

FUNÇÕES DAS LINGUAGENS DE INDEXAÇÃO E RECUPERAÇÃO DA INFORMAÇÃO NOS SISTEMAS NACIONAIS E INTERNACIONAIS DE INFORMAÇÃO CIENTÍFICA E TÉCNICA *

L. Kofnovec
Central Office of Scientific Technical and Economic Information
Praga, Checoslováquia

RESUMO

Discute-se a possibilidade de estabelecer interrelações entre linguagens de recuperação em sistemas nacionais e internacionais de informação científica e técnica (ICT). Mútua convertibilidade é possível e somente útil quando há uma interseção dos assuntos processados por diferentes linguagens de recuperação. Diversos tipos básicos de linguagem de recuperação, os quais estão em uso ou podem ser usados em sistemas nacionais e internacionais de informação são discutidos, junto com as funções que eles deveriam desempenhar.

Descritores:
Linguagens de indexação; Thesaurus; Classificação decimal universal

As linguagens de indexação e recuperação da informação exercem função muito importante na construção dos sistemas nacionais e internacionais de Informação Científica e Técnica (ICT). A qualidade e o nível dos atuais serviços de informação dependem, em grande escala, da qualidade e nível das linguagens de indexação e recuperação da informação. É necessário ter em mente que existem hoje diversos tipos de linguagem de indexação e recuperação da informação e que é também possível determinar as respectivas funções primárias e secundárias. Merecem especial atenção os problemas relacionados com a possível interconexão de tais linguagens, já que a solução destes problemas relaciona-se diretamente com a possibilidade da interconexão temática dos diferentes subsistemas de ICT, ou seja, daqueles cuja interconexão é razoável e desejável.

A partir deste ponto de vista, já nos é possível fazer um juízo das funções primárias que deverão ser exercidas pela Classificação Decimal Universal (CDU), ou pela nova CDU reestruturada (NCDU), no conjunto das várias linguagens de indexação e recuperação da informação utilizadas hoje. O pleno reconhecimento das funções principais e secundárias de cada uma das linguagens (inclusive da CDU) será de grande ajuda para a formulação correta dos princípios

e critérios básicos para o planejamento e/ou abordagem conceitual e/ou replanejamento ou desenvolvimento posterior.

É também muito importante ter em mente que de acordo com tendências objetivamente verificáveis — a tecnologia do computador está crescendo em importância nos sistemas de informação, mesmo admitindo que os sistemas que utilizam computador ainda não constituem maioria em relação aos sistemas convencionais.

Se abordarmos o problema do ponto de vista das redes nacionais de ICT, poderemos verificar a existência de diversos tesouros (linguagens utilizando descritores) especializados que já se encontram em funcionamento, de alguns em experiência e de outros em preparação. Condição importante para aumentar sua eficiência como um todo é tomar as medidas necessárias à sua mútua interconexão. Mas os esforços para atingir este objetivo só adquirem sentido se houver uma superposição temática mútua entre os tesouros individuais, como se pode ver na Figura 5 (superposição temática entre os tesouros especializados T_{11} , T_{21} e T_{23}). Noutras situações, não há necessidade de interconexão.

Certamente isto não significa que não seria razoável tomar medidas convenientes para atingir alguma consistência metodológica ou formal na preparação dos tesouros, o que é possível pela observância dos padrões internacionais, regionais e nacionais, assim como de regras obrigatórias na elaboração dos mesmos.

* Traduzido do inglês por Regina Maria Soares de Oliveira, DAI/DOP/IBICT, e revisto por Pe. Astério Tavares Campos, UnB.
Tradução autorizada pelo autor.

A unificação da superposição das partes temáticas de cada tesouro pode ser atingida com a ajuda de um macrotesouro nacional (Figura 1) ou de um grupo menor de macrotesouros parciais. Os macrotesouros poderiam unificar os termos até determinado nível de generalidade, sendo o seu uso obrigatório em cada um dos tesouros individuais especializados. Isto tornaria possível utilizar dados (p.e. informações provenientes de centros orientados para disciplinas) para atender às necessidades de informação daqueles usuários interessados na informação num nível relativamente alto de generalidade. Esta relação entre macrotesouros e tesouros especializados ajudará também na transferência de blocos de informação entre os diferentes componentes das redes nacionais de ICT, embora sob este aspecto a função dos macrotesouros seja secundária e mais ou menos auxiliar. Esta abordagem hoje não é apenas uma possibilidade teórica, pois existem muitos esforços práticos neste sentido; o exemplo a ser citado em primeiro lugar é o do sistema Internacional de ICT que está sendo construído sob a orientação do Centro Internacional de ICT de Moscou⁽¹⁾ e também na República Democrática Alemã⁽²⁾. O problema das possibilidades de interconexão em micronível (nível de microindexação) serão discutidos mais tarde.

A ligação entre as redes nacionais e a rede internacional é apresentada na Figura 1. Há, em primeiro lugar, ligações mútuas entre os diferentes tesouros (especializados) na mesma área e o tesouro internacional especializado nessa área (Figura 4). Isto já foi realizado no caso da divisão internacional de tarefas através da distribuição dos processamentos importantes de informações pelos centros de informação localizados em países diferentes. Este é o caso de tesouros idênticos ou quase idênticos (este é o caso normal) em diferentes línguas; seu desenvolvimento e manutenção são coordenados e controlados por um só centro. Num caso ideal, já a preparação de tesouros internacionais especializados pode ser coordenada e influenciada pela existência de um macrotesouro internacional ou de um tesouro regional supranacional. A criação de um macrotesouro internacional não é apenas uma possibilidade teórica, como já se viu através dos esforços realizados pelo Centro Internacional de ICT em Moscou⁽¹⁾.

Merece especial atenção o problema da transferência de blocos de informação. Esta é a situação quando encontramos a superposição temática de tesouros individuais (como pode ser visto na Figura 2) e especializados T_{11} , T_{12} e T_{13} . O instrumento básico para a transferência de blocos de informação pode ser uma classificação bem ampla (p.e. o BSO — Broad System of Ordering — que está sendo preparado para o UNISIST, ou uma versão padrão abreviada da CDU ou de qualquer outra classificação universal). Os blocos de informação necessários deverão ser tematicamente definidos pelo BSO ou suas versões nacionais, ou mesmo por meio dos respectivos macrotesouros nacionais (que podem ser compatíveis com os tesouros nacionais especializados mais importantes). Deste modo será possível também obter.

em caso de necessidade, a automação da transferência dos Blocos de informação (como se mostra na Figura 2) entre os tesouros T_{11} e T_{12} mutuamente superpostos, ou entre os tesouros T_{12} e T_{13} se enire eles houver a mesma relação. Uma situação semelhante pode ser vista noutro exemplo (Figura 3) para uma rede internacional de ICT (tesouros T_{11} e T_{12} por um lado, e tesouros T_{22} e T_{23} por outro).

Se existirem três ou mais superposições mútuas de tesouros (de acordo com a Figura 5 encontram-se nesta situação os tesouros T_{11} , T_{21} e T_{23}) é possível passar automaticamente de uma estrutura de indexação para outra, quer diretamente por meio de concordância, quer pela aplicação de descritores idênticos (combinação de descritores) — e isto será possível se for aceita uma estrutura idêntica das partes superpostas dos respectivos tesouros — ou ainda por meio de uma linguagem intermediária na forma de uma classificação universal (microclassificação), p.e. a CDU. No segundo caso, como também parcialmente no primeiro, a aplicação de um macrotesouro não é suficiente para a perfeita compatibilização das estruturas dos tesouros porque o macrotesouro coordena a preparação dos tesouros (especializados) só até determinado nível de generalidade. Aqui é necessário aplicar outro mecanismo que torne possível uma coordenação direta (que conduza a uma identidade parcial) de dois ou mais tesouros (isto é, das partes superpostas) em micronível, do ponto de vista da sua futura aplicação em nível de microindexação, que é a base do processo de recuperação da informação.

A tabela anexa mostra o resultado de uma análise das várias funções preenchidas por cada um dos tipos de linguagem de recuperação da informação. Acham-se incluídos os seguintes tipos: macrotesouros internacionais (MTI), macrotesouros nacionais (MT), tesouros internacionais especializados (TI), tesouros nacionais especializados (T), o BSO (incluindo possíveis modificações nacionais), uma classificação universal (p.e. a CDU) e classificações especiais (CE). Como funções principais entendem-se as seguintes: a microindexação (incluem-se aqui a indexação de documentos (informação) e os pedidos de recuperação da informação, assim como a sistematização dos blocos de informação), a macroindexação (formação de blocos de informação e definição dos escopos temáticos dos centros de informação assim como das organizações), a função da interconexão mútua das diferentes linguagens em micro e macronível e, finalmente, a unificação (função de padronização dos tesouros especializados). Além disto, a tabela distingue as funções primárias e secundárias que são preenchidas por cada uma das linguagens de recuperação da informação.

É evidente que, para os fins da microindexação de documentos (informação) e dos pedidos de recuperação (formulação de pedidos de recuperação), são previstos em primeiro lugar os tesouros especializados (nacionais e estrangeiros) e também as classificações especiais (isto é, microclassificações); o papel das classificações universais, p.e. a CDU, é neste caso.

secundário. Para a sistematização dos blocos de informação, p.e. nos periódicos de resumos (abstracts), os descritores normalmente não são os melhores instrumentos; preferem-se as classificações universais ou especiais ou mesmo o BSO como classificação ampla.

A classificação universal e o BSO exercem papel principal na macroindexação. As classificações especiais e os macrotesauros podem exercer funções complementar na definição mais precisa dos blocos de informação.

Uma classificação universal (quando possui o caráter de microclassificação) pode ser utilizada para a interconexão das diferentes linguagens no micronível (nível de microindexação). Pode também ser usada na forma de versão padronizada abreviada da CDU, além do BSO que foi preparado exatamente com esta finalidade, de interconexão das diferentes linguagens em macronível. Os macrotesauros e as classificações especializadas podem também exercer uma função adicional.

Para maior eficiência das redes nacionais e supranacionais, e dos sistemas e subsistemas de ICT, para incrementar a boa utilização das coleções e das bases de dados, e também para aumentar neste sentido

a total eficiência do trabalho de informação em geral, é necessário partir da análise das funções que devem ser preenchidas por cada um dos tipos de linguagens.

REFERÊNCIAS BIBLIOGRÁFICAS

- 1 - KOFNOVEC, L. - Retrieval Languages in International System of STI. *Ceskoslov. informat.*, 77(7/8)200-4, 1975. (Em Checo).
- 2 - ROSEMBAUM, A.D. - Die Koordinierung der Thesaurusarbeiten in der DDR-ein Beitrag zur Verbesserung der Informationsleistungen. *Informatik*, 79(1):16-21, 1972.

ABSTRACT

Possibility of establishing interrelations between the retrieval languages in national and international scientific information systems. Mutual convertibility is possible and useful only when there is an intersection of the subject fields processed by different retrieval languages. A number of basic types of retrieval language which are in use or might be used in national and international information systems are discussed, together with the functions they should perform (ASLIB).

Funções das Diferentes Linguagens de Indexação e Recuperação

Função	MTI	MT	TI	T	BSO	C	CE
MICROINDEXAÇÃO							
Indexação de documentos e pedidos de recuperação			1	1		2	1
Sistematização de blocos de informação					2	1	1
MACROINDEXAÇÃO							
Formação de blocos de informação	2	2			1	1	2
Definição temática dos escopos dos centros e organismos de informações					1	1	
INTERCONEXÃO							
Interconexão das diferentes linguagens em micronível						1	
Interconexão das diferentes linguagens em macronível	2	2			1	1	
UNIFICAÇÃO							
Função padronizadora em relação aos tesaurus especializados	1	1	1				

Rede Internacional

Rede Nacional 1

Rede Internacional
(mundial, regional)

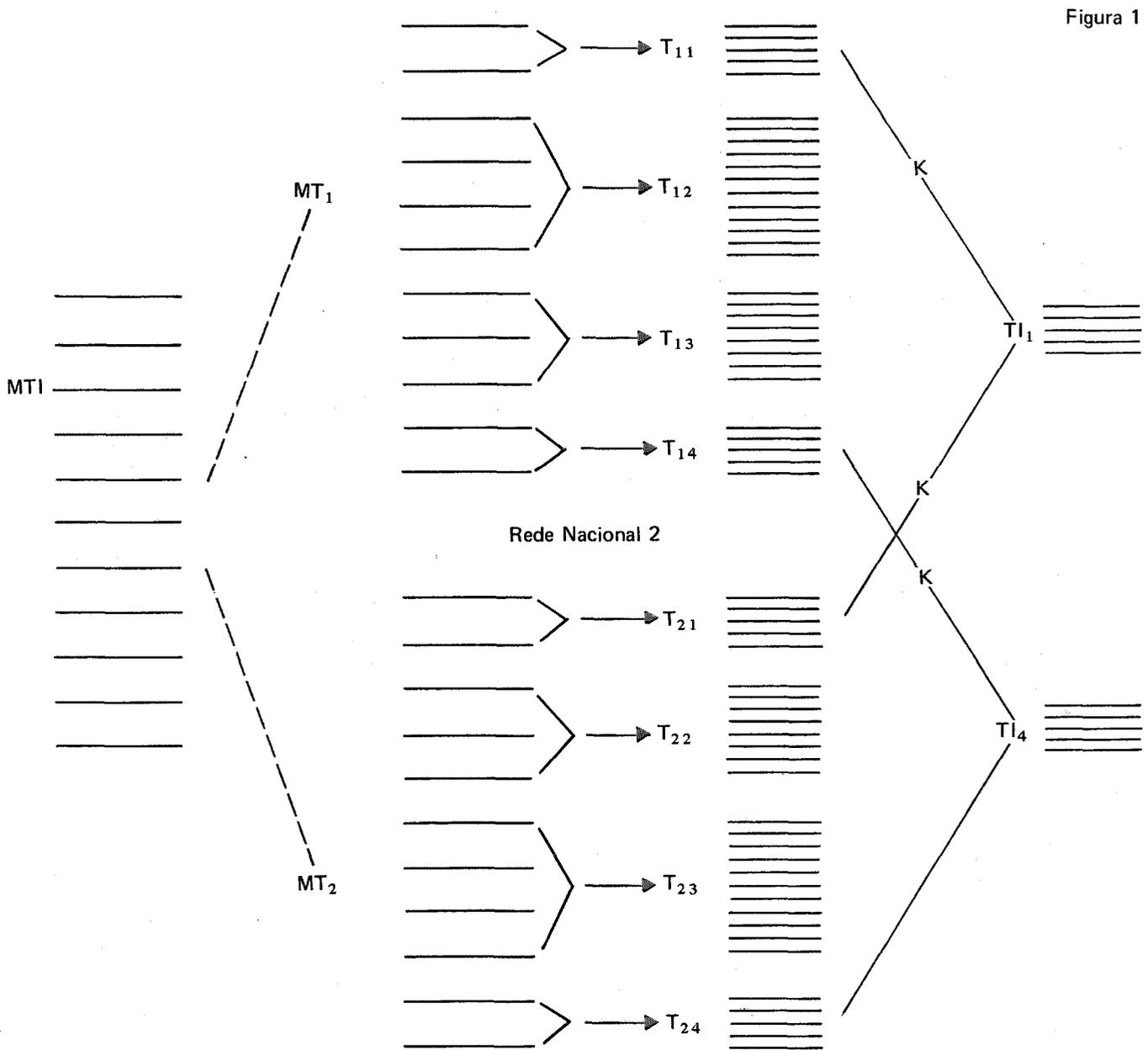


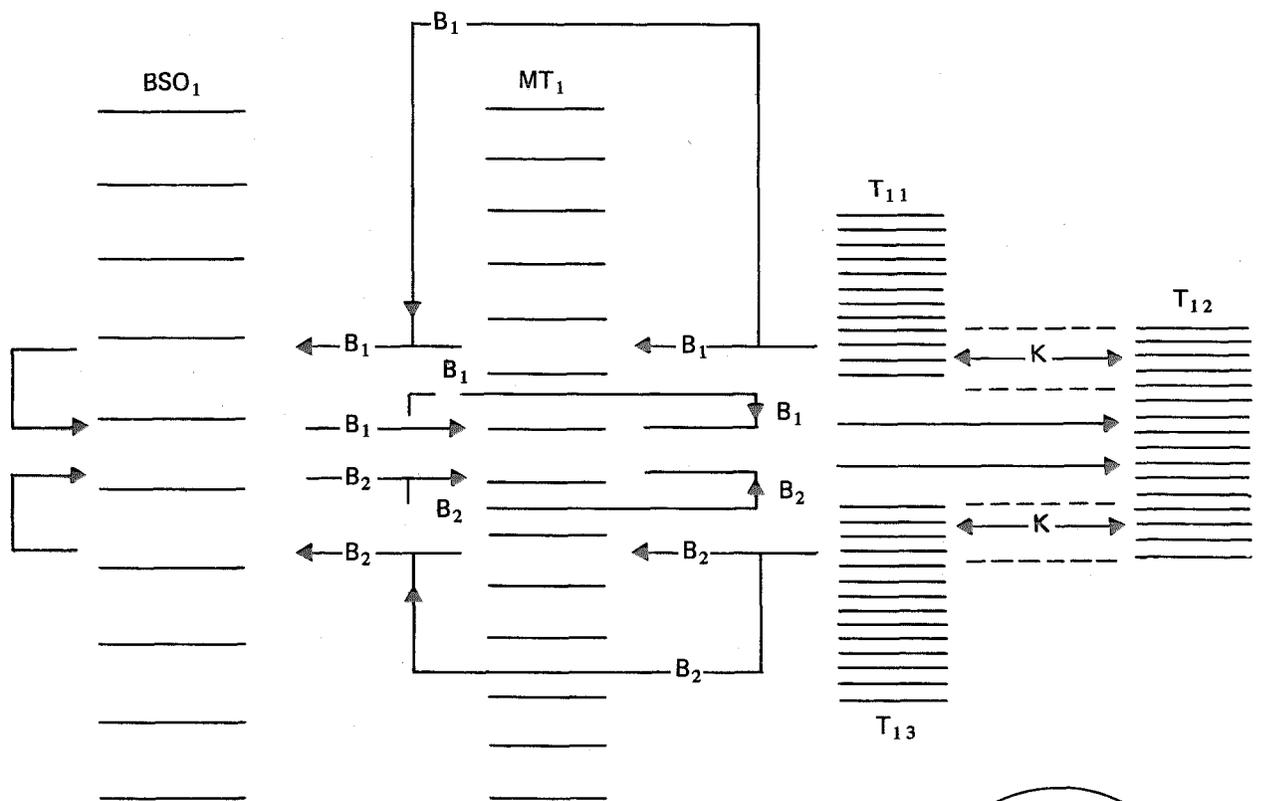
Figura 1

MTI
 MT
 T₁₁, T₁₂
 TI₁, TI₄
 K

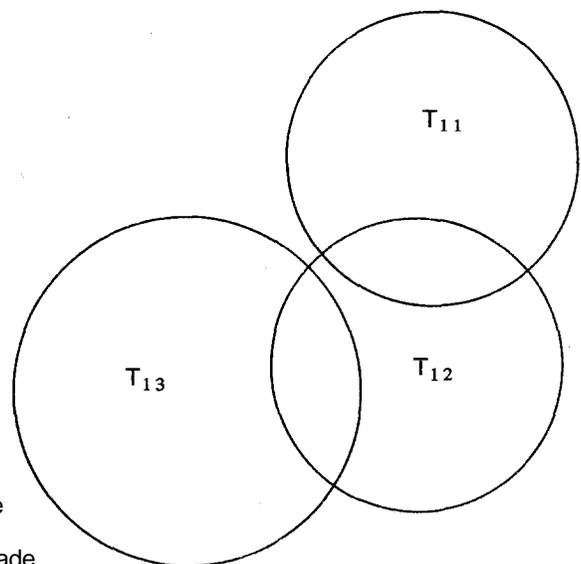
Macrotesauros internacionais
 Macrotesauros nacionais
 Tesauros nacionais especializados
 Tesauros internacionais especializados
 Relação de compatibilidade ou identidade

TRANSFERÊNCIA DE BLOCOS DE INFORMAÇÃO
(Rede Nacional)

Figura 2

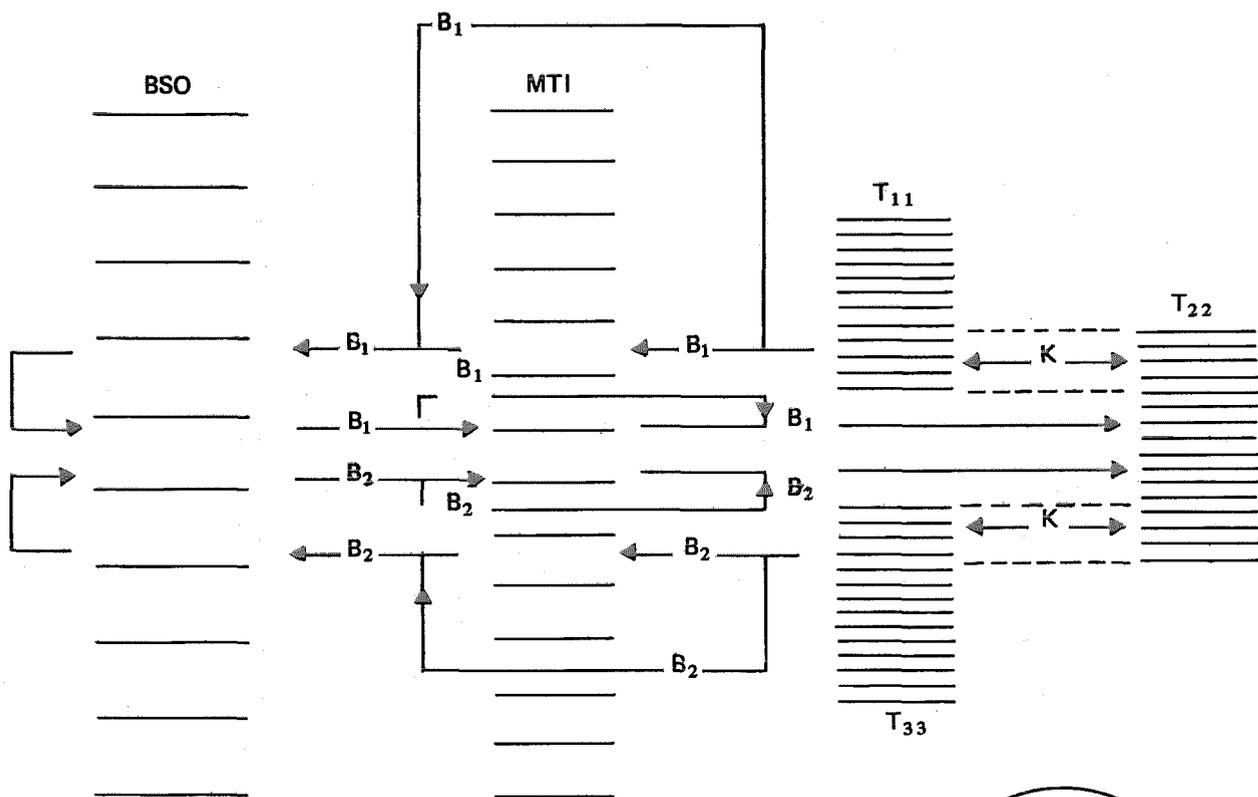


BSO₁ Versão nacional do BSO (modificado)
 ou CDU abreviada padrão
MT₁ Macrotesauros nacionais
T₁₁, T₁₂ Tesaurus nacionais especializados
B₁, B₂ Transferência (indexação) de blocos de
 informação
K Relação de compatibilidade ou identidade

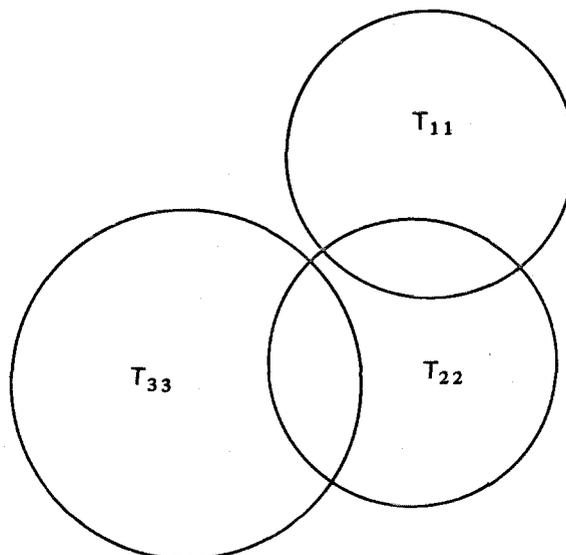


TRANSFERÊNCIA DE BLOCOS DE INFORMAÇÃO
(Rede Internacional)

Figura 3

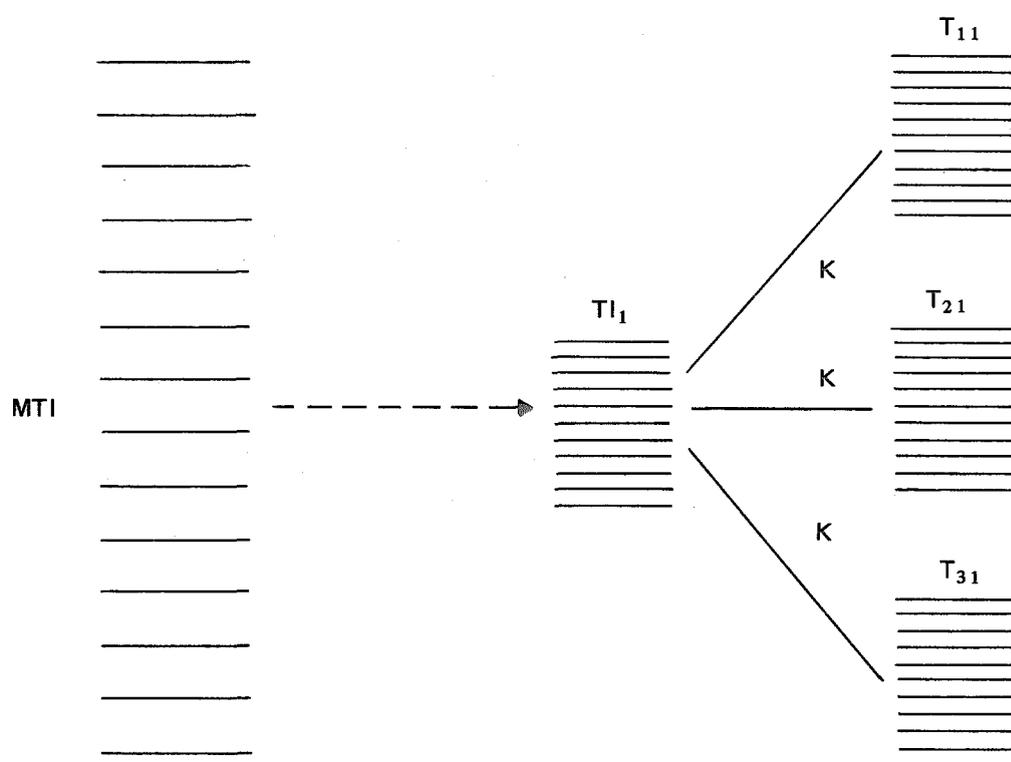


- BSO Sistema Amplo de Ordenação (BSO) ou CDU abreviada padrão
- MTI Macrotesauros internacionais
- T₁₁, T₂₂ Tesauros nacionais especializados
- B₁, B₂ Transferência (indexação) de blocos de informação
- K Relação de compatibilidade ou identidade



ORIGEM DE TESAuros NACIONAIS ESPECIALIZADOS

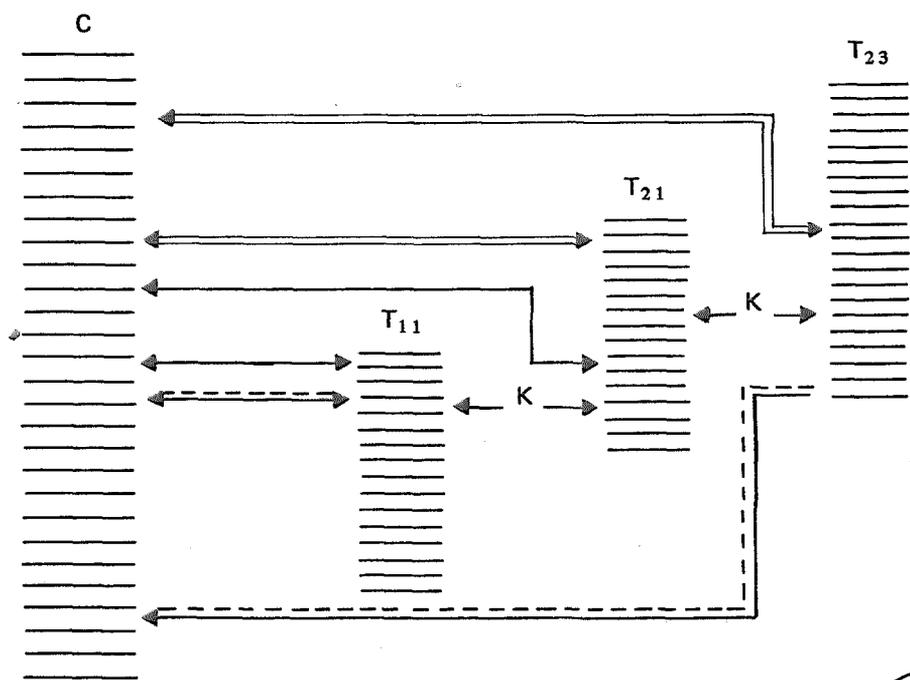
Figura 4



MTI	Macrotesauros internacionais
T ₁₁	Tesauros internacionais especializados
T ₁₁ , T ₂₁	Tesauros nacionais especializados
K	Relação de compatibilidade ou identidade

REALIZAÇÃO DE COMPATIBILIDADE ENTRE TERMOS

Figura 5



T_{11}, T_{21} Tesouros nacionais especializados
 \hat{U} Classificação universal, p. e. CDU
 K Relação de compatibilidade ou de identidade (realizado por meio de concordâncias mútuas ou identidade das partes)

