Principles for the Development of Domain Conceptual Models for Knowledge Organization Systems: An Analysis of Methodologies for Developing Learning Paths in the Field of Corporate Education †

Gercina Ângela de Lima*, Maria Luiza de Almeida Campos**, Patrícia Lopes Ferreira França***

*Universidade Federal de Minas Gerais (UFMG), Programa de Pós-Graduação Gestão & Organização do Conhecimento (PPG-GOC), Av. Antônio Carlos, 6.627, Belo Horizonte, MG, 31.270-901, Brazil, <glima@eci.ufmg.br>

**Universidade Federal Fluminense (UFF), Programa de Pós-Graduação em Ciência da Informação (PPGCI), Rua Tiradentes, 148, Ingá, Niterói, RJ, 24210-510, Brazil, <maria.almeida@pq.cnpq.br>

***Serviço Federal de Processamento de Dados, Av. José Candido da Silveira, 1200, Belo Horizonte, MG, 31035-572, Brazil, <patricialopesvirtual@gmail.com>

Gercina Ângela de Lima is a full professor at the Information Science School of Federal University of Minas Gerais, Brazil. She holds a PhD in information science from UFMG (Brazil) and a master's degree in library science and services from Clark Atlanta University (USA). She is the coordinator of the Research Group MHTX Research Group—Conceptual Modeling for Organization and Representation of Hypertextual Information and has a research scholarship from The Brazilian Research Council (CNPq) since 2010. Her current research interests include organization and representation of knowledge, hypertext, digital library, and web semantic.

Maria Luiza de Almeida Campos holds a PhD in information science and is Researcher and Professor at the postgraduate programs, at Universidade Federal Fluminense and Universidade Federal da Bahia, Brazil. She is the coordinator of the Research Group EOOCI—Ontic and Ontological Studies in Informational Contexts and has a research scholarship from The Brazilian Research Council (CNPq) since 2005. Her areas of interest are knowledge organization, models and theories of knowledge representation, terminology, semantic interoperability, and foundational ontologies.

Patrícia Lopes Ferreira França is a PhD student at Federal University of Minas Gerais where she also holds a master's degree in information science. She worked as Pedagogical Supervisor and Instructional Designer at Senac Minas. She is a member of the MHTX Research Group—Conceptual Modeling for Organization and Representation of Hypertextual Information.

Lima, Gercina Angela de, Maria Luiza de Almeida Campos and Patricia Lopes Ferreira Franca. 2020. "Principles for the Development of Domain Conceptual Models for Knowledge Organization Systems: An Analysis of Methodologies for Developing Learning Paths in the Field of Corporate Education." *Knowledge Organization* 47(7): 592-603. 41 references. DOI:10.5771/0943-7444-2020-7-592.

Abstract: This article presents a set of principles for knowledge modeling in knowledge organization systems in specific domains. It discusses the representational problem, comparing the abstraction mechanisms present in the theories related to representation in concept systems, taken from foundational authors of information science, computer science, and terminology approaches. Parallel to this context, several representational possibilities arise to assist the modeler in the activity of elaborating models of representation. It describes the application of theoretical and methodological principles when organizing, representing, and managing navigation on learning paths in the corporate education field. As a concept proof, it exposes a conceptual model of learning paths and discusses a literature review on this subject to verify to what extent these principles are being applied. It concludes that we can consider the principles discussed in this study as relevant, since they expand the modelers' freedom, not making him hostage to a specific model.

Received: 16 April 2020; Revised: 6 August 2020, 2 September 2020, 9 October 2020, 13 October, 2020; Accepted: 19 October 2020

Keywords: knowledge representation, knowledge organization, learning paths, principles

† The authors with to thank the National Development Council Scientific and Technological (CNPq / Brazil) for the granting of the Research Productivity-PQ.

1.0 Introduction

This study proposes a set of principles for the elaboration of domains' conceptual models, based on the theories of information science, computer science, and terminology. Such principles allow the representation of knowledge in different domains to find a common core of essential principles to model knowledge. In this approach, domain, according to Hjørland (2002; 2004), makes up a discipline, a community, an application field, or even an action within a community.

The conceptual domain models are present in several knowledge organization systems (KOS) and builds a systematic structure. In this sense, this investigation discusses the representational problem, comparing the abstraction mechanisms present in theories related to the representation of concepts' systems, from authors linked to the areas of information science, computer science, and terminology.

In information science, Ranganatha's faceted classification theory (1951; 1967) and the concept theory formulated by Dahlberg (1978a; 1978b) stand out; both allow the representation of knowledge domains. In computer science, authors investigate the representational models associated with the modeling of fundamental ontologies (Guarino 1998a and 1998b; Smith 2004). Based on Terminology, established principles are used to determine concepts and their relations presented by Wuester (1981). With the foundations proposed by these authors, in the three areas of knowledge, this study lists principles that aim to assist the modeler in his activity of elaborating models of representation.

In recent years there were quite a few systematic studies of the theories underlying the areas related to the construction of conceptual models. These models allow the development of documentary languages, computer systems, hypertexts, and systems aimed at building knowledge bases also called specialized systems and ontologies (Campos 2001; Lima 2004; Lopes 2018). Thus, it should add several representational possibilities to think about a given reality, from a theoretical-methodological stance that allows the modeler to overcome specific models of representation and reflect on the principles underlying the modeling process.

In this study, we understand modeling from the basis developed by Le Moigne (1977) with the theory of the general system also called modeling theory. Le Moigne affirms that to know is to model; the knowledge process is equivalent to the construction of the world/domain models that will be built, making it possible to describe and provide explanations about the observed phenomena. The author, when proposing the development of a methodological stance, asks the researcher not to think about the diversity of models but mainly about the principles that make modeling possible.

Thus, the information professional must also be aware of the principles adopted when determining the modus operandi for the elaboration of a KOS. This is because, according to Le Moigne (1977), in any modeling theory used in the elaboration of representation models, the model, product of this representation, cannot be a definitive and immutable truth but must result from a given process, of a particular construction, or of a worldview linked to space and time. Therefore, knowledge is the action of its building, a process in which the principles adopted for this construction must be clear.

Therefore, aiming at the validation of the established principles, namely the reasoning method used to analyze the domain, the representation object, the relationships between objects for the representation of conceptual structures, and the forms of graphic representation, these principles were applied in corporate education. In this domain, the focus was on the organization and representation of learning paths based on two procedures: 1) the first investigated, in this subject literature, to what extent such principles were usable; and, 2) the second was the development of a conceptual model of a learning path for corporate education as a concept proof.

We can understand learning paths as a systematic and multimodal set of learning units, containing different navigation schemes. These schemes can vary from linear and prescriptive models, passing through more hierarchical models and reaching network models in which navigation is freer (Lopes and Lima 2019). In corporate education, these paths tend to contribute to better use of the resources invested in lifelong learning, besides making it possible to overcome the teaching approach called "one size fits all" (Adesina and Molloy 2011; Freitas and Brandão 2006; Lubchak et al. 2012; Suazo et al. 2012; Yang 2012).

This article consists of four sections, including this introductory part. In Section 2, with the title "Knowledge representation and principles for domain modeling," the principles regarding domain modeling is discussed based on foundational authors taken from the concept of knowledge representation. In Section 3, entitled "The representation of learning paths in the context of corporate education," we present the definition of learning paths and to what extent the literature on procedures for the development of learning paths uses the principles of modeling explored. In addition, as a concept proof, the third section applies these principles in the development of a conceptual model. Finally, in Section 4, we present our conclusions based on the results achieved throughout the study.

2.0 Knowledge representation and principles for domain modeling

Several scientific fields define the term knowledge representation differently. A well-known and updated definition by Sowa (2000), conceived under a theoretical approach (philosophical and logical), considers knowledge representation as the application of logic and ontology in the tasks of building computer models for some domains. Artificial intelligence research discusses knowledge modeling or knowledge engineering methods (Studer et al. 1998). We obtain another point of view through linguistic studies and their elucidation in knowledge representation (Löbner 2002). In cognitive science, knowledge representation is based on the individual's world model, the context and the way one sees the world individually, comprising the sum of different knowledge structures (Gardner 1996; Mey 1982). From the perspective of information science, the authors Cleveland and Cleveland (2001), Lancaster (2003), and Stock and Stock (2013) treat knowledge representation as an approach to solving problems, such as structuring, storing information, and how to find and recover them accurately and effectively. It is important to note that among these definitions presented in several domains of science, Sowa's (2000) view is restricted to the analysis proposed in this article.

Knowledge representation is a recurring theme in several sciences, looking into how the processing of knowledge occurs in the human mind and how to materialize this knowledge. The symbolic form of this representation is something that concerns the scientific scope of documentation since its origin. Such representation associates with ways of expressing information, as highlighted by Vickery (1978). The problem is relevant and extends to many other situations besides documents and indexes. It is necessary to decide how to represent knowledge so that these representations can be manipulated (Vickery 1986, 145) in the structure of software data, the syntactic and semantic structures of natural language, the representation of knowledge in artificial intelligence, and the models of human memory.

In this sense, San Segundo (2004, 109) highlights that if knowledge is an integration process, knowledge representation will be "concepts, theories, models, formats, descriptions and structures that have a meaning of the information symbolization and, more recently, electronic information." Knowledge organization is a necessary process, which becomes more urgent as the volume of information increases. The major aim of this process is to organize the knowledge of a domain and make it available for later retrieval. The organization and retrieval of information face the same challenge to meet the needs of potential users.

In Lima's view (2020), to organize knowledge in any area, from representation to recovery, we first need to study the concepts that make up this knowledge field and the relationships between them. According to Lima (2020), we can modify the concept based on prior knowledge on the topic after a mental elaboration, transforming from the information unit to a conceptual unit for communication purposes. Thus, knowledge representation is a process that uses terminological instruments as products. So, it is necessary to establish guiding principles in the construction of these systems as described in the following section.

2.1 Principles for modeling knowledge domains in knowledge organization systems

Conceptual modeling consists of the creation of models of representation/description from a worldview (conceptualization) that exists over a domain of knowledge. This practice focuses on "identifying, analyzing, and describing the essential concepts and restrictions of a domain" (Guizzardi et al. 2002, 65). For Campos (2004), the modeling process thus requires the displacement of the "phenomenal world" to a space where knowledge representation mechanisms allow formalization processes of objects and their relationships in pre-defined representation contexts.

In order to provide the modeler or classificationist with a range of representational tools for modeling domains and, based on a systematic investigation of methods and theories, we arrived at the determination of five fundamental principles for the act of modeling domains of knowledge, which are present in studies developed within the scope of representing learning paths in the corporate education context. The first principle concerns the "domain of knowledge to be modeled." The second concerns the "reasoning method" used for the knowledge organization within a domain. The third one analyzes the definition of "the object of representation," identifying the knowledge unit to be represented. The fourth concerns the "relationships between objects," aiming to verify the possibilities of semantic connection/separation between the concepts of a given domain. And the fifth shows the "forms of graphic representation" that a model can adopt. In this section, we will present each of these principles, establishing a dialogue between the fields of information science, computer science, and terminology.

2.1.1 Knowledge domain to be modeled

The first principle to consider when developing more appropriate methodologies for conceptual models should be the determination of the knowledge domain. Therefore, it is important to check if the subject covers disciplines that have a paradigmatic or syntagmatic cut in the meeting of knowledge units. It is necessary to verify if there is evidence of the matter in an already established canonical area or if it is part of an area that gathers knowledge from different canonical areas.

When you insert the domain of the treated subject in a paradigmatic cut, it presents the units of knowledge, in most cases, hierarchically. If it is the syntagmatic type domain (having over one subject area of a canonical cut), its units of knowledge will form several hierarchical trees that will come together from generic planning. It is important to emphasize that the development of a conceptual model is complex, which requires knowledge about the subject and about the community that will benefit from the modeling. Thus, it is important that methodology developers for conceptual models establish a stage that can provide discussions about the type of subject thematic section treated in the conceptual model that is being created and the target audience of the intended mapped domain.

2.1.2 Reasoning method

The second principle concerns the adoption of a reasoning method used for the knowledge organization within a domain. Every knowledge organization, according to Morin (2000, 24), follows principles and rules in which organization can include connection operations (conjunction, inclusion, and implication) and separation (differentiation, opposition, selection, and exclusion). The process is circular and cyclical, carried out in stages, moving from separation to connection, from analysis to synthesis, and from synthesis to analysis. Therefore, knowledge involves simultaneously, separation and connection, analysis and synthesis.

Although there is an understanding that the organization of knowledge must occur from a systemic perspective, it is possible to identify the development of distinct organizational models. Traditionally, in computer science, models were developed using two different methods: 1) the inductive method, also called bottom-up; and, 2) the deductive method, also called top-down.

The inductive method allows the elaboration of models that begin from the attributes` context. Therefore, it starts by representing the elements/objects and the relationships of a context, as to obtain an organized structure of concepts, which is grouped according to the classes they belong to, according to their attributes. The deductive method, on the other hand, proposes the elaboration of abstraction mechanisms in order to first think about the domain/context regardless of the consideration of elements and their relations (this would be a later step). In this sense, a process of knowledge logical division follows, starting from general categories, containing generic concepts to more specific concepts. In effect, it generates a set of hierarchical classes in which each subclass of elements is a set of the immediately preceding class. We can associate these methods with theories defended in different knowledge areas.

2.1.3 Representation object

The third principle focuses on the represented object. In the context of the analyzed theories (general terminology theory, formal ontology theory, faceted classification theory, and concept theory), the object is defined as the smallest unit of manipulation/representation in each context. In each of these theories and in the methodologies used, it is possible to identify what can be considered an object of representation and the importance of that object as the smallest representation unit in addition to the implications that result from this definition.

Regarding terminology theory, Wuester defines this object of representation as a concept, which is a mental unit. However, in concept theory, the definition of concept is a knowledge unit. Dahlberg in her theory disagrees that the concept can be a mental unit, because such understanding can vary from person to person; according to the mind of each person, it is necessary to have an agreement between people to allow communication. Hence, it defines concept as a unit of knowledge, where an ontological commitment is established between the members of a given community.

Through the analysis of the faceted classification theory and the concept theory, it is clear that these theories support one of the first methodological efforts, in information science, to highlight the representation unit issue. Ranganathan, in his Prolegomena, introduces the concept of "isolated" as the classification system's minimum and manipulative unit. Dahlberg (2014) presents the "concept" as a minimum unit and defines it as a triad, composed of the elements referent, characteristics, and name. In this triad, the object is the referent, classified as an individual or general object that, when circumscribed to a context, requires the appropriation of characteristics, being designated a linguistic sign—a name.

In the context of formal ontology, in computer science, Guarino (1998a, 1998b) proposes that objects or "particulars" be classified as concrete and abstract. Comparatively, in concept theory, these are considered general objects, as they represent a class of objects, not a particular entity in the world. Concrete objects, within the scope of formal ontology, are classified as continuous and occurring. These discussions aim to highlight the multiplicity of interpretations and end up allowing flexibility for observation and work in different domains. In addition, they make the modeler feel the need to build the assumptions where to start from or even try to understand where he starts from in order to develop a representation model.

2.1.4 Relations between objects

The fourth principle has the purpose of verifying the possibilities of semantic connection/separation between the concepts of a domain. The relationships between objects within a context form the conceptual structure of that context and are of different natures. The analysis is presented in the movements of the act of modeling. These movements reflect groups of relationships between concepts, as described, in an exemplary manner, below.

The first movement is the verification, in the presented theories and methods, of the existence of categorical relations. This type of relationship brings together, in a first major grouping, objects by their nature, that is, entities, processes, among others. This relationship often makes it possible to reduce logical errors in establishing links between concepts, as it determines the nature of the object.

Having verified the reunion or not of objects by nature, the second movement, not in order of precedence, but of necessity, is to verify how objects of the same nature are related. In this form of relationship, there is a determined hierarchical relation. Consequently, theories and methods present relations that determine "what is the object," the understanding of the concept. Accordingly, the theory of the concept and the theory of classification call this relationship hierarchical; the general theory of terminology calls it a logical relation, and formal ontology defines this relation as "one (ISA)," considered a property in a minimal ontology of universals for the structuring of a domain.

Another movement that implies the relationships between the concepts is the analysis of how "the object is constituted," that is, its parts and elements. In this form of relationship, partitive relations are determined. Like the other relationships, this one is named differently in the theories and methods analyzed: 1) the concept theory is called partitive relationship; 2) the classification theory places hierarchical and partitive relations in the same group, calling them hierarchical relations; 3) in terminology theory, it is called the ontological relationship of partitive coordination; and, 4) in formal ontology, partitive relations are named all-part theory, also called "mereotopology." Formal ontology presents more explicit criteria for determining the type of partition. These criteria help the modeler to distinguish the nature of what will be considered as part and as a whole and function as "informational meta-categories" to reflect on the relations of a domain.

The next movement is to see how objects of different natures relate to each other and to represent that relationship more consistently. That is, in this process, some prescriptive criteria are determined that allow for more judicious connections. The relationship between concepts of different natures can be observed in the concept theory as a syntagmatic functional relationship. These relationships, unlike paradigmatic ones (logical and partitive), can be recognized as relationships that make evident a specific demand or function between objects in the phenomenal world and do not aim to explain the object and its properties.

Finally, another movement to address no longer has to do with comparing or the relation between concepts but between how to express these concepts. In other words, this movement happens in the language sphere, the so-called equivalence relation. Once again, a space for analyzing these relationships builds on the different theories and methods that encompass the thematic scope of domain modeling.

2.1.5 Forms of representation

This principle contemplates the forms of graphic representation that a model can adopt. A conceptual model must also be a communicational space where the phenomenal world transposes into a space of representation. Therefore, it is important to investigate graphic forms of representation, because sometimes these forms impair representational possibilities. In the theories and methodologies analyzed, it appears that the area of information science, despite having well-grounded theories about the concept and conceptual relationships, is scarce in terms of models that assist in the development of graphic representations. In the terminology context construction practices, where there is a configuration of general theory of terminology and ontology, graphic representations are consistently present, although there are some fundamental differences.

Even if graphic representations are not addressed, it is often, in some areas of knowledge or professional fields, important to mention that they lack, much more than any other point analyzed, a space for investigation. It associates the relevance of studying them with their contribution by making the twinned work processes in the representation of knowledge more transparent. In corporate education, for

Foundational theories of Information Science				
Principles	**Classification Theory**	**Concept Theory**	**Formal Ontology**	**Terminology Theory**
1. Reasoning Method	Synthetic Analytical Method	Synthetic Analytical Method	Inductive Method	Inductive Method
2. Object/Unit of Representation	Isolated	Concept (unit of knowledge)	Continuous object sand occurring objects	Concept (unity of thought)
3. Relations between objects	-	-	-	-
3.1 Categorical Relationship	Category	Formal Categorical Relationship	Universal Theory	Does not exist
3.2 Hierarchical/Logical Relationship	Hierarchical Relationship	Hierarchical Relationship (Abstraction and Lateral Relationship)	ISA (is a)	Logical Relationship (Comparison RelationshipandCombination Relationship)
3.3 Partitive Relationship	Hierarchical Relationship	Partitive Relationship	AllPart Theory	Ontic Relationship of Partitive Coordination
3.4 Relationship between objects of different natures	Does not exist	Syntagmatic Functional Relationship	Dependency Theory	Ontic Linking, Ontic Causality Relationship, Genealogical Relationship
3.5 Equivalence Relationship	Verbal Plan	Equivalence Relationship	Does not formally specify	Equivalence Relationship
4. Forms of Graphic Representation	Has no principles for graphic representation	Has no principles for graphic representation	Has principles for graphic representation	Has principles for graphic representation

Table 1. Summary Table of principles for domain modeling.

example, these representations may contribute to improvements in the management of learning paths.

Next, we will discuss the representation of learning paths in a corporate education context, where we consider the theories discussed in this section. In this sense, as a conclusion of the principles exposed here, we present the summary table of the principles for domain modeling (Table 1). In the summary table, we will highlight the theories that we discussed for the reasoning method, for the identification of the unit of representation, and for the relationships between objects.

3.0 Representation of learning paths in the context of corporate education

According to Lopes (2018), learning paths can be understood as a systematic and multimodal set of learning units, containing different navigation schemes. These schemes can range from linear and prescriptive models to more hierarchical models and to network models, where navigation is freer and with the purpose of developing skills. The navigation schemes in learning paths can be customized, based on variables such as objectives, student profile, and learning characteristics.

In the education field, a learning path is fundamental to the teaching-learning process, since it integrates a set of activities in an appropriate sequence, allowing the student to learn the contents more effectively. A learning path, according to Yang (2012), comprises different activities aimed at learning, and these activities can relate to different teaching approaches, seeking whenever is possible a learning path that is closer to the performance and preferences of the student.

From the librarianship and information science's view, we can analyze the learning paths and hypertext systems similarly, as highlighted in the study by Lopes (2018). To a certain extent, in Campos' (2001) view, hypertexts can be compared to a system of concepts, as they are conceptual nodes linked to other nodes through a level of relationship. Campos (2001) points out that, coupled with "being hypertextual," there is a classificatory action since the connections between the nodes implement from a network of associations.

The existence of a route in a learning path, whether previously shown or created by the learner himself, is close to the hypertextual relationships that can manifest within the same text, the so-called microtexts, or between different texts, called macrotexts. These two classifications are terms proposed by Rada (1991) in hypertexts studies.

In corporate education, the educational solutions organization learning paths is a reality in many institutions since this type of strategy can contribute to the processes of planning, execution, and evaluation of training actions as explored in the researches advocated by Brandão (2012), Carbone (2013), Freitas and Brandão (2006), and Murashima (2011). The results revealed, in general, that learning paths can be organized by areas, positions, competences, and work processes, among other categories.

In this section, based on the five fundamental principles for the act of modeling knowledge domains, discussed in Section 2, it is sought, based on studies from the fields of information science, computer science, and terminology, to verify in literature the extent to which these principles are applied. Furthermore, as a concept proof, a conceptual model of learning paths is elaborated, applying the determined principles.

3.1 Application of the principles of domain modeling in the analysis of learning paths construction methodologies

In this section, we seek to present the methodological procedures adopted in the comparative analysis carried out between the domain modeling principles mentioned in Section 2 and the strategies for building learning paths presented in the scientific literature in the area of corporate education. The research strategy adopted was exploratory and qualitative. To achieve the proposed objective, three stages of data collection and analysis were established, considering the content analysis method proposed by Bardin (2011), namely: 1) pre-analysis; 2) exploration of the material; and, 3) treatment, inference, and interpretation of results.

First, in the pre-analysis stage, we carried out an exploratory reading of the literature review publications (Lopes and Lima 2019). Subsequently, we selected forty-three publications as a corpus of analysis considering as a criterion the studies whose authors explained the stages of the conceptual modeling process. We assume that in the publications on organization and representation of learning paths, the authors possibly did not maintain adherence to the principles of the conceptual models of hyperdocuments proposed in Section 2 of this article. In the stage referring to the exploration of the material, the KOS knowledge domain modeling principles addressed in this study are categories of analysis.

Although Bardin (2011) shows an inductive logic for establishing the categories of analysis, starting from the coding from the exploratory reading, the deductive logic was chosen, since the principles presented in Section 2 were representative and convergent to the objective established for this stage of the research. Thus, for the organization and coding of the data, we structured a matrix, contemplating the referred principles, and the publications selected analysis corpus. For each publication, there was a new reading, and each principle's presence or absence inserted in the matrix. For each principle, there were subcategories of analysis related to each of the principles associated with the modeling of knowledge being raised.

Finally, in the stage of treatment, there was an inference and interpretation of results, a comparative analysis of publications and each principle was identified. It should be noted that, although they have different characteristics, each step was performed in an integrated manner in view of enabling the complementation of the results from the literature, which are presented below.

Through the analysis of the "domain of knowledge to be modeled" principle, we consider three subcategories of analysis, namely: 1) if there was a paradigmatic cut; 2) if there was a syntagmatic cut; or, 3) if it was not possible to identify the type of thematic cut adopted. In 39.5% of the publications, the authors explain a paradigmatic outline, focusing mainly on disciplines of academic curricula. To a lesser extent, in only 13.95% of publications analyzed, the authors predicted syntagmatic cuts. As explained by Campos (2001), when a paradigmatic cut is identified, the knowledge units are rigorously hierarchized. On the other hand, from the syntagmatic cut, the knowledge units form different hierarchical trees, brought together globally. In this sense, we observed that the syntagmatic cut meets the interdisciplinary and polychotomous nature of the themes identified in an organizational context, pointing to the formation of complex networks of knowledge units. Finally, it was evident that most publications in the analysis corpus (46.5%) did not foresee or did not explain how the thematic cut of the content was done. There is an assumption that this result may have been influenced by the different focuses of publications, which do not necessarily explain the adopted methodology.

Through the analysis of the "reasoning method" principle, we envisaged the following subcategories of analysis: 1) application of the deductive method; 2) use of the inductive method; 3) use of the analytical-synthetic method; and, 4) no possibility of identifying the reasoning method. In 41.86% of the publications in the analysis corpus, the authors explained the application of only the deductive reasoning method. Publications that address the inductive method start from lists of disciplines, themes, learning objectives, courses, and skills to be developed. Specialists determine these lists previously, and generally, they are linear or hierarchical. There is also a need to highlight the prevalence of prior validation of conceptual schemes by specialists in the thematic domain, characteristic of deductive approaches. Regarding publications that explained only the inductive method, in 37.2% of the publications in the analysis corpus, the authors are directed to this focus. Regarding the analytical-synthetic reasoning method, only 13.95% of publications can be classified as such. Finally, in only 6.97% of publications, it was not possible to identify the reasoning method adopted.

In the analysis of the "definition of the representation object" principle's view, publications that mentioned the requirements of: 1) granularity; 2) descriptive representation,; 3) unit of knowledge be self-contained; and, 4) publications that did not mention this principle. In the scope of this study, the concept of granularity is the smallest unit of information. The subcategory "descriptive representation" refers to so-called linguistic labels or linguistic signs used to represent units of knowledge. Regarding the quality of a self-contained knowledge unit, it relates to the importance of defining the scope of content treated in the so-called conceptual node, seeking to contemplate what the concept is and its constituent elements. In the English language, there is a specific term to refer to this quality, which would be the term self-contained. The data revealed that 27.90% of publications mention granularity, while 39.53% address descriptive representation. While in only 13.95% of publications, the authors refer to questions of granularity and descriptive representation together. Only one identified publication referred to the importance of the knowledge unit being self-contained. Finally, 39.53% did not explicitly indicate this principle. Based on the analysis of publications that did not mention the nature of the knowledge units' content, it was possible to verify that, generally, the proposals for modeling learning paths are restricted to the representation of categories and subcategories with a more global and not so specific focus that could contemplate the nature of the content. Regarding the granularity item, we observed that there is no standard as to the determination of which is the smallest node on the learning paths. Regarding the descriptive representation, there was also a lack of clarity regarding the standards adoptable for the different documentary typologies. Finally, in relation to the item "to be self-contained," in only one publication was explicit reference made that the knowledge unit is self-contained, which could favor reuse and interoperability.

In view of the analysis of the "relationships between objects" principle, the types of relationships provided as subcategories of analysis, namely: categorical, hierarchical, partitive, between categories, and equivalence. It was also considered as a subcategory of "publications whose principle was not addressed by the authors." A hierarchical relationship was identified in 51.16% of the publications, while in 83.72% there is a relationship between different categories, considering that in 46.51% of the publications the authors mention two or more types of relationships. The categorical and the equivalence relation had only one mention each. The partitive relationship was never mentioned. In only three of the publications, the authors did not explicitly describe the establishment of relationships between knowledge units. Among the publications whose authors explain relationships, these can be established a priori, by a specialist in the field, and by the student himself, according to learning objectives and preferences. Relationships can also establish automatically, considering several variables such as prior knowledge, learning style, performance, student profile, and learning objectives, among others. We emphasize the identification of "ontological chaining" relationship in which the learning path's construction considers the predecessor and successor units of knowledge. There is also an emphasis on the application of data mining techniques for the automatic generation of learning paths.

From the analysis of the principle "forms of representation," it was possible to identify the following analysis subcategories: concept map, flowcharts, business process management notation (BPMN), mind map, website or system, and publications that did not address this principle. The representations in the flowchart formats were identified in two publications; there was only one mention of the BPMN notation and mind map, each. In 11.62% of the publications, the authors covered both websites and systems and concept maps, each. We also identified that in 69.76% of the publications, the authors do not clearly mention the graphical representation of the nodes and relationships from the point of view of the information visualization by the user. However, in 46.51% of the publications, the authors explicitly address some form of representation of the conceptual model adopted for the learning paths. Of these representations, the ontologies, taxonomies, and architectures of the systems were used to implement the learning paths.

Through comparative analysis, the results revealed, in relation to the "domain of knowledge to be modeled" principle, that there is adherence to this principle in the methodologies for building the learning paths identified in the consulted literature. However, there was partial adherence, since almost half of the publications (46.5%) do not mention this principle. It is estimated that the focus on the paradigmatic outline does not portray the polytomous nature of the learning paths, which can be composed of knowledge units from different domains, not being restricted to a canonical area.

Regarding the "reasoning method" principle, considering the recommendation to adopt the synthetic-analytical method in order to privilege a systemic perspective for the organization of knowledge units, it was observed that there is low adherence to this principle in the identified methodologies. This finding is justified, since in only 13.95% of the publications the authors made explicit the adoption of an analytical-synthetic approach.

Regarding the "definition of the object of representation" principle, 60.47% of publications were identified. However, the data revealed that the lack of a standard for determining granularity and documentary typology may compromise the possibility of reusing the units of knowledge provided for in a learning path.

In relation to the principle "relations between objects," adherence was identified, considering that 93.03% of publications mention some sort of relationship between knowledge units. This result was verified, even considering that it was not possible to identify among the publications any whose authors addressed a clear orientation regarding the description and possibility of applying each type of relationship. Consequently, it is a research opportunity to provide guidance on these relationships' types and to apply them in a learning paths context. Regarding the principle "forms of representation," there was adherence, since 79.05% of publications explicitly address some graphic representation of the conceptual model or some form of learning paths visualization, however, only from the users' point of view.

In general, the principles for modeling knowledge domains presented in Section 2 are applicable to the context of creating methodologies for the organization and representation of learning paths. In this specific context, the inclusion of a principle focused on data storage and management systems is an expandable opportunity, considering the current society's scenario, sustained by a large volume and diversity of information. The application of these principles, as a concept proof, in the formulation of a conceptual model is presented below. Such a model can be understood as the manifestation of a graphic representation for a better visualization of the knowledge representation processes.

3.2 Proposal of a conceptual model for organization and representation of learning paths

The conceptual modeling stage is essential in the knowledge representation process. However, in the context of creating hyperdocuments, it is not a stage always contemplated. The proposal for a conceptual model aims to assist the authorship of hyperdocuments in order to establish a communicational space whose content author and other professionals involved in the process of creating the hyperdocument can interact and jointly represent the conceptual model (Campos 2001). In this sense, a model is proposed for the organization and representation of the learning paths, especially for the area of corporate education. This model allows exemplification of the modeling principles presented in this study, and that is applicable in other contexts.

Regarding the "domain of knowledge to be modeled" principle, there are two structure possibilities, both under a paradigmatic cut and under a syntagmatic cut. In the first possibility, the knowledge units are part of the same canonical area. For example, in learning paths on librarianship and information science, the tendency is for knowledge units to group in this knowledge field or domain. In the second possibility, knowledge units have more than one subject area. For example, a learning path composed of knowledge units from a business process and units of knowledge from a canonical area.

About the "reasoning method," in the organization of knowledge units that will compose learning paths, it is possible to adopt deductive, inductive reasoning methods, or even the analytical-synthetic method. When using the deductive method, the general categories that will comprise the units of knowledge must be pre-defined. These general categories should be exhaustive, with the capacity to accommodate all units of knowledge that intend to be covered within the learning paths coverage. In the inductive reasoning method, it starts from the analysis of knowledge units and their respective attributes, looking to identify similarities that may indicate categories of subjects that allow the grouping of these knowledge units. Finally, but not least relevant, in the analytical-synthetic method, constant movements of analysis and synthesis are carried out, using the deductive and inductive methods in an integrated manner. The analytical-synthetic approach is the most appropriate for the context of the learning paths organization.

Concerning the "definition of the object of representation," in theory, this object in the context of learning paths is the competence intended to develop. This competence represents a concept. Therefore, the smallest unit of manipulation/representation in the context of learning paths is the concept. Based on the premise that learning is a process, the representation object can be classified as an abstract object, occurring, or a process within the formal ontology's scope or else an action in the concept and terminology theories contexts.

Regarding "relationships between objects," starting from the movements proposed in Section 2, it was possible, through the proposition of a conceptual model of organization and representation of learning paths, to identify categorical relationships. These categories are like those proposed by Ranganathan in the theory of faceted classification or even those suggested by Dahlberg in the theory of concept.

In order to identify the way in which knowledge units relate to each other on learning paths, we found that the hierarchical relationship is essential, and it is present at different times in the organizing process of learning paths. In this context, we seek to identify the nature of the knowledge unit, providing for analyzes on how these educational solutions are hierarchized, in which types of categories these solutions are classified, among other possible analyzes. In a subsequent movement, the aim is to identify which parts may compose the same educational solution. A learning path is formed by different educational solutions, composed of learning objects, learning activities, and assessment activ-

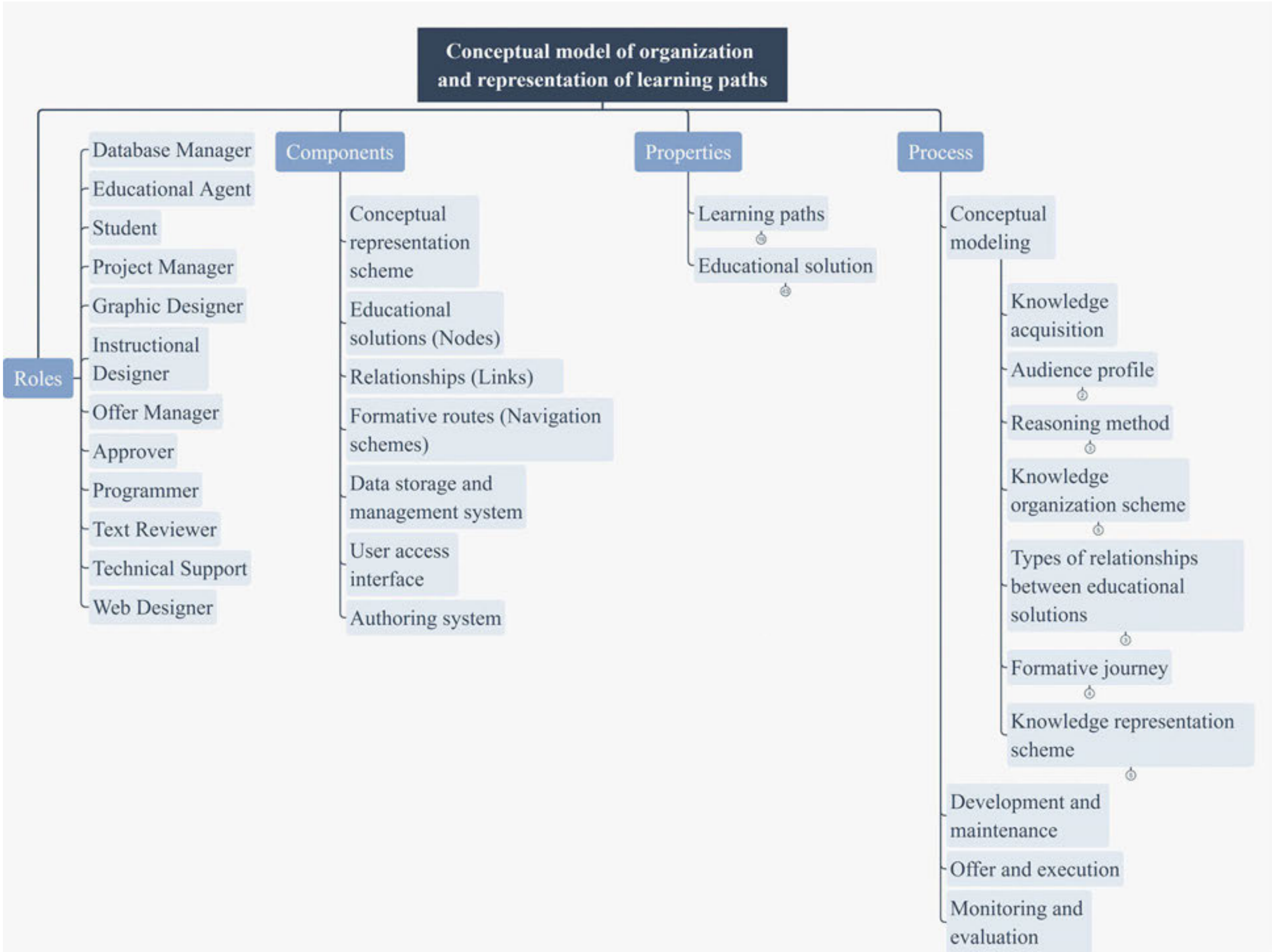

Figure 1. Conceptual model of organization and representation of learning paths. Adapted from Lopes (2018).

ities. A learning object can form from different units of knowledge, represented by videos, texts, and podcasts, among other forms of representation. In a complementary way, in the fourth movement, we seek to verify how objects of different natures relate and to represent this relationship in a more consistent way. In this sense, we identify the possibility of relationships between these objects, for example, in the relationship between an educational agent (tutor, teacher, mentor, etc.) and a given educational solution (course, video, seminar, tutorial, etc.). In the fifth and last movement, to express the different units of knowledge that may compose a learning path, there is a use of linguistic signs that enable equivalence relations between different units of knowledge.

Finally, the "forms of representation," identifies that the learning paths may be represented, graphically, by a simple or hierarchical list of knowledge units, mind maps, concept maps, specific notations, or graphs, among others. As in information science, graphic representations of learning paths also demand greater space for investigation. With this, it becomes possible to identify forms of representation of the knowledge units that mainly contain the diversity and the volume of possible units of knowledge to compose learning paths.

In Figure 1, the conceptual model is presented, as concept proof, for organization and representation of learning paths, structured based on the modeling principles of knowledge domains described in Section 2.

The model presented in Figure 1 is organized into four major categories: actors, components, properties, and processes, showing, in this first level, the relationship of the categorial type. In each of these major categories, same nature concepts are grouped. In the first category (actors), the grouped concepts have a hierarchical coordination relationship. In the second and third categories (components and properties, respectively), the partitive relationship is identified, considering that the components and properties are part of learning paths. In the fourth category (processes), there is a relationship of the hierarchical type of coordination. In the conceptual model, we opted to hide part of the predicted subcategories, considering the focus of this work, which is on the knowledge domains' modeling and also due to the space limitation.

4.0 Final considerations

This study addressed a set of principles for the development of domains' conceptual models and the verification of the application of these principles in a set of methodologies for building learning paths. We identified the methodologies in a literature review process, and, as a proof of concept, we proposed another conceptual model of learning paths based on the identified principles. It was observed that the principles for modeling knowledge domains addressed are partially applied in the methodologies for building learning paths identified in the consulted literature. Starting from the assumption that each proposed principle bases itself on fundamental theories in the areas of information science, computer science, and terminology and also considering that such principles have yet to be fully applied in the modeling processes of knowledge domains, as we presented in this study, one of our findings is that there is an opportunity for greater dissemination of these principles among professionals working in the modeling of knowledge domains.

Based on guiding principles, the professional who works in the modeling of knowledge domains has the possibility of overcoming pre-established models, being able to understand the metaconceptions involved in the modeling process, used as abstraction mechanisms. This observation is aligned with the reflection proposed by Ranganathan, still in the twentieth century, on the importance of having more classificacionists, those who elaborate classifications, and not just classifiers, those who use classifications.

Considering that the application of the principles of knowledge domains modeling addressed enabled the systematization of a conceptual model, the application of these principles in the field of corporate education, especially in the construction of learning paths, was of fundamental importance. Certainly, these principles contribute to the quality of the results achieved in addition to the possibility of proposing a new model for building learning paths. Therefore, it is possible to assume that the principles discussed in this study can be considered relevant as they enable the modeler's freedom, not leaving him hostage to a specific model. Thus, the application of the proposed principles in other fields of knowledge is expected, aiming the creation of new proposals for meta-representations that can contribute to the processes of knowledge organization and representation.

References

Adesina, Ayodeji and Derek Molloy. 2011. "A Business Process Management Based Virtual Learning Environment: Customised Learning Paths." In *CSEDU*: *Proceedings of the Third International Conference on Computer Supported Education 6-8 May 2011 Noordwijkerhout, NL, CSEDU 2011,* ed. Alexander Verbraeck, Markus Helfert, José Cordeiro and Boris Shishkov. Setúbal: Scitepress, 365-68.

Bardin, Laurence. 2011. *Análise de Conteúdo*. Lisboa: Edições.

Brandão, Hugo Pena. 2012. *Mapeamento de competências: Métodos, técnicas e aplicações em gestão de pessoas*. São Paulo: Atlas.

Campos, Maria Luiza de Almeida. 2001. *A Organização de Unidades de Conhecimento em Hiperdocumentos*. PhD diss., Instituto Brasileiro de Informação, Ciência e Tecnologia.

Campos, Maria Luiza de Almeida. 2004. "Modelização de Domínios de Conhecimento: Uma Investigação de Princípios Fundamentais." *Ciência da Informação* 33, no. 1: 22-32.

Carbone, Pedro Paulo. 2013. "Gestão por Competência e Educação Corporativa: Caminhos para o Desenvolvimento de Competências." *Inclusão Social* 7, no. 1: 44-55.

Cleveland, Donald and Ana Cleveland. 2001. *Introduction to Indexing and Abstracting*. 3rd ed. Englewood, CO: Libraries Unlimited.

Dahlberg, Ingetraut. 1978a. *Ontical Structures and Universal Classification*. Bangalore: Sarada Ranganthan Endowment.

Dahlberg, Ingetraut. 1978b. "A Referent-Oriented, Analytical Concept Theory for INTERCONCEPT." *International Classification* 5: 142-50.

Dahlberg, Ingetraut. 2014. *Wissensorganisation*: *Entwicklung, Aufgabe, Anwendung, Zukunft*. Würzburg: Ergon.

Freitas, Isa Aparecida and Hugo Pena Brandão. 2006. "Trilhas de Aprendizagem como Estratégia de TD&E." In *Treinamento, desenvolvimento e educação em organizações e trabalho: Fundamentos para a gestão de pessoas*, ed. Jairo Borges-Andrade, Gardênia da Silva Abbad and Luciana Mourão, 97-113. Porto Alegre: Artmed.

Gardner, Howard. 1996. *A nova ciência da mente: Uma história da revolução cognitiva*. São Paulo: EDUSP.

Guarino, Nicola. 1998a. "Formal Ontology and Information Systems." In *Formal Ontology in Information Systems: Proceedings of the First International Conference (FOIS'98) 6-8 June 1998 Trento, IT*, ed. Nicola Guarino. Padova: IOS Press, 3–15.

Guarino, Nicola. 1998b. "Some Ontological Principles for Designing Upper Level Lexical Resources." In *Conference on Language Resources and Evolution*: *Proceedings of the First International Conference on Language Resources and Evolution 20-30 May 1998 Granada, ES*, ed. Rosa Castro, Natividad Gallardo, Antonio Rubio and Antonio Tejada. Paris: European Language Resources Association, 1-8.

Guizzardi, Giancarlo, Herre Heinrich and Gerd Wagner. 2002. "On the General Ontological Foundations of Conceptual Modeling." In *Conceptual Modeling – ER*

2002: Proceedings of the Twenty-First International Conference on Conceptual Modeling Tampere 7-11 October 2002 Finland, *FIN*, ed. Stefano Spaccapletra, Salvatore March and Yahiko Kambayashi. Berlin: Springer, 1-14.

Hjørland, Birger. 2002. "Domain Analysis in Information Science: Eleven Approaches Traditional as Well as Innovative." *Journal of Documentation* 4: 422-62.

Hjørland, Birger, 2004. "Domain Analysis: A Socio-Cognitive Orientation for Information Science Research." *Bulletin of the American Society for Information Science and Technology* 30, no. 3: 17-21.

Lancaster, Frederic Wilfrid. 2003. *Indexação e resumos: Teoria e prática*. 2nd ed. Brasília: Briquet de Lemos.

Le Moigne, Jean-Louis. 1977. *A Teoria do sistema geral: Teoria da modelização*. Lisboa: Instituto Piaget.

Lima, Gercina Ângela. 2020. "Organização e representação do conhecimento e da informação na web: Teorias e técnicas." *Perspectivas em ciência da informação* 25: 57-97.

Lima, Gercina Ângela Borém de Oliveira. 2004. "Mapa hipertextual (MHTX): Um Modelo para organização hipertextual de documento." PhD diss., Universidade Federal de Minas Gerais.

Löbner, Sebastian. 2002. *Understanding Semantics*. London: Arnold.

Lopes, Patrícia. 2018. *Proposta de metodologia para organização e representação de trilhas de aprendizagem no contexto da educação corporativa na administração pública.* Masters' thesis. Universidade Federal de Minas Gerais.

Lopes, Patrícia and Gercina Ângela Lima. 2019. "Estratégias de organização, representação e gestão de trilhas de aprendizagem: Uma revisão sistemática de literatura." *Perspectivas em Ciência da Informação* 24: 165-95.

Lubchak, Vladimir, Olena Kupenko and Borys Kuzikov. 2012. "Approach to Dynamic Assembling of Individualized Learning Paths." *Informatics in Education* 11, no. 2: 213-25.

Mey, Marc de. 1982. *The Cognitive Paradigm: An Integrated Understanding of Scientific Development*. Chicago: University of Chicago Press.

Morin, Edgar. 2000. *A cabeça bem feita: Repensar a reforma, reformar o pensamento*. Rio de Janeiro: Bertrand Brasil.

Murashima, Mary Kimiko Guimarães. 2011. "Universidades corporativas: As trilhas em meio a novos caminhos." *Revista FGV Online* 1, no. 2: 4-23.

Rada, Roy. 1991. "Trends in: A Critical Review." *Information Processing & Management* 27: 659-77.

Ranganathan, S.R. 1951. *Philosophy of Library Classification*. Copenhagen: E. Munksgaard.

Ranganathan, S.R. 1967. *Prolegomena to Library Classification*. Bombay: Asia Publishing House.

San Segundo, Rosa. 2004. "A New Conception of Representation of Knowledge." *Knowledge Organization* 31: 106-11.

Smith, Barry. 2004. "Beyond Concepts: Ontology as Reality Representation." In *Formal Ontology in Information Systems: Proceedings of the Third International Conference (FOIS) 4-6 November 2004 Turin, IT*, ed. Achille Varzi and Laure Vieu. Amsterdam: IOS Press, 73-84.

Sowa, John. 2000. *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Pacific Grove, CA: Brooks/Cole.

Stock, Wolfgang and Mechtild Stock. 2013. *Handbook of Information Science*. Berlin: De Gruyter Saur.

Studer, Rudi, Richard Benjamins and Dietel Fensel. 1998. "Knowledge Engineering: Principles and Methods." *Data & Knowledge Engineering* 25, nos. 1-2: 161-97.

Suazo, Irma Gámez, César Garita Rodríguez and Mario Chacón Rivas. 2012. "Generating Adaptive Learning Paths in e-Learning Environments." In *XXXVIII Latin America Conference on Informatics (CLEI 2012): Conference Proceedings 1-5 October 2012, Medellin, Colombia,* ed. Yezid Enrique Donoso. Piscataway, NJ: IEEE. doi: 10.1109/CLEI.2012.6427219

Vickery, Brian. 1978. "Concepts of Documentation." *Journal of Documentation* 34: 279–87.

Vickery, Brian. 1986. "Knowledge Representation: A Brief Review." *Journal of Documentation* 423: 145-59.

Wuester, Eugen. 1981. "L'étude scientifique générale de la terminologie, zone frontalière entre la linguistique, la logique, l'ontologie, l'informatique et les sciences des choses." In *Textes choisis de terminologie. I. Fondéments théoriques de la terminologie.,* ed. Guy Rondeau and Felber Helmut, 57-114. Québec: Girsterm.

Yang, Fan. 2012. "Learning Path Construction in e-Learning–What to Learn and How to Learn?" PhD diss., Durham University.