Association for Information Systems
AIS Electronic Library (AISeL)

ACIS 2013 Proceedings | Australasian (ACIS)

2013

The Genealogy of Knowledge: Introducing a tool and method for tracing the social construction of knowledge on Wikipedia

Friedrich Chasin
University of Muenster, f_chas01@uni-muenster.de

Uri Gal
The University of Sydney, uri.gal@sydney.edu.au

Kai Riemer
The University of Sydney, kai.riemer@sydney.edu.au

Follow this and additional works at: https://aisel.aisnet.org/acis2013

Recommended Citation

Chasin, Friedrich; Gal, Uri; and Riemer, Kai, "The Genealogy of Knowledge: Introducing a tool and method for tracing the social construction of knowledge on Wikipedia" (2013). *ACIS 2013 Proceedings*. 63.
https://aisel.aisnet.org/acis2013/63

This material is brought to you by the Australasian (ACIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ACIS 2013 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Information Systems: Transforming the Future

24th Australasian Conference on Information Systems, 4-6 December 2013, Melbourne

Proudly sponsored by

ACIS 2013 Principal Sponsor

The Genealogy of Knowledge Introducing a Tool and Method for Tracing the Social Construction of Knowledge on Wikipedia

Friedrich Chasin
Department of Information Systems
The University Münster, Germany
Email: f_chas01@uni-muenster.de

Uri Gal
Discipline of Business Information Systems
The University of Sydney Business School, Australia
Email: uri.gal@sydney.edu.au

Kai Riemer
Discipline of Business Information Systems
The University of Sydney Business School, Australia
Email: kai.riemer@sydney.edu.au

Abstract

The study of the social construction of knowledge is an important topic in information systems. While the emergence of online social media platforms has brought about multiple new ways of knowledge co-creation by large groups of users, such processes are not yet well understood. We investigate how processes of knowledge co-creation can be studied in online platforms. We utilize data from Wikipedia to explore how knowledge is created and evolves over time. We draw on the theory of social representations (SRT), which views knowledge as a product of collective work carried out by social groups to make sense of their environments. We develop a method for studying social representations on Wikipedia that builds on WikiGen - an analytical tool - and qualitative analysis. Its usefulness is demonstrated with two illustrative case studies, the Cloud Computing and iPad Wikipedia articles. We contribute to IS a new tool and method for studying online knowledge co-creation processes.

Keywords

Knowledge, Knowledge creation, Social representations, Social constructivism, Wikipedia, WikiGen.

INTRODUCTION

Knowledge is inherently dynamic (Marková 1996). The study of knowledge, its creation, and ramifications is of significance to the field of information systems (IS) because knowledge and technology are intimately intertwined. Firstly, knowledge assumptions can be inscribed into IS in ways that shape the way users interact with the technology and with each other. Secondly, the way expert knowledge, as well as commonly held views on certain technological developments, come into being and get accepted is important to understanding the reasons behind the acceptance or rejection of such developments. Thirdly, certain types of IS are used as a medium through which groups and individuals communicate to express their views, construct their understanding of different phenomena and co-create new knowledge.

In this paper we focus on the third phenomenon and investigate the way in which knowledge emerges and evolves within online social platforms, a phenomenon that to date is not yet well understood. In doing so, we subscribe to a social constructivist view of knowledge (Berger and Luckmann 1966). From this perspective knowledge is a product of human interaction and communication involving different interests, rather than a neutral description of an external reality (Duveen 2000). Knowledge is dynamic, evolving and constantly negotiated in a social discourse. It is thus always historical in that interpretations of present events and phenomena are shaped by past ideas and experiences and can only be understood in relation to them.

To study the social process of knowledge creation we draw on social representations theory (SRT) (Moscovici 2000). The theory's primary assertion is that objects and events in the world only become known to us by being socially represented. When we experience a new event or encounter an unfamiliar phenomenon we engage in a social process of representing: a collective sense-making to name and situate the unfamiliar into the existing stock of knowledge. Only by being represented by a group of people does a phenomenon become part of that group's social reality and can be used in its members practice and communication (Moscovici, 1984).

Online social platforms, such as the online encyclopaedia Wikipedia, are particularly appropriate for studying the social construction, co-creation and evolution of knowledge. These platforms are designed to facilitate communication among multiple participants, approximating Habermas' idea of rational discourse (Habermas 1984 Hansen et. al. 2009). Furthermore, they allow public access to the knowledge that is produced through this process, as well as to the discussions among the participants, thus affording a view into the process of knowledge construction.

Consequently, the objective of this paper is to explore, through the lens of SRT, the social process of knowledge construction, and thus the resulting genealogy of knowledge, on online social platforms. We choose Wikipedia as an appropriate platform for studying this phenomenon. This leads us to formulate the following research question: "How can the evolution of social representations be studied on Wikipedia?"

While Wikipedia has been framed as a knowledge platform (Suchecki et al. 2012) and a number of studies exist that examine collaboration and consensus making in Wikipedia (e.g., Kaltenbrunner and Laniado 2012), there have been no attempts to systematically study the patterns of knowledge creation on the platform. At the same time Wikipedia presents an opportunity to study the social construction of knowledge and unpack some of its complexity. The structure of Wikipedia, where historical data on every article edit can be accessed, along with the transparency of its collaboration processes, make it ideal for exploring the genealogy of knowledge.

By answering the research question, we will introduce a method for studying the emergence and evolution of social representations on the Wikipedia platform. Specifically, we will demonstrate the effectiveness of an analytical tool that we developed for deriving suitable data for studying the emergence of social representations in Wikipedia articles. More generally, we will unpack and illustrate with data the process through which knowledge is socially generated and shaped.

Our method and analytical tool makes a methodological contribution for studying the evolution and social construction of emerging concepts. This will be particularly for scholars in fast moving fields such as IS, where new ideas, technologies and concepts emerge at a fast pace and scholars want to learn about how social groups make sense of, bring about and come to grips with these changes. Moreover, we contribute to the IS knowledge management field a practical application of SRT that demonstrates its usefulness in studying emerging online phenomena, such as knowledge co-creation and crowdsourcing.

We begin by presenting the theory of social representation and describing how we apply it to Wikipedia. Next we introduce WikiGen, an analytical tool which we developed to study how social representation emerge and change over time on the platform. We then apply the tool to two exemplary case studies - the *Cloud Computing (CC)* and *iPad* Wikipedia articles – and conduct quantitative and qualitative analyses of the data. This is followed by a discussion of the exemplary findings. We conclude with suggestions for further research.

THEORY OF SOCIAL REPRESENTATIONS ON WIKIPEDIA

The central idea of the theory of social representations is that people's relationship with the world is invariably mediated by a layer of socially constructed and continuously-evolving symbols, or representations, which serve to render the world meaningful for social actors (Gal and Berente 2008). Things in the world do not have an inherent meaning. Rather, they acquire meaning when they are represented by social groups in an ongoing communicative process. Social representations are semiotic devices that members of a social group construct to render their world meaningful (Valsiner 2003, Wagner et al. 1999). Phenomena or events only become *social* reality by virtue of their representations that the community forms. Only by being represented by a group of people by means of familiar conceptual devices can an event or phenomenon become a *social* object that can be perceived, characterized, compared to other social objects, and used in language and action.

The Process of Representing

Representational activity is a social process that is often triggered when something disruptive threatens socially shared perceptions of reality (Moscovici 2000). This can be a new and unfamiliar phenomenon, or an unexpected characteristic of a familiar phenomenon, which group members lack the cognitive vocabulary to describe and name. This unfamiliarity creates a sense of incompleteness and emphasises the "actuality of something absent" (Moscovici 2000 p.38). To familiarise the unfamiliar, group members form new social representations in a process that has two stages: *anchoring* and *objectification*.

Anchoring. Whenever something unfamiliar is experienced, a process of familiarizing is triggered by means of classification (Wagner et al. 1999). Classification is the positioning of something unknown in familiar conceptual categories. The choice of a suitable class of categories is based on a comparison of the unfamiliar to prototypes considered to represent the corresponding class (Moscovici 1984). For example, in its early stage, the unfamiliar phenomenon of HIV/AIDS (before acquiring this name) was anchored in terms of a 'gay plague' or

'gay cancer' (Farr 1993). Thus, the new phenomenon of HIV/AIDS was initially understood in terms of, and took on the qualities associated with, the plague or cancer. Importantly, the process of anchoring is dynamic and reflects group members' changing perceptions of different aspects in their environment. Anchors are thus an integral part of thinking in general; "there is no thought or perception without anchor" (Moscovici 2000, p.48).

Objectification. After anchoring a new phenomenon and placing it in familiar categories, further communicative activities among group members lead to an objectified representation in the form of a metaphor, symbol, or image (Wagner et al, 1999). Objectification is the process whereby socially represented knowledge receives its concrete and distinct form, or representation. Objectification involves the development of a signifier that stands for the phenomenon or object that it represents. The representation captures the essence of the phenomenon and weaves it into the social fabric of the group's common stock of knowledge.

The choice of a representation is not arbitrary. It typically is related to the knowledge, vocabulary and imagery that group members have in common and which reflect their shared identity, history, and everyday 'social terrain' (Moscovici, 2000). Accordingly, different groups may develop different, sometimes conflicting, representations of the same phenomenon, depending on their socio-historical contexts (Gal and Berente, 2008). At the same time, the complexity and interdependencies between social and individual factors in the process of forming a representation make any attempt to unveil its lineage a challenging task (László 1997).

Social Representations on Wikipedia

Wikipedia is not just an encyclopaedia that aims to hold knowledge about the world. It is also the place where this knowledge is created; a place of discourse and social engagement, where contributors collaborate, but also intensely debate and disagree about how to represent events, objects, people and other phenomena in their world[1]. Wikipedia users can be seen as a collective 'decision committee', in that they 'vote' and express opinions by changing Wikipedia content and participating in discussions on the corresponding talk pages. They can know how others have voted by accessing the revision history and observing previous discussions regarding the topic at hand. In doing so, Wikipedia users build upon the efforts of others. Consequently, reaching a temporary agreement on Wikipedia, is similar to the type of consensus building when forming social representations.

Consequently, articles in Wikipedia can be interpreted as social representations. Social representations provide a code for unambiguously naming various aspects of the world (Moscovici 2008). Furthermore, they help individuals to orientate themselves and allow communication. Those characteristics are inherent to Wikipedia articles. Each article gives a unique name to a particular aspect of the world and aims to provide meaning to it (Siorpaes and Bachlechner 2006).

Having established that Wikipedia articles can be seen as social representations and that the process of article creation corresponds to the process of knowledge construction in SRT, it is important to unpack what constitutes anchoring and objectification on Wikipedia. The most natural 'candidate' for an anchor role within an article is an internal link to another article. SRT states that anchors are themselves social representations (Moscovici 2000). An internal link within one Wikipedia article to another article satisfies this requirement. Creating internal links to other articles aims at identifying existing concepts that are perceived to be relevant to understanding the phenomenon at hand. The link anchors the current representation in terms of an existing one. However, not every internal link contained in an article is an anchor. Links contained in the first section of an article (the 'definition' section), are the most relevant anchor candidates[2]. Consequently, changing internal links in the definition section of over time reflect the continuous process of anchoring and re-anchoring.

Objectification is best operationalized as the act of linking other Wikipedia articles to the current article. Wikipedia articles that contain a reference to the social representation under study in form of a link are therefore relevant for tracing the objectification process. Objectification indicates that socially represented knowledge gains a distinct and concrete form. Therefore a growing number of links that point to the social representation under study would reflect an ongoing objectification process, as it would indicate that other phenomena are explained in terms of the representation under study.

---

[1] Recent coverage in the media provides a vivid illustration of this point: "Topics that spark Wikipedia 'edit wars' revealed", http://www.bbc.co.uk/news/technology-23354613

[2] The scope and structure of articles are highly divergent and dynamically defined by users. Therefore, different sections in an article can vary greatly and be semantically removed from the subject of the article. Hence, it is likely that links contained in these sections are not used to anchor the phenomenon under study. The definition section is different because every article starts with this section and because it contains information that pertains to the phenomenon under study. Therefore, links contained there are likely to be relevant anchors.

WIKIGEN: AN ANALYSIS TOOL FOR GENERATING WIKIPEDIA STATISTICS

In order to trace anchoring and objectification in Wikipedia, we have developed web-based analytical tool, WikiGen[3]. By connecting to Wikipedia's live databases over Wikipedia's Application Programming Interface (API), WikiGen generates multiple interactive statistics and navigation maps based on the historical revisions of any chosen article (in 196 languages). The tool can generate collaboration, anchoring, and objectification statistics. Next we will outline the tool's most salient features, before we illustrate how they can be used as part of a SRT case study for tracing the genealogy of social representations on Wikipedia.

Collaboration Statistics

To study how the editing of an article unfolds through collaboration among multiple people over time, it is important to gain an overview of the editing process. To this end, WikiGen derives a number of measures for generating collaboration statistics. WikiGen provides aggregate data on the *number of edits* and *number of editors* in different periods of time, as well as a combined measure of *edits per editor*. To capture the unfolding collaboration processes in greater detail we distinguish between different editor types in WikiGen according to user attributes provided by the Wikipedia platform: anonymous, registered and bot editors. *Overall editors* includes all three types. Similarly, we divide overall edits into *distinct edits* and *major distinct edits*. This distinction is based on the data provided by Wikipedia regarding major or minor edits as well as the existence of series of subsequent revisions that are created by the same user while he or she is saving the article several times during the editing procedure. We consider such edit series a *single distinct edit*. Figure 1 illustrates an output of the WikiGen tool for editing statistics plotted over time for the article "Kevin Rudd".

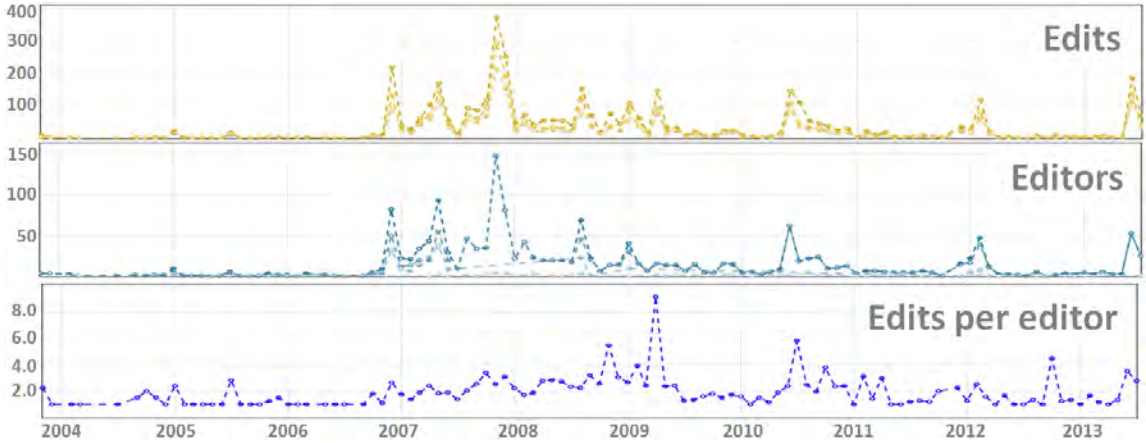

Figure 1: WikiGen output for different types of edits over time for the article "Kevin Rudd".

Anchoring Statistics

To measure anchoring activity, WikiGen generates statistics on internal links in the definition part of Wikipedia articles, which point to other Wikipedia articles. Analysis is centred on the changing composition of internal links, indicated by historical revisions.

Anchor Snapshots. Anchor snapshots allow comparisons between present anchors and their attributes, in different periods of time. To illustrate, figure 2 shows an anchor snapshot for the "iPad" article. WikiGen generates four attributes for each anchor:

- *Days survived*: The cumulative number of days an anchor was present in the definition part of the article, which takes into account that anchors can be introduced or removed throughout the period.
- *Revisions survived*: The number of major distinct edits an anchor has survived without being removed.
- *Re-introductions*: Indicates how many times an anchor was introduced to the definition part of the article.
- *Anchor Strength*: a linear combination of *days survived* and *revisions survived* that produces a measure between '0' to '1'. This measure captures the fact that some anchors appear more important, as they survive longer or are quickly reintroduced back after being removed by another user. A score of '1' indicates that an anchor has survived all revisions and stayed in the article for the whole period of time.

[3] The tool is available online: http://wikigen.info.net.ua

Anchor	Days survived	Revisions survived	(Re)Introductions	Anchor strngth
apple inc.	363.37	2576	10	0.99
apple a4	338.22	2429	7	0.93
xga	338.18	2345	10	0.91
accelerometer	338.17	2321	6	0.91
flash memory	338.17	2321	6	0.91
multi-touch	338.15	2253	4	0.90
ipod touch	338.14	2201	4	0.89
iphone	338.13	2197	2	0.89
app store	363.00	2003	6	0.88
light-emitting diode	338.11	2141	2	0.88

Showing 1 to 10 of 160 entries Previous Next

Figure 2: Ten first anchors in a yearly anchor snapshot table for ”iPad” ordered by *Anchor Strength.*

New and Obsolete Anchors. WikiGen generates a bar chart in order to indicate the number of newly introduced and removed anchors in a particular period of time. The chart, such as in figure 3, helps identifying periods with significant changes in the anchoring in which anchors are either remover or introduced, or both. Such periods can then be studied more closely in the qualitative part of the analysis.

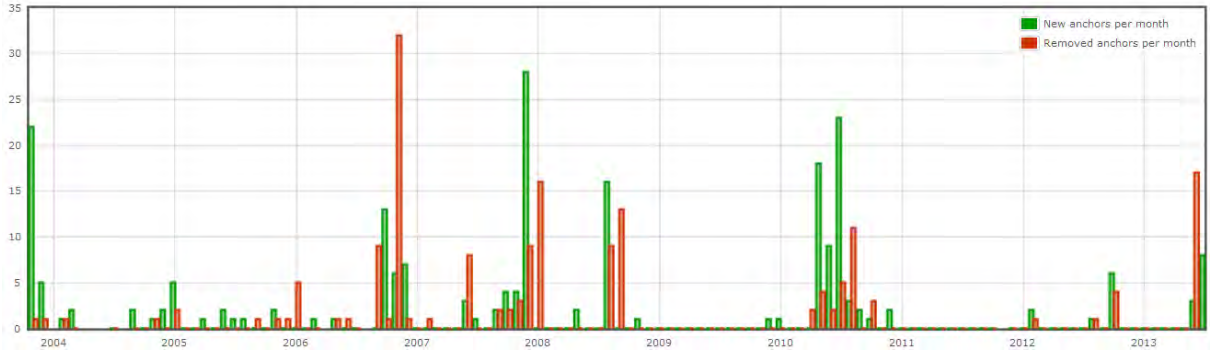

Figure 3: Exemplary output for new and obsolete anchors in WikiGen for the article “Kevin Rudd”.

Anchor Dissimilarity. This measure captures the extent to which anchors in time period *t* are dissimilar to anchors in the previous period *t-1*. This measure allows comparing periods over time to identify where significant relative change in anchoring happens. WikiGen allows flexibility in selecting time periods for analysis, from years to months The basis for this measure is the anchor attribute *days survived* which is the cumulative time an anchor stayed in the definition part of the article in the given period of time *t*. To remove influence of anchors that enter the article due to vandalism, only anchors that were present in the corresponding period for at least one day are considered. The dissimilarity score ranges from 0 (anchors are absolutely dissimilar) to 1 (anchors are absolutely similar) across time periods. The logic behind the calculation is to take the sum of least common days survived for all anchors present in periods *t* and *t-1* and set it in relation to the total sum of maximum days anchors survived for every anchor in periods t and *t-1*. The corresponding formula and an exemplary monthly output are illustrated in figure 4.

$$d(t,t-1) = 1 - \frac{\sum_{i=1}^{n} \min\left(f(a_{k_i},t), f(a_{k_i},t-1)\right)}{\sum_{j=1}^{n} \max\left(f(a_{k_j},t), f(a_{k_j},t-1)\right)}, \quad \text{where}$$

$\boldsymbol{a_k} = a_t \cup a_{t-1} = \begin{pmatrix} a_{k_1} \\ a_{k_2} \\ \dots \\ a_{k_p} \end{pmatrix}$: union of anchor name vectors a_t and a_{t-1} in periods t and $t-1$ with $\boldsymbol{a_t} = \begin{pmatrix} a_{t_1} \\ a_{t_2} \\ \dots \\ a_{t_i} \end{pmatrix}$ being a vector of anchor names a_{t_i} in period t,

$f(x,t) = z$, *with* $x \in a_t$ and z being the days survived value for anchor x in t.

Figure 4: Anchor dissimilarity formula and exemplary output for the article “Google”, March-October 2006.

Average Anchor Durability[4] measures the average time (e.g. number of days/month) each anchor was present in the definition part of an article in a particular time frame *t*. Again, we consider only those anchors that have stayed in the article at least for one day. The measure can again shed light on how stable the anchoring is during a particular time frame *t*. The formula and an exemplary output are illustrated in figure 5.

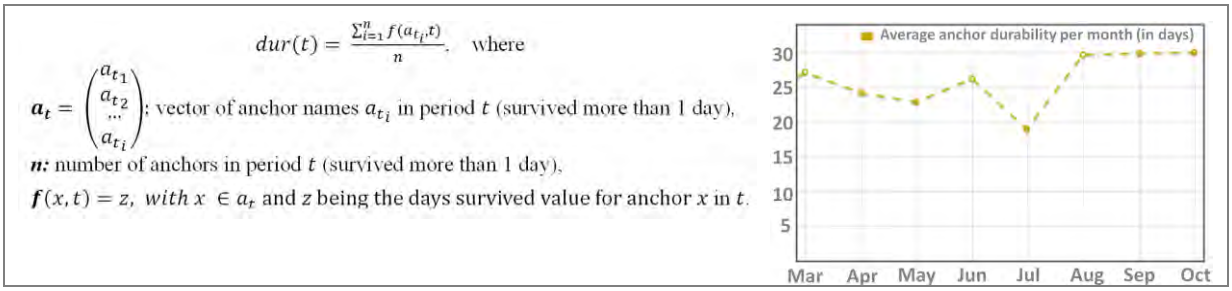

Figure 5: Anchor durability formula and exemplary output for the article "Google", March-October 2006.

Anchor edit-war level measures the level of disagreement among users by relating the number of introductions and disappearances of anchors to the total number of unique anchors in a period of time. Higher values of this measure indicate intense re-anchoring and significant changes in the social representation[5] (see figure 6).

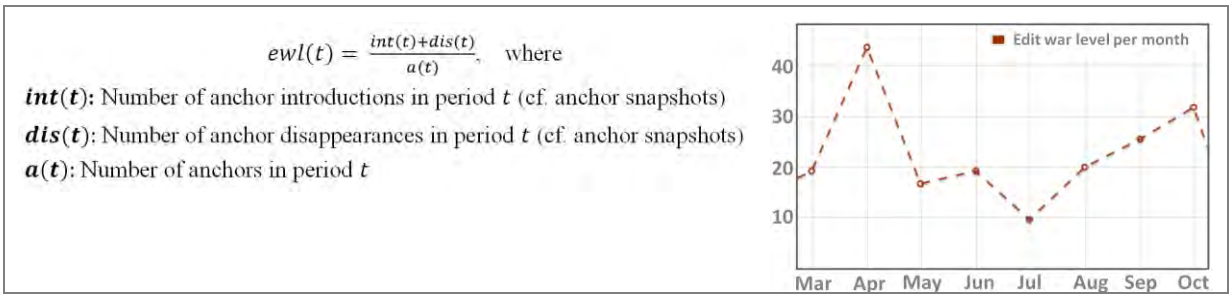

Figure 6: Anchor edit war level formula and exemplary output for the article "Google", March-October 2006.

Objectification Statistics

WikiGen provides data on how many anchors (i.e., links) that point to the analysed article were added in past periods, which allows plotting an overview of how the number of incoming links grew to make up the present number (see figure 12 later in the paper). For example, if a link to 'Serge Haroche' was introduced to the 'Physics' article in 2012, WikiGen would add one reference for the year 2012. If an article contains several links pointing at the article of interest, it is considered as only one reference.

COMPLEMENTING WIKIGEN STATISTICS WITH QUALITATIVE ANALYSIS

Quantitative analysis alone is not sufficient to describe the evolution of a social representation on Wikipedia. Firstly, the meaning of anchors can only be interpreted in the context they are used. Secondly, changes in the evolution of a social representation cannot be described in terms of single anchors alone. Rather, anchors should be semantically categorised to identify changes across different anchor groups over time. Therefore, data generated by WikiGen are utilised for subsequent qualitative analysis. Finally, to understand how a social representation changes over time, the overall analysis time frame is divided into distinct phases that reflect significant changes in the composition of anchor categories. We explain these in detail below.

Anchor Categorisation

Categorisation is done in three steps. In the first, all the different text passages in which the anchor appears are examined and their relationship to the phenomenon is established. For example, the relationship of the anchor

[4] The range is between 0 and the maximum time in the period. For the monthly output it therefore varies between 0 and 28 – 31 days (depending on the month).

[5] The 'edit-war' measure ranges from 0 (no introductions or disappearances of anchors) to infinity (x anchors are introduced y times where y -> inf)

'iPhone' to the social representation of 'iPad' can be one of similarity in one context, e.g., "a device with which iPad shares some features" or one of separation in another context, e.g., "iPad is different to iPod because it has a larger screen". When the relationships vary across multiple revisions of an anchor, all different relationships are considered for analysis.

Navigation between the relevant revisions for every anchor is allowed by WikiGen's *Anchor Map* functionality (figure 7). The map visualises periods of time a chosen anchor was present or absent in the article. Clicking on the graph links the user to the historical version of the text in which the anchor disappeared or (re)appeared. This way it is possible to navigate between relevant historical revisions of the article in which the analysed anchor is contextualised. Figure 7 provides an extract from the anchor map for the 'user interface' anchor in the iPad article. There are two time periods in which the anchor was present in the article. Every point in the map (see dots in the figure) corresponds to a revision in which an anchor was introduced or removed. Clicking on the point opens the corresponding revision text and shows the content of the historical revision of the article.

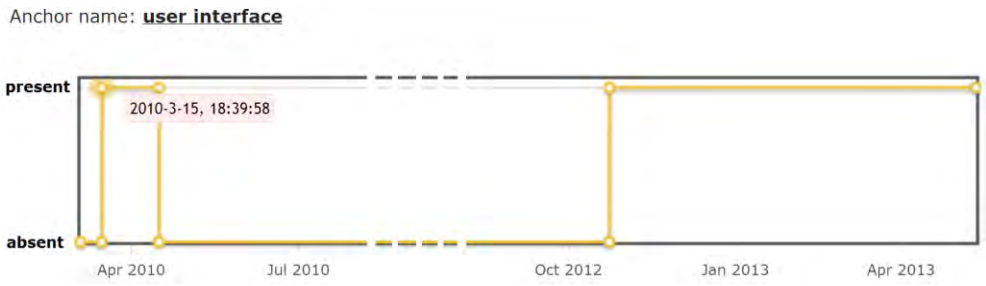

Figure 7: Anchoring map in WikiGen for the 'user interface' anchor in the iPad article

Given the context for every anchor, it is possible to merge redundant, and delete irrelevant, anchors in step two. The reasons for anchors to be considered irrelevant are manifold. For example, the definition section of an article can contain a quote from an article that is available as a pdf document. It happens that a link underlies the word 'pdf' and links to the article about the .pdf file format. Clearly this is not an anchor for the representation under study. Redundancy on the other hand is always due to either the renaming of an article during the time-frame of analysis or due to the existence of multiple articles that have different names, yet describe the same phenomenon.

The resulting list of relevant anchors corresponds to open codes in qualitative content analysis and is subject to theoretical coding in step three, which is achieved by assigning and reassigning anchors to categories in several iterations until a minimal set of mutually exclusive and collectively exhaustive categories is derived.

Identification of Anchor Evolution Phases

In the next step, the overall anchoring time frame is divided into distinct phases that reflect significant changes between the identified anchor categories. Several WikiGen statistics can be used to achieve this aim. First, data on new and old anchors for each period help identify time periods in which significant anchor movements are observed. Second, the key indicators introduced above further point to phases with significant changes in the social representation under study: A *high anchor dissimilarity* means that the representation has a high number of different anchors when compared to previous periods. Due to new or removed anchors the average time that anchors are present in the article decreases which shows as *decreased average anchor durability*. If the changes are accompanied by disagreement among users, then the high amount of reintroduced anchors will be reflected in a *high anchor edit war level*.

These quantitative indications can then be used to undertake a closer examination of the anchor snapshots in the identified time frames. Given the anchor categories that were defined, it is possible to see whether quantitative indications correspond to an actual change in the meaning of the social representation in this time period. Such a change is represented as a strengthening or weakening of one or several anchor categories. We present a brief example of the application of WikiGen and the associated qualitative analysis in the next section.

EXEMPLARY APPLICATION OF WIKIGEN AND QUALITATIVE ANALYSIS

We provide illustrative findings from applying WikiGen and the associated qualitative analysis to the 'Cloud Computing (CC)' and 'iPad' Wikipedia articles. Please note that due to space restrictions this illustration is brief, and excludes some WikiGen statistics and figures.

Collaboration Analysis

Consistent with SRT, we observe that when an unfamiliar phenomenon is encountered collaborative representational activity becomes more intense and centralised. In the early phase of both CC and iPad, termed 'initial familiarisation' (see figure 10), collaboration intensity (*number of edits* and *number of editors*) and centralisation (number of *edits per editor*) reach their highest levels for the entire time period. This indicates high engagement by users to establish initial familiarity with the new concept.

Anchoring Analysis

The collaborative activity in the initial phases of CC and iPad is characterised by the introduction of multiple anchor categories (see figure 9 for category overview). The high amount of new and removed anchors in figure 8 illustrates the initial unfamiliarity (grey areas) for both phenomena. As long as more anchors are removed than added, we interpret this as a sign of existing unfamiliarity.

In addition to the initial familiarisation phase of CC and iPad, we have identified additional phases in the evolution of anchoring. Figure 10 gives an overview over the remaining phases.

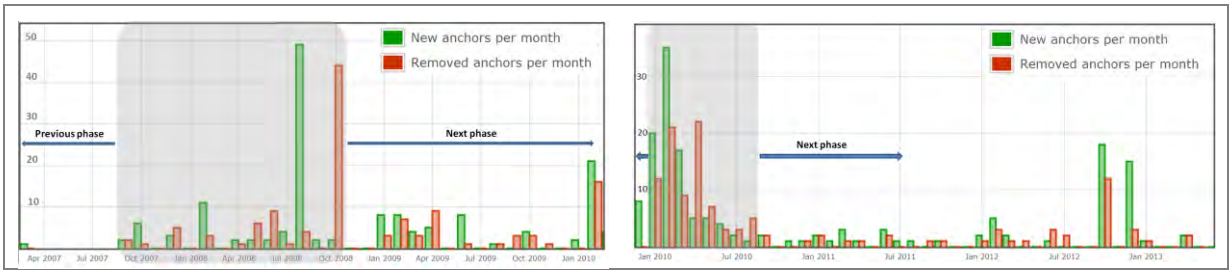

Figure 8: Monthly anchor dissimilarity of CC (left) and of iPad (right)

CC categories	Technical aspects of CC	CC use cases	Generalising (Anchors that extend the scope of CC)	Sub concepts of CC	CC solutions and providers
iPad categories	Technical aspects of the iPad	iPad use cases	Products and technologies that are similar to the iPad	Origins of the iPad	Products and companies that compete with iPad and Apple

Figure 9: Anchor categories for CC and iPad

CC phases	Infancy (Mar07-Aug07)	Initial familiarisation attempts (Sep07-Oct08)	Establishment of the generalising core (Nov08-Jan10)	First concretisation attempt (Feb10-Feb11)	Collapse and reestablishment of a generalising core (Mar11-Jan12)	Second concretisation attempt (Feb12-May13)
	CC is represented as another name for utility computing	*Initial search for different ways by which the unfamiliar phenomenon of CC can be made more tangible.*	*Establishment of the generalising nature of CC according to which cloud computing is almost everything that is computed in the internet*	*A search for suitable ways through which the representation of CC can be made more tangible, while preserving its generalising nature.*	*Attempt to combine a generalising and concrete representation of CC leads to a collapse and reestablishment of the generalising view as a dominant view on CC*	A phase that is similar to the "first concretisation attempts" phase in its interpretation.

iPad phases	Infancy (Dec09)	Initial familiarisation attempts (Jan10-Aug10)	Objectification (Sep10-Sep12)	Individualisation (Oct12-Jun13)
	A phase in which the iPad representation is based purely on rumours and speculations about a new Apple device	*The unfamiliarity associated with the released iPad leads to introduction of anchors from all categories without any of the categories being dominant*	*A phase with an intense objectification process taking place throughout the period. The iPad becomes part of the social reality*	*The representation of the iPad shifts towards a distinct device with its own use cases and its own tablet computer market that is defined in terms of the iPad device itself*

Figure 10: Evolution phases for CC and iPad social representations on Wikipedia

Our analysis revealed distinct differences in the evolution of the CC and iPad representations. Whereas the evolution of the iPad is rather stable and 'smooth' (reflecting little disagreement among users), the CC social representation is marked by a more fluctuating anchoring process. On the one hand, anchors in the 'generalising' category extend the scope of CC by generalising it to additional possible services or resource delivery modes. On the other hand, this view is repeatedly challenged by more specific anchors from the categories of 'technical aspects of CC', 'sub concepts of CC', 'CC use cases' and 'CC solutions and providers'. This tension represents

two diverging views on CC: a general, and a more specific one. This is shown in an ongoing oscillation between more and less specific representations of CC across the different phases of the representation's evolution. This can be observed in the left chart of figure 11 where peaks in the anchor dissimilarity correspond to fluctuations between general and specific perspectives on CC.

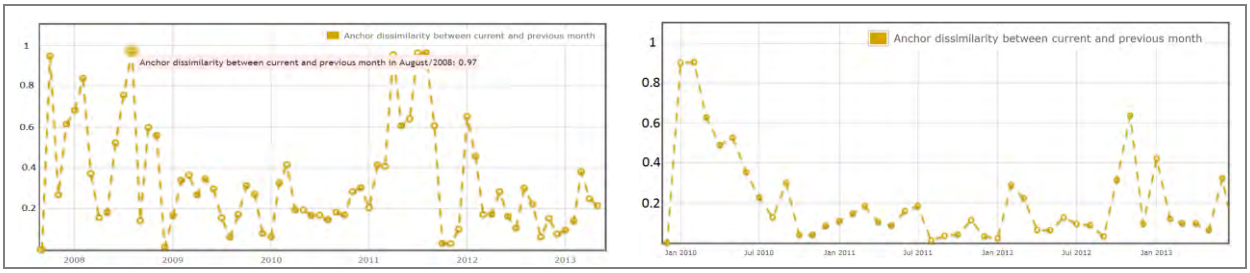

Figure 11: Dissimilarity measures for 'CC' (left) and 'iPad' (right) on a monthly basis.

Overall, and consistent with SRT, we find that both social representations were initially anchored in existing, familiar, and similar concepts; CC in *utility computing*, iPad in *tablet computer* and *iPhone*. In the course of their evolution the articles shift towards more independent representations. In doing so, both representations gradually evolve from being anchored in terms of technical characteristics toward typical ways of use. For example, CC's initial anchors emphasised technical aspects such as *data*, *virtualization*, *computer cluster*, *multi-core* and *parallel computing*. Later anchors focused on how it can be used as application software to achieve economies of scale. As for the iPad, early anchors stressed technological characteristics such as *gigabytes*, *led backlight*, *pixel* and *gigahertz.* Anchors gradually shifted towards possible usages: as a *camera*, *gaming console*, *GPS navigation software*, or *portable media player*[6].

Objectification Analysis

The objectification process of both representations shows a similar consistently increasing trend in the cumulative amount of references from other Wikipedia articles. This observation is particularly interesting given the differences in the two anchoring processes. While the anchoring of the iPad is rather stable, the corresponding process for the CC is volatile. Yet still the social representations of both phenomena increasingly become part of the social reality. Figure 12 displays the total number of new references to the two articles we studied in a corresponding year while the blue line indicates the total (cumulative) number of references over time.

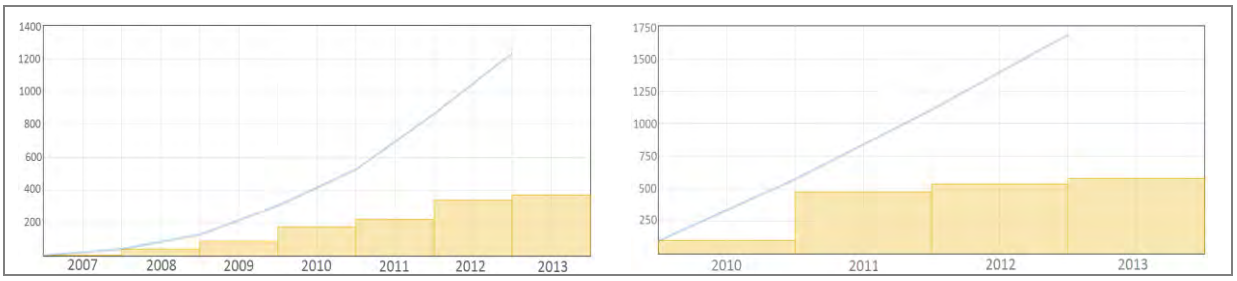

Figure 12: Objectification of CC (left) and of iPad (right)

CONCLUSION

This paper explored the genealogy of knowledge on Wikipedia - how knowledge is socially constructed and shaped over time - using SRT as a theoretical lens. We have done so by developing a tool to analyse the changes in anchoring, objectification, and collaboration patterns among users over time. We have complemented the quantitative analysis with a qualitative analysis to categorise anchors according to their contextual meaning and identify patterns in the evolution of social representations.

[6] We would assume that this pattern might be typical for IT-related innovations that are first given meaning by their technical properties and only over time in terms of their emergent applications.

We have illustrated the applicability of the tool and method with illustrative findings using the CC and iPad representations. First, our analysis provided evidence for continuous anchoring (and re-anchoring) and objectification as posited by SRT. Second, we have shown the usefulness of the quantitative data provided by WikiGen as it allows identifying the critical episodes in a social representation's evolution process. Third, the qualitative analysis was found to be necessary to meaningfully interpret the changes that are indicated by the quantitative data. Finally, we uncovered similarities and differences in the patterns of change of CC and the iPad.

These findings contribute to the IS literature in several ways. First, WikiGen and the associated methodology we developed enable a sophisticated and comprehensive analysis of the formation and development of knowledge on Wikipedia. While existing studies have examined different aspects of this phenomenon (Kaltenbrunner and Laniado 2012), our tool and methodology can generate more refined insights than has been previously possible. This will be useful for scholars in fast moving fields such as IS for studying how social representations of new technologies and concepts emerge and change over time. Second, the application of WikiGen can also contribute to a better understanding of online interaction patterns among users, by providing a view into their intensity and levels of centralisation and disagreement. This allows studying new phenomena such as knowledge crowdsourcing in new ways.

Our findings also contribute to the literature on social representations. Most importantly, while there has been a multitude of studies that examined the formation of social representations of different phenomena (e.g., Pawlowski et al. 2007; Wagner et al. 1996), none has systematically analysed the formation and change of representations by leveraging the transparency afforded by the Wikipedia platform. By applying WikiGen to large amounts of data across different Wikipedia articles, it is possible to identify and compare patterns in the evolution and change of different representations in a way that has not been previously possible.

Our study has several limitations. First, we focused on data contained in the 'article pages' of Wikipedia. However, user discussions on the 'talk pages' of the platform may contain further aspects of the social representation or even distinct representations that are not expressed in the actual article. A comprehensive application of WikiGen and associated method would require examining these pages as well. Second, the generalisability of results from two case studies is limited. Although the case studies demonstrate the applicability of the method and illustrate an effective utilisation of WikiGen, a larger sample of social representations is required to enhance the validity of our findings.

The limitations provide directions for further research. First, WikiGen can be extended to identify anchors in Wikipedia's talk pages to provide a more comprehensive analysis of the formation process of social representations, as well as of the discussions that inform this process. Second, WikiGen can be applied to identify patterns in the evolution of social representations across a group of semantically-related concepts such as a set of IT artefacts (groups of social networking websites, operating systems, etc.). Such a study could shed light on the similarities and differences in the way people understand concepts that are generally considered similar. Third, WikiGen can be applied to compare the social representations of groups of phenomena whose nature is thought to be inherently different: for instance, positive events vs. negative events, or tangible objects vs, intangible concepts.

REFERENCES

Berger, P. L., and Luckmann, T. 1966. The social construction of reality: a treatise in the sociology of knowledge, Doubleday.

Breakwell, G. M., and Canter, D. V 1993. Empirical approaches to social representations, Clarendon Press.

Duveen, G. 2000. "The Power of Ideas," in Social Representations: Explorations in Social Psychology, S. Moscovici and G. Duveen (eds.), Polity Press Cambridge.

Farr, R. M. 1993. "Common sense, science and social representations," Public Understanding of Science (2:3), pp. 189–204.

Gal, U., and Berente, N. 2008. "A social representations perspective on information systems implementation: Rethinking the concept of 'frames'," IT & People (21:2), pp. 133–154.

Habermas, J. 1984. The theory of communicative action: Reason and the rationalization of society, Boston: Beacon.

Hansen, S., Berente, N. & Lyytinen, K. (2009) "Wikipedia, Critical Social Theory, and the Possibility of Rational Discourse", The Information Society, 25, 1, pp. 38-59.

Kaltenbrunner, A., and Laniado, D. 2012. "There is No Deadline - Time Evolution of Wikipedia Discussions," CoRR (abs/1204.3453).

László, J. 1997. "Narrative organization of Social Representation," Papers on Social Representations, J. László and W.S. Rogers (eds.), pp. 155–172.

Marková, I. 1996. "Towards an Epistemology of Social Representations," Journal for the Theory of Social Behaviour (26:2), pp. 177–196.

Moscovici, S. 1984. "The phenomenon of social representations," Social representations, Moscovici, Serge Farr, R.M., p. 69.

Moscovici, S. 2008. Psychoanalysis: Its Image and Its Public, G. Duveen (ed.), Wiley.

Moscovici, S. 2000. Social representations: Explorations in social psychology, G. Duveen (ed.), Polity Press Cambridge.

Moscovici, S. 1988. "Notes towards a description of Social Representations," European Journal of Social Psychology (18:3), pp. 211–250.

Pawlowski, S. D., Kaganer, E. A., and Cater, J. J. 2007. "Focusing the research agenda on burnout in IT: social representations of burnout in the profession," European Journal of Information Systems (16:5), pp. 612–627.

Siorpaes, K., and Bachlechner, D. 2006. "Harvesting Wiki Consensus - Using Wikipedia Entries as Ontology Elements," in IEEE Internet Computing, pp. 54–65.

Suchecki, K., Salah, A. A. A., Gao, C., and Scharnhorst, A. 2012. "Evolution of Wikipedia's Category Structure," Advances in Complex Systems (15:supp01), p. 19.

Valsiner, J. 2003. "Beyond social representations: A theory of enablement," Papers on Social Representations (12), pp. 7.1–7.16.

Wagner, W., Duveen, G., Farr, R., Jovchelovitch, S., Cioldi, F. L., Marková, I., et al. 1999. "Theory and Method of Social Representations," Asian Journal Of Social Psychology (2:1), pp. 95–125.

Wagner, W., Valencia, J., and Elejabarrieta, F. 1996. "Relevance, discourse and the 'hot' stable core social representations -A structural analysis of word associations," British Journal of Social Psychology (35:3), pp. 331–351.

COPYRIGHT

[Chasin, Gal, Riemer] © 2013. The authors assign to ACIS and educational and non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full and this copyright statement is reproduced. The authors also grant a non-exclusive licence to ACIS to publish this document in full in the Conference Papers and Proceedings. Those documents may be published on the World Wide Web, CD-ROM, in printed form, and on mirror sites on the World Wide Web. Any other usage is prohibited without the express permission of the authors.