão contempladas nesses grandes distribuidores privados, deixando a área de humanidades de lado, isso resultará numa brutal perda na História da humanidade.

Tais conseqüências da privatização, tanto no campo da infra-estrutura quanto do conteúdo propriamente, exigem a atuação governamental sobre novas políticas de acesso à informação, o que deveria estar contemplado ao lado do direito a educação para todos. Caso não se reverta esse quadro a tempo tem-se um futuro pouco promissor, produzido pelo acesso limitado, caro e complicado das informações.

Considerações finais

A Internet poderia ser um instrumento de socialização e democratização da informação. Por meio dela, se poderia exercer plenamente a cidadania, tanto política quanto científica, resolver problemas com pesquisas mais práticas e precisas, reduzir a desigualdade de acesso à informação e social, econômica e política entre pessoas, países e empresas. Poderia se dar mais oportunidade para instituições de ensino e pesquisa, que têm menor quantidade e qualidade de produções científicas, disseminando informações geradoras de conhecimento para todos os cientistas e pesquisadores do mundo, sem exceções.

Sabe-se que não basta toda a sua utilidade e relevância, o reconhecimento de seus serviços de comunicação e informação e o poder de compra do usuário. É preciso eliminar as barreiras técnicas, de conhecimento do funcionamento da Web, de vários aplicativos diferenciados e outras línguas com a finalidade de fazer o crescimento do acesso à informação e as inovações e obsolescências do conhecimento e da tecnologia mais velozes do que são.

Contudo, hoje no estágio de desenvolvimento de um capitalismo monopolista, o grande oligopólio das bases de dados científicas apenas ampliou alienação e mais-valia, a submissão da transmissão do conhecimento pelos critérios econômicos de rentabilidade, impacto e vida útil, a elevação dos preços e a conseqüente perda de acesso da maioria da população à educação de qualidade.

Todos devem se unir contra esses males, os quais atingem todos os conhecimentos, confrontando e retaliando com medidas enérgicas as grandes bases de dados privadas, quais sejam: recusando a compra das informações pelos preços abusivos que praticam, criando formas alternativas de transmissão e recuperação dessas informações por bases de dados públicas, editoras acadêmicas, que as disponibilizam por um preço muito menor e asseguram maior acesso para os cidadãos.

A exclusão digital é somente um item de toda a exclusão econômica e social que se tem em curso no Brasil (Mattos, 2003). Não se trata apenas de possuir um computador com banda larga, mas de dar uma educação eficiente que através da reflexão, organização e reformulação da produção e distribuição da riqueza simbólica, possa efetivamente constituir profissionais, cientistas, pesquisadores trabalhando em benefício da sociedade.

A proposta viável para essa situação, deve levar em consideração três aspectos, necessários em qualquer decisão governamental: o político, o econômico e o legal. O primeiro diz respeito à vontade política, a saber, a urgência que legislativo e executivo têm para formular e colocar em exercício leis que possibilitem um maior acesso à informação pelos indivíduos; o segundo leva em consideração que a verba destinada à educação, pesquisa e acesso à informação deve ser ampliada, o percentual do PIB brasileiro destinado a esses quesitos é muito baixo comparado a outros países emergentes; a terceira seria aplicar leis e regulamentações que coíbam o poder dos grandes oligopólios, democratizem a informação sem lesar o direito autoral e de propriedade intelectual. Qualquer projeto que almeje ser aprovado deverá necessariamente para por esses três aspectos, os quais podem não ser suficientes, mas são indispensáveis na construção de uma solução.

Ao lado disso, por ser tratar de uma Política Nacional específica de informação, cabe também envolver no debate tecnólogos e cientistas da informação, os quais conhecem com mais profundidade as particularidades do conhecimento e da informação, no que toca seu registro, armazenamento, tratamento e recuperação.

Atualmente o movimento ao acesso livre à informação científica, envolvendo o Instituto Brasileiro de informação para Ciência e Tecnologia – Ibict -, apresenta algumas propostas: construção de repositórios digitais gratuitos, arquivos abertos – *open archives* – e *Eprints*.

A atitude é louvável, mas esbarra em alguns problemas: a necessidade da revisão pelos pares para a indexação nos *open archives*, o que nem sempre acontece, e o fato de que os repositórios gratuitos devem buscar deter o mesmo poder de reconhecimento que os provedores e as bases de dados privadas internacionais, mas isso vai demorar um pouco se for possível.

Os aumentos de preços exagerados para os orçamentos das universidades das assinaturas das revistas científicas, comprovam o conflito de interesses entre o oligopólio privado da informação científica e a comunidade científica, que, ao mesmo tempo, demanda os serviços das bases de dados privadas e produz o conteúdo das mesma, gerando uma crescente exclusão no mercado de produção e disseminação do conhecimento, confirmando o que foi salientado na hipótese deste trabalho.

De acordo com a lógica deste mercado, para estar incluído é preciso pagar “duas vezes”: por um lado, pagar (via impostos) o custo do Estado para financiar as pesquisas; em segundo lugar, arcar com os recursos cobrados para a assinatura das bases de dados. Deve-se destacar,

contudo, que, no primeiro caso, todos os cidadãos arcam com as despesas decorrentes do financiamento público e, no segundo caso, apenas as pessoas físicas e jurídicas mais abastadas têm recursos para realizar o acesso às informações organizadas nas bases de dados privadas.

Não se pode desprezar a produção intelectual de países emergentes como Índia e China, os quais não estão alinhados com as grandes bases de dados privadas.

Em benefício do desenvolvimento da ciência, o que contraria o poder de mercado desse grande oligopólio, algumas medidas devem ser tomadas: envolver entidades políticas nessa briga, unir as pessoas interessadas na quebra do monopólio da informação por essas grandes empresas, juntar e conjugar os países numa política de retaliação e confronto as práticas abusivas dessas empresas, principalmente os emergentes, aproveitar os blocos dos países emergentes já montados, como o G20, boicotar a compra de informações disponibilizadas nessas grandes bases de dados, além de seus editores e colaboradores.

Referências

BARRETO, A. M. Informação e conhecimento na era digital. Revista **Transinformação**, Campinas, 17(2): 111-122, maio/ago., 2005.

BOLAÑO, C. **Indústria cultural, informação e capitalismo**. São Paulo: Hucitec, 2000.

CAMPELLO, B. S.; CENDÓN, B. V.; KREMER, J. M.; **Fontes de Informação para Pesquisadores e Profissionais**. Belo Horizonte: Ed. UFMG, 2003.

CHAUMIER, J. **Systemes marche et technologies D' information**. Paris: Editora Entreprise Moderne d' Édition, 1986.

COYLE; D. **Sexo, drogas e economia**. São Paulo: Editora Futura, 2003.

DANTAS, M. **A lógica do capital-informação: a fragmentação dos monopólios e a monopolização dos fragmentos num mundo de comunicações globais**. Rio de Janeiro: Editora Contraponto, 2002.

EDEY, H, C; PEACOCK, A. T. **Renda Nacional e Contabilidade Social**, Zahar, Rio de Janeiro, 1963. Capítulo III, item 1, p.51.

FERIGLE, S. A. La privatizacion em los sistemas telematicos de informacion. **Revista Espanhola de Documentação Científica**. v. 15, nº3, 1992.

HICKS, J. R. **The Social Framework**, Oxford University Press, Oxford, 1960. CapítuloII, item 1, p.p. 43-44.

KURAMOTO, H. Informação científica, proposta de um novo modelo para o Brasil. **Ciência da Informação**. Brasília, V.35 , n.2. Brasília, 2006.

LANCARSTER, F. W. **Indexação e Resumos: teoria e prática**. Brasília, Briquet de Lemos, 1993.

MARX, K. **O Capital: Crítica da Economia Política**. São Paulo: Paz e Terra, 1980.

MATTOS, F. A . M. (2003). Exclusão digital e exclusão social: elementos para uma discussão”. **Revista Transinformação (programa de pós-graduação em Ciência da Informação da PUC de Campinas)**. Volume 15, número 3. Edição especial de 2003.

MEADOWS, A.J. **A comunicação científica**. Brasília: Briquet de Lemos Livros, 1999.

MORENO, F. P; ARELLANO, M. Á. M. Publicação científica em arquivos de acesso aberto. **Arquivística.net**. Rio de Janeiro, v.1, n.1, p.76-86, jan./jun., 2005. Disponível em: <[http:// www.arquivistica.net](http://www.arquivistica.net)> Acesso em: 02/12/2006.

OLIVEIRA, E. B. P. M; NORONHA, D.P. **A comunicação científica e o meio digital**. Departamento de Biblioteconomia e Documentação, Escola de Comunicações e Artes, Universidade de São Paulo. São Paulo, 2006.

POSSAS, M. S. Concorrência e elementos subjetivos. **Revista de Economia Política**, São Paulo, vol. 18. nº4 (72), out./dez., 1998.

SOARES, G.A.D. O Portal de Periódicos da Capes: dados e pensamentos. **Revista Brasileira de Pós-Graduação**, Brasília, n.1, p. 10-25, julho, 2004.

TENOPIR, C.; KING, D.W. A publicação de revistas eletrônicas: economia da produção, distribuição e uso. **Ciência da Informação**, Brasília, v.27, n.2, p. 176-182, maio/ago.1998. Disponível em: < <http://www.scielo.br/pdf/ci/v27n2/king.pdf> >. Acessado em: 10 jan. 2007.

VALENTIN, M. L. P. **Política Nacional de Informação para a Produção de Bases de Dados em C&T**: estabelecimento de critérios de avaliação. São Paulo: Departamento de Biblioteconomia e Documentação da USP, 2001.