

Shall I trust a recommendation?

(Towards an evaluation of the trustworthiness of recommender sites)

G. Lenzini, Y. van Houten, W. Huijsen, and M. Melenhorst

Novay, Brouwerijstraat 1
NL-7523 XC – Enschede, The Netherlands

Abstract. We present a preliminary study the aim of which is to provide a high level model for the evaluation of the trustworthiness of recommender systems as e-commerce services. We identify and comment on relevant trustworthiness indicators for the following different perspectives: user interface, content information and quality of recommendations.

Keywords: Trustworthiness, Recommender Systems, Review sites.

Introduction

The way in which consumers shop online is changing due to the rise of customer reviews, recommender systems, and the increasing influence of social communities on shopping behaviour. Consumers book their holidays on Tripadvisor (www.tripadvisor.com), on Amazon (www.amazon.com) they receive recommendations for products they might also want to buy, and people share their tastes, experiences and suggestions on Kaboodle (www.kaboodle.com). These examples show that more and more online shopping activities are becoming a social process, just as it often happens in a regular shopping mall. Moreover, there is a growing need for information re-education and classification, as the abundance of reviews, suggestions and products is overwhelming to consumers. Hence, recommender systems are introduced to help users cope with the options they can choose from.

As shopping in the real world involves trusting advices and recommendations from other people, trust is also an important factor that must be considering when designing an online shopping service. In this paper we address trust factors within e-commerce recommender sites.

Background: Trust and Recommender Systems in an E-Commerce Context

Despite being welcomed with some reserve years ago, the concept of trust in digital information is nowadays widely accepted. Only few years ago trust was considered possible only among people, but successive studies in human-computer interaction have indicated that people can have relations (including the social relation of trust) also with computer technology. According to [18] “*individuals can be induced to behave as if computers warranted human treatment, even though users know that the machines do not actually warrant this treatment*”. In other words, modern computer technology has the ability to influence human perception so that computers can be perceived *as if* they were human partners [3]. Therefore, comprehensive models of trust, originally designed to be applied in social science, have been adapted to fit digital information. Trust models can now describe trust in digital information as well as trust in information systems, in e-commerce, and in on-line relationships [10]. It is therefore legitimate to talk about trust in pieces of information, in on-line business services, and in virtual identities on the web.

What is the role of trust in this case? Trust is what influences the relationship between quality (e.g., of a piece of information, or of a service) and usage (e.g., of that piece of information, or of that service). In other words, trust is the intervening variable that directly affects the use of a certain technology [17]. This statement is actually even truer for web sites, which today offer their users a wide range of features, some of these explicitly pushing social aspects.

In this paper, we focus on recommender sites within the e-commerce domain. Modern recommender sites act both as review sites that gather and offer customer reviews, and as recommender systems, by presenting (or suggesting) items that are likely to be of interest to the user. In the e-commerce domain social aspects become more and more important, as it is witnessed by the growing importance of

customer reviews, and by the rise of shopping communities clustered around life-styles, products, or mutual interests.

These types of sites embrace many of the aforementioned “objects” of trust, such as digital information (they provide users with ratings, written reviews, and any kind of relevant information), on-line service (they offer searching among a set of products, they compose ad-hoc recommendations, they visualize and array items according to users’ interest), e-commerce (they allow users to buy items and to rate the service) and on-line entities (they keep identities of users, of reviewers, of sellers, and of goods and use these records to build a so called “Web of Trust”).

The aim of our research is to compose a criterion for the evaluation of the trustworthiness of e-commerce and recommender sites. Part of this evaluation can be done automatically, yielding trustability indicators that may support a higher-level trust evaluation. In its simpler version, this criterion is a checklist that arranges the relevant aspects that a recommender site has to possess to be considered trustworthy. In its most advanced version, it can lead to the development of a tool that automatically evaluates a web site and gives back the site’s trustworthiness rank. We underline that our approach to trust does not target the improvement of the accuracy of reviews and recommendations as it happens, for instance, in [19] where the trustworthiness of partners is proposed as a mean to improve the accuracy of the recommendation’s predictions. We are not interested in this computational approach to trust. Instead, we look at a recommender site as an e-commerce service, and when we talk about the trustworthiness of a recommender site we mean the perception of trust that a user has when using that service. We aim to improve the understanding of the aspects that influence the perception of trust that a user has when choosing a certain recommender site instead of another. From this point of view, the accuracy of the predictions that a recommender system is able to provide, is only one of the possible variables that can affect the overall trustworthiness of the service.

Trust Model

Among the different natures of trust (individual, interpersonal, relational, societal) we are interested in the interpersonal trust. “Interpersonal” means that trust is treated as a tie between two entities, the *trustor* and the *trustee*. This tie is defined in terms of an attitude that the trustor holds towards the trustee. Among the existing trust models we have selected that in Figure 1, due to Kelton *et al* [10]. It has been presented as a trust model for digital information and we consider it a candidate to describe trust in recommender sites. In the following paragraph we describe the main elements of this model

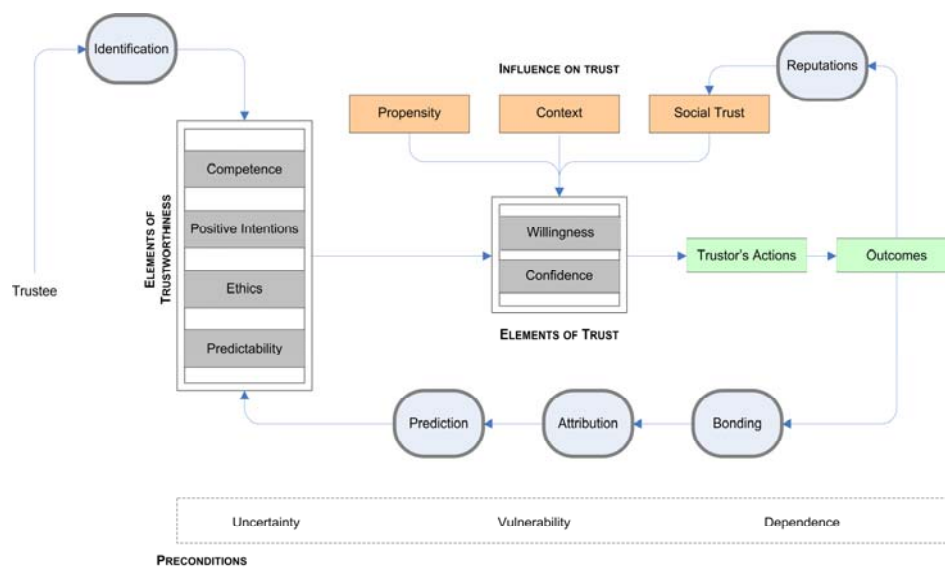


Figure 1 Trust model (adapted from [10]). At the bottom of the picture there are the preconditions of trust; in double squares, the elements of trustworthiness and the elements of trust; in the squares at the top of the figure, there are the factors that influence trust; in round boxes, the processes of trust.

Preconditions of Trust. It is recognized that trust can arise only in case of uncertainty, vulnerability, and dependence. The need of trust exists because the trustor is vulnerable to the trustee’s actions on which the success of the trustor’s goals depend upon. Moreover, the trustor is unable to predict the trustee’s behavior with certainty.

Elements of Trust. Central elements of trust are the *confidence* in positive outcomes and the *willingness* to modify one's behavior in expectation of these outcomes. Confidence expresses the perception that the likelihood of positive outcomes is high. Willingness is the freedom of accepting or rejecting the risk that, indeed, outcomes might not be as expected.

Elements of Trustworthiness. These are elements that influence the perceived likelihood that a trustee will uphold the trustor's trust. *Competence* implies that the trustee has the knowledge or the skill to carry on the activities on which the trustor depends upon. *Positive intentions* (also called benevolence, goodwill, motivations) are associated with the affective dimension of the trustee's trust; they are inferred on a cognitive basis. *Ethics* are the general intentions (directed not only towards the trustor), which include integrity, honesty, fairness, fiduciary obligation, and predisposition of putting other's interest first. *Predicability* is the degree to which the trustee's behavior conforms to the expectations; it is usually inferred from past behavior and from the role or the function of the trustee.

Processes of Trust. Trust develops itself through several processes. *Prediction* (called also knowledge-based trust or relational trust) is based on consistency of the trustee's past behavior. *Attribution* (called also dependability and calculus-based trust) is based on words or actions of the trustee or on other credible information about the trustee. *Bonding* refers to the development of emotional relations between trustor and trustee. *Reputation* emerges from the opinions that a community has about the trustee. *Identification* arises from the extent to which trustor and trustee share common identities, goals, and values.

Influence of Trust. Several factors influence the establishment and the development of trust. *Propensity* to trust is a purely psychological attribute that regulates a person's slant to trust or to distrust. *Context* is the situation in which the trustor, the trustee, and their relations are embedded. *Social trust* is an external factor that brings people to trust a recipient who is already trusted by others people which are part of well identified community.

The trust model herein described can be instantiated to obtain models of trust in particular scenarios. In the following section we instantiate the model to describe trust in our objective of research: the recommender sites.

Trust in Recommender Sites

This section specifies several elements of the aforementioned trust model in the context of recommendation sites. These elements provide a basis for the development of our criterion.

First we observe that the *preconditions of trust* are fulfilled in the specific context of recommender sites. Because there is no standard procedure that helps users evaluate the quality of information, users are forced to work in a situation of uncertainty. Moreover, users may suffer from potential harm when relying upon reviews and recommendations: A misinformation (i.e., faulty or incomplete information) may bring to a loss of time, to a loss of money, or both. Therefore users find themselves in a position of dependence when resorting to a recommender site and to the information it provides, and they have to balance the risk of a loss and the probability of a gain before taking a decision to trust and proceeding with a purchase. The relationship between risk and trust, out of scope in this paper, is far from been simple and it is matter of study (cf. [9]). Users that have confidence in the proper functioning and in the reliability of a recommender site have, nevertheless, the possibility of refusing a recommendation, of quitting the process of purchasing, and of switching to a different site. In some cases, even after having paid, users have the option to return the good back to the seller and to be refund of what paid. The *elements of trust* (i.e., confidence and willingness) are then well specified in the context of recommender sites.

In the following sections we provide an overview of the variables that, according to the literature, are related with the elements of trustworthiness of a recommendation service. Where possible, we also indicate how they support and influence the processes of establishing trust.

We start with an analysis of user interfaces, the users' gateways to services. In fact, elements of trustworthiness emerge, and are influential, already at this level: For instance, a badly designed interface may be perceived as a sign of lack of professionalism. Then we proceed with an analysis of the linguistic contents of a site, whose quality influences trust in manifold ways: In fact, a correct or an incorrect use of language supports or compromises, respectively, the credibility of the service. We also report on the tools that today's language and web-spidering technologies make available for an automatic trustworthy analysis of content. Finally, we move towards a qualitative analysis of recommendations and of reviews such as, e.g., the existence of tools for evaluating the reputation of reviewers themselves, or the possibility of having personalized recommendations that fit users' taste

and context. As far as possible, all the elements we are going to analyze are related to the model presented in Figure 1.

User Interface Analysis

Users' willingness to engage with a website is strongly dependent on the extent to which the site succeeds in conveying trustworthiness. Hertzum *et al.* [7] indicate that trust emphasizes that even seemingly objective factors are, actually, perceived factors. Trust does not reside in sources as a label that can be read, but has to be established actively by the individual person: it is ultimately the information seeker's responsibility to assess the source trustworthiness. Thus, websites – including recommender sites - must be accessible in ways that enable information seekers to form an opinion about their trustworthiness. The available elements in a user interface thus influence the subjective perception of trust.

Shneiderman [26] provides a number of guidelines for ensuring users that they are engaging in a trust relationship. First, a website should disclose patterns of past performance. For example, airlines should report on the percentages of flights that landed without delays. Information about the organization and its management, employees and history may also engage and assure skeptical users. Second, a website should provide references from past and current users. Opinions and experiences of other users provide valuable information on a product or a service. Third, a website should get certifications from third parties. For example, seals of approval from consumers and professional groups (e.g. displayed in the form of logos) help to establish trust. Of course, means for verifying the reliability of the information provided should also be present. Fourth, it should be easy to locate, read, and enforce policies involving privacy and security: Privacy policies that are difficult to find and that are incomprehensible or awkward to read only undermine trust.

Of course, users may question whether all this information eliciting trust is actually trustworthy information. What if comments of satisfied costumers were actually written from employees of that company? One way to improve the users' perception of trust is to allow users to compare information, for example by allowing users to have easy access to different sources of information and to check whether the information is consistent. The system the user interacts with should facilitate making these kinds of comparisons. In the case of recommendations, multiple recommendations should be available, and the source of the recommendations should be explicitly visible.

Tseng and Fogg[27]distinguish four types of trust (they called it credibility) by means of the evidence on which the trust is founded [27]. Namely:

- Experienced credibility, which refers to how much a person believes someone or something based on first-hand experience (e.g., interacting with people over time assesses their expertise and trustworthiness).
- Reputed credibility, which expresses how much a person believes someone on the basis of what third parties have reported (e.g., by asking someone for advices based on having her recommended by a colleague).
- Surface credibility, which refers to how much a user believes someone or something by simple inspection of surface attributes (e.g., by assessing people by the way they dress or the language they use).
- Presumed credibility, which refers to how much a user believes someone or something because of general assumptions in the perceiver's mind and stereotypes (e.g., believing that your friends tell the truth, whereas car salespeople do not).

When applied to recommender sites, first-hand experiences are related to a person's own experience with the trustworthiness of the system as a whole, or of a reviewer in particular. Reputations of reviewers can be managed by the systems, e.g., by granting top reviewers a visible top status and their characterizing attributes. Surface attributes refer to user interface elements, including the name and use of words by the reviewer. General assumptions and stereotypes concern expectations about the interests and the competence of a reviewer. In sum, the user interface elements described by Shneiderman [26]can not only be applied to the recommender system as a whole, but also to reviewers providing a recommendation. All these aspects are related to the elements of trustworthiness as presented in the model displayed in Figure 1: *competence*, positive *intention*, ethics, and *predictability*.

Automatic Trustability Analysis of Content

The trustability of an individual or of a piece of content (e.g., a written review) can be gauged from a number of indicators, such as the text's internal consistency, and the author's reputation. The indicators can be considered input to the trustor. This process can be supported programmatically by

analyzing the content of the trustee. In the following, we also discuss how language technology and web spidering can be used in performing trustability analysis.

Trustability Indicators

Trustability indicators are measures of aspects of the extent to which a text can be trusted. These indicators provide a rough sense of trustability: a text may score poor on an indicator and still be trustworthy. It is by an overview of a number of indicators that a general trustability may be determined. The following list gives some indicators that help to establish the trustability of texts:

- *Contact information*: the availability of contact information.
- *Grammar and spelling*: the extent to which a text adheres to grammar and spelling rules.
- *Duplication*: the extent to which (parts of) the text occur in other texts and locations.
- *Publication date*: the date of publication of the text.
- *Source Reputation*: the reputation of the source (author, author's organization, publisher).
- *Subjectivity*: the extent to which the text uses subjective language.

The following sections will discuss, for each of these indicators, how language and web spidering technologies can be applied to derive measures. Language technology is a wide area of research. Many of its techniques can be used for analyzing the trustability of a text from a linguistic point of view. Web spidering is the programmatic analysis of the content and of the links between websites. Web spidering is used whenever structured information is required that is published on web sites and when that information is not available in a structured format (because it does not exist in such a format, because no programmatic access to an underlying structured format is offered, or because even though programmatic access is available, it is unfeasible because of the wide variety of different structured formats).

Contact Information. The availability of contact information indicates openness and verifiability. Contact information can be available in the text itself, on the same webpage, or elsewhere on the site on which the text is published.

Grammar and Spelling. This indicator concerns the extent to which a text adheres to grammar and spelling rules. The rationale is that texts that adhere to grammar and spelling rules more closely are also more likely to have been written with greater care, and may therefore be more trustworthy.

Duplication. This indicator concerns the extent to which (parts of) the text occurs in other texts and locations. The rationale is that texts of which multiple copies exist have been copied, and must therefore be considered valuable and trustworthy. Automatic plagiarism detection is a research area within language detection that addresses this issue.

Example 1 (Verification of Natural Language Text). Melanie Martin proposes a system for the verification of natural-language texts [15] [16]. The goal of her research is to test the hypothesis that information on the Web can be verified automatically with minimal user interaction. Text verification is done by building and evaluating an interactive system. Reliability is defined as a measure of the extent to which information on a given web page can be trusted, and verification is defined as a reasonable determination of the truth or correctness of a statement by examination, research, or comparison with similar text. Given a document to be verified, the system retrieves and clusters other relevant documents, and offers the user an interactive process in which the user can verify the contents of the text based on the other texts. The system consists of a number of components:

- **Document Clustering:** Clusters relevant documents based on whether or not they agree or disagree with the statement to be verified.
- **Document Retrieval:** Uses standard IR techniques.
- **Query Processing:** Uses standard IR techniques.
- **Reliability Ranking:** Estimate the likelihood that a statement on the Web can be trusted using standards developed by information scientists, as well as linguistic aspects of the page, and the link structure of associated web pages.

Publication Date. The publication date is the date on which the text was published. Related information that may be of interest is the date of the latest change. For some types of content, this may have implications for the trustability. Also, this information may be used to determine whether the information is out-of-date. E.g., information published a decade ago on topics that have seen a lot of changes in the last few years are less likely to be trustworthy. The publication date may be obtained directly from the text itself or by spidering the web page.

Source Reputation This indicator concerns the reputation of the source: the author, the author's organization, and/or the publisher of the text. If the source is known and trusted, then the text may be trusted as well. In order for the name of the author, organization, or publisher to be determined, one may use named entity recognition. This is a technique that analyzes a text to extract proper names.

Example 2 (WikiScanner). WikiScanner (wikiscanner.virgil.gr) consists of a publicly searchable database that links millions of anonymous Wikipedia edits to the organizations where those edits apparently originate from. The information is compiled by cross-referencing the edits with data on the owners of the associated block of IP addresses. WikiScanner does not work on edits made under a username. WikiScanner also cannot distinguish edits made by authorized users of an organization's computers from edits made by unauthorized intruders, or by users of public-access computers that may be part of an organization's network. The information offered by WikiScanner can be used to determine the organization related to an anonymous Wikipedia edit.

Subjectivity. Natural-language texts convey emotions and opinions about people, politicians, products, companies, *et cetera*. A text's subjectivity may be used to determine whether one will trust the information. A text may have a negative subjectivity or a positive subjectivity on a given topic, and the level of subjectivity may vary from low to high. Sentiment analysis (a.k.a. opinion mining) [20] analyses texts for their emotional content, mostly to determine whether this is positive or negative attitude. Interesting works in the area of subjectivity have been done by [8] and [29]. [8] derives a "Dutch subjectivity lexicon" automatically, i.e., a set of Dutch words that express subjectivity, using an English subjectivity lexicon, an online translation service, and a Dutch general purpose thesaurus. [29] presents a way to automatically determine subjective words and collocations (e.g., "unwise in", "ad hoc", and "drastic as") from corpora. Unique words we found to be subjective more often than expected.

Example 3 (Vox-Pop). The Dutch website Vox-Pop (<http://vox-pop.nl/>) performs sentiment analysis on thousands of Dutch-language web pages with news and opinions on news. On the basis of this analysis, it presents the *vox populi* on several topics. In Figure 2, for example, the graphic at the top shows the positive and negative score gained by the Dutch Prime Minister according to the news reported by 11 sources in one month period. Users may create personalized widgets for the persons or the topics that they want to monitor.

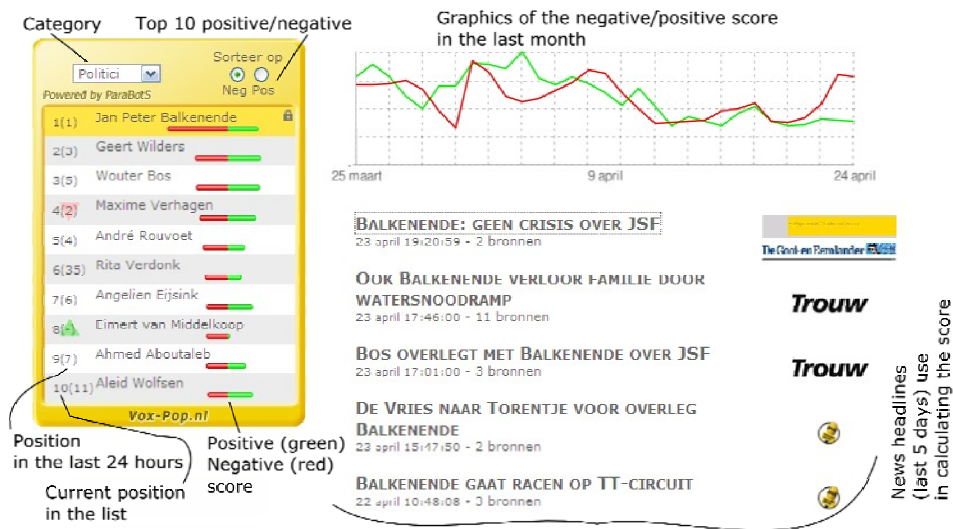


Figure 2 Example of analysis in the Dutch web site vox-pop.nl that automatically analyses the positive and negative feelings towards/against politicians according to the news. (Bronnen = sources, Trouw = Trust)

Quality of Recommendation Analysis

Modern recommender systems must cope with an increasing demand of complexity; for instance, a recommendation application for restaurant should take into account the contextual information (e.g., has the restaurant been recommended for a romantic dinner or for a business lunch?). Adomavicius and Tuzhilin suggest that recommender systems can provide better recommendation if they are extended according to the following criteria [1]:

- *A more comprehensive understanding of users and items:* this criterion suggests extending the simple keyword-based techniques with advanced profiling techniques based on data mining rules, sequences, and signatures can be used to build user profiles and to describe items.
- *Multi-dimensional recommendations:* this criterion suggests extending recommendations also using contextual information. The utility of a certain item or product to a user may strongly depend on time, place, and situational factors.
- *Multi-criteria ratings:* this criterion suggests extending the single-criterion ratings with multi-criteria ratings (e.g., restaurant rated according to food, decor, and service).
- *Non-intrusiveness:* this criterion suggests that recommender systems may use nonintrusive techniques (e.g., by adding geographical identification metadata to media).

Advanced recommender systems already incorporate some of the previous extensions. For example Beauté-test (www.beaute-test.com), an on-line shopping service in the cosmetics business, uses multi criteria ratings and quite detailed user profiles. To provide a guideline for the evaluation of how recommendations provided by the recommender system can be trustworthy for the users, we have analyzed a large number of recommender sites and we have identified a list of indicators for the trustworthiness of their recommender or review services. The following list comments on the most relevant indicators, according to our experience. We also considered the solutions that have been presented in relevant conferences in the field (e.g., [21][22][23][24]). The list is still incomplete and it will be extended as future work.

Robustness of Rating System: This criterion focuses on the strategies used to provide recommendations and ratings, and on the solutions that the web-site maintainers apply to avoid attacks that tries to subvert the fairness of reviews and consequently to bias recommendations.

Multi-criteria reviews. This criterion focuses on the number and on the type of the criteria used to rate the quality of a product or a service. More and more web site are now moving from offering only one general overall “quality” rating towards the use of multi-criteria. For example E-bay (www.ebay.com) has changed its rating system by asking buyers to enter four ratings corresponding to four different criteria about sellers (communication, shipping, speed & charges, and description adequacy).

Recognition and Roles. This criterion focuses on evaluating the availability of different roles in the users. For example, recommenders can be classified depending on their activities (e.g., sellers or buyers) or depending on their recognized reliability (e.g., experts or simple users) as CNET.com (www.cnet.com) does. The criterion also describes if the recognition mechanism specifies how roles are managed, for instance, how a user changes role, if a certification is required to be addressed as an expert etc. SierraTradinPost (www.sierratradinpost.com), a web site for outdoor clothing and accessories can certify its recommenders and identifies them with a label.

Example 4 (Dooyoo). The role management of Dooyoo (www.dooyoo.com) relies on special members called Guides: guides are skilled reviewers of the Dooyoo community. They are experts in their categories and provide advice to new members by helping them to write the most helpful reviews as possible. They read reviews and they recommend Crowns. A Crowned Review is an exceptionally well-written review. Guides also alert Dooyoo about potential abuses or frauds. To be a good Dooyoo Guide an identity needs to show that he (i) has been rating reviews fairly and consistently (ii) has a passion for helping others (iii) understands how Dooyoo works (iv) tries to read all the reviews in a chosen category (v) is prepared to volunteer his time and energy (vi) has written a good number of Crowned Reviews.

Rewards. This criterion evaluates the presence of rewarding mechanisms that stimulate recommender in leaving ratings and textual comments. We have already explained the Crown awards in Dooyoo. Another example is given by Epinions (www.epinions.com) that encourages members leaving good quality reviews, by rewarding them with Eroyalties. Eroyalties, which are redeemable in U.S. dollars tracks how much a member earns for writing reviews. These bonuses are not tied directly to product purchases, but are based instead on more general use of reviews by consumers when they making decisions. Thus a member could potentially earn as much for helping someone make a buying decision with a positive review as he could for helping someone avoid a purchase with a negative review.

Personalization. This criterion evaluates the presence of mechanisms for the personalization of ratings, i.e., the presence/absence of solutions that help users understand whether a given opinion matches their taste, scope, context, etc. For example, users can block certain opinions because inappropriate, or unconvincing. Alternatively, users can perform searches based on their profiles or preferences. Explainability, that is the presence of a written explanation why a certain good has been recommended, is also a way to improve the degree of personalization [14].

Web of Trust. This criterion evaluates the presence of solutions that facilitate the establishment of a user's trusted network of recommenders, or alternatively the presence of solutions for the evaluation of the quality of the source of recommendations. A trusted network can be constructed indirectly, e.g., by giving the users the possibility of evaluating the trustworthiness of recommenders depending on their recommendations, or it can be a social network imported by some already existing community. It also comments whether the management of the trust network is left to the user or is controlled by the service provider. Many review systems start offering their users an ad-hoc Web of Trust. In this case the Web of Trust of a member is the network of reviewers whose reviews and ratings have been consistently found to be valuable by that member. The Web of Trust mimics the way people share word-of-mouth advices every day, and it is based on the following sociological concepts:

- Friends have a proven track record. If a friend consistently gives you good advice, you're likely to believe that person's suggestions in the future.
- You know which preferences you and your friend share. If you both like the same types of films, you're more likely to trust your friend's recommendations on what to see.

The scientific community is devoting more and more attention to the role of trusted recommenders in improving the accuracy of recommendations (cf. [13]). It has been proved that trust-based recommendations are both competent and well intentioned; the resulting recommender system is more accurate and proved to be resistant to Sybil attacks [5].

Conclusion and Future Work

We started with an important question in mind: how can we to provide a scheme for the evaluation of the trustworthiness of a recommender site? As reported in [6], not all the word-of-mouth is equal and consumers need to distinguish between 'good' (honest, true) and 'bad' (dishonest, false) information before deciding upon a purchase. Therefore the issue "*how consumers choose their information source and the mechanisms that help them find trusted information sources will be of particular interest for future research*" is of paramount importance for the success of a recommender site.

Recommender sites were born with the goal of helping users to cope with the information complexity typical of the Web 2.0 paradigm, but recommender systems' qualities and efficiencies also depend upon different factors. In this preliminary study we have identified and commented some of the most common indicators of trustworthiness. We addressed the following different "perspective" upon a recommender site: user interfaces, automatic content analysis, and quality of recommendations. We also found that techniques from language technology and web spidering can be used to support higher-level trustability analysis. Existing systems such as Melanie Martin's system and WikiScanner implement some of these ideas. The following table summarized our preliminary check list of relevant indicators for trustworthy recommender sites, according to the criteria we have introduced so far.

| | User Interface | Content Analysis | Recommendations |
|-----------------------------------|---|--|--|
| Trustworthiness indicators | Certification from trusted parties; Management infos; Past reviews patterns; References from past and current users; Offer/Compare reviews; Source of reviews present and easy to check; Reputation of the reviews managed. | Contact info; Grammar/Spelling; Duplication; Publication date; Source reputation; Subjectivity. | Robustness; Multi-criteria reviews; Recognition and roles; Rewards mechanisms; Personalization; Web-of-Trust. |

Table 1 Preliminary check list for trustworthiness evaluation of recommender sites.

This checklist should be considered an early attempt to classify the variables or the factors that affect the perception of trust in the design of trustworthy e-commerce and recommender websites. Hence, validation is required to assess its applicability in different contexts and to test its comprehensiveness. Such a validation exercise would be the next step of our research. Nevertheless, knowing what factors determine trust does not suffice: We need to research ways to visualize these aspects in order to aid users in their assessment of the trust-related variables mentioned in Table 1. How can we present them? What effect does this have on purchasing behavior and on perceived ease of use? And what role can social communities play with regard to trust? Finally, an interesting future work is to understand how to incorporate these variables into the design of recommender systems. Here, we suggest the need of tools for web design that support software engineers in automatically including "trust-enhancing" features in their products during the development phase.

References

- [1]. Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Trans. on Knowledge and Data Engineering*, 17(6), 734–749.
- [2]. Balabanovic M., & Shoham., Y. (1997). Fab: content-based, collaborative recommendation. *Comm. of the ACM*, 40(3), 66–72 Chopra, K. and Wallace, W. A., Trust in Electronic Environments, Proc. 36th Hawaii Int. Conference on Systems Science, January 6-9, 2003, Big Island, HI, USA, 2003, IEEE Computer Society
- [3]. Cassell, J., & Bickmore, T. (2000). External manifestations of trustworthiness in the interface. *Comm. of the ACM*, 43 (12), 50-56.
- [4]. Clements, M., de Vries, A. P., Pouwelse, J. A., Wang, J., and Reinders, M. H. T., Evaluation of Neighbourhood Selection Methods in Decentralized Recommender Systems, *Proc. ACM SIGIR Workshop on Large Scale Distributed Systems for Information Retrieval, LSDS-IR 07, Amsterdam, The Netherlands, 2007*
- [5]. Dell'Amico, M., and Capra, L. (2008). SOFIA: Social Filtering for Robust Recommendations. In *IFIP Int. Federation for Information. Processing, 263, Trust Management II*, 135-150. Boston, Springer.
- [6]. Duan, W., Gu, B., and Whinston, A. B., Do online reviews matter? An empirical investigation of panel data, *Decision Support Systems*, 45, 1007-1016, 2008.
- [7]. Hertzum, M., Andersen, H. H. K., Andersen, V. & Hansen, C.B. (2002). Trust in information sources: seeking information from people, documents, and virtual agents. *Interacting with Computers 14*, 575–599
- [8]. Jijkoun, V. and Hofmann, K., *Task-Based Evaluation Report: Building A Dutch Subjectivity Lexicon (D-14)*, Technical Report, Version 1, 2 September 2008
- [9]. Josang, A. and Lo Presti, S. (2004) Analyzing the Relationship Between Risk and Trust. In: Second International Conference on Trust Management (iTrust 2004), March 29 - April 1st 2004, Oxford, UK.
- [10]. Kelton, K., Fleischmann, K. R. and Wallace, W. A., Trust in Digital Information, *J. of American Society for Information Science and Technology*, 59(3), 363-374, 2008.
- [11]. Chopra K., & Wallace, W. A. (2003), Trust In Electronic Environments, *Proc. of the 36th Hawaii Int. Conf. of Systems Science (HICCS 03)*, IEEE Computer Society, 331.1
- [12]. Kramer, M. A., Self-selection Bias in Reputation Systems, *Proc IFIPTM 2007: Joint iTrust and PST Conf. on Privacy, Trust Management and Security, July 30 - August 2, 2007, New Brunswick, Canada*, Trust Management, 255-268, Springer-Verlag
- [13]. Lathia, N., Hailes, S., & Capra, L. (2008). Trust-Based Collaborative Filtering. In *IFIP Int. Federation for Information. Processing, 263, Trust Management II*, 119-123. Boston, Springer
- [14]. Koren, Y., Tutorial on Recent Progress in Collaborative Filtering, *Proc. the 2008 ACM Conf. on Recommender Systems (RecSys) 2008*, October 23-25, Lausanne, Switzerland
- [15]. Martin, M. J., *Reliability and Verification of Natural Language Text on the World Wide Web*, dissertation, New Mexico State University, 2005
- [16]. Martin, M. J. (2004), *Reliability and Verification of Natural Language Text on the World Wide Web*, ACM-SIGIR Doctoral Consortium, July 25, 2004, Sheffield, England.
- [17]. Muir, B.M., (1994). Trust in automation (Part I): Theoretical issues in the study of trust and human intervention in automated systems. *Ergonomics*, 37(11), 1905- 1922.
- [18]. Nass, C.I., Fogg, B.J., & Youngme, M. (1996), Can Computer be teammates? *Int. J. of Human-Computer Studies*, 45 (6), 669-678.
- [19]. O'Donovan, J., & Smith, B. (2005), Trust in Recommender Systems, *Proc. of the 10th Int. Conf. On Intelligent User Interface (IUI), San Diego, CA, USA, 10-13 Jan 2005*, ACM, 167-174.
- [20]. Pang, B., & Lee, L., Opinion Mining and Sentiment Analysis (2008). *Foundations and Trends in Information Retrieval*, Vol. 2, No 1–2, pp. 1–135, 2008
- [21]. *Proc. of 2007 ACM Conf. on Recommender Systems (RecSys07)* Minneapolis, MN, USA.
- [22]. *Proc. of 2008 ACM Conf. on Recommender Systems (RecSys 08)*, October 23-25, Lausanne, Switzerland.
- [23]. *Proc. of IFIP Joint Conf. on Security, Trust Management, and Security (IFIPTM) 2007*, New Brunswick, CA
- [24]. *Proc. of IFIP Joint Conf. on Security, Trust Management, and Security (IFIPTM) 2008*, Trondheim, Norway
- [25]. Resnick, P. and Sami, R. 2007. The influence limiter: provably manipulation-resistant recommender systems. In *Proc. of the 2007 ACM Conference on Recommender Systems*, USA, October 19 - 20, 2007, ACM, 25-32
- [26]. Shneiderman, B., Designing trust into online experiences, (2004). *Comm. of the ACM*, 43(12), 57-59.
- [27]. Tseng, S. & Fogg, B.J. (1999). Credibility and computing technology. *Comm.s of the ACM*, 42(5), 39-44.
- [28]. Veltmaat, L., (2009) *Taaltechnologie voor betrouwbaarheid van informatie*, TR2009-01-12, Telematica Instituut.
- [29]. Wiebe, J., Wilson, T., Bruce, R, Bell, M., and Melanie, M. (2004), *Learning Subjective Language*, in: *Computational Linguistics*, Vol. 30, Issue 3 pp: 277-308.
- [30]. Yang, J., Wang, J., Clements, M., Pouwelse, J. A., de Vries, A. P., & Reinders, M. H. T. (2007), An Epidemic-based P2P Recommender System, *Proc. ACM SIGIR Workshop on Large Scale Distributed Systems for Information Retrieval (LSDS-IR 07), Amsterdam, The Netherlands, 2007*.