



GRAPPLE

D6.2a Version: 1.0

Integrated user model rating and reputation system

Document Type	Deliverable
Editor(s):	Dominikus Heckmann (DFKI)
Author(s):	Dominikus Heckmann(DFKI), Erwin Leonardi(TUD), Rafael Math(DFKI), Fabian Abel(I3s), Eelco Herder(I3s), Geert-Jan Houben(TUD)
Reviewer(s):	Fabrizio Giorgini(GILABS), Christian Glahn(OUNL)
Work Package:	WP6
Due Date:	31-10-2009
Version:	1.0
Version Date:	31-01-2010
Total number of pages:	12

Abstract: This document describes the design of an integrated user model rating and reputation system. It bases on the higher goal to reach the level of "trust" between the different participating LMSs or "trust in the World WideWeb" in general. We present an overview over the scientific background and discuss design decisions that have to be realized in the follow up deliverable D6.2b).

Keyword list: Rating, Reputation, User Modelling

Summary

In a distributed user modelling scenario the shift to Web 2.0 leads to the idea that the quality, the trustworthiness and the correctness of statements about users can be judged by the so called “wisdom of the crowds”. In a distributed user model scenario, the quality of the user model data can vary and make decisions based on user models especially difficult. The idea of this deliverable is to enrich the user models with quality ratings to support the trust-idea of the Semantic Web approach as well as a social computing approach to improve the user models. We present the design of a rating system for GrappleStatements as well as the design of a reputation system for information providers.

Authors

Person	Email	Partner code
Dominikus Heckmann	heckmann@dfki.de	DFKI
Erwin Leonardi	e.leonardi@tudelft.nl	TUD
Rafael Math	Rafael.Math@dfki.de	DFKI
Fabian Abel	abel@l3s.de	L3S
Eelco Herder	herder@l3s.de	L3S
Geert-Jan Houben	g.j.p.m.houben@tudelft.nl	TUD

Table of Contents

AUTHORS	2
TABLES AND FIGURES.....	3
1 INTRODUCTION	4
2 RELATED WORK	4
2.1 Collective Election Systems	4
2.2 Rating Systems	5
2.3 Trust & Reputation.....	6
3 DESIGN OF THE RATING SYSTEM.....	7
3.1 What can be rated in GRAPPLE?	7
3.2 Who is allowed to rate in GRAPPLE?	8
3.3 Design of the Database Schema	8
3.4 Design of the User Interface	9
4 DESIGN OF THE REPUTATION SYSTEM	9
4.1 Design of the Trust Model.....	9
4.2 Design of the Database Schema	10
4.3 Design of the User Interface	10
5 DISCUSSION & SUMMARY	10

Tables and Figures

List of Figures

Figure 1: Instantiated Database Schema for Testing Basic Ratings 9

Figure 2: User interface widget for a 5-scale rating 9

Figure 3: Multi-dimension ration (in analogy to [4]) 9

List of Acronyms and Abbreviations

GALE	GRAPPLE Adaptive Learning Environment
GRAPPLE	Generic Responsive Adaptive Personalized Learning Environment
GEB	GRAPPLE Event Bus
GrappleBroker	A service to store and retrieve partial user models represented as GrappleStatements
GrappleStatement	A Unit of information about a user together with contextual meta data.
GUMF	GRAPPLE User Model Framework
LMS	Learning Management System
OWL	Ontology Web Language
RDF	Resource Description Framework
TEL	Technology-Enhanced Learning
UM	User Model (or sometimes “User Modelling”) a set of GrappleStatements
WP	Work Package

1 Introduction

In a distributed user modelling scenario the shift to Web 2.0 leads to the idea that the quality, the trustworthiness and the correctness of statements about users can be judged by the so called "wisdom of the crowds". In a distributed user model scenario, the quality of the user model data can vary and make decisions based on user models especially difficult. The idea of this small deliverable is to enrich the user models with quality ratings to support the trust-idea of the Semantic Web approach as well as a social computing approach to improve the user models. The integration into the rest of the GRAPPLE system is simply realized by including the ratings into the basic information unit, namely the GrappleStatements.

This deliverable consists of a design of extensions to the D6.1a/b services, taking into account that different (parts of) distributed user models must be rated differently, which leads to ranking of the reliability of user model values. The implementation will include technical documentation. The implementation will follow in deliverable (D6.2b).

2 Related Work

In this section, we briefly discuss several approaches that are related to rating, trust, and reputation systems on the Web.

In general, a rating is the *subjective* evaluation or assessment of something, in terms of quality, quantity, or some combination of both. Most rating systems centre around rating content, often user-contributed content and they frequently help apply community values and acclaim to that content. Users can rate a service by using positive/neutral/negative values (e.g. eBay rating system), 5-point scales (e.g. using stars), or 10-point scales.

In [5], there are three important elements of a good rating system: granularity, distinction, and a statistical basis. With granularity it means that the users have more options such that the users a greater ability to differentiate their ratings. For example, 5-scale rating that allows half-points (e.g. 0.5, 1.0, 1.5, ..., 5.0) gives the user more options for expressing their ratings compared to the one without half-points (e.g. 1, 2, 3, 4, and 5). Further, a problem that could occur when a user gives ratings is inconsistency in his/her ratings over time. If a user rates everything low for a while, and then rates everything high, then he/she has a greater chance of biasing the overall rating pool. Worse, his/her individual ratings aren't meaningful, because, for example, we cannot look at two items, see that one is a "6" and another is an "8", and truly believe that the user likes the "8" a fair amount more than the "6". Therefore, the criteria of the ratings have to be distinct in order to help the users to stay consistent and give more meaningful ratings. The following is an example of distinct rating systems with 5-scale rating for a game.

- 5 – Excellent game. Always want to play.
- 4 – Good game. I like to play.
- 3 – Average game, slightly boring, take it or leave it.
- 2 – Bad game, likely won't play this again although could be convinced.
- 1 – Extremely annoying game, won't play this ever again

Finally, a good rating system should be statistically sound. This means that the average of all given ratings of a certain item should be meaningful to the users. Of course, the best way to make the ratings statistically sound is with volume. By having, for example, thousands of ratings for an item, any anomalies are going to become noise. One way to compute the average ratings that are statistically sound is by applying Bayesian average. The idea behind a Bayesian average is to normalize ratings by pushing them toward the average rating for an item.

2.1 Collective Election Systems

Collaborative rating seen under the social Web 2.0 community approach, points towards collective election systems. The long history of collective election systems continues today in the World Wide Web. According to Allen and Appelcline [6], in particular people of the western culture appreciate a high participation in election processes. The reason for this syndrome is the belief that the fairest result can be achieved if a great diversity of people takes part in the election process.

In the course of time, three distinct kinds of election systems evolved, which differ in the intended results: subjects should either make a choice between objects or tell their opinion about objects or compare different objects.

So-called selection systems allow the target-oriented selection between several elements. Selection systems include representative systems and consultative systems, both of which are political election systems. In representative systems, the election is performed through a delegate (representative) like in the republican election system. Consultative systems constitute the real idea of democracy because the election is done by direct ballot with a decision by majority. A variant of the consultative systems is formed by the consensus systems (e.g. the Westminster system), in which the power of the masses is stronger than the one of an individual; thus, the "tyranny of the majority" can be avoided.

A subcategory of the selective systems is built by opinion systems, which mainly act as an indicator of how people have decided in an election or how they will decide in the near future. That is, these systems constitute surveys before an election or first projections during an election.

The third type of election systems is comparison systems, which support the measurement of individual elements among themselves. They can be divided into three categories: comparing ranking systems, which are mainly objective and classify participants of a competition; comparing rating systems, which mix up objective and subjective opinions and are used to especially rate entities (e.g. the product rating system of Amazon.com); reputation systems mix up objective and subjective opinions as well, but are rather used to rate persons than things.

2.2 Rating Systems

This section describes the concept of rating systems in general as well as different criteria that are essential for a good rating system by means of eBay's rating system as an example

The online auction house eBay was one of the first web sites to make use of rating systems in an extensive way. Their rating system was considered as innovation in the field of e-commerce and should be used to determine the trustworthiness of the users by providing facilities to give a positive, negative, or neutral feedback for each transaction. The feedback score is calculated as the number of unique positive feedback minus the number of unique negative feedback. Unique feedback means that repeated ratings performed by the same user in the same week are excluded. The positive feedback score is calculated by dividing the number of positive feedback ratings by the total number of feedback ratings obtained, resulting in a percentage value.

Additionally, the user can add an arbitrary textual description of the transaction rating.

As almost all ratings given by eBay users are positive, however, the feedback scores have a very large expressiveness because they do not tell much more than how many transactions have been performed by the respective member.

Allen and Appelcline investigated several rating systems [6], among others also the one of eBay. They defined the following criteria that are essential for the quality of a rating system:

- A rating system has to be granular. Statistics show that people tend to ignore the lower part of the scale of rating systems. Consequently, the decimal places of mean feedback scores are of decisive importance. The rounding of mean scores or the simplification of ratings to a "thumbs up" or "thumbs down" would be the wrong strategy.
- One major issue with rating systems is the one of inconsistent ratings performed by the user. If a user expresses the same opinion in two different rating processes with two distinct ratings, the feedback statistic is influenced. That's why the user should be supported to remain consistent by creating clearly defined meaning which feedback value corresponds to which opinion of the object. The larger the rating scale is the larger is the risk that the user gets confused.
- Rating systems must be statistically reliable. The best way to realize this is to collect a large number of ratings. If there are only a few unique ratings for an object, the risk of an unreliable overall rating increases. The more ratings are performed for an object, the smaller is the effect of outliers. Optimally, one should assign a smaller significance to objects with very few ratings.

- If users rate each other, there is a high risk that there are almost only positive ratings. The reason for this effect is that users are afraid of getting a negative rating in revenge for a negative feedback. That's why rating system should not be bilateral.
- User ratings should have a concrete usage within a web site. One possibility to achieve this is to display rankings, in which the best and the worst rated objects can be investigated.
- A clear user interface is essential for a rating system as bad designed, confusing graphical user interfaces discourage the user; thus, the quality of the ratings can suffer.

Although being part of eBay's success story, Allen and Appelcline pointed out that the rating system of eBay in its early version shows the following weaknesses:

- It is not granular because it only supports the three choices of positive, neutral and negative feedback.
- There are no clear guidelines which user behaviour should result in which rating.
- The rating system is statistically not reliable. The rating only depicts a rough amount of performed transactions rather than a real, subjective measure for the trustworthiness of the user.
- As the users rate each other, they hesitate to give negative feedback because they are afraid of getting a negative rating in revenge.
- The rating system basically does not have significance because there is no possibility to exclude users with a low rating score from auctions, for instance.

In May 2007, eBay introduced a reworked version of the rating system [11]. The so-called detailed rating makes it possible to rate further aspects of a transaction rather than only providing the three categories positive, neutral and negative. Using the following questions, the transparency of the auction platform shall be increased:

1. Was the item delivered as described?
2. How would you rate the communication with the seller?
3. How long took the shipping time?
4. Were the shipping and handling charges adequate?

Using the new five-star rating system, users can give detailed statements within these additional rating dimensions. The mean value of these rating categories is shown in the user profile, among the standard feedback score.

2.3 Trust & Reputation

In the real world, a group of users provides ratings for items (in our case, also the information providers, apart from the GrappleStatements, see Section 3.1). However, not all of users who rate the information provider are trustworthy. Hence, the reputation of this group of users has to be determined. Note that not all given ratings are fair, reliable, and trustworthy. Reputation is the opinion (more technically, a social evaluation) of the public toward a person, a group of people, an organization, or an information source. Reputation systems are often useful in large online communities in which users may frequently have the opportunity to interact with users with whom they have no prior experience or in communities where user generated content is posted. In many cases, reputation is used to establish trust. The past interactions or performance of an information source captured in its reputation are combined to assess its future behaviour. Once the reputation is determined, trust can be established

Zhang and Cohen [18] proposed a method to determine trustworthiness of *advisors* - users who provides ratings for the information provider. Note that the existence of malicious advisors is discussed in [19]. This method determines the trustworthiness of advisors by computing and combining two values: *private* and *public reputation values*. The private reputation of an advisor is based on ratings that are given by the advisors to information providers with whom the user has already had some experience. The intuitions is that if the advisor is reputable and has similar preferences as the user, then the ratings given by the user and advisor are likely similar. The public reputation value is important in the case of lacking private knowledge about the advisor. This reputation value is based on the advisor's ratings of all information providers in the system. The combination of the weighted private and public reputation values can be used to represent the trustworthiness of the advisor. The weight is significance in specifying the reliability of reputation values. For

example, the user who relies on private reputation value will give more weight to private reputation value and consider less the public reputation value.

In [19], Richardson et al. also addressed the problem of possibly unfair or unreliable ratings. The approach in [19] provides a means of merging trust that is robust to noise and emphasizes on personalized trust. One interesting feature of this approach is the propagation of trust through users' network. Note that this approach is described as a generalization of PageRank [20] to the Semantic Web. In this approach, a user is required to build his/her Web of Trust, a small set of users whom he/she trusts. Based on this Web of Trust, the trust values of all other users are composed. Then, the trust value of an information provider is derived and computed using some aggregation functions along each possible chain of trust from the user to the information provider. The result of this computation is not agglomerate "trustworthiness" of each user. Instead, each user receives a personalized set of trusts, which may vary widely from person to person.

Furthermore, the domain of knowledge (context) can be considered while computing the trust of information provider. Ding et al. [21] addressed this issue and elaborated on several types of trust: *domain expert trust* (trust in an agent's domain knowledge), *recommendation expert trust* (trust in an agent's ability to refer other agents), *similar trusting trust* (two agents having similar trust in other agents), and *similar cited trust* (two agents being similarly trusted by other agents). This context consideration is also supported by Bizer and Oldakowski [22]. In [22], Buzer and Oldakowski made several claims: (1) any statements contained in the Semantic Web must be considered as claims rather than facts until trust can be established, (2) it is too much of a burden to provide trust information that is current, (3) context-based trust matters, and (4) it is possible to use "content-based trust" using common sense rules of the world to make a trust decision.

3 Design of the Rating System

After the literature survey in the related work, we decided to introduce a distinct five-scale rating model because the most prominent and most effective social web pages and related work implement such a five-scale rating. We take the path of the wisdom of the crowds. Now we have to look closer at questions that meet the ideas of the GRAPPLE scenarios.

3.1 What can be rated in GRAPPLE?

First of all, the basic unit of the GRAPPLE user modelling framework GUMF is subject to be rated: the so called GrappleStatement, as defined and introduced in deliverable D6.1a:

Definition 1: *A GrappleStatement is defined as an information unit, maybe enriched with meta-data about it.*

The final syntactical representations of GrappleStatements are defined in deliverable D2.1 and described in D7.2 in a Semantic Web language. For this deliverable D6.2a), it is only of importance that we can treat each GrappleStatement as a resource, which means each statement can be referenced by a unique identifier.

Examples of such GrappleStatements in pure textual form are:

S1 := Mary likes chemistry

S4 := Mary prefers learning style A to learning style B

S7 := Mary is colour-blind

S8 := Peter likes mathematics

S9 := Peter is a good Java programmer (claimed by Mary)

S10 := Peter knows the concept of object oriented programming (learned with the AHA! system)

S12 := Peter knows the Theorem of Pythagoras (learned with Moodle, last year)

S13 := Mary does not like chemistry very much (claimed by Peter, yesterday)

These examples as shown above refer to pure statements (S1-S8) and statements with metadata information (S9-S13). An interesting conflict holds between S1 and S13: both claim opposite values about

Mary's interest in chemistry. The whole motivation of this work package task D6.2 is to allow everybody, or selected people, to "rate" statements, in order to compare them and in order find out the truth, or better in terms of Web 2.0: the wisdom of the crowds.

Now the question arises, what part of the statement should be rated? Since the GrappleStatements consist of a main part and a meta part, both can be the subject to the rating. Interesting dimensions are for example:

- Actuality of the (main part) information
- Correctness of the (main part) information
- Trustworthiness of the information (influenced by meta parts like "creator of the statement")

The second subject that could be rated directly – apart from GrappleStatements – is the creator of the statements, the so-called information provider directly. This will be discussed in Chapter 4 under the topic of "Trust and Reputation".

3.2 Who is allowed to rate in GRAPPLE?

According to [23], peer production is gaining acceptance for the creation of structured knowledge and semantic metadata. Peer production refers to a web-based model of production where a large number of contributors work on a common project.

In this deliverable, we propose rating-based incentive mechanisms to assure the quality of structured knowledge created collaboratively. Users review contributions created by other users and LMSs and rate them according to the quality perceived. To motivate users to contribute to the community, we reward them with points according to the extent and quality of their contributions. We compute the quality of contributions, in turn, based on their ratings.

The weight of the rating can be adjusted according to who is rating. This leads to the question:

Who should we allow in GRAPPLE to rate statements and the information provider? Possible candidates are for example:

- The creator (can be everybody)
- The described person (user/student)
- The peer of the described person
- The teacher/tutor of the student

Everybody, in Web 2.0 terms: "the crowd". For sure, for most situations, we will allow the creator, the described person and the teacher of the student to rate statements. For testing purposes, we will allow the peer and the crowd in some selected situations to rate statements. A clear discussion about "who is allowed to read and to rate which student information" in form of GrappleStatements has to be undertaken as soon as the second years prototype is running and the issue can be discussed with concrete scenarios.

Before we can come up with a decent weight system, we have to gain a clear understanding of the different roles in relation to rating. We expect that this weight system will not be domain independent. Thus, we plan to solve this issue in deliverable D6.2b, once we have implemented the rating module and once we have received domain dependently rated GrappleStatements.

3.3 Design of the Database Schema

The proposed rating system is a 5-scale rating, storing user ratings in a relational database table as RDF-triples (s-p-o), whereas s denotes the ID of a statement to rate, p the predicate (IS..rated) and o (object) a digit between 1 (bad) and 5 (very good). Further information to store will be the Dublin Core dimensions for creator, date of creation and possibly the user's individual rating history, which can be provided as comma separated list that contains digits 1-5 in order of their appearance. The following figure depicts this table at the example of a statement rating.

s	p	o	creator	created
STATEMENT..9B9BQkED5PD9eIagcIPWCj	IS..rated	4	USER..rmath	2009-11-23 16:25:58
STATEMENT..9B9BQkVd9HOKuLXwPv2D4j	IS..rated	3	USER..rmath	2009-11-19 12:51:17
STATEMENT..9B9BQk00Az3EYR2qhjutpj	IS..rated	4	USER..rmath	2009-11-19 12:51:21
STATEMENT..9B9BQPLMsDQjPcblykLwoj	IS..rated	4	USER..rmath	2009-11-19 12:51:25
STATEMENT..9B9BQPZ9J5CGcbNpsOnKuj	IS..rated	4	USER..rmath	2009-11-19 12:51:31

Figure 1: Instantiated Database Schema for Testing Basic Ratings

From the user-statement relations of this table the average rating will be computed by adding up all the users' ratings for a certain statement divided by the number of users as it can be counted from the table. Moreover, to save computing power, the result of the averaging process and the total number of ratings will be stored for every statement individually in another database table. In order to identify the related statement the statementID from Figure 1 is applied. As soon as a user submits a rating, or changes a previous rating, the average value and – if necessary – the number of raters will be updated. Due to this strategy expensive database requests resulting from the averaging process can be reduced to submitting ratings. Displaying certain ratings occur much more frequent and can be accomplished by a less expensive select-request on the pre-computed values.

3.4 Design of the User Interface

The main user interface will be an AJAX-based widget to enter and to display the individual as well as the community rating of the GrappleStatements. This widget, see Figure 2, will be developed together with the GVIZ component of WP 4.

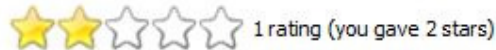


Figure 2: User interface widget for a 5-scale rating

In the case of a fine-grained multi-dimension rating, a listing of independent dimensions as shown in Figure 3 can be used to allow for multi-dimension rating in one row.



Figure 3: Multi-dimension rating (in analogy to [4])

Of course, apart from this direct method to enter ratings via a user interface, we will allow to update a rating via the standardized GRAPPLE web services, to be sent via the GRAPPLE event bus GEB.

4 Design of the Reputation System

4.1 Design of the Trust Model

The GRAPPLE Reputation System should be realized by a notion of "Trust in terms of reliability of the information provider".

In the first prototype the level of trust can be set manually by the different information consumers, (to say, by the LMSs using information from other LMSs, GRAPPLE or GALE) by choosing different decorations and medals for the different information providers. In a second step, the trust of an information provider will be computed by counting the number of excellent rated statements in the past. In this sense, the reputation system will partly base on the results of the rating system of Chapter 3. A system's reputation comprises the count of positive and negative ratings in that system's history. The concrete formula has to be set up together with all LMS partners, as soon as the first critical amount of ratings has been collected.

4.2 Design of the Database Schema

The database schema for the first prototype will be pure RDF triples that realize the relations

/ system A (as information consumer) | level of trust to | system B (as information provider) |

while the level of trust will be a symbolic value of the range “low – medium –high”.

4.3 Design of the User Interface

The design of the user interface will be aligned with the activities in WP4.5, our suggestions are the following:

- the three levels of trust that can be given to information provider by information consumer will be represented by three different medals: “gold, silver, and copper” with red ribbons, as shown below.



- on the other way round, the information provider can give reputation models to the information consumers with blue ribbons, as shown below



This user interface works for the manually given medals as well as the rating-based calculated reputation model.

We expect this medal-based rewarding system to motivate the different partners to provide high-quality user model information to each other. If this turns out (after the first round of evaluations) not to work properly, alternative rewarding ideas (like monetary models) have to be analyzed to ensure the exchange of user model data among different GRAPPLE partners.

An issue that finally arises is the question how to integrate both, the rating and the reputation system, into the rest of the GRAPPLE system. Here, the modular approach and especially the design decisions of the basic information unit, namely the GrappleStatements, help a lot. We can simply include the ratings into the metadata part of the GrappleStatements. By this means they are available to all GRAPPLE modules via the GEB interface. The integration of the reputation system is a bit more complex, if a certain information provider, in our case the different LMSs

5 Discussion & Summary

One major issue with structured knowledge built in a collaborative manner as given in GRAPPLE is the assurance of high-quality user-generated content, especially because there is normally no central authority giving instructions and controlling the content. The GrappleBroker can only control the syntactical appearance of GrappleStatements, but not the semantic correctness. Apparently, special ways of controlling the quality of newly created data like GRAPPLE statements are needed.

In this deliverable, we discussed the important questions arising in this context:

- How can we ensure the generation and the control of high-quality content for our distributed TEL system?
- How can we motivate members of our learner and LMS community to actively participate, to contribute, and to ensure quality?

In a first step, realized already in the GrappleStatement definition in deliverable D6.1a, we store the identifier of the creator who is responsible for the parts of the user models. By doing so, we are able to find out whether a user or system constantly inserts low quality content.

In a second step, we introduce the community to participate in the quality ensures process of our data. For this reason, we design a five-star rating system for GRAPPLE, with which users can rate the statements of our user models.

In a third step, even more, it is important to make it possible to rate also the raters as proposed by Noy et al. [24]. For that we have introduced and designed a reputation model.

The rewarding of users depending on the amount and the quality (determined via the ratings) of their participation seems to be a promising approach. Users identified as very active and trustworthy participants could be rewarded through GRAPPLE privileges (e.g. being promoted on the projects web site), for instance.

The concrete next steps in WP6.2 will be to implement and integrate the five-star rating system and to implement and integrate the medal-based reputation system.

References

1. Baldonado, M., Chang, C.-C.K., Gravano, L., Paepcke, A.: The Stanford Digital Library Metadata Architecture. *Int. J. Digit. Libr.* 1 (1997) 108–121
2. Bruce, K.B., Cardelli, L., Pierce, B.C.: Comparing Object Encodings. In: Abadi, M., Ito, T. (eds.): *Theoretical Aspects of Computer Software. Lecture Notes in Computer Science*, Vol. 1281. Springer-Verlag, Berlin Heidelberg New York (1997) 415–438
3. van Leeuwen, J. (ed.): *Computer Science Today. Recent Trends and Developments. Lecture Notes in Computer Science*, Vol. 1000. Springer-Verlag, Berlin Heidelberg New York (1995)
4. Michalewicz, Z.: *Genetic Algorithms + Data Structures = Evolution Programs*. 3rd edn. Springer-Verlag, Berlin Heidelberg New York (1996)
5. Allen, Christopher ; Appelcline, Shannon: Collective Choice: Rating, Systems. December 2005. – http://www.lifewithalacrity.com/2005/12/collective_choi.html
6. Allen, Christopher ; Appelcline, Shannon: Systems for Collective Choice. December 2005. – http://www.lifewithalacrity.com/2005/12/systems_for_col.html
7. Berners-Lee, Tim ; Hendler, James ; Lassila, Ora: The Semantic Web. In: *Scientific American* 284 (2001), Mai, Nr. 5, pp. 34–43.
8. Appelcline, Shannon: Collective Choice: More Thoughts About Ratings. December 2006. – http://www.skotos.net/articles/TTnT /TTnT_198.phtml
9. Wolfgang Wahlster: Kap. 1 In: Bullinger, Hans-Jörg (Hrsg.): *Trendbarometer Technik: Visionäre Produkte, Neue Werkstoffe, Fabriken der Zukunft*. München, Wien : Hanser, 2004, pp. 62–63. – http://www.dfki.de/~wahlster/Publications/Semantisches_Web.pdf
10. Anderson, Chris: *The Long Tail - der lange Schwanz: Nischenprodukte statt Massenmarkt ; das Geschäft der Zukunft*. Hanser, March 2007. – ISBN 3446409904
11. eBay: Mehr Transparenz für noch mehr Vertrauen: eBay erweitert sein Bewertungssystem. Mai 2007. – http://presse.ebay.de/news.exe?typ=SU&news_id=101149, date: 14.08.2009
12. Fensel, Dieter ; Wahlster, Wolfgang ; Lieberman, Henry; Hendler, James: *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*. MIT Press, Februar 2003. – ISBN 0262062321

13. Heckmann, Dominikus: Ubiquitous User Modeling. Berlin : Akademische Verlagsgesellschaft Aka GmbH, 2006. – ISBN 3898382974
14. Recktenwald, Pascal: Ein Web 2.0 Bewertungssystem für UbisWorld, Bachelor Thesis, Saarland University. 2007
15. O'Reilly, Tim: What Is Web 2.0. September 2005. – <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>
16. Schwarzkopf, Eric; Heckmann, Dominik; Dengler, Dietmar; Kröner, Alexander: Mining the Structure of Tag Spaces for User Modeling. In: Complete On-Line Proceedings of the Workshop on Data Mining for User Modeling at the 11th International Conference on User Modeling. Corfu, Griechenland, 2007, pp. 63–75
17. Wahlster, Wolfgang ; Dengel, Andreas: Web 3.0: Convergence of Web 2.0 and the Semantic Web. In: Telekom Technology Radar II, June 2006, pp. 1–23
18. Zhang, J. and Cohen, R.: A Trust Model for Sharing Ratings of Information Providers on the Semantic Web. In: Canadian Semantic Web. Semantic Web and Beyond: Computing for Human Experience, Springer, 2006, pp. 45-61.
19. Richardson, M., Agrawal, R., and Domingos, P.: Trust Management for the Semantic Web. In the Proceeding of Second International Semantic Web Conference (ISWC 2003), Sanibel Island, FL, USA, October 20-23, 2003.
20. Brin, S., and Page, L.: The Anatomy of a Large-Scale Hypertextual Web Search Engine. Computer Networks and ISDN Systems, 30 (1-7), pp. 107-117, 1998.
21. Ding, L., Kolari, P., Ganjugunte , S., Finin, T., and Joshi, A.: Modeling and Evaluating Trust Network Inference. In the Proceeding of Seventh International Workshop on Trust in Agent Societies at AAMAS, 2004.
22. Bizer, C., and Oldakowski, R.: Using context- and content-based trust policies on the semantic web. In the Proceedings of the 13th international conference on WorldWide Web - Alternate Track Papers & Posters, WWW 2004, New York, NY, USA, May 17-20, 2004.
23. Christian Hütter, Conny Kühne, Klemens Böhm: Peer Production of Structured Knowledge – an Empirical Study of Ratings and Incentive Mechanisms. Proceedings at CIKM'08, October 26–30, 2008, Napa Valley, California, USA., pp. 827-835
24. Noy, N., Guha, R., and Musen, M. (2005). User Ratings of Ontologies: Who will Rate the Raters? Proceedings of the AAAI Spring Symposium on Knowledge Collection from Volunteer Contributors, 2005