

# Presentation

Dagobert Soergel  
PhD, University of Maryland

What follows are some remarks on the topics covered in the papers in this issue, culminating in a call for stronger collaboration among the communities dealing with issues of ontologies and other knowledge organization systems and an effort towards unification of the thinking and the results achieved by these communities.

I begin with recalling some distinctions to place the papers in this issue in context. These distinctions are pragmatic and useful, but they can be — and have been — debated endlessly, sometimes on logical grounds. Most of these debates I find decidedly not useful.

In dealing with ontologies and their use in representing data there are three levels:

The papers in this issue focus on Entity Value Ontologies. One gets the impression that for some authors what makes a system worthy of the name “ontology” depends not so much on the content or on the level (in the sense presented in the table), but rather on the degree of formality, such as being expressed in RDF, OWL, or some more expressive logic format. This stance should be revisited

what is important the content a system has to offer and how precisely that content is expressed (Things can be expressed very precisely in natural language).

Three levels in representing data		
<b>Data model ontologies</b>	Define entity types and relationship types. Entity types are classes, but not all classes are entity types. For example, taxon is an entity type and also a class the members of the class, or the entity values of the entity type, are taxa, such as <i>Rosa multiflora</i> <i>Rosa multiflora</i> is, in turn, a class whose members are individual plant specimens. Examples: Basic Format Ontology (BFO, including BFO relationships), Unified Foundational Ontology (UFO), FOAF, Dublin Core	Tbox Terminological axioms
<b>Entity value ontologies</b>	These systems deal with entity values – biological taxa, body parts and body systems (anatomy plus), genes, biological functions, diseases and disorders, places, historical periods, cultures, artistic styles, cultural objects, concepts – and terms used to name them. Such systems are also known as classification schemes, taxonomies, subject heading lists, or thesauri, collectively known as knowledge organization systems. Systems that use the term “ontology” in their name usually have a more elaborate structure with differentiated relationships. Examples: The Universal Decimal Classification (UDC), the Gene Ontology, (GO) SnoMed, the International Classification of Diseases (ICD), the Medical Subject Headings (MeSH), AGROVOC, the Thesaurus of Psychological Index Terms, the Getty vocabularies, including the Art and Architecture Thesaurus (AAT).	
<b>Data sets, data files, databases</b>	Data sets are here construed as collections of assertions about reality (a term used here naively without taking any philosophical stance) or representing thought or imagination and dreams. Assertions are formed from relationship types by filling in entity values for the relationship arguments. In their simplest form assertions take the form of triples, but representing the complexity of reality and thought requires multi-argument relationships. The examples in Hick’s paper on ontology evaluation (Paper I-2) are mostly assertion, indicating that she considers data sets as ontologies, at least if they are presented in a formal way.	Abox Assertional axioms
<b>Boundaries between levels are porous</b>	The boundaries between these levels are by no means strict. For example, the Unified Medical Language System (UMLS) Semantic Types list animal (a class) and even mammal and human as an entity type, while to me animal is a value of the entity type organism, albeit high up in the hierarchy of these entity values. Many thesauri, such as MeSH or the Alcohol and Other Drug Thesaurus include relationships that are assertions about reality. For example, in the MeSH hierarchy there is a heading anti-infective agents listing a chemical substance under this heading is an assertion about reality. MeSH includes such relationships because they are useful for searching. In a broad view, all three levels could be combined into one big data base without worrying which level an element belongs to. On the other hand, distinction between the levels is useful. Systems must decide pragmatically what information to include at what level	

The papers discuss design and structure, development (closely linked to design), evaluation, and application of ontologies, broadly construed.

Lima and Maculan (*Comparative study of semantic structures in different knowledge organization systems*) and Maculan and Aganette (*Ontologia aplicada a um modelo de gestão organizacional: contribuições da Ciência da Informação*) discuss the relational structure of ontologies and other knowledge organization systems. Seppälä, Ruttenberg and Smith (*Guidelines for Writing Definitions in Ontologies*), as well as Maria Campos (*Elaboration of domain models in ontologies: the onomasiological approach and the function of the definition*), discuss a specific but important aspect of design, namely the bringing together concepts and terms and writing of good definitions.

Mendonça (*Bulding ontologies through the OntoForInfoScience Methodology: a detailed approach to perform activities of the ontology development*), and Soares, Farinelli and Elkin (*Ontology building in practice: a case study applied to the obstetric domain*), deal with the process of developing ontologies.

Evaluation of ontologies is the topic of Hicks (*Metrics and methods for comparative ontology evaluation*). This paper makes excellent points, but some of the metrics are more suitable at the assertion level than at the data model and the entity value levels, where the focus of evaluation should be. The literatures on the evaluation of data models and on the evaluation of thesauri, classifications, and similar knowledge organization systems have useful insights to offer.

The remaining papers, except *Ontologies in the context of Information Science: a bibliometric study in Brazil*, deal with various applications of ontologies, by function – harmonizing sensor data (*Semantic Middleware for Industrial Sensors*), indexing (*Automatic indexing and ontologies: identification of convergent contributions in Information Science*) and by domain, for example,

music (*Application and use of music ontologies*)

architecture, engineering, and civil construction (*Ontologies as support for information modeling in architecture, engineering and construction*), and, on a deep philosophical level, organizations and corporations (*An ontological approach to the normative dimension of organizations: an application of Documents Acts Ontology*).

Figueiredo and Almeida report on a bibliographic analysis of the literature on ontologies in Brazil. This is a good starting point, but any such analysis is limited by the source(s) covered and the terms used for searching. The unification of thinking and results from several fields called for in the next paragraph could be supported by extending this work, using a much broader group of sources and a broader approach to the selection of search terms.

It is welcome that this issue combines contributions from *information science* (a label commonly used for the community best represented by organizations such as the Association for Information Science and Technology (ASIST), the International Association for Knowledge Organization (ISKO), and, to some extent, library organization) and *ontology* (a label commonly used for the community best represented by the International Association for Ontology and its Applications (IAOA) and the Ontology Summit). These communities deal with much the same issues – from a different perspective perhaps, drawing on different scientific traditions, using different terminology and different methods – but the same issues nonetheless. As this issue demonstrates, there is growing recognition of this commonality of interests and that joining forces would be better for making progress than working in parallel and separate universes.

Much closer collaboration between OntoBras, ISKO Brasil, and the Brasil chapter of ASIST in organizing conferences and in other applications and closer collaboration between university

faculty from both communities (often across the boundary information science – computer science – would be beneficial for both sides. There are two more communities that should be involved: data modeling and linguistics.

Linguists have developed their own upper level ontologies (not too dissimilar from BFO, for example) for their purpose of explaining language structure and use. Unifying all these contributions may require for data modeling, ontologies, and other knowledge organization systems and effort comparable in scope and insight to what the group known under the name Bourbaki did for mathematics

# Apresentação

Dagobert Soergel

PhD, Universidade de Maryland

A seguir estão algumas observações sobre os tópicos abordados nos trabalhos desta questão, culminando em um apelo a uma colaboração mais forte entre as comunidades que lidam com questões de ontologias e outros sistemas de organização do conhecimento e um esforço para a unificação do pensamento e os resultados alcançados por estas comunidades.

Começo por recordar algumas distinções para colocar os documentos nesta edição em contexto. Essas distinções são pragmáticas e úteis, mas podem ser — e têm sido — debatidas sem parar, às vezes por motivos lógicos. A maioria desses debates é decididamente inútil.

Ao lidar com ontologias e seu uso na representação de dados, existem três níveis:

Os trabalhos nesta edição se concentram em Ontologias de Valor de Entidade. Pode-se ter a impressão de que, para alguns autores, o que faz um sistema digno do nome "ontologia" não depende tanto do conteúdo quanto do nível (no sentido apresentado na tabela), mas sim do grau de formalidade, como sendo expresso em RDF, OWL, ou algum formato de lógica mais expressiva. Esta posição deve ser revisada

O que é importante o conteúdo que um sistema tem para oferecer e com que precisão esse conteúdo é expresso (as coisas podem ser expressas de forma muito precisa em linguagem natural).

Três níveis de representação de dados

Três níveis de representação de dados		
<b>Ontologias de modelo de dados</b>	Defina tipos de entidades e tipos de relacionamento. Os tipos de entidades são classes, mas nem todas as classes são tipos de entidade. Por exemplo, o taxon é um tipo de entidade e também uma classe os membros da classe, ou os valores da entidade do tipo de entidade, são taxa, como Rosa multiflora Rosa multiflora é, por sua vez, uma classe cujos membros são espécimes de plantas individuais. Exemplos: Ontologia de Formato Básico (BFO, incluindo relacionamentos BFO), Ontologia Fundamental Unificada (UFO), FOAF, Dublin Core	Tbox Axiomas terminológicos
<b>Ontologias de valor de entidade</b>	Esses sistemas lidam com valores de entidades - taxa biológica, partes do corpo e sistemas corporais (anatomia mais), genes, funções biológicas, doenças e distúrbios, lugares, períodos históricos, culturas, estilos artísticos, objetos culturais, conceitos - e termos usados para nomeá-los. Tais sistemas também são conhecidos como esquemas de classificação, taxonomias, listas de títulos de assunto ou tesouros, conhecidos coletivamente como sistemas de organização do conhecimento. Os sistemas que utilizam o termo "ontologia" em seu nome geralmente possuem uma estrutura mais elaborada com relacionamentos diferenciados. Exemplos: A Classificação Decimal Universal (UDC), a Ontologia de Gênero, (GO) SnoMed, a Classificação Internacional de Doenças (ICD), as Seções de Assunto Médico (MeSH), AGROVOC, o Tesouro de Termos de Índice Psicológica, os vocabulários de Getty, incluindo o Thesaurus de Arte e Arquitetura (AAT).	
<b>Conjuntos de dados, conjuntos de arquivos, bases de dados</b>	Os conjuntos de dados são aqui interpretados como coleções de afirmações sobre a realidade (um termo usado aqui ingenuamente sem tomar qualquer posição filosófica) ou representando pensamento ou imaginação e sonhos. As afirmações são formadas a partir de tipos de relacionamento preenchendo valores de entidade para os argumentos de relacionamento. Na sua forma mais simples, as afirmações assumem a forma de triplos, mas representar a complexidade da realidade e do pensamento requer relacionamentos multi-argumentos. Os exemplos no artigo de Hick sobre a avaliação da ontologia (Paper I-2) são principalmente afirmativos, indicando que ela considera os conjuntos de dados como ontologias, pelo menos, são apresentados de forma formal.	Abox Axiomas assertivos
<b>Limites entre níveis são porosos</b>	Os limites entre esses níveis não são estritos. Por exemplo, os tipos semânticos do Sistema de linguagem médica unificada (UMLS) listam animais (uma classe) e até mesmo mamíferos e humanos como um tipo de entidade, enquanto que eu animal é um valor do organismo de tipo de entidade, embora alto na hierarquia destes valores da entidade. Muitos tesouros, como MeSH ou o Thaurus de álcool e outros medicamentos, incluem relacionamentos que são asserções sobre a realidade. Por exemplo, na hierarquia MeSH existe um título de agentes anti-infecciosos Listar uma substância química sob este título é uma afirmação sobre a realidade. A MeSH inclui tais relacionamentos porque são úteis para pesquisa. Em uma visão ampla, todos os três níveis poderiam ser combinados em uma grande base de dados sem se preocupar com qual nível pertence um elemento. Por outro lado, a distinção entre os níveis é útil. Os sistemas devem decidir pragmaticamente quais informações incluir em que nível	

Os trabalhos discutem *design* e estrutura, desenvolvimento (estritamente ligado ao design), avaliação e aplicação de ontologias, amplamente interpretadas.

Lima e Maculan (*Comparative study of semantic structures in different knowledge organization systems*) e Maculan e Aganette (*Ontologia aplicada a um modelo de gestão organizacional: contribuições da Ciência da Informação*) discutem a estrutura relacional das ontologias e outros sistemas de organização do conhecimento. Seppälä, Ruttenberg e Smith (*Guidelines for Writing Definitions in Ontologies*), bem como Maria Campos (*Elaboration of domain models in ontologies: the onomasiological approach and the function of the definition*) discutem um aspecto específico, mas importante, do *design*, a saber, reunir conceitos e termos e escrever boas definições.

Mendonça (*Bulding ontologies through the OntoForInfoScience Methodology: a detailed approach to perform activities of the ontology development*) e Soares, Farinelli e Elkin (*Ontology building in practice: a case study applied to the obstetric domain*) lidam com o processo de desenvolvimento de ontologias.

A avaliação de ontologias é o tópico de Hicks (*Metrics and methods for comparative ontology evaluation*). Este artigo faz pontos excelentes, mas algumas métricas são mais adequadas ao nível de afirmação do que no modelo de dados e os níveis de valor da entidade, onde o foco da avaliação deve ser. As bibliografias sobre a avaliação de modelos de dados e a avaliação de tesouros, classificações e sistemas semelhantes de organização de conhecimento têm informações úteis a serem oferecidas.

Os documentos restantes, exceto *Ontologies in the context of Information Science: a bibliometric study in Brazil*, lidam com várias aplicações de ontologias, por função - harmonizando dados de sensores (*Semantic Middleware for Industrial Sensors*), indexação (*Automatic indexing and ontologies: identification of convergent contributions in Information Science*) e por domínio, por exemplo, música (*Application and use of music ontologies*)

arquitetura, engenharia e construção civil (*Ontologies as support for information modeling in architecture, engineering and construction*), e, em nível filosófico profundo, organizações e corporações (*An ontological approach to the normative dimension of organizations: an application of Documents Acts Ontology*).

Figueiredo e Almeida relatam uma análise bibliográfica da literatura sobre ontologias no Brasil. Este é um bom ponto de partida, mas qualquer análise é limitada pela(s) fonte(s) coberta(s) e os termos usados para pesquisa. A unificação do pensamento e os resultados de vários campos exigidos no próximo parágrafo podem ser apoiados ao ampliar esse trabalho, usando um grupo muito mais amplo de fontes e uma abordagem mais ampla para a seleção de termos de pesquisa.

É bem-vinda que esta edição combine contribuições da *ciência da informação* (um rótulo comumente usado para a comunidade melhor representada por organizações como a Associação de Ciência da Informação e Tecnologia (ASIST), a Associação Internacional para a Organização do Conhecimento (ISKO) e, para alguns extensão, organização da biblioteca) e ontologia (um rótulo comumente usado para a comunidade melhor representada pela Associação Internacional para Ontologia e suas Aplicações (IAOA) e a Cúpula de Ontologia). Essas comunidades lidam com os mesmos problemas — de uma perspectiva diferente talvez, com base em diferentes tradições científicas, usando diferentes terminologias e diferentes métodos —, mas os mesmos problemas, no entanto. Como esta edição demonstra, há um crescente reconhecimento dessa comunidade de interesses e que unir forças seria melhor para fazer progressos do que trabalhar em universos paralelos e separados.

Uma colaboração muito mais estreita entre OntoBras, ISKO Brasil e o ASIST Brasil na organização de conferências e em outras aplicações e uma colaboração mais estreita entre professores universitários de ambas as comunidades (muitas vezes na ciência da informação de fronteira —

ciência da computação — seria benéfico para ambos os lados. São duas comunidades mais que devem estar envolvidas: modelagem de dados e linguística.

Os linguistas desenvolveram suas próprias ontologias de nível superior (não muito diferentes da BFO, por exemplo) para o propósito de explicar a estrutura e uso da linguagem. Unificar todas essas contribuições pode exigir modelagem de dados, ontologias e outros sistemas e esforços de organização do conhecimento comparáveis em alcance e visão ao que o grupo conhecido como Bourbaki fez para a matemática

# Presentación

Dagobert Soergel

PhD, Universidad de Maryland

Lo que sigue son algunos comentarios sobre los temas tratados en este documento, que culminan en un llamado a una colaboración más fuerte entre las comunidades que se ocupan de cuestiones de ontologías y otros sistemas de organización del conocimiento y un esfuerzo hacia la unificación del pensamiento y los resultados logrados por estas comunidades.

Comienzo por recordar algunas distinciones para contextualizar los documentos sobre este tema. Estas distinciones son pragmáticas y útiles, pero pueden ser, y han sido, debatidas interminablemente, a veces por motivos lógicos. La mayoría de estos debates los considero decididamente no útiles.

Al tratar con las ontologías y su uso en la representación de datos, existen tres niveles:

Los documentos en este número se centran en Entity Value Ontologies. Uno tiene la impresión de que para algunos autores lo que hace que un sistema digno del nombre "ontología" dependa no tanto del contenido o del nivel (en el sentido presentado en la tabla), sino más bien del grado de formalidad, como expresándose en RDF, OWL o algún formato de lógica más expresiva. Esta postura debe ser revisada

lo importante es el contenido que un sistema tiene para ofrecer y la precisión con que se expresa ese contenido (las cosas se pueden expresar con mucha precisión en el lenguaje natural).

Tres niveles en representación de datos		
<b>Ontologías de modelo de datos</b>	Defina tipos de entidades y tipos de relación. Los tipos de entidades son clases, pero no todas las clases son tipos de entidad. Por ejemplo, el taxón es un tipo de entidad y también una clase los miembros de la clase, o los valores de la entidad del tipo de entidad, son tasa, como Rosa multiflora La rosa multiflora es, a su vez, una clase cuyos miembros son especímenes de plantas individuales. Ejemplos: Ontología de Formato Básico (BFO, incluyendo relaciones BFO), Ontología Fundamental Unificada (UFO), FOAF, Dublin Core	Tbox Axiomas terminológicos
<b>Ontologías de valor de entidad</b>	Estos sistemas se ocupan de valores de entidades - taxa biológica, partes del cuerpo y sistemas corporales (anatomía más), genes, funciones biológicas, enfermedades y trastornos, lugares, períodos históricos, culturas, estilos artísticos, objetos culturales, conceptos - y términos usados para nombrarlos. Estos sistemas también se conocen como esquemas de clasificación, taxonomías, listas de títulos de asunto o tesauros, conocidos colectivamente como sistemas de organización del conocimiento. Los sistemas que utilizan el término "ontología" en su nombre generalmente poseen una estructura más elaborada con relaciones diferenciadas. Los siguientes son los siguientes: La Clasificación Decimal Universal (UDC), la Ontología de Género, (GO) SnoMed, la Clasificación Internacional de Enfermedades (ICD), las Secciones de Asunto Médico (MeSH), AGROVOC, el Tesoro de Términos de Índice Psicológico, los vocabularios de Getty, incluyendo el Thesaurus de Arte y Arquitectura (AAT).	
<b>Conjuntos de datos, conjuntos de archivos, bases de datos</b>	Los conjuntos de datos se interpretan aquí como colecciones de afirmaciones sobre la realidad (un término usado aquí ingenuamente sin tomar ninguna posición filosófica) o representando pensamiento o imaginación y sueños. Las afirmaciones se forman a partir de tipos de relaciones que rellenan valores de entidad para los argumentos de relación. En su forma más simple, las afirmaciones asumen la forma de triples, pero representar la complejidad de la realidad y del pensamiento requiere relaciones multi-argumentos. Los ejemplos en el artículo de Hick sobre la evaluación de la ontología (Paper I-2) son principalmente afirmativos, indicando que considera que los conjuntos de datos como ontologías, al menos, se presentan de forma formal.	Abox Axiomas asertivos
<b>Límites entre niveles son porosos</b>	Los límites entre estos niveles no son estrictos. Por ejemplo, los tipos semánticos del sistema de lenguaje médico unificado (UMLS) enumeran animales (una clase) e incluso mamíferos y humanos como un tipo de entidad, mientras que el animal es un valor del organismo de tipo de entidad, aunque alto en jerarquía de estos valores de la entidad. Muchos tesauros, como MeSH o el Thaurus de alcohol y otros medicamentos, incluyen relaciones que son aserciones sobre la realidad. Por ejemplo, en la jerarquía MeSH existe un título de agentes antiinfecciosos Listar una sustancia química bajo este título es una afirmación sobre la realidad. MeSH incluye tales relaciones porque son útiles para la investigación. En una visión amplia, los tres niveles podrían combinarse en una gran base de datos sin preocuparse de qué nivel pertenece un elemento. Por otro lado, la distinción entre los niveles es útil. Los sistemas deben decidir pragmáticamente qué información incluir en qué nivel	

Los trabajos discuten *design* e estrutura, desenvolvimento (estritamente ligado ao design), avaliação e aplicação de ontologias, amplamente interpretadas.

Lima y Maculan (*Comparative study of semantic structures in different knowledge organization systems*) y Maculan y Aganette (*Ontologia aplicada a um modelo de gestão organizacional: contribuições da Ciência da Informação*) discuten la estructura relacional de ontologias y otros sistemas de organización del conocimiento. Seppälä, Ruttenberg y Smith (*Guidelines for Writing Definitions in Ontologies*), así como Maria Campos (*Elaboration of domain models in ontologies: the onomasiological approach and the function of the definition*) discuten aspecto específico, pero importante, del *design*, a saber, reunir conceptos y términos y escribir buenas definiciones.

Mendonça (*Building ontologies through the OntoForInfoScience Methodology: a detailed approach to perform activities of the ontology development*) y Soares, Farinelli y Elkin (*Ontology building in practice: a case study applied to the obstetric domain*) se ocupan con el proceso de desarrollo de ontologias.

La evaluación de ontologias es el asunto de Hicks (*Metrics and methods for comparative ontology evaluation*). Este artículo levanta excelentes puntos, pero algunas métricas son más adecuadas al nivel de afirmación que en el modelo de datos y los niveles de valor de entidad, donde debe ser el foco de la evaluación. Las bibliografías sobre evaluación de modelos de datos y evaluación de tesauros, clasificaciones y sistemas semejantes de organización del conocimiento tienen informaciones útiles a ofrecer.

Los documentos restantes, excepto *Ontologies in the context of Information Science: a bibliometric study in Brazil*, tratan de varias aplicaciones de ontologias, por función — armonizando datos de sensores (*Semantic Middleware for Industrial Sensors*), indexación (*Automatic indexing and ontologies: identification of convergent contributions in Information Science*) y por dominio, por ejemplo, música (*Application and use of music ontologies*)

arquitectura, ingeniería y construcción civil (*Ontologies as support for information modeling in architecture, engineering and construction*), y, en nivel filosófico profundo, organizações e corporações (*An ontological approach to the normative dimension of organizations: an application of Documents Acts Ontology*).

Figueiredo y Almeida relatan análisis bibliográfico de la literatura sobre ontologias en Brasil. Este es un buen punto de partida, pero cualquier análisis es limitado por la(s) fuente(s) cubierta(s) y los términos utilizados para investigación. La unificación del pensamiento y los resultados de varios campos exigidos en el siguiente párrafo pueden ser apoyados al ampliar el trabajo, usando un grupo mucho más amplio de fuentes y un enfoque más amplio para la selección de términos de investigación.

Es bienvenida que esta edición combine contribuciones de la *ciencia de la información* (rótulo comúnmente usado para la comunidad mejor representada por organizaciones como la Asociación de Ciencia de la Información y Tecnología (ASIST), la Asociación Internacional para la Organización del Conocimiento (ISKO) y, para algunos extensión, organización de la biblioteca) y ontología (rótulo comúnmente usado para la comunidad mejor representada por la Asociación Internacional para Ontología y sus Aplicaciones (IAOA) y la Cúpula de Ontología). Esas comunidades se ocupan con los mismos problemas — de una perspectiva diferente talvez, con base en diferentes tradiciones científicas, usando diferentes terminologias y diferentes métodos —, pero los mismos problemas, entretanto. Como esta edición demuestra, hay creciente reconocimiento de esa comunidad de intereses y que unir fuerzas seria mejor para hacer progresos que trabajar en universos paralelos y separados.

Una colaboración mucho más estrecha entre OntoBras, ISKO Brasil y el ASIST Brasil en la organización de conferencias y en otras aplicaciones una colaboración más estrecha entre profesores universitarios de ambas comunidades



(muchas veces en la ciencia de la información de frontera — ciência de la computación — sería benéfico para ambos lados. Son dos comunidades más que deben estar involucradas: modelaje de datos y lingüística.

Los lingüistas desarrollaron sus propias ontologías de nivel superior (no muy distintas de la BFO, por ejemplo) con el propósito de explicar la estructura y uso del lenguaje. Unificar todas esas contribuciones puede exigir modelaje de datos, ontologías y otros sistemas y esfuerzos de organización del conocimiento comparables en alcance y visión a lo que el grupo conocido como Bourbaki hizo para la matemática.