

# INTEROPERABILIDADE SEMÂNTICA: proposta metodológica para o mapeamento semântico entre SOCs em sistemas heterogêneos

## SEMANTIC INTEROPERABILITY: methodological proposal for the semantic mapping between KOS in Heterogeneous Systems

*Maria Luiza de Almeida Campos<sup>1</sup>  
Nilson Theobald Barbosa<sup>2</sup>*

### RESUMO

Na atualidade, com todo o avanço tecnológico e a diversidade de repositórios, bases de dados e usuários e com a variedade de terminologias para o tratamento e a recuperação da informação, questões relacionadas à interoperação semântica se colocam. No âmbito da Ciência da Informação, um dos procedimentos para solucionar estes problemas de heterogeneidade, no que tange à diversidade de vocabulários existentes nos vários repositórios e bases de informações é o mapeamento semântico. O mapeamento é o processo de estabelecer relações entre os conceitos de um vocabulário para outros e é usado largamente para obter a interoperabilidade sintática e semântica nos SOCs. Neste artigo, a fim de endereçar tal problema, apresentamos a proposta metodológica de I. Dahlberg, na qual o mapeamento é o primeiro passo para que se possa identificar o nível de compatibilidade entre os SOCs, denominados por essa autora de Sistemas Ordenados, com o propósito de permitir o uso comum dessas linguagens, quando forem efetuadas buscas em diferentes bases de dados ou quando se deseja intercambiar informações indexadas com estas linguagens.

**Palavras-chaves:** Interoperabilidade Semântica. Mapeamento Semântico. SOCs. Sistemas Heterogêneos.

### ABSTRACT

Nowadays, issues related to semantic interoperation arise in the context of all the technological advancement and the diversity of repositories, databases and users along with the variety of terminologies for the treatment and retrieval of information. In the scope of Information Science, semantic mapping is one of the procedures to solve these problems of heterogeneity, when related to the diversity of vocabularies existing in the various repositories and information bases. Mapping is the process of establishing relationships between concepts from one vocabulary to others and is used widely to achieve syntactic and semantic interoperability in KOS. In this article, to address such a problem, we present I. Dahlberg's methodological proposal, in which mapping is the first step in order to identify the level of compatibility between KOS, that are named by this author of Ordering Systems. There is the purpose of allowing the common use of these languages, when searches in different databases are made or when it is desired to exchange information indexed with these languages.

**Keywords:** Semantic interoperability. Semantic Mapping. KOS. Heterogeneous Systems.

*Artigo submetido em 20/09/2020 e aceito para publicação em 16/10/2020*

- 
- 1 Docente permanente no Programa de Pós-Graduação em Ciência da Informação. Universidade Federal Fluminense, Brasil. Bolsista de Produtividade do CNPq. ORCID <https://orcid.org/0000-0002-9253-3706>. E-mail: [marialuizalmeida@gmail.com](mailto:marialuizalmeida@gmail.com)
  - 2 Doutorando no Programa de Pós-Graduação em Ciência da Informação. Universidade Federal Fluminense, Brasil. ORCID <https://orcid.org/0000-0003-1497-313X>. E-mail: [nilson@tbarbosa.org](mailto:nilson@tbarbosa.org)

## 1 INTRODUÇÃO

Apesar do grande avanço computacional, a ampla diversidade de novas tecnologias disponíveis e os esforços para a criação de repositórios interligados semanticamente, podemos ver que a questão da interoperabilidade é um problema ainda em aberto. Agora não apenas num patamar físico, no qual as tecnologias atualmente já apresentam soluções que endereçam de modo bastante consistente o problema, com a integração de servidores, repositórios, bases de dados, e os usuários conectados, mas em um patamar fundamentalmente semântico.

Neste ambiente, onde existem bases heterogêneas e uma variedade de terminologias para a representação e a recuperação das informações geradas, questões relacionadas à reutilização do conhecimento e de interoperação semântica se colocam como essenciais para que seja possível lidar com o avassalador crescimento dessas informações em todos os setores e áreas do conhecimento e da atuação humana, visando a uma recuperação de informações consistentes. Uma delas é a necessidade de interoperabilidade semântica entre os diversos vocabulários que são utilizados para descreverem os recursos informacionais existentes nos variados repositórios e bases de informações.

No âmbito da Ciência da Informação, tal problema é endereçado na área de estudos que tratam de mapeamento entre vocabulários controlados, tesouros, ontologias, ou seja, mapeamento entre Sistemas de Organização do Conhecimento - SOC em geral (ZENG; CHAN, 2004; ZENG, 2019).

SOC é o termo utilizado em língua portuguesa para o termo em inglês *Knowledge Organization System* - KOS, proposto pelo grupo de trabalho *Networked Knowledge Organization Systems* - NKOS<sup>3</sup> em sua reunião de formação, na conferência ACM Digital Libraries de 1998, em Pittsburgh, Pensilvânia (EUA), e, desde então, vem sendo discutido sob essa denominação o seu papel na representação e na organização do conhecimento sob a luz do crescimento da Internet, da *web* semântica, da interoperabilidade entre sistemas heterogêneos, dos sistemas ligados abertos, dos sistemas de big data, dos sistemas de inteligência artificial, entre outros.

Na década de 1980, século passado, Dahlberg (1983) nomeava de Sistemas Ordenados o que hoje, guardadas as devidas proporções temporais e tecnológicas, denominamos de SOC, sendo definidos por ela como “[...] qualquer instrumento usado na organização, descrição (indexação) e

---

3 O NKOS se dedica à discussão do modelo funcional e de dados para habilitar Sistemas de Organização do Conhecimento - KOS, como sistemas de classificação, tesouros, dicionários geográficos e ontologias, como serviços de informação interativos em rede para apoiar a descrição e a recuperação de diversos recursos de informação através da internet (NETWORKED KNOWLEDGE ORGANIZATION SYSTEMS, 2020).

recuperação do conhecimento, composto por expressões verbais ou notacionais para conceitos e suas relações, dispostos de forma ordenada” (DAHLBERG, 1983, p. 5).

Já desde os anos 1980, antes mesmo das possibilidades computacionais hoje existentes, Dahlberg (1981, 1983, 1996) propõe procedimentos metodológicos para mapear as potencialidades semânticas dos conceitos que fazem parte de tais Sistemas Ordenados. Este mapeamento é o primeiro passo para identificarem-se possibilidades de compatibilização semântica entre SOCs.

Para o encaminhamento de tais questões, consideramos aqui neste artigo apresentar na seção 2 uma discussão sobre o conceito de interoperabilidade e suas diversas concepções, em que ressaltamos a importância do papel da interoperabilidade semântica. Na seção 3, apresentamos o conceito de mapeamento como um primeiro passo para ser possível o estabelecimento de potencialidades semânticas entre os diversos SOCs. Na seção 4, a proposta metodológica de Dahlberg, que visa ao mapeamento em nível semântico entre SOCs, ou Sistemas Ordenados, é discutida com o objetivo de identificar o nível de compatibilização semântica entre conceitos de diversos SOCs. Finalizamos apontando, na seção 5, a importância e a atualidade do estudo das metodologias de compatibilização da Ciência da Informação para os processos automatizados modernos.

## 2 A INTEROPERABILIDADE E SUAS DIVERSAS CONCEPÇÕES

Nesta seção, nosso objetivo é conceituar a interoperabilidade, suas diferentes visões em variados contextos. A partir do entendimento das potencialidades do que denominaremos de interoperabilidade sintática, estrutural e semântica, vamos nos concentrar nos problemas existentes no que concerne à interoperabilidade semântica.

Já na década de 1990, século passado, Paepcke *et al* (1998) mostram que a interoperabilidade vem sendo um problema crítico há algum tempo e continuará a ser num futuro previsível, uma vez que o número de sistemas computacionais, repositórios de informação, aplicações e usuários se multiplicam em uma taxa explosiva.

Segundo a *National Information Standard Organization*, “Interoperabilidade é a capacidade de múltiplos sistemas com diferentes plataformas de *hardware* e *software*, diferentes estruturas de dados e diferentes interfaces trocarem dados entre si com uma perda mínima de conteúdo e funcionalidade”<sup>4</sup>

---

4 Tradução livre dos autores. No original: “Interoperability is the ability of multiple systems with different *hardware* and *software* platforms, data structures, and interfaces to exchange data with minimal loss of content and functionality”.

(NATIONAL INFORMATION STANDARDS ORGANIZATION, 2004; CHAN; ZENG, 2006). Já o Comitê de Catalogação, Descrição e Acesso da *Association for Library Collections & Technical Services*, numa definição a partir de seu grupo de trabalho sobre metadados, após considerar diversas definições de diferentes fontes e autores, compila sua própria definição: “Interoperabilidade é a capacidade de dois ou mais sistemas ou componentes de trocar informação entre si e utilizar a informação trocada sem um gasto significativo de recursos em quaisquer dos sistemas”<sup>5</sup> (COMMITTEE ON CATALOGING: DESCRIPTION & ACCESS, 2000).

Interoperabilidade é um conceito que pode ter diferentes níveis de aplicação. Esses níveis podem ser relativos aos sistemas, à sintaxe, à estrutura e à semântica. Em cada um desses níveis, temos diferentes problemas a serem enfrentados. No nível da interoperabilidade de sistemas, podemos dizer que enfrentamos incompatibilidades básicas entre os diferentes *hardwares*, sistemas operacionais e *softwares* em geral, assim como na parte estrutural, nos diferentes modelos e estruturas de dados e esquemas. Já com relação aos problemas da interoperabilidade sintática, podemos resumir as diferenças relativas às codificações e representações. Finalmente, a interoperabilidade semântica apresenta seus problemas com relação às inconsistências na terminologia e significados (SHETH, 1998; OUKSEL; SHETH, 1999).

Outra classificação para tipos de interoperabilidade é apresentada pela Organização Europeia de Interoperabilidade (*European Interoperability Framework* ou EIF), que destaca quatro níveis ou camadas de interoperabilidade, ou seja, interoperabilidade legal, interoperabilidade organizacional, interoperabilidade semântica e interoperabilidade técnica (EUROPEAN COMMISSION, 2017).

Nesse contexto, a interoperabilidade legal aponta para uma coerência ou alinhamento entre as questões legais que, no caso em questão, difundem-se entre diferentes países, estados-membros e regiões. A interoperabilidade organizacional se refere a como são alinhados ou integrados os processos de negócios entre as diferentes organizações e como são trocadas as informações relevantes. A interoperabilidade técnica cobre as aplicações e infraestruturas que ligam sistemas e serviços. Os aspectos da interoperabilidade técnica incluem especificações de interface, serviços de interconexão e integração, entre outros (EUROPEAN COMMISSION, 2017).

Por fim, no modelo do EIF, temos a discussão da interoperabilidade semântica, que trata do aspecto específico que iremos enfrentar, conforme definido pelo *European Interoperability Framework*.

---

5 Tradução livre dos autores. No original: “Interoperability is the ability of two or more systems or components to exchange information and use the exchanged information without special effort on either system”.

Nesse caso, esta camada de interoperabilidade procura garantir o preciso formato e significado dos dados trocados, de forma que a informação seja preservada e compreendida entre as partes. Nesta visão, a interoperabilidade semântica cobre ambos os aspectos, sintáticos e semânticos. O aspecto semântico se refere ao significado dos elementos e suas relações. Inclui desenvolver vocabulários e esquemas para descrever troca de dados e procura garantir que os dados sejam compreendidos da mesma forma por todos os envolvidos na comunicação. O aspecto sintático se refere à capacidade de descrever a forma exata da comunicação a ser trocada nos termos de sua gramática e formato (EUROPEAN COMMISSION, 2017).

O *European Interoperability Framework* (2017) dispõe, ainda, que a base para aperfeiçoar a interoperabilidade semântica (considerando os pressupostos acima do que seja esta camada de interoperabilidade) é perceber dados e informações como poderosos ativos. Os acordos relativos aos dados de referência, na forma de vocabulários controlados, taxonomias e tesouros são pré-requisitos para a obtenção da interoperabilidade semântica. Da mesma forma que os padrões técnicos promoveram a interoperabilidade técnica por décadas atrás, padrões e especificações coerentes, robustas e universais são necessários para os desafios de permitir as trocas de informações com significado e semântica. Esta camada semântica de interoperabilidade, dadas as significativas diferenças linguísticas, culturais etc., é que apresenta maiores desafios. Por isso, sem que existam esforços na maturação na camada de interoperabilidade semântica será difícil garantir troca de informação, portabilidade e movimento livre de dados.

Neste ponto, começamos a discutir uma gradação de interoperabilidade que tenha por objetivo não os sistemas e ou suas estruturas, mas sim entre SOCs. Enquanto até então nos preocupávamos com a compatibilização entre dados ou metadados, agora precisamos explicitar como se dá esta definição do que consideramos interoperabilidade ao nível dos SOCs utilizados nos repositórios heterogêneos de dados, com vistas a um tratamento e recuperação inteligente de dados e informações.

Para uma definição de interoperabilidade no escopo dos SOCs e reconhecida como um padrão, podemos citar a ISO 25964, nomeada como 'Tesouros e Interoperabilidade com outros Vocabulários'. Esta norma define interoperabilidade como "a habilidade de dois ou mais sistemas trocarem informação e terem a capacidade de usar a informação que foi trocada" (INTERNATIONAL STANDARD ORGANIZATION, 2013).

Marcia Zeng, em seu artigo Interoperabilidade (ZENG, 2019), relata que, nos domínios relacionados, a Organização do Conhecimento, o conceito de interoperabilidade aborda três pontos

focais: sintático, estrutural e semântico e, citando Joudrey e Taylor (2017, p. 189), afirma que “sem interoperabilidade em todos os três níveis, os metadados não podem ser compartilhados sem esforço, de forma eficiente, ou lucrativamente”. E ressalta que, sem interoperabilidade sintática, dados e informações não podem ser tratados adequadamente em relação aos formatos, codificações, propriedades, valores, tipologia de dados; e, portanto, eles não podem ser integrados e nem trocados. Sem interoperabilidade semântica, o significado da linguagem, terminologia e valores de metadados usados não pode ser negociado ou corretamente compreendido. Vários graus de expressividade semântica podem ser combinados com diferentes tipos: baixo na interoperabilidade sintática, médio em interoperabilidade estrutural, e alto ou muito alto na interoperabilidade semântica.

Ao tratarmos de vocabulários controlados, tesouros, taxonomias e sistemas de classificação, percebemos que, do mesmo modo, os diferentes sistemas de organização do conhecimento vão invariavelmente diferir uns dos outros pela sua estrutura, pela semântica, pelos aspectos léxicos, ou mesmo pelos aspectos notacionais. Essas diferenças poderão abranger diferentes domínios de assunto, diferentes escopos e diferentes abrangências, poderão ter diferenças semânticas devido a variações em sua estrutura conceitual. Além disso, o uso da terminologia poderá variar, assim como seus níveis de especificidade. As chances são tais como: se existirem dez diferentes sistemas, existirão dez diferentes orientações e especificidades (IYER; GIGUERE, 1995).

Chan e Zeng (2002) defendem que essas incompatibilidades apresentam problemas básicos para a compatibilização e a interoperabilidade de vocabulários e repositórios, ou seja, estabelecer tradução entre tesouros, sistemas de classificação em vários sistemas diferentes, para estes autores, pode se tornar extremamente desafiador ou até mesmo impossível.

Podemos identificar, a partir daqui, então, dois níveis de interoperabilidade que são de especial interesse para nosso estudo, cada um com características específicas com relação às estratégias utilizadas e com as implicações que têm como relação as características específicas de recuperação, a interoperabilidade baseada em palavras ou a interoperabilidade baseada em conceitos, ou seja, a interoperabilidade sintática e a interoperabilidade semântica, respectivamente (HUBRICH, 2010).

No caso da interoperabilidade sintática, é necessário que os dados e as informações sejam tratados de acordo com o respeito aos formatos, às codificações, às propriedades, aos valores e aos tipos de dados. Sem essa compatibilização, esses dados não poderão ser combinados e nem ao menos trocados entre fontes diferentes. A interoperabilidade sintática, baseada em palavras, pauta-se basicamente pela comparação de cadeias de caracteres. A similaridade de cadeias de caracteres

pode ser analisada e determinada por um tipo de alinhamento léxico, responsável por permitir afirmar que duas cadeias de caracteres se equivalem. O uso de dicionários e vocabulários controlados pode aumentar a efetividade desse ambiente. Mas, devido à sua alta imprecisão, esse tipo de relação entre sistemas baseada em palavras não pode atender adequadamente àqueles que procuram explorar conhecimento entre diferentes bases de dados.

O problema da heterogeneidade semântica é comum em praticamente todas as áreas do conhecimento humano e deve ser a principal questão a ser considerada ao se discutir a interoperabilidade semântica, que surge da necessidade de se compartilhar conteúdos e dados entre ecossistemas<sup>6</sup> digitais diferentes e dentro deles. Nesses sistemas, uma combinação de diferentes aplicações e fontes de dados é utilizada, tendo como característica normalmente serem descritos em múltiplos esquemas, com vocabulários com semânticas diferentes e desenvolvidas localmente, implementados com diferentes linguagens de marcação e baseados em modelos que espelham diferentes conceituações. Além disso, os dados costumam ser representados em diferentes níveis de granularidade, cujas semânticas são baseadas em diferentes modelos, resultando que a informação não possa ser compartilhada entre aplicações de *software* e as fontes de dados não possam ser combinadas (FRITZSCHE *et al*, 2017).

Portanto, podemos dizer que alcançar a interoperabilidade semântica significa habilitar diferentes agentes, serviços e aplicações para trocar informação, dados e conhecimento, de forma significativa, tanto na *Web* como fora dela. Para que seja possível habilitar esse tipo de interoperabilidade, os agentes envolvidos precisam compartilhar o mesmo vocabulário, compreendido por todos, ou criar correspondências ou mapeamentos entre diferentes vocabulários (USHOLD; MENZEL, 2005). Ou seja, temos duas estratégias básicas a percorrer: uma pela criação de um vocabulário comum, quando há possibilidade por meio de uma decisão de autoridade, e outra por meio do estabelecimento de correspondências.

Nesse sentido, temos que a interoperabilidade semântica a ser alcançada é definida e obtida pela identidade ou por um forte alinhamento no significado dos conceitos. Dessa forma, diferentemente da estratégia baseada em palavras e cadeias de caracteres, podemos obter uma alta precisão nas relações intersistemas predefinidas. É importante lembrar que neste modelo de estabelecimento de correspondências e mapeamentos temos ainda uma alta demanda na experiência e no conhecimento das pessoas responsáveis na criação dos vínculos.

---

<sup>6</sup> Utilizamos aqui ecossistema num significado ampliado em relação à biologia e à ecologia, como vem sendo largamente empregado, aplicando-se a qualquer sistema ou rede de elementos e partes que interagem.

Nesse sentido, como citado acima, visando a uma recuperação com mais precisão, a Parte 2 da Norma ISO 25964-2, *Information and documentation — Thesauri and interoperability with other vocabularies*, trata de questões que envolvem tesauros e outros tipos de vocabulários que são comumente usados para recuperação de informações, descrevendo, comparando e contrastando os elementos e as características desses vocabulários que são implicados quando a interoperabilidade é necessária. Ela apresenta recomendações para o estabelecimento e a manutenção de mapeamentos entre vários tesauros, ou entre tesauros e outros tipos de vocabulários. O principal objetivo da interoperabilidade entre vocabulários é permitir que uma expressão formulada usando um vocabulário seja convertida em uma expressão correspondente em um ou mais vocabulários, tenham ou não o mesmo idioma. Para a realização desse processo, a chave é o mapeamento, ou seja, se cada um dos conceitos em um vocabulário foi mapeado para o conceito ou conceitos correspondentes em outro vocabulário, torna-se possível interconectar os termos representando este conceito em cada um deles. Podemos dizer, portanto, que essa interoperabilidade é habilitada ao se estabelecer mapeamentos interconceituais, em especial, mas não apenas equivalências (INTERNATIONAL STANDARD ORGANIZATION, 2013).

Entretanto, apesar da norma ISO citada ter papel importante na definição dos diferentes processos de mapeamento, por outro lado não apresenta procedimentos e metodologias para a realização desses processos. Nesse sentido, na seção 4, vamos tratar de uma proposta de procedimento apresentada por Dahlberg em seus estudos sobre Sistemas Ordenados (DAHLBERG, 1981, 1983, 1996). No entanto, antes, na próxima seção, vamos apresentar como o conceito de Mapeamento se insere nas discussões que envolvem a identificação de potencialidades semânticas entre SOCs.

### **3 O MAPEAMENTO EM SOCS**

Estabelecer mapeamento entre SOCs possibilita verificar o grau de compatibilidade semântica desses Sistemas. O mapeamento pode envolver diferentes estratégias para estimar este grau de compatibilidade nos termos que fazem parte de tais Sistemas e permite que se possa ter uma compreensão mais precisa das possibilidades de interoperabilidade semântica, revelando similaridades conceituais e ou mesmo novos relacionamentos para os termos em análise.

Neste sentido, Zeng (2019) afirma que o mapeamento é o processo de estabelecer relações entre os conceitos de um vocabulário e outros e é usado largamente para obter a interoperabilidade semântica nos SOCs existentes, sendo que o termo “*mapeamento*” pode ser usado para se referir ao

processo de estabelecer estas relações como também pode ser usado para registrar o produto do processo de mapeamento, ou seja, uma declaração das relações encontradas entre um vocabulário e os outros participantes.

O mapeamento é o primeiro passo para se estabelecer a interoperabilidade semântica entre SOCs e, a partir dele, verifica-se a possibilidade de integração desses vocabulários ou mesmo de estabelecimento de um dispositivo de comutação que permita navegar de um SOC a outro, que foi denominado por Horsnell (1975) de Linguagem intermediária.<sup>7</sup>, com a função de recuperar informação em bases distribuídas indexadas por vocabulários heterogêneos.

A norma ISO 25964-2 apresenta diferentes modelos estruturais para mapeamento entre vocabulários, a saber, Unidade estrutural, Ligação direta e Estrutura central<sup>8</sup> (INTERNATIONAL STANDARD ORGANIZATION, 2013).

No modelo de unidade estrutural, todos os vocabulários participantes do mapeamento têm a mesma estrutura hierárquica e relações associativas entre os conceitos, podendo ser representada em diferentes idiomas ou notações. Este modelo, pela sua simplicidade, normalmente não envolve mapeamentos, mas é incluído pela norma para compreensão e comparação com outros modelos.

O modelo de ligação direta estabelece ligações entre dois ou mais vocabulários que não têm a mesma estrutura. Esses vocabulários podem diferir, além de sua estrutura, em seu escopo e idioma, e podem incluir diversos tipos, como: esquemas de classificação, lista de autoridades de nomes, bem como outros tesouros. Os mapeamentos são estabelecidos entre os conceitos de um determinado vocabulário e cada um dos outros participantes, podendo o processo ser estendido para qualquer número de vocabulários. A partir desse processo, qualquer um deles pode ser usado para realizar uma busca na coleção indexada por qualquer outro.

O modelo de estrutura central é aquele em que, em vez de se estabelecerem mapeamentos entre todos os pares de vocabulários, elege-se um vocabulário como central, para o qual todos são mapeados. Neste modelo, cada conceito no vocabulário central deve ser mapeado para os possíveis correspondentes conceitos nos outros vocabulários e, de mesma forma, cada conceito nos vocabulários satélites deve ser mapeado para o vocabulário central. Este modelo é capaz de permitir que o vocabulário central possa ser usado para buscar quaisquer dos recursos indexados pelos vocabulários satélites, seja (i) pela conversão da frase de busca ou (ii) pela conversão dos termos indexados no vocabulário

7 Para um maior entendimento sobre Linguagem Intermediária, o seguinte trabalho é recomendado: HORSNELL, Verina. The intermediate lexicon: an aid to international co-operation. **Aslib Proceedings**, v. 27, n. 2, p. 57-66, feb. 1975.

8 No original Structural unity, Direct-Linked e Hub structure.

destino. De modo análogo, este processo pode ser realizado num vocabulário satélite e sofrer a conversão para realizar a busca na coleção do vocabulário central. Sem prejuízo desta análise, podemos ainda ter uma busca iniciada em um vocabulário satélite que precise de realizar duas conversões para acessar coleções em outro vocabulário satélite. Nesse caso, se os dois passos envolvem equivalências simples e exatas, é possível que o mapeamento forneça resultados aceitáveis, mas caso envolva equivalências compostas ou mapeamentos inexatos, a qualidade dos resultados não pode ser prevista (INTERNATIONAL STANDARD ORGANIZATION, 2013).

Dáí o problema que é possível identificar neste processo, ou seja, para obter mapeamentos de boa qualidade que atendam a todo o conjunto de vocabulários e necessidades envolvidas, o vocabulário central precisa incorporar todos os conceitos presentes nos vocabulários satélites. Portanto, quando não há um vocabulário que se preste a realizar esta função dentre aqueles que precisam participar do processo de interoperabilidade, é necessário construir um novo completo que sirva a este propósito, podendo também se realizar uma extensão de um deles, de forma a atender ao escopo e à especificidade dos outros combinados.

Quanto aos tipos de mapeamento, a norma apresenta os mapeamentos por equivalência, os hierárquicos, os associativos e os exatos, inexatos e parciais. Os mapeamentos por equivalência são estabelecidos quando conceitos correspondentes são encontrados em dois (ou mais) vocabulários. Os hierárquicos são os que podem ser estabelecidos entre conceitos quando um é mais abrangente que outro, podendo ser de tipo genérico, parte-todo ou de instância. Os associativos são aqueles quando não é possível estabelecer mapeamentos por equivalência ou por hierarquia, mas os conceitos são semanticamente associados, de forma que objetos indexados por um deles são de mesma forma relevantes se indexados pelo outro. Os mapeamentos, variando com o grau de heterogeneidade dos vocabulários envolvidos, podem ser mais ou menos perfeitos. Para isso, podem ser divididos em exatos, inexatos ou parciais.

Portanto, como vimos, todas as relações que são estabelecidas entre SOCs são chamadas de mapeamentos, e estes são os mais comumente utilizados.

Para Zeng (2019), detalhando mais as situações vistas pela norma ISO mostrada, durante o processo de mapeamento e mesclagem de equivalências conceituais, são procurados *nomens*<sup>9</sup>

---

9 *Nomem* é definido como sendo uma associação entre uma entidade e uma designação que se refere a ela e é a denominação usada para se referir a uma instância de qualquer entidade no universo do discurso. Pode ser qualquer signo ou sequência de signos (caracteres alfanuméricos, símbolos etc.) pelo(a) qual uma entidade é referida, conhecida ou endereçada (ZENG; ZUMER; SALABA, 2010).

específicos que representam os conceitos, formando sintaxes definidas. Desse processo, podem surgir vários tipos de relações semânticas, englobando situações como: correspondência exata, específico para geral, geral para específico, muitos para um, mapeamentos cíclicos, e não correspondência. Aitchison, Gilchrist e Bawden (1997) já apontavam também diferentes tipos de equivalências definidos pelos padrões: equivalência exata, equivalência inexata (ou quase-equivalência), equivalência parcial, equivalência um para muitos e não equivalência. Já nos meados do século XX, alguns autores da CI e da CC diziam que, apesar de ser possível a criação de algoritmos computacionais que realizem de forma automática pelo menos parte desse processo, mapeando inicialmente as combinações exatas, e posteriormente as variações gráficas e palavras variantes na forma, a tarefa intelectual de conversão por este método é altamente consumidora de tempo e de pessoal especializado, em especial no caso de vários SOCs (LANCASTER, 1986; WALL; BARNES, 1969).

Nesse sentido, é que consideramos investigar a proposta metodológica de Dahlberg (1981, 1983), que se pauta em apresentar procedimentos visando promover o mapeamento em nível semântico que requer um grau de análise conceitual e compreensão que, na época de sua criação, era avançado demais para ser realizado de forma automatizada, mas que hoje pode contribuir de forma significativa na formulação de processos e algoritmos automatizados de correspondência e compatibilização de vocabulários. A seguir, apresentaremos os procedimentos metodológicos propostos por Dahlberg.

#### **4 MAPEAMENTO DAS POTENCIALIDADES SEMÂNTICAS ENTRE SOCS: uma investigação dos estudos de Dahlberg**

Mapear as potencialidades semânticas de diversos SOCs, ou como denominado por Dahlberg, Sistemas Ordenados<sup>10</sup>, requer procedimentos para a compreensão do significado dos termos usados nesses Sistemas, como, por exemplo: sua definição, o conceito mais alto na hierarquia e a categoria na qual este conceito pode estar inserido, que indica se é um objeto, um processo, uma qualidade. O mapeamento das potencialidades semânticas é um mecanismo que permite compatibilizar, ou seja, correlacionar os elementos desses diferentes Sistemas, como nos diz Dahlberg (1981), com o propósito de permitir o uso comum dessas linguagens, quando forem efetuadas buscas em diferentes bases de dados ou quando se deseja intercambiar literatura indexada com essas linguagens.

---

<sup>10</sup> Nesta seção, iremos utilizar a terminologia proposta à época por Dahlberg, o que atualmente denominamos de SOC, ou seja, Sistema Ordenado. De mesma forma, nas citações diretas, iremos utilizar a terminologia original da autora.

Como reportado acima, o mapeamento é o primeiro passo para que se possa identificar o nível de compatibilidade entre Sistemas Ordenados. Segundo Dahlberg (1981, p. 87), uma vez estabelecida a compatibilidade, seremos capazes de

... procurar sem problemas em uma grande variedade de arquivos diferentes, com quaisquer dos termos compatibilizados;

recuperar a informação de qualquer local, desde que rotulada com um conceito específico e indexada pelos termos dos sistemas envolvidos na compatibilização;

informar ao usuário de uma certa Linguagem de Indexação da disponibilidade de informação no tópico de seu interesse, indexada com outras Linguagens de Indexação.

obter os equivalentes de um sistema em uma linguagem natural (por exemplo, inglês) em outra linguagem natural (por exemplo, alemão); dessa forma o sistema de comutação pode também atuar como um tesouro multilíngue.

Nesse sentido, no âmbito da compatibilização de linguagens, Dahlberg (1983) alerta para a importância de se trabalhar não só no plano do casamento verbal ou linguístico dos termos, mas também na análise dos conceitos, para que uma correspondência semântica possa ser estabelecida.

Devemos ter em mente que a correlação de sistemas ordenados não pode ser realizada apenas no nível da linguagem, ou seja, no nível dos termos e seus descritores. [...] Os termos são apenas os portadores da informação e só podem exibir a informação sobre o conteúdo do conceito representado se as relações necessárias tiverem sido estabelecidas através de uma definição. [...] O que é necessário, então, é a fixação da relação entre os objetos do mundo externo e/ou nossos construtos mentais de tais objetos, e os elementos da linguagem que os denotam. [...] é o conhecimento de nossos conceitos que está envolvido ao lidar com os problemas de comparações e compatibilidade entre sistemas ordenados (DAHLBERG, 1983, p. 5).

Dahlberg apresenta em seus trabalhos (DAHLBERG, 1981, 1983) um caminho metodológico para a realização de mapeamentos entre Sistemas Ordenados, que nós consideramos apresentar aqui visando ressaltar a importância desta etapa inicial. É a partir do mapeamento das potencialidades linguísticas e semânticas que consideramos ser possível a tomada de decisão quanto à possibilidade de integrar tais Sistemas Ordenados ou mesmo mantê-los interligados através de Léxicos Intermediários, como citado na seção 3, ou mesmo como nos fala Levy (2014), através de um Sistema de Coordenadas Semânticas.

Nesse sentido, Dahlberg (1981) apresenta procedimentos metodológicos que se configuram em etapas para analisar as potencialidades de compatibilização entre Sistemas Ordenados sob dois aspectos: 1. Procedimentos para o estabelecimento de comparações verbais entre Sistemas Ordenados; 2. Procedimentos para o estabelecimento de comparações conceituais entre Sistemas Ordenados. A

seguir, apresentaremos cada um deles apoiados e adaptados a partir da proposta de Dahlberg (1981, 1983), que sofreram inclusões e alterações a partir das pesquisas realizadas nos últimos anos, em que consideramos diferentes aplicações desta metodologia (BARBOSA; CAMPOS, 2018; CAMPOS, 2005; CAMPOS *et al*, 2009).

#### 4.1 Procedimentos para o Estabelecimento de Comparações Verbais entre Sistemas Ordenados

O primeiro movimento para iniciar uma atividade de compatibilização entre Sistemas Ordenados é entender que tal atividade é de natureza analítico-conceitual, necessitando que o profissional que se dedique a ela deva ter, além de um bom conhecimento do domínio temático abrangido pelos Sistemas Ordenados, conhecimento da estrutura conceitual desses Sistemas.

Assim, antes de começar qualquer procedimento de compatibilização, é necessário obter uma visão geral dos termos que se sobrepõem entre os Sistemas Ordenados em análise. Para tanto, é realizada uma análise do casamento verbal dos termos, o qual pode ser feito por meios computacionais. Se apenas uma pequena percentagem dos termos se sobrepõe, então a comparação por inteiro deve ser reconsiderada, pois pode-se verificar que os Sistemas Ordenados em análise não possuem semelhanças a nível do domínio de conhecimento que estão tratando, no qual a terminologia é diferente, possuindo aspectos temáticos que os diferenciam.

Para a tomada de decisão acima, é necessário calcular o grau de compatibilidade dos termos entre os diversos Sistemas Ordenados. Para tanto, Dahlberg propõe a geração de uma Matrix Alfabética que deverá ser constituída de colunas com os seguintes conteúdos: 1. Número sequencial, 2. Termo (lista de todos os termos existentes nos Sistemas Ordenados considerados), 3. Sistema Ordenado em análise, no qual será marcada a coincidência entre os termos ocorrida nestes Sistemas (aqui serão quantas colunas quantos forem os Sistemas Ordenados), 4. Totalização de termos coincidentes encontrados nos Sistemas Ordenados (chamada de CV = Coincidência Verbal), assim como indicado na Tabela 1.

**Tabela 1 – Matriz Alfabética**

Número Sequencial	Termo	Marcação do Conceito			Número de Coincidências Verbais
		Sistema 1	Sistema 2	Sistema n	

**Fonte:** elaborada pelos autores com dados extraídos de Dahlberg (1983).

Considerando-se a matriz alfabética gerada, Dahlberg propõe, então, uma medida de quantificação para estabelecer o grau de compatibilidade entre os sistemas envolvidos no processo. Esta medida, chamada por ela de taxa de coincidência verbal, pode ser calculada da seguinte forma: o somatório do total de valores da coluna referente ao item 4 acima (CV = coincidência verbal) dividido pelo total de possíveis coincidências (total de termos multiplicado pelo número de sistemas ordenados em análise). Valores iguais ou superiores a 0,50 para esta taxa indicam um grau relativamente alto de compatibilidade entre os sistemas analisados.

Entretanto, esta compatibilidade verbal, que só pode ser medida para Sistemas Ordenados de um mesmo idioma, não nos diz se: (1) os termos coincidentes significam realmente os mesmos conceitos (pode haver homonímia e polissemia); (2) os Sistemas Ordenados podem usar modos ligeiramente diferentes de nomear; (3) pode haver um grau mais alto de compatibilidade entre Sistemas Ordenados se considerarmos os conceitos em vez dos nomes no sistema de posições; (4) os conceitos que diferem em especificidade poderiam entrar na comparação para serem usados como substitutos para algum conceito que esteja ausente.

Devido a estas razões, uma matriz de compatibilidade puramente verbal não será suficiente para ajudar a estabelecer compatibilidade, ainda que ela certamente sirva como uma etapa preliminar necessária.

Nesse sentido, após considerarmos a taxa de coincidência verbal, se a sobreposição for considerável, então será necessário estabelecer procedimentos que visem verificar o nível de coincidência conceitual entre os termos, que veremos adiante.

## **4.2 Procedimentos para Comparação Conceitual entre Sistemas Ordenados**

O primeiro movimento para a comparação conceitual entre Sistemas Ordenados é realizar a descrição conceitual dos termos. Essa descrição permite a verificação do nível de semelhança conceitual existente entre os termos e seus conteúdos conceituais. Nesse sentido, Dahlberg apresenta uma lista de elementos que devem ser anotados sobre o termo em análise, que ela irá denominar de Registro de Conceito, para que dois Sistemas Ordenados possam ser mapeados. O Registro de conceitos pode ser definido como um dispositivo que mapeia as potencialidades semânticas de um termo em um dado Sistema Ordenado.

Ao estabelecer um Sistema Ordenado, “é preciso acumular o conhecimento necessário sobre os referentes através do registro do conceito. Se um sistema ordenado não foi desenvolvido com ajuda de um registro de conceitos, é necessário estabelecê-lo ao se comparar os sistemas ordenados” (DAHLBERG, 1983, p. 6).

Consideramos, a partir dos estudos realizados, que o Registro do Conceito, com algumas semelhanças com fichas terminológicas, deva ser constituído dos campos vistos no Quadro 1.

**Quadro 1 – Registro do Conceito**

00	Código do Conceito
01	Denominação do conceito
02	Fonte do conceito (nome do Sistema Ordenado ou um código definido para indicá-lo)
03	Notação do conceito (quando for o caso de se compatibilizar Sistemas de Classificação)
04	Definição do conceito e com indicação das fontes
05	Denominação do próximo conceito mais amplo
06	Denominação do conceito mais alto na hierarquia
07	Indicação do nível hierárquico do conceito (formando uma cadeia de conceitos): Nível mais alto, próximo nível mais baixo, terceiro nível etc.
08	Denominação do conceito mais alto na hierarquia
09	<p>Categoria da forma do conceito</p> <p>(O) Objeto, entidade</p> <p>(P) Processo, atividade</p> <p>(Q) Quantidade, qualidade</p> <p>(R) Relação</p> <p>(S) Conceito relacionado a espaço</p> <p>(T) Conceito relacionado a tempo</p> <p>(W) Campo de assunto ou disciplina</p>
10	Outros nomes do conceito (sinônimos)
11	Abreviações do Termo
12	Comentários relativos aos conceitos correspondentes em outros Sistemas Ordenados

**Fonte:** elaborada pelos autores com dados extraídos de Dahlberg (1983).

Após a realização do Registro do Conceito de cada termo dos Sistemas Ordenados em comparação, é necessário realizar uma análise desses Registros para verificar o grau de semelhança conceitual entre os termos de mesma grafia identificados nesses Sistemas. Esta comparação, que deve ser feita apoiada por especialistas, permite eliminar as discrepâncias resultantes da análise verbal feita anteriormente e mesmo identificar semelhanças conceituais entre termos que não possuem a mesma grafia. Um exemplo: no caso de ocorrer semelhança verbal entre termos de dois Sistemas ordenados, é possível, por meio da análise da estrutura hierárquica e de sua definição, concluir se os termos são ou não correspondentes ao mesmo conceito.

Dessa forma, para a realização da análise dos Registros dos Conceitos, Dahlberg (1983) apresenta três critérios que podem ser adotados e que permitem identificar o nível de compatibilidade entre os Sistemas Ordenados em análise:

- i. coincidência conceitual – quando dois conceitos combinam suas características – nesse caso, são equivalentes e possuem um grau de equivalência;
- ii. correspondência conceitual – quando dois conceitos combinam a maior parte de suas características – nesse caso, são similares;
- iii. correlação conceitual – quando dois conceitos são correlacionados quando possuem diferentes níveis de detalhe, ou quando a relação entre eles não é de semelhança.

O tipo de correlação pode ser apontado por meio de símbolos, por exemplo: “<” e “>” para indicar diferentes níveis de detalhamento, “C” para indicar que um conceito em uma linguagem equivale a uma combinação de conceitos na outra.

A partir de tais critérios, é possível elaborar uma Matriz de Compatibilidade Semântica (DAHLBERG, 1981), em que é possível medir o grau de compatibilidade entre os Sistemas Ordenados em análise. A Matriz de Compatibilidade Semântica é formada pelos seguintes campos:

1. Nome dos conceitos do Sistema Ordenado considerado como o mais abrangente (também chamado por Dahlberg de Mestre) ou Nome dos Conceitos gerados a partir da comparação entre os Sistemas Ordenados que são considerados os de nível mais alto na hierarquia (quando não se considera nenhum dos Sistemas Ordenados como mais abrangente);

2. Nome dos Sistemas Ordenados em análise (aqui serão elaboradas tantas colunas quantos forem os Sistemas Ordenados e, para cada sistema ordenado, será colocado o termo equivalente identificado); 3. Número total de coincidências conceituais identificadas relativo ao termo em análise e à presença nos demais Sistemas Ordenados, assim como indicado na Tabela 2.

**Tabela 2** – Matriz de Compatibilidade Semântica

Nome do Conceito	Conceitos Coincidentes			Número de Coincidências Conceituais
	Sistema 1	Sistema 2	Sistema n	

**Fonte:** elaborada pelos autores com dados extraídos de Dahlberg (1983).

A Matriz de Compatibilidade Semântica, segundo Dahlberg (1981), funcionaria como uma “caixa preta”, em que se deve entrar com um termo em um Sistema Ordenado para obter o termo correspondente nos outros Sistemas Ordenados envolvidos.

Ainda quanto ao processo de mapeamento das potencialidades semânticas aqui apresentado, consideramos que ele possa permitir uma remodelação e atualização dos SOCs em análise no que tange à compatibilidade no nível verbal. Cada SOC pode ser completado pelas alternativas de termos sinônimos identificados nos outros Sistemas Ordenados e assim melhorar o acesso verbal a essas classes e seus conceitos, por exemplo.

Atualmente, os recursos tecnológicos existentes permitem que esses processos de mapeamento semântico possam ser cada vez mais automatizados, realizados por agentes de *software*, e é esta investigação que estamos empenhados em verificar em nossa linha de pesquisa, ou seja, em que ponto estamos deste processo e quais caminhos precisamos seguir. Considerando a premissa de que a solução para a interoperabilidade de sistemas heterogêneos passa por uma visão interdisciplinar, especialmente, mas não só, entre a ciência da informação e a ciência da computação, colocamos em nossas considerações finais uma perspectiva para futuras pesquisas e estudos nessa área.

## 5 CONSIDERAÇÕES FINAIS

Desde o início dos estudos voltados para a compatibilidade de vocabulários, tesouros, taxonomias, sistemas ordenados, como denomina Dahlberg, ou SOCs, como denominamos atualmente, tivemos um crescimento enorme da capacidade computacional, que pode contribuir tanto com a capacidade de processamento quanto com o desenvolvimento de algoritmos, métodos e técnicas para automatizar este processo. Mas, se por um lado as tecnologias da informação nos dão essas facilidades, por outro, este mesmo crescimento computacional permite a disseminação e o crescimento exponencial de bases de dados, de diversos tipos e com enorme diversidade de linguagens de indexação, tornando o problema da interoperabilidade semântica de solução complexa e difícil.

A busca pela interoperabilidade semântica é uma necessidade que se estende a todos os domínios do conhecimento e impacta diretamente na capacidade de se recuperar informações de forma inteligente na *Web*, restringindo a possibilidade de uso amplo das informações distribuídas em bases heterogêneas. A Ciência da Computação tem expressiva contribuição neste processo com a construção de algoritmos e sistemas que permitam extrair este conhecimento das relações explicitadas nos sistemas de organização do conhecimento e imitar a capacidade humana de estabelecer o grau de relação entre diferentes ‘coisas’ de acordo com as evidências semânticas disponíveis, conseguindo definir, por exemplo, que em um domínio de Educação, as palavras escola, professor e aluno são mais relacionadas semanticamente do que escola, tijolo e cimento, mas poderia ser exatamente o contrário em um domínio de Construção Civil.

Destacamos aqui as pesquisas atualmente realizadas neste caminho para o estudo das medidas semânticas que podem ser usadas nos processos que buscam integrar ontologias heterogêneas, como alinhamento e mapeamento, e que podem se voltar para determinar similaridades semânticas em diferentes ontologias. Estas medidas semânticas são ferramentas matemáticas utilizadas para estimar a força da relação semântica entre unidades de linguagem, conceitos ou instâncias, através de uma descrição numérica obtida de acordo com a comparação de informações que sustentam seu significado e seu estudo está fortemente presente em um grande número de trabalhos da CC, entre os quais podemos destacar, a título de exemplo, Lu *et al* (2020), Deepak e Santhanavijayan (2020), Martinez-Gil e Chaves-Gonzalez (2020) e Hussain *et al* (2020).

Michael Buckland (BUCKLAND, 1999) já afirmava em 1999 que as transições e os mapeamentos entre vocabulários são uma questão de interesse central para a Ciência da Informação e, ao explicitarmos neste artigo uma proposta de metodologia e procedimentos para a compatibilização de SOCs proposta por Dahlberg, estamos afirmando que o estudo dos métodos propostos pela CI têm no mundo de hoje grande relevância na compreensão e na formulação de algoritmos, procedimentos e requisitos para um processo de compatibilização de linguagens que caminhe para a interoperabilidade e para uma visão semântica das informações produzidas pelos diversos domínios na *Web* e na Internet.

## **Agradecimentos**

Agradecemos ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq pelo apoio financeiro, e ao Grupo de Pesquisa Estudos Ônticos e Ontológicos em Contextos Informacionais -EООCI/UFF pelas discussões e trocas durante esta fase de nossa pesquisa.

## REFERÊNCIAS

AITCHISON, Jean; GILCHRIST, Alan; BAWDEN, David. **Thesaurus construction and use: a practical manual**. 3. ed. London: ASLIB, 1997.

BARBOSA, Nilson Theobald; CAMPOS, Maria Luiza de Almeida. A interoperabilidade semântica em repositórios institucionais de dados de pesquisa. *In: CAMPOS, Maria Luiza de Almeida et al. (org.). **Produção, tratamento, disseminação e uso de recursos informacionais heterogêneos**: diálogos interdisciplinares*. Niterói: IACS/UFF, 2018. p. 236-240. (Série Estudos da Informação, 5). Disponível em: <http://ppgci.uff.br/wp-content/uploads/sites/86/2019/11/PPGCI-VSEI-livro5.pdf>. Acesso em 02 set. 2020.

BUCKLAND, Michael. Vocabulary as a central concept in library and information science. *In: THIRD INTERNATIONAL CONFERENCE ON CONCEPTIONS OF LIBRARY AND INFORMATION SCIENCE (CoLIS3)*, 3., 1999, Dubrovnik. **Proceedings** [...]. Lokve: Benja Publishing, 1999. p. 3-12. Tema: Digital libraries: interdisciplinary concepts, challenges and opportunities.

CAMPOS, Linair Maria et al. Aspectos metodológicos no reuso de ontologias: um estudo a partir das anotações genômicas no domínio dos tripanosomatídeos. **RECIIS: Revista Eletrônica de Comunicação, Informação & Inovação em Saúde**, v. 3, n. 2, p. 64-75, 2009. Disponível em: <https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/820>. Acesso em: 02 set. 2020.

CAMPOS, Maria Luiza de Almeida. Integração de ontologias: o domínio da bioinformática e a problemática da compatibilização terminológica. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 6., 2005, Florianópolis. **Anais** [...]. Florianópolis: ANCIB, 2005.

CHAN, Lois Mai; ZENG, Marcia Lei. Ensuring interoperability among subject vocabularies and knowledge organization schemes: a methodological analysis. *In: IFLA GENERAL CONFERENCE AND COUNCIL*, 68., 2002, Glasgow. **Proceedings** [...]. Glasgow : IFLA, 2002. p. 1-8. Disponível em: <https://archive.ifla.org/IV/ifla68/papers/008-122e.pdf>. Acesso em: 07 set. 2020.

CHAN, Lois Mai; ZENG, Marcia Lei. Metadata interoperability and standardization: a study of methodology part I. **D-Lib Magazine**, v. 12, n. 6, jun., 2006. Disponível em: <https://www.dlib.org/dlib/june06/chan/06chan.html>. Acesso em: 02 set. 2020.

COMMITTEE ON CATALOGING: DESCRIPTION & ACCESS. **Task force on metadata**: final report. [S. l.: s. n.], 2000. (CC:DA/TF/Metadata/5). Disponível em: <https://www.libraries.psu.edu/tas/jca/ccda/tf-meta6.html>. Acesso em: 02 set. 2020.

DAHLBERG, Ingetraut. Compatibility and integration of order systems 1960-1995: an annotated bibliography. *In: COMPATIBILITY AND INTEGRATION OF ORDER SYSTEMS: RESEARCH SEMINAR OF THE TIP/ISKO MEETING*, 1995, Warsaw. **Proceedings** [...]. Warsaw: Wydawnictwo, 1996.

DAHLBERG, Ingetraut. Conceptual compatibility of ordering systems. **International Classification**, v. 10, n. 1, p. 5-8, 1983. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-1983-1-5/conceptual-compatibility-of-ordering-systems-volume-10-1983-issue-1>. Acesso em: 02 set. 2020.

DAHLBERG, Ingetraut. Towards establishment of compatibility between indexing languages. **International Classification**, v. 8, n. 2, p. 88-91, 1981.

- DEEPAK, Gerard; SANTHANAVIJAYAN, A. OntoBestFit: a best-fit occurrence estimation strategy for RDF driven faceted semantic search. **Computer Communications**, v. 160, p. 284-298, jul. 2020. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0140366420301572?via%3Dihub>. Acesso em: 07 set. 2020.
- EUROPEAN COMMISSION. **Communication from the commission to the european parliament, the council, the european economic and social committee and the committee of the regions: European Interoperability Framework – implementation strategy**. Brussels: COM, 2017. Disponível em: [https://eur-lex.europa.eu/resource.html?uri=cellar:2c2f2554-0faf-11e7-8a35-01aa75ed71a1.0017.02/DOC\\_1&format=PDF](https://eur-lex.europa.eu/resource.html?uri=cellar:2c2f2554-0faf-11e7-8a35-01aa75ed71a1.0017.02/DOC_1&format=PDF). Acesso em: 07 set. 2020.
- FRITZSCHE, Donna *et al.* Ontology summit 2016 communiqué: ontologies within semantic interoperability ecosystems. **Applied Ontology**, v. 12, n. 2, p. 91-111, 2017. Disponível em: <https://content.iospress.com/articles/applied-ontology/ao181>. Acesso em: 07 set. 2020.
- HORSNELL, Verina. The intermediate lexicon: an aid to international co-operation. **Aslib Proceedings**, v. 27, n. 2, p. 57-66, feb. 1975. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/eb050493/full/html>. Acesso em: 2 set. 2020.
- HUBRICH, Jessica. Intersystem relations: characteristics and functionalities. *In*: COLOGNE CONFERENCE ON INTEROPERABILITY AND SEMANTICS IN KNOWLEDGE ORGANIZATION, 2010, Cologne. **Proceedings** [...] Würzburg: Ergon-Verlag, 2010. p. 69-90. Tema: Concepts in Context.
- HUSSAIN, Muhammad Jawad *et al.* An approach for measuring semantic similarity between Wikipedia concepts using multiple inheritances. **Information Processing & Management**, v. 57, n. 3, p. 1-9, 2020. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0306457319307162>. Acesso em: 07 set. 2020.
- INTERNATIONAL STANDARD ORGANIZATION. **ISO 25964-2: Information and documentation — Thesauri and interoperability with other vocabularies — Part 2: Interoperability with other vocabularies**. Geneve: International Standard Organization, 2013.
- IYER, Hermalata; GIGUERE, Mark. Towards designing an expert system to map mathematics classificatory structures. **Knowledge Organization**, v. 22, n. 3/4, p. 141-147, 1995.
- JOUDREY, Daniel N.; TAYLOR, Arlene G. **The organization of information**. 4. ed. Santa Barbara: Libraries Unlimited, 2017.
- LANCASTER, Frederick Wilfrid. **Vocabulary control for information retrieval**. 2. ed. Arlington, Va.: Information Resources Press, 1986.
- LÉVY, Pierre. **A esfera semântica: computação, cognição, economia da informação**. São Paulo: Annablume, 2014. t. 1.
- LU, Wenpeng *et al.* Deep hierarchical encoding model for sentence semantic matching. **Journal of Visual Communication and Image Representation**, v. 71, p. 1-9, aug. 2020.

MARTINEZ-GIL, Jorge; CHAVES-GONZALEZ, Jose M. A novel method based on symbolic regression for interpretable semantic similarity measurement. **Expert Systems with Applications**, v. 160, 113663, p. 1-14, 2020.

NATIONAL INFORMATION STANDARDS ORGANIZATION. **Understanding metadata**. Bethesda: NISO Press, 2004.

NETWORKED KNOWLEDGE ORGANIZATION SYSTEMS. **Networked knowledge organization systems/services/structures**. [S. l.: s. n.], 2020. Disponível em: <https://nkos.slis.kent.edu/>. Acesso em: 02 set. 2020.

OUKSEL, Aris M.; SHETH, Amit. Semantic interoperability in global information systems. **ACM SIGMOD Record**, v. 28, n. 1, p. 5-12, mar. 1999.

PAEPCKE, Andreas *et al.* Interoperability for digital libraries worldwide. **Communications of the ACM**, v. 41, n. 4, apr. 1998. DOI: <https://doi.org/10.1145/273035.273044>

SHETH, Amit P. Changing focus on interoperability in information systems: from system, syntax, structure to semantics. *In*: GOODCHILD, Michael *et al.* **Interoperating geographic information systems**. Boston, Springer, 1998. p. 5-29.

USHOLD, Mike; MENZEL, Christopher (ed.). **Semantic integration & interoperability using RDF and OWL**. W3C Editor's Draft 3 November 2005. Disponível em: <https://www.w3.org/2001/sw/BestPractices/OEP/SemInt/>. Acesso em: 02 set. 2020.

WALL, Eugene; BARNES, J. M. **Intersystem compatibility and convertibility of subject vocabularies**. Technical Report n° 1582-100-TR-5. Philadelphia: Auerbach, 1969. (NTIS Document PB 184 144).

ZENG, Marcia Lei. Interoperability. **Knowledge Organization**, v. 46, n. 2, p. 122-146, 2019.

ZENG, Marcia Lei; CHAN, Lois Mai. Trends and issues in establishing interoperability among knowledge organization systems. **Journal of the American Society for Information Science and Technology**, v. 55, n. 5, p. 377-395, 2004.