

O PHL quebrando paradigmas em serviços e processos biblioteconômicos

Elysio Mira Soares de Oliveira

Apresentado no "I Encontro Nacional de Usuários do PHL"
(22/23 julho 2005 – Curitiba-PR)

Meu primeiro contato com o "computador" se deu em 1973. Estava cursando o primeiro ano de Biblioteconomia e numa visita à biblioteca do Instituto de Matemática da EESC-USP em São Carlos conheci um catálogo de cartões perfurados. Completamente fora de todos os padrões que estávamos aprendendo na escola, mas muito funcional (os usuários adoravam). Em cada cartão havia somente uma linha (80 caracteres) com duas informações: entrada (70 caracteres) e localização (9 caracteres). Havia uma gaveta com cartões de títulos, outra de assuntos e outra de autores. No final de cada dia, a bibliotecária inseria os novos cartões nas respectivas gavetas, em qualquer ordem e colocava todo o conteúdo da gaveta em uma máquina classificadora de cartões. A agilidade da "classificadora de cartões" na ordenação dos cartões me impressionou.

Na disciplina de Documentação ministrada pelo saudoso Prof. Alfredo Américo Hamar, fomos inteirados do padrão de registros criado pela Biblioteca do Congresso Americano (LC) que permitia armazenar em fitas magnéticas, dados coletados a partir de cartões perfurados ou fitas perfuradas. Os cartões eram lidos em um leitor e os dados gravados em uma fita magnética. Um software formatava estes dados para impressão das fichas catalográficas, iguais àquelas que éramos obrigados a datilografar e depois intercarlar manualmente no fichário. Ainda achava o método do Instituto de Matemática mais eficiente. Não via nenhuma outra inovação neste processo a não ser o de manter uma cópia do catálogo, gravado na fita magnética, para possível intercâmbio. Foi meu primeiro contato com o formato (MARC - Catálogo Legível por Máquina).

Em 1978, quando passei a integrar a equipe encarregada da mecanização do catálogo da biblioteca do antigo IEA (Instituto de Energia Atômica), hoje IPEN, me inteirei sobre um projeto da FGV-Rio da adoção do formato CALCO, uma variante brasileira do MARC. Nesta mesma ocasião a DIDC/IEA já tinha seus catálogos mecanizados utilizando outro padrão. Para a mecanização, adotou-se um sistema denominado ILS (Integrated Library System). Este sistema utilizava as regras do AACR2 e um formato semelhante ao MARC mas com uma novidade sobre o MARC, ele não era tão preso às regras de pontuação e sinais requeridos pelo MARC e permitia a impressão dos índices dos catálogos, em papel contínuo em um formato mais linear e econômico. Eliminava o trabalho de alfabetização manual e intercalação das fichas no fichário. Aí sim, senti algum avanço sobre os métodos tradicionais de catalogação. Para a ordenação, não era mais necessário despejar a gaveta de cartões na classificadora. Bastava encaixar o rolo de fita magnética na leitora e digitar um simples comando que a impressora se incumbia da impressão das listas que iriam compor o catálogo.

Neste mesmo ano, no IEA, tive o primeiro contato com um outro formato, o formato ANSI, criado pelo "American National Standard Institute", não mais para armazenamento

de registros em fitas magnéticas para impressão de catálogos mecanizados, como o MARC ou o ILS, mas sim para armazenamento em disco magnético (bases de dados) para buscas on-line. O novo catálogo "eletrônico" permitia que a informação de interesse fosse encontrada por qualquer palavra ou por uma expressão booleana composta por conjuntos de palavras.

Nesta nova mídia, a redundância na transcrição de dados, necessárias para a geração de entradas para os índices dos catálogos impressos, passam a ser desnecessárias. O conceito de entrada primária, secundárias, título uniforme, formas variantes do título, título para alfabetação, entre outros perdem o sentido face aos novos recursos de acesso aos dados nesta nova mídia.

A IBM passa a comercializar o STAIRS©IBM (Storage and Information Retrieval System). Surgem os primeiros softwares de gestão de bases de dados, novos formatos de armazenamento de dados e com eles novas estruturas e formatos de registro bibliográfico. O foco do processamento sofre uma mudança radical, os detalhes necessários para a descrição do suporte na catalogação convencional, voltada à produção de índices e catálogos impressos dão lugar à descrição de conteúdo. O registro bibliográfico é incrementado com resumos, sumários e descritores. Novas chaves de acesso, até então, desprezadas nos padrões tradicionais de catalogação voltados para a produção de índices impressos.

As vantagens oferecidas por esta nova mídia, levaram os grandes sistemas de informação bibliográfica, a adotarem formatos mais eficientes para a alimentação de suas bases de dados. Órgãos do próprio governo norte-americano, como a USAEC, adotava o formato ANSI para organização do acervo de informações nucleares (Nuclear Science Data Base). Bibliotecas e sistemas de informações de órgãos das Nações Unidas também decidiram por padrões mais eficientes face às facilidades apresentadas por esta nova mídia.

A concepção da catalogação tradicional, voltadas à produção de catálogos impressos começa a ser questionada. Algumas bibliotecas que ainda não haviam iniciado seu processo de automação passam a desenvolver sistemas fundamentados em novos formatos. A morte do formato MARC (MARC is dead?), passa a ser tema e polêmica na maioria dos eventos internacionais promovidos pelas associações nacionais e internacionais de bibliotecas.

As empresas de desenvolvimento de softwares comerciais se mantinham alheias a estas discussões e persistiam na utilização do MARC como padrão em seus softwares. Não pela sua eficiência, mas pelo simples fato de que seus consumidores potenciais eram bibliotecas públicas que já vinham produzindo seus catálogos mecanicamente a partir de registros armazenados em fitas magnéticas utilizando o MARC.

Nenhum outro software de automação, naquela época, teria comércio garantido se não considerasse o investimento já feito por estas bibliotecas na mecanização de seus catálogos. O peso dado pelos bibliotecários à manutenção de compatibilidade com o

passado era bem maior que o peso dado, talvez até por desconhecimento, aos novos recursos de processamento que iam surgindo. O argumento das facilidades da aquisição de catálogos já processados que podiam ser adquiridos, também facilitava a venda destes softwares. Estes vendedores se beneficiavam destes dois fatores.

Encontraram nas bibliotecas públicas e universitárias, um terreno fértil para a venda de seus produtos. Esta facilidade não era a mesma entre as bibliotecas especializadas, com acervos mais específicos e que optavam por soluções de processamento mais ágeis e eficientes.

A OIT inicia um projeto de desenvolvimento de um software de gestão de bases de dados textuais. Este projeto é transferido para a Unesco e se dá o surgimento do ISIS (Integrated Set for Information System). Um novo conceito de base de dados criado para gestão de informação textual, permitindo registros de tamanho variável, campo de dado repetitivo, acesso através de índice no formato de árvore balanceada (B-tree file), e exportação de registros em um padrão de intercâmbio, o ISO-2709, mais tarde utilizado também para intercâmbio de registros MARC. Com o surgimento do ISIS, o UNISIST lança um manual de referência para registros legíveis por máquina, com recomendações que passaram ser adotadas pela quase totalidade dos sistemas de informação e bibliotecas especializadas e de órgãos das Nações Unidas (IAEA, OIT, Unesco, PAHO, FAO, CEPAL, etc.).

Surgem os micros computadores e com eles a popularização dos serviços de automação de catálogos nas bibliotecas. A Unesco desenvolve e distribui gratuitamente o MicroISIS para uso em microcomputadores (MicroISIS e o Winisis). Dezenas de milhares de bibliotecas passaram a adotar esta nova tecnologia para gestão de seus acervos com base nas recomendações da UNISIST. Outras, talvez pela dificuldade da conversão de seus dados, passaram a utilizar o ISIS na gestão de suas bases de dados, mas continuando presas as antigas regras ainda voltadas para impressão de catálogos.

Surge então outro novo componente, a internet e com ela o hipertexto trazendo profundas alterações no processo de transmissão do conhecimento. A catalogação, a descrição bibliográfica e os meios e padrões de armazenamento da informação, sofre, mais uma vez, uma enorme transformação. Nesta nova mídia, e com o recurso da linguagem de marcação HTML, o texto ganha a hiper-dimensão permitindo ao leitor caminhar nestas dimensões através de ligações pré-estabelecidas.

Junto com o HTML surge o conceito de meta dado, criando regras que permitem ao próprio autor assinalar os elementos de dados que descrevem a sua criação, permitindo que robôs criem automaticamente os pontos e acesso à massa de informação armazenada (catálogos). Nasce em Ohio, Estados Unidos, um novo formato de descrição bibliográfica baseado em meta dados, o Dublin Core adotado, pelo W3C como formato "standard" para descrição de informação na internet. Hoje, os catálogos que consultamos através do Yahoo, Google, Altavista, etc. utilizam estes recursos.

Em maio de 2001, Martin Dillon em um texto intitulado "The Role of Cataloging in the Future of Libraries" fala sobre o futuro das bibliotecas e especificamente sobre as mudanças nas técnicas, regras, formatos e métodos da catalogação face ao crescimento dos recursos da internet e do surgimento de uma nova linguagem e de um novo formato de catalogação, o formato XML. Baseado na linguagem HTML e ideal para a catalogação. Coincidentemente, naquele mesmo mês, o futuro que ele se referia, já se fazia presente na biblioteca Dr. Romeu Ritter dos Reis em Porto Alegre, onde o PHL já estava sendo utilizado em caráter experimental. Nascia ali a primeira experiência brasileira na automação de serviços e processos de bibliotecas através da Web.

Já nesta época, havia a corrente de pensamento negativa sobre questões de segurança a respeito do uso da internet para controle de serviços e processos. Mas também, havia uma corrente contrária a este pensamento negativo e que mesmo ciente dos riscos, apostaram no uso desta mídia e começaram a desenvolver aplicações comerciais (E-commerce) e bancárias (E-bank).

Em alguns meses de teste, já se podia perceber o poder desta nova tecnologia onde muitos paradigmas começavam a ser quebrados.

Esta nova concepção no tratamento da informação e as facilidades oferecidas por esta nova mídia, vieram proporcionar, não só aos bibliotecários mas também a profissionais de outras áreas, a possibilidade de organizarem seus acervos e serviços de forma simples, eficiente e intuitiva, se beneficiando de todas as facilidades do ambiente Web.

Se analisarmos atentamente, podemos perceber que no PHL são tantos os novos conceitos incorporadas e paradigmas quebrados que até entre a maioria dos profissionais de informática se percebe dificuldade de compreendê-los.

O PHL não é um "software", é uma aplicação Web que gera hipertextos a partir de informações armazenadas em bases de dados CDS/ISIS©Unesco. Pode ser comparado ao Google, Altavista, ou Yahoo, com algumas diferenças, sendo uma delas, a de oferecer formulários que permitem modificar ou incrementar seus conteúdos.

O PHL está constituído por um conjunto de roteiros que tem a função de interpretar as requisições feitas através de um software de navegação (IE Explorer, Mozilla, Netscape, Opera) do usuário e repassar estas instruções para um software servidor de bases de dados (WWWisis©Bireme), que por sua vez executa estas instruções e devolve para seja interpretada pelo software do usuário.

Foi concebido com base na tecnologia cliente/servidor e incorporado de recursos que permitem simular este ambiente em um computador isolado, tornando-o de uso nos três ambientes conhecidos (local/intranet/internet). Estes recursos permitem que seja utilizado em computadores isolados porque fazem com que parte do processador seja dedicado ao processamento de requisições recebidas pela porta de comunicação 80 (HTTP) e a outra parte dedicada ao processamento de demandas locais.

Não tem exigências de hardware. Pode ser utilizado em sistemas operacionais leves (linux, freebsd, etc.) e utiliza padrões de bases de dados não relacionais e software de gestão não residente. O WWWisis(c)Bireme é um software servidor de bases de dados que foi concebido fora dos padrões convencionais, além de ser minúsculo ele só é carregado na memória do servidor no momento que recebe e processa uma requisição, ao terminar é imediatamente liberado da memória. Normalmente, os softwares de gestão de bases de dados (SQLserver, MySQL, Oracle, etc.), ficam sempre carregados na memória, mesmo quando não estão sendo requisitados e por este motivo acabam impondo limitações de acessos simultâneos e requerem altos recursos de hardware (processador).

Todas as bases de dados que constituem o PHL podem ser manipuladas por qualquer "software" ou ferramenta de gestão de bases de dados ISIS. (Microisis©Unesco, Winisis©Unesco, Cisis©Bireme, WWWisis©Bireme, etc.);

Todas as bases de dados que constituem o PHL podem ser exportadas no padrão de intercâmbio ISO-2709 ou no novo formato de bases de dados com estrutura XML;

Qualquer usuário que domine a linguagem de formato ISIS (Print Formating Language) pode interagir com qualquer uma das bases de dados do PHL e produzir qualquer tipo de relatório no formato desejado.

Todas as saídas geradas pelo PHL são arquivos texto com marcações HTML ou XML e podem ser manipulados com qualquer "software" compatíveis com este padrão.

Todos os roteiros que constituem o PHL são arquivos textos (fontes abertos) e podem ser editados com qualquer editor de texto.

Vou encerrar esta palestra, com uma frase de um de um naturalista inglês que viveu no século XVIII e que acho bastante oportuna face aos novos meios tecnológicas que foram surgindo a partir do final do século passado.

“Na natureza, não são os mais fortes nem os mais inteligentes que sobrevivem e sim, aqueles com maior capacidades de se adaptarem às mudanças” (Charles Darwin - 1859)