Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 July 2018 doi:10.20944/preprints201807.0560.v1

Peer-reviewed version available at Publications 2018, 6, 44; doi:10.3390/publications6040044

Article

The evolution of the concept of Semantic Web in the context of Wikipedia: An exploratory approach to study the collective conceptualization in a digital collaborative environment

Luís Miguel Machado [1]*, Maria Manuel Borges [2] and Renato Rocha Souza [3]

[1] Univercity of Coimbra; luismachado@fe.uc.pt; Tel.: +351-964-152-983
[2] Univercity of Coimbra; mmb@fl.uc.pt
[3] Fundação Getúlio Vargas; renato.souza@fgv.br
* Correspondence author

Abstract: Wikipedia, as "social machine", is a privileged place to observe the collective construction of concepts without central control. Based on Dahlberg's theory of concept, and anchored in the pragmatism of Hjørland - in which the concepts are socially negotiated meanings - the evolution of the concept of Semantic Web (SW) was analyzed in the English version of Wikipedia. An exploratory, descriptive and qualitative study was designed and we identified 26 different definitions (between 7-12-2001 and 12/31/2017) of which 8 are of particular relevance for their duration, with the latter being the two recorded at the end of the analyzed period. According to them, SW: "is a extension of the web"and "is a Web of Data"; the latter, used as a complementary definition, links to Berners-Lee's publications. In Wikipedia, the evolution of the SW concept appears to be based on the search for the use of non-technical vocabulary and the control of authority carried out by the debate. As a space for collective bargaining of meanings, the Wikipedia study may bring relevant contributions to a community's understanding of a particular concept and how it evolves over time.

Keywords: Semantic Web; Wikipedia; conceptual evolution; negotiated meanings.

1. Introduction

Wikipedia can be described as one of the "abstract social machines" advocated by Berners-Lee and Fischetti [1], in processes enabled by the World Wide Web (WWW) where people do the creative work and the machines do the administrative counterparts. The concept includes the software and systems framework that supports it, as well as the rules, policies and organizational structure governing the participation of the actors in the same "machine" [2]. In the case of Wikipedia, the massive number of collaborators (more than 32 million registered users [1]) contributes to the hypothesis that it is the most comprehensive project in the scope of Digital Humanities [3]. Its dynamics makes it used as a field for investigation of the interaction between humans and computational artifacts under several foci, such as sociological [4,5], informational [6,7] or educational. In this last perspective, we highlight the work done by Cress and Kimmerle [8], who developed a theoretical framework based on the combination of Luhmann's social systems theory with Piaget's theory of equilibrium to study the construction of collaborative knowledge using Wikis. This model was refined and tested in later work [9-13]. Besides this, other approaches were made to the study of *collective knowledge construction* or *collaborative learning* [10,14-17]. In another perspective,

[1] In addition to registered users (34,115,228 as of 26-07-2018) numerous other unregistered users participate in the Wikipedia, (https://en.wikipedia.org/wiki/Wikipedia:About).

© 2018 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 July 2018 doi:10.20944/preprints201807.0560.v1

Peer-reviewed version available at Publications 2018, 6, 44; doi:10.3390/publications6040044

some comparative studies are also mentioned by the points of contact with the present study: between Wikipedia and other sources [18-21]; between Wikipedia articles on the same topic in different languages [22]; or on the contribution of the editors of Wikipedia [13,23].

Considering that Wikipedia presents itself as a free encyclopedia where any Internet user can edit, it is pertinent to investigate how this *collective construction of knowledge* takes place. Collective knowledge understood in the sense given by Scardamalia and Bereiter, that is, the public knowledge available to be managed and used by others [24]. In this theme, the present work intends to place the focus of the analysis on the evolution of a concept constructed in a collaborative way, represented in the respective entry in Wikipedia.

In the same way that the *collective construction of knowledge* was limited to its observable exteriorization, also the collective construction of a concept will be restricted here to its verbal definition represented in the form of written statements in a collaborative way. Although it can be understood as a reductionist view, we believe that these verbal externalizations are the ways in which a body of people can work through building a concept. We consider that this point of view fits in with Hjørland's pragmatic view that "Concepts are dynamically constructed and collectively negotiated meanings" [25]. With a distinct epistemological position, Dahlberg, in her “referent-oriented, analytical concept theory”, does not consider the influence of the social context in the formation of concepts, but takes it into account when it comes to their organization and representation [26]. In this perspective, Dahlberg's theory of concept approaches the position of Hjørland with respect to the representation of concepts, so that the theory provides a reference for the characterization, categorization and decomposition of concepts [27,28].

The semantic web concept was chosen for analysis because it did not present a stable and consensual definition over time, even in the community directly related to its provenance, the Computer Science field. This condition, verified in previous research [29], confers on the collective construction space a context conducive to the debate and negotiation of different personal perspectives. The existence of the previous study, focused on the statements of the World Wide Web Consortium (W3C) and its director (Berners-Lee) works, allows a comparison between this perspective and that of the editors of the Wikipedia article under analysis.

In this way, we intend to analyse the evolution of the semantic web concept in the English version of Wikipedia, treating this as a context of collective knowledge construction. For this purpose, the objective is to: i) collect the different definitions presented in the "introductory section" of the Semantic Web article, from December / 2001 (date of creation of the article) to December / 2017; (ii) to analyse the definitions collected in relation to the concept in question; iii) to diachronically compare the concepts among each other and between them and the analysis of the same concept based on the publications of Berners-Lee and W3C.

2. Materials and Methods

In order to fulfill the defined objectives, an exploratory / descriptive qualitative study was designed, following an observational / comparative methodology [30]. For the operationalization of the empirical component of the study we chose the English version of Wikipedia, since this is the language used in the W3C and Berners-Lee reference documents on the subject of Semantic Web. Thus, the "history" of the Semantic Web entry was mapped to identify the semantic changes made to the support statement of the respective definition, presented in the "introduction" of the different versions of this Wikipedia article. During the analysis, it was used, whenever deemed necessary, to

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 July 2018 doi:10.20944/preprints201807.0560.v1

Peer-reviewed version available at Publications 2018, 6, 44; doi:10.3390/publications6040044

the "discussion" page[2] in order to obtain contextual information to help clarify the definitions presented.

As analytical technique, the categorization was applied "by collection", i.e., the categories resulted from the analogous and progressive classificatory process performed [31]. Subsequently, a procedure based on "pattern-matching time-series" [32] was used to the content units considered in each category, for diachronic comparison. The conceptual analysis focused on the identification of generic terms and their specifying characteristics, in order to compare the definitions collected [27,28]. In the determination of generic terms, we sought for the non-use of compound terms, for the sake of simplicity.

In situations where the definitions use evaluative terms or contextual interpretation ("discussion" page and descriptions appended to the respective changes), we used the contributions provided by the analysis of the discursive strategies, in particular the predicative, of intensification and of attenuation, as long as they provide indicators on the valuing of characteristics and the attitudes and positions of stakeholders [33].

3. Results

There were 129 changes in the introductory part of the Wikipedia´s entry titled Semantic Web, in which 26 definitions with some degree of semantic difference were identified (the corresponding statements are found in Appendix A). Table 1 presents definitions grouped within each category, according to the respective generic term.

Table 1. Generic terms and respective content units retrieved from the identified definitions.

rf.GT[1]	Generic Terms (GT)	Content Units
	category 1. Main definitions	
a)	vision	(#01) is Tim Berners-Lee's vision of the future of the WWW; (#02) is a vision of the future of the WWW.
b)	project	(#03) is a current project; (#04) is a project underway; (#05) is a project.
c)	evolution	(#06) is an evolution of the current Web; (#08) is an evolution of the WWW.
d)	framework	(#09) is a loosely defined and evolving framework of WWW based technologies; (#10) is a loosely defined and evolving framework.
e)	initiatives	(#12) is a set of loosely-defined and evolving initiatives.
f)	extension	(#13) is an evolving extension of the WWW.
g)	development	(#14) is an evolving development of the WWW.
h)	methods and technologies	(#15) it describes methods and technologies; (#18) is a group of methods and technologies.
i)	web of data	(#20) is a "web of data"; (#22) is a "man-made woven web of data".
j)	roadmap	(#23) is the roadmap of a "man-made woven web of data".
k)	movement	(#24) is a collaborative movement.
l)	extension	(#26) is an extension of the Web.
	category 2. Complementary definitions - sub-category 2.1. Assigned	
m)	manifestation	(#07) is a manifestation of Tim Berners-Lee's vision of the Web.
n)	(something)	(#11) it derives from Tim Berners-Lee's vision of the WWW.

[2] Record of all changes made to the page with the possibility of comparing the different versions, (https://en.wikipedia.org/w/index.php?title=Semantic_Web&action=history).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 July 2018 doi:10.20944/preprints201807.0560.v1

Peer-reviewed version available at Publications 2018, 6, 44; doi:10.3390/publications6040044

o)	term	(#16) is a term coined by Tim Berners-Lee.
p)	web of data	(#19) Tim Berners-Lee defined the Semantic Web as "a web of data"; (#25) the term was coined by Tim Berners-Lee for a web of data.
category 2. Complementary definitions - sub-category 2.2. Common use		
q)	model and technologies	(#17) it is mainly used to describe the W3C's model and technologies.
r)	formats and technologies	(#21) is often used to refer to the formats and technologies that enable it.

[1] rf.GT – references associated to the generic terms (GT). The numerical references of the content units relate to the 26 definitions identified (see Apendix A).

In generic terms list there is an exception for the use of the compound term "web of data", which was considered necessary because of the syncronogenematic nature of the element "of data" [34] and its necessity for the meaning intended with the term in question.

The option for two categories, "main definition" and "complementary definition", was necessary since in some of the versions of the Semantic Web entry two or three definitions coexisted. In these cases, the analysis of their statements revealed two patterns: in one, an assignment of the definition to Berners-Lee (subcategory 2.1.), and, on the other, a relation to the common usage of the term (subcategory 2.2.). Units #01 and #02 (rf.GT a)) were considered within the category 1., despite their close relationship with Berners-Lee, given that these initial versions of the article are the only definitions, as the main definition.

The groupings, by generic term, presented in Table 1 constitute the time series used in the diachronic analysis, as shown in Figure 1.

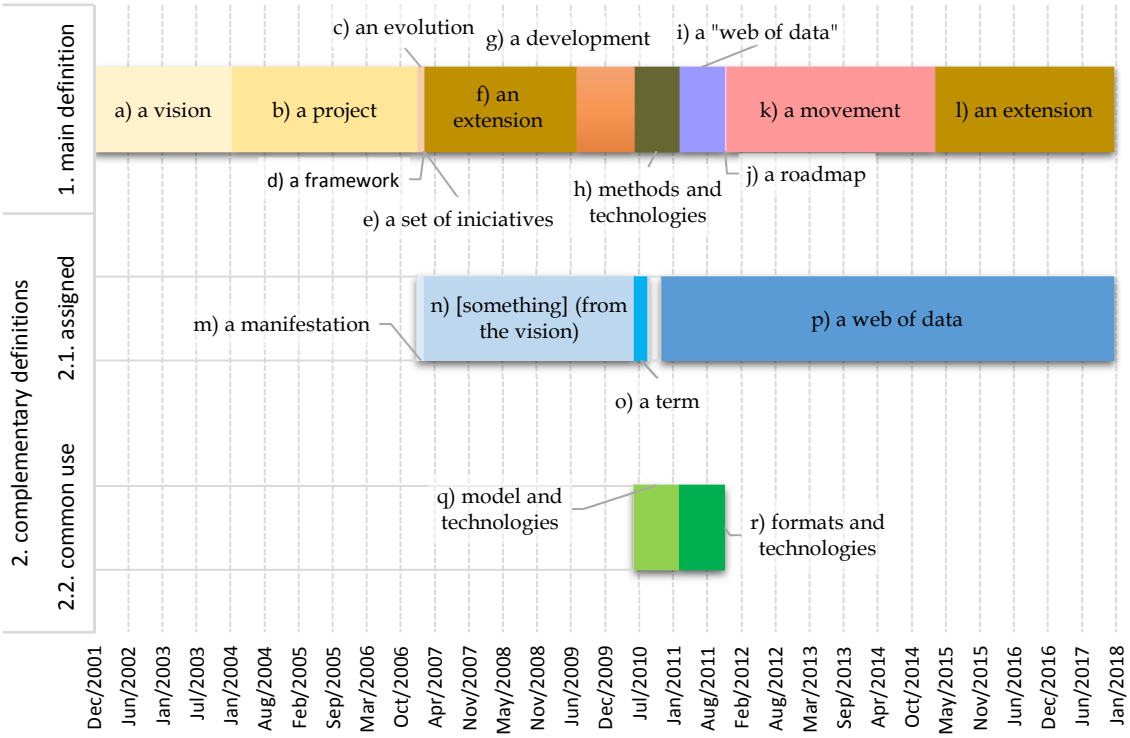

Figure 1. Temporal destribuition of the definitions (group by the respective Generic Terms).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 July 2018 doi:10.20944/preprints201807.0560.v1

Peer-reviewed version available at Publications 2018, 6, 44; doi:10.3390/publications6040044

The diachronic visualization presents an enlightening overview of the evolution of the Semantic Web concept in Wikipedia's context. Given the extended time span (December 2001 to December 2017) it is natural that definitions with little longevity are less noticeable, as is the case with those referred to with d), e) and j), whose duration is less than 10 days.

The analysis of the definitions revealed conceptual variations due to the introduction or alteration of the specific characteristics attributed to the generic term (see Table 2).

Table 2. Specific characteristics of generic terms.

rf.GT	Specifiers pre-GT	Generic Terms (GT)	Specifiers post-GT
a)		(#01; #02) vision	(#01) of Berners-Lee of the future of the WWW; (#02) of the future of the WWW
m); n)		(#07) manifestation; (#11) it (derives from)	(#07) of Berners-Lee's vision of the future of the WWW; (#11) Berners-Lee's vision of the WWW
b)	(#03) a current	(#03; #04; #05) project	(#04) underway
c)		(#06; #08) evolution	(#06) of the current WWW; (#08) of the WWW
d)	(#09; #10) a loosely defined and evolving	(#09; #10) framework	(#09) of WWW based technologies
e)	(#12) a loosely defined and evolving set of	(#12) initiatives	
f); g); l)	(#13; #14) an evolving	(#13; #26) extension; (#14) development	(#13; #14; #26) of the WWW
h)	(#18) a group of	(#15; #18) methods and technologies	
q); r)		(#17) model and technologies; (#21) formats and technologies	(#17) proposed by W3C; (#21) that enable it [the SW]
i)	(#22) a man-made woven	(#20; #22) web of data	
j)		(#23) roadmap	(#23) of a man-made woven web of data
k)	(#24) a collaborative	(#24) movement	

In some cases, the conceptual drift only occurs in the supplements, as is the case in group b) of Table 2, with a single generic term, "project", which includes three variations: first the project is adjectivized as being "current" (#03), then it is "underway" (#04) and, finally, it loses its adjectivation (#05).

In an inverse situation are the supplements that serve as a link between the different generic terms, as occurs in groups c) to g) of Table 2. The variation between the five terms becomes gradual when framed by the specifiers that are maintained or little altered, such as pertaining to WWW membership in these groups. Another example is visible in the change from the term "evolution" (#08) to "framework" (#09 and #10), where the former becomes part of the specifying characteristics

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 July 2018 doi:10.20944/preprints201807.0560.v1

Peer-reviewed version available at Publications 2018, 6, 44; doi:10.3390/publications6040044

of the second, an "evolving framework". This specifier, "evolving", accompanies the following three terms: "set of initiatives" (#12), "extension" (#13), and "development" (#12).

The comparison between the definitions of the semantic web concept, identified in Wikipedia, with those resulting from the analysis of the same concept based on the publications of Berners-Lee and 3WC, was also carried out in a diachronic perspective. For the sake of clarity and representativeness, we have opted to restrict the analysis to variations with a duration of more than 90 days, and not to include the two complementary definitions of common use (subcategory 2.2), since they would only add "noise " to this comparison. Applying these criteria result in eight main definitions and two complementary definitions (Figure 2).

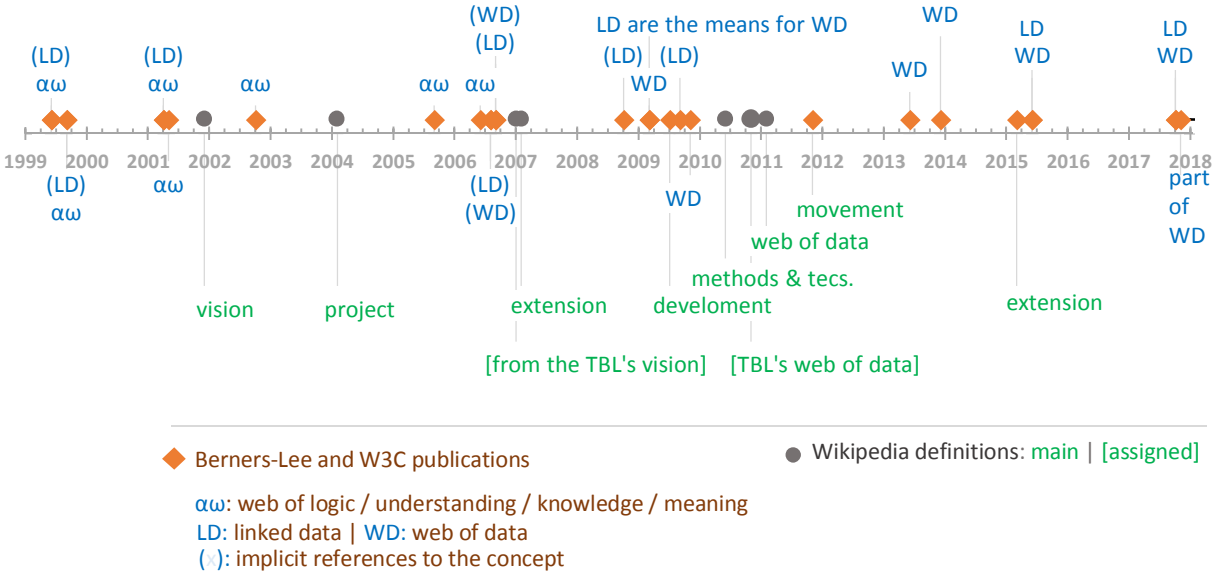

Figure 2. Comparative temporal distribution between the definitions of Semantic Web from the two sources (Wikipedia and publications of Berners-Lee/W3C).

From the observation of the temporal distribution, presented in Figure 2, two situations stand out, being the first related to the variations of the main definition with the generic term "vision" and "project", to coincide with the period in which publications with definitions that have terms like "logic", "understanding", "knowledge" or "meaning" (αω). The second situation concerns to the term "web of data", both in the main definition (in 2011) and in the complementary (in 2010), after this term is used explicitly (in 2009) in the analysed Berners-Lee / W3C publications.

Another potential relation is to verify if we take into account the descriptions present in the Berners-Lee and W3C publications previously analysed. For this matter, we repeat in Table 3 the content units of the cited study [29].

Table 3. Groups and respective content units considered in the analyses of the publications of Berners-Lee and W3C.

Groups	The Semantic Web is…
1. Descriptions that include "semantics"	a. The Web of understanding (1999 Jun.7); A universal web of knowledge (2001 Apr.26).
	b. An extension of the current Web in which information is given well-defined meaning (2001 May; 2002 Oct.); A web of logic (2005 Sep.13); A Web of actionable information derived from data through a semantic theory for interpreting the symbols (2006 Jun.).
	c. A web of data with meaning (1999 sep.22).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 July 2018 doi:10.20944/preprints201807.0560.v1

Peer-reviewed version available at Publications 2018, 6, 44; doi:10.3390/publications6040044

2. Linked Data	a. The Web of linked data (2006 Aug.11; 2015). b. A new data model to support the linking of data from many different models (1999 Jun.7); The web of connections between different forms of data (1999 Sep.22); A world of trusted information shared along collaborating groups of users (2001 Apr.26); An open web of inter-referring resources (2006 Aug.11); A type of extension of the Web to extend the Web to cover linked data (2006 Sep.); A network of data on the Web (2008 Oct.); The world of linked data (2009 Oct.22). c. Linked Data provides the means (2009 Mar.).
3. Web of Data	a. A Web of Data (2009 Mar.; 2009 Nov.12; 2013 Jun.27; 2013 Dec.11; 2015; 2017 Oct.11). b. One extension of the Web moving from text documents to data resources (2006 Aug.11); Is intended to function in the context of the relational model of data (2006 Sep.). c. Part of the Web of Data (2016).

Referring to the Table 3, we can note that the term "extension" is used to define the Semantic Web in two moments. Initially, it appears in two documents (of 2001 and 2002, subgroup 1.b.) very close to the beginning of the article in Wikipedia (December, 2002) and then (August and September, 2006; subgroups 3.b. and 2.b., respectively). The same term was used in the Wikipedia definitions in February, 2007 ("an evolving extension"), very close, though, of the second occurrence in the publications.

Unlike the definition of the Semantic Web as the "Web of Data", verified in the two sources, we did not find in the definitions of Wikipedia mentions that could be understood as the "Web of Linked Data", as it appears explicitly in two publications in table 3, for 2006 and 2015 (sub-group 2.a.).

4. Discussion

The concept of Semantic Web, presented in the respective entry of Wikipedia, shows an evolution that seems to oscillate between the search for a more concrete definition and the use of terms accessible to the common layman. Although the evolution of this concept presents points of contact and similarity between the two scopes (Wikipedia and the publications of Berners-Lee and W3C), the differences detected go beyond that imposed by the type of support, *continuum* in the first and composed by *discrete units* in the second. The present study leads to the conclusion that the search for adaptation to non-specialist readers by Wikipedia editors marks a significant difference between the two scopes. The adaptation referred to above may also give rise to the need for additional definitions, since it is thus possible to present in an integrated form more than one point of view concerning the same concept.

The search for a clearer and more specific definition is, we believe, responsible for the elimination of dubious expressions or buzzwords[3]. In some changes made to the article, this attempt to promote clarification is explicitly stated, as in 11/21/2011, where the segment "that facilitates machines to understand the semantics, or meaning, of information on the World Wide Web" was taken from the definition and classified as "obscure"[4]. Also, in the change from the generic term "project" to "framework", as well as in the change from the latter to "extension", we can identify this

[3] A buzzword is a word or expression that has become fashionable in a particular field and is being used a lot by the media, (https://www.collinsdictionary.com/dictionary/english/buzzword).

[4] Retrieved from *Semantic Web: Revision history*, (dynamic URL).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 July 2018 doi:10.20944/preprints201807.0560.v1

Peer-reviewed version available at Publications 2018, 6, 44; doi:10.3390/publications6040044

double intention of clarification and adapting to non-specialist readers. This belief is reinforced by the debate around this last change (from "framework" to "extension"), shown in the respective discussion page, where it is possible to find, in the editors' debate, the search for the balance between the personal understandings of the given concepts and the adequacy to the general readers. The discussion we are referring to is not the unique example of negotiation processes for the terms to be used in the definitions, detected on the discussion page. On the other hand, there were no occurrences in the history of the Semantic Web entry, of the repeated and systematic alternation between versions, known as "edit wars" [35], as we can see in several entries of Wikipedia.

In fact, regarding the authorship of the changes to the definition presented in the Semantic Web article, they are characterized by debate and diversity. In the 26 definitions registered, there are 16 different users registered and 4 unregistered. In addition, users with more than one definition make their contribution in the same edition and with definitions that fall into different categories; one main and one attributed and / or common use (see Appendix A). The only exception, reported on 20/02/2007, occurred in the context of what could have originated an "edit war" between two editors (Dreftymac and Cygri). However, the debate was transferred to the appropriate channel, the discussion page, where the predominant position of the two editors was the negotiation of a consensus between the two different visions. A negotiation, where the perception of the multiple meanings that the Semantic Web concept can take for different people is present: "we deal with a much-hyped term that is used to mean quite different things by different people." (Cygri, Feb .21, 2007)[5].

In spite of this, the last definition ("is an extension of the WWW") has remained stable for almost three years, in parallel with the definition attributed to Berners-Lee: "The term was coined by Tim Berners-Lee for web of data that can be processed by machines". The scope of this term, "extention", may contribute to the stability of the definition, but does not contribute to a specification of the concept that it intends to define. From this point of view, the Semantic Web concept can be seen as being in a "pseudo-concept" phase which, according to Vygotsky [36], is characterized by an intermediate stage between the general or complex notions and the fully developed concept.

Another issue that may create some kind of restraint in changing the definition is the link (academic and professional) of the author of the last definition to the Semantic Web. However, we are not giving to this influence too much weight because, in Wesch's words: "Authorized information is not beyond discussion on Wikipedia, information is authorized through discussion" [37].

5. Conclusions

Given the characteristics of Wikipedia, described and discussed throughout this paper, we can consider it as a place for collective bargaining of meanings, and it is therefore important to take it as an object of study for a community's understanding of any concept in particular. This position is aligned with Hjørland quote: "Concepts have been understood as socially negotiated meanings that should be identified by studying discourses rather than by studying individual users or a priori principles." [25].

Despite Wikipedia's relevance to this study of the collective construction of meanings, others will be necessary to understand the importance of this same role in a more comprehensive process of *dictionaryization* where the content of a concept is fixed by its definition [38]. It is possible, however, to draw a parallel between the conceptual evolutionary dynamics inherent in the workings of Wikipedia and Derqui's assertion, that says that: "a social system is organized around definitions and redefinitions" [38].

[5] See: https://en.wikipedia.org/wiki/Talk:Semantic_Web.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 July 2018 doi:10.20944/preprints201807.0560.v1

Peer-reviewed version available at Publications 2018, 6, 44; doi:10.3390/publications6040044

Author Contributions: Conceptualization, Luís Miguel Machado; Writing – original draft, Luís Miguel Machado, Maria Manuel Borges and Renato Souza; Writing – review & editing, Luís Miguel Machado, Maria Manuel Borges and Renato Souza.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Definitions extracted from the "Semantic Web: Revision history" inserted in the context units (column *Context units*) with bold emphasis of the content units; the respective authors (column *Users*); the date in which the definition was entered (column *Start*); the date it was withdrawn (column *End*); the reference relative to 129 statements collected (column *ref.Tt.*) and the reference assigned to units of content (column *ref.Df.*), composed of the chronological number, followed by the generic term identifier and the category (main (1.) or complementary definitions (2.1 or 2.2)) where this has been classified (see Table 1).

ref.Df	ref.Tt.	Start	End	Context units	Users
#01.a cat.1	t001	07-12-2001	13-01-2004	*is **Tim Berners-Lee's vision of the future of the WWW**.*	The Anome
#02.a cat.1	t002	13-01-2002	10-02-2004	*is **a vision of the future of the WWW** proposed by Tim Berners-Lee,*	65.2.226.xxx
#03.b cat.1	t008	10-02-2004	23-07-2004	*is **a current project** under the direction of Tim Berners-Lee of the W3C to extend the ability of the WWW.*	ShaunMacPherson
#04.b cat.1	t012	23-07-2004	26-07-2004	*is **a project underway** that intends to create a universal medium for the exchange of information by giving meaning, in a manner understandable by machines, to the content of documents on the WWW.*	Mjb
#05.b cat.1	t013	26-07-2004	11-01-2007	*is **a project** that intends to create a universal medium for the information exchange by giving meaning, in a manner understandable by machines, to the content of documents on the WWW.*	Lou Quillio
#06.c cat.1	t020	11-01-2007	26-01-2007	*is **an evolution of the current Web** that seeks to provide granular access to the underlying data that fuels the WWW.*	KingsleyIdehen
#07.m cat.2.1	t020	11-01-2007	20-02-2007	*It's **a manifestation of the** W3C chairman **Tim Berners-Lee's vision of the Web** as a universal medium for Data, Information, and Knowledge exchange.*	KingsleyIdehen
#08.c cat.1	t026	26-01-2007	19-02-2007	*is **an evolution of the WWW** in which information is machine processable (rather than being only human oriented),*	71.68.198.237
#09.d cat.1	t028	19-02-2007	20-02-2007	*is **a loosely defined and evolving framework of WWW based technologies** that seek to augment human readable content with information that is machine processable,*	Numskll
#10.d cat.1	t029	20-02-2007	20-02-2007	*is **a loosely defined and evolving framework** intended to augment web content with machine processable metadata,*	Dreftymac
#11.n cat.2.1	t029	20-02-2007	12-06-2010	***It derives from** W3C director **Tim Berners-Lee's vision of the WWW** as a universal medium for data, information, and knowledge exchange.*	Dreftymac
#12.e cat.1	t031	20-02-2007	21-02-2007	*is **a set of loosely-defined and evolving initiatives** to extend web content into a framework that can be processed and interpreted by automata,*	Dreftymac

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 July 2018 doi:10.20944/preprints201807.0560.v1

Peer-reviewed version available at Publications 2018, 6, 44; doi:10.3390/publications6040044

ref.Df	ref.Tt.	Start	End	Context units	Users
#13.f cat.1	t034	21-02-2007	13-07-2009	*is **an evolving extension of the WWW** in which Web content can not only be expressed in natural language, but also in a form that can be understood, interpreted and used by software agents,*	Cygri
#14.g cat.1	t057	13-07-2009	12-06-2010	*is **an evolving development of the WWW** in which web content can not only be expressed in natural language, but also in a form that can be understood, interpreted and used by software agents,*	Andy Dingley
#15.h cat.1	t067	12-06-2010	01-09-2010	***It describes methods and technologies** to allow machines to understand the meaning - or "semantics" - of information on the WWW.*	Averell23
#16.o cat.2.1	t067	12-06-2010	01-09-2010	*is **a term** coined by W3C director Sir Tim Berners-Lee.*	Averell23
#17.q cat.2.2	t067	12-06-2010	23-02-2011	*it is mainly used to describe **the model and technologies** proposed by the W3C.*	Averell23
#18.g cat.1	t073	01-09-2010	23-02-2011	*is **a group of methods and technologies** to allow machines to understand the meaning - or "semantics" - of information on the WWW.*	Wikidemon
#19.p cat.2.1	t074	13-11-2010	21-11-2011	*Tim Berners-Lee defined the Semantic Web **as "a web of data** that can be processed directly and indirectly by machines".*	99.41.179.96
#20.i cat.1	t085	23-02-2011	03-09-2011	*is **a "web of data"** that enables machines to understand the semantics, or meaning, of information on the WWW.*	Michael A. White
#21.r cat.2.2	t085	23-02-2011	12-11-2011	*is often used more specifically to refer to **the formats and technologies** that enable it.*	Michael A. White
#22.i cat.1	t093	03-09-2011	12-11-2011	*is **a "man-made woven web of data"** that facilitates machines to understand the semantics, or meaning, of information on the WWW.*	Wireless friend
#23.j cat.1	t096	12-11-2011	21-11-2011	*is **the roadmap of a "man-made woven web of data"** that facilitates machines to understand the semantics, or meaning, of information on the WWW.*	Karima Rafes
#24. cat.1k	t097	21-11-2011	06-06-2013	*is **a collaborative movement** led by the W3C that promotes common formats for data on the WWW.*	24.69.174.26
#25.o cat.2.1	t113	06-06-2013	-	*The term was coined by Tim Berners-Lee for **a web of data** that can be processed by machines.*	Nigelj
#26.l cat.1	t117	09-03-2015	-	*is **an extension of the Web** through standards by the W3C.*	Denny

References

1. Berners-Lee, T.; Fischetti, M. Weaving the Web: The original design and ultimate destiny of the World Wide Web; Harper Collins: New York, **1999**; ISBN 006251587X
2. Hendler, J.; Shadbolt, N.; Hall, W.; Berners-Lee, T.; Weitzner, D. Web Science: An interdisciplinary approach to understanding the web. Commun. ACM **2008**, 51, 60, doi:10.1145/1364782.1364798.
3. Flores, P. Is Wikipedia the largest-ever digital humanities project? Exploring an emerging relationship Available online: https://blog.wikimedia.org/2016/08/17/wikipedia-largest-digital-humanities-project/ (accessed on Nov 16, 2016).
4. Iñiguez, G.; Török, J.; Yasseri, T.; Kaski, K.; Kertész, J. Modeling social dynamics in a collaborative environment. EPJ Data Sci. **2014**, 3, 7, doi:10.1140/epjds/s13688-014-0007-z.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 July 2018 doi:10.20944/preprints201807.0560.v1

Peer-reviewed version available at Publications 2018, 6, 44; doi:10.3390/publications6040044

5. Westerman, W. Epistemology, the Sociology of Knowledge, and the Wikipedia Userbox Controversy. In Folklore and the Internet: Vernacular Expression in a Digital World; Blank, T. J., Ed.; University Press of Colorado, Urban Institute, **2009**; pp. 123–158 ISBN 978-0-87421-750-6.
6. Kleeb, R.; Gloor, P. A.; Nemoto, K.; Henninger, M. Wikimaps: dynamic maps of knowledge. Int. J. Organ. Des. Eng. **2012**, 2, 204, doi:10.1504/IJODE.2012.047575.
7. Biuk-Aghai, R. P. Visualizing Co-Authorship Networks in Online Wikipedia. In International Symposium on Communications and Information Technologies; IEEE, **2006**; pp. 737–742.
8. Cress, U.; Kimmerle, J. A theoretical framework of collaborative knowledge building with wikis: a systemic and cognitive perspective. In CSCL'07: Proceedings of the 8th iternational conference on Computer supported collaborative learning; **2007**; pp. 156–164.
9. Kimmerle, J.; Moskaliuk, J.; Cress, U. Using Wikis for Learning and Knowledge Building: Results of an Experimental Study. J. Educ. Technol. Soc. **2011**, 14, 138–148.
10. Kump, B.; Moskaliuk, J.; Dennerlein, S.; Ley, T. Tracing knowledge co-evolution in a realistic course setting: A wiki-based field experiment. Comput. Educ. **2013**, 69, 60–70, doi:10.1016/j.compedu.2013.06.015.
11. Kimmerle, J.; Moskaliuk, J.; Harrer, A.; Cress, U. Visualizing co-evolution of individual and collective knowledge. Information, Commun. Soc. **2010**, 13, 1099–1121, doi:10.1080/13691180903521547.
12. Moskaliuk, J.; Kimmerle, J.; Cress, U. Collaborative knowledge building with wikis: The impact of redundancy and polarity. Comput. Educ. **2012**, 58, 1049–1057, doi:10.1016/j.compedu.2011.11.024.
13. Halatchliyski, I.; Moskaliuk, J.; Kimmerle, J.; Cress, U. Explaining authors' contribution to pivotal artifacts during mass collaboration in the Wikipedia's knowledge base. Int. J. Comput. Collab. Learn. **2014**, 9, 97–115, doi:10.1007/s11412-013-9182-3.
14. Yap, K. C.; Chia, K. P. Knowledge construction and misconstruction: A case study approach in asynchronous discussion using Knowledge Construction – Message Map (KCMM) and Knowledge Construction – Message Graph (KCMG). Comput. Educ. **2010**, 55, 1589–1613, doi:10.1016/j.compedu.2010.07.002.
15. Aviv, R.; Erlich, Z.; Ravid, G.; Geva, A. Network analysis of knowledge construction in asynchronous learning networks. J. Asynchronous Learn. … **2003**, 7, 1–23.
16. Sultan, N. Knowledge management in the age of cloud computing and Web 2.0: Experiencing the power of disruptive innovations. Int. J. Inf. Manage. **2013**, 33, 160–165, doi:10.1016/j.ijinfomgt.2012.08.006.
17. Ren, R. The evolution of knowledge creation online: Wikipedia and Knowledge Processes. In Proceedings of the 11th International Symposium on Open Collaboration - OpenSym '15; ACM Press: San Francisco, **2015**; pp. 1–3.
18. Holman Rector, L. Comparison of Wikipedia and other encyclopedias for accuracy, breadth, and depth in historical articles. Ref. Serv. Rev. **2008**, 36, 7–22, doi:10.1108/00907320810851998.
19. Giles, J. Internet encyclopaedias go head to head. Nature **2005**, 438, 900–901, doi:10.1038/438900a.
20. Blumenstock, J. E. Size matters: Word Count as a Measure of Quality on Wikipedia. In Proceeding of the 17th international conference on World Wide Web - WWW '08; ACM Press: Beijing, **2008**; p. 1095.
21. Townsend, S.; Osmond, G.; Phillips, M. G. Wicked Wikipedia? Communities of Practice, the Production of Knowledge and Australian Sport History. Int. J. Hist. Sport **2013**, 30, 545–559, doi:10.1080/09523367.2013.767239.
22. Gottschalk, S.; Demidova, E. Analysing Temporal Evolution of Interlingual Wikipedia Article Pairs. In Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval - SIGIR '16; ACM Press: Pisa, **2016**; pp. 1089–1092.
23. Zhao, S. J.; Zhang, K. Z. K.; Wagner, C.; Chen, H. Investigating the determinants of contribution value in Wikipedia. Int. J. Inf. Manage. **2013**, 33, 83–92, doi:10.1016/j.ijinfomgt.2012.07.006.
24. Scardamalia, M.; Bereiter, C. Knowledge building. Encycl. Educ. **2003**, 1370–1373.
25. Hjørland, B. Concept theory. J. Am. Soc. Inf. Sci. Technol. **2009**, 60, 1519–1536, doi:10.1002/asi.21082.
26. Arboit, A. E. O processo de (re) construção da teoria do conceito no domínio de Organização do Conhecimento: uma visão dialógica. Scire **2012**, 2, 129–134.
27. Dahlberg, I. Teoria do Conceito. (A. T. Campos, Trans.). Ciência da Informação **1978**, 7, 101–107.
28. Dahlberg, I. A referent-oriented, analytical concept theory for INTERCONCEPT. Int. Classif. **1978**, 5, 143–151.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 July 2018 doi:10.20944/preprints201807.0560.v1

Peer-reviewed version available at Publications 2018, 6, 44; doi:10.3390/publications6040044

29. Machado, L. M. O.; Rocha, R. S.; Simões, M. da G. de M. Semantic Web or Web of Data? A diachronic study (1999 to 2017) of the publications of Tim Berners-Lee and the World Wide Web Consortium. J. Assoc. Inf. Sci. Technol. in press.
30. Gil, A. C. Métodos e Técnicas de Pesquisa Social; 6th ed.; Atlas S.A.: São Paulo, **2008**; ISBN 9788522451425.
31. Bardin, L. Análise de conteúdo; Almedina: São Paulo, **2011**; ISBN 9788562938047.
32. Yin, R. K. Case study research: Design and methods; 5th ed.; SAGE Publications: London, **2014**; ISBN 9781452242569.
33. Manual de análise do discurso em Ciências Sociais; Iñiguez, L., Ed.; Vozes: Petrópolis, **2004**; ISBN 9788532630049
34. Stock, W. G. Concepts and semantic relations in information science. J. Am. Soc. Inf. Sci. Technol. **2010**, 61, 1951–1969, doi:10.1002/asi.21382.
35. Viégas, F. B.; Wattenberg, M.; Dave, K. Studying cooperation and conflict between authors with history flow visualizations. In Proceedings of the 2004 conference on Human factors in computing systems - CHI '04; ACM Press: Vienna, **2004**; Vol. 6, pp. 575–582.
36. Vigotski, L. S. A construção do pensamento e da linguagem; Martins Fontes: São Paulo, **2001**; ISBN 9788533624306
37. Hacking the Academy: New Approaches to Scholarship and Teaching from Digital Humanities; Cohen, D. J., Scheinfeldt, T., Eds.; University of Michigan Press, **2013**; ISBN 9780472071982.
38. Derqui, P. M. Da informação à categorização: a formação sistêmica dos conceitos, Universidade de São Paulo: São Paulo, **2014**.