Communications of the Association for Information Systems

Volume 43 Article 37

12-2018

A Historical Perspective on Information Systems: A Tool and Methodology for Studying the Evolution of Social Representations on Wikipedia

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

Kai Riemer
The University of Sydney

Friedrich Chasin
University of Münster

Communications of the Association for Information Systems

Research Paper DOI: 10.17705/1CAIS.04337 ISSN: 1529-3181

A Historical Perspective on Information Systems: A Tool and Methodology for Studying the Evolution of Social Representations on Wikipedia

Uri Gal

The University of Sydney
Discipline of Business Information Systems
Australia
uri.gal@sydney.edu.au

Kai Riemer

The University of Sydney
Discipline of Business Information Systems
Australia
kai.riemer@sydney.edu.au

Friedrich Chasin

University of Münster
Department of Information Systems
Germany
friedrich.chasin@ercis.uni-muenster.de

Abstract:

In recent years, scholars' interest in developing historically informed explanations of information systems has surged. Several scholars have suggested that doing so can help information systems scholars to examine shifts in the academic nature of our discipline, trace the origins of prominent information systems phenomena, and reflect on and critique their own work. To enable such inquiry, we draw on the theory of social representation to build an analytical tool, WikiGen, and develop a methodology for examining the evolution of collective knowledge on Wikipedia. We demonstrate the usefulness of the tool and methodology by applying it to an illustrative case study, the Wikipedia article on cloud computing. After presenting the results of the analysis, we discuss the applicability of the tool and methodology, the contributions of our study, and possibilities for future research.

Keywords: Social Representations Theory, History, Wikipedia, WikiGen.

This manuscript underwent peer review. It was received 05/15/2017 and was with the authors for 8 months for 3 revisions. Nik Hassan served as Associate Editor.

1 Introduction

In recent years, scholars' interest in historical perspectives on information systems (IS) has surged. A growing number of scholars have explored long-standing trends, fashions, streams of research, methodologies, discourses, and objects of inquiry that have characterized the IS discipline since its inception (e.g., Baskerville & Myers, 2009; Bernroider, Pilkington, & Cordoba, 2013; Elbanna & Newman, 2013; Belanger & Carter, 2012; Porra, Hirschheim, & Parks, 2014; Zhang, 2015). Recognizing the importance of these efforts and their potential contribution to the IS discipline, the Association for Information Systems established the History Task Force in 2013 to foster the "collection, preservation, interpretation, writing, and dissemination of the historical information in and about the IS field" (Zhang, 2015, p. 478).

Several factors have spurred this interest in historical approaches. Some researchers claim that it behooves any maturing academic discipline to construct and articulate a historical account of itself in order to gain legitimacy and be recognized by external stakeholders (Bryant, Black, Land, & Porra, 2013). Others maintain that outlining the history of the IS discipline and tracing its roots can help define the discipline, delineate its boundaries, and foster a shared understanding of its core identity. Many believe that doing so will help bridge communication gaps among the discipline's varied subcommunities (Hirschheim & Klein, 2012), and others emphasize that adopting a historical approach provides an opportunity for the IS community to take a reflexive and critical stance toward its own work in order to ensure that its members do not apply ideas and theories without recognizing how they emerged and how scholars understood and used them in the past (Mitev & De Vaujany, 2012).

Importantly, historical inquiries involve more than simply "documenting a parade of personalities, events and dates, a nostalgic trip down memory lane" (Hassan, 2018). They involve describing events and structuring them in a meaningful way so as to create a coherent narrative (Hirschheim, Saunders, & Straub, 2012) or storyline that sequences events in time and creates a "plot" (Greenhalgh et al., 2005). Such "plots" should both interpret the nature of the events discussed and explain (and offer insight about) their evolution over time (Sylvester, Tate, & Johnstone, 2011).

Historical accounts can follow different paradigms; in IS, the two most notable paradigms include the objectivist, or Rankean, paradigm and the social science paradigm (Hassan, 2018). While the former follows a realist ontology and strives for an objective, rigorous, and "correct" reconstruction of past events, the latter sees the world as socially constructed and recognizes that any historical account is always subject to cultural viewpoints in its interpretation of past events. In this paper, we subscribe to the social science paradigm.

This paradigm always views historical accounts as coming from a particular vantage point rather than as "objectively accurate". Based on engagement with and interpretation of historical evidence, historians create a situated narrative of past events rather than reconstruct an account of how things happened "in actuality" (Hirschheim & Klein, 2012). The social science paradigm stipulates that, from this process, one invariably produces a contextual outcome in the sense that any historical narrative is—to a certain degree—a product of its author's culture and time and of those that characterize its subject matter.

However, we stress that, while situated and contextual, not all historical narratives are equally valid, and historians strive to construct narratives that constitute authentic representations of experienced events. Thus, when examining evidence, historians focus on considering the significance of past phenomena and events as people experienced them at the time that they occurred rather than taking things at face value or uncritically imposing their current frames of reference to understand past phenomena or events. To do so, historians must seek access to primary sources of rich data, determine whether they should consider them as historical (Hassan, 2017), and develop narratives that reflect the local significance of events. In doing so, they should bear in mind the ways in which people from different cultural, social, or demographic backgrounds experienced phenomena or events (Bryant et al., 2013).

In the IS context, historical inquiries can have two broad applications (Bryant et al., 2013). The first provides accounts that chart shifts in the academic nature of the IS discipline over time by examining the dominant theories researchers used, technologies they studied, and methodologies they employed (e.g., Grover, London, & Craig, 2016; Iivari, 2016; Vaezi, Mills, Chin, & Zafar, 2016). By observing these changes, we can gain insight into the origins of the issues that currently occupy researchers in the discipline, understand how they came to hold the position in the discipline that they currently do, and identify potential research areas.

The second application traces the origins of prominent IS phenomena and broadens our understanding of them. IS historians can account for the technological and intellectual roots of significant technologies and managerial practices, such as social networking applications, ERP systems, or business process re-engineering. They can explore the conditions for their emergence and endurance as socio-technical phenomena and explain the reasons that other phenomena have dissipated, lost traction, and disappeared along the way. Furthermore, researchers can examine how such phenomena have impacted the lives of people from different backgrounds and in different social settings, how these people have understood and used then, and how these uses and understandings have changed over time (Bryant et al., 2013).

Despite the potential usefulness of a historical approach to IS research, scholars in the discipline have not widely applied it. Some notable exceptions include the special issues on the topic in *MISQ* (Mason, McKenney, & Copeland, 1997), *JAIS* (Hirschheim et al., 2012) *JIT* (Bryant et al., 2013), and *CAIS* (Zhang, 2015, 2016), Mitev and De Vaujany's (2012) analysis of published papers on history and IS, and Porra, Hirschheim, and Parks' works (2005, 2006, 2014). These contributions notwithstanding, historical research in the IS discipline still represents a highly unconventional form of inquiry and rarely appears in our most prestigious journals and conferences.

To address this paucity in research and to echo the message that the authors above have voiced, we focus on encouraging IS academics to introduce elements of historical inquiry into their research. To do so, we demonstrate the use of Wikipedia as one potential source of data for historical inquiry into how IS phenomena have evolved over time.

Wikipedia, an online open platform, allows users to collectively create knowledge. Unlike most encyclopedias, anyone can create and edit Wikipedia articles in a collaborative process[1]. Importantly, Wikipedia stores and makes freely available detailed information about edits, editors and their discussions, and past versions of each article. Due to its accessibility and sheer size (both in terms of the quantity of information that it holds and the number of contributors who actively create and edit this information), Wikipedia serves as an excellent source of historical data.

To study online knowledge creation, we draw on the theory of social representations (Moscovici, 2000). The theory provides a rich vocabulary to describe the formation, change, and content of knowledge, and, thus, one can use it to develop historical explanations of current phenomena. The theory adopts the premise that any phenomena, objects, and events become meaningful to people only when people socially represent them. When we experience a new event or encounter an unfamiliar phenomenon, we engage in a social process of representing: a collective sensemaking to name and situate the unfamiliar in the existing stock of knowledge. Only when a group of people represent a phenomenon does it become part of the group's social reality and the group's members can use it in their practice and communication (Moscovici, 1984).

Wikipedia serves as a particular appropriate phenomenon to study the social process of representing because it does not merely contain knowledge about the world but involves social collaboration whereby people make sense of phenomena through creating and continually shaping social representations. Its design facilitates communication among multiple participants, which approximates Habermas' idea of rational discourse (Habermas, 1984; Hansen, Berente, & Lyytinen, 2009). Wikipedia's structure, which allows one to access historical data on every article, and the transparency of its collaboration processes make it ideal for exploring the social process of representing and studying how phenomena unfold historically over time.

To conduct our exploration, we developed the Wikipedia Genealogy Generator (WikiGen), a freely available, Web-based analytical tool[2]. We specifically designed WikiGen to examine the collective knowledge-creation process by tracing the history of social representations on Wikipedia. It does so by generating a set of statistical analyses based on Wikipedia data. To understand this data and its contextual significance in depth, one needs to integrate WikiGen's quantitative output with a qualitative analysis. The combination of quantitative and qualitative analyses constitutes a novel methodology for studying the historical construction of knowledge on Wikipedia. In this paper, we outline and demonstrate the methodology and its usefulness for historical analyses in IS, which also serves as the paper's main contribution.

[1] While individuals can edit most articles without needing to log in, some articles do require one to log in.
[2] See http://wikigen.org

The paper proceeds as follows: in Section 2, we review Wikipedia as a source of information for historical studies. In Section 3, we describe social representations theory and how we apply it to Wikipedia. In Section 4, we introduce WikiGen and outline the methodology we used to study the history of social representations on Wikipedia. In Section 5, we demonstrate the usefulness of the methodology by applying it to an illustrative case study on the Wikipedia article on cloud computing (CC). In Section 6, we discuss the nature of our methodology as a historical research method as Porra et al. (2014) define it and its usefulness and applicability for IS researchers. In Section 7, we outline our contributions and limitations and provide suggestions for future research. In Section 8, we summarize the paper's contributions and, in Section 9, conclude the paper.

2 Wikipedia as a Source for Historical Studies

Created in January, 2001, Wikipedia is the largest non-print encyclopedia in the world. As of December, 2017, it contains almost 47 million articles in 291 languages that more than 72 million contributors have created and edited over 2.4 billion times ("List of Wikipedias", n.d.). The English Wikipedia alone has over 5.5 million articles and over 31,000 active contributors (i.e., with more than five edits per month) ("English Wikipedia", n.d.). As of March, 2017, Wikipedia ranked as the fifth most popular site on the Internet (Alexa, 2018).

Professional historians have recognized the usefulness of Wikipedia for historical studies (Rosenzweig, 2006; Wolff, 2013; Phillips, 2016). These historians have shown considerable interest in Wikipedia as a medium for historical writing because it challenges traditional research methods: while trained historians have individually mostly authored conventional historical texts, a large number of amateur writers can and do openly contribute to Wikipedia articles (Rosenzweig, 2006). Therefore, some historians have approached Wikipedia and the historical validity of its content with a degree of skepticism and trepidation (Phillips, 2016). Some historiographers have stressed that historical articles on Wikipedia may resemble collective constructions of memory more than rigorous historical narratives; the former reflects attempts to represent the past from in a specific worldview, and the latter methodically reconstructs past events with a skeptical approach to human motive and action (Wolff, 2013).

At the same time, comparative reviews of Wikipedia have revealed that, despite some noticeable procedural differences between the online encyclopedia and some of its conventional counterparts, historical content on Wikipedia largely has a degree of accuracy similar to professionally written historical texts (Rosenzweig, 2006). While historical narratives on the platform may be synthetic, dull, and verbose, its neutral point of view (NPOV) policy helps ensure that articles on even controversial topics present differing perspectives and opinions, which lends the platform an encyclopedic voice. Consequently, Wikipedia presents comprehensive and well-balanced descriptions of historical events (Rosenzweig, 2006).

However, while the historiography discipline and professional historians have found interest in Wikipedia as a medium to study history writing, we point out that it has a broader significance for historical studies and has relevance to a variety of disciplines such as IS because it contains historical data on the creation and change of *all* of its content. Its appeal as a source for historical studies stems from fact that it allows one to study any and all articles in its repository from a historical perspective.

Wikipedia serves as a particularly useful source to conduct historical studies on IS phenomena. Such studies can examine the discipline's history and underpinning theories and/or analyze how particular phenomena of interest to IS scholars have evolved over time. IS is one of the youngest academic disciplines, and many IS researchers focus on contemporary technological trends. Their origins often fall well within the 17-year period of Wikipedia's existence; as such, one can appropriately study such trends using Wikipedia data. We further note that the kind of historical analyses that Wikipedia affords have particular significance for a discipline that engages with fast-paced phenomena, such as technological innovations, which may emerge, progress, mature, and fade over short periods. One can investigate the dynamics of such technologies not only for documentary purposes in hindsight but also to better understand the phenomenon as such.

3 Social Representations Theory

Social representations theory (SRT) adopts the premise that a layer of socially constructed and continuously evolving symbols, or representations, that serve to render the world meaningful for social actors invariably mediates people's relationship with the world (Gal & Berente, 2008). As such, SRT stands in the tradition of social constructionism, which posits that we construct our understanding of things and events in the world through processes of social negotiation that form the basis for our shared assumptions about reality (Berger & Luckmann, 1966). Concretely, SRT provides a well-elaborated set of concepts that show how exactly such a social meaning-making process unfolds.

According to the theory, things in the world do not have an inherent meaning. Rather, they acquire meaning when social groups represent them in an ongoing communicative process. Phenomena or events only become social reality by virtue of their representations that the community forms. Only when a group of people represent a phenomenon or event via familiar conceptual devices can it become a social object that people can perceive, characterize, and compare to other social objects and use in language and action (Wagner et al., 1999).

3.1 The Process of Representing

Representational activity refers to a social process that often occurs when something disruptive threatens socially shared perceptions of reality (Moscovici, 2000). This disruption can be a new and unfamiliar phenomenon or an unexpected characteristic of a familiar phenomenon that group members lack the cognitive vocabulary to describe and name. This unfamiliarity creates a sense of incompleteness and emphasizes the “actuality of something absent” (Moscovici, 2000, p. 38). To familiarize the unfamiliar, group members form new social representations in a process that has two components: anchoring and objectification.

3.1.1 Anchoring

Whenever people experience something unfamiliar, such as a radical societal change or technological innovation, they initiate a process of familiarizing via classification (Wagner et al. 1999). Classification entails positioning the 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, people anchored the unfamiliar phenomenon of HIV/AIDS in its early stages (before it acquired the name HIV/AIDS) in terms of a “gay plague” or “gay cancer” (Farr, 1993). Thus, people initially understood the HIV/AIDS phenomenon in terms of (and it took on the qualities associated with) a plague or cancer. The process of anchoring is dynamic and reflects group members’ changing perceptions of different aspects in their environment. Thus, anchors are an integral part of thinking in general: “there is no thought or perception without anchor” (Moscovici, 2000, p. 48).

3.1.2 Objectification

In addition to anchoring a new phenomenon and placing it in familiar categories, group members engage in further communicative activities that lead to an objectified representation in the form of a metaphor, symbol, or image (Wagner et al., 1999). Objectification refers to the process whereby socially represented knowledge acquires a concrete and distinct form (i.e., a representation). Objectification involves the development of a signifier that stands for the phenomenon or object that it represents (Gal & Berente, 2008). The representation captures the essence of the phenomenon and weaves it into the social fabric of the group’s common stock of knowledge. For example, that people today readily and widely recognize HIV/AIDS as a medical condition in its own right indicates that they have objectified this representation[3].

The choice of a representation is not arbitrary. It typically relates to the knowledge, vocabulary, and imagery that group members have in common and that reflect their shared identity, history, and everyday “social terrain” (Moscovici, 2000). Accordingly, different groups often develop different, sometimes conflicting, representations of the same phenomenon depending on their socio-historical contexts (Gal & Berente, 2008).

3.2 Social Representations in Wikipedia

Wikipedia is not merely an encyclopedia that holds knowledge about the world. It is also the place where this knowledge is created: a place of dynamic discourse and social engagement. In Wikipedia, contributors collaborate but also intensely debate and disagree about how to represent events, objects,

[3] “Anchoring” refers to the primary process that occurs when an event, person, or object is socially represented: that which is objectified inevitably becomes so through the process of anchoring. In turn, objectification provides the evidence that the anchoring process has occurred.

people, and other phenomena in their world. Thus, one can see Wikipedia contributors as a "decision committee" in that they "vote" and express their opinions by changing article 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 and can engage with others' contributions. In doing so, contributors build on others' efforts in order to reach, if only temporarily, a collective agreement about the meaning of the topic at hand. This description reflects the social process of representing wherein groups of people collectively and continually shape knowledge through anchoring and objectification.

To examine the history of social representations on Wikipedia, one needs to operationalize the theory and identify corresponding Wikipedia concepts. We maintain that each Wikipedia article reflects a social representation. Social representations name various aspects of the world and, thereby, orientate people and provide a basis for their communication. Wikipedia articles serve this function: they give a unique name to different domains of the world and reflect a socially negotiated meaning of that domain in a particular point in time (Hepp, Siorpaes, & Backlechner, 2007).

Anchoring and objectification processes occur in and across Wikipedia articles. One can trace anchoring by examining internal links in one article that point to other articles. Anchoring involves explaining an unfamiliar phenomenon in terms of a familiar one whose nature people deem relevant to understanding the unfamiliar phenomenon (Moscovici, 2000). On Wikipedia, internal links serve this anchoring role: they create an association between a topic that requires elaboration and explanation with other known topics that pertain to understanding the topic at hand. An internal link anchors the current representation in terms of an existing one. For example, when Apple first released the iPad, Wikipedia articles anchored it to the "iPhone" and "Tablet computer" (to mean the "iPad is like a large iPhone" and the "iPad is a particular kind of tablet computer").

However, not every internal link in an article constitutes an anchor. Wikipedia articles' scope and structure, which users define dynamically, diverge highly; different sections in an article can even include information that semantically differs to a large degree from the article's subject. However, every Wikipedia article starts with a definition section that contains information that pertains to the phenomenon of interest. As such, links in this first section will most likely anchor the topic, and links in other sections will most likely not. Consequently, changes in internal links in the definition section of an article over time will likely reflect the article's anchoring history.

Objectification indicates that socially represented knowledge gains a distinct and concrete form. One can best operationalize objectification in Wikipedia as the act of linking other articles to the article at hand. Thus, one can use Wikipedia articles that contain a reference to a particular social representation in the form of a link to trace the objectification process for that social representation. A growing number of links that point to a particular social representation would reflect an ongoing objectification process because it would indicate that other phenomena are explained (i.e. anchored) in terms of the representation under study.

4 WikiGen: An Analytical Tool for Studying Social Representations on Wikipedia

As we state in Section 1, in order to trace the social process of representing in Wikipedia, we developed a Web-based analytical tool called the Wikipedia Genealogy Generator or WikiGen for short. By connecting to Wikipedia's live databases over the platform's application programming interface (API), WikiGen generates collaboration, anchoring, and objectification statistics based on the historical revisions of any chosen article. These statistics help one to study how social representations emerge and evolve over time on Wikipedia. We briefly outline them below[4]:

Collaboration statistics trace the following over time in a given article:

- The number of edits
- The number of editors, and
- The number of edits per editor.

[4] In Appendix A, we describe these statistics in detail and provide a comprehensive WikiGen user manual.

Anchoring statistics track both the strength of individual anchors and the dynamics of multiple anchors over time.

The strength of individual anchors includes the following metrics for a chosen period:

- The number of days an anchor has existed
- The number of revisions an anchor has survived
- The number of anchor re-introductions, and
- Anchor strength—a linear combination of days present and revisions survived.

The dynamics of multiple anchors include the following metrics for a chosen period:

- The number of new and removed anchors
- Anchor dissimilarity—a score that gauges the level of dissimilarity between anchors present in two consecutive periods
- Anchor durability—a score that gauges the level of durability of anchors across two consecutive periods, and
- Edit-war level—a score that gauges the level of disagreement between editors in a period.

Objectification statistics count the variation in the number of links that lead back to a chosen article from other Wikipedia articles.

4.1 WikiGen as a Temporal Information Analysis Tool

WikiGen follows other analysis tools that researchers have developed in the past to examine online collaborative processes in general and/or collaborative interactions on Wikipedia in particular. Researchers have also used other tools and techniques over the years to investigate the social dynamics in Wikipedia and their temporal evolvement over time.

For instance, Viégas, Watternberg, and Dave (2004) created the history flow-visualization tool that can show relationships between multiple Wikipedia article versions to surface patterns of collaboration and conflict. By applying the tool to over 70 Wikipedia article histories, the authors revealed patterns of vandalism, negotiation, and content stability (Viégas et al., 2004).

Flöck and Acosta (2015) developed whoVIS, a Web tool for investigating the collaborative writing process on Wikipedia. whoVIS can track several dynamics such as which editors make changes to an article, the degree to which editors belong to two opposing groups, and authorship distribution among editors.

Borra et al. (2015) developed Contropedia to analyze interaction patterns on Wikipedia. By associating edits to sentences that contain links across periods, Contropedia can generate a controversy score to illuminate patterns in how controversial topics have developed in Wikipedia articles.

Similarly, DeDeo (2016) used finite-state modeling of interactions between Wikipedia editors to examine the logical structure of conflict and cooperation on the platform. By applying hidden Markov models to 62 of the most edited articles on Wikipedia, he identified a punctuated, decentralized, and extended war/peace process characterized by editors' ongoing revert actions.

Keegan, Lev, and Arazy (2016) used sequence analysis to examine temporal dynamics in online collaboration. To do so, they related different types of Wikipedia article edits (e.g., content addition, text rephrase, and vandalism) to multiple editors over time in a given article, which enabled them to identify several behavioral patterns around content production on Wikipedia.

Finally, Twyman, Keegan, and Shaw (2017) developed a Python script to examine the revision history of Wikipedia pages related to the black lives matter (BLM) movement. By doing so, they traced behavioral patterns that unfolded on the BLM-related Wikipedia pages and that reflected how people constructed collective memory and mobilized collective action.

These studies share an analytical focus on how temporal dynamics have unfolded on Wikipedia. Using various tools and statistical techniques, these studies focus on surfacing structural or thematic patterns that underlie the process of knowledge construction on Wikipedia. Therefore, to the extent that one considers Wikipedia a historical encyclopedia in that it provides more historical content than any other website or traditional encyclopedia (Wolff, 2013), one can apply these temporal information-analysis tools and techniques to facilitate historical research. They do so by allowing their users to examine, from different angles and at different levels of granularity, the social dynamics involved in constructing historical narratives.

While the above studies do not make historical inquiries their explicit objective, we developed WikiGen expressly to facilitate historical research. Our use of SRT to inform the design of the tool, and our focus on links as anchors, which emphasizes the temporal dimension in the social construction of historical knowledge, makes WikiGen particularly suitable for conducting historical inquiries on Wikipedia. In Section 5, we outline our proposed methodology for facilitating such inquiries.

5 A Methodology for Studying Social Representations on Wikipedia

In this section, we describe how one can use WikiGen to conduct a historical study of Wikipedia articles. In doing so, we describe how a qualitative analysis can complement the tool's quantitative output to produce a robust methodology for studying the history of social representations on Wikipedia.

One cannot fully understand the evolution of social representations with the quantitative data that WikiGen provides alone for two reasons. First, one can interpret the meaning of each anchor that WikiGen identifies only in the context in which it is used and in relation to other present anchors. While WikiGen quantifies the presence of anchors in a chosen period and tracks changes in anchors, it cannot reveal their contextual meaning. Therefore, one needs to conduct a qualitative analysis to semantically group anchors together in order to understand the broader concepts that anchor a representation. Second, to understand how a social representation changes over time, one should trace shifts in the composition of anchor categories and their dominance over time to identify and describe distinct stages in a representation's evolution. While WikiGen can help one to identify the start and end points of particular stages (e.g., by using anchoring dynamics and collaboration measures), this identification alone does not allow one to understand the meaning and significance of stages.

Figure 1 (next page) outlines the steps in our methodology for studying social representations on Wikipedia. We elaborate on them below.

5.1 Timeframe Selection and Collaboration Analysis

5.1.1 Timeframe Selection

To study social representations, one first delineates the timeframe for analysis, which one can do in two ways. First, once one has chosen a Wikipedia article, one can analyze its entire lifespan (from the day of its creation to the present time). One can do so when one has an interest in a particular person, event, or phenomenon and wishes to observe changes in its representation over time. One can perform this kind of inquiry for articles on people or events whose existence predated Wikipedia (e.g., "Johann Sebastian Bach" or "World War II") and for articles that describe contemporary events or phenomena (e.g., the 2016 "United States presidential election", "Brexit", or "Blockchain").

One can also focus on particular periods in the evolution of a representation around which significant changes have occurred—changes that researchers believe will be reflected in substantive representation shifts. For example, researchers may choose to investigate the way in which the article on LinkedIn has changed after Microsoft acquired it in June, 2016.

5.1.2 Collaboration Statistics

One can use collaboration statistics to identify periods of significant representational change. SRT postulates that heightened representational activity will occur around times of change as people attempt to familiarize unfamiliar events or phenomena. These can be estimated by using collaboration statistics. Specifically, we expect that, during periods of change, we will observe an increased number of edits and editors.

5.2 Anchor Categorization and Coding

The next step involves categorizing the anchors that have existed in the article throughout the timeframe one has chosen to analyze. This process comprises two steps: 1) anchor cleaning and 2) anchor categorization and coding.

Step	Type	Details	WikiGen utilization
1. Timeframe selection and collaboration analysis			
1.1. Timeframe selection	Qual.	Two options: analyze the entire lifespan of the article; analyze a limited timeframe around times of change.	--
1.2. Collaboration analysis	Quant.	Examine edits / editors / edits per editor statistics to identify periods of change.	Collaboration statistics
2. Anchor categorization and coding			
2.1 Anchor cleaning	Mixed	Generate list of anchors (quant); Remove irrelevant, duplicate, and weak anchors (qual).	Anchor table
2.2. Anchor categorization and coding	Qual.	Examine the contextual meaning of each valid anchor across all versions of the article; Iteratively allocate anchors into semantically homogeneous categories.	Anchor table; Anchor map
3. Identification of Anchor Evolution Stages			
3.1. Quantitative indications	Quant.	Identify variations in quantitative statistics that reflect significant shifts in anchoring and collaboration patterns.	Anchoring dynamics statistics; Collaboration statistics
3.2. Qualitative interpretation of quantitative indications	Qual.	Examine and interpret changes in anchors around time periods identified by quantitative indications to determine if they reflect semantic shifts and transition between evolution stages; Construct a narrative that describes each stage as well as differences between them.	Anchor table; Anchor map
4. Objectification			
4.1. Objectification analysis	Quant.	Examine objectification pattern (increase / decrease over time); Identify change to pattern around transition in anchoring stages	Objectification statistics

Figure 1. Steps in Methodology for Studying Social Representations on Wikipedia

5.2.1 Anchor Cleaning

After WikiGen has identified anchors, one needs to delete the irrelevant anchors, merge the duplicate anchors, and remove the weak anchors. Anchors may be irrelevant for various reasons. For example, the definition section of an article may contain a quote from an external PDF document. In such a case, a user might have linked the word “PDF” to the external file even though PDF does not anchor the representation under study. Duplicate anchors occur when multiple anchors with different spelling link to articles with the same meaning (e.g., “Web Service” and “Web Services”). In cases where WikiGen identifies a particularly high number of anchors, one can remove weak anchors by using the anchor strength statistic.

5.2.2 Anchor Categorization and Coding

One should qualitatively examine the text passages in which each valid anchor appears to establish their contextual meaning. Anchors are identified in the WikiGen anchor resilience table[5] (Figure 2).

[5] WikiGen can generate anchor resilience tables for both months and years. Figure 2 is an anchor resilience table for 2016.

Anchor	Days survived	Revisions survived	(Re)Introductions	Anchor strength
internet	30.00	52	0	1.00
software as a service	30.00	52	0	1.00
virtualization	30.00	52	0	1.00
salesforce	30.00	52	0	1.00
google apps	30.00	52	0	1.00
platform as a service	30.00	52	0	1.00
infrastructure as a service	30.00	52	0	1.00
intranet	30.00	52	0	1.00
mainframe computer	30.00	52	0	1.00
rackspace	30.00	52	0	1.00

Month: June

Anchor	Days survived	Revisions survived	(Re)Introductions	Anchor strength
internet	30.74	69	1	0.98
computer network	30.74	69	1	0.98
utility computing	28.64	54	1	0.84
computing	28.64	54	1	0.84
service (economics)	28.64	54	1	0.84
product (business)	28.64	54	1	0.84
electrical grid	28.64	54	1	0.84
hybrid cloud	11.83	18	1	0.32
software as a service	2.36	16	0	0.15
virtualization	2.36	16	0	0.15

Month: July

Figure 2. WikiGen Anchor Resilience Tables in June and July, 2014, for the CC Article[6]

The tool links each anchor in the table to an "anchor map" that displays the periods in which the chosen anchor existed and did not exist in the article (Figure 3). Each dot in the map represents a point in time when someone either introduced or removed the anchor from the article. Clicking on a dot connects the user to the version of the article at that point in time.

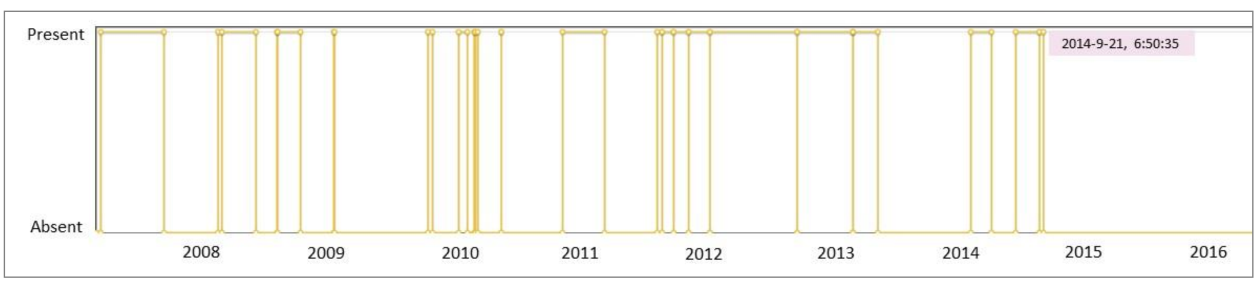


Figure 3. WikiGen Anchor Map for the Anchor Utility Computing in the CC Article

The qualitative process entails examining each version of the article in which the anchor existed. One does so by clicking on each dot on the top part of the map that marks the beginning of a period in which the anchor existed. These dots signify that someone has introduced or re-introduced the anchor to the article. For example, as Figure 3 shows, someone last introduced the utility computing anchor to the CC article on 21 September, 2014. Dots on the bottom of the map indicate anchor removals.

Based on this examination, one then places the anchors into semantically homogeneous categories in a process of theoretical coding and classification (Glaser, 1978), which entails assigning each anchor in the anchor resilience table either to existing categories or to new ones depending on their contextual meaning. Thus, one assigns and reassigns valid anchors to categories in several iterations until one

[6] We highlight anchors that existed in June and but not in July in red. We highlight anchors that existed in July but not in June in green

derives a minimal set of collectively exhaustive categories. For example, the social representation Xbox 360 contains multiple and semantically different anchors: the anchors PlayStation 3 and Wii describe competitors of Xbox 360, whereas the anchors 802.11 b/g/n and TOSLINK S/PDIF describe its technological features. Thus, one may subsequently place these identified anchors into the competitors and technological features categories or use them to create new ones.

When the meaning of an anchor varies across multiple revisions of an article, one needs to consider all different meanings. For instance, in the context of the iPad article, the anchor iPhone may describe similar products to the iPad in one version of the article (“a device with which iPad shares some features”) or emphasize a difference between iPad and other products in another version (“iPad is different to iPhone because it has a larger screen”). In this case, one might place the anchor iPhone in two categories.

5.3 Identification of Anchor Evolution Stages

Identifying evolution stages comprises two interlaced components (i.e., quantitative and qualitative) that one should use iteratively. The quantitative component relies on WikiGen to identify statistical indications that are likely to demarcate semantically distinct evolution stages in the article. One should verify these indications through qualitatively analyzing changes in the meaning of anchors in each evolution stage. Insights from the qualitative analysis may lead to one to conduct further quantitative analyses to reconsider the timeframes for the evolution stages. The process stops when the qualitative analyses provide results that concur with the quantitative indications, which we explain in detail below.

5.3.1 Quantitative Indications

Quantitative statistics that indicate significant shifts in anchoring and collaboration patterns help one trace anchor evolution stages. Such shifts may reflect distinct stages in the evolution of the social representation under study. Anchoring dynamics statistics are particularly important for tracing possible points in time where a transition between stages occurs:

- Extensive addition or removal of anchors may indicate a substantial change in the anchoring of the representation and requires further qualitative analysis.
- High anchoring dissimilarity means that the representation has a high number anchors in a given period that differ from the anchors that existed in the preceding period. Therefore, it warrants further qualitative analysis.
- Low anchoring durability indicates that, due to a high number of new or removed anchors, the average time that anchors are present in the article decreases. Therefore, it warrants further qualitative analysis.
- A high edit-war score signifies disagreement among contributors regarding the anchors in the article. Intense disagreements may indicate semantic changes to the representation and warrants further qualitative analysis.

In addition to anchoring stability statistics, one may also use collaboration statistics to detect possible shifts in evolution stages:

- Increased number of edits, editors, and edits per editor reflect increased efforts by contributors to (re-)articulate the representation under study. Therefore, such periods warrant further qualitative analysis.

5.3.2 Qualitative Interpretation of Quantitative Indications

Unusual quantitative measures do not necessarily imply a semantic change and transition between evolution stages. New anchors may simply replace old ones with a similar meaning. Thus, quantitative indications can result from an inconsequential restructuring of an article that does not impact its meaning. Therefore, one needs to examine the quantitative indicators in the context of the identified anchor categories to see whether they reflect a change in the meaning of the social representation in a given period. Such a change would manifest as: 1) an increase or decrease in the dominance of one or several categories (i.e., as an increase or decrease in the number or strength of anchors in one or several categories), or as 2) the disappearance of old or the appearance of new categories.

WikiGen’s anchor resilience table (see Figure 3) provides a particularly useful way to trace the changes in anchors’ strength over time. The table tracks monthly (and yearly) changes in anchors and, thus, enables one to compare the composition and strength of multiple anchors over different periods. In this way, one can observe changes in anchors and their strength and in the contextual meaning of each anchor (by following the links from each anchor to an anchor map as we explain above) during periods that quantitative indications have previously identified as potentially transitional.

Qualitatively examining the identified differences in the representation across evolution stages results in a historical narrative that describes each evolution stage and the differences between them[7]. These stages reflect the genealogy of knowledge in the sense that they illuminate the (changing) relationship between the social representation under study and those from which it inherits its characteristics.

5.4 Objectification Analysis

In the last part of the methodology, one examines objectification level of the social representation by using the WikiGen statistics for the number of Wikipedia articles that contain references to the social representation under study. In particular, we focus on the pattern of objectification, the rate of its increase or decrease over time, and observed changes to this pattern in and around times of transitions between evolution stages.

6 Illustration of the Methodology for Studying the History of Social Representations

To demonstrate the methodology's usefulness for studying social representations, we applied it to the cloud computing (CC) Wikipedia article. We chose CC because it is a contemporary IS phenomenon that has garnered significant attention from academics and practitioners alike as its corresponding Wikipedia article evidences. Since its creation in March, 2007, 4,374 contributors have edited the article 9,903 times and people have viewed it over 21.7 million times. Furthermore, some of the biggest players in the IT world such as, Amazon, Salesforce, Microsoft, Oracle, and Google engage in providing different CC applications, which indicates that CC is an important issue for many of today's large IT and business organizations.

6.1 Timeframe Selection and Collaboration Analysis

6.1.1 Timeframe Selection

Because we focus on illustrating how one can use the methodology rather than answering specific research questions concerning the history of CC, we chose to examine the article and observe changes in the way CC was represented over the first four years of its existence (an arbitrary timeframe rather than the entire lifespan of the article or a period during which significant real-life events that may have influenced the CC article occurred). Accordingly, our investigation ranged from March, 2007, when the CC Wikipedia article was created, until February, 2011.

A contemporary phenomenon, CC entered the general discourse (outside of Wikipedia) around the same time that the corresponding Wikipedia article was created, which we approximated by observing the search popularity for the term "cloud computing" in Google Trends (Figure 4), which indicates that searches for the term started in early 2007 and peaked around mid-2011.

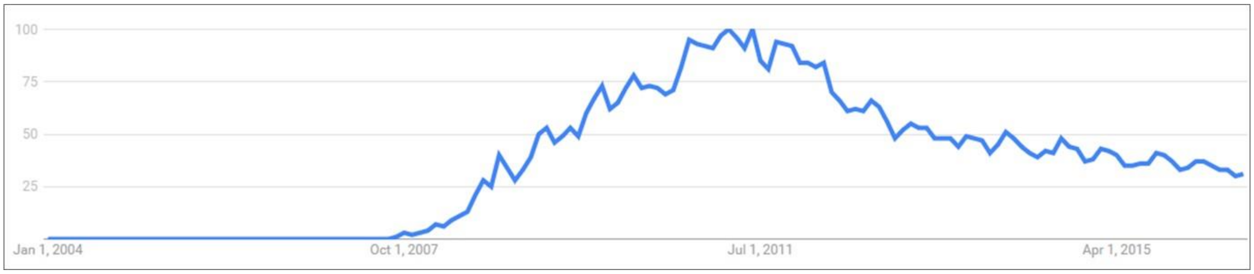


Figure 4. Search Popularity for "Cloud Computing" in Google Trends

---
[7] If it is not possible to qualitatively interpret the identified evolution stages, one needs to perform another iteration using a different quantitatively derived set of timeframes for evolution stages.

6.1.2 Collaboration Analysis

Collaboration statistics present the distribution of edits and editors over the period we examined. We can observe a couple of corresponding peaks in the number edits and editors around July, 2008, and January, 2010 (Figures 5 and 6). These peaks indicate heightened representational activity, which may signify changes in the way that people collectively understood CC on Wikipedia.

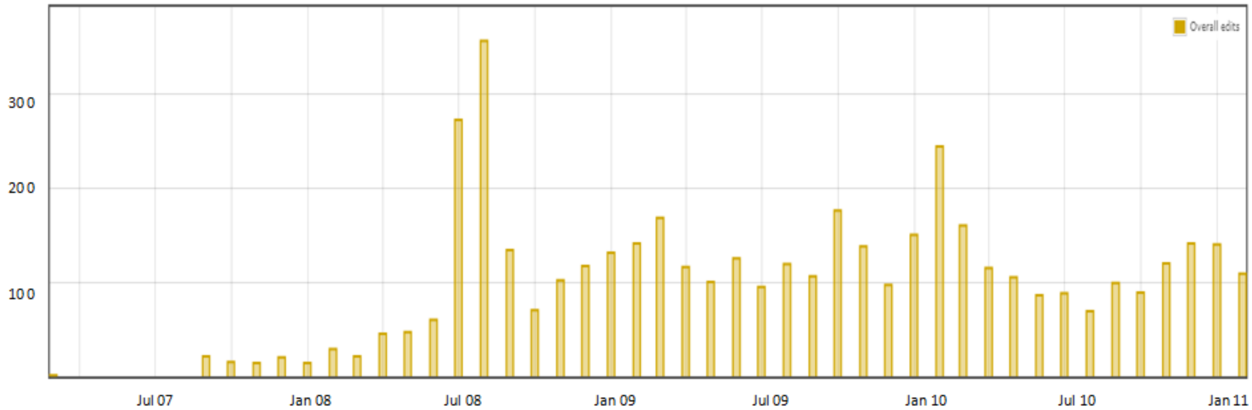


Figure 5. Distribution of Edits During the Lifespan of the CC Article

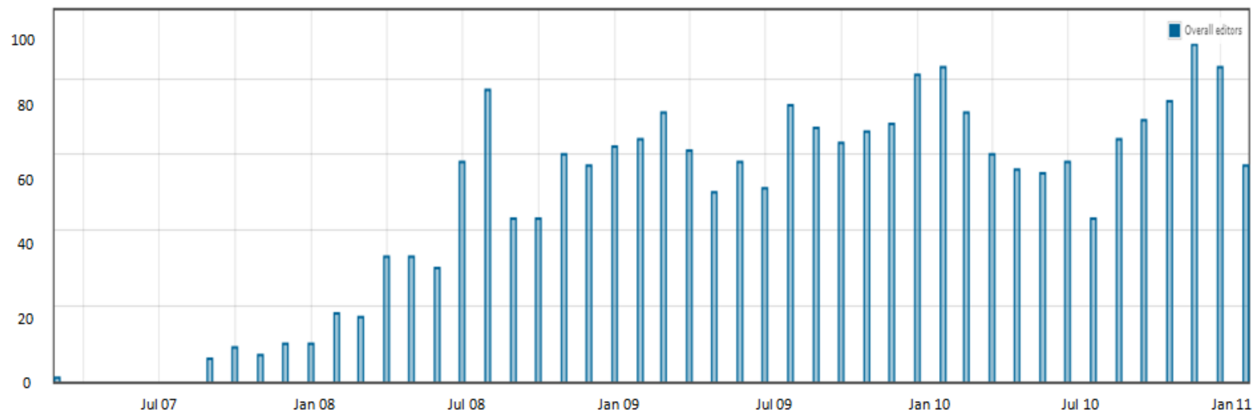


Figure 6. Distribution Of Editors During The Lifespan Of The CC Article

6.2 Anchor Categorization and Coding

6.2.1 Anchor Cleaning

WikiGen initially identified 223 anchors. Given the high number of anchors (and the high number of relatively weak anchors), we set an anchor strength of 0.15 per individual year as a threshold below which we excluded anchors from further analysis. Of the remaining 122 anchors, we removed three (pdf, nist, gartner) as they did not pertain to the CC representation. We also identified 12 additional anchors as duplicates due to different spellings and subsequently removed them (e.g., “Yahoo” and “Yahoo!”). The final list included 107 anchors. The anchor strength comparison table in (which we display in Figures 7 and 8 below) show the 25 strongest anchors ordered by their total strength (a cumulative measure of anchor strength for each year from 2007 to 2016).

Anchor	Strength in 2007	Strength in 2008	Strength in 2009	Strength in 2010	Strength in 2011	Strength in 2012
internet	0	0.72	0.81	0.99	0.35	1.00
computing	0	0.53	0.29	0.98	0.53	0.99
computer network	0	0	0	0	0.57	1.00
electrical grid	0	0	0	0	0.28	0.81
infoworld	0	0	0.89	0.99	0.22	0.05
economies of scale	0	0	0	0	0	0.67
web browser	0.29	0.45	0.99	0.99	0.39	0.97
utility computing	0.25	0.47	0.04	0.26	0.53	1.00
converged infrastructure	0	0	0	0	0	0.97
shared services	0	0	0	0	0	0.97
software as a service	0	0.70	0.90	0.11	0.09	0.74
service-oriented architecture	0	0	0	0.15	0.22	0
hardware virtualization	0	0	0	0.06	0.22	0
autonomic computing	0	0.21	0	0	0.10	0
data center	0.29	0.66	0	0.34	0	0.01
software	0.29	0.45	0.99	0.99	0.22	0.33
platform as a service	0	0	0.73	0.11	0	0.53
server (computing)	0	0	0.54	0.99	0.26	0.58
infrastructure as a service	0	0	0.60	0	0	0.52
application software	0	0	0	0	0.04	0.97
data	0	0.45	0.99	0.99	0.22	0
computer network diagram	0	0.25	0.99	0.99	0.22	0
distributed computing	0.14	0.27	0	0	0	0
scalability	0	0	0.94	0.99	0.22	0.05
mobile app	0	0	0	0	0	0.92

Figure 7. Anchor Strength Comparison Table Showing the Strongest 25 Anchors for the CC Article

Anchor	Strength in 2013	Strength in 2014	Strength in 2015	Strength in 2016	Strength (sum)	Strength (sum, adjusted)
internet	0.99	0.76	0	0.05	5.67	2.909
computing	0.94	0.56	0	0	4.82	2.550
computer network	0.99	0.81	0.34	0	3.71	2.532
electrical grid	0.37	0.22	0.45	0.82	2.95	2.286
infoworld	0	0	0.45	0.83	3.43	2.038
economies of scale	0.37	0	0.45	0.83	2.32	1.896
web browser	0.37	0	0	0	4.45	1.848
utility computing	0.37	0.40	0	0	3.32	1.679
converged infrastructure	0.37	0	0.45	0.34	2.13	1.586
shared services	0.37	0	0.45	0.34	2.13	1.586
software as a service	0.18	0.50	0	0	3.22	1.469
service-oriented architecture	0	0	0.45	0.83	1.65	1.405
hardware virtualization	0	0	0.45	0.83	1.56	1.369
autonomic computing	0	0	0.45	0.83	1.59	1.327
data center	0	0	0.16	0.83	2.29	1.277
software	0.13	0.04	0	0	3.44	1.243
platform as a service	0.18	0.50	0	0	2.05	1.107
server (computing)	0	0.02	0	0	2.39	1.052
infrastructure as a service	0.18	0.50	0	0	1.80	1.018
application software	0.37	0	0.15	0	1.53	0.996
data	0	0	0	0	2.65	0.893
computer network diagram	0	0	0	0	2.45	0.853
distributed computing	0.61	0.43	0	0	1.45	0.839
scalability	0	0	0	0	2.20	0.818
mobile app	0.37	0	0	0	1.29	0.811

Figure 8. Anchor Strength Comparison Table Showing the Strongest 25 Anchors for the CC Article

6.2.2 Anchor Categorization and Coding

Consistent with the process that we outline in Section 4.2.2, we used WikiGen's anchor resilience table and anchor map functionalities to qualitatively code the 107 anchors to group them into the following categories:

1) The category impact on IT practice contained seven anchors that described the various practices, products, and services that CC has influenced. Anchors such as product and software demonstrated that something that was previously delivered as a software product became a service delivered through CC. Similarly, the anchor client-server described the paradigm shift from using rich clients towards using thin clients or Web browsers in order to access IT resources. Figures 9 to 11 concretely illustrate the coding process for the anchor software (we followed the same process for all other 106 anchors). Its listing in the anchor resilience table (Figure 9) links to the anchor's anchor map (Figure 10).

Show 100 entries Search:

Anchor	Days survived	Revisions survived	(Re)Introductions	Anchor strength
internet	250.62	475	5	0.72
software as a service	260.00	440	5	0.70
data center	240.44	417	5	0.66
computing	157.37	397	5	0.53
everything as a service	155.66	392	4	0.52
web application	212.39	244	3	0.48
utility computing	179.94	286	4	0.47
virtualization	178.95	284	5	0.47
web 2.0	148.99	341	3	0.47
software	145.10	321	3	0.45
web browser	145.10	321	3	0.45
google apps	145.10	321	3	0.45

Figure 9. Anchor Resilience Table for the CC Article (Ordered by Anchor Strength) with the Anchor Software Highlighted

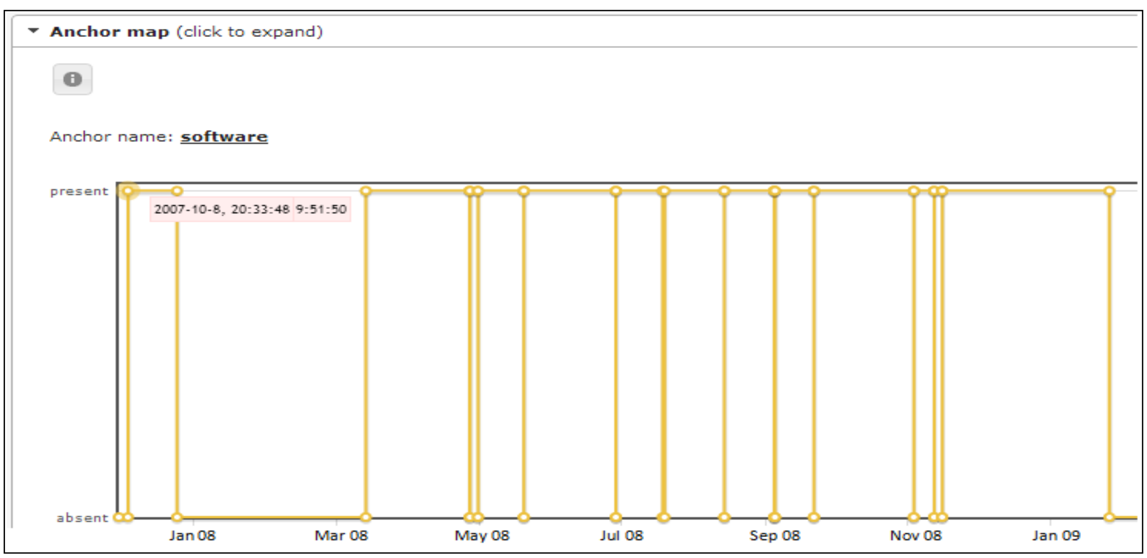


Figure 10. Anchor Map for the Anchor Software in the CC Article

Dots on the top part of the map indicate when the anchor re-entered the article (as we explain in Section 4.2.2). Clicking on these dots links to the version of the article in that point in time. Thus, clicking on the dot in which someone first introduced the anchor links to the article as it appeared on 8 October, 2007 at

8:33 pm (Figure 11). In the article, we can see that the anchor software (highlighted in yellow) serves as an example of the impact of CC on IT practices[8].

Figure 11. The CC Article on 8 October, 2007

2) The category figurative aspects contained six anchors that used metaphors, abstraction, or analogies to describe CC. For instance, the anchors electrical grid and electricity served as analogies for resource delivery in CC that drew on a perceived similarity between service delivery in CC and the way electricity is delivered through electrical networks.

3) The category relationship to broader concepts contained 13 anchors that illustrated the broad scope of CC, such as Internet and computing. Both anchors described CC as anything that involves computations on the Internet. Other anchors such as shared services or converged infrastructure served to position CC as a consequence of global trends in IT. Anchors such as services or utility extended CC even further by generalizing it to any possible services or resource delivery type over the network.

4) The category technical aspects contained 37 anchors that described various technical components and functionalities of CC or technologies used in CC. Some examples include infrastructure as a service, software as a service, data as a service remote server, and parallel computing.

5) The category providers and users contained 33 anchors to examples of corporations that provided (e.g., Google, NetSuite, and Salesforce) or used (e.g., General Electric, L'Oréal and Procter & Gamble) CC.

6) The category benefits contained five anchors that focused on the financial and operational advantages that can result from CC. Some examples include economies of scale (which are inherent for CC), capital expenditure (which emphasizes the possibility to forego capital expenditures and only pay for the use of resources), and quality of service (which stresses the quality of commercial CC solutions that legal agreements guarantee).

Despite the differences between them, the first three anchor categories bear a semantic resemblance. Each of them extended the scope of CC either by outlining its impact on other IT phenomena or by using broad lingual metaphors or concrete examples to describe it. Therefore, we combined them to create one category that we called generalizing. Thus, the final list included four anchor categories: 1) generalizing, 2) technical aspects, 3) providers and users, and 4) benefits.

6.3 Identification of Anchor Evolution Stages

Next, we examined the fluctuations in anchors from the different categories throughout the first four years of the CC article in order to understand how the CC social representation changed over time and to identify different stages in its history. We conducted the quantitative and qualitative aspects of this

[8] Note that, in the vast majority of cases, we did not find any semantic differences in anchors across multiple versions of the article. Thus, for the most part, anchors did not move across different categories.

examination (see Sections 5.3.1 and 5.3.2) concurrently as we explain in Section 4.3 above. We outline the evolution of stages below.

6.3.1 CC as Utility Computing (March to August, 2007)

The CC article was created on 3 March, 2007. The first version of the article explained CC in terms of a single anchor: utility computing (part of the technical aspects category). At that time, utility computing was defined on Wikipedia as: “[a] business model whereby computer resources are provided on-demand and on pay-per-use basis” and couched it in technical anchors, such as on-demand computing and grid computing. Thus, CC was predominantly understood in technical terms.

6.3.2 An Elaboration of the Technical Features of CC (September, 2007, to July, 2008)

In September, 2007, contributors started to more fully describe CC in a way that went beyond the utility computing concept and that made CC more concrete by specifying its technical features. Their contributions introduced anchors from the technical aspects category, such as Web application, Web browser, and rich Internet application.

In the following months, anchors from the technical aspects category continued to dominate the CC representation. Contributors used grid computing, autonomic computing, and distributed computing to explain the technical nature of CC in a more refined way. One anchor in particular (software as a service) underlined one of the most important aspects of CC at the time; namely, the cloud’s service orientation and the software migration from clients’ devices into the Web. Additional anchors from the technical aspects category during this stage included computer cluster, multi-core, and parallel computing.

This anchoring activity had quantitative manifestations that one can see in the introduction and removal of anchors, which started to increase in September, 2007 (see Figure 12), and in the dissimilarity score that started to fluctuate in the same month (see Figure 13).

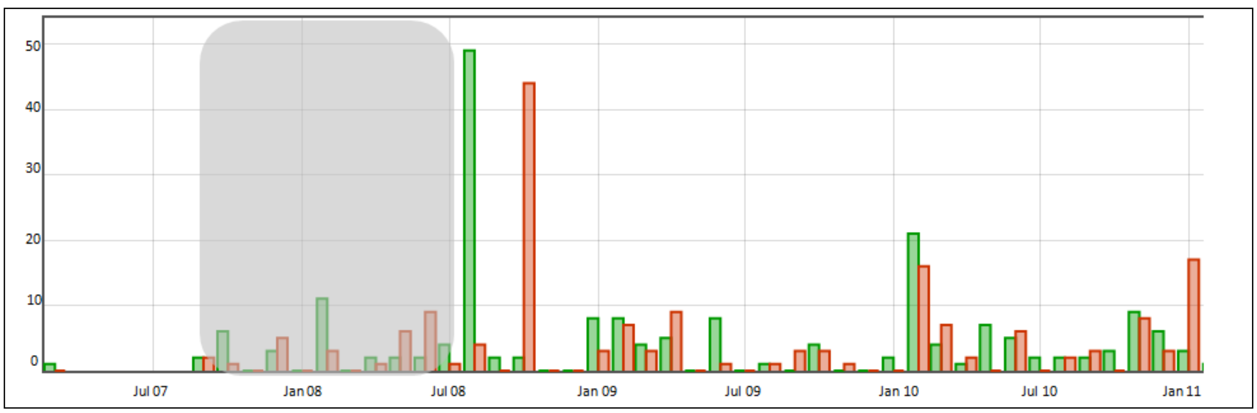


Figure 12. New and Removed Anchors for CC

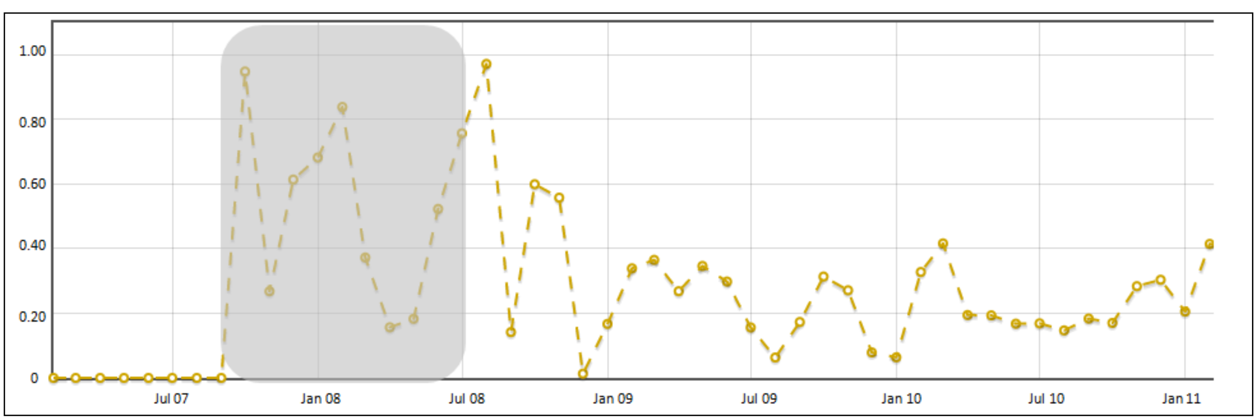


Figure 23. Anchoring Dissimilarity for CC

6.3.3 Application of CC (August to October, 2008)

Starting in August, 2008, contributors introduced multiple new anchors from the providers and users category to provide examples of different companies that used or provided CC applications (e.g., Google, Salesforce, IBM, Microsoft, and General Electric). One can see these anchors in the anchor resilience table for August, 2008 (highlighted in yellow in Figure 14).

Month: August

Show 100 entries Search:

Anchor	Days survived	Revisions survived	(Re)Introductions	Anchor strength
internet	31.00	119	0	1.00
the cloud	31.00	119	0	1.00
utility computing	31.00	117	1	0.99
data center	31.00	117	1	0.99
virtualization	31.00	117	1	0.99
computing	31.00	117	1	0.99
everything as a service	31.00	117	1	0.99
quality of service	31.00	117	1	0.99
service level agreement	31.00	117	1	0.99
pdf	31.00	117	1	0.99
capital expenditure	31.00	117	1	0.99
utility	31.00	117	1	0.99
electricity	31.00	117	1	0.99
subscription	31.00	117	1	0.99
multitenancy	31.00	117	1	0.99
google	31.00	117	1	0.99
salesforce	31.00	117	1	0.99
yahoo!	31.00	117	1	0.99
hewlett packard	31.00	117	1	0.99
ibm	31.00	117	1	0.99
microsoft	31.00	117	1	0.99
general electric	30.58	110	2	0.96
l'oréal	30.58	110	2	0.96
valeo	30.58	110	2	0.96
grid computing	28.57	94	1	0.86

Figure 34. Anchor Resilience Table for August, 2008

One can clearly observe the introduction of these anchors in the new and removed anchors graph (the tall green bar in Figure 15) and in the dissimilarity graph (see Figure 16), which peaks in August, 2008, to indicate that anchors in the article in that month almost completely differed from the anchors in it in July. The heightened collaboration activity around that time also evidences these changes (see Figures 4 and 5).

At the end of the stage in October, 2008, a high number of anchors disappeared, which included all the providers and users anchors that the article contained earlier in this stage with the exception of Google apps. The tall red bar in Figure 15 represents these changes. One can also observe them in the low anchor durability score in October, which was both preceded and followed by more stable periods (see Figure 17).

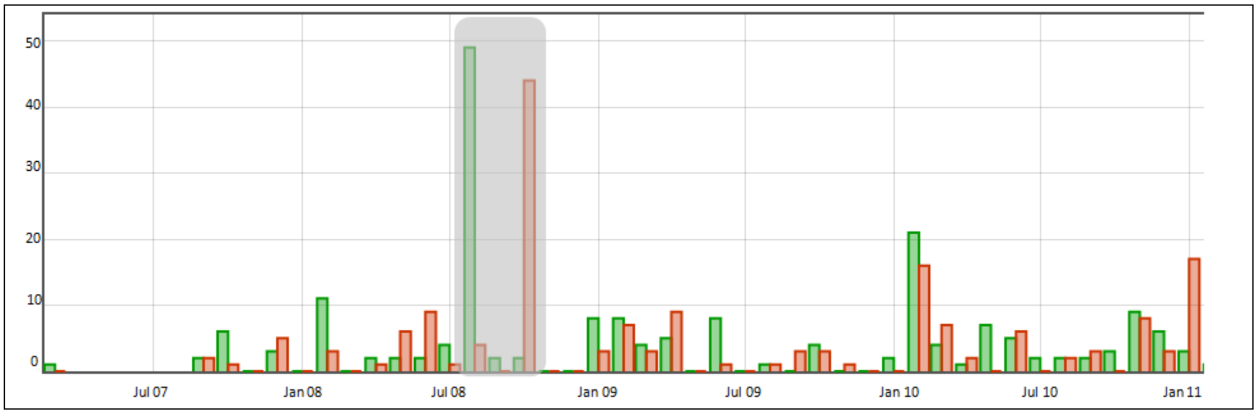


Figure 45. New and Removed Anchors for CC

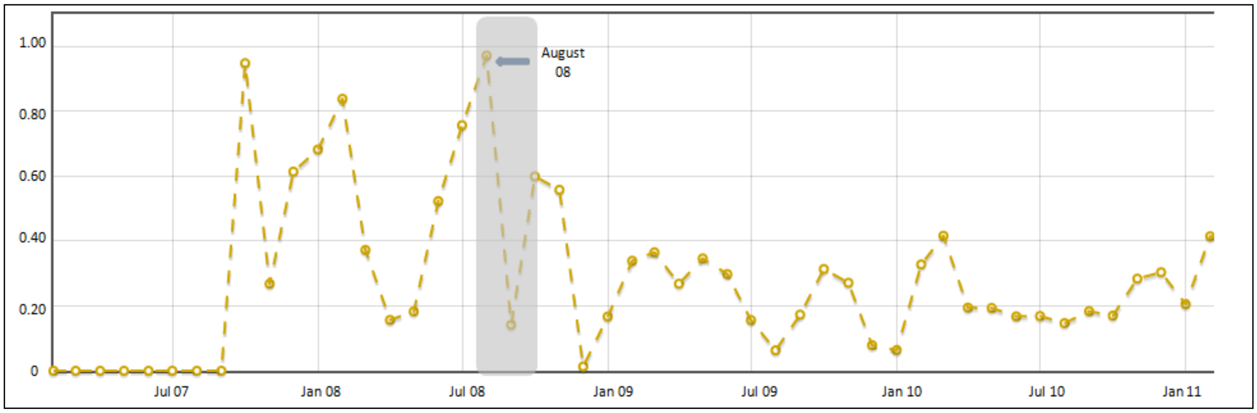


Figure 56. Anchoring Dissimilarity for CC

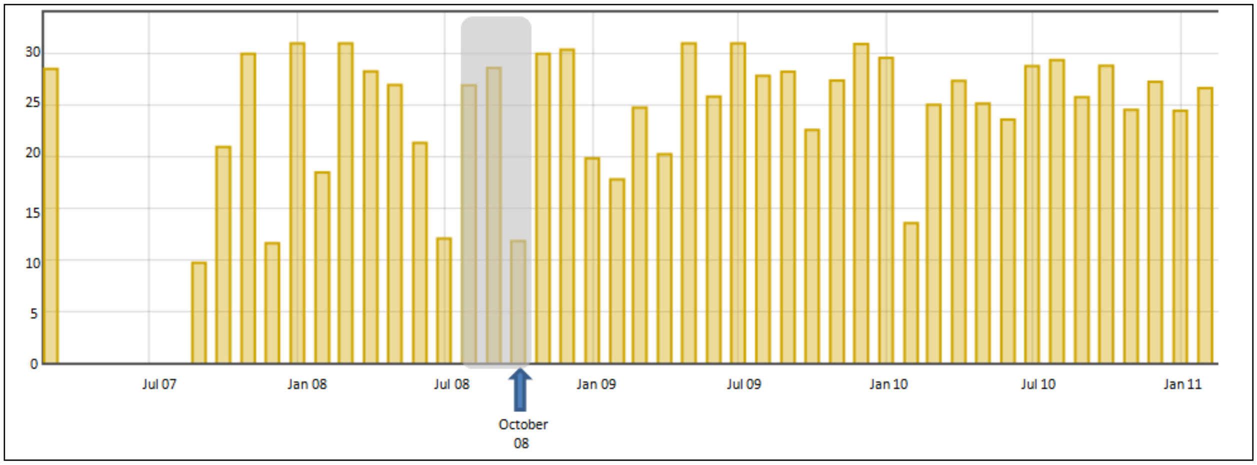


Figure 67. Anchoring Durability for CC

6.3.4 Generalizing CC (November, 2008, to January, 2010)

After self-referential anchoring in the second stage (i.e., CC explained in terms of its technical features) and an introduction of anchors from providers and users category in the third stage, the fourth stage further broadened the scope of CC. For instance, one can see that contributors introduced anchors from the generalizing category, which emphasized the relationship of CC to other concepts and practices and its impact on them.

While contributors introduced some generalizing anchors in the early months of 2008 (i.e., Internet and computing), anchors from this category increased in number and strength toward the end of 2008 and in 2009. Some of the strongest anchors in this period were data, software, metaphor, and abstraction. The last two anchors described CC as an analogy for the complexity of the Internet as the version of the article from 10 November, 2009, shows: "The term *cloud* is used as a metaphor for the Internet, based on how the Internet is depicted in computer network diagrams and is an abstraction of the underlying infrastructure it conceals" ("Cloud computing", n.d.).

Contributors also used another generalizing anchor, paradigm shift*,* to explain the underlying transformation in the delivery of computing services that CC implies and linked it to other similar transformations. One can see as much in the version of the article from 12 January, 2010: "cloud computing paradigm shift is similar to the displacement of electricity generators by electricity grids early in the 20th century" ("Cloud computing", n.d.).

The anchoring statistics for the generalizing stage reveal an interesting picture. Anchoring dissimilarity (see Figure 18) hovered around 30 percent—a relatively low figure compared to the previous periods. Similarly, the intensity of the anchoring process decreased, which one can observe in the relatively low numbers of new and removed anchors (see Figure 19) and in the relative stability of anchors during this period (see Figure 20). Thus, it appears that, after the intensive anchoring phase that occurred in response to the novelty of CC, anchoring activity become relatively stable.

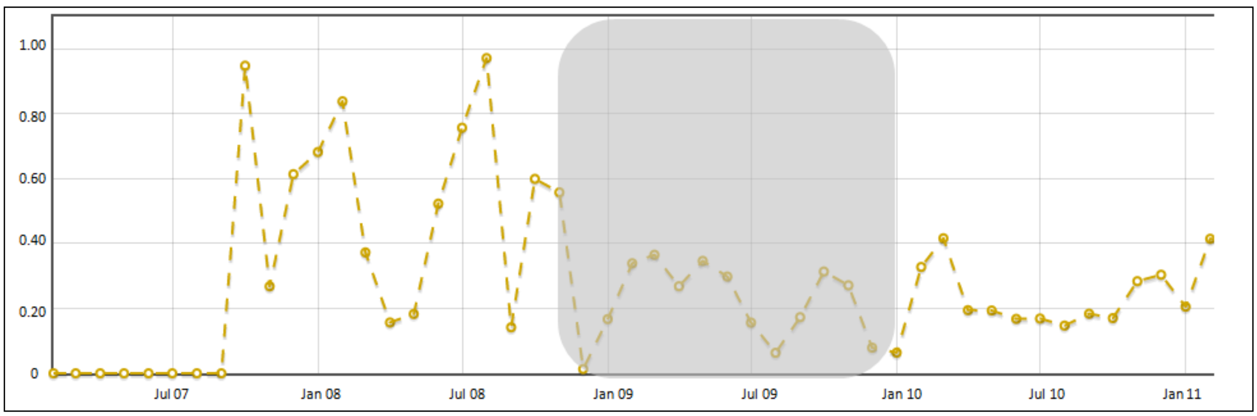


Figure 78. Anchoring Dissimilarity for CC

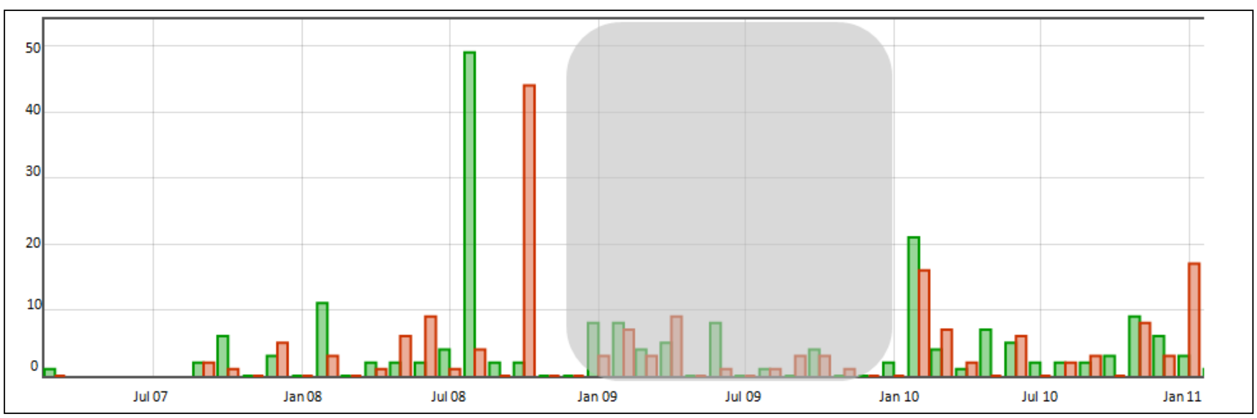


Figure 89. New and Removed Anchors for CC

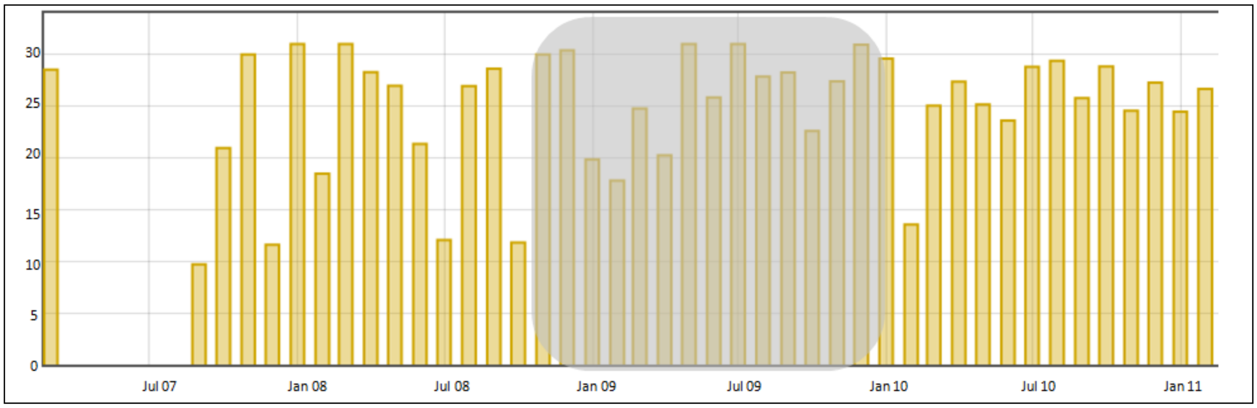


Figure 20. Anchoring Durability for CC

6.3.5 Concretization (February, 2010, to February, 2011)

In general, the anchoring activity in this stage was moderate, which one can observe in the relatively low dissimilarity scores (see Figure 21), low number of new and removed anchors (see Figure 22), and in the consistently high anchoring durability (see Figure 23).

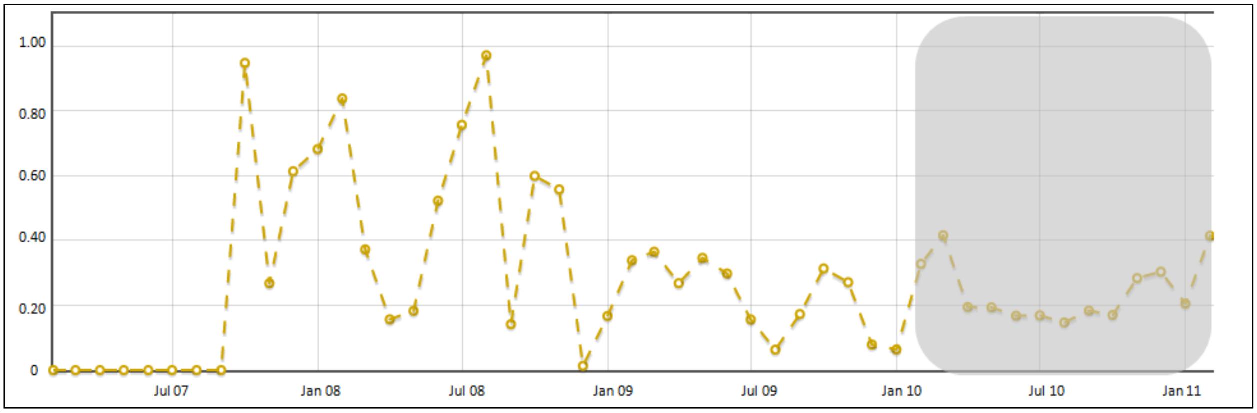


Figure 21. Anchoring Dissimilarity for CC

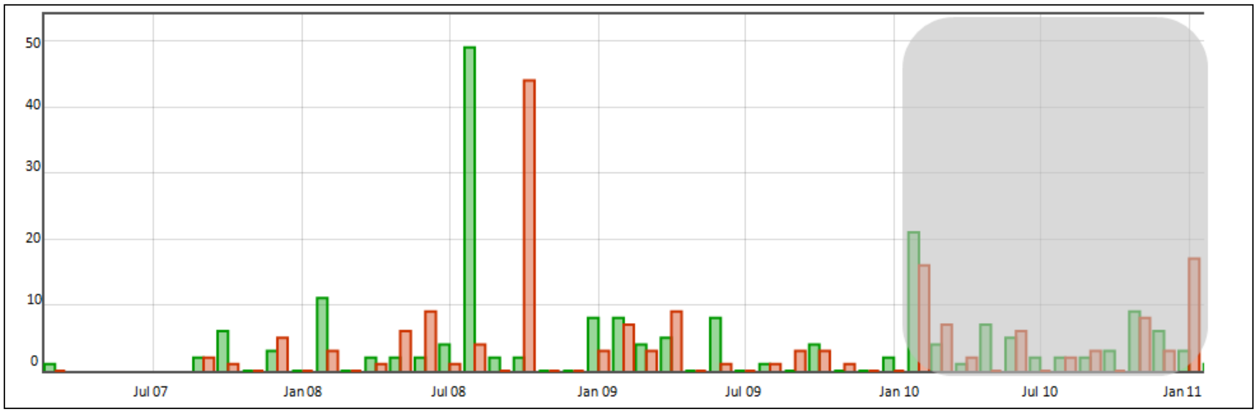


Figure 22. New and Removed Anchors for CC

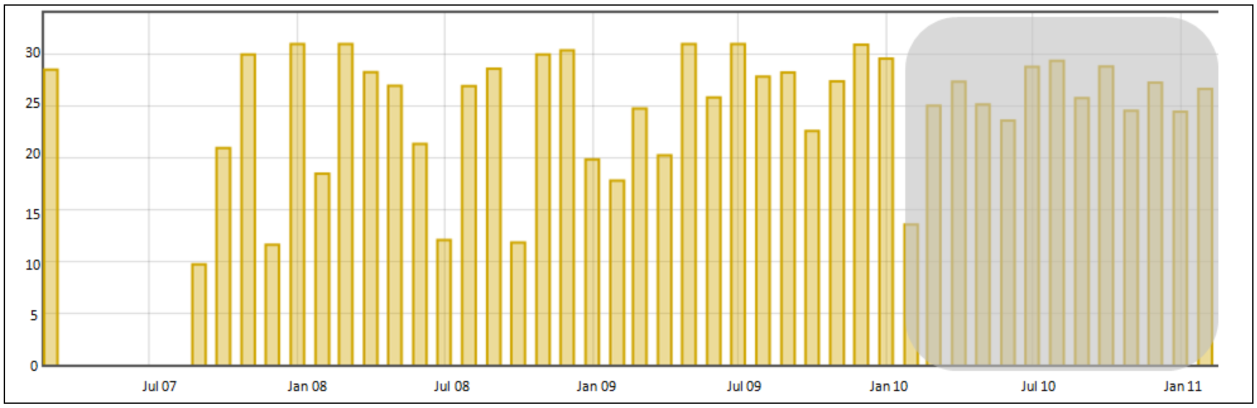


Figure 23. Anchoring Durability for CC

2010

Show 100 entries Search:

Anchor	Days survived	Revisions survived	(Re)Introductions	Anchor strength
software	364.99	983	6	0.99
web browser	364.99	983	6	0.99
internet	362.90	971	6	0.99
data	364.99	983	6	0.99
computer network diagram	364.99	981	7	0.99
scalability	364.87	978	8	0.99
infoworld	364.99	981	7	0.99
business application	364.99	983	6	0.99
metaphor	364.99	981	7	0.99
abstraction	364.99	981	7	0.99
server (computing)	364.99	981	7	0.99
computing	360.38	969	7	0.98
virtualization	346.75	912	8	0.93
quality of service	311.97	825	5	0.84
paradigm shift	310.95	785	15	0.82
google	227.00	568	4	0.60
mainframe computer	238.62	541	11	0.60
service level agreement	208.02	505	5	0.54
salesforce	200.30	497	4	0.52
electricity grid	203.24	485	8	0.52
client-server	203.30	452	9	0.51
pdf	180.76	457	4	0.48
microsoft	180.70	429	6	0.46
ibm	168.39	433	8	0.45
nist	155.08	408	3	0.42
computer	149.63	403	6	0.41
amazon web services	141.84	382	4	0.39

Figure 24. Anchor Resilience Table Showing the Strongest Anchors in 2010[9]

[9] We highlight generalizing anchors in yellow, providers and users anchors in blue, and benefits anchors in red.

However, from qualitatively examining the anchors, we found that a marked shift occurred in the nature of the CC representation during this stage. On the one hand, anchors from the generalizing category continued to be strong with software, Internet, data, metaphor, abstraction, and paradigm shift all present in the top 15 places of the strongest anchor list for most of the concretization stage.

On the other hand, contributors once again introduced anchors from the providers and users category to familiarize CC by using concrete examples of companies that either used or provided CC functionalities (e.g., Google, Salesforce, Microsoft, IBM, and Amazon Web Services) (see Figure 24 above).

In addition to the providers and users category, anchors from the benefits category gained strength during this stage. Some of the strongest anchors in this category were Web browser, business application, quality of service, and service-level agreement. The first anchor described the mechanism that ensures the reliability of the CC services that the second and third anchors described, which resulted in the overall outcome that the last anchor described.

6.3.6 Summary of Anchor Evolution Stages

Figure 25 summarizes the evolution stages in the CC article.

Stage	1. CC as utility computing	2. Technical features of CC	3. Application of CC	4. Generalizing CC	5. Concretization
Timeframe	Mar 2007 – Aug 2007	Sep 2007 – Jul 2008	Aug 2008 – Oct 2008	Nov 2008 – Jan 2010	Feb 2010 – Feb 2011
Description	CC explained in terms of a single anchor from the "*Technical aspects*" category, *Utility computing*.	Introduction of additional "*Technical aspects*" anchors to ground CC in terms of its technical features.	Introduction of multiple anchors from the "*Providers and users*" category to emphasize the applicability of CC.	Removal of most "*Providers and users*" anchors and introduction of multiple anchors from the "*Generalizing*" category; focus on the relationship of CC to other concepts and practices.	Re-introduction of anchors from the "*Providers and users*" category and introduction of anchors from the "*Benefits*" category.

Figure 25. Evolution Stages for the CC Article

As we may expect when people represent new, unfamiliar, and relatively abstract phenomena (as opposed to familiar and unambiguous people, events, or phenomena), the CC article saw multiple representational shifts and transitions in its anchoring activity in the four years after it first appeared. The process moved slowly at first, which the low numbers of edits and editors and the single anchor in the first few months of the article's evolution evidence.

To ground the unfamiliar and intangible phenomenon, contributors anchored CC in existing, recognizable, and tangible technical terms in the second stage (Web browser, Web Application, distributed computing, and grid computing were already familiar concepts in 2007). Focusing on the tangible technical aspects of a new and unfamiliar phenomenon in order to make sense of it makes for a more attainable task than trying to explain its general meaning, significance, or consequences. Such anchoring may further indicate that more technically inclined audiences initially had an interest in and engaged with the topic, whereas concrete business applications emerged only over time.

Accordingly, as contributors refined the technical side of CC, the scope of the representation broadened and the focus of anchoring shifted externally to CC's applicability. One can see as much between August and October, 2008, when contributors introduced multiple anchors that described users and providers of CC applications.

This trend of externalization continued during 2009 when contributors introduced multiple generalizing anchors into the article. These anchors described CC in more general terms than before (e.g., as a metaphor for the complexity of the Internet) and related the paradigmatic shift that it implied to other such shifts in history.

The focus on the generalization of CC lasted until the beginning of 2010 when contributors introduced multiple anchors that explained the concept in terms of its benefits and applications in addition to its general implications. These anchors signified a certain maturity in the concept as its understanding became more specific in terms of various fields of application and their associated benefits.

6.4 Objectification Analysis

From examining the objectification process of CC, we found an increasing trend in the amount of references from other Wikipedia articles (Figure 26). This trend began on 5 September, 2007 (less than six months after the CC Wikipedia article was created) when the article on cloud applications referenced CC.

This trend likely reflects the significant attention given to CC in popular media, which introduced and made it understandable to wide audiences. Many Wikipedia contributors were plausibly exposed to information regarding the CC phenomenon from sources outside of Wikipedia. Therefore, they felt it appropriate to anchor different phenomena (i.e., different articles they contributed to) in terms of CC.

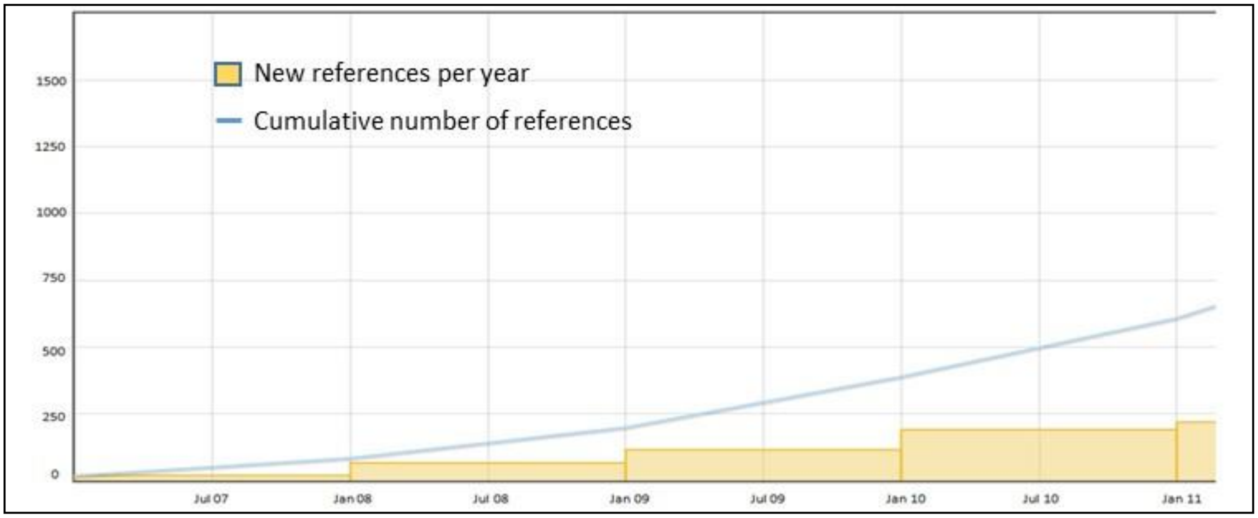


Figure 26. Total Number of New References Per Year to the CC Article (the Blue Line Indicates the Total Cumulative Number of References Over Time)

7 Discussion

In Section 1, we emphasize the importance of adopting a historical perspective in IS research and advocate the use of Wikipedia as an online source of historical data. To enable researchers to engage in inquiry informed by a historical lens, we drew on SRT to build the analytical tool WikiGen and developed a mixed methodology for examining the process of knowledge creation on Wikipedia. In this section, we initially position our methodology as a historical method in relation to the framework that Porra et al. (2014) posit and then discuss its usefulness and implications for IS research.

7.1 Positioning WikiGen Methodology as a Historical Method

Porra et al. (2014) outline the “historical method” as a research method in information systems. To this end, the authors introduce a four-tier framework for characterizing research methods that comprise 1) paradigmatic assumptions, 2) research approach, 3) methods that guide research, and 4) concrete techniques that should be applied with the method. Subsequently, they use the framework to unpack and position the historical method as a legitimate IS research practice. Thus, the framework can demonstrate that our methodology qualifies as a particular kind of historical method. We note that, unlike the examples that Porra et al. provide, our methodology does not focus on reconstructing and explaining past events or concrete organizational change processes but rather on tracing the social construction processes of particular concepts or the collective sensemaking of public events. Nonetheless, our methodology complies with the distinctions that the authors make in the framework:

1) **Paradigmatic assumptions**: Porra et al. (2014) posit that reality, its social institutions, and "historical narratives are socially constructed" (p. 543), which aligns with interpretivist and constructionist understandings of reality and the social science paradigm that Hassan (2018) outlines. Our methodology, through its explicit grounding in SRT, belongs firmly to a social constructionist paradigm in that it assumes concepts by which we understand the world to result from social negotiation and communicative processes.
2) **Research approach**: Porra et al. (2014) stress the pragmatist nature of historical research. Accordingly, the historical method should at once heed 1) the agency of the historian who conducts research from a particular point of view and justifies methodological choices and 2) the openness required to accommodate diverse perspectives in dealing with the subject matter (Porra et al., 2014). Our methodology explicitly focuses more on highlighting the diverse perspectives and viewpoints that go into defining and constructing particular understandings (e.g., through the WikiGen collaboration and anchor analyses). It also stresses the necessity for the historian to make certain choices in applying WikiGen to follow interesting and potentially significant lines of enquiry (e.g., in identifying anchor evolution stages).
3) **Research methods**: Porra et al. (2014) take a broad view of method and highlight the role that particular theories of change play in historians' work. True to their pragmatist stance, the authors propose that historians should adopt a change theory appropriate to the matter under study. Given that we focus on studying how particular IS concepts or phenomena have come to be socially understood over time, we argue that our methodology embodies an interpretivist theory of change because we do not assume that the anchoring process and its emergent narratives are true in any objective sense. Rather, we "view the resulting narratives as socially constructed plausible scenarios of the past events based on the evidence and the change theory" (Porra et al., 2014, p. 549)
4) **Techniques**: Porra et al. (2014) argue that pragmatist historians should focus on employing useful techniques and not on following an "anything goes" attitude since "even the most renowned pragmatists follow guidelines when these work" (p. 552). The authors demonstrate how they themselves used guidelines for explaining past events, such as breaking up events into steps, asking focusing questions, and gathering and critiquing evidence. We outline our methodology as a set of actionable guidelines and techniques for using WikiGen. Yet, while we demonstrate its applicability and usefulness with an example, only the adoption and use of the tool and methodology in future studies will confirm whether or not are guidelines prove useful from a pragmatic point of view.

We conclude that the WikiGen methodology complies with the characteristics of a historical research method as Porra et al. (2014) outline. In Section 7.2, we propose ways in which it might benefit IS researchers.

7.2 Usefulness and Implications of the WikiGen Methodology for IS research

By way of example, we demonstrate the usefulness of WikiGen and our methodology by applying them to the CC Wikipedia article. In doing so, we found that contributors to the article engaged in continuous anchoring as SRT posits. By using the methodology, we identified distinct shifts in the evolution of the CC representation by tracking fluctuations in collaboration, anchoring, and objectification patterns. Further, by conducting qualitative analyses, we could semantically interpret the shifts in the collective meaning of CC. This interpretation showed that the representation of CC gradually became more sophisticated and layered (i.e., it initially couched CC in a single technical anchor before it eventually used multiple anchors from different categories) and externally oriented (i.e., it initially explained CC's internal technical features before it eventually provided examples of its use, relationships to other practices and concepts, and benefits).

While we applied WikiGen and the methodology to a single illustrative article, we believe that one can use them to inform various studies concerning the collective creation of knowledge.

First, work that examines the popularity and use of different phenomena in IS research (e.g., Baskerville & Myers, 2009; Belanger & Carter, 2012; Petter, DeLone, & McLean, 2012; Elbanna & Newman, 2013) can benefit from gaining an insight into how people outside the immediate IS discourse understand such phenomena and the way these understandings originated and changed over time. IS academics do not work in an intellectual vacuum. We engage in an ongoing dialogue with governmental agencies, the business community, and practitioners through conferences, teaching, and research activities. Our work both influences and is influenced by societal agendas, knowledge trends, and trendsetters such as business leaders, consulting firms, and mass media (Baskerville & Myers, 2009). However, the vast majority of studies that have adopted a historical perspective in IS have used work published in academic and practitioner journals as an exclusive data source. For the most part, they have not leveraged data that has appeared in the popular press, industry publications and white papers, and non-refereed online sources such as user-generated content.

Given Wikipedia's size, popularity, and dynamism (i.e., the fact that contributors regularly update content to reflect their shifting opinions and interactions), we maintain that it represents an important part of the trendsetting and consumption processes that researchers have not systematically examined in IS research to date.

Second, one can use WikiGen and the methodology to gauge the relevance of IS research topics. Various authors in the IS literature have long debated the relevance of our research argued that IS research should have relevance to industry and practitioner groups outside our academic community (e.g., Baskerville, 1999; Benbasat & Zmud 1999; Rosemann & Vessey, 2008). Importantly, relevance does not simply concern IS scholars' willingness and ability to engage with research topics that business and practitioner groups have an interest in. It is also a matter of engaging with these topics in a way that external stakeholders find meaningful. Thus, for our research to have relevance, we need to not only research technologies, practices, or concepts that outside groups have an interest in but also approach them in a way that is compatible with (but not necessarily identical to) the way they outside groups approach and understand them.

To enable a meaningful discussion, it is important to ascertain that there is a basic agreement about the nature of the phenomena that are discussed (although not necessarily about their significance, usefulness, or implications) or an acknowledgement of the disagreements. Using WikiGen and the methodology to analyze phenomena of interest can help IS researchers understand how they people outside the IS academic community interpret them and, based on this understanding, attempt to build semantic bridges across disciplinary and professional boundaries to increase the relevance of our research.

Third, one can employ WikiGen and the methodology to examine similarities and differences in the patterns of collaboration, anchoring, and objectification across different types of IS topics, such as:

- Tangible and abstract (e.g., iPad & CC): do audiences interpret IS phenomena that have a concrete and identifiable physical manifestation differently compared to inherently immaterial phenomena whose exact nature one cannot easily and clearly articulate and pin down? Can one observe differences in how fast the representations of these phenomena converge or become objectified, in the degree of their sophistication (i.e., number of anchors and anchor categories), or in the level of disagreement among contributors?
- Successes and failures (e.g., Google Docs & Google Wave): does the collective interpretation of IS phenomena that large audiences have accepted and used exhibit different patterns than that of IS phenomena that users have not taken up, that users have outright abandoned, or that failed to perform as expected? Do we employ distinct social and interpretative mechanisms to make sense of phenomena that are generally considered a success or failure?
- Contemporary and retrospective (e.g., Internet of things & ARPANET): do the representational patterns of phenomena that audiences collectively interpret as they occur or appear in the public consciousness differ to the representational patterns of phenomena whose initial nature and impact one can access only in a mediated fashion and recall on by drawing on personal or social memory? This kind of inquiry can help to shed light on the mechanisms of collective memory and the way that they inform the construction of shared knowledge.
- Different phenomena that are generally considered to occupy the same IS category (e.g., Facebook & Google+, eBay & Amazon, OS X & Windows): how do different social networking applications, online retailers, or operating systems compare in terms of the way that audiences represent them over time? Are there any noticeable difference and similarities between the representations of different companies and technologies that operate or are used in the same industry or serve a similar purpose?

Finally, one can apply WikiGen and the methodology to examine how different online communities outside the academic IS discourse understand prominent IS phenomenon. For example, one can analyze

Wikipedia articles in different languages that correspond to the same phenomenon in order to explore representational differences and similarities between the analyzed communities.

8 Contributions

Our study makes several contributions. First, the tool and methodology that we outline in this paper provide a robust and accessible way for scholars to engage in historically informed IS research. Despite the potential benefits of such research, researchers have not widely acknowledged or adopted it in the IS discipline. We believe that our work can help researchers to conduct studies a historical lens and substantially develop this stream of literature.

Second, our study makes a methodological contribution to the IS literature. By formalizing an ordered and theoretically informed process around a Web-based and freely available tool, we provide a substantive methodology that researchers can employ to examine how people collectively shape knowledge in one of the biggest online environments on the Internet. Furthermore, applying the methodology can help researchers reveal how collective understandings of different phenomena emerge and change over time, which allows them to access the origins of collectively accepted knowledge and the content and structure of the processes through which this knowledge evolves.

Third, our study contributes to the literature on social representations. Research using SRT has had a long tradition in the social sciences, particularly in the sociology and social psychology disciplines. Researchers have applied the theory to study myriad phenomena ranging from public acceptance of psychoanalysis (Moscovici, 1961) and biotechnology (Bauer & Gaskell, 2002) to the collective understanding of superbugs (Washer & Joffe, 2006), mental illnesses (Morant, 2006), risk (Barnett & Breakwell, 2003), climate change (Höijer, 2010), aggression (Campbell, Muncer, Guy, & Banim, 1998), and others. In the IS literature, only a handful of researchers have used SRT to study IT implementation (Gal & Berente, 2008), burnout in the IT profession (Pawlowski, Kaganer, & Cater, 2007), change in work practices (Vaast & Walsham, 2005), and misalignment of knowledge management systems with organizational strategy (Dulipovici & Robey, 2013).

To the best of our knowledge, researchers have never applied SRT to empirically study how people collectively interpret phenomena on online collaborative platforms in general or on Wikipedia in particular. We believe that, due to Wikipedia's transparency, applying SRT to it can illuminate in detail the way collective knowledge processes unfold by examining how social representations emerge and change over time. Furthermore, by applying WikiGen to large amounts of data across different Wikipedia articles, one can identify and compare patterns in the evolution and change of different representations in a previously impossible way.

Finally, out study makes a contribution to practice by providing companies with a mechanism to track the way that Wikipedia contributors perceive them and their products and services. In recent years, commercial companies have increasingly begun to systematically collect and analyze online data to inform business and marketing strategies and product development and to improve customer relationships and retention rates. Drawing on various forms of data analytics, they have leveraged different techniques to analyze data from blogs, social networking websites, and online communities (e.g., sentiment analysis and trend analysis). WikiGen can serve as an additional tool for companies to gain a nuanced understanding of how public perceptions of them are formed and change over time.

9 Limitations and Future Research

Our study has several limitations. First, in structuring WikiGen and the methodology to examine how social representations emerge and change (i.e., how Wikipedia contributors interpret phenomena online), we acutely recognized that the language with which we describe this process (namely, SRT) is in itself a neither neutral nor objective representational layer. Indeed, we could have described the collective creation of online knowledge using different constructs and conceptual categories in a way that would have produced different explanations of the process. Consequently, we do not claim to describe *the* reality of collective knowledge creation but rather *one* reality, or interpretation, of it—a representation of a representation. One should understand the results that come from using the tool and methodology accordingly.

Second, when developing WikiGen, we focused on data contained in Wikipedia's "article pages" and did not incorporate data from its "talk pages". In the talk pages, contributors discuss and negotiate how to develop and change the corresponding Wikipedia article as it appears in the article pages. Therefore, the talk pages may contain an additional layer of the representational activity that WikiGen currently does not tap into. Future research could extend WikiGen to identify anchors in Wikipedia’s talk pages to more comprehensively analyze how social representations form and evolve.

Third, while Wikipedia provides a rich and dynamic platform for historical analyses, future research could extend the scope of such analyses to include other forums where representational activity likely occurs. Some examples include social media networks (e.g., Reddit, Twitter, and Instagram), blogs, social forums, professional publications, popular press, and ads. Indeed, future research could conduct an SRT-informed analysis of representational processes on Wikipedia in conjunction with a similar analysis of representational processes on any of these forums.

Finally, future research could further develop WikiGen to incorporate high-level analyses of social representations that go beyond those that the tool currently incorporates. Some examples include relational analysis (i.e., a systematic examination of the changing relationships between different categories of anchors), advanced objectification analysis (i.e., an ongoing examination of the nature of other Wikipedia articles that refer to the article under study to better understand how audiences understand and use it), and social dynamics analysis (i.e., an examination of the relationship between patterns of social interactions among contributors—based on collaboration statistics and edit-war levels—and the characteristics of the emergent representation—based on the size of the article and the number of anchor and anchor categories).

References

Alexa. (2018). *Wikipedia.org traffic statistics*. Retrieved from http://www.alexa.com/siteinfo/wikipedia.org

Barnett, J., & Breakwell, G. (2003). The social amplification of risk and the hazard sequence: The October 1995 OC pill scare. *Health, Risk and Society, 5*(3), 301-13

Baskerville, R, L. (1999). Investigating information systems with action research *Communications of the AIS, 2,* 1-32.

Baskerville, R, L., & Myers M, D. (2009). Fashion waves in information systems research and practice. *MIS Quarterly, 33*(4), 647-662.

Bauer, M. W., & Gaskell, G. (2002). *Biotechnology: The making of a global controversy*. Cambridbe, UK: Cambridge University Press.

Belanger, F., & Carter, L. (2012). Digitizing government interactions with constituents: An historical review of e-government research in information systems. *Journal of the Association of Information Systems, 13*, 363-394.

Benbasat, I., & Zmud R, W. (1999). Empirical research in information systems: The practice of relevance. *MIS Quarterly, 23*(1), 3-16.

Berger, P. L., & Luckmann, T. (1966). *The social construction of reality: A treatise in the sociology of knowledge*. New York, NY: Doubleday.

Bernroider, E. W. N., & Pilkington, A., & Cordoba, J. (2013). Research in information systems: A study of diversity and inter-disciplinary discourse in the AIS basket journals between 1995 and 2011. *Journal of Information Technology, 28*, 74-89.

Borra, E. & Weltevrede, E., & Ciuccarelli, P., & Kaltenbrunner, A., & Laniado, D., & Magni, G., & Mauri, M., & Rogers, R., & Venturini, T. (2015). Societal controversies in Wikipedia articles. In *Proceedings of the Conference on Human Factors in Computing Systems.*

Bryant, A., & Black, A., & Land, F., & Porra, J. (2013). Information systems history: What is history? What is IS history? What IS history? …and why even bother with history? *Journal of Information Technology, 28*, 1-17.

Campbell, A., & Muncer, S., &, Guy, A., & Banim, M. (1998). Social representations of aggression: Crossing the sex barrier. *European Journal of Social Psychology*, *26*(1), pp. 135–147.

Cloud computing. (n.d.). In *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Cloud_computing

Dulipovici, A., & Robey, D. (2013). Strategic alignment and misalignment of knowledge management systems: A social representation perspective. *Journal of Management Information Systems, 29*(4), 103-126.

Elbanna, A., & Newman, M. (2013). The rise and decline of the ETHICS methodology of systems implementation: Lessons for IS research, *Journal of Information Technology*, *28*, 124-136.

English Wikipedia. (n.d.). In *Wikimedia*. Retrieved from http://stats.wikimedia.org/EN/SummaryEN.htm

Farr, R. M. (1993). Common sense, science and social representations. *Public Understanding of Science,* 2(3), 189-204.

Flöck, F., & Acosta. M. (2015). whoVIS: Visualizing editor interactions and dynamics in collaborative writing over time. In *Proceedings of the 24th International Conference on World Wide Web.*

Gal, U., & Berente, N. (2008). A social representations perspective on information systems implementation: Rethinking the concept of "frames". *Information Technology & People*, *21*(2), 133-154.

Glaser, B. G. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory*. MilL Valley, CA: Sociology Press.

Greenhalgh, T., Robert, R., MacFarlane, F., Bate, P., Kyriakidou, O., & Peacock, R. (2005). Storylines of research in diffusion of innovation: A meta-narrative approach to systematic review. *Social Science and Medicine*, *61*(2), 417-430.

Grover, V., London, J., & Craig, K. (2016). A historical observation of the intellectual and institutional structures of the field. *Communications of the Association for Information Systems,* 38, 432-476.

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

Hansen, S., Berente, N., & Lyytinen, K. (2009). Wikipedia, critical social theory, and the possibility of rational discourse*. The Information Society, 25*(1), 38-59.

Hassan, N. R. (2017). The history and philosophy department. *Communications of the Association for Information Systems*, *41*(15), 319-333.

Hassan, N. R. (2018). Taking IS history seriously. In R.D. Galliers & M.-K. Stein (Eds.), *The Routledge companion to management information systems* (pp. 5-29). London, UK: Routledge.

Hepp, M., Siorpaes, K., & Bachlechner, D. (2007). Harvesting wiki consensus: Using Wikipedia entries as vocabulary for knowledge management*. IEEE Internet Computing*, *11*(5), 54-65.

Hirschheim, R., & Klein, H., K. (2012). A glorious and not-so-short history of the information systems field. *Journal of the Association of Information Systems,* 13(4), 188-235.

Hirschheim, R., & Saunders, C., & Straub, D. (2012). Historical interpretations of the IS discipline: An introduction to the special issue*. Journal of the Association of Information Systems, 13*(4), i-viii*.*

Höijer, B. (2010). Emotional anchoring and objectification in the media reporting on climate change. *Public Understanding of Science, 19*(6), 717-731.

Iggers, G. G. (2005). *Historiography in the twentieth century*. Middletown, CT: Wesleyan University Press.

Iivari, J. (2016). Making sense of the history of information systems research 1975-1999: A view of highly cited papers. *Communications of the Association for Information Systems, 35*, 515-561.

Keegan, B. C., & Lev, S., & Arazy, O. (2016). Analyzing organizational routines in online knowledge collaborations: A case for sequence analysis in CSCW. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 257-266).

List of Wikipedias. (n.d.). In *Wikipedia*. Retrieved from http://meta.wikimedia.org/wiki/List_of_Wikipedias#Grand_Total

Mason, R., O., McKenney, J., L., & Copeland, D., G. (1997). Developing an historical tradition in MIS research. *MIQ Quarterly, 21*(3), 257-278.

Mitev, N., & De Vaujany, F. (2012). Seizing the opportunity: Towards a historiography of information systems*. Journal of the Association of Information Systems, 27*, 110-124.

Morant, N. (2006). Social representations and professional knowledge: The representation of mental illness among mental health practitioners. *British Journal of Social Psychology*, *45*(4), 817-838.

Moscovici, S. (1961). *La psychanalyse, son image et son public*. Paris, France: PUF.

Moscovici, S. (1984). The phenomenon of social representations. In R. M. Farr & S. Moscovici (Eds.), Social representations (pp. 3-69). Cambridge, UK: Cambridge University Press.

Moscovici, S. (2000). The phenomenon of social representations*.* In S. Moscovici & G. Duveen (Eds.), *Social representations: Explorations in social psychology.* Cambridge, UK: Polity Press.

Pawlowski, S. D., & Kaganer, E. A., & 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), 612-627.

Pawlowski, S. D., Kaganer, E. A., & 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*, 612-627*.*

Petter, S., DeLone, W., & McLean, E, R. (2012). The past, present, and future of "IS success". *Journal of the Association for Information Systems, 13*, 341-362.

Phillips, M. G. (2016). Wikipedia and history: A worthwhile partnership in the digital era? *The Journal of Theory and Practice*, *20*(4), 523-543.

Porra, J., & Hirscheim, R., & Parks, M. S. (2005). The history of Texaco's corporate information technology function: A general systems theoretical interpretation. *MIS Quarterly, 29*(4), 721-746*.*

Porra, J., & Hirscheim, R., & Parks, M. S. (2006). Forty years of the corporate information technology function at Texaco Inc.—a history. *Information and Organization, 16*(1), 82-107.

Porra, J., Hirschheim, R., & Parks, M. (2014). The historical research method and information systems research. *Journal of Association of Information Systems, 15*(9), 536-576.

Rosemann, M., & Vessey, I. (2008). Toward improving the relevance of information systems research to practice: the role of applicability checks. *MIS Quarterly, 32*(1), 1-22.

Rosenzweig, R. (2006). Can history be open source? Wikipedia and the future of the past. *Journal of American History*, *93*, 117-146.

Sylvester, A., Tate, M., & Johnstone, D. (2011). Beyond synthesis: Representing heterogeneous research literature. *Behaviour & Information Technology*, *32*(12), 1199-1215.

Twyman M., Keegan B, C., & Shaw A. (2017). Black lives matter in Wikipedia: Collaboration and collective memory around online social movements. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 1400-1412).

Vaast, E., & Walsham, G. (2005). Representations and actions: The transformation of work practices with IT use. *Information and Organization, 15*(1), 65-89*.*

Vaezi, R., Mills, A., Chin, W., & Zafar, H. (2016). User satisfaction research in information systems: Historical roots and approaches, *Communications of the Association for Information Systems, 38*, 501-532.

Viégas, F., Watternberg, M., & Dave, K. (2004). Studying cooperation and conflict between authors with history flow visualizations. In *Proceedings of the Conference on Human Factors in Computing Systems* (pp. 575-582).

Wagner, W., Duveen, G., Farr, R., Jovchelovitch, S., Cioldi, F. L., & Marková, I. (1999). Theory and method of social representations. *Asian Journal of Social Psychology,* 2(1), 95-125.

Washer, P., & Joffe, H. (2006). The hospital "superbug": Social representations of MRSA. *Social Science and Medicine*, *63*(8), 2141-2152*.*

Wolff, R. S. (2013). The historian's craft, popular memory, and Wikipedia. In K. Nawrotzki & J. Dougherty (Eds.), *Writing history in the digital age* (pp. 64-74). Ann Arbor, MI: University of Michigan Press.

Zhang, P. (2015). The IS history initiative: Looking forward by looking back. *Communications of the Association for Information Systems, 36*, 477-514.

Zhang, P. (2016). The IS history initiative: Continued efforts and results. *Communications of the Association for Information Systems, 38*, 420-431.

Appendix A: WikiGen Manual

In Section 3, we briefly describe some of WikiGen's main statistics. Here, we provide a step-by-step guide for generating these statistics and describe their underlying formulas. For illustration purposes, we use the English Wikipedia article on the Amazon Kindle.

Selecting a Wikipedia Article for Analysis

Users can select an article to analyze by navigating to http://wikigen.org and entering its exact name in the search field (see Figure A1).

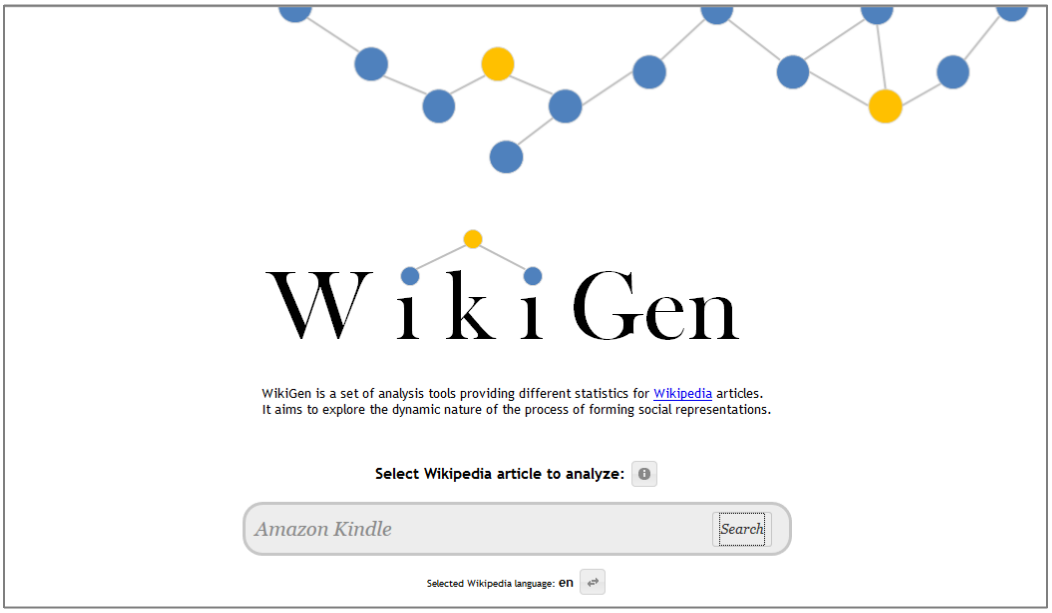


Figure A9. WikiGen Landing Page

When clicking the search button, users can see the current Wikipedia version of the article and use the various WikiGen statistics that appear in tabs on the top of the page (see Figure A2).

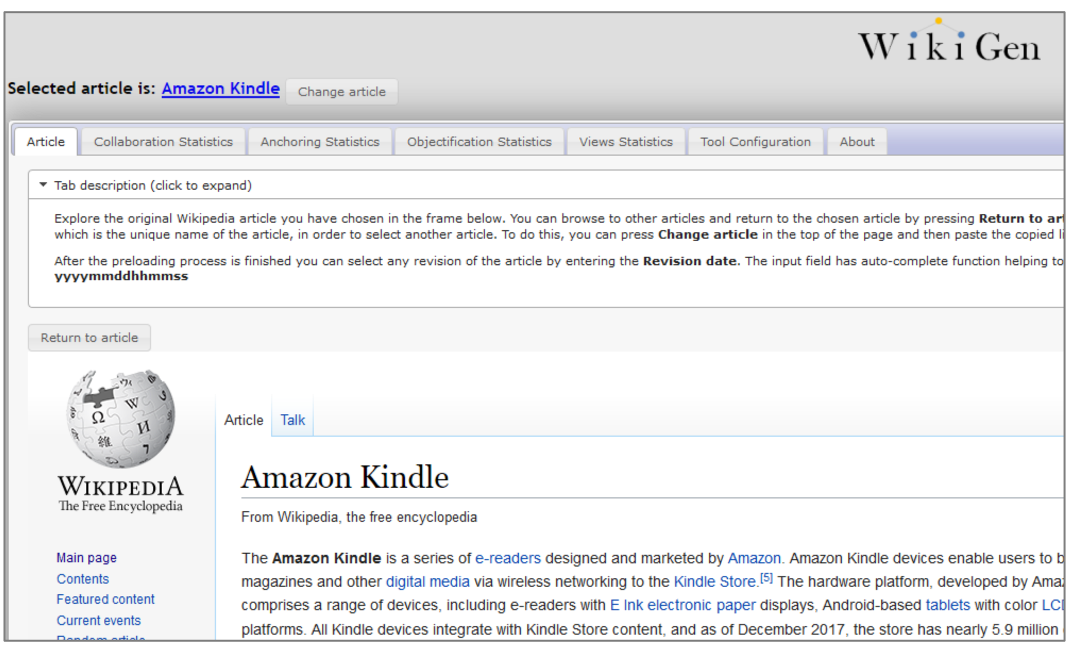


Figure A2. Article View

Using Collaboration Statistics

Users can access "collaboration statistics" through the second tab from the left. This tab provides information about the number of edits and editors over time and the number of edits per editor over time. Users can apply these analyses to all or part of an article's lifespan. Users can define the period in the "date from" and "date until" fields (see Figure A3).

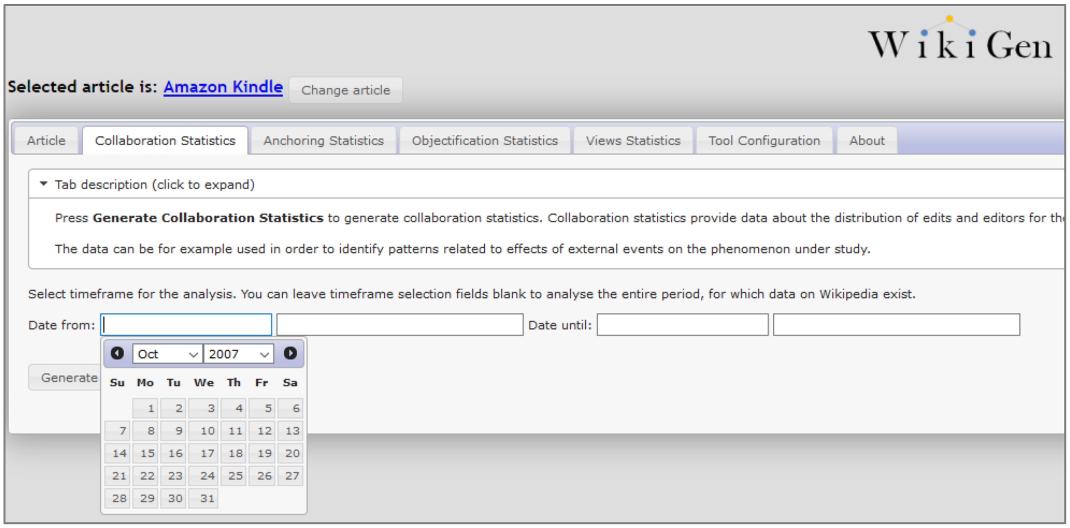


Figure A3. Timeframe Analysis

For example, by clicking on the "edit statistics" tab (see Figure A4), users can choose to view "detailed revision statistics" (see Figure A5), and, by clicking on the "editors statistics" tab (Figure A6), users can view "detailed editors statistics" (see Figure A7). Moreover, by clicking on "edits statistics: measures" in the "Edits statistics" tab, users can see the date of the first revision (i.e., when the article was created) and its total number of edits (see Figure A8).

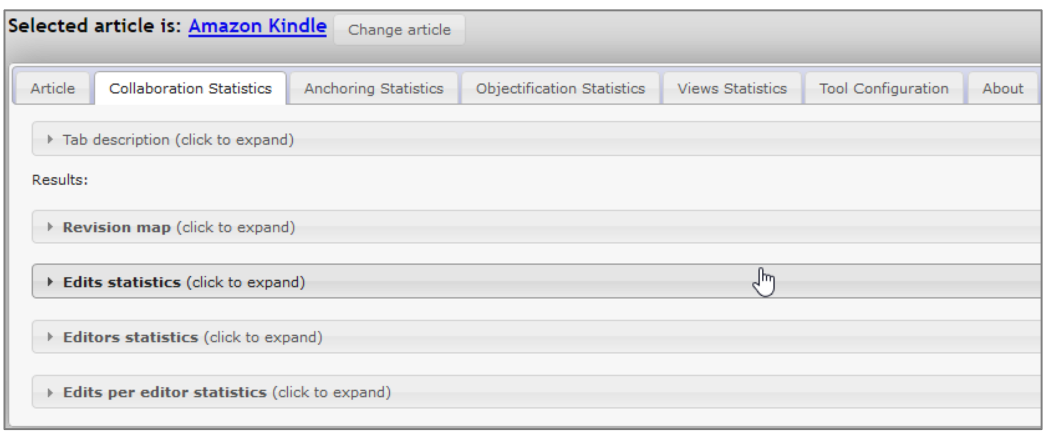


Figure A4. Collaboration Statistics View

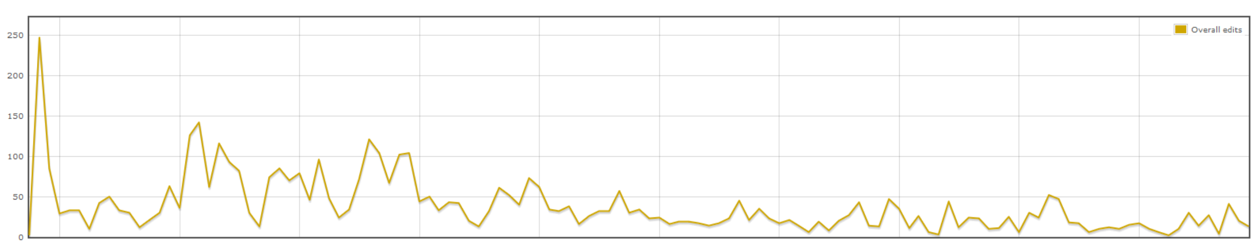


Figure A5. Detailed Revisions Statistics for the Article on the Amazon Kindle

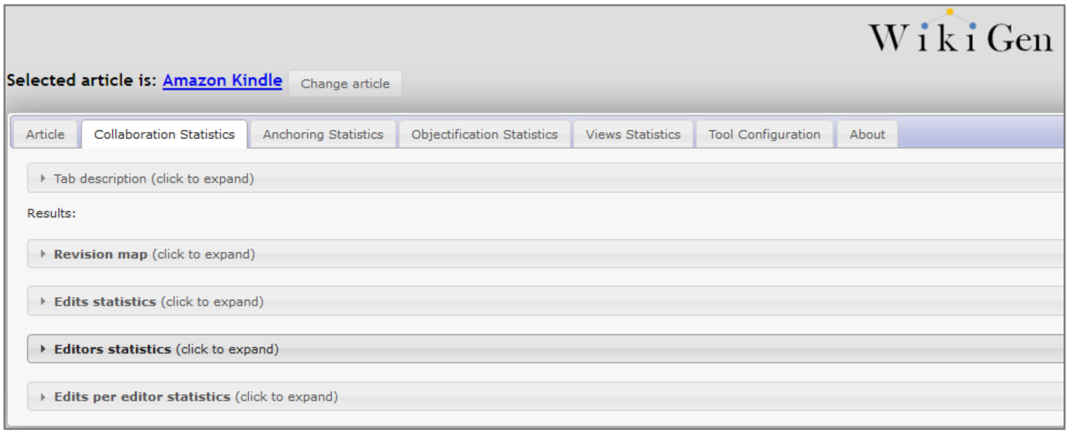


Figure A6. Collaboration Statistics View, Editors statistics tab open

Figure A7. Detailed Editors Statistics for the Article on the Amazon Kindle

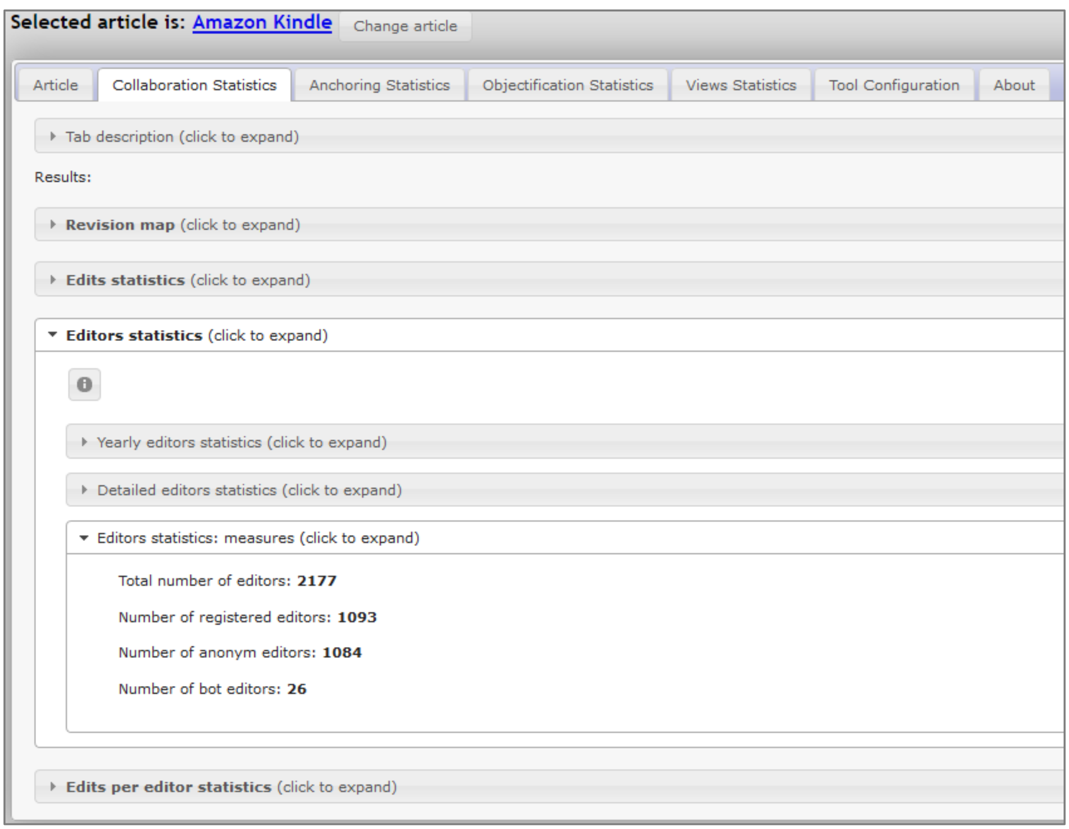


Figure A8. Edit Statistics Measures

Using Anchoring Statistics

Users can access anchoring statistics through the third tab from the left. Clicking the "generate links statistics" button (see Figure A9) produces a range of metrics, which users can access through the relevant tabs in the middle of the page (see Figure A10).

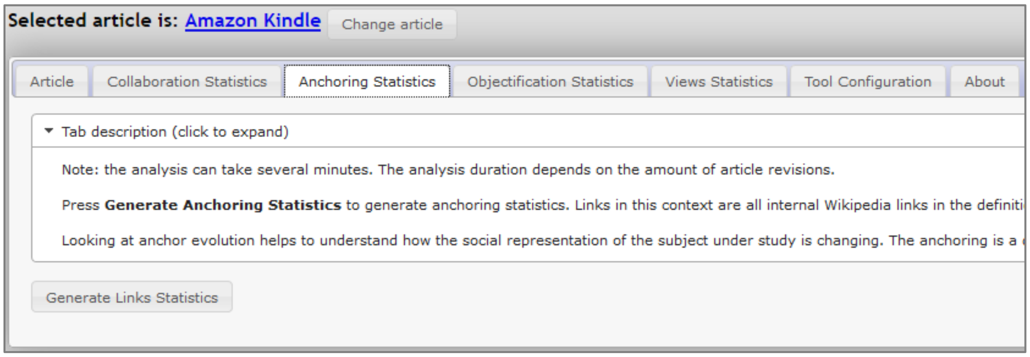


Figure A9. Anchoring Statistics View

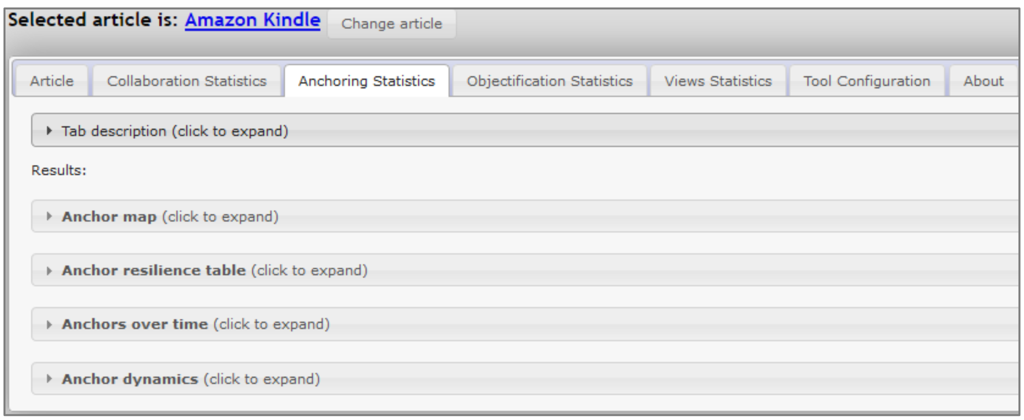
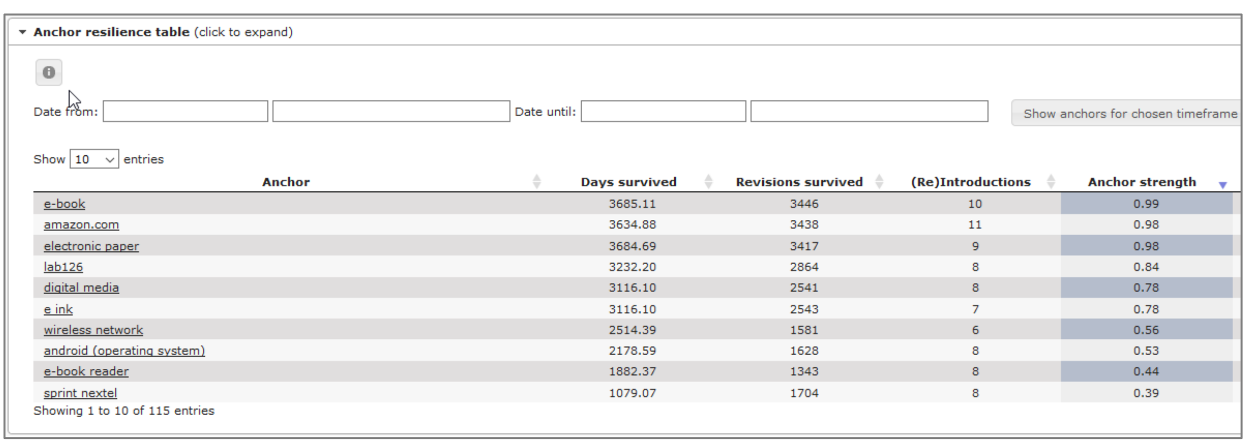


Figure A100. Anchoring Statistics after clicking the Generate Links Statistics button

For example, clicking on the "anchor resilience table" allows users to view, sort, and search the strongest anchors in different periods (see Figure A11). "Anchor strength" is a linear combination of "days survived" and "revisions survived" (as we explain in Section 3) that produces a measure between 0 and 1; anchors that survive for long periods or that are quickly reintroduced after being removed will have a high score. A score of 1 indicates that an anchor has survived all revisions and stayed in the article for the entire period of time.


Anchor resilience table (click to expand)

Date from: Date until: Show anchors for chosen timeframe

Show 10 entries

Anchor	Days survived	Revisions survived	(Re)Introductions	Anchor strength
e-book	3685.11	3446	10	0.99
amazon.com	3634.88	3438	11	0.98
electronic paper	3684.69	3417	9	0.98
lab126	3232.20	2864	8	0.84
digital media	3116.10	2541	8	0.78
e ink	3116.10	2543	7	0.78
wireless network	2514.39	1581	6	0.56
android (operating system)	2178.59	1628	8	0.53
e-book reader	1882.37	1343	8	0.44
sprint nextel	1079.07	1704	8	0.39

Showing 1 to 10 of 115 entries

Figure A111. Anchor Resilience Table for the Article on the Amazon Kindle

In addition to tracing individual anchors, WikiGen can capture changes in the dynamics of multiple anchors across selected periods. It does so through four statistics accessible through the “anchor dynamics” tab (see Figure A12).

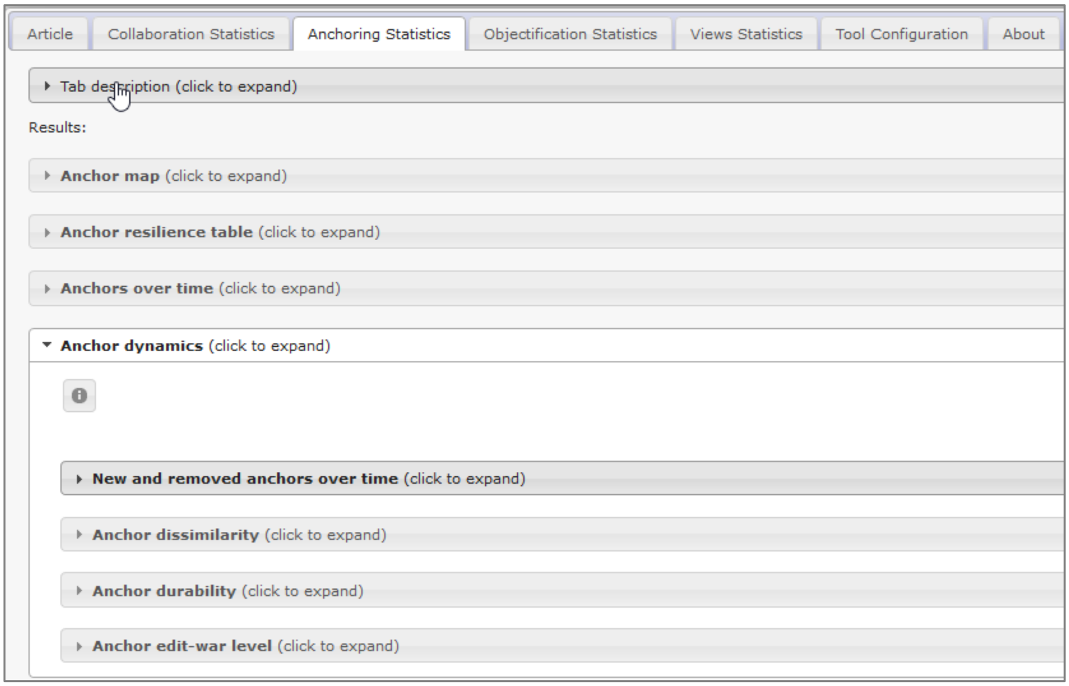


Figure A122. Anchor dynamics

1) **New and removed anchors**: WikiGen displays in a bar chart the number of introduced and removed anchors in a chosen period in a bar chart. The chart helps users to identify periods with extensive anchoring activity (i.e., periods in which many anchors were either removed or introduced (or both), which indicate potentially major shifts in how a topic is understood). Figure A13 below shows new and removed anchors for the article on the Amazon Kindle.

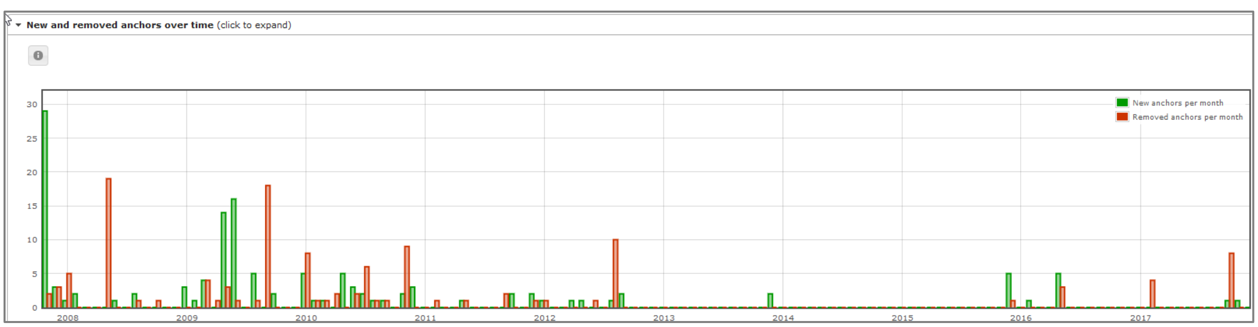


Figure A133. New and Removed Anchors for the Article on the Amazon Kindle

2) **Anchoring dissimilarity**: this measure captures the extent to which anchors in period t are dissimilar to anchors in the previous period t-1. It allows users to compare periods over time to identify significant changes in anchoring. The measure is based on the anchor attribute “days survived” and only considers anchors that existed in the chosen period for at least one day to eliminate the temporary influence of anchors that entered the article due to vandalism. The dissimilarity score ranges from 1 (anchors are completely dissimilar) to 0 (anchors are completely similar) across periods. The calculation takes the sum of least common days survived for all anchors present in periods t and t-1 and sets it in relation to the total sum of maximum days anchors survived for every anchor in periods t and t-1. Figures A14 and A15 illustrate the corresponding formula and an example output, respectively.

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

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

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

Figure A144. Anchor Dissimilarity Formula

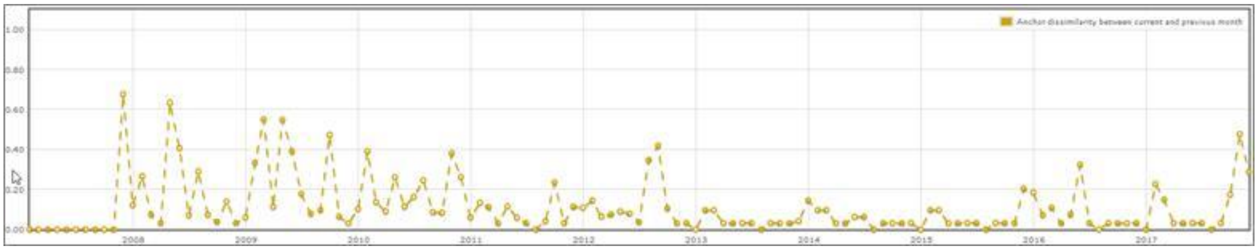


Figure A155. Anchor Dissimilarity for the Article on the Amazon Kindle

3) **Anchoring durability**: measures the average time (e.g., number of days) anchors have existed in the definition part of an article in a chosen period (again, the calculation considers only those anchors that existed in the article for at least one day). The measure ranges between 0 and the maximum time for that period: low scores indicate anchoring instability, whereas high scores indicate anchoring stability. Figures A16 and A17 illustrate the formula and an example output, respectively.

$$dur(t) = \frac{\sum_{i=1}^{n} f(a_{t_i},t)}{n}, \quad \text{where}$$

$\boldsymbol{a_t} = \begin{pmatrix} a_{t_1} \\ a_{t_2} \\ \cdots \\ a_{t_i} \end{pmatrix}$: vector of anchor names a_{t_i} in period t (survived more than 1 day),

$\boldsymbol{n}$: number of anchors in period t (survived more than 1 day),

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

Figure A166. Anchor Durability Formula

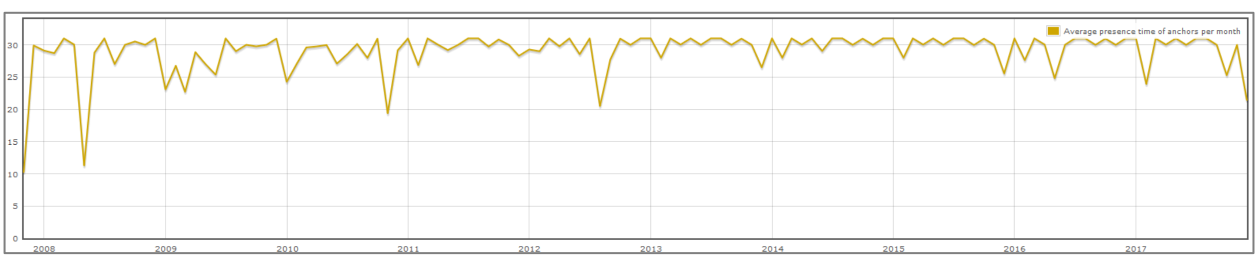


Figure A17. Anchor Durability for the Article on the Amazon Kindle

4) **Edit-war level**: measures the intensity of disagreement among users by relating the number of anchor introductions and disappearances of anchors to the total number of unique anchors in a period. Edit-war ranges from 0 (no introductions or disappearances of anchors), which indicates no disagreement among contributors, to infinity (x anchors introduced and removed y times where y -> inf), which indicates intense disagreement among contributors. Figure A18 displays the formula for edit-war level and Figure A19 displays the edit-war levels for the article on the Amazon Kindle.

$$ewl(t) = \frac{int(t)+dis(t)}{a(t)}, \quad \text{where}$$

$\boldsymbol{int(t)}$: Number of anchor introductions in period t (cf. anchor snapshots)

$\boldsymbol{dis(t)}$: Number of anchor disappearances in period t (cf. anchor snapshots)

$\boldsymbol{a(t)}$: Number of anchors in period t

Figure A178. Edit-war Level Formula

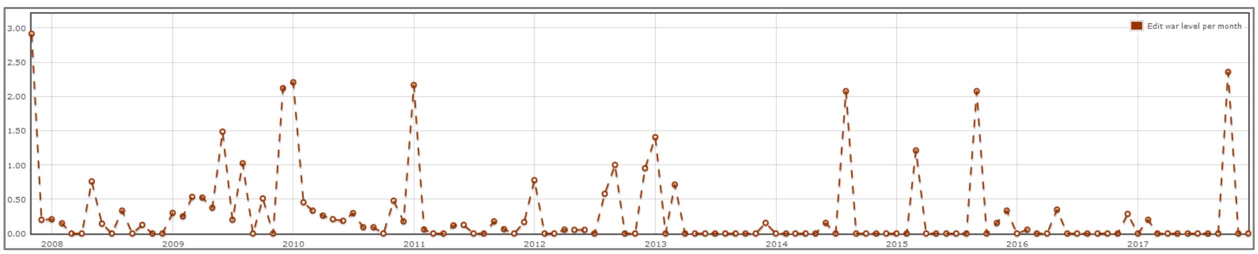


Figure A189. Edit-war Levels for the Article on the Amazon Kindle

Using Objectification Statistics

Users can access objectification statistics through the fourth tab from the left (see Figure A20). Clicking the "Generate Objectification Statistics" button produces a reference graph (see Figure A21) that shows the number of links that lead back to the article per year and the cumulative number of back links.

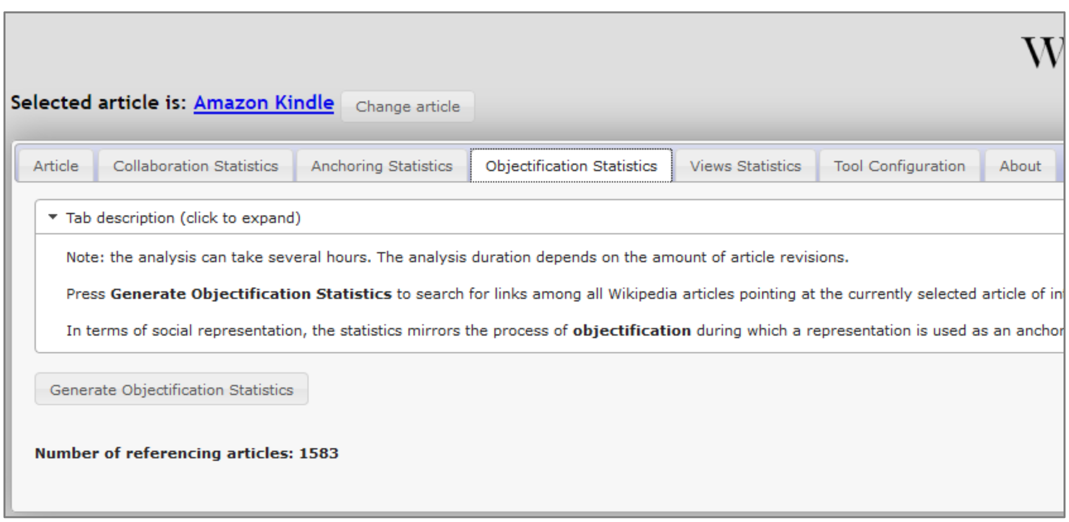


Figure A20. Objectification Statistics view

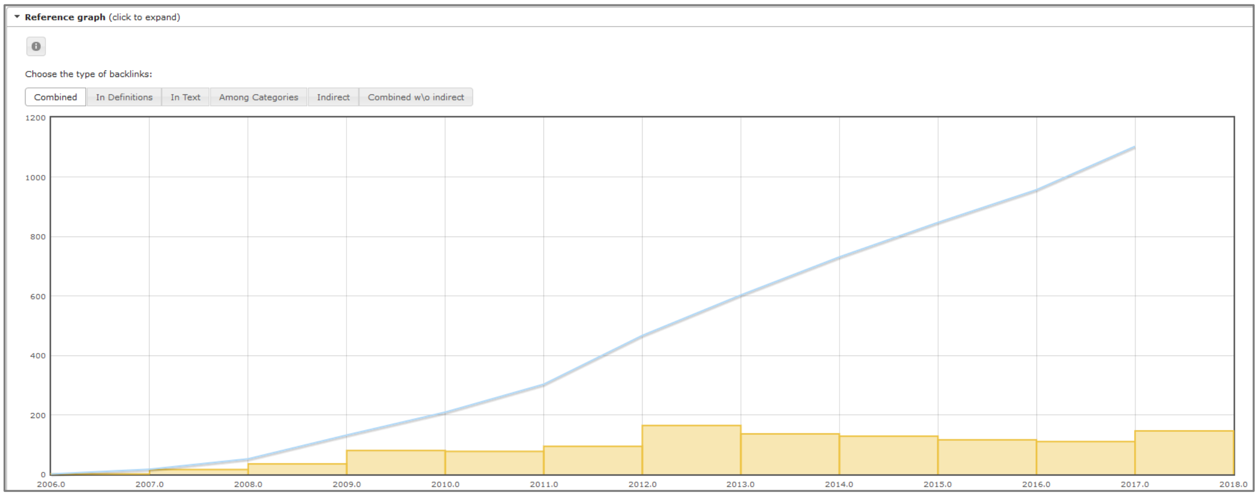


Figure A219. Objectification Reference Graph for the Article on the Amazon Kindle

About the Authors

Uri Gal is an Associate Professor in Business Information Systems at the University of Sydney Business School, where he has been since 2010. He received his PhD in Information Systems from Case Western Reserve University in the USA. His expertise is in IT-enabled change processes and his research focuses on the effect of new digital technologies on people and organisations. He is particularly interested in the use and impact of Workforce Analytics on managerial decision-making, employees, and work practices. His work has been published in some of the leading Information Systems journals and conferences.

Kai Riemer is a Professor of Information Technology and Organisation in the Discipline of Business Information Systems. He joined the University of Sydney in August 2009 from Münster University in Germany, where he held a position as Assistant Professor. He received his Diploma in Information Systems (Dipl. Wirt-Inform.) and his PhD in Management from Münster University, Germany, where he also finished his Postdoctoral thesis (Habilitation). His research interest is in Disruptive Technologies, Enterprise Social Media, Virtual Work, Collaborative Technologies and the Philosophy of Technology. In one current research stream he focuses on the introduction and management of Enterprise Social Networking into organisations and the elicitation of associated benefits. The research has resulted in the development of a number of management frameworks that are widely used among practitioners in this field.

Friedrich Chasin is a postdoctoral researcher at the University of Muenster and a private lecturer at the South Westphalia University and the Osnabrueck University of Applied Sciences. His main research interests include digital innovation and in particular green information systems, and IT-enabled sharing. For the past two years, he has been studying the phenomenon of peer-to-peer sharing by examining corresponding businesses in Brazil, Germany and South Korea. His special focus is on the design-oriented research where he has been developing peer-to-peer sharing services for the electric vehicle domain since 2013.