Association for Information Systems

AIS Electronic Library (AISeL)

Establishing Information Quality Guidelines in Social Information Systems: Comparison and Discussion of Two Approaches

Roman Tilly
University of Cologne, tilly@wim.uni-koeln.de

Robert Wintermeyer
University of Cologne, rwinterm@smail.uni-koeln.de

Oliver Posegga
University of Bamberg, oliver.posegga@uni-bamberg.de

Establishing Information Quality Guidelines in Social Information Systems: Comparison and Discussion of Two Approaches

Completed Research Paper

Roman Tilly
University of Cologne
Albertus-Magnus-Platz
D-50923 Cologne
tilly@wim.uni-koeln.de

Robert Wintermeyer
University of Cologne
Albertus-Magnus-Platz
D-50923 Cologne
rwinterm@smail.uni-koeln.de

Oliver Posegga
University of Bamberg
An der Weberei 5
D-96052 Bamberg
oliver.posegga@uni-bamberg.de

Abstract

Social Information Systems (SocIS) enable many people to interact digitally and collaboratively create and share digital content. Nevertheless, the large and heterogenous SocIS communities make it challenging to ensure information quality (IQ) because members' interpretation and evaluation of content might be very different. As a remedy, many platforms explicitly state normative IQ guidelines. Guidelines can be developed either by the community members themselves or by the platform provider (and imposed on the community). It is unclear, however, which of these two approaches members agree with more strongly and which produces the more satisficatory IQ guidelines. Through an empirical survey study covering 15 different SocIS platforms, we find that members do agree more and are more satisfied when guidelines have been developed by the community. These findings are important for platform providers to improve IQ and retain members, and also inform research on IQ in SocIS.

Keywords: Social Information Systems; Social Media; Information Quality; Guidelines

Introduction

Social Information Systems (SocIS) enable many people to interact digitally and collaboratively create and share digital content (see Kaplan and Haenlein, 2010; Schlagwein, Schoder and Fischbach, 2011; Schoder, Gloor and Metaxas, 2013). Over the last two decades, SocIS have been extremely successful in terms of member engagement in various applications contexts, such as social media, collaborative content projects, citizen science, review platforms, and many more (Statista, 2018). SocIS have the potential to facilitate digital social interaction among millions of people and leverage contributions from numerous voluntary members for shared projects, such as curating open encyclopedias (most famously, Wikipedia, but also various other platforms built upon the common MediaWiki artifact) or building digital maps (e.g., OpenStreetMaps). However, large and heterogenous SocIS communities make it challenging to ensure quality of data and information. This is the case, for example, in citizen science projects that count animal species in the wild (Lukyanenko, Parsons and Wiersma, 2014), but also in everyday digital interactions,

such as when members comment on, describe, or label content on social media platforms such as YouTube or LastFM (Figueiredo et al., 2013).

While the term "quality of data and information" is discussed in more detail below, it can first be thought of very broadly as any evaluative aspect a human interpreter may associate with a piece of data or information, such as whether it is understandable, timely, comprehensive, correct, or relevant. Since people have different expectations, perceptions, and evaluations of information, there is a huge potential for disagreement and even conflict in SocIS when it comes to contribution and curation of content. For examples, some Wikipedia editors prefer comprehensive articles (referred to as "inclusionism") while others (following "deletionism") prefer that articles be concise (Kostakis, 2010; "Deletionism and inclusionism in Wikipedia," 2019), often resulting in lengthy "edit wars" in which multiple editors mutually revert their changes to an article back and forth (Sumi et al., 2011). Hence, the general question arises of how SocIS—as systems of collaborative content production and consumption by large, heterogenous communities—can manage to mediate different perceptions of quality with respect to data/information and establish shared perceptions of quality among community members. This is important not only to researchers and businesses interested in data from SocIS, but also affects the interactions and collaboration of SocIS members and thus the survival and success of such platforms (Lin and Lee, 2006; Lin, 2008; Hew, 2009).

We can distinguish conceptually between three dimensions of quality of data/information in a SocIS. First, there can be agreement or disagreement between community members about the quality of *a particular piece* of data/information, that is, whether it is complete, accurate, relevant, and so forth. We call this the *descriptive dimension*. Second, there can be agreement or disagreement about the very *definition of what constitutes quality*, that is, about the general normative standards that should be applied to evaluate any piece of data/information and that should be adhered to when providing data/information. We call this the *normative dimension*. Finally, it must be acknowledged that SocIS—like any IS—are systems of semiotic (i.e., sign-based) communication and interaction (cf. Lyytinen, 1985; Mingers, 1995; Stamper, 1996; Mingers and Standing, 2018) in which not only information itself is distributed, but also signs (or, more generally, data) from which individuals can only subjectively construct information and meaning (Mingers and Standing, 2018). Hence, there can also be agreement or disagreement about the *information interpreted from data*, which we call the *semiotic dimension*. Conceptually, all three dimensions are relevant when thinking about and trying to improve the overall assessment of information quality (IQ) in a SocIS, because communication and collaboration within the community might fail due to misinterpretation of data (semiotic dimension), different views of what generally constitutes good information (normative dimension), or different evaluations of a particular piece of information in a specific situation (descriptive dimension).

SocIS platforms take different steps to address these problems. Many platforms have introduced features such as rating, voting, or recommending through which they collect community members' quality assessments of particular pieces of content (Chen, Xu and Whinston, 2011). For example, online reviews at Amazon.com can be rated as "helpful" and answers at StackExchange.com can voted up or down. These and similar features work primarily in the descriptive dimension. Some platforms provide space deliberately devoted to discuss, among others, issues of interpretation of content (the semiotic dimension), such as, Talk Pages at Wikipedia ("Help Talk pages," 2019) and other wikis. With respect to the normative dimension, many platforms explicate normative guidelines (also referred to as "policies" or "community standards") stating what constitutes high-quality content as opposed to low-quality content. Wikipedia's "Policies and Guidelines" ("Wikipedia Policies and guidelines," 2019) or Facebook's "Community Standards" ("Community Standards | Facebook," 2019) are examples. These guidelines can influence the behavior of members in creating and evaluating content and, even more important, in developing a shared normative understanding of IQ (Stvilia, Twidale, Gasser and Smith, 2005; Beschastnikh, Kriplean and McDonald, 2008; Butler, Joyce and Pike, 2008; Forte and Bruckman, 2008). Shared normative IQ standards should generally have a positive effect on members' continued participation in a SocIS.

However, SocIS platforms differ markedly with respect to how these normative guidelines are developed. On the one hand, community members themselves may formulate and maintain the guidelines; that, for example, is the case for Wikipedia's policy environment. On the other hand, an organizational platform provider may develop the guidelines and impose them on the community, as Facebook and other commercial SocIS do. While there are examples of both approaches, it is still unclear which of the different

approaches to guideline development (i.e., community-developed vs. provider-imposed) is associated with higher agreement and greater satisfaction among members, thus indicating a stronger agreement among members regarding normative IQ standards. Therefore, this study investigates empirically whether there are systematic differences in guideline acceptance/satisfaction for community-developed vs. provider-imposed approaches across different platforms. Specifically, we ask the following two research questions:

RQ1: Is a community-developed or a provider-imposed approach to guideline development associated with higher levels of *agreement* of community members with the guidelines?

RQ2: Is a community-developed or provider-imposed approach to guideline development associated with higher levels of *satisfaction* of community members with the guidelines?

The remainder of this paper is structured as follows. Section 2 presents the conceptual background on SocIS and quality of data/information and discusses existing research on how SocIS try to maintain quality of content. Section 3 describes our methodological approach to answer the research questions. Results are presented in Section 4 and discussed in Section 5. We conclude in Section 6.

Background

In this section, we introduce the central concepts of SocIS and quality of data and information.

Social Information Systems

SocIS are information systems in which information and communication technology (ICT) allows digital social interaction among multiple actors who participate primarily voluntarily and in often new, digitally transformed ways, thus creating virtual communities (cf. Schlagwein et al., 2011). A related concept is social media, which can be defined as "a generic term for social interactions built on a multitude of digital media and technologies, which allow users to create and share content and to act collaboratively" (Schoder et al., 2013, p. 10); or as the "group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content" (Kaplan and Haenlein, 2010, p. 61). While the concept of social media is in many ways similar to SocIS, social media are often portrayed as media in which everyone can contribute and consume user-generated content (UGC), as opposed to traditional media (e.g., newspapers, television, radio broadcasting, corporate websites). We see SocIS as the more comprehensive concept, which includes social media as well as social networking sites, review portals, bulletin boards, personal blogs, short messaging services, collaborative projects, content-sharing platforms, and virtual worlds (Kaplan and Haenlein, 2010; Kietzmann, Hermkens, McCarthy and Silvestre, 2011).

The concept of SocIS can be defined more specifically with respect to three key characteristics. First, a SocIS is a socio-technical system (STS), like an IS in the tradition of the STS approach (Trist, 1981) that has long been adopted in IS research (Cherns, 1976; Bostrom and Heinen, 1977; Mumford, 1995, 2003; Clegg, 2000; Avgerou, Ciborra and Land, 2004). An STS comprises a social subsystem that includes, for example, individuals, groups, and their relationships and hierarchies, and a technical subsystem that includes, for example, technology, tasks, and processes. These social and technical subsystems are "intertwined in a complex web of mutual causality" (Trist, 1981, p. 13): they interact and influence each other; changes in one subsystem probably evoke changes in the other; and an IS is the phenomenon that emerges from the continuous interaction of the two subsystems (Lee, 2004). Second, the social subsystem is constituted by people that actually participate in the SocIS, primarily voluntarily and because of the affordances of digital communication and social interaction for various purposes rather than due to their role and related tasks in a formal organization. Third, the technical subsystem is oriented towards affording such digital communication and social interaction to an open group of people, rather than supporting organizational tasks and activities and being under the control of an organization (cf. Butler, 2001; Bagozzi and Dholakia, 2002; Schlagwein et al., 2011). SocIS are usually open in the sense that their members are not predetermined and required to use a SocIS, but rather have different individual motivations to use a SocIS and be a member of the community (Rheingold, 1993; Sproull and Arriaga, 2007). Affordances of social interaction of SocIS are typically provided continuously (Bagozzi and Dholakia, 2002), that is, they are not limited to a certain occasion, project, or otherwise predefined timeframe. Due to their openness and continuity, SocIS social subsystems or communities are potentially large, heterogeneous, and changing (Gu,

Konana, Rajagopalan and Chen, 2007; Ma and Agarwal, 2007; Agarwal, Gupta and Kraut, 2008; Xu, Yang, Cheng and Lim, 2014).

Research has identified a wide range of motives for people to use SocIS and become community members, for example, to experience identity with and belonging to the community (Bagozzi and Dholakia, 2002; Chiu, Hsu and Wang, 2006; Fang and Neufeld, 2009; Bateman, Gray and Butler, 2011); receive information, help, and to learn (Nov, 2007; Fang and Neufeld, 2009; Jin, Li, Zhong and Zhai, 2015); for self-expression/self-presentation and to gain the attention of others (Rui and Whinston, 2012; Toubia and Stephen, 2013); participate in a collaborative project (Shah, 2006; Bitzer, Schrettl and Schröder, 2007); exchange social support (Rheingold, 1993; Ridings and Gefen, 2004); find and maintain friendships (Rheingold, 1993; Wasko and Faraj, 2000; Ellison and boyd, 2013); for fun and recreation (Wasko and Faraj, 2000; Bitzer et al., 2007; Nov, 2007); for reasons of cooperation and altruism (Bitzer et al., 2007; Anthony, Smith and Williamson, 2009); to build reputation and recognition within a community (Wasko and Faraj, 2005; Roberts, Hann and Slaughter, 2006; Anthony et al., 2009; Fang and Neufeld, 2009); due to feelings of obligation to remain and reciprocate contributions (Chiu et al., 2006; Bateman et al., 2011); and/or as a result of commitment to the community's goals (Anthony et al., 2009; Budhathoki and Haythornthwaite, 2013). Different types of SocIS correspond differently to these needs, motivations, and goals, and as members are typically free to leave at any time, communities in SocIS always face the issue of discontented members dropping out (boyd and Ellison, 2007; Ren et al., 2012; Xu et al., 2014; Bock, Ahuja, Suh and Yap, 2015, p. 2015).

When trying to explain IS adoption/acceptance and (continued) use, IS research has often distinguished between *utilitarian IS* (those that provide instrumental value to the user) and *hedonic IS* (those that provide self-fulfilling value) (van der Heijden, 2004; Lowry et al., 2013). This distinction has been adopted from the distinction between utilitarian and hedonic products in consumer behavior research (Hirschman and Holbrook, 1982; Holbrook and Hirschman, 1982; van der Heijden, 2004). However, given such a wide range of motives that can explain SocIS use and that include motivations from all three main categories of hedonic, intrinsic, and extrinsic motivations identified by Lowry and colleagues (Lowry, Gaskin and Moody, 2015), it is difficult to categorize SocIS generally as either hedonic IS or utilitarian IS. Rather, SocIS are "mixed-motivation systems" (Lowry et al., 2013, p. 618), that is, their adoption and use depend on a mix of hedonic, intrinsic, and extrinsic motivations that can vary among users of the same SocIS platform as well as across platforms.

Information Quality in SocIS

Although "information" is central to IS research, the term is used with various different connotations and authors are often not explicit about their specific understanding of the concept of information and its delineation from other concepts such as "data" and "meaning" (McKinney and Yoos, 2010; Boell, 2017; Mingers and Standing, 2018). Hence, there have repeatedly been calls for and attempts at more thorough engagement with "information" at a conceptual level (Lee, 2004, 2010; Baskerville, 2010; McKinney and Yoos, 2010; Boell, 2017) and several authors have compared the different conceptualizations of "information" (e.g., Mingers, 1996; Bates, 2006; McKinney and Yoos, 2010; Floridi, 2011), of which the works by Boell and Cecez-Kecmanovic (Boell and Cecez-Kecmanovic, 2015; Boell, 2017) and Mingers and colleagues (Mingers, 2013; Mingers and Willcocks, 2014; Mingers and Standing, 2018) are probably the most recent and comprehensive in the IS field. We follow Mingers and Standing in defining information as "the *relationship* between a token, sign or message and the event(s) that caused it" (Mingers and Standing, 2018, p. 87; emphasis in original). Tokens, signs, messages or—more generally—physical differences are the manifestation of some causal event and thus signify their cause (Mingers and Standing, 2018). Data are "a collection of signs, derived from differences, put together for a purpose [and if] well-formed and correct … it may carry information" (Mingers and Standing, 2018, p. 87). Thus, information exists independent of humans, and is objective and veridical in the sense that, ontologically, physical differences and signs unequivocally carry the information about the events that have caused them (Mingers and Standing, 2018). However, regarding the epistemology of information (i.e., how humans can experience and know about information carried by signs), Mingers and Standing argue that "[t]he information *that is available for a particular observer* depends on the prior knowledge of the observer. If the observer does not know the code or language, the sign may carry no information for that observer" (Mingers and Standing, 2018, p. 87; emphasis in original). And further, "*[i]mport* is the meaning for, or effect on, a receiver of a particular

dataset or message. It depends on the receiver's knowledge and state of mind at the time. Thus, information is objective, while import is subjective" (Mingers and Standing, 2018, p. 87; emphasis in original).

Questions and problems related to the quality of data and information have attracted the attention of many researchers, research programs, and government efforts (Batini, Cappiello, Francalanci and Maurino, 2009; Madnick, Wang, Lee and Zhu, 2009; Illari and Floridi, 2014). Several definitions of quality of data/information have been proposed and even more ways to measure it in a practical or research context (Batini et al., 2009; Madnick et al., 2009; Sadiq, Yeganeh and Indulska, 2011; Jayawardene, Sadiq and Indulska, 2013; Xiao, Lu, Liu and Zhou, 2014). For example, Wang and Strong explicitly adopted a concept of quality from the broader literature on quality and conceptualized data quality in terms of "fitness for use," that is, high-quality data are "data that are fit for use by data consumers" (Wang and Strong, 1996, p. 6). Other less use-oriented conceptualizations include, for example, that data quality is the degree to which data correspond to the respective entities in the real world, proposed, for example, by Wand and Wang (1996) and Orr (1998). These and other definitions often explicitly view (organizational) IS as information product manufacturing systems for which consumers assess the quality of their (information product) output (e.g., Wang, Storey and Firth, 1995; Ballou, Wang, Pazer and Tayi, 1998; Wang, 1998; Kahn, Strong and Wang, 2002). In a recent review of concepts of data and information quality (DIQ) in IS research and their applicability to SocIS, two of the present authors identified several different DIQ concepts, many of which explicitly take the perspective of data/information consumers to conceptualize DIQ (Tilly, Posegga, Fischbach, and Schoder, 2017). However, we found that most of the DIQ concepts—while also applied in research on SocIS—make assumptions, often attributable to the organizational context, that conflict conceptually with the nature and characteristics of SocIS because SocIS rely on mutual interaction rather than fixed roles of producers and consumers; because perceptions of DIQ in SocIS are inherently heterogeneous and dynamic, just as the SocIS' communities are; and because members' perceptions of DIQ are often implicit to their actions in the SocIS rather than explicit (Tilly et al., 2017).

Based on the concept of information as developed by Mingers and Standing (2018), IQ can generally be conceptualized as how well someone can construct a subjective signifying relationship between a particular sign (or, more generally, data) and some causal event (i.e., obtain information) and how valuable or important the meaning derived from such information is, based on some normative evaluative standard. This implies that for SocIS with potentially large and heterogenous communities of voluntary members from different backgrounds and with different motivations that engage in mutual social interaction, members might exhibit different shared interpretations of data on the semiotic dimension, different IQ standards on the normative dimension, and divergent evaluations of information on the descriptive dimension. Therefore, to improve the overall IQ as perceived by their members, SocIS need to take measures to facilitate the development of shared interpretations, shared normative standards, and complementary evaluations of information. Thus, facilitating good overall IQ is very different from many traditional concepts of DIQ and related management approaches that originated in the context of organizational IS (Tilly et al., 2017). According to their perspective of IS as information product manufacturing systems within a particular organization, providing high-quality information basically means tailoring IS to the needs and interpretations of predefined information consumers. In contrast, the approach for SocIS providers is to provide the community with the means and mechanisms to negotiate and approximate effectively along the descriptive, normative, and semiotic dimensions.

Approaches to Develop Information Quality Guidelines in SocIS

We can distinguish between two fundamentally different approaches of SocIS to develop IQ guidelines. First, guidelines may be *developed by the community members themselves*, that is, by those people that participate in the SocIS and interact with each other by creating, sharing, and consuming content. Some members may, over time, attain more influence on the development of guidelines than others (e.g., due to meritocratic mechanisms), but participation is, in principle, open to anyone. In this case, members participate in developing the community's understanding of IQ on both the descriptive and normative dimensions, that is, they can provide their assessment of quality of a particular piece of information based on their personal standards and those of the community (descriptive dimension) as well as participate in the development of the normative standards themselves (normative dimension). Thus, a community-developed approach would allow many members of the SocIS to voice their opinions on fundamental questions of quality, discuss and negotiate them, and integrate them into the community's standards. Existing research has shown, for example, that over time the set of policies on Wikipedia grows, becomes

more differentiated, and is adopted by community members, thus helping to settle individual disputes over article quality (Stvilia et al., 2005; Beschastnikh et al., 2008; Butler et al., 2008). Discussions of individual cases can result in broader changes to existing policies or the creation of new policies (Schneider, Passant and Decker, 2012), thus preserving normative standards for similar cases in the future. Guidelines also provide a shared language and knowledge to the community, which are also directly available to new members, to assess and negotiate quality (Beschastnikh et al., 2008; Jones, 2008). Conversely, one can assume that arguing with (many) other people over the Internet will not always result in positive outcomes. For example, Wikipedia has become known for its so-called "edit wars" ("Wikipedia Edit warring" 2019) in which multiple editors disagree about changes that should or should not be made to an article, and mutually edit and revert other editors' edits back and forth (Sumi et al., 2011; e.g., Kalyanasundaram, Wei, Carley and Herbsleb, 2015). There is no guarantee that large-scale participatory development approaches successfully result in agreed-upon and satisfying guidelines and similar evaluations of information.

Second, guidelines may be *set by the platform provider and imposed on the community and its members*. The platform provider (person or organization) has, *as a matter of course*, more influence than community members on the SocIS in general and on the development of IQ guidelines in particular. The provider decides whether to grant members of the SocIS community more or less influence on the normative guidelines and their application, thus limiting their participation in developing the community's understanding of IQ primarily to the descriptive dimension. In other words, members may individually assess whether a particular piece of information complies with the normative standards of the SocIS, and probably suggest its promotion if it does or report it to the provider if not does not. In fact, in addition to editorial content moderation teams and algorithms for automatic detection of problematic content, many platform providers rely on their communities to identify and report content that may violate normative standards (Crawford and Gillespie, 2016; Gillespie, 2018). But community members typically have limited or virtually no influence on setting the normative standards applied in such evaluations. This approach is common for commercial SocIS platforms such as Facebook ("Community Standards | Facebook," 2019), Twitter ("Rules and policies | Twitter," 2019), and YouTube ("Policies | YouTube," 2019) (Gillespie, 2018).

The approach of provider-imposed guidelines often provokes disagreement and dissatisfaction among community members, probably because their views are not represented in the guidelines and they disagree with the providers' assessments; such was the case, for instance, when Facebook removed and later restored a well-known, Pulitzer Prize-winning Vietnam War photo (Scott and Isaac, 2016). When guidelines are carefully developed by experienced online editors or community managers on behalf of a platform provider and enforced by a well-versed team of professionals in a transparent and standardized process, however, the outcome could—at least theoretically—be better than if thousands of community members argue with each other. In the end, community members might be more satisfied with a provider-imposed approach when the resulting guidelines actually help settle their conflicts with other members and thus improve IQ overall, although they did not participate in the process of setting the guidelines. Some platforms are working on making their content moderation processes more transparent and open in an effort to increase members' trust and acceptance ("Facebook makes its community guidelines public and introduces an appeals process | The Verge," 2018; "Facebook releases content moderation guidelines—rules long kept secret | The Guardian," 2018).

Although existing research indicates some causal effects with respect to different approaches to developing and applying IQ guidelines in SocIS, much is still unknown. With respect to the effects of community-developed versus provider-imposed approaches on the acceptance of and satisfaction with the guidelines, the research suggests mixed implications, and which of the above-mentioned hypotheses hold is unclear. There also seems to be no common theoretical framework for exploring these questions. Hence, our research sets out to investigate in an exploratory way whether there are systematic differences in guideline acceptance/satisfaction for community-developed versus provider-imposed approaches across different platforms.

Methodology

We conducted online surveys on multiple SocIS platforms to investigate how the two different approaches to guideline development—namely, community-developed vs. provider-imposed—are associated with SocIS members' agreement to and satisfaction with the guidelines for DIQ. The approach is structured into

three steps: (1) selecting platforms for both approaches; (2) developing and conducting platform-specific surveys; and (3) analyzing survey data. These are described in the following subsections.

Selecting Platforms

The selection of SocIS platforms for this study was based on their popularity in terms of size and page traffic. Popularity was assessed based on Wikipedia's lists of most popular websites ("List of most popular websites," 2019), social networking websites ("List of social networking websites," 2019), and wikis ("List of wikis," 2019), as well as the Alexa ranking of page traffic ("Website Traffic, Statistics and Analytics - Alexa," n.d.). Several filter criteria were further applied. Platforms had to exhibit characteristics of SocIS (see Section 2): they needed to state normative guidelines regarding DIQ explicitly; and, for practical reasons, they needed to offer the possibility to disseminate links to online surveys and needed to be accessible either in English or German. To ensure relevance of the study and to increase the number of survey participants, only the most popular SocIS (after applying the filter criteria) were considered. In descending order of popularity, platforms were screened regarding their approach to guideline development and categorized as either community-driven or provider-imposed. The screening and selection of platforms was continued until 10 platforms had been identified for each of the approaches (20 in total).

Developing and Conducting Platform-specific Surveys

The surveys were designed to investigate the differences between community-driven and provider-imposed approaches to guideline development with respect to two constructs: the agreement of participants to normative statements regarding DIQ made in the guidelines; and the satisfaction of participants with the given guideline framework. It is important to note that the surveys were designed, phrased, and distributed in such a way as to ensure that participants answered with respect to the platform they used (see also the next subsection on how the surveys were conducted), as is explained in the following. Appendix A provides a template of the survey questionnaire.

The first construct (agreement to guidelines) was measured by eleven items. For each platform, we derived ten platform-specific items from the guidelines stated on that particular platform. These were formulated as normative statements to which participants were asked to indicate their agreement. Some platforms already offered brief summaries of guidelines that could be included directly as survey items. For example, Wikipedia provides brief "in a nutshell" summaries of otherwise quite extensive policies.[1] For those platforms that did not offer such summaries of guidelines, survey items were built from the full texts by means of qualitative content analysis (Mayring, 2010). That is, full texts of guidelines were paraphrased and categorized inductively under broader normative statements. For example, YouTube's guideline on providing proper context for content ("The importance of context - YouTube Help," n.d.) was thus converted to the following survey item: "Videos with terrorist content may be allowed if sufficient contextual information has been provided and the purpose is understandable." Participants were asked to indicate their agreement to these ten statements on a 7-point scale, "1" being "disagree" and "7" being "agree." Although results for these ten items cannot be compared directly to each other across platforms, the average of the ten items for each survey participant should provide a measure of the overall agreement of participants to their platform's guideline framework. As a control item, participants were also asked directly the degree (from "not at all" to "completely") to which they agreed with the platform's guidelines.

The second construct (satisfaction with guidelines) was measured by four items asking participants whether they think guidelines are missing (reverse coded) and whether existing guidelines are mature, applicable, and sufficient.

Further, data on general attitudes and behavior regarding content quality, content creation and consumption, and Internet and social media use, as well as demographic data, were collected as control variables. As explained earlier, a modified version of the survey questionnaire template was created and distributed specifically to members of each platform. Depending on the specifics of the platform, different ways of distributing the links to the online surveys had to be employed. Some platforms allowed us to

---

[1] For example, https://en.wikipedia.org/wiki/Wikipedia:Neutral_point_of_view

contact randomly selected members directly, some had internal or external discussion forums separated from the primary content, and for other platforms it was only possible to post the link publicly.

Analyzing Survey Data

Based on the raw survey data, the latent constructs of "agreement to guidelines" (average of items 1 through 10) and "satisfaction with guidelines" (average of items 12 through 15; item 12 reversed) were constructed and construct reliability was assessed. We analyzed the survey data with respect to the research question in two ways. First, group comparisons (community-developed vs. provider-imposed) were conducted for the two constructs "agreement to guidelines" and "satisfaction with guidelines," both visually (box and scatter plots) and statistically (t-test for unpaired samples). Second, to assess the relationship of different guideline development approaches while controlling for other possible influences, different linear models were fitted for both constructs. The minimal models included the approach to guideline development as well as demographic data (education, sex, age). Control models also included measures of general Internet and social media use, attitudes and behavior towards content creation and consumption, and quality, as well as knowledge about community guidelines. Finally, full models were estimated in which the dependent constructs "agreement to guidelines" and "satisfaction to guidelines" were included as explanatory variables for the each other.

Results

From December 2018 through March 2019, online surveys for 20 SocIS platforms were designed and disseminated on the platforms themselves or on related platforms. In total, 227 participants responded to the surveys; the distribution of participants across platforms, however, is quite skewed. Table 1 provides an overview of selected platforms, categorization in community-developed or provider-imposed guideline development, and the number of survey participants. Participants were mostly German (71.8%), male (71.4%), between 18 and 29 years old (42.3%), and have a bachelor's or equivalent degree (33.5%). The three most frequent platforms in the data set are Facebook (23.3%), German Wikipedia (15.0%), and English-language Wikipedia (11.5%). No responses were collected on five of the platforms due to platform policies or administrator blocks. With respect to the distribution of answers between the two approaches to guideline development, the data set is quite balanced (provider-imposed: 52.9%; community-developed: 47.1%). Appendix A provides descriptive statistics of the constructs/items.

Provider-imposed Approach		Community-developed Approach	
Facebook	53	Wikipedia DE	34
YouTube	10	Wikipedia EN	26
Twitch	23	Wowpedia	8
Pinterest	3	Dota2pedia	0
Twitter	14	Wikibooks	5
Tumblr	0	Wiktionary	12
Yelp	8	Wikiquote	0
Google Plus	4	Wikisource	0
Instagram	4	wikiHow DE	0
LinkedIn	1	wikiHow EN	22
Total	120	Total	107

Table 1. Platforms Included in the Study and Number of Responses

Direct comparisons between the two approaches to guideline development with respect to "agreement to guidelines" reveal that members of platforms that follow a community-developed approach (M = 5.936, SD = 0.845) on average agree significantly stronger ($t(224.983)$ = -6.522, p = .000) to the guidelines than do members of platforms with a provider-imposed approach (M = 5.157, SD = 0.956). The results are similar for member satisfaction with guidelines. Members of platforms with community-developed guidelines (M

= 5.414, *SD* = 1.110) are significantly more satisfied (t(220.14) = -8.680, p = .000) with the guidelines than are members of platforms with provider-imposed guidelines (*M* = 4.152, *SD* = 1.074). Visual comparisons by boxplots and scatterplots (see Figure 1) corroborate these findings, although they also show differences between platforms and between members of one platform as well as overlap between the two approaches.

Participant self-assessment regarding agreement to guidelines (item 11) confirms the higher levels of agreement for the community-developed approach. Participants from these platforms stated more often that they agree "completely" or "for the most part" with the guidelines than participants from the group of provider-imposed guidelines (χ^2(6) = 40.843, p = .000).

Linear models support the findings from group comparisons. The community-developed guidelines approach is significantly positively in all three models (minimal, control, full) with respect to agreement to guidelines (β = .572, p < .01) as well as regarding satisfaction with guidelines (β = .758, p < .01)[2]. Further, participants who stated that they adhere to guidelines "approximately," "predominantly," or "as accurately as possible" when creating content also tend to agree with those guidelines (β = .568/.516/.540, respectively; p < .1); participants with a doctoral degree tend to disagree with them (β = -.625, p < .1). Table 2 reports the results from linear models (for reasons of space, category variables are reported only as either "included" or not).

**Figure 1. Comparison of Agreement and Satisfaction for
Different Approaches to Guideline Development**
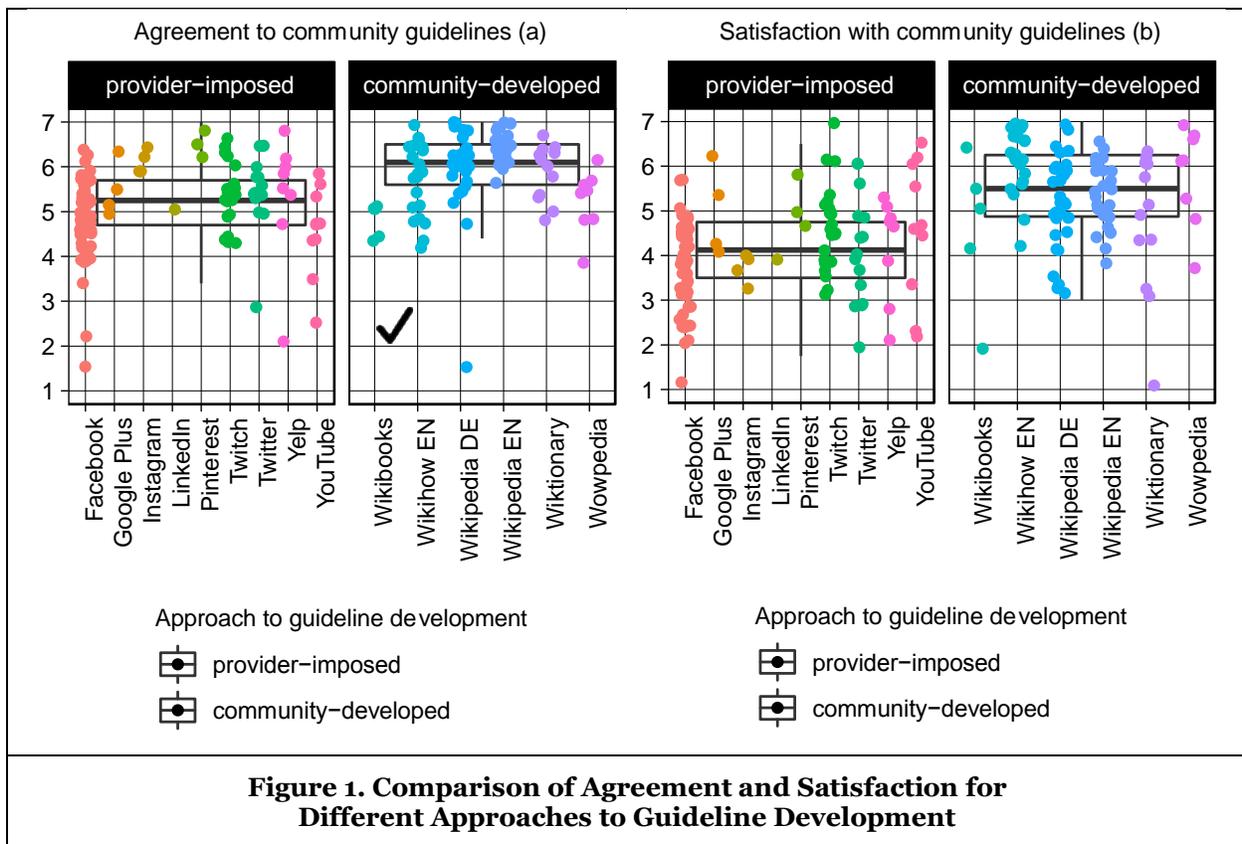
Discussion

With respect to our two research questions, the empirical findings indicate that members of platforms with community-developed guidelines agree more strongly with their platforms' guidelines (RQ1) and are more satisfied with them (RQ2) than are members of platforms with guidelines developed and imposed by the platform provider. These differences also persist when controlling for differences in social media use

---

[2] All model coefficients refer to the full models.

behavior and demographic differences among members. These results support the hypothesis that more participatory approaches to guideline development are more effective in establishing shared normative standards of IQ among community members.

Dependent variable	Agreement to community guidelines			Satisfaction with community guidelines		
Model	minimal	control	full	minimal	control	full
Approach to guideline development: community-developed	.852‡	.658‡	.572‡	1.256‡	.836‡	.758‡
	(.140)	(.178)	(.187)	(.164)	(.191)	(.198)
The quality of the content I create is important to me		.085	.086		-.016	-.026
		(.053)	(.053)		(.057)	(.057)
The quality of the content I consume is important to me		.081	.084		-.034	-.044
		(.057)	(.057)		(.061)	(.061)
I know the community guidelines for creating content		-.023	-.037		.133†	.135†
		(.055)	(.056)		(.059)	(.059)
Satisfaction with community guidelines			.103			
			(.068)			
Agreement with community guidelines						.118
						(.079)
Constant	4.659‡	3.727‡	3.484‡	4.078‡	2.368†	1.926*
	(.409)	(.917)	(.928)	(.479)	(.985)	(1.024)
Frequency of Internet use		included	included		included	included
Frequency of social media use		included	included		included	included
Frequency of {platform} use		included	included		included	included
I create new content/edit content myself		included	included		included	included
I regularly consume content		included	included		included	included
I am consciously implementing the community guidelines for creating content		included	included		included	included
Highest degree (yet)	included	included	included	included	included	included
Age	included	included	included	included	included	included
Gender	included	included	included	included	included	included
Observations	227	227	227	227	227	227
R²	.235	.353	.361	.358	.544	.549
Adjusted R²	.185	.210	.215	.316	.442	.446
Note:	*p<0.1; †p<0.05; ‡p<0.01					

Table 2. Linear Models for Agreement to and Satisfaction with Community Guidelines

Further, while the overall agreement of members to specific guidelines was rather positive for both approaches, satisfaction with the current state of the guideline framework was rated lower and also diverged more between the two approaches, with members of SocIS with community-developed guidelines more satisfied than members of SocIS with provider-imposed guidelines. This supports the hypothesis that even

though both approaches result in IQ guidelines that are somewhat agreed upon by the respective communities, integrating community members directly in the process of negotiating IQ guidelines in SocIS leads to higher satisfaction, probably due to stronger identification with the resulting guidelines.

Our findings suggest that platform providers that currently impose guidelines on their communities should seek ways to provide opportunities for community members to participate not only in their application (descriptive dimension), but also in the development of guidelines (normative dimension). A combined approach, such as an editorial team working on the platform provider's behalf but that also offers systematic ways for the community to participate in the development of guidelines, might even outperform either individual approach. Established crowdsourcing approaches such as ideation platforms (Huang, Vir Singh and Srinivasan, 2014) could be tested for this purpose.

Given that there has been little research to date on how community guidelines can be developed and the effects of different approaches on, for example, members' agreement and satisfaction, these initial yet promising findings call for further research on the normative dimension of IQ in SocIS. First, as our study was limited to 20 platforms (with participants from 15 platforms), the empirical research should be extended to more platforms. Another limitation to empirical data collection is that all platforms in the community-developed group were some kind of wiki. This was probably due to our sampling of platforms by popularity and the fact that the MediaWiki software, which affords community-driven guideline development by design, is widely used. In other words, when following a community approach to guideline development, MediaWiki seems to be the leading software. Different sampling approaches should ensure that other variants of community-developed guidelines are included in future research.

Second, because of our research methodology, causal conclusions cannot be drawn from our results to explain the higher levels of agreement and satisfaction for community-developed guidelines. Therefore, further research should investigate potential causal drivers for this empirical pattern. A straightforward hypothesis would be that more inclusive and participatory development approaches can resolve many normative issues and thus lead to agreed-upon and satisfying guidelines that reflect shared normative standards of IQ. However, the empirical pattern could also be driven, for example, by selection effects (i.e., people with similar normative IQ standards choose more participatory platforms). These and other explanations could be evaluated by means of qualitative and/or experimental research methodologies.

Finally, our research focused on *explicit* normative standards as laid down in the communities' guidelines, while normative standards also exist in the minds of SocIS members. Further, the internal/mental normative standards are probably more important to everyday interactions of members and their assessment and creation of content. Hence, further research should investigate the degree to which these mental normative standards overlap between members and with the explicit normative standards.

Conclusion

SocIS have potentially large, heterogeneous communities of voluntary members from different backgrounds and with different purposes and goals that engage in mutual social interaction. Hence, their interpretations and evaluations of information shared in communication and collaboration can be very different, meaning it is difficult for SocIS to achieve high levels of IQ for all their members. This is very different from IQ in traditional organizational IS, in which IQ is typically defined and tailored with respect to the needs of specific consumers of "information products." SocIS, in contrast, need to take steps to facilitate the development of shared interpretations, shared normative standards, and complementary evaluations of information among their members. There are different ways SocIS can do so, one of which is the explicit statement of normative IQ guidelines. However, these guidelines can be developed either by the community itself or by the platform provider, and the question is which approach is associated with higher member agreement to and satisfaction with the guidelines. Based on an empirical survey involving 227 participants from 15 different SocIS platform, this study found that members of platforms that employ the community-developed approach both agree more strongly and are more satisfied with their platform's guidelines.

Hence, we assume that the particular approach to develop normative IQ guidelines has an effect on the degree to which platform members develop shared normative standards of IQ, which would positively affect their assessment of IQ and thus their continued use of the platform (see Lin and Lee, 2006; Lin, 2008; Hew, 2009). These findings are relevant for platform providers, as they point towards some ways to

establish shared IQ standards that might be more effective than others. Our study thus contributes to the research on IQ in SocIS, both empirically and conceptually.

References

Agarwal, R., A. K. Gupta and R. Kraut. (2008). "The Interplay Between Digital and Social Networks." *Information Systems Research*, *19*(3), 243–252.

Anthony, D., S. W. Smith and T. Williamson. (2009). "Reputation and Reliability in Collective Goods - The Case of the Online Encyclopedia Wikipedia." *Rationality and Society*, *21*(3), 283–306.

Avgerou, C., C. Ciborra and F. Land (Eds.). (2004). *The social study of information and communication technology: innovation, actors and contexts*. Oxford ; New York: Oxford University Press.

Bagozzi, R. P. and U. M. Dholakia. (2002). "Intentional social action in virtual communities." *Journal of Interactive Marketing*, *16*(2), 2–21.

Ballou, D., R. Y. Wang, H. Pazer and G. K. Tayi. (1998). "Modeling Information Manufacturing Systems to Determine Information Product Quality." *Management Science*, *44*(4), 462–484.

Baskerville, R. L. (2010). "Knowledge lost and found: a commentary on Allen Lee's "retrospect and prospect."" *Journal of Information Technology*, *25*(4), 350–351.

Bateman, P. J., P. H. Gray and B. S. Butler. (2011). "The Impact of Community Commitment on Participation in Online Communities." *Information Systems Research*, *22*(4), 841–854.

Bates, M. J. (2006). "Fundamental forms of information." *Journal of the American Society for Information Science and Technology*, *57*(8), 1033–1045.

Batini, C., C. Cappiello, C. Francalanci and A. Maurino. (2009). "Methodologies for Data Quality Assessment and Improvement." *ACM Computing Surveys*, *41*(3), 16:1–16:52.

Beschastnikh, I., T. Kriplean and D. McDonald. (2008). "Wikipedian Self-Governance in Action: Motivating the Policy Lens." In: *Proceedings of the 2008 AAAI International Conference on Weblogs and Social Media*.

Bitzer, J., W. Schrettl and P. J. H. Schröder. (2007). "Intrinsic motivation in open source software development." *Journal of Comparative Economics*, *35*(1), 160–169.

Bock, G.-W., M. K. Ahuja, A. Suh and L. X. Yap. (2015). "Sustainability of a Virtual Community: Integrating Individual and Structural Dynamics." *Journal of the Association for Information Systems*, *16*(6), 418–447.

Boell, S. K. (2017). "Information: Fundamental positions and their implications for information systems research, education and practice." *Information and Organization*, *27*(1), 1–16.

Boell, S. K. and D. Cecez-Kecmanovic. (2015). "What is "Information" Beyond a Definition?" *Proceedings of the 36th International Conference on Information Systems (ICIS)*.

Bostrom, R. P. and J. S. Heinen. (1977). "MIS Problems and Failures: A Socio-Technical Perspective. Part I: The Causes." *MIS Quarterly*, *1*(3), 17–32.

boyd, danah m. and N. B. Ellison. (2007). "Social Network Sites: Definition, History, and Scholarship." *Journal of Computer-Mediated Communication*, *13*(1), 210–230.

Budhathoki, N. R. and C. Haythornthwaite. (2013). "Motivation for Open Collaboration Crowd and Community Models and the Case of OpenStreetMap." *American Behavioral Scientist*, *57*(5), 548–575.

Butler, B. S. (2001). "Membership Size, Communication Activity, and Sustainability: A Resource-Based Model of Online Social Structures." *Information Systems Research*, *12*(4), 346–362.

Butler, B. S., E. Joyce and J. C. Pike. (2008). "Don't Look Now, but We've Created a Bureaucracy: The Nature and Roles of Policies and Rules in Wikipedia." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1101–1110). New York, NY, USA: ACM.

Chen, J., H. Xu and A. B. Whinston. (2011). "Moderated Online Communities and Quality of User-Generated Content." *Journal of Management Information Systems*, *28*(2), 237–268.

Cherns, A. (1976). "The Principles of Sociotechnical Design." *Human Relations*, *29*(8), 783–792.

Chiu, C.-M., M.-H. Hsu and E. T. G. Wang. (2006). "Understanding Knowledge Sharing in Virtual Communities: An Integration of Social Capital and Social Cognitive Theories." *Decision Support Systems*, *42*(3), 1872–1888.

Clegg, C. W. (2000). "Sociotechnical principles for system design." *Applied Ergonomics*, *31*(5), 463–477.

"Community Standards | Facebook." (2019). Retrieved from https://www.facebook.com/communitystandards/introduction/

Crawford, K. and T. Gillespie. (2016). "What Is a Flag for? Social Media Reporting Tools and the Vocabulary of Complaint." *New Media & Society*, *18*(3), 410–428.

"Deletionism and inclusionism in Wikipedia." (2019, April 27). In: *Wikipedia*.

Ellison, N. B. and danah m. boyd. (2013). "Sociality Through Social Network Sites." In: *The Oxford Handbook of Internet Studies*.

"Facebook makes its community guidelines public and introduces an appeals process | The Verge." (2018, April 24). Retrieved from https://www.theverge.com/2018/4/24/17270910/facebook-community-guidelines-appeals-process

"Facebook releases content moderation guidelines – rules long kept secret | The Guardian." (2018, April 24). Retrieved from https://www.theguardian.com/technology/2018/apr/24/facebook-releases-content-moderation-guidelines-secret-rules

Fang, Y. and D. Neufeld. (2009). "Understanding Sustained Participation in Open Source Software Projects." *Journal of Management Information Systems*, *25*(4), 9–50.

Figueiredo, F., H. Pinto, F. Belém, J. Almeida, M. Gonçalves, D. Fernandes and E. Moura. (2013). "Assessing the quality of textual features in social media." *Information Processing & Management*, *49*(1), 222–247.

Floridi, L. (2011). *The Philosophy of Information*. Oxford University Press.

Forte, A. and A. Bruckman. (2008). "Scaling Consensus: Increasing Decentralization in Wikipedia Governance." In: *Proceedings of the 41st Annual Hawaii International Conference on System Sciences*. Washington, DC, USA: IEEE Computer Society.

Gillespie, T. (2018). *Custodians of the Internet: platforms, content moderation, and the hidden decisions that shape social media*.

Gu, B., P. Konana, B. Rajagopalan and H.-W. M. Chen. (2007). "Competition Among Virtual Communities and User Valuation: The Case of Investing-Related Communities." *Information Systems Research*, *18*(1), 68–85.

"Help Talk pages." (2019, March 1). Retrieved from https://en.wikipedia.org/w/index.php?title=Help:Talk_pages&oldid=885651521

Hew, K. F. (2009). "Determinants of Success for Online Communities." *Behaviour & Information Technology*, *28*(5), 433–445.

Hirschman, E. C. and M. B. Holbrook. (1982). "Hedonic Consumption: Emerging Concepts, Methods and Propositions." *Journal of Marketing*, *46*(3), 92–101.

Holbrook, M. B. and E. C. Hirschman. (1982). "The Experiential Aspects of Consumption: Consumer Fantasies, Feelings, and Fun." *Journal of Consumer Research*, *9*(2), 132–140.

Huang, Y., P. Vir Singh and K. Srinivasan. (2014). "Crowdsourcing New Product Ideas Under Consumer Learning." *Management Science*, *60*(9), 2138–2159.

Illari, P. and L. Floridi. (2014). "Information Quality, Data and Philosophy." In: L. Floridi & P. Illari (Eds.), *The Philosophy of Information Quality* (pp. 5–24). Cham: Springer.

Jayawardene, V., S. Sadiq and M. Indulska. (2013). *An analysis of data quality dimensions* (ITEE Technical Report No. 2013– 01) (pp. 1–32). Retrieved from http://espace.library.uq.edu.au/view/UQ:312314

Jin, J., Y. Li, X. Zhong and L. Zhai. (2015). "Why users contribute knowledge to online communities: An empirical study of an online social Q&A community." *Information & Management*.

Jones, J. (2008). "Patterns of Revision in Online Writing A Study of Wikipedia's Featured Articles." *Written Communication*, *25*(2), 262–289.

Kahn, B. K., D. M. Strong and R. Y. Wang. (2002). "Information Quality Benchmarks: Product and Service Performance." *Communications of the ACM*, *45*(4), 184–192.

Kalyanasundaram, A., W. Wei, K. M. Carley and J. D. Herbsleb. (2015). "An Agent-based Model of Edit Wars in Wikipedia: How and when is Consensus Reached." In: *Proceedings of the 2015 Winter Simulation Conference* (pp. 276–287). Piscataway, NJ, USA: IEEE Press.

Kaplan, A. M. and M. Haenlein. (2010). "Users of the World, Unite! The Challenges and Opportunities of Social Media." *Business Horizons*, *53*(1), 59–68.

Kietzmann, J. H., K. Hermkens, I. P. McCarthy and B. S. Silvestre. (2011). "Social media? Get serious! Understanding the functional building blocks of social media: SPECIAL ISSUE: SOCIAL MEDIA." *Business Horizons*, *54*(3), 241–251.

Kostakis, V. (2010). "Identifying and Understanding the Problems of Wikipedia's Peer Governance: The Case of Inclusionists versus Deletionists." *First Monday*, *15*(3).

Lee, A. S. (2004). "Thinking about Social Theory and Philosophy for Information Systems." In: J. C. Mingers & L. Willcocks (Eds.), *Social Theory and Philosophy for Information Systems* (pp. 1–26). Chichester, West Sussex, England ; Hoboken, NJ: J. Wiley.

Lee, A. S. (2010). "Retrospect and prospect: information systems research in the last and next 25 years." *Journal of Information Technology*, *25*(4), 336–348.

Lin, H.-F. (2008). "Determinants of Successful Virtual Communities." *Information & Management*, *45*(8), 522–527.

Lin, H.-F. and G.-G. Lee. (2006). "Determinants of success for online communities." *Behaviour & Information Technology*, *25*(6), 479–488.

"List of most popular websites." (2019, April 28). In: *Wikipedia*.

"List of social networking websites." (2019, April 26). In: *Wikipedia*.

"List of wikis." (2019, April 26). In: *Wikipedia*.

Lowry, P., J. Gaskin and G. Moody. (2015). "Proposing the Multimotive Information Systems Continuance Model (MISC) to Better Explain End-User System Evaluations and Continuance Intentions." *Journal of the Association for Information Systems*, *16*(7), 515–579.

Lowry, P., J. Gaskin, N. Twyman, B. Hammer and T. Roberts. (2013). "Taking "Fun and Games" Seriously: Proposing the Hedonic-Motivation System Adoption Model (HMSAM)." *Journal of the Association for Information Systems*, *14*(11), 617–671.

Lukyanenko, R., J. Parsons and Y. F. Wiersma. (2014). "The IQ of the Crowd: Understanding and Improving Information Quality in Structured User-Generated Content." *Information Systems Research*, *25*(4), 669–689.

Lyytinen, K. (1985). "Implications of Theories of Language for Information Systems." *Management Information Systems Quarterly*, *9*(1).

Ma, M. and R. Agarwal. (2007). "Through a Glass Darkly: Information Technology Design, Identity Verification, and Knowledge Contribution in Online Communities." *Information Systems Research*, *18*(1), 42–67.

Madnick, S. E., R. Y. Wang, Y. W. Lee and H. Zhu. (2009). "Overview and Framework for Data and Information Quality Research." *ACM Journal of Data and Information Quality*, *1*(1), 2:1–2:22.

Mayring, P. (2010). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (11th ed.). Beltz.

McKinney, E. H. and C. J. Yoos. (2010). "Information About Information: A Taxonomy of Views." *MIS Quarterly*, *34*(2), 329–344.

Mingers, J. C. (1995). "Information and meaning: foundations for an intersubjective account." *Information Systems Journal*, *5*(4), 285–306.

Mingers, J. C. (1996). "An evaluation of theories of information with regard to the semantic and pragmatic aspects of information systems." *Systems Practice*, *9*(3), 187–209.

Mingers, J. C. (2013). "Prefiguring Floridi's Theory of Semantic Information." *Communication, Capitalism & Critique*, *11*(2).

Mingers, J. C. and C. Standing. (2018). "What is information? Toward a theory of information as objective and veridical." *Journal of Information Technology*, *33*(2), 85–104.

Mingers, J. C. and L. Willcocks. (2014). "An integrative semiotic framework for information systems: The social, personal and material worlds." *Information and Organization*, *24*(1), 48–70.

Mumford, E. (1995). *Effective systems design and requirements analysis: the ETHICS approach*. Basingstoke: Macmillan.

Mumford, E. (2003). *Redesigning human systems*. Hershey, Pa: IRM Press.

Nov, O. (2007). "What Motivates Wikipedians?" *Communications of the ACM*, *50*(11), 60–64.

Orr, K. (1998). "Data Quality and Systems Theory." *Communications of the ACM*, *41*(2), 66–71.

"Policies | YouTube." (2019). Retrieved from https://www.youtube.com/yt/about/policies

Ren, Y., F. M. Harper, S. Drenner, L. Terveen, S. Kiesler, J. Riedl and R. E. Kraut. (2012). "Building Member Attachment in Online Communities: Applying Theories of Group Identity and Interpersonal Bonds." *MIS Quarterly*, *36*(3), 841–864.

Rheingold, H. (1993). *The Virtual Community: Homesteading on the Electronic Frontier*. Reading, Mass: Addison-Wesley Pub. Co.

Ridings, C. M. and D. Gefen. (2004). "Virtual Community Attraction: Why People Hang Out Online." *Journal of Computer-Mediated Communication*, *10*(1).

Roberts, J. A., I.-H. Hann and S. A. Slaughter. (2006). "Understanding the Motivations, Participation, and Performance of Open Source Software Developers: A Longitudinal Study of the Apache Projects." *Management Science*, *52*(7), 984–999.

Rui, H. and A. Whinston. (2012). "Information or attention? An empirical study of user contribution on Twitter." *Information Systems & E-Business Management*, *10*(3), 309–324.

"Rules and policies | Twitter." (2019). Retrieved from https://help.twitter.com/en/rules-and-policies

Sadiq, S., N. K. Yeganeh and M. Indulska. (2011). "20 Years of Data Quality Research: Themes, Trends and Synergies." *Proceedings of the Twenty-Second Australasian Database Conference (ADC)*, 153–162.

Schlagwein, D., D. Schoder and K. Fischbach. (2011). "Social Information Systems: Review, Framework, and Research Agenda." In: *Research-in-Progress Proceedings of the 32nd International Conference on Information Systems (ICIS)*.

Schneider, J., A. Passant and S. Decker. (2012). "Deletion Discussions in Wikipedia: Decision Factors and Outcomes." In: *Proceedings of the 8th Annual International Symposium on Wikis and Open Collaboration* (pp. 17:1–17:10). New York, NY, USA: ACM.

Schoder, D., P. Gloor and P. T. Metaxas. (2013). "Social Media and Collective Intelligence—Ongoing and Future Research Streams." *KI - Künstliche Intelligenz*, *27*(1), 9–15.

Scott, M. and M. Isaac. (2016, September 9). "Facebook Restores Iconic Vietnam War Photo It Censored for Nudity." *The New York Times*.

Shah, S. K. (2006). "Motivation, Governance, and the Viability of Hybrid Forms in Open Source Software Development." *Management Science*, *52*(7), 1000–1014.

Sproull, L. and M. Arriaga. (2007). "Online Communities." In: H. Bidgoli (Ed.), *Handbook of Computer Networks* (pp. 898–914). John Wiley & Sons, Inc.

Stamper, R. (1996). "Signs, Information, Norms and Systems." In: B. Holmqvist, P. B. Andersen, H. Klein, & R. Posner (Eds.), *Signs of Work: Semiosis and Information Processing in Organisations* (pp. 349–397). Walter de Gruyter GmbH & Co KG.

Statista. (2018). *Social Media Usage Worldwide* (No. did-12393-1). Statista. Retrieved from https://www.statista.com/study/12393/social-networks-statista-dossier/

Stvilia, B., M. B. Twidale, L. Gasser and L. C. Smith. (2005). "Information Quality Discussions in Wikipedia." In: *Proceedings of the 2005 International Conference on Information Quality*.

Sumi, R., T. Yasseri, A. Rung, A. Kornai and J. Kertesz. (2011). "Edit Wars in Wikipedia." In: *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom)* (pp. 724–727).

"The importance of context - YouTube Help." (n.d.). Retrieved from https://support.google.com/youtube/answer/6345162?hl=en&ref_topic=2803176

Tilly, R., O. Posegga, K. Fischbach and D. Schoder. (2017). "Towards a Conceptualization of Data and Information Quality in Social Information Systems." *Business & Information Systems Engineering*, *59*(1), 3–21.

Toubia, O. and A. T. Stephen. (2013). "Intrinsic vs. Image-Related Utility in Social Media: Why Do People Contribute Content to Twitter?" *Marketing Science*, *32*(3), 368–392.

Trist, E. (1981). "The Evolution of Socio-technical Systems." *Occasional Paper*, *2*.

van der Heijden, H. (2004). "User Acceptance of Hedonic Information Systems." *MIS Quarterly*, *28*(4), 695–704.

Wand, Y. and R. Y. Wang. (1996). "Anchoring Data Quality Dimensions in Ontological Foundations." *Communications of the ACM*, *39*(11), 86–95.

Wang, R. Y. (1998). "A Product Perspective on Total Data Quality Management." *Communications of the ACM*, *41*(2), 58–65.

Wang, R. Y., V. C. Storey and C. P. Firth. (1995). "A framework for analysis of data quality research." *IEEE Transactions on Knowledge and Data Engineering*, *7*(4), 623–640.

Wang, R. Y. and D. M. Strong. (1996). "Beyond Accuracy: What Data Quality Means to Data Consumers." *Journal of Management Information Systems*, *12*(4), 5–33.

Wasko, M. M. and S. Faraj. (2000). ""It Is What One Does": Why People Participate and Help Others in Electronic Communities of Practice." *The Journal of Strategic Information Systems*, *9*(2–3), 155–173.

Wasko, M. M. and S. Faraj. (2005). "Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice." *MIS Quarterly*, *29*(1), 35–57.

"Website Traffic, Statistics and Analytics - Alexa." (n.d.). Retrieved from https://www.alexa.com/siteinfo

"Wikipedia Edit warring." (2019, March 6). In: *Wikipedia*.

"Wikipedia Policies and guidelines." (2019, April 22). Retrieved from https://en.wikipedia.org/w/index.php?title=Wikipedia:Policies_and_guidelines&oldid=893669639

Xiao, Y., L. Y. Y. Lu, J. S. Liu and Z. Zhou. (2014). "Knowledge diffusion path analysis of data quality literature: A main path analysis." *Journal of Informetrics*, *8*(3), 594–605.

Xu, Y. (Calvin), Y. Yang, Z. Cheng and J. Lim. (2014). "Retaining and attracting users in social networking services: An empirical investigation of cyber migration." *The Journal of Strategic Information Systems*, *23*(3), 239–253.

Appendix A: Survey Questionnaire and Descriptive Statistics of Data

#	Construct / Item	Continuous Variable: Scale | M | SD Categorical Variable: Levels | N | %		
1-10	Agreement with community guidelines *(constructed as average of items 1-10; items being, e.g., e.g., 'Articles should be written from an impersonal perspective'; Cronbach's alpha: .73)*	7-point scale, 'disagree' (1) to 'agree' (7)	5.524	.984
11	I agree with the community guidelines...	not at all	8	3.5
		for the most part not	3	1.3
		predominantly not so	2	0.9
		partly	34	15.0
		predominantly so	56	24.7
		for the most part	100	44.1
		completely	24	10.6
12-15	Satisfaction with community guidelines *(constructed as average of items 12-15; items 'Guidelines for the preparation of content are still missing' (reverse coded), 'The community guidelines are mature', 'The community guidelines are applicable', and 'The guidelines for content creation are sufficient'; Cronbach's alpha: .79)*	7-point scale, 'disagree' (1) to 'agree' (7)	4.747	1.258
16	The quality of the content I create is important to me	7-point scale, 'disagree' (1) to 'agree' (7)	6.075	1.432
17	The quality of the content I consume is important to me	7-point scale, 'disagree' (1) to 'agree' (7)	5.974	1.320
18	Frequency of internet use	rarer	0	.0
		several times per month	2	.9
		at least once a week	2	.9
		several times per week	8	3.5
		daily	215	94.7
19	Frequency of Social Media use	rarer	31	13.7
		several times per month	6	2.6
		at least once a week	8	3.5
		several times per week	38	16.7
		daily	144	63.4
20	Frequency of {platform} use	rarer	6	2.6
		several times per month	4	1.8
		at least once a week	7	3.1
		several times per week	58	25.6

		daily	152	67.0
21	I create new content / edit content myself	rarer	56	24.7
		several times per month	30	13.2
		at least once a week	30	13.2
		several times per week	52	22.9
		daily	59	26.0
22	I regularly consume content	rarer	8	3.5
		several times per month	17	7.5
		at least once a week	10	4.4
		several times per week	63	27.8
		daily	129	56.8
23	I know the community guidelines for creating content	7-point scale, 'not at all' (1) to 'very well' (7)	4.925	1.807
24	I am consciously implementing the community guidelines for creating content	not at all	24	10.6
		rarely	10	4.4
		sometimes	7	3.1
		approximately	29	12.8
		predominantly	76	33.5
		as accurate as possible	81	35.7
25	Highest degree (yet)	Bachelor or equivalent	76	33.5
		Completed vocational training	35	15.4
		Doctoral or equivalent	8	3.5
		Lower secondary education	5	2.2
		Master or equivalent	41	18.1
		Other	1	0.4
		Still in school education	12	5.3
		University entrance qualification	31	13.7
		Upper secondary education	18	7.9
26	Age	under 18 years	16	7.0
		18-29 years	96	42.3
		30-50 years	48	21.1
		over 50 years	67	29.5
27	Gender	female	63	27.8
		male	162	71.4
		other	2	.9
28	Origin (Country)	free text input	*n/a*	*n/a*

Table 3. Descriptive Statistics of Survey Data