

Increasing Information Auditability for Social Network Users

Alexandre Pinheiro¹, Claudia Cappelli¹, and Cristiano Maciel²

¹ Universidade Federal do Estado do Rio de Janeiro - UNIRIO, Rio de Janeiro, Brazil
{alexandre.pineiro, claudia.cappelli}@uniriotec.br

² Universidade Federal de Mato Grosso - UFMT, Mato Grosso, Brazil
cmaciel@ufmt.br

Abstract. This paper sets out the challenge of how to provide information auditability to citizens regarding social networks. The aim is to discuss the issues concerning information published on social networks and specifically to describe how to verify the veracity of this information. It is based on the idea that it urgent improvement of interface requirements is necessary for this kind of software to provide users with ways to account for, validate, verify and control information. The paper reports the construction of a catalog of characteristics for information auditability in social networks. An exploratory study suggests mechanisms that can be used to implement these characteristics into this kind of software interface. The conclusion presents remarks on challenges and future studies.

Keywords: Auditability, Social Networks, Transparency, and Information Auditability.

1 Introduction

The concept of transparency was strengthened and began to be discussed internationally following initiatives in the public and private sectors. These sectors encourage the provision of information and seek for the participation of society in issues of common interest. To meet the demand for transparency, not only in the governmental area, but also in many areas where transparency is applicable, it is necessary for information systems that connect information senders and receivers to also apply transparency. Therefore, usability, accessibility and credibility of information have become some of the most essential requirements for information systems that aim to support transparency.

The participation of end users in content production on the Internet has impacted the relationship with information because it is no longer only the newspapers, television channels, radios and popular websites that provide mainstream information. . . The way we/ interact with information has changed. The content generated by end users is not subjected to the same rules of traditional media. The collaborative and dynamic environment, full of interactions, does not allow us to evaluate information conventionally. The various users who interact with this environment can interpret

information available in social networks differently. This information can be modified, lose context and change meaning during the interaction process. What happens when a curious user finds difficulties to use a social network interface and disseminates misinformation? What are the consequences of spreading this misinformation? As an example, we can mention Facebook®, which offers numerous features encouraging the sharing of information [1]. This range of features combined with people's excitement in various social network activities and usability problems can contribute to the spreading of information the content of which has not been evaluated by the user. Offering easy-to-use tools and clean interfaces is a requirement for a good user-system interaction design [2] and it allows the evaluation of information (auditability). Therefore, it contributes to the evolution of the social network and ensures the spreading of correct information.

Auditability is the ability to conduct a critical evaluation of information provided by users. To achieve this goal, information auditability in social network catalog is herein structured according to the settings of the NFR Framework [3]. The construction of this catalog is based on the Transparency Catalog developed by Cappelli [4]. The catalog for information auditability in social networks is composed of characteristics that contribute to achieving the concept of auditability in social networks. The operationalizations and mechanisms arise from the HCI area. They have been defined based on a literature review related to user experience in the use and interaction with web interfaces [2] [5], in the construction of quality systems [6] and in recommendations for best practices defined by international standards [7] [8] [9].

The purpose of this paper is to present this complete catalog to support the evaluation of information auditability in social networks. We hope to help users to distinguish reliable information from information without references, validity or source. Other social networks interested in offering their users the possibility of accessing and sharing reliable information can adopt the suggestions on auditability presented in this work.

2 The Challenge of Information Auditability in Social Networks

It is hard to ensure that conventional approaches to the concept of quality of information fit into a context characterized by interactivity, collaboration and constant symbolic exchanges mediated by the architecture of social networks in which the user starts to behave as an actor that not only uses, but also produces, remakes and qualifies the informational content [10]. When a user wants to upload a photo, the elements displayed by the system must be prepared for the action to be completed successfully. To publish a photo, some elements are required like a field to upload the image, the possibility of adding a heading, the possibility of identifying who is in the photo and other features. All those features are very common in social network software, but mechanisms to guarantee accountability, validity, verifiability, controllability and traceability are very rare. Auditing information in a social network seems to be a difficult job. Some examples, such as the paradox of information relevance from each

user's point of view, the usage of colloquial language, information overload and user anxiety considering the dynamism of information [11] have been pointed out.

Social networks group a set of autonomous participants, linking ideas and resources around shared values and interests. Research on social networks evaluates behaviors or opinions of individuals according to the structures in which they are inserted. In these studies, the units of analysis are not the individual attributes like class, age, gender, but the set of relations that individuals establish through their interactions [12]. In social networks, users share the same space and are seeking to interact, share opinions, preferences and interests in various subjects. The diversity and large audience as well as the low segmentation make social networks spaces where each can publish without commitment to the truth. According to Lee [13], unverified and unclear information is daily broadcasted on the Internet. In his research he mentioned, as an example, a hoax that circulated during 2012 about the fact that former U.S. President Abraham Lincoln had registered a patent for Facebook in 1845. With an image of a supposed old newspaper and some elaborated history, in less than 2 days the link posted on Facebook containing the false news had already 16,000 Facebook "likes" and a lot of people believing that the prank was a real historical fact.

2.1 Related Work

Some related studies were visited. Most of them recognize the lack of information credibility in social networks as a big problem. Budak et al. [14] recognize the problem of the spread of misinformation in social networks and studied a way to decrease the number of people who believe in incorrect information. The research describes the use of prediction algorithms to verify the network nodes that can be affected by inaccurate information and then disseminate appropriate information to minimize the effects of misinformation for users connected by these nodes. His point is to deliver quality information before the user accesses some content without credibility, preventing the adoption and spreading of "bad information".

Vedder [15] analyzed misinformation through the Internet with an ethical and epistemological approach, highlighting the consequences of the use of unreliable information found on the web. An example used to show the misinformation problem is the adoption of pseudoscientific information by Internet users to determine a health care treatment. The decisions based on this kind of information can have consequences for the user's life. The survey also reports the use of inaccurate information for educational purposes that may harm the already established truths. The incorrect information may be dangerous to the receiver and is more dangerous when it reaches other people due the spreading capacity of the Internet. The suggested solutions to the problem of misinformation through the Internet are presented as strategies. The first strategy is based on the development of the user's ability to evaluate information. Training users to identify elements that can help to define the information's degree of credibility can implement this strategy. Another strategy mentioned is related to some factors such as disclosing the source of the published information and the existence of an institution with a recognized reputation which ratifies the information. The use of

credibility-conferring systems can help to enforce these strategies so they can have a global reach.

Flanagin and Metzger [16] showed that people increasingly rely on Internet and web-based information although it is potentially inaccurate and biased. They asked 1,041 people whether they verified Internet information before using it. Overall, respondents reported they considered Internet information as credible as that obtained from television, radio, and magazines. Credibility among the types of information sought, such as news and entertainment, varied across media channels. Respondents said they rarely verified web-based information, which can be a great problem because most of these respondents all received information as being true.

In all these studies we can see the concern and need for information auditability circulating on the Internet mostly on social networks. Some have only cited the problem, others presented solutions but none of them defined mechanisms to effectively implement auditability in social networks. This paper seeks to systematize the implementation of these mechanisms.

3 A Catalog for Information Auditability in Social Networks

With the increase of transparency, social networks can be more auditable. Cappelli [4] defined the concept of transparency. This concept was modeled as a catalog with five degrees. All of them composed of a set of characteristics. The union of those characteristics contributes to transparency and represents the Transparency Catalog. The five degrees are: i) Accessibility (access to the information), ii) Usability (use of the information provided), iii) Informativeness (providing information with quality), iv) Understandability (understanding of the processes and information provided), and v) Auditability (ability to conduct a critical examination of the information provided). Each of these five sets of characteristics includes other characteristics, totaling 33 characteristics in all. Auditability is one of these degrees and is composed of accountability, controllability, verifiability, validity, and traceability.

In Software Engineering, catalogs have been used to store features of quality and elements for defining non-functional requirements. Catalogs represent a systematic way to decompose non-functional requirements and are a method to prioritize, operationalize and treat interdependencies between it [3]. Catalogs also gather operationalization, which is a set of practices to be followed in order to put certain characteristics in the desired context. Operationalizations become reality through the implementation of mechanisms.

The Transparency Catalog was created according to the SIG (Softgoal Interdependency Graph) notation. The SIG is a hierarchical structure, where characteristics and types are represented by clouds (softgoals). Softgoals are connected by links of interdependence, where a child softgoal contributes to the achievement of the parent softgoals. A SIG can also show the types of contribution among softgoals. These types of assistance may be BREAK (a negative contribution sufficient to prevent the top characteristic from being met), HURT (a partially negative contribution where the

top characteristic is not met), UNKNOWN (it's unclear whether the contribution is negative or positive), HELP (partial positive contribution to meeting the top characteristic) and MAKE (a positive contribution sufficient for the top feature to be met).

3.1 Information Auditability in Social Networks Catalog construction

As previously stated, a catalog is composed of characteristics, operationalizations and implementation mechanisms. Our approach to building the catalog of information auditability in social networks is based on Transparency and HCI areas. First, to identify the characteristics each of the 33 characteristics of the transparency catalog [4] was analyzed.

In the transparency catalog, the characteristic of *accountability* is defined as the ability to inform the reason for something. The importance of accountability to information auditability in social networks is related to the necessity of providing information about its operation, operation of their tools, guidance on how users should proceed in a situation (troubleshooting) and presentation of questions and answers. The environment of social networks should allow users to guide other users who have doubts about interaction and use of social networking as a whole. The characteristic of *controllability* is defined as the ability to control. We chose this characteristic due to its relationship with user control, what users are doing and the information they are accessing. We can make important inferences from analyzing actions, social presence, reputation and influencing the skills of a user. The *traceability* characteristic is defined as the ability to follow the development of an action or the construction of information. Traceability considers the changes and the justifications of the information transformation and it directly helps with the tasks related to audits once information is extremely changeable in a collaborative environment such as social networks. Tracking a page or a user that started spreading of a piece of information and analyzing the sources of this information are necessary resources to ensure auditable activities. Traceability can be an effective method for gathering data from users and can be used to support analysis of how users are using the system and what they are achieving [5]. Besides helping in transparency, the tracking of user information allows the analysis of their goals and actions. This kind of evaluation by users is a characteristic that web applications should implement to follow human-computer interaction international standards [7] [8] [9]. Another characteristic that contributes to auditability is *verifiability*. This characteristic is related to the necessary ability to legitimize something. The social network should allow the user to check whether information is legitimate. Some social networks such as Facebook have verified profiles that indicate that a user or a page is exactly what it is said to be. This kind of feature supports the legitimacy of identity. The characteristic of *validity* was also used in our approach because it is related to the ability of testing something by experiment or observation to determine if what is being done is correct. In the analysis of this characteristic, social networks should allow tests to validate the veracity of information. The result of these tests helps to ensure auditability.

In addition to the characteristics required for auditability already specified by the transparency catalog, there are other characteristics of this catalog that also contribute positively towards our goal of increasing information auditing in social networks. Among these characteristics there is adaptability. *Adaptability* is the capacity to change according to circumstances or needs. During the analysis we need to reorganize information in a different way from which it comes to us. The information will be adapted to satisfactorily suit different types of social network users. The ability to use something when it is needed is the definition of *availability*. To analyze information it is necessary to have access to it as well as the implementation mechanisms to validate it. In the case of social networks, availability is also a mainstay of the system itself, otherwise the lack of information would not be an incentive to attract users to use such a network. *Clarity* is an essential characteristic to check information. Some malicious pieces of information spread on social networks take advantage of the lack of clarity to trick users oblivious to the informational content but eager for information sharing. The dynamism of the interaction environment on social networks is another factor that makes the spreading of unclear information easy. The characteristic of *completeness* is related to the ability not to miss elements that the information could or should have. Comprehensive information allows better analysis by the users. The lack of correctness of information strengthens the need for tools that support the concept of auditability. The characteristic of *correctness* is related to the ability to provide information free of errors. *Composability* is the ability to construct information with different partitions. To audit information it is necessary to make inferences between existing data from the information and other gathered data. The characteristic of *dependability* is related to the unfolding of information. To evaluate information we need to know about the relationships of each part to the whole. When we share information, it should provide knowledge about all other parties involved in the process. Information with *extensibility* contributes substantially to auditability. In social networks information is often published without detail, which complicates its verification and validation. The characteristic of *decomposability* is related to the capability of partitioning something. During an audit we may need to split the information into smaller pieces and by doing this we can carefully evaluate each piece of information. The characteristic of *uniformity* is defined as the capacity to maintain a unique form. When information is uniformly provided it makes verification tasks easy and allows the user to make a relationship between how the information is presented and its content. The SIG of Information Auditability in Social Networks is represented in Figure 1.

After identifying those characteristics, we applied the foundations and practices of HCI area to define those operationalization and implementation mechanisms. The operationalization and implementation mechanisms have been defined based on a literature review related to user experience in the use and interaction with web interfaces [2] [5], in the construction of quality systems [6] and in recommendations for best practices defined by international standards [7] [8] [9].

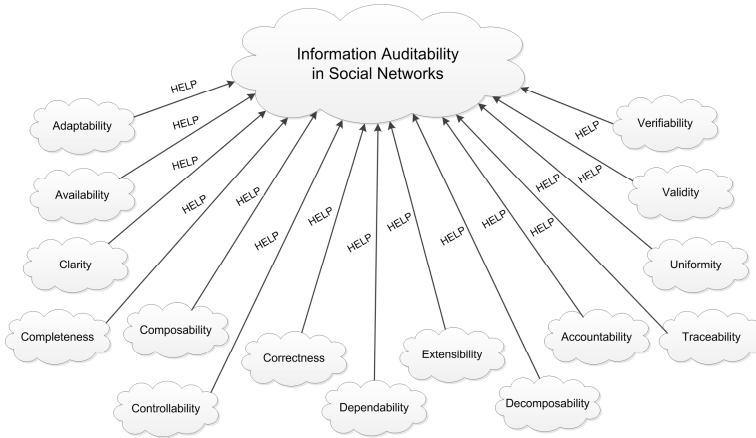


Fig. 1. Information Auditability in Social Networks Characteristics

For each characteristic identified to compose the information auditability in the social networks catalog, we identified elements to operationalize them and mechanisms to implement them in social networks. Table 1 shows some operationalization and implementation mechanisms defined for traceability.

An activity that contributes to misinformation in social networks is link sharing (URLs) together with unreliable stories, usually brought from outside the social network. These kinds of posts are spams, hoaxes and biased news that trick many users.

Table 1. Operationalization and implementation mechanisms of traceability

<i>Traceability</i>	
Operationalizations	Implementation Mechanisms
1) Identification of the origin, availability and security of the published URLs	<ul style="list-style-type: none"> • Check availability of the URL • Show the original URL • Emphasize Secure URLs (HTTPS)
2) Show Metadata from URL information	<ul style="list-style-type: none"> • Allow the URL information preview • Make hashtags from keywords defined in the metadata of the URL • Check page rank of the URL
3) Preprocessing the URL	<ul style="list-style-type: none"> • Apply web crawlers to analyze the construction and content of the website • Check if the website was developed according to web standards by reading its code.
4) Identification of nodes in the social network connected to the source of information	<ul style="list-style-type: none"> • Use reputation systems to validate user or page interaction with information.

An action to curb this type of activity begins with the availability of complete information about the spreaded URL, tracing its origin and allowing the user to make a comparative analysis with the post that contains the suspicious URL. In social networks, URLs must be accessible and the system needs to ensure this, maintaining its content up to date to enable the traceability of information and knowing the URL source. One way to do this is automating scripts such as web spiders, which test the links available in the environment. Another way can be through collaborative features using feedback from users on broken links or a links that represent a threat (e.g. a suspicious download). It is important that the user be able to refine the system by giving his opinion about the operation.

A common practice in social networks is the use of redirected or shortened links that do not allow viewing of the true URL. This makes it difficult to trace the origin of information and to analyze the URLs using this type of practice. A possible implementation is to show the original URL. Showing the full link also helps in users' direct memory retrieval. When the user knows the URL, he can make a mental association between the content of the website, where the URL leads to and the information containing or surrounded by the URL [5]. Another action that can be done is highlighting secure URLs to increase user safety because the link is already validated and authenticated. The choice of browsing through secure URLs should be displayed whenever a website presents this possibility.

By reading metadata from URLs published or recommended in social networks, the system can show information about the website before users access the link. Users will have the opportunity to evaluate the type of information they will encounter when accessing the link. When a website is built according to web standards there is concern on the part of the developer with the usability and the end users. Unreliable websites with dubious intentions and supporting spam, viruses, false information and all sort of unwanted content are normally developed without following web standards and do not show quality features. Another way to use the information found in metadata of websites is related to the use of the meta element called keywords. Meta elements are machine parsable and one of them is involves using keywords that define the content of the website. The characteristic of explicit indexing of keywords metadata can be associated with features of tagging. Tags represent terms related to information and thier use can help users to link information from different sources with the same content. With a variety of information related to the same content, the users can analyze how the information was constructed and check for characteristics that add value and credibility to it. We can also use indexing systems like Google's Page Rank that evaluate URLs according to the quantity and quality of links pointing to it. A website with a large number of links pointed at it has a positive contribution when we analyze its reliability. Although the analysis of metadata content helps with information evaluation, a website can still be manipulated to trick users and support misinformation. To make a deep evaluation of the content of a website we can submit the URLs posted on social networks to a crawler. Crawlers can visit webpages and trace the links on those pages. They will analyze each link traced and collect data about those websites. With the information gathered by the crawler, we can do a comparative analysis and check the information spread with the content of the website.

To improve the quality and to check credibility of information, social networks have a native characteristic, which is their collaborative capability. The collaboration capability allows network members to evaluate information. In the case of a positive consensus on the content, source and other aspects, individuals will have a reliability gain in the information trustability. To implement this feature as traceability operationalization some behaviors must be observed. Individuals can transform information in social networks and even original content can be turned into misinformation. Reputation systems can help to solve problems that happen when a large number of users interact with information, facilitating content creation and consumption. An example of reputation system that can be used in social networks relies on a chronological analysis of user contributions to information, increasing or decreasing reputation whenever new contributions are made [17].

3.2 Prototyping Implementation Mechanisms

This section shows some prototypes to demonstrate how the ideas proposed in this paper can be implemented. Imagine you access your news feed in a social network and someone places a piece of news. In general this post will contain a text and a link to the source of the news. Following the approach presented in this paper, by means of the mechanisms suggested in Table 1, the piece of news could have the format presented in Figure 2, where the numbers respectively represent: 1) The title of the URL retrieved from metadata of the website; 2) the title of the URL retrieved from metadata of the website; 3) A dialog box for user feedback about the link or any factor that could indicate problems with the information, such as broken or suspicious links; 4) Page Rank according to indexing and linking system; 5) Friends of the user who are interacting with the information and their reputation; 6) Tags automatically generated from the reading of the metadata keywords.

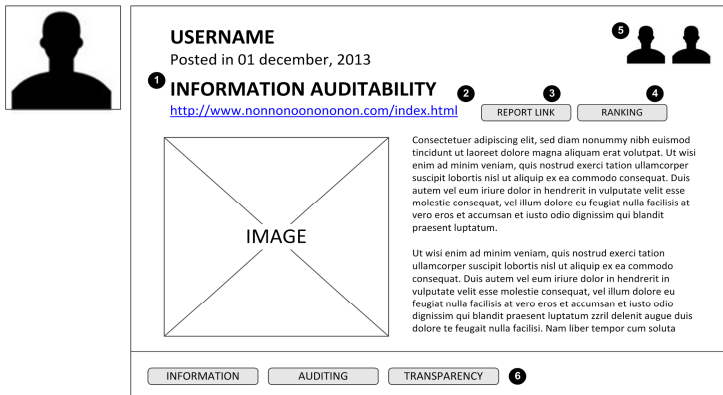


Fig. 2. An example of traceability mechanisms on a social network post

In another situation, imagine you are accessing a tag (Figure 2) about some content, for example auditability. When selecting a tag you can see your friends' list and all those who interacted with associated information based on social search. Your level of relationship and knowledge about their friends will help in the exchange and evaluation of information of some content. The amount of interactions with related information would be shown above your friend's picture as presented in Figure 3.

Misinformation and lack of security in navigation in social networks are normally associated with unknown, misunderstood or suspicious external links spread in the system. These kinds of URLs are accessed by users who don't know what the content of those links is. Figure 4 presents a feature that provides a preview based on pre-processing contents of external links that will help the users to evaluate if the website has the information it claims to have and if it is safe. Anxious users that access links and information without any concern are preferred targets of actions that reinforce the spreading of misinformation.

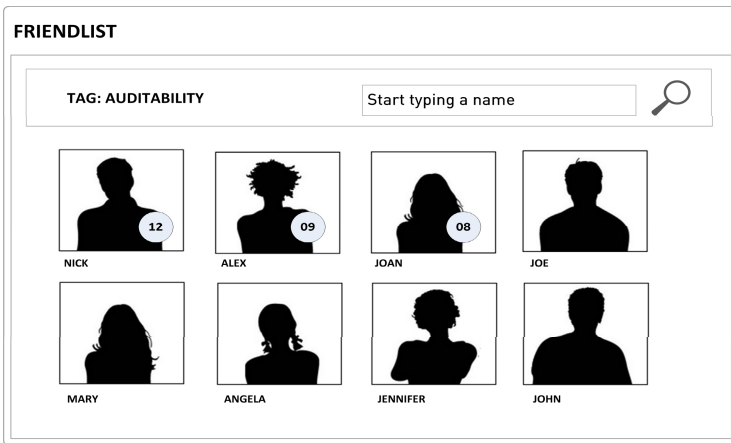


Fig. 3. An example of relation between your friends and information interaction

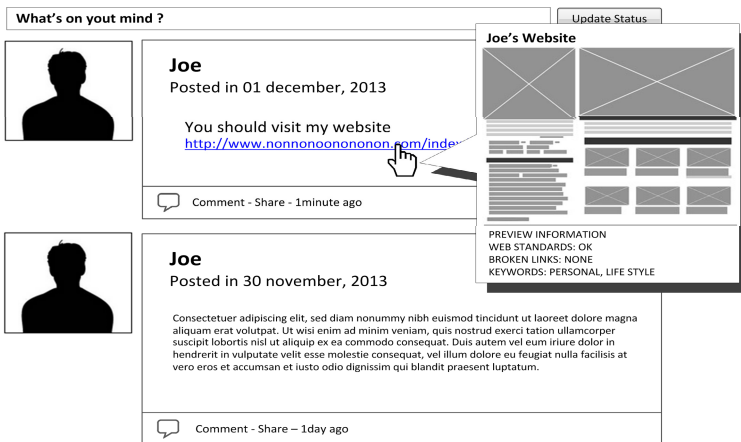


Fig. 4. Example of preview feature with information about the website

4 Conclusion

In this paper, we describe and demonstrate the construction of a catalog of characteristics for information auditability in social networks. The need for information auditability follows the current demands on transparency, especially in social networks, where everyone can post information.

The concept formalization seems to be an important step to achieve information auditability in social network. To formalize the information auditability in a social network concept, a set of characteristics, operationalization and implementation mechanisms were organized into a catalog from which design decisions could be taken.

The interface prototypes presented in this paper show that the catalog may lead to interfaces designed with more features that allow audit information. They also illustrate that the use of cited mechanisms is possible. The resulting interface has characteristics that should facilitate auditability by social network users.

The next challenges and future work for providing information auditability in social networks could be envisioned. In this paper, we presented a catalog identifying characteristics, operationalizations and implementation mechanisms to improve information auditability in social networks, but a guide is necessary to define how to apply this to interface and software construction. It is also necessary to discuss the target audience. Future research could encompass procedures and tools to design interfaces with auditable characteristics. However, the prototype interface must first undergo a usability inspection for which we initially intend to use the heuristic evaluation [18] method.

Research directions on how to leverage information auditability in social networks could comprise different ways. One of the most difficult aspects in information auditing is being sure about sources. The complexity and cost to have information about information provenience is a challenge. Another challenge is how to define some interface standards to facilitate the implementation of the same mechanisms making software auditable.

References

1. Hart, J., Ridley, C., Taher, F., et al.: Exploring the Facebook Experience: A New Approach to Usability. In: 5th Nordic Conference on Human-Computer Interaction: Building Bridges, Lund, Sweden, pp. 471–474 (2008)
2. Sharp, H., Rogers, Y., Preece, J.: Interaction Design: Beyond Human-Computer Interaction. John Wiley & Sons, England (2007)
3. Chung, L., Nixon, B., Yu, E., et al.: Non-functional Requirements. Software Engineering (2000)
4. Cappelli, C.: An Approach for Business Processes Transparency Using Aspects. Dissertation, Pontificia Universidade Católica do Rio de Janeiro (2009)
5. Blandford, A., Attfield, S.: Interacting with information. Synthesis Lectures on Human-Centered Informatics 3, 1–99 (2010)
6. Software Engineering. Product Quality, Part 1: Quality Model. NBR ISO/IEC 9126-1:2003, Brazil

7. Ergonomics of Human-system Interaction, Part 151: Guidance on World Wide Web User Interfaces. ISO 9241-151:2008(E), Switzerland
8. Ergonomics of Human-system Interaction, Part 210: Human-centered Design for Interactive Systems. ISO 9241-210:2010(E), Switzerland
9. Software engineering. Software Product Quality Requirements and Evaluation (SQuaRE). Guide to SQuaRE. ISO/IEC 25000:2005(E), Switzerland
10. Assis, J., Moura, M.: A Qualidade da Informação na Web: Uma Abordagem Semiótica. *Revista Informação & Informação* 16(3), 96–117 (2011)
11. Agarwal, N., Yiliyasi, Y.: Information Quality Challenges in Social Media. In: 15th International Conference on Information Quality, Arkansas, USA (2010)
12. Marteleto, R.: Análise de Redes Sociais – Aplicação nos estudos de transferência da informação. Programa de Pós-Graduação em Ciência da Informação UFRJ/ECO (2001)
13. Lee, N.: Misinformation and Disinformation. *Facebook Nation*, pp. 103–113. Springer, New York (2013)
14. Budak, C., Agrawal, D., El Abbadi, A.: Limiting the Spread of Misinformation in Social Networks. In: 20th International Conference on World Wide Web, Hyderabad, India (2011)
15. Vedder, A.: Misinformation through the Internet: Epistemology and Ethics. *Ethics and the Internet*, 125–132 (2001)
16. Flanagin, A.J., Metzger, M.J.: Perceptions of Internet Information Credibility. *Journalism & Mass Communication*, 515–540 (Quarterly September 2000)
17. De Alfaro, L., Kulshreshtha, A., Pye, I., et al.: Reputation Systems for Open Collaboration. *Communications of the ACM*, 81–87 (2011)
18. Nielsen, J.: 10 Usability Heuristics for User Interface Design (1995), <http://www.nngroup.com/articles/ten-usability-heuristics/> (accessed November 10, 2013)