

MODELO DE ECOSSISTEMA SEMÂNTICO DE INFORMAÇÕES CORPORATIVAS PARA PROCESSAMENTO DE OBJETOS MULTIMÍDIA

MODEL OF SEMANTIC ECOSYSTEM OF CORPORATE INFORMATION FOR PROCESSING MULTIMEDIA OBJECTS

Sergio de Castro Martins¹

RESUMO

Neste trabalho é proposto um modelo conceitual de ecossistema semântico de informações corporativas tendo como foco o processamento de objetos de informação não estruturados, em especial os multimídia, amplamente utilizados em ambientes de negócios. A metodologia baseou-se no Design Science, abordagem que visa resolver problemas práticos e, também, de problemas de conhecimento. O modelo baseia-se na aplicação de vocabulários semânticos a objetos multimídia, de modo a expandir as suas potencialidades de representação e recuperação. O modelo proposto é construído de modo a aproveitar-se da utilização de elementos estáveis no ambiente organizacional, como atores, processos, metadados de negócios, objetos de informação e, também, algumas infraestruturas básicas do ambiente informacional corporativo. As etapas previstas do modelo são: Ambientação, Declaração de parâmetros, Destilação dos dados, Enriquecimento de metadados e Armazenamento. Como resultado, em termos teóricos o modelo permite o processamento de dados heterogêneos e não estruturados segundo os recortes estabelecidos e mediante os parâmetros acima elencados. Além disso, o modelo prevê aplicação e adaptabilidade ao ecossistema informacional de variados tipos de organizações.

Palavras-chave: Enriquecimento semântico de metadados. Organização da Informação e do Conhecimento. Gestão de Conteúdo Corporativo. Web Semântica. Multimídia.

ABSTRACT

In this work, a conceptual model of the semantic ecosystem of corporate information is proposed, focusing on the processing of unstructured information objects, especially multimedia, widely used in business environments. The methodology was based on Design Science, an approach that aims to solve practical problems and also knowledge problems. The model is based on the application of semantic vocabularies to multimedia objects, in order to expand their potential for representation and retrieval. The proposed model is built in order to take advantage of the use of stable elements in the organizational environment, such as actors, processes, business metadata, information objects and, also, some basic infrastructures of the corporate information environment. The stages of the model are: Setting, Parameter declaration, Data distillation, Metadata enrichment and Storage. As a result, in theoretical terms the model allows the processing of heterogeneous and unstructured data according to the established cuttings and using some parameters. In addition, the model provides for application and adaptability to the informational ecosystem of various types of organizations.

Keywords: Metadata Semantic Enrichment. Information Organization and Representation. Enterprise Content Management. Semantic Web. Multimedia.

Artigo submetido em 01/12/2020 e aceito para submissão em 15/12/2020

¹ Doutor em Ciência da Informação pela Universidade Federal Fluminense, Brasil. Líder do Grupo de Pesquisa Tecnologias Semânticas e Inteligência Informacional. ORCID <https://orcid.org/0000-0003-2406-6648> . E-mail: smartins@id.uff.br

1 INTRODUÇÃO

As sociedades contemporâneas têm se destacado pela produção, veiculação e consumo de informações num volume, velocidade e variedade sem precedentes na história. Este fenômeno é percebido também pela reconfiguração de práticas sociais, econômicas e científicas marcadamente distintas de épocas anteriores, consolidadas pelo uso de dados massivos. Desde que a convergência das tecnologias da comunicação e da informação deram origem à chamada “sociedade da informação” (Castells, 2016), mudanças significativas passaram a afetar as formas de produção, armazenamento, representação e recuperação da informação.

Com o advento da web tal como a conhecemos, nos anos 1990, a produção e o uso de informações digitais aumentaram exponencialmente, modificando os processos de produção do conhecimento global. Se antes a forma tradicional de aquisição de conhecimento se dava sobretudo em universidades e bibliotecas, a partir de então o eixo começa a deslocar-se para a web. Neste sentido, a web passa a absorver a produção do conhecimento em diversos formatos: tanto quanto os textos, também áudios, vídeos e imagens tornaram-se os objetos de informação mais expressivos no ambiente digital. Tais objetos caracterizam-se como objetos não estruturados, visto que não possuem padrões em seu formato e, desta maneira, têm apresentados desafios consideráveis às disciplinas e áreas de processamentos e organização da informação, sobretudo as Ciências da Computação, da Informação e da Gestão da Informação.

A produção massiva de tais objetos, tendo como fatores, além do seu volume, também a velocidade e a variedade, têm dificultado o controle informacional no entorno web. Em instituições que armazenam os estoques do conhecimento, como por exemplo as bibliotecas, as formas e metodologias tradicionais de controle, como a organização, representação e recuperação dos objetos de informação – entre elas as linguagens documentárias (LD) ou sistemas de organização do conhecimento (SOC) – mostram-se bem efetivas até os dias de hoje, porém no entorno web tais controles não mostram-se muito efetivos (CHU, 2010).

Nos últimos anos, uma série de iniciativas têm sido efetivadas para o processamento e controle de informações na web, de modo a lidar com os desafios apresentados pela grande quantidade e heterogeneidade dos objetos de informação, principalmente os não estruturados, como as multimídias. Para o controle em escala dos objetos, iniciativas de inteligência artificial têm sido aplicadas ao processamento de imagens e processamentos de textos e linguagem natural. Para tanto, técnicas

como aprendizado de máquina (*Machine Learning*) e aprendizado profundo (*Deep Learning*) vêm sendo amplamente utilizadas para a recuperação de informações.

Uma outra iniciativa de controle da informação na web, de natureza distinta, vem sendo desenvolvida por organismos e instituições diversas, como por exemplo a *World Wide Web Consortium* (W3C) e a *International Standard Organization* (ISO), dentre outras. Esta iniciativa, também reconhecida como aplicação de inteligência artificial (Chu, 2010), é a Web Semântica (WS), concebida inicialmente por Tim Bernes Lee, um dos criadores dos protocolos da web tal como ela se apresenta nos dias de hoje. De acordo com Bernes Lee, *A Web Semântica não é uma Web separada, mas uma extensão da atual, em que as informações recebem um significado bem definido, permitindo que computadores e pessoas trabalhem em cooperação.*” (BERNERS-LEE, 2001, P. 29, tradução nossa).

A concepção da Web Semântica tem como estrutura uma arquitetura que engloba variados aspectos, como a adoção de um identificador único para as entidades dos objetos de informação na web, como o URI (*Uniform Resource Identifier*), além da instituição de metadados tendo como base a linguagem de marcação XML. Nesta concepção, e de modo a aproveitar os aspectos sintáticas e estruturantes destes, uma nova classe de metadados são acrescentados aos objetos de informação, com vistas a dotá-los de inteligência semântica e, assim, potencializar e tornar mais eficiente a recuperabilidade destes. Estes metadados adotam uma arquitetura de triplas referenciais e constituem o padrão de representação denominado de *Resource Description Framework* (RDF), padrão esse que visa representar sujeitos, objetos e uma dada relação entre estes. Para González,

Em 1997 aparece o RDF, linguagem para representar conhecimento na web baseado em XML. RDF não é um formato de arquivo, mas um formalismo de dados desenhado para apoiar a gestão de dados distribuídos em ambiente web, que permite realizar afirmações sobre recursos. Mediante RDF é possível apresentar qualquer tipo de propriedade de um recurso, como título, autor, copyright etc. Junto com o XML, é um dos pilares principais da chamada Web Semântica. A Web Semântica não é só um intercâmbio de dados, mas também uma organização de dados distribuída e descentralizada. (GONZÁLEZ, 2014, P. 80)

Também para a interpretação destes metadados, a WS prevê o uso de robôs ou softwares inteligentes que leem de maneira autônoma tais metadados e fazem inferências de conteúdo, possibilitando uma interoperabilidade de informações eficiente junto aos sistemas de informação preparados para suportar este tipo de tecnologia.

Além do padrão RDF, uma série de vocabulários começaram a ser instituídos e formalizados para a composição de variados critérios descritivos dos objetos de informação. De acordo com o *Ontology Engineering Group* (OEG),

Um vocabulário em *Linked Opendata Vocabulary* reúne definições de um conjunto de classes e propriedades (simplesmente chamadas de termos do vocabulário), úteis para descrever tipos específicos de coisas, ou coisas em um determinado domínio ou indústria, ou coisas em geral, mas para um uso específico. Os termos de vocabulários também fornecem links em dados vinculados, no caso acima, entre uma pessoa e uma cidade. As definições de termos fornecidas pelos vocabulários trazem uma semântica clara para descrições e links, graças à linguagem formal que utilizam (algum dialeto de RDF, como RDFS ou OWL). Em suma, os vocabulários fornecem a cola semântica que permite aos dados se tornarem dados significativos (OEG, 2019, tradução nossa)

Os vocabulários podem, ademais, ser utilizados por domínios semanticamente mapeados, constituindo-se então como ontologias. Sobre as ontologias, Chu afirma que “*podem ser vistas como um vocabulário controlado universal para representação de recursos no web*” (CHU, 2010, P. 277, tradução nossa) e, para González, “*é uma descrição explícita e formal de conceitos em um domínio de discurso, propriedades, atributos e restrições sobre as facetas*” (GONZÁLEZ, 2014, P. 76). As ontologias de domínios, de modo a representar seus objetos e relações, utilizam modelos conceituais com base em vocabulários semânticos para descrever os mais variados aspectos de forma e conteúdo dos objetos de informação. De acordo com a OEG, em 2020 mais de 700 vocabulários foram mapeados como recursos a serem aplicados aos mais diversos contextos descritivos. Com isso, estes vocabulários podem ser considerados metadados de alto nível, ou metadados semânticos, conforme a figura 1:

Figura 1 - Tipos de Metadados



Fonte: Fisher e Sheth (2004)

Vocabulários semânticos têm sido entendidos como uma nova geração de tecnologias de organização e representação de informações no entorno web. De acordo com Martins (2019),

O ambiente informacional da web é vasto, o que tornaria impossível o processo tradicional de descrever e representar os inumeráveis objetos de informação, como feito nas bibliotecas. Neste sentido, a proposta da Web Semântica tem feito esforços no sentido de estender e emular a dinâmica e a eficiência dos tradicionais métodos de organização e representação da

informação – como os vocabulários controlados e linguagens documentárias – ao ambiente web, já agora com auxílio de inteligência artificial e inteligência semântica para dar autonomia e escalabilidade a este processo. (MARTINS, 2019, P. 106)

Para as disciplinas que têm como foco a organização, representação e recuperação da informação, a arquitetura da WS fornece elementos de grande auxílio para a efetivação destas tarefas, embora a abrangência de sua aplicação represente ainda um grande desafio. Conquanto a WS ainda esteja longe de abarcar toda a web, ambientes controlados podem ser vistos como laboratórios de interesse para experimentos e aplicação de suas metodologias e tecnologias semânticas, como, por exemplo, em áreas como a *Enterprise Content Management* (ECM) ou Gestão de Conteúdo Corporativo. Isso posto, o propósito deste artigo visa o estabelecimento de um modelo de ecossistema semântico de informações no âmbito empresarial ou corporativo para o processamento – destilação e enriquecimento semântico – de objetos não estruturados, como os multimídia.

2 METODOLOGIA

O procedimento metodológico utilizado para a instituição do modelo de ecossistema semântico é o *Design Science* – ou *Design Science Research* (DSR) – abordagem utilizada desde os anos 1970 pela área de engenharia para a concepção de artefatos com vistas a melhorias de diversos tipos, principalmente nos aspectos referentes a procedimentos, e logo foi adotada também por outras áreas, como Matemática e a Computação (Rodrigues, 2018). De acordo com Valverde, Toleman e Cater-Steel, “A abordagem *Design Science* tem um histórico de fornecer resultados úteis na avaliação de construções e modelos em sistemas de informação” (VALVERDE, TOLEMAN E CATER-STEEL, 2009, P. 211, tradução nossa). Também Wieringa (2009) sustenta que esta metodologia pode ser utilizada não somente para resolver problemas práticos, mas também problemas de conhecimento.

A abordagem *Design Science* é muito pouco conhecida ou utilizada na Ciência da Informação (Barbosa e Bax, 2017; Rodrigues, 2018). Entretanto, segundo os autores, ela oferece grandes potencialidades de uso, sobretudo no que se refere ao estabelecimento de modelos de gestão da informação, visto que permite a construção de artefatos inovadores a serem aplicados na solução de problemas reais. Para Rodrigues (2018),

O principal elo entre a Ciência da Informação e a DSR está nos Sistemas de Recuperação de Informação. Em um cenário de grande acesso e de fluxos contínuos de informação pela rede, é natural que tais sistemas desempenhem papel fundamental na sociedade. No Campo da Ciência da Informação, não é diferente. Desse modo, é natural que a relação bem estabelecida da DSR com estes sistemas seja explorada e aprofundada. (RODRIGUES, 2018, P. 119)

Também para Santana (2015),

A pesquisa-design, que é a aplicação desse paradigma (*Design Science*) em pesquisa, destina-se a resolver os problemas humanos e organizacionais estendendo suas capacidades criando artefatos inovadores. Artefatos de tecnologia da informação são amplamente definidos como construções (vocabulário e símbolos), modelos (abstrações e representações), métodos (algoritmos e práticas), e instanciações (implementações e os sistemas de protótipos). (SANTANA, 2015, P. 47)

A utilização da abordagem DSR para a instituição de modelos ou artefatos de informação segue o guia instituído por Hevner et. al. (2004), que contempla as seguintes orientações:

Orientação n. 1: Design como um artefato:

Design Science deve produzir um artefato viável na forma de uma construção, um modelo, um método ou uma instanciação.

Orientação n. 2: Relevância do problema:

O objetivo da metodologia *Design Science* é desenvolver soluções baseadas em tecnologia para importantes e relevantes problemas de negócios.

Orientação n. 3: Avaliação do Design:

A utilidade, qualidade e eficácia de um artefato de design devem ser rigorosamente demonstradas através de uma avaliação executada metodicamente.

Orientação n. 4: Contribuições de pesquisa:

A pesquisa eficaz em *Design Science* deve fornecer informações claras e contribuições verificáveis nas áreas do artefato de design, fundações de design e/ou metodologias de design.

Orientação n. 5: Rigor de pesquisa:

A pesquisa com método em *Design Science* baseia-se na aplicação de métodos rigorosos tanto na construção quanto na avaliação de o artefato de design.

Orientação n. 6: Design como processo de busca:

A busca de um artefato eficaz requer a utilização de meios para atingir os fins desejados enquanto satisfaz as leis do ambiente problemático.

Orientação n. 7: Comunicação da pesquisa:

A pesquisa com método em Design Science deve ser apresentada efetivamente tanto para públicos orientados à tecnologia quanto para públicos orientados à gestão.

Para a construção do modelo, segundo estas orientações, foram primeiramente declarados os parâmetros e requisitos necessários ao seu funcionamento ou aplicação. Assim, os seguintes passos foram definidos:

- 1) Declaração dos requisitos de implantação;
- 2) Declaração dos requisitos operacionais;
- 3) Exposição da dinâmica operacional do modelo.

Uma vez instituídos esses passos, foi possível o desenvolvimento do modelo, entendido como um artefato de sistemas de informação, assim também como foi possível estabelecer sua adequação às orientações de Hevner et. al. (2004) para a abordagem DSR.

3 GESTÃO DE CONTEÚDO CORPORATIVO

O uso de sistemas de informações pelo setor corporativo tem sido empregado desde as últimas décadas do século XX para o gerenciamento tanto de documentos físicos quanto digitais. Também variadas técnicas, metodologias e tecnologias vêm sendo utilizadas desde então, consolidando subáreas temáticas especializadas nestas práticas. Com o tempo, surgem áreas especializadas como Gestão da Informação, Gestão de Documentos e Gestão de Conteúdo Corporativo, dentre outras. Segundo Tyrväinen et. al. (2006), a *Enterprise Content Management* (ECM) – ou Gestão de Conteúdo Corporativo (GCC) – integra aspectos destas outras áreas no que se refere ao processamento de conteúdo informacional das organizações. Para Päivärinta e Munkvold,

O gerenciamento de conteúdo corporativo (ECM) integra o gerenciamento de informações estruturadas, semiestruturadas e não estruturadas, código de software incorporado em apresentações de conteúdo e metadados em soluções para produção, armazenamento, publicação e utilização de conteúdo em organizações. (PÄIVÄRINTA E MUNKVOLD, 2005, P. 1)

No entorno organizacional, todos os aspectos relacionados ao ambiente de informação, incluindo processos, pessoas e tecnologias, compõem um ecossistema, tal como apontado por Inmon, Imhoff e Souza (2001). Para estes autores, esse ecossistema informacional corporativo funciona como uma fábrica

ou linha de montagem da informação, constituindo-se no que eles chamam de “Fábrica da Informações Corporativas”, ou *Corporate Information Factory* (CIF). O CIF é composto dos seguintes componentes:

Ambiente externo: constituído de empresas e pessoas que geram e consomem informação.

- Aplicações: constituem os sistemas de informações corporativos em geral.
- Armazém de dados operacionais: conjunto de dados integrados e orientado a assuntos utilizados para decisões de nível tático.
- Camada de transformação e integração de dados: camada do ecossistema corporativo em que os dados adquiridos ou gerados são refinados para a estrutura corporativa.
- Armazém de Dados: constituem os sistemas de armazenamento de dados (Data Warehouse).
- Data Mart: armazéns derivados dos Data Warehouse, geralmente orientados a assuntos ou áreas específicas da empresa, com dados mais refinados que o Data Warehouse.
- Internet e Intranet: principais linhas de comunicação entre o ambiente interno e externo de uma empresa.
- Metadados: catálogo de informação infra-estrutural da CIF.
- Armazém de exploração e mineração de dados: consiste no ambiente de exploração e manipulação de dados.
- Armazém alternativo: consiste na possibilidade de expansão do armazenamento de dados de forma infinita, onde não há preocupação com capacidade de estoque de dados.
- Sistemas de suporte à decisão: constitui no cérebro da CIF, integrando uma série de aplicações centradas no Data Warehouse.

Posteriormente (2016), os autores acrescentaram novos componentes ao CIF:

- Data Lake: ou Lago de Dados, consiste em armazém de dados construídos para suportar dados e documentos massivos e não estruturados, diferentemente do Data Warehouse.
- Data Pond: ou Lagoa de Dados, são subdivisões dos Data Lakes, sendo compostos de Lagoa de dados brutos (*Raw data ponds*), Lagoa de dados analógicos (*Analog data ponds*), Lagoa de dados do aplicativo (*Application data ponds*), Lagoa de dados textuais (*Textual data ponds*) e Lagoa de dados de arquivo (*Archival data ponds*).

No ecossistema de informações corporativo, os objetos multimídia são consideravelmente significativos, tanto quanto em número quanto em importância (Schmitz, 2002), visto que são

documentos de registro de variadas atividades comerciais, como reuniões, contato com clientes, treinamentos e capacitações, material publicitário, dentre inumeráveis outros. No entanto, tais objetos oferecem dificuldades consideráveis para os sistemas de informação, uma vez que a recuperação de seus conteúdos exigem tecnologias complexas. Como são objetos não estruturados, a leitura de seus conteúdos torna-se cada exponencialmente mais difíceis se ocorrerem em grandes quantidades no ecossistema informacional, requerendo, assim, uma interpretação humana ou, também, interpretações por sistemas de inteligência artificial.

Nesse panorama, os metadados passaram a exercer um papel cada vez mais relevante no que se refere à descrição de conteúdos não somente de baixo nível, isto é, descrição de aspectos explícitos nos objetos multimídia, como cores, texturas e formas, dentre outros aspectos similares, mas também de alto-nível, de modo a descrever aspectos implícitos como o contexto abstrato de significados destes objetos. Mesmo com o exponencial crescimento da capacidade computacional dos atuais sistemas de informações corporativos, muitos ainda carecem de alta efetividade no que se refere à representação de conteúdo de multimídia, sobretudo em relação ao emprego de tecnologias semânticas. Segundo Inmon, O'Neil e Fryman, “*infelizmente, a semântica dos sistemas de negócios tem sido amplamente esquecida e ignorada*” (INMON, O'NEIL e, FRYMAN, 2008, P. 199, tradução nossa). Com relação a isso, os autores observam que há muito poucas iniciativas de emprego de tecnologias semânticas para representação de objetos de informação, o que dificulta a eficiência na recuperação destes.

4 METADADOS E VOCABULÁRIOS SEMÂNTICOS

Os metadados são elementos que descrevem dados há décadas, embora no passado tenham desempenhado papéis pouco nobres, como descrever aspectos explícitos de objetos de informação. Entretanto, após o advento da web, sua importância tem sido crescente. De acordo com Fisher e Sheth (2004), no âmbito da Gestão de Conteúdo Corporativo, os metadados podem ser de diversos tipos, porém dividem-se nas seguintes grandes categorias:

- Metadados estruturais: refere-se à estrutura do conteúdo. Em ambiente web, pode-se entender tais metadados como o script HTML, XML ou CSS que formatam e estruturam uma página web;
- Metadados semânticos: são metadados associados implícita ou explicitamente a um conteúdo, cuja relevância é determinada pela sua posição ontológica dentro de um determinado domínio ou aplicação. Tais metadados são blocos de semântica (Fisher e Sheth, 2004), consistindo de anotações e *tags* atribuídos *a posteriori*, de modo a contextualizar o objeto;

- Padrões de Metadados: conquanto também sejam metadados semânticos, tais padrões possuem regras formais em relação aos anteriores, sendo entendidos também como ontologias (W3C, 2019). Estes tipos de metadados oferecem diversos aspectos descritivos para representação do objeto fora dele mesmo, contextualizando-o em relacionando-o aos assuntos tratados e, também, a outros diversos objetos, atores, entidades etc.

Os metadados semânticos e os padrões de metadados são tecnologias oriundas da Web Semântica, consistindo de anotações e vocabulários descritivos formalizados, de maneira a dotar os metadados com inteligência semântica para descrição e representação de objetos de informação. Segundo Martins (2019), “os metadados semânticos e o padrões de metadados são aqueles que transcendem a descrição de aspectos explícitos dos objetos, ou ofness, possuindo capacidade para representar aspectos intrínsecos ao conteúdo, isto é, o aboutness. Metadados semânticos, além de serem formalizados em seu padrão, podem ser auxiliados por vocabulários, glossários ou taxonomias corporativas de modo a descrever valores atribuídos” (MARTINS, 2019, P. 149).

Embora metadados semânticos possam ser entendidos como vocabulários semânticos, eles constituem-se também como Sistemas de Organização do Conhecimento de nova geração, ou *Simple Knowledge Organization Systems* (SKOS) (González, 2014). Como dão suporte à modelagem semântica de domínios do conhecimento, eles funcionam como ontologias. Segundo a W3C (2019b), não existe uma divisão clara sobre as diferenças entre vocabulários semânticos e ontologias, motivo pelo qual eles podem ser entendidos como a mesma coisa. Um dos primeiros padrões de representação semântica é o RDF, que descreve formalmente os objetos de informação de acordo com uma tripla que identifica sujeito, objeto e sua respectiva relação, conforme a figura 2:

Figura 2 - Declara

```
1. <?xml version="1.0"?>
2. <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3. xmlns:externs="http://www.example.org/terms/">
4. <rdf:Description rdf:about="http://www.example.org/index.html">
5. <externs:creation-date>August 16, 1999</externs:creation-date>
6. </rdf:Description>
7. </rdf:RDF> ção RDF em tripla
```

Fonte: W3C, 2004

Com base no padrão RDF, uma série de vocabulários foram instituídos, como o *Dublin Core* (DC), *Friend of a Friends* (FOAF), *Schema*, além de variados outros. De acordo com Martins (2019), o DC

[...] provê descrições tanto de nível mais elementar quanto de relações mais complexas, como relações e cobertura temporal. Além disso, permite resumos e outras informações sobre o recurso ao qual descreve. Frequentemente o DC é utilizado de maneira conjunta com outros vocabulários, enriquecendo o código com informações mais efetivas sobre conteúdos e propriedades técnicas de recursos (MARTINS, 2019, P. 160)

Por sua vez, o FOAF tem sido utilizado para descrever associação entre várias entidades, como pessoas, documentos, coisas, grupos, instituições, dentre numerosos outros aspectos. Dito de outra forma, os vocabulários FOAF funcionam como uma tecnologia de agregação que permite interligar entidades, pessoas, coisas, grupos a objetos e recursos de informação. Eles são um dos vocabulários semânticos “*mais abrangentes e completos, permitindo várias possibilidades declarativas que podem ser potencialmente úteis a ambientes informacionais corporativos*” (MARTINS, 2019, P. 162). Um outro vocabulário potencialmente útil é o *Schema*, que funciona incorporando metadados semânticos em objetos baseados na web. Neste sentido, eles permitem uma forma de estruturação descritiva de forma hierárquica, onde cada descritor define um tipo e incorpora diversas propriedades e possibilidades descritivas, como ações, eventos, pessoas ou personalidade, lugares, tempo, organização, produtos etc.

Os vocabulários semânticos podem ser incorporados aos códigos-fonte dos objetos de informação, em bases de dados, nos próprios objetos, nos diretórios de arquivos digitais e, também, em documentos RDF. Além disso, vários diferentes vocabulários podem ser utilizados simultaneamente na descrição de um objeto de informação, desde que não haja conflito entre as propriedades descritivas. Essa versatilidade no uso de múltiplos vocabulários tende a potencializar consideravelmente a representação e a recuperação de tais objetos.

No entorno corporativo as tecnologias semânticas podem ser empregadas para integrar os metadados ou vocabulários semânticos a contextos e requisitos de negócios e às infraestruturas de apoio. Com isso, é possível interligar os objetos de informação a diferentes elementos do ambiente corporativos, sobretudo àqueles que se relacionam com o ecossistema informacional.

5 DESTILAÇÃO E ENRIQUECIMENTO SEMÂNTICO DE METADADOS

Após o advento da web, objetos multimídia passaram a representar uma parcela significativa dos objetos de informação. No entorno corporativo, igualmente, estes objetos também são consideravelmente numerosos (Schmitz, 2002). Nos anos 1990 começam a proliferar as bases de dados multimídia e, ao

final daquela década, padrões automatizados para descrição do conteúdo de mídias começaram a ser instituídos. Um dos mais conhecidos destes padrões é o MPEG-7, desenvolvido pela *International Standard Organization* (ISO). Este padrão, inicialmente, descrevia conteúdos de baixo nível dos objetos multimídia, tendo como saída o formato de dados com base em XML Schema, um tipo de linguagem de esquemas de descrição – ou *Description Definition Language* (DDL). O MPEG-7 dividiu-se em duas grandes iniciativas, como o Grupo de Áudio e o Grupo de Vídeo. Em ambos os casos, originalmente somente elementos de baixo nível, como formas, texturas e cores eram reconhecidos e descritos nos formatos automáticos de saída, mas ao longo do tempo passaram a incorporar níveis semânticos elementares de descrição de conteúdo.

Para o reconhecimento de conteúdo de alto nível, sistemas de informação com base no padrão MPEG-7 devem interagir com bancos de imagens e sons, além de bases de taxonomias corporativas, efetuando então comparações visando o reconhecimento e extração de classes, entidades, pessoas, sons e demais aspectos de conteúdo. Bancos de dados com imagens e vídeos, além de termos esquematizados em taxonomias, fornecem os devidos parâmetros de comparação, possibilitando, assim, o reconhecimento de elementos no conteúdo dos objetos. Esse processo, conhecido como *destilação* (INMON, 2016), com o tempo passou a utilizar ontologias específicas de multimídia embarcadas nos sistemas. Eles também permitem a inserção de termos descritores livres – ou *tags* – por usuários do sistema, processo denominado *anotação*.

As anotações, entendidas como a adição de termos livres aos metadados destilados dos objetos, caracterizam-se como *metadados informais ou sintáticos*, visto que são adições livres, sem padrões formais. Por outro lado, a inserção de vocabulários semânticos pode ser entendida como inserção de *metadados formais e qualificados*, pois conferem inteligência semântica aos objetos. A adição de vocabulários semânticos nos metadados dos objetos é o que Zeng denomina de anotação semântica ou *enriquecimento semântico*. Brisebois, Abran e Nadembega afirmam que “*a web semântica e os dados vinculados são alguns dos conceitos mais importantes para apoiar o Enriquecimento Semântico de Metadados em um ecossistema de software*” (BRISEBOIS, ABRAN E NADEMBEGA, 2017, P. 2, tradução nossa).

O enriquecimento semântico pressupõe a inserção de ontologias e vocabulários para potencializar os metadados dos objetos, independentemente dos metadados originais destes. Esta inserção pode ser feita através de ferramentas especializadas, tanto de licença proprietária quanto de licença aberta. Gracy (2018) lembra que o enriquecimento semântico de metadados ainda é muito pouco explorado

e destaca a necessidade de se enriquecer objetos de informação – sobretudo multimídia – com um alinhamento de vocabulários, isto é, o uso concomitante de vários vocabulários. Outra vantagem do uso de vocabulários é que eles podem substituir, em muitos casos, o uso de ontologias pesadas de domínio, uma vez que estas são onerosas e especialmente morosas em seu desenvolvimento (Gargouri e Jaziri, 2011).

6 MODELO DE ECOSSISTEMA SEMÂNTICO

O modelo de ecossistema semântico proposto neste trabalho tem como foco os objetos de informação não estruturados, como os multimídia, no ambiente corporativo de empresas. Neste sentido, ele situa-se no escopo da Gestão de Conteúdo Corporativo, pretendendo aplicar inteligência semântica a estes objetos de informação para uma melhor representação e recuperação integrada e interoperável por sistemas de informações compatíveis com esta tecnologia. O modelo proposto tem como escopo geral:

- Âmbito do Gestão de Conteúdo Corporativo, sob a perspectiva do conteúdo;
- Foco no caráter informativo e dinâmico dos objetos para tomada de decisões;
- Utilização da infraestrutura informacional e tecnológica pré-existente;
- Utilização de padrões de vocabulários, metodologias e ferramentas de licença livre;
- Estabelecimento de um modelo de ecossistema semântico com base na web.

Da mesma forma, o modelo não possui como foco os seguintes aspectos:

- Curadoria e custódia de objetos;
- Segurança e preservação dos objetos;
- Aspectos de gestão arquivística de objetos com caráter probatório ou de mero registro;
- Construção de ontologias de domínio ou de aplicação;
- Utilização de ontologias prontas pré-existentes.

Em relação aos requisitos de infraestrutura, os seguintes elementos são necessários para a operacionalização do modelo, conforme o Quadro 1:

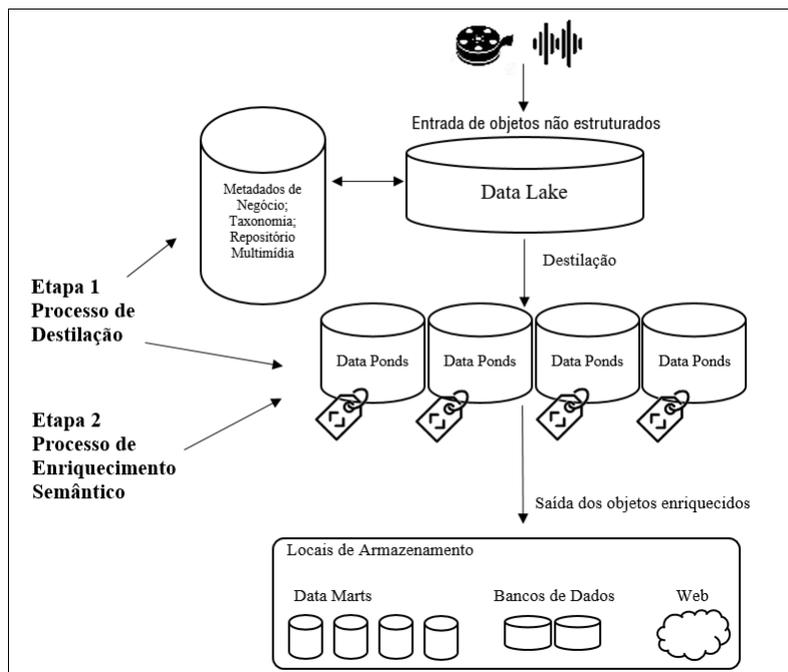
Quadro 1 - Requisitos de infraestrutura do modelo

Requisitos Técnicos de Infraestrutura para Implementação	Aplicação
Sistemas de Armazenamento : <ul style="list-style-type: none"> • Data Lake • BD NoSQL • Data Pond • Data Mart • Data Warehouse • Bases Multimídia • Base de Taxonomia Corporativa 	Armazenamento em várias instâncias dos objetos de informação operacionalizados pelo modelo, como os objetos multimídia
Ferramentas <ul style="list-style-type: none"> • Sistemas de Informação que suportem tecnologias semânticas • Ferramentas de Anotação e Enriquecimento Semântico de Objetos 	Inserção de anotações e vocabulários semânticos aos objetos; As ferramentas devem possuir funcionalidades referentes ao processamento dos padrões exigidos, como o MPEG-7 e padrões de metadados do entorno da Web Semântica

Fonte: O autor

Conforme os requisitos apontados, o modelo requer dois tipos de elementos para seu funcionamento: um relativo aos sistemas de armazenamento dos objetos de informação nos seus variados estágios; e outro relativo às ferramentas de manipulação e processamento dos objetos de informação. Além disso, estas ferramentas devem ser capazes de operar o padrão MPEG-7 e, também, os padrões da Web Semântica. Assim, modelo pode ser representado pelo seguinte esquema, conforme figura 3:

Figura 3 - Esquema geral do modelo proposto



Fonte: O autor

O modelo também pode ser exposto num sentido mais descritivo, conforme o Quadro 2 a seguir:

Quadro 2 - Etapas do processo de destilação e enriquecimento semântico dos objetos de informação

<p>Entrada</p>	<p>Entrada</p> <p>Entrada de objetos no Data Lake ou Lago de Dados: os objetos são armazenados no seu estado bruto, sem qualquer restrição de esquema ou estrutura.</p> <p>Inscrição de Metadados de baixo-nível</p> <p>Metadados automáticos inseridos no objeto: no momento da criação ou da captura dos objetos de informação, metadados de baixo-nível são associados e inseridos automaticamente nos objetos.</p>
<p>Destilação</p>	<p>Reconhecimento de entidades</p> <p>Processo de destilação ou extração de conteúdo dos objetos multimídia: Ferramentas de decomposição de padrão MPEG-7 reconhecem de maneira automática vários recursos inerentes à mídia, representando aspectos diversos nos metadados criados no momento do reconhecimento.</p> <p>Comparação</p> <p>Ferramentas de decomposição de padrão MPEG-7 são executadas; comparações são feitas com os metadados de negócios proveniente de fontes como Repositórios Multimídia ou a Taxonomia Corporativa de modo a reconhecer entidades e outros aspectos para extração de conteúdo.</p> <p>Destilação</p> <p>Reconhecimento de conteúdo: inserção de metadados de baixo-nível dos elementos reconhecidos no objeto; Saída em formato WSDL, XML Schema, dentre outros.</p> <p>Alocação</p> <p>Uma vez realizada a atribuição dos metadados de baixo-nível, o conteúdo é alocado nas lagoas de dados (Data Ponds).</p>
<p>Enriquecimento Semântico</p>	<p>Enriquecimento Semântico</p> <p>Nos Data Ponds serão realizados os procedimentos de enriquecimento semântico, com utilização dos vocabulários destacadas pelo modelo para suportar as descrições e referências semânticas relacionadas ao objeto.</p> <p>O enriquecimento semântico dos objetos é manual, isto é, necessita da participação humana para o processo.</p> <p>Serão inscritos vocabulários que representem atores, processos e outros objetos, juntamente com as anotações dos metadados de negócios</p> <p>Decomposição</p> <p>Partes dos objetos poderão ser decompostas.</p>
<p>Saída</p>	<p>Realocação</p> <p>Uma vez enriquecidos, serão decompostos nos Data Marts, bancos de dados ou na web.</p>

Fonte: O autor

Ademais, o modelo adora alguns parâmetros como elementos essenciais e comuns ao funcionamento de toda empresa, visto que estão presentes em praticamente todas as organizações. Tais elementos são os objetos de informações, os atores e os processos, conforme o Quadro 3:

Quadro 3 - Elementos referenciais do modelo

ELEMENTO	DESCRIÇÃO
Objetos de Informação	Objetos de informação são os elementos básicos não somente do ecossistema informacional corporativo, mas também um dos ativos mais essenciais ao funcionamento de uma organização. Dos insumos de informação saem produtos e serviços para o consumidor, além de vários processos internos e externos à organização.
Atores	Os atores são a razão de ser dos objetos de informação; eles são os produtores, veiculadores, processadores e consumidores de tais objetos. Também são os atores que executam os processos de negócios, mantendo uma organização funcional. Os atores podem ser pessoas, departamentos ou mesmo instituições.
Processos	Os processos são os movimentos padronizados, regulados e dinâmicos que mantêm as organizações em funcionamento. Geralmente são formalizados em procedimentos, etapas ou esquemas de modo a dar objetividade máxima à realização de tarefas. Os processos necessitam de atores e objetos de informação para serem executados.

Fonte: O autor

Em continuação, as relações entre estes elementos e as etapas do modelo são estabelecidas no Quadro 4, a seguir:

Quadro 4 - Relação entre os elementos referenciais e as etapas do modelo

ELEMENTOS	DESTILAÇÃO OU EXTRAÇÃO	ENRIQUECIMENTO SEMÂNTICO
Objetos de Informação	Reconhecimento de elementos de conteúdo de objetos	Representação do objeto para associação a outros objetos, atores, processos e metadados de negócio
Atores	Reconhecimento de atores reconhecidos no conteúdo dos objetos	Representação de atores para associação a outros atores, processos e metadados de negócio
Processos	Reconhecimento de processos reconhecidos no conteúdo dos objetos	Representação de processos para associação a outros processos, objetos, atores e metadados de negócio

Fonte: O autor

De acordo com o modelo proposto, cada um destes elementos é composto de classes, atributos e propriedades, que devem ser interligadas entre si, de modo a formar a teia de relacionamento do ecossistema de informação corporativo. Os vocabulários semânticos, por sua vez, são utilizados por suas propriedades descritivas, visando cada um dos elementos e suas propriedades. Estas classes possuem partes e relações interdependentes no ecossistema de informações corporativo, que podem assim ser definidas:

- Objeto de Informação
 - Objeto Principal (Vídeo e Áudio)
 - Partes Constituintes
- Atores
 - Pessoas
 - Departamentos
 - Organizações
- Processos
 - Processos ou Atividades de Negócio

Estes elementos e suas partes formam o escopo do ecossistema semântico a que o modelo pretende representar. Assim, pretende-se criar uma relação ou teia semântica entre esses elementos, interligando-os mediante relações. Cada um destes elementos tem classes e várias propriedades ou atributos selecionados como requisitos para representação por vocabulários semânticos e anotações. Uma vez realizada a etapa de destilação dos objetos multimídia – seja por reconhecimento automático com uso do padrão MPEG-7 ou anotação por indivíduos – ocorre a etapa de enriquecimento semântico. Assim, as seguintes classes, subclasses e atributos são então processadas por inserção manual de vocabulários semânticos com uso de sistemas apropriados. Os vocabulários que compõem os requisitos descritivos de objetos e relações do modelo apresentam-se nas Quadros abaixo e, em negrito, encontram-se os atributos e elementos obrigatórios de interlace do registro, funcionando como chaves de associação.

Quadro 5 - Atributos selecionados para classe de Vídeo

OBJETOS DE INFORMAÇÃO			
VÍDEO			
Classe/Atributo	Descrição	Sintaxe	Vocabulário
Tem Tipo	Tipo do Objeto	<i>DCMI MovingImage</i>	DC; OA Annotation
Tem Identificador	Identificador único do objeto	<i>Identifier</i>	DC; Schema.org/RDF; OA Annotation
Tem Título	Título do Objeto	<i>Title</i>	DC; OA Annotation
Tem Formato	Formato	<i>HasFormat</i>	DC; OA Annotation
Tem Duração	Tempo de exibição	<i>Duration</i>	DC; OA Annotation
Tem Data	Data de criação	<i>Date</i>	DC; OA Annotation
Tem Lugar	Local do conteúdo	<i>ContentLocation</i>	Schema.org/RDF; OA Annotation
Tem Transcrição	Transcrição do conteúdo	<i>Transcription</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do objeto	<i>Description</i>	DC; OA Annotation
Tem Período de Tempo	Cobre período de tempo	<i>TemporalCoverage</i>	Schema.org/RDF; OA Annotation
Tem Espaço Geográfico	Cobre espaço geográfico	<i>SpatialCoverage</i>	Schema.org/RDF; OA Annotation
Tem Parte	Possui objeto-filho	<i>HasPart</i>	Schema.org/RDF; OA Annotation
Faz Referência A	Referência a outra coisa	<i>Mentions</i>	Schema.org/RDF; OA Annotation
Tem Assunto Principal	Assunto Principal do Objeto	<i>PrimaryTopic</i>	FOAF; OA Annotation
Tem Assunto	Assunto do conteúdo	<i>About</i>	DC; OA Annotation
Tem Criador	Criador do objeto	<i>Creator</i>	DC; OA Annotation
Tem Contribuidor	Co-autores do objeto	<i>Contributor</i>	DC; OA Annotation
Tem Produtor	Produtor do objeto	<i>Producer</i>	DC; OA Annotation
Tem Fonte	Organização do Criador	<i>SourceOrganization</i>	Schema.org/RDF; OA Annotation
Tem Texto Associado	Conteúdo textual	<i>Text</i>	Schema.org/RDF; OA Annotation
Tem Posição	Objeto inteiro ou sequencial	<i>Position</i>	Schema.org/RDF; OA Annotation
Tem Licença	Licença de uso	<i>Rights</i>	DC; OA Annotation
Tem Comentário	Comentários ou anotações	<i>Comment</i>	DC; OA Annotation
Tem Restrições de Acesso	Controle de acesso ao objeto	<i>ConditionsOfAccess</i>	Schema.org/RDF; OA Annotation
Tem Patrocinador	Patrocinador do objeto	<i>Sponsor</i>	Schema.org/RDF; OA Annotation
Tem Pessoas	Pessoas no Objeto	<i>AgentClass – Person</i>	DC; Schema.org/RDF; OA Annotation

Tem Grupo	Grupo ou Departamento	<i>Department</i>	Schema.org/RDF; OA Annotation
Tem Assunto	Assunto ou processo	<i>SubjectOf</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do objeto	<i>Description</i>	Schema.org/RDF; OA Annotation
Tem Palavras-chave	Palavras-chave do Objeto	<i>Description/Keyword</i>	Schema.org/RDF; OA Annotation
Tem Público-alvo	Atores interessados	<i>TargetAudience</i>	Schema.org/RDF; OA Annotation
Tem Tipo de Uso	Material do uso pretendido	<i>LearningResourceType</i>	Schema.org/RDF; OA Annotation
Tem Local de Produção	Local de produção do objeto	<i>RecordedAt</i>	Schema.org/RDF; OA Annotation
Tem Propósito	Ação pretendida pelo objeto	<i>PotentialAction</i>	Schema.org/RDF; OA Annotation
Tem Evento	Se relaciona a evento	<i>Event</i>	Schema.org/RDF; OA Annotation

Fonte: O autor

Quadro 6 - Atributos selecionados para segmento de Vídeo

OBJETOS DE INFORMAÇÃO			
SEGMENTO DE VÍDEO			
Classe/Atributo	Descrição	Sintaxe	Vocabulário
Tem Tipo	Tipo do Objeto	<i>DCMI MovingImage</i>	DC; OA Annotation
Tem Identificador	Identificador único do segmento	<i>Identifier</i>	DC; Schema.org/RDF; OA Annotation
Tem Título	Título do Objeto	<i>Title</i>	DC; OA Annotation
Tem Formato	Formato	<i>HasFormat</i>	DC; OA Annotation
Tem Duração	Tempo de exibição	<i>Duration</i>	DC; OA Annotation
Tem Data	Data de criação	<i>Date</i>	DC; OA Annotation
Tem Data de Início	Data inicial	<i>StartTime</i>	Schema.org/RDF; OA Annotation
Tem Data de Fim	Data final	<i>EndTime</i>	Schema.org/RDF; OA Annotation
Tem Lugar	Local do conteúdo	<i>ContentLocation</i>	Schema.org/RDF; OA Annotation
Tem Transcrição	Transcrição do conteúdo	<i>Transcription</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do objeto	<i>Description</i>	DC; OA Annotation
Tem Período de Tempo	Cobre período de tempo	<i>TemporalCoverage</i>	Schema.org/RDF; OA Annotation
Faz Parte	Possui objeto-pai	<i>IsPartOf</i>	Schema.org/RDF; OA Annotation

Faz Referência A	Referência a outra coisa	<i>Mentions</i>	Schema.org/RDF; OA Annotation
Tem Assunto	Assunto do conteúdo	<i>About</i>	DC; OA Annotation
Tem Criador	Criador do objeto	<i>Creator</i>	DC; OA Annotation
Tem Contribuidor	Co-autores do objeto	<i>Contributor</i>	DC; OA Annotation
Tem Produtor	Produtor do objeto	<i>Producer</i>	DC; OA Annotation
Tem Fonte	Organização do Criador	<i>SourceOrganization</i>	Schema.org/RDF; OA Annotation
Tem Texto Associado	Conteúdo textual	<i>Text</i>	Schema.org/RDF; OA Annotation
Tem Posição	Objeto inteiro ou sequencial	<i>Position</i>	Schema.org/RDF; OA Annotation
Tem Licença	Licença de uso	<i>Rights</i>	DC; OA Annotation
Tem Comentário	Comentários ou anotações	<i>Comment</i>	DC; OA Annotation
Tem Restrições de Acesso	Controle de acesso ao objeto	<i>ConditionsOfAccess</i>	Schema.org/RDF; OA Annotation
Tem Patrocinador	Patrocinador do objeto	<i>Sponsor</i>	Schema.org/RDF; OA Annotation
Tem Pessoas	Pessoas no Objeto	<i>AgentClass – Person</i>	DC; Schema.org/RDF; OA Annotation
Tem Grupo	Grupo ou Departamento	<i>Department</i>	Schema.org/RDF; OA Annotation
Tem Assunto	Assunto ou processo	<i>SubjectOf</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do objeto	<i>Description</i>	Schema.org/RDF; OA Annotation
Tem Palavras-chave	Palavras-chave do Objeto	<i>Description/Keyword</i>	Schema.org/RDF; OA Annotation
Tem Público-alvo	Atores interessados	<i>TargetAudience</i>	Schema.org/RDF; OA Annotation
Tem Tipo de Uso	Material do uso pretendido	<i>LearningResourceType</i>	Schema.org/RDF; OA Annotation
Tem Propósito	Ação pretendida pelo objeto	<i>PotentialAction</i>	Schema.org/RDF; OA Annotation
Tem Evento	Se relaciona a evento	<i>Event</i>	Schema.org/RDF; OA Annotation

Fonte: O autor

Quadro 7 - Atributos selecionados para classe de Áudio

OBJETOS DE INFORMAÇÃO			
ÁUDIO			
Classe/Atributo	Descrição	Sintaxe	Vocabulário
Tem Tipo	Tipo do Objeto	<i>DCMI Sound</i>	DC; OA Annotation
Tem Identificador	Identificador único do objeto	<i>Identifier</i>	DC; Schema.org/RDF; OA Annotation DC
Tem Título	Título do Objeto	<i>Title</i>	DC; OA Annotation
Tem Formato	Formato	<i>HasFormat</i>	DC; OA Annotation
Tem Duração	Tempo de exibição	<i>Duration</i>	DC; OA Annotation
Tem Data	Data de criação	<i>Date</i>	DC; OA Annotation
Tem Lugar	Local do conteúdo	<i>ContentLocation</i>	Schema.org/RDF; OA Annotation
Tem Transcrição	Transcrição do conteúdo	<i>Transcription</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do objeto	<i>Description</i>	DC; OA Annotation
Tem Período de Tempo	Cobre período de tempo	<i>TemporalCoverage</i>	Schema.org/RDF; OA Annotation
Tem Espaço Geográfico	Cobre espaço geográfico	<i>SpatialCoverage</i>	Schema.org/RDF; OA Annotation
Tem Parte	Possui objeto-filho	<i>HasPart</i>	Schema.org/RDF; OA Annotation
É assunto Principal	Assunto Principal do Objeto	<i>PrimaryTopic</i>	FOAF; OA Annotation
Faz Referência A	Referência a outra coisa	<i>Mentions</i>	Schema.org/RDF; OA Annotation
Tem Assunto Principal	Assunto Principal do Objeto	<i>PrimaryTopic</i>	FOAF; OA Annotation
Tem Assunto	Assunto do conteúdo	<i>About</i>	DC; OA Annotation
Tem Criador	Criador do objeto	<i>Creator</i>	DC; OA Annotation
Tem Contribuidor	Co-autores do objeto	<i>Contributor</i>	DC; OA Annotation
Tem Produtor	Produtor do objeto	<i>Producer</i>	DC; OA Annotation
Tem Fonte	Organização do Criador	<i>SourceOrganization</i>	Schema.org/RDF; OA Annotation
Tem Texto Associado	Conteúdo textual	<i>Text</i>	Schema.org/RDF; OA Annotation
Tem Posição	Objeto inteiro ou sequencial	<i>Position</i>	Schema.org/RDF; OA Annotation
Tem Licença	Licença de uso	<i>Rights</i>	DC; OA Annotation
Tem Comentário	Comentários ou anotações	<i>Comment</i>	DC; OA Annotation
Tem Restrições de Acesso	Controle de acesso ao objeto	<i>ConditionsOfAccess</i>	Schema.org/RDF; OA Annotation
Tem Patrocinador	Patrocinador do objeto	<i>Sponsor</i>	Schema.org/RDF; OA Annotation

Tem Pessoas	Pessoas no Objeto	<i>AgentClass – Person</i>	DC; Schema.org/RDF; OA Annotation
Tem Grupo	Grupo ou Departamento	<i>Department</i>	Schema.org/RDF; OA Annotation
Tem Assunto	Assunto ou processo	<i>SubjectOf</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do objeto	<i>Description</i>	Schema.org/RDF; OA Annotation
Tem Palavras-chave	Palavras-chave do Objeto	<i>Description/Keyword</i>	Schema.org/RDF; OA Annotation
Tem Público-alvo	Atores interessados	<i>TargetAudience</i>	Schema.org/RDF; OA Annotation
Tem Tipo de Uso	Material do uso pretendido	<i>LearningResourceType</i>	Schema.org/RDF; OA Annotation
Tem Local de Produção	Local de produção do objeto	<i>RecordedAt</i>	Schema.org/RDF; OA Annotation
Tem Distinções	Desambiguação de sons	<i>DisambiguatingDescription</i>	Schema.org/RDF; OA Annotation
Tem Propósito	Ação pretendida pelo objeto	<i>PotentialAction</i>	Schema.org/RDF; OA Annotation
Tem Evento	Se relaciona a evento	<i>Event</i>	Schema.org/RDF; OA Annotation

Fonte: O autor

Quadro 8 - Atributos selecionados para segmento de Áudio

OBJETOS DE INFORMAÇÃO			
SEGMENTO DE ÁUDIO			
Classe/Atributo	Descrição	Sintaxe	Vocabulário
Tem Tipo	Tipo do Objeto	<i>DCMI Sound</i>	DC; OA Annotation
Tem Identificador	Identificador único do segmento	<i>Identifier</i>	DC; Schema.org/RDF; OA Annotation
Tem Título	Título do Objeto	<i>Title</i>	DC; OA Annotation
Tem Formato	Formato	<i>HasFormat</i>	DC; OA Annotation
Tem Duração	Tempo de exibição	<i>Duration</i>	DC; OA Annotation
Tem Data	Data de criação	<i>Date</i>	DC; OA Annotation
Tem Data de Início	Data inicial	<i>StartTime</i>	Schema.org/RDF; OA Annotation
Tem Data de Fim	Data final	<i>EndTime</i>	Schema.org/RDF; OA Annotation
Tem Lugar	Local do conteúdo	<i>ContentLocation</i>	Schema.org/RDF; OA Annotation
Tem Transcrição	Transcrição do conteúdo	<i>Transcription</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do objeto	<i>Description</i>	DC; OA Annotation

Tem Período de Tempo	Cobre período de tempo	<i>TemporalCoverage</i>	Schema.org/RDF; OA Annotation
Tem Espaço Geográfico	Cobre espaço geográfico	<i>SpatialCoverage</i>	Schema.org/RDF; OA Annotation
Faz Parte	Possui objeto-pai	<i>IsPartOf</i>	Schema.org/RDF; OA Annotation
Faz Referência A	Referência a outra coisa	<i>Mentions</i>	Schema.org/RDF; OA Annotation
Tem Assunto	Assunto do conteúdo	<i>About</i>	DC, OA Annotation
Tem Criador	Criador do objeto	<i>Creator</i>	DC, OA Annotation
Tem Contribuidor	Co-autores do objeto	<i>Contributor</i>	DC, OA Annotation
Tem Produtor	Produtor do objeto	<i>Producer</i>	DC, OA Annotation
Tem Fonte	Organização do Criador	<i>SourceOrganization</i>	Schema.org/RDF; OA Annotation
Tem Texto Associado	Conteúdo textual	<i>Text</i>	Schema.org/RDF; OA Annotation
Tem Posição	Objeto inteiro ou sequencial	<i>Position</i>	Schema.org/RDF; OA Annotation
Tem Licença	Licença de uso	<i>Rights</i>	DC, OA Annotation
Tem Comentário	Comentários ou anotações	<i>Comment</i>	DC, OA Annotation
Tem Restrições de Acesso	Controle de acesso ao objeto	<i>ConditionsOfAccess</i>	Schema.org/RDF; OA Annotation
Tem Patrocinador	Patrocinador do objeto	<i>Sponsor</i>	Schema.org/RDF; OA Annotation
Tem Pessoas	Pessoas no Objeto	<i>AgentClass – Person</i>	DC; Schema.org/RDF; OA Annotation
Tem Grupo	Grupo ou Departamento	<i>Department</i>	Schema.org/RDF; OA Annotation
Tem Assunto	Assunto ou processo	<i>SubjectOf</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do objeto	<i>Description</i>	Schema.org/RDF; OA Annotation
Tem palavras-chave	Palavras-chave do Objeto	<i>Description/Keyword</i>	Schema.org/RDF; OA Annotation
Tem Público-alvo	Atores interessados	<i>TargetAudience</i>	Schema.org/RDF; OA Annotation
Tem Tipo de Uso	Material do uso pretendido	<i>LearningResourceType</i>	Schema.org/RDF; OA Annotation
Tem Local de Produção	Local de produção do objeto	<i>RecordedAt</i>	Schema.org/RDF; OA Annotation
Tem Distinções	Desambiguação de sons	<i>DisambiguatingDescription</i>	Schema.org/RDF; OA Annotation
Tem Propósito	Ação pretendida pelo objeto	<i>PotentialAction</i>	Schema.org/RDF; OA Annotation
Tem Evento	Se relaciona a evento	<i>Event</i>	Schema.org/RDF; OA Annotation

Fonte: O autor

Quadro 9 - Classes e subclasses selecionados para representação de Atores

ATORES			
Classe	Subclasse	Descrição	Vocabulário
Agente			Schema.org/RDF; OA Annotation
	Organização	Subclasse de Agente; Organização ou Instituição	Schema.org/RDF; OA Annotation
	Departamento	Subclasse de Agente; Subclasse de Organização; Departamento ou Área	Schema.org/RDF; OA Annotation
Pessoa		Subclasse de Agente; Subclasse de Organização ou Instituição; Subclasse de Departamento; Indivíduo	Schema.org/RDF; OA Annotation

Fonte: O autor

Quadro 10 - Atributos selecionados para a subclasse Organização

ATORES			
ORGANIZAÇÃO			
Classe/Atributo	Descrição	Sintaxe	Vocabulário
Faz Parte De	Compõe "Agente"	<i>IsPartOf:</i> Agent	Schema.org/RDF; OA Annotation
Tem Identificador	Identificação única da organização	<i>Identifier</i>	DC; Schema.org/RDF; OA Annotation
Tem Nome	Nome da organização	<i>Name</i>	Schema.org/RDF; OA Annotation
Tem departamentos ou Areas	Divisões e departamentos internos	HasPart: Department	Schema.org/RDF; OA Annotation
Tem Pessoas	Funcionários e colaboradores	HasPerson: Pessoas	Schema.org/RDF; OA Annotation
Tem Endereço	Endereço da organização	<i>Location</i>	Schema.org/RDF; OA Annotation
Tem Email	E-mail da organização	<i>Email</i>	Schema.org/RDF; OA Annotation
Tem Página Web	Página web da organização	<i>Url</i>	Schema.org/RDF; OA Annotation
Tem Telefone	Telefone da organização	<i>Telephone</i>	Schema.org/RDF; OA Annotation
Tem Logotipo	Logomarca da organização	<i>Logo</i>	Schema.org/RDF; OA Annotation
Atua em Ramo de Negócio	Setor de atuação	Description	Schema.org/RDF; OA Annotation
Tem Palavras-chave	Palavras-chave da organização	Description/Keyword	Schema.org/RDF; OA Annotation
Tem Publicações	Publicações em mídia impressa ou digital	Publisher	DC; Schema.org/RDF; OA Annotation
Produz Informação	Produz objetos de Informação	Creator	DC, OA Annotation
Possui Processos	Possui processos de negócios	SubjectOf	DC; Schema.org/RDF; OA Annotation

Fonte: O autor

Quadro 11 - Atributos selecionados para a subclasse Departamento

ATORES			
DEPARTAMENTO			
Classe/Atributo	Descrição	Sintaxe	Vocabulário
Faz Parte De	Compõe "Agente"	<i>IsPartOf:</i> <i>Agent</i> <i>Organization</i>	Schema.org/RDF; OA Annotation
Tem Identificador	Identificação única do Departamento	<i>Identifier</i>	DC; Schema.org/RDF; OA Annotation
Tem Nome	Nome do Departamento	<i>Name</i>	Schema.org/RDF; OA Annotation
Tem Pessoas	Funcionários e colaboradores	<i>HasPerson:</i> <i>Pessoas</i>	Schema.org/RDF; OA Annotation
Tem Email	E-mail do Departamento	<i>Email</i>	Schema.org/RDF; OA Annotation
Tem Telefone	Telefone do Departamento	<i>Telephone</i>	Schema.org/RDF; OA Annotation
Atua em Área de Negócio	Área de negócio	<i>SubjectOf</i>	Schema.org/RDF; OA Annotation
Tem Publicações	Publicações em mídia impressa ou digital	<i>Publisher</i>	DC; Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do departamento	<i>Description</i>	Schema.org/RDF; OA Annotation
Tem Palavras-chave	Palavras-chave do departamento	<i>Description/Keyword</i>	Schema.org/RDF; OA Annotation
Produz Informação	Produz objetos de Informação	<i>Creator</i>	DC, OA Annotation

Fonte: O autor

Quadro 12 - Atributos selecionados para a subclasse Pessoa

ATORES			
PESSOA			
Classe/Atributo	Descrição	Sintaxe	Vocabulário
Faz Parte De	Compõe "Agente"	<i>IsPartOf:</i> <i>Agent;</i> <i>Organization;</i> <i>Department</i>	Schema.org/RDF; OA Annotation
Tem Identificador	Identificação única da pessoa	<i>Identifier</i>	DC; Schema.org/RDF; OA Annotation
Tem Nome Principal	Nome da pessoa	<i>FirstName</i>	FOAF; OA Annotation
Tem Sobrenome	SobrenomeLastName		FOAF; OA Annotation
É membro	Pertence a organização e departamento	<i>MemberOf</i>	Schema.org/RDF; OA Annotation

Se relaciona com Pessoas	Funcionários, colaboradores e outros	Knows	FOAF; OA Annotation
Tem Email	E-mail da pessoa	<i>Email</i>	Schema.org/RDF; OA Annotation
Tem Telefone	Telefone da pessoa Departamento	<i>Telephone</i>	Schema.org/RDF; OA Annotation
Tem Criador	Criador do objeto	Creator	DC, OA Annotation
É Colaborador	Colaborador na criação de objetos	Contributor	DC, OA Annotation
É Produtor	Produtor de objetos	Producer	DC, OA Annotation
Atua em Área de Negócio	Atua em projetos e processos	Project	FOAF; OA Annotation
Se relaciona com Área de Negócio	É relacionado a processos ou áreas de negócios	Project	FOAF; OA Annotation
Atua em Projeto ou Processo	Participou de Projeto ou Processo	CurrentProject	FOAF; OA Annotation
Atua em Projeto ou Processo	Participa de Projeto ou Processo	PastProject	FOAF; OA Annotation
Tem capacitação	Conhece atividades, tarefas e processos	KnowsAbout	Schema.org/RDF; OA Annotation
Tem Função	Possui cargo ou função	HasOccupation	Schema.org/RDF; OA Annotation
É assunto de Objetos	Referenciado por objetos	SubjectOf	Schema.org/RDF; OA Annotation
Tem Registro em Vídeo	Tem registro de vídeo	MovingImage	DC; Schema.org/RDF; OA Annotation
Tem Foto	Tem registro de imagem	StillImage	DC; Schema.org/RDF; OA Annotation
Tem Registro em Áudio	Tem registro de som	Sound	DC; Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do ator	Description	Schema.org/RDF; OA Annotation
Tem Palavras-chave	Palavras-chave do ator	Description/Keyword	Schema.org/RDF; OA Annotation

Fonte: O autor

Quadro 13 - Atributos selecionados para a classe Processos

PROCESSOS			
Classe/Atributo	Descrição	Sintaxe	Vocabulário
Tem Identificador	Identificação única do processo	<i>Identifier</i>	DC; Schema.org/RDF; OA Annotation
Compõe uma Organização	Compõe atividades de negócios da organização	Organization	Schema.org/RDF; OA Annotation
Compõe um Departamento	Compõe atividades de negócios de departamento	Department	Schema.org/RDF; OA Annotation

Tem Responsável	Responsáveis pelo processo	<i>Sponsor</i>	Schema.org/RDF; OA Annotation
Tem Colaborador	Possui equipes e colaboradores	<i>Contributor;</i> <i>Employee</i>	Schema.org/RDF; OA Annotation
Tem Pessoas	Pessoas envolvidas	<i>Person;</i> <i>Duns</i>	FOAF; Schema.org/RDF; OA Annotation
Tem Duração	Tempo de exibição	<i>Duration</i>	DC; OA Annotation
Tem Data	Data de criação	<i>Date</i>	DC; OA Annotation
Tem Data de Início	Data inicial	<i>StartTime</i>	Schema.org/RDF; OA Annotation
Tem Data de Fim	Data final	<i>EndTime</i>	Schema.org/RDF; OA Annotation
Se Relaciona a Outro Processo	É composto por outros processos menores	<i>SuperEvent</i>	Schema.org/RDF; OA Annotation
Se Relaciona a Outro Processo	Compõe outro processo maior	<i>SubEvent</i>	Schema.org/RDF; OA Annotation
Há apresentação de trabalhos	Atividades relacionadas ao processo	<i>WorkFeatured</i>	Schema.org/RDF; OA Annotation
Há execução de tarefas	Tarefas relacionadas ao processo	<i>WorkPerormed</i>	Schema.org/RDF; OA Annotation
Se relaciona com Área de Negócio	É relacionado a processos ou áreas de negócios	<i>Project</i>	FOAF; OA Annotation
Tem propósito	Possui função nos negócios	<i>PotentialAction</i>	Schema.org/RDF; OA Annotation
Tem assunto	Se relaciona a assunto ou trabalho	<i>SubjectOf</i>	Schema.org/RDF; OA Annotation
Tem Descrição	Descrição do processo	<i>Description</i>	Schema.org/RDF; OA Annotation
Tem Palavras-chave	Palavras-chave do ator	<i>Description/Keyword</i>	Schema.org/RDF; OA Annotation
Tem Propriedades Adicionais	Possui valores e referências adicionais do processo	<i>AdditionalProperty</i>	Schema.org/RDF; OA Annotation
É composto por objetos de informação	Necessita de objetos para ser efetuado	<i>Text;</i> <i>StillImage;</i> <i>MovingImage;</i> <i>Sound</i>	DC; Schema.org/RDF; OA Annotation
Produz Objetos de Informação	Objetos de informação como produto do processo	<i>Text;</i> <i>StillImage;</i> <i>MovingImage;</i> <i>Sound</i>	DC; Schema.org/RDF; OA Annotation

Fonte: O autor

Quadro 14 - Chaves de relação entre os elementos do ecossistema semântico

ELEMENTO	ATRIBUTOS-CHAVE DE RELACIONAMENTO
Objeto - Objeto	<i>IsPartOf</i> <i>HasPart</i> <i>Description</i> <i>Description/Keyword</i> <i>Mentions</i> <i>PrimaryTopic</i> <i>Transcription</i> <i>Annotation (Business Metadata)</i>
Objeto – Atores; Atores – Objeto	<i>SubjectOf</i> <i>MovingImage</i> <i>StillImage</i> <i>Sound</i> <i>Creator</i> <i>Contributor</i> <i>Producer</i> <i>Description</i> <i>Description/Keyword</i> <i>Mentions</i> <i>PrimaryTopic</i> <i>Transcription</i> <i>Organization</i> <i>Department</i> <i>AgentClass – Person</i> <i>TargetAudience</i> <i>Mentions</i> <i>Annotation (Business Metadata)</i>
Objeto – Processos; Processos – Objeto	<i>MovingImage</i> <i>Sound</i> <i>Description</i> <i>Description/Keyword</i> <i>Mentions</i> <i>PrimaryTopic</i> <i>Transcription</i> <i>Annotation (Business Metadata)</i>
Atores – Atores	<i>IsPartOf</i> <i>Agent</i> <i>Organization</i> <i>Department</i> <i>MemberOf</i> <i>Knows</i> <i>Description</i> <i>Description/Keyword</i> <i>Annotation (Business Metadata)</i>
Atores – Processos; Processos – Atores	<i>Project</i> <i>CurrentProject</i> <i>PastProject</i> <i>KnowsAbout</i> <i>HasOccupation</i> <i>Organization</i> <i>Department</i> <i>Sponsor</i> <i>Contributor;</i> <i>Employee</i> <i>Person;</i> <i>Duns</i> <i>Annotation (Business Metadata)</i>
Processos – Processos	<i>SuperEvent</i> <i>SubEvent</i> <i>WorkFeatured</i> <i>WorkPerformed</i> <i>Project</i> <i>SubjectOf</i> <i>Description</i> <i>Annotation (Business Metadata)</i>

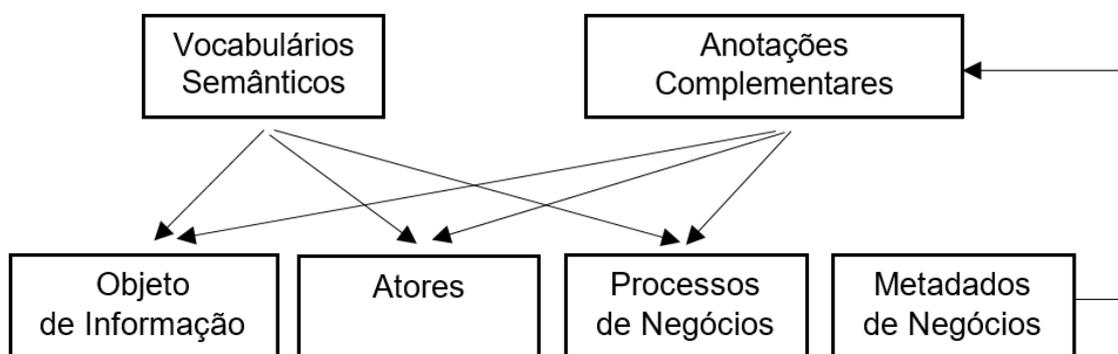
Fonte: O autor

As anotações livres, realizadas manualmente (*Annotation - Business Metadata*), devem ser baseadas nos metadados de negócios, que deverão por sua vez ser padronizados por vocabulários ou taxonomias de modo a uniformizar os termos descritores. Desta maneira, todos os metadados – sintáticos e semânticos – serão padronizados e formalizados, minimizando, assim, inconsistências e incompatibilidades entre metadados e seus elementos descritores. Por seu turno, as linguagens e vocabulários selecionados para representação dos elementos no modelo – classes, subclasses e atributos – foram os seguintes:

- RDF
- Dublin Core
- Schema.org
- FOAF
- W3C - OA Annotation Vocabulary

Uma representação ilustrativa da atuação dos vocabulários semânticos e das anotações baseadas em metadados de negócios pode ser assim concebida, conforme a figura 4:

Figura 4 - Esquema de alimentação do ecossistema semântico



Fonte: O autor

7 CONSIDERAÇÕES FINAIS

O modelo proposto visa o desenvolvimento de soluções no uso de tecnologias semânticas para o tratamento dos problemas levantados, mediante processamento de materiais multimídia, notadamente não estruturados, e que constituem grande parte do ativo informacional das organizações

contemporâneas. Como um modelo conceitual, ele é recomendatório e pretende-se como aplicável a casos típicos do ambiente informacional corporativo de maneira simulada e simplificada em relação aos cenários reais exigidos. Neste sentido, as seguintes inferências podem ser feitas, com base em reflexões lógicas e causais, sobre as possibilidades percebidas do modelo:

- Os objetos multimídia podem ser decompostos em estrutura e conteúdo;
- A operacionalização do modelo pode ser realizada em duas etapas, a fase de destilação de objetos e a fase de enriquecimento semântico destes objetos;
- No que diz respeito ao processo de enriquecimento semântico, este possui uma frente formal, com utilização de vocabulários semânticos formalizados, e uma frente informal, com utilização de metadados de negócios livres à nomeação de cada organização, embora seja necessária uma uniformização por taxonomia;
- Metadados de negócios uniformizados por taxonomia são elementos valiosos no processo de enriquecimento semântico proposto pelo modelo, de modo a facilitar agregações e a potencializar o caráter de interoperabilidade dos objetos e seus segmentos;
- Conquanto o processo de destilação seja automático, é necessária a participação humana para acompanhamento e correção de eventuais erros de interpretação das ferramentas de destilação, bem como adição daqueles metadados que não tenham sido reconhecidos pelas referidas ferramentas;
- O processo de enriquecimento semântico requer a participação humana como uma condição necessária, visto que interpretações de alto-nível obedecem a aspectos pragmáticos pertinentes contextos e interesses, fator que foge aos processos automáticos de interpretação por máquinas;
- A criação de objetos segmentários pode herdar propriedades e atributos do objeto-pai, embora por vezes seja necessária a replicação de atributos especificamente relacionados ao objeto derivado;
- Os vocabulários semânticos possuem um amplo potencial descritivo e, no universo de possibilidades, foram selecionados para o modelo aqueles que se julgou pertinentes à instituição de um ecossistema semântico satisfatório para operacionalização dos objetos e partes constituintes do ecossistema na maioria das organizações, como atores e processos;
- A utilização de vocabulários semânticos requer um ecossistema ambientado na web; ainda que o ecossistema corporativo não forneça informações para o ambiente exterior – a menos que as regras de negócios assim o defina – é possível um retorno de informações mais ricas e relevantes advindas de fontes externas à organização;

- A operacionalização do modelo proposto permite de modo pleno a composição de objetos dinâmicos, montados em tempo de execução de consulta, fornecendo informações customizadas ou personalizados reunindo segmento de vários objetos pertinentes às necessidades informacionais do interessado;
- Os vocabulários semânticos utilizados para declarações de requisitos e atributos de objetos multimídia selecionam aspectos descritivos não exaustivos; desta forma, o modelo proposto foi construído para representar contextos corporativos mais genéricos e universais, não sendo indicado para empresas jornalísticas ou meios de comunicação de massa, uma vez que estas organizações, por operarem de maneira nuclear com material audiovisual, necessitam de padrões descritivos de mídia bem mais complexos ou mesmo ontologias de mídia específicas para suas necessidades; No caso de adaptação a estes tipos de organizações, uma nova seleção de vocabulários semânticos se fará necessária de modo a atender às demandas informacionais do uso de tais objetos;
- O modelo é aberto, isto é, passível de atualização, revisões, reconfigurações e adaptações sempre que necessário; para tanto, checagens de conformidade periódicas impõem-se como necessárias; Tais procedimentos implicam em escolha de novos vocabulários ou novas classes e atributos do vocabulário utilizado.
- Em relação aos benefícios do modelo, os seguintes tópicos puderam ser percebidos:
- Flexibilidade e adaptabilidade aos mais diversos contextos de negócios, mediante a possibilidade de atualização, substituição ou adição de novos vocabulários;
- Baixo custo de implementação, por prescindir de ontologias prontas e aproveitar a infraestrutura local;
- Escalabilidade, podendo ser utilizado em ecossistemas informacionais de diferentes tamanhos;
- Fácil implementação;
- Possibilidade de revisão, atualização e melhorias contínuas.
- Também o modelo possui as seguintes limitações:
- O modelo opera de maneira mais eficiente em organizações cujo ecossistema informacional tenha base na web;
- O modelo, embora criteriosamente descritivo em seus requisitos, etapas e processos, ainda não foi implementado ou testado em cenário real. Cenários reais podem apresentar variáveis e fatores não previstos ou considerados pelo modelo;

- No tratamento de objetos multimídia, o modelo tal como desenhado não contempla os requisitos e necessidades de empresas específicas de mídia ou comunicação de massa, como redes de televisão, rádios e outros ramos de broadcast e cinematografia, que requerem especificações mais detalhadas, mas sim os objetos multimídia mais comuns do ambiente corporativo.

Considerando os fatores acima, entende-se que o modelo, em sua forma estrutural, possui grande capacidade de adaptação e flexibilidade de expansão. Devido à sua estrutura, o modelo pode ser expandido ou mesmo contraído, sem que sua dinâmica de funcionamento perca desempenho, continuando operacional em qualquer dos casos, visto que sua arquitetura permite qualquer escalabilidade e adaptabilidade. Conforme exposto por Chu (2010), o uso de inteligência artificial (IA) para a organização, representação e recuperação da informação têm sido cada vez mais empregadas, incluindo-se as tecnologias semânticas como um aspecto de IA para o processamento inteligente e autônomo da informação. Assim, técnicas híbridas combinadas e inventivas baseadas em tecnologias semânticas, como processamento automático ou assistido da informação, constituem-se em desafios motivadores para as áreas que têm como foco a informação.

REFERÊNCIAS

- BARBOSA, D.; BAX, M. A Design Science como metodologia para a criação de um modelo de Gestão da Informação para o contexto da avaliação de cursos de graduação. **RICI: R. Ibero-amer. Ci. Inf.**, Brasília, v. 10, n. 1, p. 32 -48, jan/jul, 2017.
- BERNERS-LEE, T. **The Semantic Web**. Scientific American, may 2001.
- BRISEBOIS, R.; ABRAN, A.; NADEMBEGA, A. A Semantic Metadata Enrichment Software Ecosystem based on Metadata and Affinity Models. **International Journal of Information Technology and Computer Science**, 8, 1-13, aug 2017.
- CASTELLS, M. **A Sociedade em Rede**. A era da informação: economia, sociedade e cultura. Volume I. 3. ed. rev. e amp.. São Paulo, Editora Paz e Terra, 2016.
- CHU, H. **Information representation and retrieval in the digital age**. 3. ed. Medford, NJ: ASIST Monographic Series, 2010.
- FISHER, M.; SHETH, A. Semantic Enterprise Content Management. In: SINGH, M. P. **The practical handbook of internet computing**. (Computer and Information Science Series). Boca Raton, FL: Chapman & Hall/CRC, 2004.

GARGOURI, F.; JAZIRI, W. *Ontology Theory, Management and Design: An Overview and Future Directions*. IN: **Ontology Theory, Management and Design: Advanced Tools and Models**. Information Science Reference. Hershey NY: IGI Global, 2011.

GONZÁLEZ, J. A. M. **Linguagens Documentárias e vocabulários semânticos para a web**: elementos conceituais. Salvador: EDUFBA, 2014.

GRACY, K. F. Enriching and enhancing moving images with Linked Data: an exploration in the alignment of metadata models. **Journal of Documentation**, v. 74 Issue: 2, pp.354-371, 2018.

HEVNER et. al. Design Science in Information Systems Research. **MIS Quarterly**, vol. 28 n. 1, pp. 75-105, mar, 2004.

INMON, W.; O'NEIL, B.; FRYMAN, **Business Metadata**. Burlington, MA: Elsevier, 2008.

IMHOFF, C.; SOUSA, R. **Corporation information factory**. (3. ed.). New York, NY: Wiley & Sons, 2001.

IMHOFF, C.; SOUSA, R. **Data Lake Architecture**: Designing the Data Lake and Avoiding the Garbage Dump. Basking Ridge NJ, Technics Publications, 2016.

IMHOFF, C.; SOUSA, R.; LINSTEDT, D. **Data Architecture**: a primer for the data scientist. Waltham, MA, Elsevier, 2015.

MARTINS, S. de C. **Modelo conceitual de ecossistema semântico de informações corporativas a ser aplicado em objetos multimídia**. Tese de Doutorado. Universidade Federal Fluminense. Programa de Pós-Graduação em Ciência da Informação da Universidade Federal Fluminense. Niterói: UFF, 2019.

OEG. **LOD - Linked Open Vocabularies**. 2019. <Disponível em: <https://lov.linkeddata.es/dataset/lov/about>>.

RODRIGUES, D. D. Design Science Research como caminho metodológico para disciplinas e projetos de Design da Informação. **InfoDesign: Revista Brasileira de Design da Informação**. v. 15, n. 1, p. 111-124, 2018.

SANTANA, A. F. **BPMG: um modelo conceitual para governança em BPM**. Tese de Doutorado. Universidade Federal de Pernambuco. Programa de Pós-Graduação em Ciência da Computação do Centro de Informática da (CIn/UFPE). Recife, UFPE, 2015.

SCHMITZ, P. Multimedia goes corporate. **IEEE Multimedia**, 9(3), 18–19. 2002.

TYRVÄINEN, P. *et. al.* Characterizing the evolving research on enterprise content management. **European Journal of Information Systems**, v. 15, 627–634, 2006.

SCHMITZ, P.; PÄIVÄRINTA, T. On rethinking organizational document genres for Electronic Document Management. HICSS-32. Abstracts and CD-ROM of Full Papers. **Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences**. 1999.

VALVERDE, R.; TOLEMAN, M.; CATER-STEEL, A. Design Science: A Case Study in Information Systems Re-Engineering. in: CATER-STEEL, A., AL-HAKIM, L. **Information Systems: Research Methods, Epistemology, and Applications**. Hershey, NY: Information Science Reference, 2009.

WIERINGA, R. Design science as nested problem solving, **Proceedings of the 4th int. conf. on design science research in information systems and technology**. ACM, p. 8, 2009.

WIERINGA, R. **Design Science Methodology for Information Systems and Software Engineering**. Berlin: Springer-Verlag, 2014.

W3C. **Best Practice Recipes for Publishing RDF Vocabularies**. W3C Working Group Note 28 August 2008. <Disponível em: <https://www.w3.org/TR/swbp-vocab-pub/>>. Acesso em 15 de abril de 2019.

W3C.. **Ontologies**. W3C Working Group. 2019b. <Disponível em: <https://www.w3.org/standards/semanticweb/ontology>>. Acesso em 15 de abril de 2019.