Knowledge Graph and "Semantization" in Cyberspace: A Study of Contemporary Indexes†

Silvana Drumond Monteiro*, Maria Aparecida Moura**

*Department of Information Science. Universidade Estadual de Londrina (Paraná—Brazil), Rodovia Celso Garcia Cid, n. 445, KM 380, Campus Universitário, CEP 86.057-970, Londrina, PR, Brazil, <silvanadrumond@gmail.com>

**School of Information Science, Universidade Federal de Minas Gerais, Minas Gerais, Brazil <cidamoura@gmail.com>

Silvana Drumond Monteiro holds a Ph.D. in communication and semiotic and postdoctoral degree in information sciences. She is currently an associate professor at the Department of Information Science at the State University of Londrina, Brazil, working on undergraduate and graduate programs. She is also the Coordinator of post-graduation program (*Stricto Sensu*) in information science at the same institution (PPGCI/UEL). Her most recently works focus on representation and knowledge organization in cyberspace and semiotics.

Maria Aparecida Moura holds a degree in library science, a Ph.D. in communication and semiotic and a post doctorate in cognitive semiotics and new media. She is a titular professor from the School of Information Studies of the Federal University of Minas Gerais (ECI/UFMG). She is Director of Information Governance in UFMG. Her range of interests comprehends semiotics and collaborative media, social and theoretical foundations of information and knowledge organization in collaborative environments.

Monteiro, Silvana Drumond, Moura, Maria Aparecida. **Knowledge Graph and "Semantization" in Cyberspace: A Study of Contemporary Indexes.** *Knowledge Organization. 41(6)*, 429-439. 27 references.

Abstract: Knowledge Graph (KG) is a semantics-based search performed by intelligent agents, developed and implemented by Google in December 2012. Though it appears to be just another Google feature, it is actually a huge investment in artificial intelligence in the field of information retrieval. The paper addresses how meaning production occurs in cyberspace, as well as questions about the various stages of web development. Questioning how meaning is made on the web requires us to reframe our understanding of the social, semantic and pragmatic webs, moving towards a ubiquitous web for desktop research. To understand KG, a document analysis was performed, based on an analysis of a Google search engine result page (SERP) and also based on DBpedia and Freebase results. The "semantization" process occurs through the convergence of meaning-making technologies, which Google KG has been implementing, namely: autosuggest, semantic tags, entity collections, geosearch collections, topical weblinks and reference sources. Currently, indexes acquire a more complex meaning. It is in the increasingly intertwined web of sign filigrees, interpreters, interpretants and interpretations that we use to represent the world in cyberspace.

† This research was supported by a grant from Conselho Nacional de Pesquisa (CNPq).

Received: 8 June 2014; Revised: 6 September 2014; Accepted 14 October 2014

Keywords: semantic web, Knowledge Graph, Cyberspace

1.0 Introduction

At the end of 2012, a new tool called Knowledge Graph (KG) was launched by Google. This tool presents a summary representation of how Google arranges information about people, books, movies, etc. On Google's official blog, it is described as "a huge map of real world elements and connections within in order to provide more relevant results" (Google, 2012). We realized a semantic process was operating in KG's search results. As longtime researchers of cyberspace, search engines, and language, we asked, *what is behind KG?* We discovered that machines were involved in interpretation, which immediately suggested an association with Peircean interpreters in contemporary indexes. Hybrid and semiotic processes of interpretation are occurring, in this case, from a machine, but exactly which metadata, algorithms, and intelligent agents are at work? In cyberspace, nothing is simple. This paper addresses the questions above, as well as a more complex question that has been unearthed: about which web are we talking or studying?

The socio-cultural and socio-technical movements began in the 90s with young Americans who took advantage of the creation of computer networks in the 70s and the new visual representations of information in cyberspace by means of the web, created by Tim Berners-Lee, a European, in 1989. The web has become a large digital platform on a global scale for what Lévy (1993) has called "collective intelligence" (cited in Teixeira and Silva, 2013). Meanwhile, in the topography of cyberspace, the web has been the site of major innovations and specializes in visible and invisible semiotic layers. We have conducted studies on the invisible layers, but they have not been conclusive (Monteiro and Fidencio, 2013, Vignoli and Monteiro, 2013). Nevertheless, the specific object of this study is to understand, as much as possible, the environment surrounding KG, namely, the pragmatic and semiotic webs. We hope to understand how the semantic emergences, language games between information and search terms, and dynamics are part of this new index that is at the same time social and machine-made. They are all evidence from which we draw to reflect on information science. The article discusses, in the context of KG, elements of meaning that occur within search results in cyberspace. It is not a specific study on the semantic web, but rather on the processes of "semantization" that occur in cyberspace through web semantics and pragmatic web technologies. By "semantization," we mean large-scale, semantic indexing by Google, using the technologies of semantic web and KG in addition to the terms of research made by users. As in information retrieval, "semantization" means semantically-related terms in the results offered in the index.

Search engines, which in this article will be called "contemporary indexes," can be defined as enormous databases of information about sites on the web. As Battelle (2006) states, indexes are also populated with tags, another type of metadata. Soares (2008) uses the metaphor of postmodern oracles to describe search engines. In current times, according to Fragoso (2007), the search engine or index is the focal point of the online search experience. To this end, we aim to illuminate the context of KG throughout the evolution of the web—the main construct of cyberspace—in which the processes of "semantization" become possible as intelligent agents and the search practices of subjects are operationalized by contemporary indexes.

2.0 Web semantic waves: of which web are we speaking?

New terminology has emerged that characterizes and labels the various stages of the web's technological development. When KG was launched, it seemed to signal the emergence of the "pragmatic" web (Schoop et al., 2006; Singh, 2002), as an analogous dimension of language, which it is, but this concept needs to be further discussed and understood. What did not seem obvious at its inception is that the pragmatic web is a convergence of all other web versions, especially of the semantic web. But there are other disruptive versions, such as Web 4.0, the ubiquitous web, the social and the semantic webs, and the semiotic web. In particular, the emergence of the semiotic web reinforces assumptions about how language functions and how meaning is inferredby intelligent agents in Google's new innovation.

It is impossible to discuss KG's operating procedures without mentioning, even briefly, the evolution of the web's key technological attributes since its creation, because KG has been using the cyberspace platform to aggregate and link information about the real world. Figure 1 summarizes the path of the web's evolution.

According to Mills (2008), the semantic wave in cyberspace is comprised of four stages. The first stage, known as Web 1.0, aims to link and retrieve information in the network. According to Vechiato (2013), its greatest advantages revolve around hypermedia and hypertext. Like the other stages, it is associated with language dimensions, in this case, a syntactic web, with only languages that are machine-readable and understood by humans. To Santaella (2012, 37) "the characteristic verbs of Web 1.0 are: to be available, seek, access and read."

In the next stage, Web 2.0, the focus is to provide a platform to connect people, "putting the *I* in user interface and the *we* in the social participation" (Mills 2008, 3). Following this reasoning, Santaella (2012) notes other characteristic practices and verbs: to expose yourself, to exchange, to collaborate and to interact. This web is for social uses, and by adding a human perspective, the imposition of the context of searchers can initially be associated with a pragmatic dimension, which is not entirely wrong. However, Vechiato (2013) states that a connection with a semantic dimension is needed for "semantization" to emerge in cyberspace, i.e., the next step after Web 3.0,

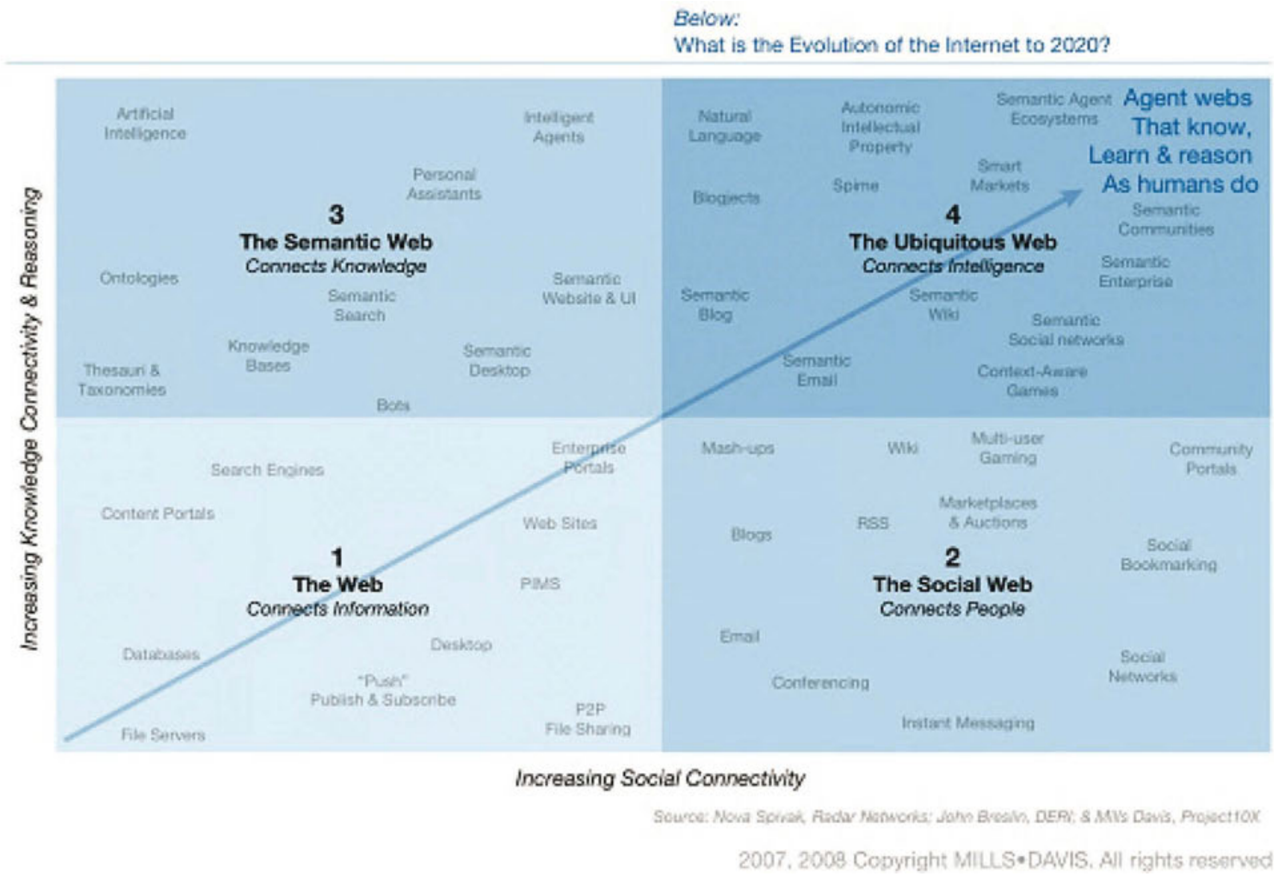

Figure 1. What is the evolution of the Internet to 2020?

perhaps because in Web 2.0, machines do not interpret meanings and human senses.

Nevertheless, digital footprints leave contextual clues (or profiles) about people, things, and concepts, which offer the potential for significant customization services. Increasing the capacity for e-commerce, it enables search engines to provide search suggestions and related topics (through historical searches or encapsulated logs), which is already considered an early semantic-based search. Vechiato (2013) considers Web 1.0 and 2.0 syntactic because resources are connected by hyperlinks, but they are not semantically related.

The next stage, Web 3.0, as shown in Figure 1, is in full swing. In a technological context, "semantics" means building an adequate infrastructure for intelligent agents to roam the web to extract knowledge about something or someone (entities), and perform complex actions. This is to represent meanings, which link knowledge and put them to work in order to make the experience more relevant, useful, and enjoyable in cyberspace because:

> The key notion of semantic technology is to represent meanings and knowledge (e.g., knowledge of something, knowledge about something, and knowledge how to do something, etc.) separately from content or behavior artifacts, in a digital form that both people and machines can access and interpret. As a platform, Web 3.0 will embrace all semantic technologies and open standards that can be applied on top of the current Web. It is not restricted just to current Semantic Web standards. Web 3.0 will encompass a broad range of knowledge representation and reasoning capabilities including pattern detection, deep linguistics, ontology and model based inferencing, analogy and reasoning with uncertainties, conflicts, causality, and values (Mills 2008, 5).

Returning to Santaella (2012, 38), Web 3.0's salient features are:

- the semantic web;
- crowdsourcing or collaborative production mediated by the web;
- more sophisticated social network platforms;
- mobility and cloud computing technologies;
- the web as a space to mediate services.

Following this, the semantic web can be viewed from four different perspectives:

– an advance of Web 2.0;
– metadata technology for business software;
– a social movement for open-source data;
– a new generation of artificial intelligence.

Web 4.0 represents the emergence of digital semiotics; it connects intelligences in the ubiquitous web. According to Mills (2008), the ubiquitous web will, in the near future, unite technologies and services, enabling the growth of social connectivity with the social web (or 2.0) and the growth of knowledge and reasoning of the semantic web (or 3.0). If the current stage connects to semantics, the next will connect to intelligence. Thus, smart agents, artificial intelligence, semantic search, ontologies, and knowledge bases (among others) will prepare ground to establish meaning, not only by humans, but also by machines. Therefore, the semiotic web will emerge as well as a correct use of the term "pragmatic web," as recommended in scientific literature.

Note that KG meets the criteria listed above, i.e., it is an intelligent agent that searches metadata and algorithms: structured search and linked knowledge (Resource Description Frame - RDF - and DBpedia linked data). It recognizes entities and queries (Freebase) knowledge reference databases (CIA World Factbook, Wikipedia) and therefore became the object of our research, simply because as a quasi-interpreter, KG popularizes some postulations about a machine's ability to interpret, not just read, meanings that populate cyberspace.

Mills (2008) also describes the concepts of weak and strong semantics, as shown in Figure 2:

Although not all instruments are unique to the semantic web, it is interesting to observe in the semantic context, the arrangement of language organization and knowledge representation in cyberspace. Almeida and Souza (2011, 40) consider "the strong and weak semantic are characterizations of the more or less expressiveness of the instrument, which makes it possible to better represent reality, considering the operation of an automated system." Almeida, Souza and Fonseca (2011, 196), who are situated in the field of knowledge organization, state, "the instruments presented in spectrum are distinct and varied including thesauri, data-base schemas, modeling languages, and Web-oriented declarative language." They propose separating them into groups, according to their use (see Figure 3):

It is worth noting that the semantic web controls the meaning process that occurs in cyberspace. Our view is that cyberspace is a setting for knowledge organization, in which practices of information retrieval are added to new technologies, intelligent agents, and new languages, for indexing and presentation of indexes. Perhaps we are at the beginning of the emergence of a semantic phenomenon, through the operations of Resource Description Framework (RDF), Ontology Web Language (OWL) and First-Order Logic (FOL). A finite set of axioms appear in search results, but the metadata growth process, knowledge context and representation can lead to higher-order logic, i.e., strong semantics (Mills, 2008, 5). From the semiotic perspective, Pietarinen (2003) argues that since strong semantics go beyond referring to a direct translation of a sign, signal or data (something the semantic web

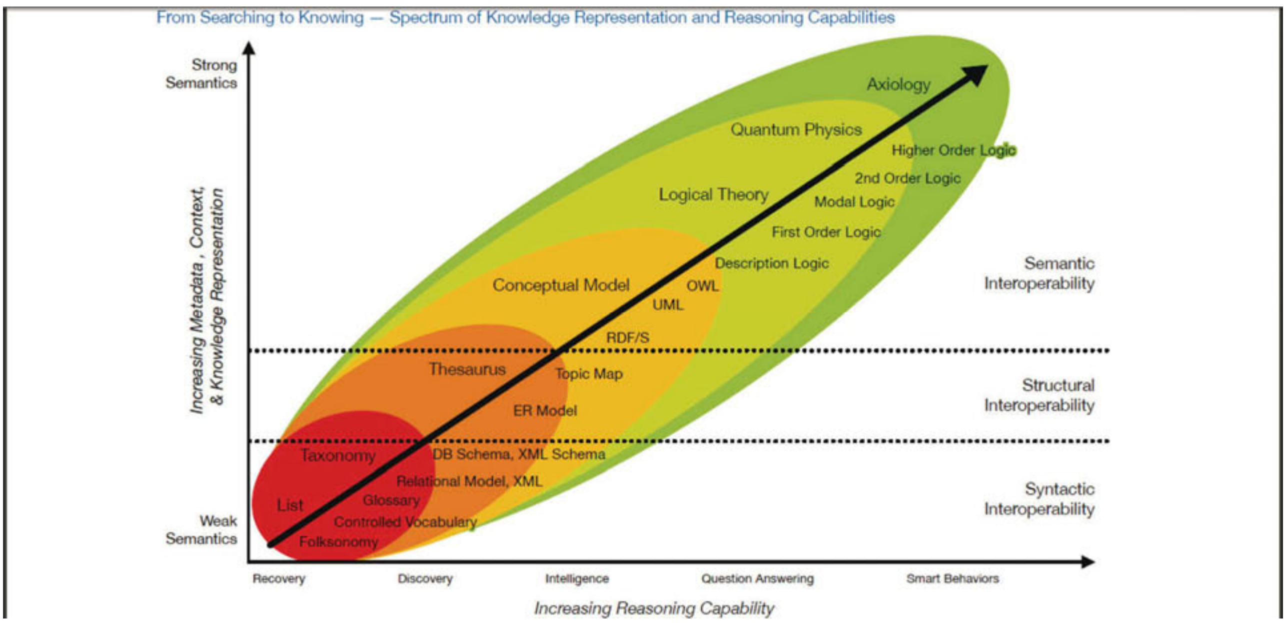

Figure 2. From searching to knowledge representation and reasoning capabilities (Mills and Davis 2008, 5) Retrieved from https://cs.uwaterloo.ca/~j55wu/pub/swwave2008.pdf

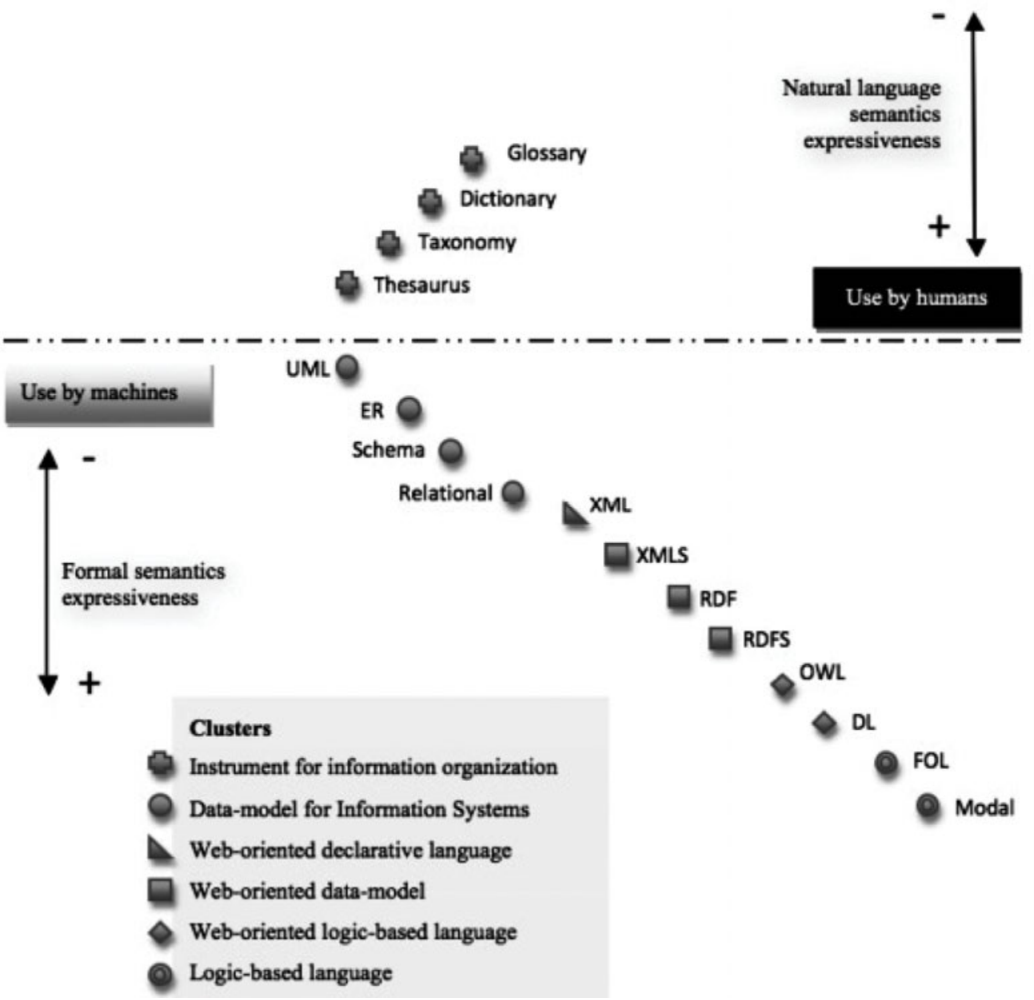

Figure 3. Proposal for a new spectrum

has bothered with) it should also incorporate interpreters' contexts and meanings. Enunciators, interpreters and interpretants should be in a dialogical relationship in which intelligent agents are particularly important:

> From the perspective of multi-agent systems ... the challenge involves an attempt to build advanced agents, which are designed to reproduce the various roles of quasi-semiotic enunciators and quasi-interpreters ("almost" because they are applicable to artificial systems).

Studies have pointed out the possibility of improvement in the indexing of web resources through the contents stemmed from other information sources, such as the investigation of natural language use and processing, attributed by users to digital resources (Castro and Santos 2007, 26). In this sense, Vaidhyanathan (2011, 74) states that Google embodies a pragmatic theory in the strict sense:

> Theoretically, the way Google relies on the collective assessment of the millions of users of the network seems to realize one of the most influential theories of epistemology: the American pragmatism. As Charles Sanders Peirce and William James formulated in the 1890s ... the truth is generated by a process of experimentation, discovery, feedback and consensus.

In the semiotic web, shared meanings and knowledge are human- and machine-made constructs. At the end of his report, Mills (2008, 14) refers to it as the social semantic web 4, in which:

– content is generated by users;
– humans are in synergy with machines;
– it increasingly returns to scale, and
– knowledge emerges.

From the contemporary, symbolic context that cyberspace provides, language dimensions return to the study of signs and to the object in question, i.e. knowledge and information representation and organization through KG, which will be analyzed in the next section, through indexical traces in the semantic/semiotic processes of interpretation. For both, some search results were frozen and observed in a search engine results page (SERP) in an attempt to establish and review the process of meaning resulting from a search in a contemporary index.

3.0 On Knowledge Graph: method and discussion of results

Almost nothing has been written on KG in scientific literature for several reasons: first, because it is a commercial enterprise, and second, because it is a recent innovation. In this section, several blogs written by technologists and Google have been reviewed in an attempt to reveal which technologies run behind KG, in order to understand its structure and functioning. Thus, we conducted a document analysis in KG, based on technical and scientific literature, aiming to scrutinize the modus operandi of the search results on a SERP as well as from related sources, DBpedia and Freebase.

KG uses a new approach to engineering information retrieval, as a result of artificial intelligence that aims for machine learning and semantic-based search results. According to one of its creators, KG is a first step toward building the next generation of research on web collective intelligence (Singhal 2012). In the release video, Google product manager Johanna Wright states that "we are in the early phases of moving from being an information engine to becoming a knowledge engine" (Google 2012). At the launch, three goals were presented:

- find the right thing;
- get the best abstract;
- go broader and deeper.

According to Slawski (2013), KG's patent was granted in the U.S. in 2013, with the following data:

> Abstract: Methods, systems, and apparatus, including computer programs encoded on a computer storage medium, for providing knowledge panels with search results. In one aspect, a method includes obtaining search results that are responsive to a received query. A factual entity referenced by the query is identified. Content is identified for display in a knowledge panel for the factual entity. The content includes at least one content item obtained from a first resource and at least one second content item obtained from a second resource different than the first resource. Data is provided that causes the identified search results and the knowledge panel to be presented on a search results page. The knowledge panel presents the identified content in a knowledge panel area that is alongside at least a portion of the search results.

The patent describes different models that can be used for different types of entities and relationships, which

Figure 4. Introduction to search results in Google Knowledge Graph. Retrieved from http://goo.gl/iOlIyD

Figure 5. Google search result DBpedia. Retrieved from http://dbpedia.org/page/Cambridge

include images, result extension, location on the SERP, and the ranking of search subject request, among other elements. For each entity, at minimum, a title, an image, a description and an associated fact are shown. Figure 4 shows KG's release image.

Since then, KG has been incorporated into an overall search on the right corner of the results page, a "knowledge panel" that identifies people, personalities, cities, books, movies, and other nouns. Figure 5 shows the search results for the various meanings of the word "Cambridge": the dictionary meaning, the encyclopedic meaning (city), and university, etc. A mechanism allows a natural language query (search string based on a question or complete sentence) that returns more contextualized results, i.e., act as knowledge seekers (semantic and pragmatic) rather than just keyword matchers (syntax). Exploring the results contained in the KG panel, we started the analysis.

The Figure 5 example shows a result from DBpedia. DBpedia is a multi-domain knowledge database in English that contains descriptions of 3,770,000 things, 2,350,000 million classified in an ontology, including 764,000 persons, 573,000 places (including 387,000 populated), 333,000 creative works (including 112,000 music albums, 72,000 films and 18,000 video games), 192,000 organizations (including 45,000 companies and 42,000 educational institutions), 202,000 species and 5,500 diseases, besides having versions in 111 languages/dialects (http://wiki.dbpedia.org/Datasets). In DBpedia, the encyclopedia entry page is long, but a short abstract can be updated on mobile platforms in several languages, along with brief semantic or descriptive entity markers. DBpedia is a valuable source for geolocation-based applications, as it interconnects with other data sources such as Geonames, Eurostat, U.S. Census Bureau and the CIA World Factbook. The focus is on data binding, including entity description.

The real world entities and connections have been made by means of structured data or linked data. Wikipedia links to other sources as not merely a knowledge source, but a structured database that operates with semantic markup tools, such as Resource Description Framework (RDF) and Query Language (SPARQL) to Query Data. Table 1 contains some examples of a set of data entities proposed by DBpedia. It means the semantic web is not adequate to describe the entity or resource, but rather connects it to a semantic network to increase expressiveness. The patented information retrieval system may include models for different types of entities, which contains placeholders for images, titles, descriptions and associated events, such as people, places, landmarks, films, businesses, games, sports teams, and sports events, and disambiguates between them.

However, one cannot attribute KG results' semantic effort only to open and linked data from DBPedia. In 2010, Google acquired Freebase (Metaweb System), which is configured as a community-curated database of well-known people, places and facts with an absurd number of described entities. It also contains structured data in RDF Java Script Object Notation (JSON) format, which, in print, contains 76 fields.

Class	Examples
City	Cambridge, Berlin, Manchester
Country	Spain, Iceland, South Korea
Politician	George W. Bush, Nicolas Sarkozy, Angela Merkel
Musician	AC/DC, Diana Ross, Röyksopp
Music album	Led Zeppelin III, Like a Virgin, Thriller
Director	Woody Allen, Oliver Stone, Takashi Miike
Film	Pulp Fiction, Hysterical Blindness, Breakfast at Tiffany's
Book	The Lord of the Rings, The Adventures of Tom Sawyer, the Bible
Computer Game	Tetris, World of Warcraft, Sam & Max hit the Road
Technical Standard	HTML, RDF, URI

Table 1. DBpedia categories of described entities Wikipedia. Retrieved from http://wiki.dbpedia.org/Datasets

Database construction consists of organizing and storing a database schema or graph in a field that designates a common name or category, for example, "media." "Type" defines a person, place or thing, in this case, "author"; "property" defines the quality of type, for example, writer. Thus, the below statement would be established within the domain *media*, as follows:

> William Shakespeare is a (type) author, (property) written work (value) Hamlet.

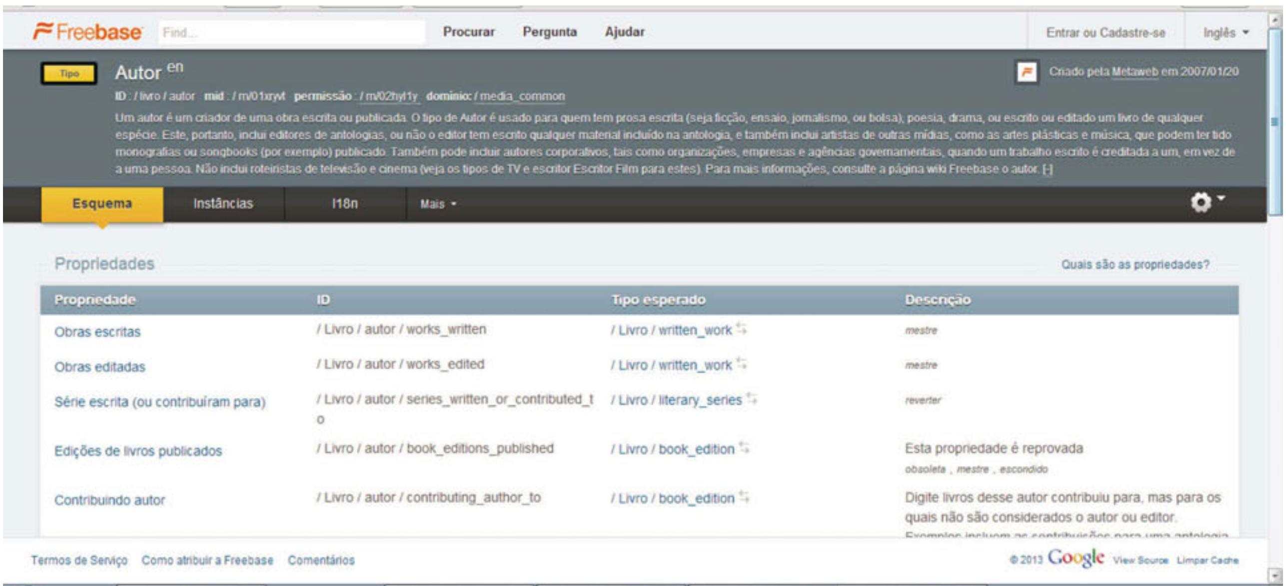

Figure 6. Examples of the Freebase scheme Freebase. Retrieved from http://goo.gl/ocu2J1

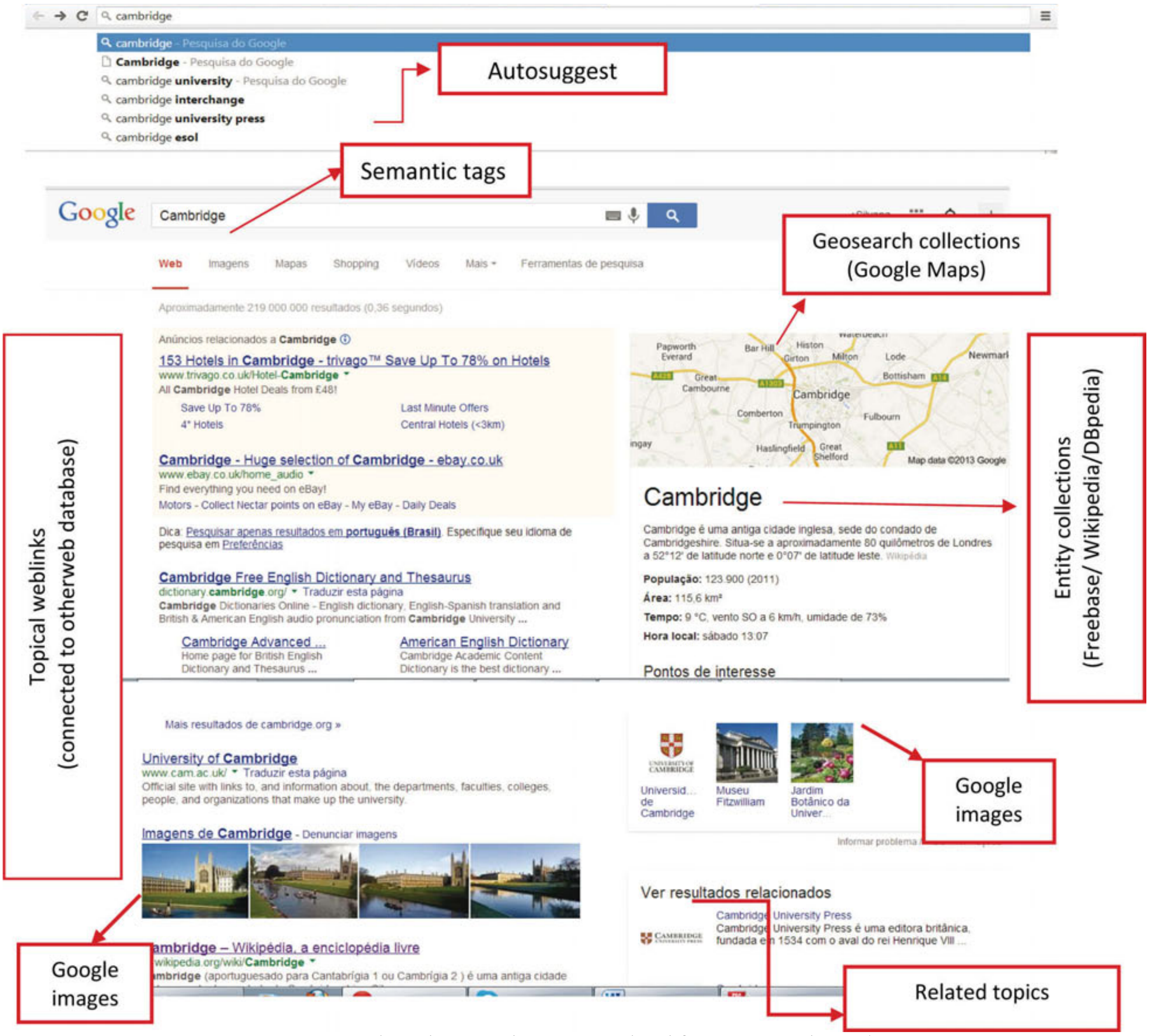

Figure 7. Search result at Google Source: Retrieved from www.google.com.br

Likewise, properties are grouped into types, and types are grouped into domains. Each contains a description, as shown in Figure 6.

Google is building a semantic network, and Freebase especially adds the following semantic and pragmatic possibilities: autosuggest (turning strings into things), semantic tags (adding meaning to content), entity collections (exploring related topics), geosearch collections (exploring locational topics) and topical web links (web links to external internet pages) (Simister 2013). Following Simister's framework, KG's semantic effects are illustrated in Figure 7, including other aspects that appear in results such as topics, related sources, and Google images.

Google has been recording the queries of searchers and mines search logs left by searchers. Thus, Google uses a large amount of corporate information. For example, "see related results" stems from historical searches; at the same time, it offers a disambiguation. By providing related results to the same expression, Google provides index links to homonyms for other entities. For other types of searches, such as for authors, for example, results appear with images from Google Books. There is also, in some cases, arising from Google+ forum discussions, finally, an incredible semantic convergence that includes structured data, entities, and an indexed database captained by KG as an intelligent agent. Even items that appear in KG's space, referring to consulted entities, can change in relation to users' requests. To illustrate, Slawski (2013) uses the example of a search for former U.S. president Abraham Lincoln. In the results, his height is shown because it is an often-requested feature. In searches for other presidents, this information is not included.

The ranking algorithm, pageRank, was developed about 11 years ago, based on a metric that is updated by entities (relationships), authorRank, agentRank, and graphics, so they can be designated as an entity search engine (Fiorelli 2013). In other words, from the indexing of texts, this index evolved toward semantic representations in complex encyclopedic form. So far, we can deduce that semantically-modeled, machine-executable knowledge allows us to connect information about people, events, places, time, among others, from different content sources and various application processes of the sematic web. Note, depending on the query, that other Google sources can be activated, such as YouTube, Google+, and other sources that could be categorized into topical weblinks.

It is important to note that the panel containing KG results is a document generated from a description and interconnection of the searched object with the respective sources. For each entity, subject, and search pattern, there are possibilities to increase the network of interpretant signs. To this extent, searching and indexing is to generate interpretant signs. Analyzing KG allows us to affirm that meaning processes occur from logical data analysis (sematic web) and from a human perspective and understanding (pragmatic web). Of course, all complex semantic networks are marked by human-machine hybridization: a symbolic universe of Google *corpora*.

4.0 Closing remarks

It is the responsibility of researchers to ask how language functions in contemporary indexes in cyberspace. In this sense, it seems KG has discovered how to connect metadata while humans capture the pragmatic dimension in the indexes. It was the task of this article to present some reflections on contemporary indexes and in particular on Google's intelligent agent KG. To Mostafa (2013), "The index is more akin to tables because to index is to relate." Therefore, we have designated search engines as contemporary indexes. While it is assumed that indexing is a means of organizing content in cyberspace, in the context of information science, it is also assumed that these indexes are of another order or degree: that of mechanical and human hybrid assemblages, typical of the current state of technology.

An index for information science is an instrument of both organization and information retrieval. If the indexes are historically defined as the "relations of words or phrases, arranged according to certain criteria, which locates and refers to information contained in a text" (ABNT 2004, 1), now they can additionally be defined as "sign that has relative temporal or spatial contiguity with the object to which it refers" (Infopédia 2013), i.e., it acquires a semiotic meaning, even if it is under an information arrangement. These issues are challenging, particularly for information science: to understand contemporary indexes that are strongly marked by semantic technologies and the experience or the patterns of searchers. Are we prepared for social and machine-based indexing? From the ontological point of view, have we considered the phenomena of hybrid meaning-making processes? These are the questions that make it a challenge to understand contemporary digital objects.

References

Almeida, Maurício Barcellos and Souza, Renato Rocha. 2011. "Avaliação Do Espectro Semântico De Instrumentos Para Organização Da Informação." *Encontros Bibli: Revista Eletrônica de Biblioteconomia e. Ciência da Informação*, Florianópolis, 16 n. 31: 25-50.

Almeida, Maurício Barcellos, Souza, Renato Rocha and Fonseca, Fred. 2011. Semantics in the Semantic Web: a critical evaluation. *Knowledge organization* 38: 187-203.

Associação Brasileira de Normas Técnicas (ABNT). 2004. *NBR 6034 - Preparação de índice de publicações*. Rio de Janeiro: Author.

Batelle, John. 2006. *A Busca: Como O Google E Seus Competidores Reinventaram Os Negócios E Estão Transformando Nossas Vidas*. Rio de Janeiro: Elsevier.

Castro, Fabiano Ferreira de and Plácida Leopoldina Ventura Amorim da Costa Santos. 2007. *Catalogação E Metadados: Interlocuções Nos Ambientes Informacionais Digitais*. Accessed September 26, 2014. http://iibi.unam.mx/publicaciones/235/iv_encuentro_catalogacion_19_ferreira_de_castro_fabiano.html.

Fragoso, Suely. 2007. Quem Procura Acha? O Impacto Dos Buscadores Sobre O Modelo Distributivo Da Word Wide Web. *Eptic: revista de economia política de las tecnologias da la información y comunicación* 9, n.3, sept/dic. 2007. Accessed March 11, 2014. http://ufrgs.academia.edu/SuelyFragoso.

Freebase. 2013. Accessed October 08, 2013. http://www.freebase.com/.

Fiorelli, Gianluca. 2013. *Search in the knowledge era*. Accessed November 11, 2013. http://www.stateofdigital.com/search-in-the-knowledge-graph-era/.

Google. 2012. *The knowledge graph*. Accessed December 20, 2012. http://www.google.com/insidesearch/features/search/knowledge.html.

Infopédia. 2013. Accessed July 08, 2013. http://www.infopedia.pt/lingua-portuguesa/%C3%ADndice.

Lévy, Pierre. 1993. *As tecnologias da inteligência*. São Paulo: 34 Press.

Mills, Davis. 2008. *Semantic wave 2008 report: industry roadmap to web 3.0 & multibillion dollar market opportunities: Project*

10X's. Accessed September 08, 2013. https://cs.uwater loo.ca/~j55wu/pub/swwave2008.pdf.

Monteiro, Silvana Drumond and Fidencio, Marcos Vinicius. 2013. As dobras semióticas do ciberespaço: da web visível à invisível. *Transinformação*, Campinas 25, n.1: 35-46. Accessed October 02, 2013. http://periodicos.puc-campinas.edu.br/seer/index.php/trans info/article/view/1786/1702.

Mostafa, Solange Puntel. 2013. *Filosofia da diferença e a ciência da informação.* Rio de Janeiro: e-papers.

Pietarinen, Ahti-Veikko. 2003. The semantic + Pragmatic web = the semiotic web. *International conference www/internet 2003*. Accessed December 28, 2012. http://www.iadis.net/dl/final_uploads/200302C149.pdf.

Santaella, Lúcia. 2012. A Tecnocultura Atual E Suas Tendências Futuras. *Signo y pensamiento*, **n.** 30: 30-43.

Schoop, Mareik, Moor, Aldo de and Dietz, Jan Leonardus Gerardus. 2006. The pragmatic web: a manifesto." *Communications of the ACM* 49, no. 5: 75-76.

Simister, Shawn. 2013. The freebase api: tapping into Google's knowledge graph. In: Google Developers. *Freebase API.* Accessed November 15, 2013. https://developers.google.com/freebase/.

Singh, Munidar P. 2002. *The pragmatic web*: *preliminary thoughts*. Accessed August 20, 2012. http://www.csc.ncsu.edu/faculty/mpsingh/papers/positions/nsf-02.pdf.

Singhal, Amit. 2012. *Knowledge graph***.** Google Official Blog. Accessed January08, 2013. http://googleblog.blogspot.com.br/2012/05/introducing-knowledge-graph-things-not.html.

Slawkis, Bill. 2013. *How Google decides what to know in knowledge graph results.* Accessed November 08, 2013. http://www.seobythesea.com/2013/05/google-knowledge-graph-results/.

Soares, Cláudio. 2008. *a anatomia da busca.* Pontolit. Brazil: Author.

Teixeira, Marcelo Mendonça and da Silva, Marina Helena de Oliveira. 2013. Hiperligações No Ciberespaço: Interatividade, Comunicação E Educação. *Revista temática* 9 n. 10: 1-13. Accessed November 08, 2001. http://www.insite.pro.br/2013/Outubro/hiperligacoes_ciber espaco_interatividade.pdf.

Vaidhyanathan, Siva. 2011. *A Googlelização de tudo* (*e por que devemos nos preocupar).* São Paulo: Cultrix.

Vechiato, Luis Fernando. 2013. *Encontrabilidade: contributo para uma conceituação no campo da ciência da informação.* Tese de Doutorado, Ciência da Informação, Universidade Júlio Mesquita, Marília, Brazil.

Vignoli, Richele Grenge and Monteiro, Silvana Drumond. 2013. *A topografia da dark web e seus não lugares*. Encontro Nacional de Pesquisa em Ciência da Informação, Florianópolis. Accessed December 08, 2013. http://enancib.sites.ufsc.br/index.php/enancib2013/XIVena ncib/paper/viewFile/204/312.

Wikipedia. 2013. *DBpedia.* Accessed September 02, 2013. http://wiki.dbpedia.org/Datasets