
O Impacto do OSI na Permuta Internacional de Registos UNIMARC e CCF *

CHRISTINE BOßMEYER

Die Deutsche Bibliothek, Frankfurt am Main (Alemanha)

ESTA comunicação cobre duas áreas fascinantes em que há vários anos estou interessada: OSI e UNIMARC. Até agora, nunca encontrei nenhuma relação entre as duas mas, uma vez que tenho de preparar esta comunicação, vou tentar estabelecê-la. De facto, e muito honestamente, tive algumas dificuldades para preparar esta comunicação. Primeiro, o tema pode levar a pensar que o OSI é uma função automatizada de bibliotecas muito bem definida e estabelecida, como o são a circulação ou a catalogação em linha. A verdade, porém, é que a interconexão entre sistemas ainda se encontra na sua infância. Além disso, alguns projectos e experiências no âmbito do OSI não têm progredido assim tanto.

* Tradução da comunicação original em inglês apresentada ao UNIMARC/CCF Workshop, organizado pela CEE e UNESCO, Florence, June 1991.

Ainda não dispomos de grande experiência com a tecnologia OSI.

Uma segunda reserva a propósito deste tópico tem a ver com a minha própria desconfiança de que as questões de troca de dados que hoje enfrentamos não se alterarão num ambiente OSI utilizando o UNIMARC ou o CCF. Seja como for, resolvi ultrapassar as minhas dúvidas e acrescentar um ponto de interrogação ao título desta comunicação, e abordar o todo como uma pergunta. Considerarei, pois, se o OSI terá ou não um impacto na permuta de dados em geral, e sobre o UNIMARC e o CCF em particular. A resposta a esta pergunta tanto pode ser sim, como não.

1. Análise da situação presente

Qualquer consultor digno desse nome começa o seu trabalho com um levantamento da situação presente. Seguirei este método, e

começarei com uma análise da situação actual no que diz respeito à troca de dados. Realçarei duas áreas:

1. programas para conversão de dados;
2. serviços de troca de dados.

Alguns outros aspectos relacionados com a permuta e partilha de registos bibliográficos também têm interesse como, por exemplo, as questões legais, a licença por contrato entre o produtor de registos e o comprador, direitos de autor e modificação de registos, custos, etc. Regra geral, os responsáveis pelas bibliotecas ou os departamentos jurídicos preocupam-se com estas questões. Mas estas questões não têm praticamente nada a ver com o processamento da informação e com a permuta de registos. Por esse motivo, não as considerarei nesta comunicação.

1.1 Conversão de dados

Desde que existe informação bibliográfica legível por máquina, os programas para conversão de dados de um formato para outro passaram a desempenhar papel importante. A disponibilidade de um programa de conversão é a condição prévia mais importante para permitir a inclusão de informação bibliográfica de origem diversa no sistema de uma qualquer biblioteca. Os interfaces que permitem a execução de

import/export e a possibilidade de converter dados passaram a constituir um importante critério de avaliação dos sistemas para bibliotecas disponíveis no mercado.

A verdade, porém, é que a interconexão entre sistemas ainda se encontra na sua infância. Além disso, alguns projectos e experiências no âmbito do OSI não têm progredido assim tanto. Ainda não dispomos de grande experiência com a tecnologia OSI.

Os métodos para criação de programas e para criação de programas de conversão foram aparecendo lenta e gradualmente. De momento, podem considerar-se três categorias de programas de conversão:

1. Programas dedicados de conversão

Trata-se de programas capazes de transformar dados de um determinado tipo de formato para outro. Estes programas garantem uma total correspondência na conversão. Muitas vezes a definição dos dados relativamente à fonte e ao formato final vêm incluídos no programa-fonte da própria aplicação (segundo o PICA designa-se esta técnica por *hard-coded*).

2. Programas de conversão por tabela

Significa isto que os programas podem ser parametrizados ou con-

trolados por tabelas de conversão. A definição de campos ou de elementos, assim como as regras da conversão, são externas ao programa. Portanto, estas tabelas podem ser alteradas ou acrescentadas de forma independente, mantendo-se estável o próprio programa.

Os programas de conversão por tabela são geralmente muito flexíveis, i. e. um programa pode ser usado com diversos formatos MARC partindo do princípio que estes não apresentam diferenças substanciais. Cada formato de conversão requer a sua própria tabela de conversão.

A disponibilidade de um programa de conversão é a condição prévia mais importante para permitir a inclusão de informação bibliográfica de origem diversa no sistema de uma qualquer biblioteca.

3. Conversores

Os conversores representam um progresso relativamente aos programas de conversão parametrizados. Em relação às tabelas de conversão, estes programas ainda aceitam processar informação a partir do «exterior», indicando como é que o conteúdo de um campo deve ser processado. Nos últimos anos têm sido desenvolvidos e introduzidos com êxito conversores para MARC.

Há uma questão que contudo permanece: será teoricamente possível desenvolver um conversor capaz de transferir registos bibliográficos de um formato-fonte para um formato final qualquer?

Melhoria de qualidade através do controlo de autoridade

É muito comum entre os bibliotecários a ideia de que a conversão de registos tem a ver com a falta de qualidade dos dados. Durante o processo de conversão alguns programas recorrem agora a mecanismos para fazer corresponder os registos já convertidos com entradas de autoridade ou tabelas de validação. Sem dúvida nenhuma que através do controlo de autoridade se pode melhorar a qualidade dos registos convertidos. É óbvio que estas funções podem ser utilizadas para traduzir diferentes formas estabelecidas para entradas de autoridade e para tornar a permuta internacional mais fácil.

Processamento em diferido

Os programas clássicos de conversão aplicados em grandes computadores correm em diferido. Para cada tarefa os ficheiros são convertidos de acordo com o seguinte esquema:

- leitura de um registo de um ficheiro;

- processamento do registo e execução da conversão;
- integração do registo no ficheiro.

Este ciclo prossegue até o ficheiro ter sido lido na totalidade.

Outra característica de um procedimento em diferido é que normalmente este não é residente em memória. Os programas em diferido são carregados durante uma tarefa, e as tarefas em diferido iniciam-se de acordo ou com um horário pré-definido ou controlados por uma pessoa.

Conversão de registos em sistemas locais

O desenvolvimento e utilização de um programa de conversão já não é uma tarefa que interesse apenas a bibliotecas nacionais, agências bibliográficas de carácter comercial ou grandes bibliotecas.

Com o aparecimento dos sistemas locais no mercado e o rápido crescimento de entidades várias, aumentou a necessidade de programas de conversão. Também a invenção dos CD ROM's e a disponibilização de bases de dados bibliográficos de grandes dimensões neste suporte são outras das razões para a banalização dos programas de conversão.

Existe nos sistemas locais uma tremenda necessidade de identificar e converter registos bibliográficos a partir quer de bases de dados [em

linha*] quer de CD ROM's, modificando-os e, pelo menos, garantindo a importação desses registos para bases de dados locais. O reverso também é verdadeiro: os registos de sistemas locais podem ser transferidos (*uploaded*) para grandes sistemas. Para este procedimento nos dois sentidos é necessário um programa flexível e que possa ser definido pelo pessoal do sistema local. Para este tipo de ambientes, os programas de conversão deverão operar em modo de transacção e deverão ser residentes e executados imediatamente.

O desenvolvimento e utilização de um programa de conversão já não é uma tarefa que interesse apenas a bibliotecas nacionais, agências bibliográficas de carácter comercial ou grandes bibliotecas.

É também necessário que os programas de conversão disponham de *interfaces* e sejam capazes de aceitar registos e enviar registos convertidos para o sistema de aplicação.

1.2 Procedimentos para troca

Troca de suportes de armazenamento (Figura 1)

* Esclarecimento da responsabilidade da tradutora.

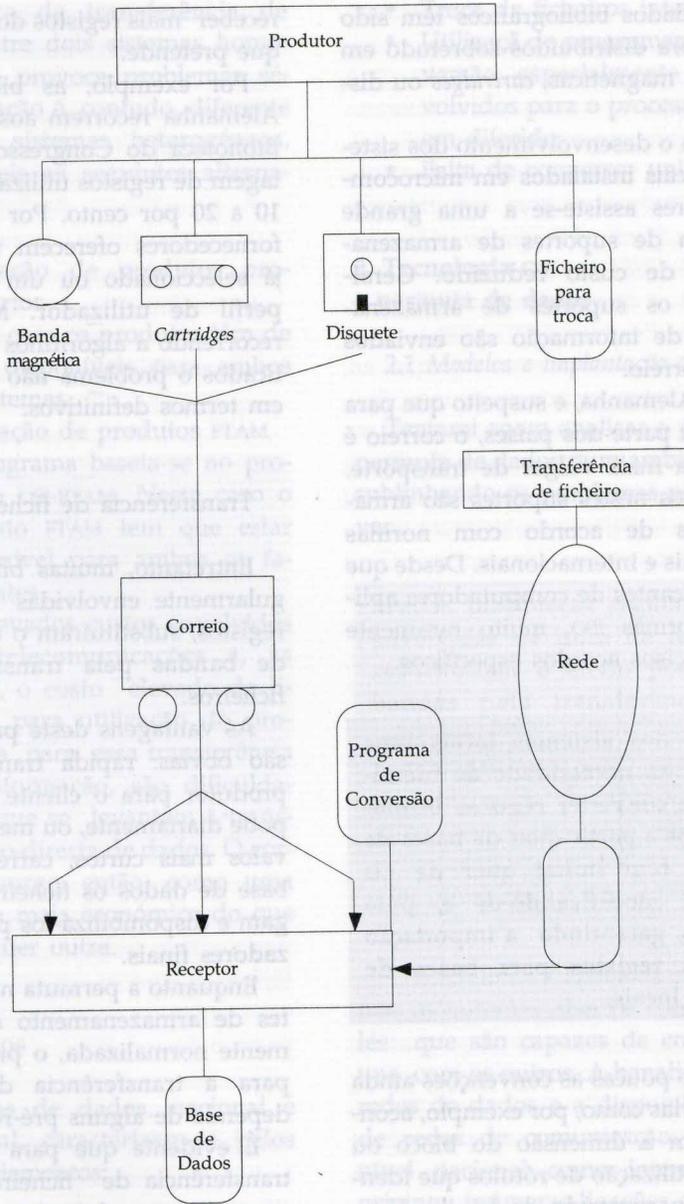


Figura 1
Modelo para permuta de dados

Os dados bibliográficos têm sido até agora distribuídos sobretudo em bandas magnéticas, *cartridges* ou disquetes.

Com o desenvolvimento dos sistemas locais instalados em microcomputadores assiste-se a uma grande procura de suportes de armazenamento de custo reduzido. Geralmente, os suportes de armazenamento de informação são enviados pelo correio.

Na Alemanha, e suspeito que para a maior parte dos países, o correio é a forma mais vulgar de transporte. Os dados nestes suportes são armazenados de acordo com normas nacionais e internacionais. Desde que os fabricantes de computadores aplicam normas ISO, muito raramente são precisos acordos específicos.

Existe nos sistemas locais uma tremenda necessidade de identificar e converter registos bibliográficos a partir quer de bases de dados [em linha] quer de CD ROM's, modificando-os e, pelo menos, garantindo a importação desses registos para bases de dados locais.

Já são poucas as convenções ainda necessárias como, por exemplo, acontece com a dimensão do bloco ou com a utilização de rótulos que identificam o conteúdo da banda. Uma das características da banda ou da disquete é que o cliente acaba por

receber mais registos do que aqueles que pretende.

Por exemplo, as bibliotecas na Alemanha recorrem aos serviços da Biblioteca do Congresso. A percentagem de registos utilizados varia de 10 a 20 por cento. Por isso, muitos fornecedores oferecem um produto já seleccionado ou um serviço por perfil de utilizador. Mas mesmo recorrendo a algoritmos muito sofisticados o problema não tem solução em termos definitivos.

Transferência de ficheiros

Entretanto, muitas bibliotecas regularmente envolvidas na troca de registos, substituíram o envio postal de bandas pela transferência de ficheiros.

As vantagens deste procedimento são óbvias: rápida transmissão do produtor para o cliente. O receptor pode diariamente, ou mesmo a intervalos mais curtos, carregar na sua base de dados os ficheiros que chegam e disponibilizá-los para os utilizadores finais.

Enquanto a permuta nestes suportes de armazenamento está amplamente normalizada, o procedimento para a transferência de ficheiros depende de alguns pré-requisitos.

É evidente que para proceder à transferência de ficheiros se torna necessário que dois sistemas estejam ligados, por exemplo, numa rede X.25.

A prática de transferência de ficheiros entre dois sistemas homogéneos não provoca problemas sérios. A situação é, contudo, diferente entre dois sistemas heterogéneos, sendo viáveis as seguintes alternativas:

1. Utilização de produtos proprietários —

Neste caso os produtos têm de estar disponíveis para ambos os sistemas.

2. Utilização de produtos FTAM

O programa baseia-se no protocolo OSI-FTAM. Neste caso o produto FTAM tem que estar disponível para ambos os fabricantes.

Os elevados custos envolvidos nas telecomunicações e, às vezes, o custo elevado da licença para utilização do programa para essa transferência de informação, são dificuldades que se levantam à transmissão directa de dados. O correio surge, então, como uma forma mais económica do que qualquer outra.

Resultados

As trocas de dados, nacional e internacional, caracterizam-se pelos seguintes elementos:

- Troca em suporte de armazenamento;

- Troca de ficheiros inteiros;
- Utilização de programas de conversão especialmente desenvolvidos para o processamento em diferido;
- Falta de conversor universal.

2. Tecnologia OSI e permuta de dados

2.1 Modelos e implantação OSI

Tentarei agora analisar o papel da permuta de dados num ambiente OSI, sublinhando os problemas por resolver.

Muitas bibliotecas regularmente envolvidas na troca de registos, substituíram o envio postal de bandas pela transferência de ficheiros.

É evidente que para proceder à transferência de ficheiros se torna necessário que dois sistemas estejam ligados, por exemplo, numa rede X.25.

O modelo de 7 camadas da OSI define os sistemas abertos como aqueles que são capazes de comunicar uns com os outros. A banalização de redes de dados e a disponibilização de redes de comunicação, tanto a nível nacional como internacional, originou inúmeras ligações entre sistemas e redes. Contudo, o conceito OSI ultrapassa em muito a simples

ligação entre sistemas heterogêneos de computadores.

Num cenário OSI, as aplicações instaladas e operacionais em diferentes equipamentos, e os sistemas em linha serão capazes de intercomunicar uns com os outros. A tecnologia OSI torna possível a ligação entre sistemas de bibliotecas. O acesso aos dados e a sua utilização sofrerão, pois, grandes transformações.

De facto, o OSI oferece apenas as pré-condições, os princípios e as normas indispensáveis para este diálogo entre sistemas. O OSI não muda o ambiente do sistema de cada biblioteca. Isto significa que ainda se manterão as seguintes diferenças:

- equipamentos diferentes;
- programas diferentes;
- sistemas diferentes de gestão; de bases de dados;
- formatos diferentes.

O diálogo OSI entre duas aplicações ocorre em linha. A troca de suportes de armazenamento, e a literatura acompanhante chegaram ao fim.

Há ainda outro aspecto a ter presente: um serviço OSI não pode ser organizado como uma tarefa idêntica àquela que hoje executamos quando uma banda é carregada na nossa base de dados. O OSI exige condições para a qual os nossos actuais programas de conversão e troca não estão (ou ainda não estão) preparados.

Todo o tráfico entre diferentes serviços de processamento será, a mé-

dio e longo prazo, alterado com o OSI. A sua instalação afectará os seguintes serviços:

- transferência de pesquisas;
- permuta de dados;
- actualização de ficheiros;
- empréstimo interbibliotecas (ILL);
- encomenda de obras e facturação entre bibliotecas e livrarias (*Electronic Data Interchange EDI*).

Num cenário OSI, as aplicações instaladas e operacionais em diferentes equipamentos, e os sistemas em linha serão capazes de intercomunicar uns com os outros.

O OSI oferece apenas as pré-condições, os princípios e as normas indispensáveis para este diálogo entre sistemas. O OSI não muda o ambiente do sistema de cada biblioteca.

Seria demasiado extenso abordar neste artigo todas as aplicações OSI. Referirei apenas, pois, a permuta electrónica de registos, deixando de parte o ILL e o EDI.

2.2 Transferêncis de registos e os protocolos de aplicação OSI

A transferência de registos inclui o manuseamento e a transmissão de

registos de uma aplicação para outra. O manuseamento de registos inclui a conversão de registos de um formato interno para um formato de troca. A transferência de registos é a função-base exigida para os seguintes serviços:

- captura de registos numa base de dados estrangeira, por exemplo, depois de uma pesquisa bem sucedida;
- sincronização de ficheiros, por exemplo, de ficheiros de autoridade;
- actualização de bases de dados, por exemplo, na troca de registos entre uma agência bibliográfica de cariz comercial e um sistema local.

O diálogo entre um sistema produtor e um sistema final decorre num modo de ligação orientada. Isto significa, por exemplo, que o sistema produtor — Sistema A — envia uma pesquisa para o sistema final — Sistema B — e fica a aguardar resposta. O procedimento no Sistema A fica à espera até que o Sistema B confirme o pedido ou responda — de forma positiva ou negativa — ao pedido de busca.

Este modo é também conhecido por serviço confirmado.

A questão que se levanta a seguir tem a ver com os princípios ou convenções a definir para a transferência de registos aplicando os protocolos OSI. A resposta é fácil e breve.

A provisão só pode ser reconhecida numa biblioteca que tenha adoptado os princípios OSI:

- Pesquisa e Recuperação (*Search and Retrieve* — SR);
Norma Internacional Provisória DIS 10 162 e 10 163;
- Empréstimo interbibliotecas (ILL) Norma Internacional Provisória DIS 10 160 e 10 161.

Outras normas importantes para a permuta de dados como a do Acesso Remoto a Bases de Dados (*Remote Database Access*, RDA) e como a da Ordenação e Recuperação de Documentos (*Document Filing and Retrieval*, DFR) não oferecem a possibilidade de mapear os elementos e os atributos dos dados bibliográficos.

Em princípio, existem duas possibilidades no protocolo OSI para a definição de dados bibliográficos:

- a definição dos campos do registo a ser permutado;
- a referência a formatos de dados já existentes.

A primeira alternativa tem a vantagem de que o protocolo explicita o atributo, por exemplo, um campo para o autor, um campo para o lugar de publicação, etc. Se o número de atributos a permutar for limitado, então o método usado é geralmente este.

Por exemplo, o protocolo ILL refere a lista de atributos indis-

pensável para executar o pedido de empréstimo interbibliotecas.

A segunda alternativa é a seguida pelo protocolo SR. Este protocolo define uma lista das sintaxes possíveis do registo, mas não define uma lista completa dos campos. Duas entidades, por exemplo, um fornecedor de dados e um receptor de dados podem negociar em que formato os registos deverão ser transferidos (SR: Síntaxe Preferencial do Registo). Parte-se do princípio, contudo, que ambos os sistemas conhecem a sintaxe do registo, ou seja, o formato dos dados. Na prática, um registo num ambiente OSI não é mais nada do que uma cadeia de caracteres que passa de uma máquina para outra. A aplicação-receptora, pelo menos, no computador reconhece a estrutura da sequência de dados, pode distinguir campos e manusear o conteúdo dos mesmos.

O que significa um conversor universal de dados já foi explicado anteriormente, ou seja, do meu ponto de vista, será uma espécie de compilador capaz de converter diferentes formatos-fonte para um formato universal de dados, por exemplo, o UNIMARC.

Para a transferência de registos, nem o conhecidíssimo EBCDIC nem os códigos ISO são utilizados. Os dados estão codificados de acordo com a

Abstract Syntax Notation, ASN.1, outra norma ISO.

A assunção de que a norma ASN.1, ou antes, de que os compiladores ASN.1 são capazes de converter a estrutura de dados, está errada.

2.3 O impacto do OSI na conversão de dados

Todos os que sonharam com soluções óptimas para a conversão de registos bibliográficos no âmbito dos protocolos OSI têm de ser realistas. Repito, a conversão de um registo tem que ser definida quando um registo inicia o seu percurso OSI em direcção à meta, e a conversão continuará à chegada do registo.

Este exemplo parte do princípio que ambas as aplicações não utilizam o mesmo formato de dados.

Quais são, pois, as consequências desta diferença de formatos para o programa de conversão?

1. Os programas de conversão devem operar em tempo real e não, como hoje, serem apenas carregados quando uma banda chega pelo correio;
2. Os programas de conversão devem estar sempre prontos a actuar, ou seja, quando a execução do programa é solicitada, por exemplo, pela chegada de registos, o programa começa a correr e a conversão realiza-se.

O programa para garantir esta execução, deve:

- dispor de um *interface* programa-a-programa;
- ser programado de uma forma múltipla, ou seja, um mesmo código poderá ser partilhado por diferentes grupos.

É uma questão essencial saber como, e se, podem os programas de conversão existentes ser modificados para que sejam utilizados em implementações OSI. Qualquer pessoa que esteja a pensar em implementar o OSI tem de responder a esta questão.

Enquanto não dispusermos de tal compilador, só poderemos desenvolver os nossos próprios programas e converter dados para UNIMARC ou para CCF seja qual for o formato universal preferido.

Outro problema relativo à transferência de registos deverá ser ilustrado com uma interligação puramente teórica de sistemas (Figura 2).

O cenário lida com quatro sistemas interligados:

- O Sistema A desempenha o papel principal. Permuta registos com três outros sistemas e cada um deles usa um formato diferente. O Sistema A pode enviar e receber registos (esta seria,

aliás, uma configuração fantástica);

- O Sistema B, C e D recebem registos do Sistema A. O Sistema C também envia registos para o Sistema A. O trabalho mais pesado de conversão passa-se no Sistema A. Partindo do princípio que o Sistema A utiliza o seu próprio formato interno, quatro programas de conversão têm que estar activos para servir as conexões OSI:
 - Formato interno para UNIMARC;
 - Formato interno para CCF;
 - Formato interno para MAB;
 - MAB para formato interno.

Mesmo quatro processos de conversão em modo transacção são demasiados processos para qualquer aplicação. O resultado será uma perda significativa de qualidade de execução. Pensando nos muitos e diferentes formatos e variantes MARC que existem apenas usados pelas agências bibliográficas nacionais, não será precisa muita imaginação para reconhecer que uma conexão OSI e uma transferência de registos com numerosos formatos de dados não é possível. Na minha opinião, há duas soluções para este problema:

- a utilização de conversores universais de dados;
- a troca através de um formato universal de dados.

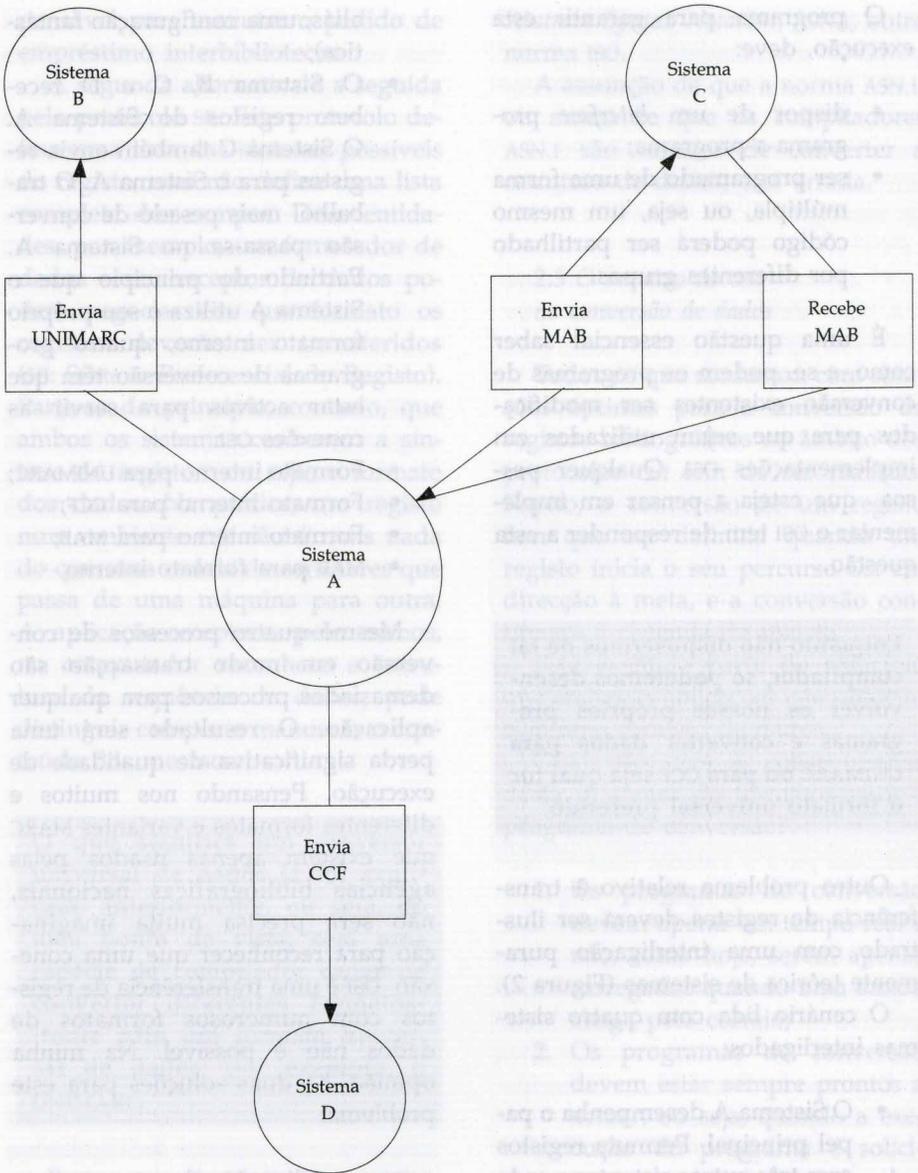


Figura 2
 Cenário para transferência de registros

O que significa um conversor universal de dados já foi explicado anteriormente, ou seja, do meu ponto de vista, será uma espécie de compilador capaz de converter diferentes formatos-fonte para um formato universal de dados, por exemplo, o UNIMARC. E este compilador pode também garantir a direcção inversa e criar, a partir de um formato universal, vários formatos finais. Não sei se as actuais técnicas de programação são suficientemente boas para construir um compilador com esta funcionalidade.

A implementação do CCF ou do UNIMARC vale sempre a pena. a longo prazo, a implementação do OSI e a transferência de registos exigirão um formato universal de dados.

Não pretendo especular, mas estou convencida de que este será o caminho no futuro com sistemas baseados no conhecimento. Enquanto não dispusermos de tal compilador, só poderemos desenvolver os nossos próprios programas e converter dados para UNIMARC ou para CCF seja qual for o formato universal preferido.

A implementação do CCF ou do UNIMARC vale sempre a pena. E, com uma visão geral destes problemas, certamente que um formato universal reduz o número de programas de conversão. E, estou convencida que,

a longo prazo, a implementação do OSI e a transferência de registos exigirão um formato universal de dados.

Conclusão

Para mim, já deixou de estar em aberto a questão do impacto do OSI na troca internacional de dados. A resposta é um «sim» bem claro. A concluir, o OSI não só modifica profundamente a permuta de dados, como provoca uma nova procura de programas de conversão. Sem a utilização de um formato universal de dados, a troca de dados através de uma ligação OSI é impossível. O UNIMARC pode, pois, servir como uma ponte para ultrapassar as diferenças entre os vários sistemas heterogéneos.

Glossário

- CCF — Common Communication Format
- EBCDIC — Extended Binary — Coded Decimal Interchange Code
- FTAM — File Transfer Access and Management
- MAB — Formato de recolha de dados alemão
- MARC — Machine Readable Cataloging
- OSI — Open Systems Interconnection
- PICA — Centre for Library Automaton, Leiden (Holanda)

Direitos de tradução e publicação gentilmente cedidos pela IFLA e pela autora. A tradução e o glossário são da responsabilidade de Maria Luísa Cabral com a colaboração de António Manuel Freire.