

DOI: 10.5748/9788599693131-14CONTECSI/PS-4475

METHODOLOGIES FOR CONSTRUCTION OF ONTOLOGIES: A SYSTEMATIC REVIEW OF LITERATURE

Tâmara Batista Reis de Oliveira (Universidade Salvador, Bahia, Brasil) – tamarabr38@hotmail.com

Marcio A. Pereira da Silva (Universidade Salvador, Bahia, Brasil) – marcio.alexandre83@gmail.com

Paulo Caetano da Silva (Universidade Salvador, Bahia, Brasil) – paulo.caetano@pro.unifacs.br

Nacles Bernardino Pirajá Gomes (Universidade Salvador, Bahia, Brasil) – nacles@gmail.com

The constant growth of the volume of data stored in the most diverse databases results in difficulties for an information retrieval. In this context, as ontologies are presented in areas of Computer Science and Information as instruments of knowledge representation, the banks are able to assist the activities of modeling, organization and information retrieval. Nevertheless, the lack of methodological standardization for the development of ontologies makes its construction difficult. As a way to identify as best practices for an ontology construction, this work presents a Systematic Literature Review to identify methodologies focused on ontology construction.

Keywords: Ontology, Systematic literature review, Ontology building process, methodology for ontology building.

METODOLOGIAS PARA CONSTRUÇÃO DE ONTOLOGIAS: UMA REVISÃO SISTEMÁTICA DA LITERATURA

O constante crescimento do volume de dados armazenados nos mais diversos bancos de dados resulta em dificuldades para a recuperação da informação. Neste contexto, as ontologias são apresentadas pelas áreas da Ciência da Computação e da Informação como instrumentos de representação do conhecimento, os quais são capazes de auxiliar nas atividades de modelagem, organização e recuperação de informação. Não obstante, a falta de padronização metodológica para o desenvolvimento de ontologias dificulta a sua construção. Como forma de identificar as melhores práticas para a construção de ontologias, este trabalho apresenta uma Revisão Sistemática da Literatura para identificação de metodologias voltadas à construção de ontologias.

Palavras-Chave: Ontologia, Revisão Sistemática de Literatura, processo de construção de ontologias, metodologia para construção de ontologias.

1. INTRODUÇÃO

A partir do século XX, o desenvolvimento de novas tecnologias de informação proporcionou um aumento considerável na produção e intercâmbio de registros de informações em inúmeras áreas. Ao longo do tempo, o crescimento do volume de dados armazenados vem provocando dificuldades para a recuperação da informação. Assim sendo, a preocupação em organizar, armazenar e recuperar informações de forma simples, porém eficaz, tornou-se um desafio para estudiosos e profissionais nas mais diversas áreas do conhecimento e tem conduzido pesquisadores, a exemplo de Lancaster (1986), Gruber, (1995), Berners-Lee, Hendler e Lassila (2001) e Rautenberg *et al.* (2010), a se dedicarem ao desenvolvimento de mecanismos para tratamento da informação, a fim de melhorar a eficácia dos Sistemas de Informação (SI). Segundo Iske e Boersma (2005) apud Rautenberg *et al.* (2010), tais estudos efetivamente têm contribuído para a geração de novos instrumentos capazes de auxiliar na organização e recuperação da informação de forma mais eficiente. Sob a perspectiva de Marcondes *et al.* (2008), esta dedicação tem prosseguido em direção à representação do conhecimento em formatos legíveis por computador. É dentro deste contexto que estudiosos como Moreira (2002), Gottschalg-Duque (2005) e Almeida (2006) apresentam as ontologias como um instrumento de utilização da informação que tem despertado o interesse para as atividades de modelagem e recuperação de informação. Almeida e Bax (2003), Silva (2008) e Rautenberg *et al.* (2010) observam que um grande número de pesquisas têm demonstrado a importância e a utilização das ontologias na tarefa de organização de informações. Almeida (2006), Marcondes *et al.* (2008) e Borsato (2014) afirmam que as ontologias proporcionam melhorias na recuperação de informações através da organização do conteúdo das fontes de dados que compõem o domínio. Marcondes *et al.* (2008) e Trajano (2014) observam que por utilizarem formas de representação baseadas em lógica, as ontologias possibilitam o uso de mecanismos de inferência para criar novo conhecimento a partir do existente.

O desenvolvimento de ontologias requer a utilização de uma metodologia para apoiar a construção e proporcionar mais qualidade ao processo de seu desenvolvimento. Entretanto, essa busca é dificultada pela diversidade de métodos e metodologias existentes, o que dificulta a escolha de um padrão adequado.

Alguns pesquisadores que se dedicam ao estudo de ontologias concordam que existem dois fatores principais que dificultam sua construção:

- a) Falta de consenso na padronização das metodologias de desenvolvimento: Rautenberg *et al.* (2010), Trajano (2014) e Reis e Silva (2015) aduzem que as abordagens para desenvolvimento e construção de ontologias, na maioria das vezes são específicas e limitadas. Corroborando com Fernández (1999) e Uschold e Gruninger (1996), esses autores apontam que o principal problema do ponto de vista metodológico é a falta um padrão para o processo de construção, pois não existe consenso sobre uma metodologia mais adequada para o processo de construção, o que pode resultar na ausência de atividades padronizadas, sendo a construção conduzida de forma artesanal e não como uma atividade científica. Para Almeida e Bax (2003) o desenvolvimento de uma ontologia é uma atividade que precisa ser apoiada de alguma forma, por exemplo, por meio de ferramentas e métodos;

b) Falta de explicações sobre as abordagens adotadas: este problema é apontado nos estudos de Silva (2008) e Reis e Silva (2015) e consiste na falta de explicações sistemáticas de como, onde e sob quais perspectivas as abordagens teóricas são utilizadas no processo de elaboração de ontologias.

Diante dos entraves identificados para a construção de ontologias, na tentativa de dirimir as dificuldades encontradas para a seleção de uma metodologia de apoio ao processo de desenvolvimento, este trabalho realiza um estudo analítico na literatura, especializada por meio de uma Revisão Sistemática de Literatura (RSL) sobre as metodologias e métodos para construção de ontologias, para o levantamento do estado da arte sobre este tema.

A escolha da RSL como ferramenta metodológica para apoiar este trabalho está respaldada pela observação de estudos que consideram a relevância desta prática, a exemplo de King e He (2005) e Kitchenham *et al.* (2010), os quais afirmam que o acúmulo de conhecimentos depende cada vez mais da integração entre estudos e descobertas anteriores, levando-se em consideração que um percentual considerável de estudos em Sistemas de Informação (SI) utilizam a RSL como metodologia de pesquisa. Para Dyba e Dingsoyr (2008) e Kitchenham *et al.* (2010), a RSL é considerada o pilar da prática baseada em evidências da literatura especializada. Portanto, a escolha da realização da Revisão Sistemática de Literatura é fundamentada na necessidade de busca por estudos, que servirão como alicerces para a identificação das principais metodologias para a construção das ontologias.

A seguir, na Seção 2 é mostrada a metodologia adotada neste trabalho para a execução de Revisão Sistemática da Literatura. Na Seção 3, alguns resultados encontrados são discutidos. Nas Seções 4 e 5, algumas limitações e perspectivas são discutidas com a finalidade de nortear trabalhos futuros.

2. METODOLOGIA: PROCEDIMENTOS ADOTADOS PARA A EXECUÇÃO DA REVISÃO SISTEMÁTICA DE LITERATURA

Para garantir um nível satisfatório na recuperação dos estudos primários, a execução desta Revisão Sistemática da Literatura adotou o processo proposto por Biolchini *et al.* (2005), que obedeceu três etapas distintas: Planejamento da Revisão, Condução da Revisão e Publicação dos Resultados.

2.1. Planejamento da revisão

Nesta fase foi definido um protocolo para a execução da Revisão Sistemática da Literatura (RSL), o qual seguiu as etapas descritas na sequência.

- 1) Descrição dos objetivos da Revisão Sistemática: identificar estudos que contenham as principais metodologias e métodos dedicados ao desenvolvimento de ontologias, a fim de identificar o(s) que melhor(es) se adapta(m) ao processo de construção de uma ontologia de domínio.
- 2) Elaboração das questões de pesquisa e estrutura PIO: conforme Marconi e Lakatos (2003), este trabalho considerou os elementos PIO para cada uma das questões de pesquisa: a) população; b) intervenção; e c) resultado. Neste contexto, foi realizada uma busca a partir das seguintes questões de investigação:
 - a) Questão de pesquisa 1: Quais são as metodologias mais utilizadas

para a construção de uma ontologia de domínio?

- População (P): publicação que aborde a representação do conhecimento através de ontologias.
- Intervenção (I): propostas que apresentem metodologias para a construção de ontologias.
- Resultados (O): as principais publicações que propõem metodologias para a construção de ontologias devidamente identificadas.

b) Questão de pesquisa 2: Quais são as etapas necessárias ao processo de construção de uma ontologia de domínio?

- População (P): publicação que aborde a construção de ontologia de domínio.
- Intervenção (I): métodos ou metodologias que abordem as etapas do desenvolvimento de ontologias.
- Resultados (O): os principais métodos ou metodologias que tratam das etapas para a construção de ontologias foram identificados.

3) A estratégia de busca utilizadas neste trabalho baseia-se em dois critérios, os quais são descritos a seguir:

a) Seleção das fontes de pesquisa: Conforme sugerem Marconi e Lakatos (2003), as bases de dados que serviram como fontes para a obtenção dos estudos foram selecionadas, considerando-se as entidades que disseminam literatura notavelmente reconhecidas pela comunidade científica como de caráter voltado à pesquisa especializada. São elas: *Digital Library* do Portal ACM, o Portal CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) e *IEEEExplore Digital Library*;

b) Elaboração de *Strings* de busca: para auxiliar na recuperação dos estudos foram elaboradas *strings* de busca que funcionaram como motor de busca nas fontes selecionadas. As *strings* foram construídas através da composição de termos identificados a partir das estruturas PIO das questões de pesquisa, são eles: *ontologia, construção e metodologia*. Estes termos e seus sinônimos foram traduzidos para a língua inglesa a fim de se definir a *string* de busca padrão, que ficou assim definida: "*((ontology) AND (construction OR building OR development OR implementation) AND (methodology OR method OR rule OR standard OR best practices OR guideline))*". Para apoiar a recuperação dos documentos, à partir da *string* de busca padrão, uma segunda *string* foi elaborada acrescentando-se termos "*design*" e "*framing*", sinônimos de "*construction*": "*((ontology) AND (construction OR building OR development OR implementation OR design OR framing) AND (methodology OR method OR rule OR standard OR best practices OR guideline))*". Para todas as bases de dados eletrônicas utilizadas, as *strings* de busca foram revisadas conforme a particularidade ou limitação da máquina de busca, visando observar a importância dos

termos usados e a pertinência das publicações retornadas.

- 4) Adoção de critérios para inclusão e exclusão de estudos: tais critérios foram definidos para auxiliar a análise primária de seleção dos estudos:
 - a) Critérios para inclusão dos estudos: As publicações devem estar disponíveis na *Web*, na *Digital Library* do Portal ACM, no Portal CAPES ou na *IEEEExplore Digital Library*; estudos que apresentem palavras da *string* de busca no seu título ou no seu resumo; estudos escritos em inglês, espanhol ou português; publicações que representem metodologias ou métodos que tratam de construção de ontologias e abordam o assunto de forma analítica e não apenas como uma citação; será levada em consideração a ocorrência de quantificação de citações identificadas; estudos que respondam a qualquer uma das questões de pesquisa;
 - b) Critérios para exclusão dos estudos: estudos que apresentavam ontologias, mas não abordavam metodologias para a construção; estudos que não respondam a nenhuma das questões de pesquisa.

2.2. Condução da revisão

O processo de condução da Revisão Sistemática ocorreu da seguinte forma: (i) as *strings* de busca foram executadas nas fontes selecionadas; (ii) os estudos primários foram identificados; (iii) os documentos retornados foram selecionados de acordo com os critérios de inclusão/exclusão e avaliados segundo os critérios de qualidade estabelecidos durante o planejamento da revisão.

2.2.1. Processo para a recuperação e pré-seleção dos estudos primários

Nesta etapa foram recuperados os estudos preliminares e realizado o primeiro processo de seleção dos estudos. Após a execução das *strings* de busca nas respectivas bases de dados selecionadas, os estudos retornados foram inseridos na ferramenta JABREF (Jabref, 2009) versão 2.4.2, a qual é uma ferramenta utilizada para gerenciar referências, a fim de apoiar a análise dos documentos. Os resultados quantitativos de documentos recuperados são apresentados na Tabela 1.

Tabela 1 - Resultado quantitativo dos documentos recuperados

Para cada estudo retornado foram extraídas as seguintes informações, consideradas relevantes à análise do documento: autor(es); título do documento; ano de publicação; procedência; pertinência do resultado da busca. Este processo pode ser observado pela captura de tela da JABREF, ilustrada na Figura 1.

QuickSearch: <input type="text"/> <input type="button" value="clear"/>		Number of matching entries: 67/67		Search Settings	
Author	Title	Year	Journal/Proceedings	Review	DOI/URL
BRONZA, M. C. S. F.	<i>NeOn Methodology for building ontology networks: specification, scheduling and reuse.</i> [Abstract] [Review]	2010	268 f.. Tese (Doutorado) - Universidade Politécnica de Madrid	Ok - conteúdo pertinente	URL
USCHOLD, M.; GRUNINGER, M.	<i>Ontologies: principles, methods an applications</i> [Abstract] [Review]	1996	Knowledge Engineering Review, v. 11, n. 2.	ok - conteúdo pertinente	URL
RAUTENBERG, S., FILHO, A. C., TODESCO, J. L., OSTUNI-GAUTHIER, F. A.	<i>Ferramenta ontoKEN: Uma Contribuição à Ciência da Informação para o Desenvolvimento de Ontologias</i> [Abstract] [Review]	2010	Perspectivas em Ciência da Informação, v.15, n.1, p. 239-258	não - não aborda metodologia/propõe uma ferramenta	URL
GRUNINGER, M.; FOX, M. S.	<i>Methodology for the design and evaluation of ontologies.</i> [Abstract] [Review]	1995	Department of Industrial Engineering University of Toronto, Toronto, Canada M5S 1A4	ok - conteúdo pertinente	DOI

Figura 1. Formulário para o gerenciamento dos estudos retornados.

A ferramenta *JabRef* fornece algumas vantagens auxiliares na seleção dos estudos, permitindo: a) recuperar o documento completo; b) categorizar referências; c) mapear nos documentos os critérios de inclusão e exclusão; d) ler e extrair informações dos artigos; e) utilizar os campos disponíveis para comentários e revisões; f) customizar os dados de entrada conforme a necessidade da pesquisa; h) identificar documentos repetidos. A análise dos documentos ocorreu pela leitura do título e do resumo, observando-se os critérios de inclusão/exclusão (vide alínea d, Seção 2.1). Uma avaliação inicial excluiu as referências sem disponibilidade de acesso via *Web*, os estudos repetidos e os estudos que não atenderam diretamente aos critérios de inclusão/exclusão, o que reduziu o número de estudos selecionados de sessenta e sete para trinta e uma publicações.

2.2.2 Processo de seleção dos estudos

Neste estágio, os artigos recuperados foram analisados em seu conteúdo a fim de serem selecionados os estudos para compor os resultados desta RSL. Marconi e Lakatos (2003) e Mafra e Travassos (2006) observam que uma avaliação mais detalhada faz-se necessário, para garantir a qualidade do conteúdo dos estudos. Portanto, os trinta e um documentos resultantes da avaliação inicial sofreram uma segunda avaliação, em que foram considerados critérios de qualidade do conteúdo dos estudos, conforme apresentado na sequência.

a) Segunda avaliação: análise qualitativa do conteúdo dos estudos selecionados: o segundo processo de avaliação dos documentos recuperados estabeleceu procedimentos qualitativos específicos para definir os objetos para compor os resultados desta Revisão Sistemática, o que resultou na seleção quantitativa de treze artigos, os quais foram documentados e incluídos no quadro de *Estudos Selecionados*, conforme ilustra o Quadro 1. Uma análise preliminar nos conteúdos das publicações mostrou que a maior parte das metodologias - e.g. *Methontology* (Fernández et al., 1997), *On-to-Knowledge* (Sure et al., 2002) e NeOn (Baonza, 2010) - apresentam semelhanças entre as fases de desenvolvimento de ontologias com as fases relativas ao processo de desenvolvimento de *software*. Algumas dessas semelhanças foram identificadas também por outros autores a exemplo de Fernández (1999) e Silva (2008), os quais recomendam a utilização do padrão aceito internacionalmente para desenvolvimento de *software*, a norma IEEE-1074 (IEEE, 1997), como instrumento de análise qualitativa no desenvolvimento de ontologias. Para Fernández (1999), o processo de desenvolvimento de ontologias deve ser fundamentado no padrão internacional IEEE-1074 por ser esta uma norma relativa à Engenharia de *Software* e pelo fato das ontologias serem consideradas componentes de produtos de *software*. Corroborando, Silva (2008) observa que Fernández, Gómez-Pérez e Juristo (1997), Uschold e Gruninger (1996), e Noy e MacGuinnes (2001) apresentam as categorias para o desenvolvimento de ontologias como sendo equivalentes às *fases de ciclo de vida software*, observando que elas foram adaptadas diante aos processos extraídos da norma IEEE-1074 e características particulares às ontologias, de *formalização e integração*. Desse modo, pode-se conferir a maturidade de cada metodologia em relação ao modelo de *ciclo de vida*, e assim, avaliar a qualidade da metodologia. Com base nas experiências destes autores e mediante a perspectiva deste trabalho, definiu-se pela utilização da norma IEEE-1074 como referencial de qualidade para a análise das metodologias de desenvolvimento de ontologia, por descrever um processo estruturado e metódico para o desenvolvimento de produtos de *software*, conforme Pressman (2002). Assim sendo, conforme proposto por Silva (2008), a fundamentação de cada categoria de análise - ou fases do ciclo de vida, que servirá como medida de maturidade e qualidade para a seleção de metodologias de desenvolvimento de ontologia - é apresentada na sequência:

- **Gerenciamento do projeto:** ocorrem atividades relacionadas ao início de um projeto como criação do processo, planejamento da gestão, monitoramento e controle bem como todo o ciclo de vida do *software*;
- **Pré-desenvolvimento:** fase que inclui atividades de estudo de viabilidade e análise de requisitos do sistema;
- **Especificação de requisitos:** fase de definição das necessidades do requisitante e determina as restrições ou regras que o *software* (ou ontologia, no caso deste trabalho) deverá cumprir. Serve como documento inicial para a realização das tarefas de modelagem e prototipação;

- **Modelagem conceitual:** consiste em desenvolver uma representação do sistema capaz de satisfazer os requisitos de *software* especificados na fase anterior;
- **Formalização:** utiliza técnicas oriundas da área de Inteligência Artificial a fim de transformar o modelo conceitual da ontologia em um modelo formal para definir de forma eficiente o seu significado;
- **Implementação:** objetiva a transformação da representação do projeto de arquitetura do *software* em uma linguagem de programação. Aplicando este processo às ontologias, a implementação consiste em mapear o modelo formal em uma linguagem adequada às demandas, e.g. OWL (*Ontology Web Language*), XML (*Extensible Markup Language*);
- **Manutenção:** etapa pós-desenvolvimento, consiste em identificar problemas e promover melhorias nos produtos;
- **Integração:** considera a reutilização de conceitos existentes em meta ontologias, ou ontologias de alto nível, buscando integrar a ontologia em questão a ontologias existentes. Tais atividades podem ser realizadas durante a fase de modelagem conceitual e implementação, sendo considerada, portanto, um processo integral;
- **Avaliação:** esta fase é executada simultaneamente às atividades de desenvolvimento do *software* como, por exemplo: condução de revisões e auditorias nos processos, execução de testes e avaliação de resultados;
- **Documentação:** trata das atividades de desenvolver e distribuir documentação para desenvolvedores e envolvidos nos processos, a fim de fornecer informações sobre a ontologia.

2.2.3. Extração de informações dos documentos selecionados

Após a obtenção da amostragem documental selecionada, foi realizada a identificação das metodologias propostas nos documentos. Nesta etapa foram extraídos e sintetizados os dados relevantes de cada um dos estudos selecionados, para serem publicados durante a etapa de Publicação dos Resultados, conforme indicado por Mafra e Travassos (2006). Os documentos selecionados foram incluídos no quadro de Estudos Selecionados, conforme ilustrado no Quadro 1.

É importante observar que alguns estudos retornados também abordavam normas para a construção de vocabulários controlados e tesouros, o que pode ser de grande proveito para a elaboração de vocabulários pertencentes às ontologias. Assim sendo, a norma ANSI/NISO Z39.19-2005-R2010 (NISO, 2014) e o Manual BITI (Campos et al., 2004), foram selecionados por contemplarem quase todas as fases de ciclo de vida, além de terem sido os mais citados pela maioria dos autores, a exemplo de Almeida e Bax (2003), Silva (2008) e Rautenberg *et al.* (2010). O Quadro 1 apresenta o resultado da seleção dos estudos.

REFERÊNCIA	CARACTERÍSTICAS RELEVANTES
Gruninger e Fox (1995)	O estudo apresenta a Metodologia de Gruninger e Fox, dedicada a criar modelos de senso comum sobre empresas através de um conhecimento compartilhado sobre o negócio que conduza a deduções de respostas sobre questões acerca do domínio.
Uschold e Gruninger (1996)	Os autores propõem a <i>Enterprise Ontology</i> , uma extensão do Método de Uschold e King, que considera como estágios para o desenvolvimento de uma ontologia: i) identificação do propósito; ii) construção; iii) avaliação; iv) documentação.
Bernaras, Laresgoiti e Corera (1996)	É proposto o Método Kactus que utiliza um método recursivo consistente em uma proposta inicial para uma base de conhecimento; quando é necessária uma nova base em domínio similar, generaliza-se a primeira base em uma ontologia adaptada a ambas as aplicações; quanto mais aplicações, mais genérica a ontologia.
Swartout, Patil, Knight e Russ (1996)	O trabalho apresenta o método <i>SENSUS</i> , baseado em uma ontologia que leva o mesmo nome, propõe processos para estabelecer as ligações entre os termos específicos e os termos da ontologia de alto nível.
Férrnandez, Gómez-Pérez e Juristo (1997)	O trabalho apresenta a metodologia <i>Methontology</i> que se baseia em um conjunto de estágios de desenvolvimento (especificação, conceitualização, formalização, integração, implementação e manutenção), um ciclo de vida que compreende evolução de protótipo e técnicas para realizar as atividades de planejamento, desenvolvimento e suporte.
Falbo, Menezes e Rocha (1998)	Os autores apresentam a metodologia SABIO que estabelece uma abordagem sistemática para construção de ontologias através da incorporação de características das metodologias existentes até então, ao passo em que propõem novas características: utilização de linguagem gráfica para representar a ontologia; classificação axiomática e algumas diretrizes para a captura, formalização, avaliação e documentação de ontologias.
Sure, Staab e Studer (2002)	A metodologia <i>On-to-Knowledge</i> foi elaborada visando o desenvolvimento de ontologias a serem empregadas em Sistemas de Gestão do Conhecimento, e compreende cinco fases: i) estudo de viabilidade; ii) fase inicial; iii) refinamento; iv) avaliação; v) manutenção e evolução.
Noy e McGuinness (2001)	É proposto o Método 101 que adota quatro atividades para o desenvolvimento de uma ontologia: i) definir as classes; ii) organizar as classes em uma taxonomia; iii) definir as propriedades para as classes e descrever seus valores permitidos; iv) adicionar valores de propriedades para as instâncias.
Reed e Lenat (2002)	A obra estuda a metodologia Cyc, cuja base de conhecimento foi desenvolvida compreendendo três processos: i) extração do

	conhecimento de senso comum; iii) extração auxiliada por computador; e iii) extração gerenciada por computador.
Campos, Gomes e Motta (2004)	Os autores apresentam o Manual BITI que trata da elaboração de um tipo de vocabulário controlado, os tesouros, mostrado sequencialmente e em detalhes, os procedimentos necessários à sua construção.
Silva (2008)	Este trabalho apresenta uma metodologia para o desenvolvimento de ontologias de domínio com base em metodologias já existentes, através de uma perspectiva interdisciplinar entre as Ciências da Informação e da Computação.
ANSI/NISO Z39.19-2005-R2010 (2010)	O padrão ANSI/NISO Z39.19-2005-R2010 permite construir diversos tipos de vocabulário controlado como tesouros, taxonomias, listas e anel de sinônimos, em uma ordem conhecida e estruturada com o intuito de disponibilizar claramente os relacionamentos de equivalência, associativos e hierárquicos.
Baonza (2010)	A metodologia NeOn foi elaborada com o objetivo de dirimir as principais limitações das metodologias conhecidas, ao passo em que procurou se beneficiar das vantagens oferecidas por cada uma, além de cobrir um novo paradigma de desenvolvimento de ontologia, baseado em reuso, desenvolvimento colaborativo e redes de ontologias.

Quadro 1 - Estudos selecionados através da Revisão Sistemática de Literatura

2.3. Publicação dos resultados

Conforme observado no Quadro 1, o material selecionado para compor os resultados da RSL priorizou 11 metodologias/métodos para construção de ontologias, uma metodologia para a construção de tesouros e uma norma para construção de vocabulários controlados, totalizando 13 objetos de estudo. Com base em Marconi e Lakatos (2003), a publicação dos resultados compreendeu a apresentação de uma síntese da análise do conteúdo dos estudos selecionados, o que possibilitou a classificação dos mesmos mediante as categorias de análise para o desenvolvimento de ontologias, conforme discutido na Seção 2.2.2.

2.3.1 Classificação do conteúdo dos estudos selecionados

As categorias de análise para o desenvolvimento de ontologias foram denominadas e adaptadas aos processos extraídos da norma IEEE-1074, de modo que estas categorias representam as fases que descrevem o *ciclo de vida* do desenvolvimento da ontologia. Para ilustrar melhor o processo de classificação das metodologias selecionadas foi elaborado um quadro de classificação dos estudos selecionados, conforme ilustrado no Quadros 2, em que a classificação comparativa dispõe os objetos investigados em suas colunas e as *fases do ciclo de vida* de cada um dos estudos em suas linhas. Neste contexto, as fases que fazem parte da proposta do objeto investigado tiveram a cédula preenchida, ao passo em que as fases que não fazem parte da proposta da metodologia em questão, a célula referente encontra-se em branco.

Fases do ciclo de vida		Gruningher e Fox	Enterprise	Methontology	Método 101	On-to-Knowledge	Silva (2008)	NeOn	Manual BITI	Norma ANSI/NISO Z39.19-2005	Processo elaborado para a construção da ontologia
Gerenciamento do projeto											
Processos Orientados	Pré-desenvolvimento										
	Desenvolvimento	Especificação de requisitos									
		Modelagem conceitual									
		Formalização									
		Implementação									
	Pós-desenvolvimento										
Processos de Integração	Integração										
	Avaliação										
	Documentação										

Quadro 2 - Classificação dos estudos selecionados

3. RESULTADOS ALCANÇADOS

Após a análise dos estudos investigados, tendo sido observados a partir das categorias de análise adotadas na Seção 2.2.2 e das informações dispostas no Quadro 2, foi possível delinear algumas considerações sobre as principais características de cada um dos estudos e realizar uma análise comparativa, a fim de proporcionar ao usuário desta Revisão Sistemática de Literatura uma visão sobre a abordagem proposta em cada um dos documentos analisados, facilitando assim a escolha da metodologia mais adequada ao processo de construção da ontologia. Esta análise obedeceu a uma síntese de critérios, conforme observado em Fernández (1999) e Silva (2008):

1) Ciclo de desenvolvimento: algumas abordagens seguem as fases de ciclo de vida, outras não: a) nesse quesito, as mais completas são a metodologia NeOn e a proposta metodológica de Silva (2008), as quais contemplam todas as fases do ciclo de vida e a *Methontology*, que não propõe apenas a fase de pré-desenvolvimento; b) o Manual BITI apresentou um ciclo de vida quase completo segundo a Norma IEEE-1074, observando-se que não foram consideradas as fases de implementação e integração, por não pertencerem ao propósito dos vocabulários controlados; c) a SABIO não possui recomendações explícitas a cerca do ciclo de vida da ontologia; d) os métodos Cyc e Método 101 se destacam na implementação da ontologia, entretanto, aspectos importantes como o gerenciamento do projeto, o estudo de viabilidade, a manutenção e a avaliação de ontologias são desprezados; e) a norma ANSI/NISO Z39.19-2005 não objetiva propor um ciclo de atividades para a construção de vocabulários controlados, e sim recomenda uma série de tratamentos em relação ao termo, em atividades que se enquadram no processo de construção.

- 2) Detalhamento das atividades e dos procedimentos: algumas metodologias e métodos mostram um baixo nível de detalhamento na elucidação dos passos para construção de ontologias como *Cyc*, *Kactus*, *Enterprise*, Gruninger e Fox (1995), Método 101 e *SENSUS*. Já a *Methontology*, a NeOn e a proposta de Silva (2008) se destacam por fornecer um nível mais alto de detalhes de como proceder na condução das atividades propostas. Entretanto, estas metodologias não abordam a possibilidade de formalização da ontologia com o apoio de uma ferramenta.
- 3) Princípios teóricos e metodológicos para identificação, definição e organização de conceitos: a) neste quesito se destacou o manual BITI, pois nos vocabulários controlados é evidente a inserção de princípios da teoria da classificação, segundo Ranganathan (1967) e da teoria do conceito, conforme Dahlberg (1978) na especificação de seus elementos; b) as metodologias e os métodos para construção de ontologia indicam recursos para a definição de conceitos e como organizá-los em uma taxonomia, porém, não fornecem clareza quanto aos princípios teóricos utilizados. Este processo foi mais bem elucidado por Silva (2008).
- 4) Dependência em relação à aplicação: as metodologias *Kactus*, Gruninger e Fox, *SENSUS* e *On-To-Knowledge* demonstraram possuir alta dependência da aplicação desenvolvida em conjunto com a ontologia.
- 5) Recomendações para a Formalização do Conhecimento: neste critério destacaram-se: Gruninger e Fox, *SENSUS*, *On-To-Knowledge*, *Methontology* e NeOn.
- 6) Técnicas Recomendadas: nas metodologias SABiO, *On-To-Knowledge*, *Methontology*, NeOn e Silva (2008) foram identificadas recomendações de técnicas em apoio à construção de ontologias.

4. LIMITAÇÕES ENCONTRADAS

Durante a execução das *strings* de busca, observou-se que algumas máquinas de busca apresentaram limitações, e.g. em relação ao tamanho da *string*, impedindo uma correta execução. Portanto, em algumas plataformas, as *strings* de busca tiveram que ser resumidas, a fim de se adaptarem às necessidades da máquina de busca em algumas fontes (e.g Portal CAPES).

Outra dificuldade detectada foi à falta de indicação de ferramentas para apoiar o processo de formalização da ontologia. Portanto, este trabalho adotou as recomendações de Fernández et al. (2004), observando que a formalização da ontologia pode contar com o apoio de uma ferramenta que utilize linguagem baseada em lógica descritiva (e.g. ferramenta *Protégé* e linguagem OWL).

5. PERSPECTIVAS FUTURAS

Como perspectivas de trabalhos futuros, os autores deste artigo esperam que esta Revisão Sistemática de Literatura possa auxiliar na escolha de uma metodologia mais adequada ao processo de construção de novas ontologias, conforme a necessidade do desenvolvedor.

Espera-se também que as metodologias aqui discutidas e que também apresentam ontologias já construídas, possam ser reaproveitadas e integradas a novos projetos, pois se tratam de

ontologias notadamente reconhecidas pela comunidade científica, a exemplo do Método 101, *Enterprise e Methontology*.

6. REFERÊNCIAS

- Almeida, M. B. (2006). *Um modelo baseado em ontologias para a representação da memória*. 2006. 345 f. Tese Doutorado. Programa de Pós-Graduação em Ciência da Informação, Escola de Ciência da Informação, Universidade Federal de Minas Gerais - UFMG.
- Almeida, M. B.; Bax, M. P. (2003). *Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção*. Ci. Inf. 32, 3 (set./dez. 2003) 7-20.
- Baonza, M. C. (2010). *NeOn methodology for building ontology networks: specification, scheduling and reuse*. Tese Doutorado. Faculdade de Informática, Universidad Politécnica de Madrid.
- Bernaras, A.; Laresgoiti, I.; Corera, J. (1996). *Building and Reusing Ontologies for Electrical Network Applications*. In: Proceedings of the European Conference on Artificial Intelligence, ECAI/96, 12, 298-302. Budapest.
- Berners-Lee, T.; Hendler, J.; Lassila, O. (2001). *The semantic Web*. Scientific American. 284, 5 (maio 2001) 34-43.
- Biolchini, J.; Mian, P. G.; Natali, A.; C.C.; Travassos, G. H. (2005). *Systematic Review in Software Engineering*. Programa de Engenharia de Sistemas e Computação, Universidade Federal do Rio de Janeiro.
- Borsato, M. (2014). *Bridging the gap between product lifecycle management and sustainability in manufacturing through ontology building*. Computers in Industry. (fever. 2014) 258–269. Amsterdam.
- Campos, M. L; Gomes, H. E.; Motta, D. F. (2004). *Elaboração de Tesouro Documentário - Tutorial*. Disponível em: <<http://www.conexaorio.com/bit/tesauro/>>.
- Dalhberg, I. (1978). *Teoria do Conceito*. Ciência da Informação. 7, 2, 101-107. Rio de Janeiro.
- Dyba, T.; Dingsoyr, T. (2008). *Strength of evidence in systematic reviews in software engineering*. In: ACM-IEEE International Symposium on Empirical Software Engineering and Measurement. 2, 178-187. New York.
- Falbo, R. A.; Menezes, C. S.; Rocha, A. R. (1998). *A Systematic approach for building ontologies*. In: Ibero-American Conference on AI, 6., 1998. Lisboa. Anais... Springer Berlin Heidelberg, 2003, 349-360.

- Fernández, M. (1999). *Overview of methodologies for building ontologies*. Laboratorio de Inteligencia Artificial, Facultad de Informática - Universidad Politécnica de Madrid.
- Fernández, M.; Gómez-Pérez, O.; Juristo, H. (1997). *Methontology: from ontological art towards ontological engineering*. Laboratorio de Inteligencia Artificial, Facultad de Informática, Universidad Politécnica de Madrid Madrid. 1997.
- Fernández, M.; Gómez-Pérez, A., Corcho, O. (2004). *Methodologies and methods for building ontologies*. In: Gomez-Perez, A., Fernandez- Lopes, M, Corcho, O. *Ontological Engineering*. Springer, 2004, 107-153. London.
- Gottschalg-Duque, C. (2005). *SiRILiCO: uma proposta para um sistema de recuperação de informação baseado em teorias da linguística computacional e ontologia*. 117 f. Tese Doutorado. Programa de Pós-Graduação em Ciência da Informação, Escola de Ciência da Informação, Universidade Federal de Minas Gerais – UFMG.
- Gruber, T. (1995). *Toward Principles for the Design of Ontologies Used for Knowledge Sharing*. International Journal Human-Computer Studies. 43. 5 (nov. 1995) 907-928, Palo Alto.
- Gruninger, M.; Fox, M. S. (1995). *Methodology for the design and evaluation of ontologies*. Department of Industrial Engineering, M5S 1A4. University of Toronto.
- IEEE. (1997). IEEE 1074 - Standard for Developing Software Life Cycle Processes. Disponível em: <http://arantxa.ii.uam.es/~sacuna/is1/normas/IEEE_Std_1074_1997.pdf>.
- Jabref. (2009.) *Open Source Bibliography Reference Manager*. Disponível em: <<http://jabref.sourceforge.net/download.php>>
- King, W.; He, J. (2005). *Understanding the roles and methods of meta-analysis in IS Research*. Communications of the AIS 16, 32 (out, 2005) 665-686.
- Kitchenham B.; Pretorius R., Budgen, D.; Brereton O. P.; Turner M.; Niazi M., Stephen L. (2010). *Systematic literature reviews in software engineering – A tertiary study*. Elsevier Information and Software Technology. 52, 8 (agos. 2010) 792-805. Newton.
- Lancaster, F. W. (1986). *Vocabulary control for information retrieval*. 2 ed., 280, IRP, Virgínea.
- Mafra, S., Travassos, G. (2006). *Primary and Secondary Studies Supporting the Search for Evidence in Software Engineering*. Technical Report [in Portuguese]. COPPE, Federal University of Rio de Janeiro.
- Marcondes, C. H. et al. (2008). *Ontologias como novas bases do conhecimento científico*. Perspectivas em Ciência da Informação. 13, 3(set/dez. 2008) 20-39. Belo Horizonte.

- Marconi, M.; Lakatos, E. (2003). *Metodologia científica: ciência e conhecimento científico*. 5. ed. Atlas, São Paulo.
- Moreira, A. (2002). *Uso de ontologia em sistemas de informação computacionais*. Perspectivas em Ciência da Informação. 7, 1(jan/ jun 2002.) 49-60. Belo Horizonte.
- NISO (2014) ANSI/NISO Z39.19-2005. (2010) Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. Bethesda: NISO Press, 2010. 172 p. Disponível em: <http://www.niso.org/apps/group_public/project/details.php?project_id=46>.
- Noy, N. F.; McGuinness, D. L. (2001). *Ontology development 101: A guide to creating your first ontology*. Knowledge Systems Laboratory - Stanford University.
- Pressman, R. S. *Engenharia de Software*. 5. ed. Rio de Janeiro: McGraw-Hill. (2002). 842p.
- Ranganathan, S. R. (1967). *Prolegomena to Library classification*. 640p. Asia Publishing House, New York, NY.
- Rautenberg, S.; Filho, A. C.; Todesco, J. L., Ostuni-Gauthier, F. A. (2010). *Ferramenta ontoKEM: uma contribuição à Ciência da Informação para o desenvolvimento de ontologias*. Perspectivas em Ciência da Informação. 15, 1(jan./abr. 2010) 239-258. Belo Horizonte.
- Reed, S. L.; Lenat, D. B. (2002). *Mapping ontologies into Cyc*. In: American Association for Artificial Intelligence. Technical Report WS-02-11.
- Reis, T. B.; Silva, P. C. (2015). *Ontology for Concepts's Integration of Corporate Sustainability Index (ISE) and G4 Guidelines of Global Reporting Initiative"*. In: 14th Internacional Conference [www/internet](http://www.internet). Maynooth.
- Silva, D. L. (2008). *Uma proposta metodológica para construção de ontologias: uma perspectiva interdisciplinar entre as ciências da informação e da computação*. Dissertação Mestrado. Programa de Pós-graduação em Ciência da Informação. Universidade Federal de Minas Gerais.
- Sure, Y.; Staab, S.; Studer, R. (2002). *Methodology for Development and Employment of Ontology based Knowledge Management Applications*. ACM SIGMOD Record, 31, 4, 18-23.
- Swartout, B.; Patil, R.; Knight, K.; Russ, T. (1996). *Toward distributed use of large-Scale ontologies*. Information Sciences Institute, University of Southern California.
- Trajano, M. (2014). *Estudo do processo de apropriação da ontologia pela Ciência da Informação no Brasil*. 2014. 270 f. Dissertação Mestrado. Programa de Pós-Graduação em

Ciência da Informação, Universidade Federal de Pernambuco.

Uschold, M.; Gruninger, M. (1996). *Ontologies: Principles, Methods an Applications*. 11, 2. Knowledge Engineering Review - University of Edinburgh.