MyChoice & Traffic Lights of Trustworthiness:
Where Epistemology Meets Ethics in Developing Tools for Empowerment and Reflexivity

Judith Simon

Institut Jean Nicod
Ecole Normale Supérieure
29, rue d'Ulm
F-75005 Paris
France
Email: judith.simon@ens.fr
Telephone: +33 (0) 1 44 32 26 46
Fax: +33 (0) 1 44 32 26 99

Abstract: One major trend in software development has been labelled *social software*. A key feature of it is that social networks or trust relationships between users of a system are used for the selection and evaluation of the quality of information provided on the web. Based on such observations, I will examine the relationship between knowledge and trust in the web from an epistemological point of view, focusing on recommender systems to elucidate my claims. I will argue that as soon as knowledge is regarded to be the result of socio-epistemic practices, as is the case on the web, epistemology has to meet ethics and politics in analyzing *and* amending these practices. In the second part of this paper, I will introduce **MyChoice** and traffic lights of trustworthiness as widgets to be included into social software applications. Their goal is to raise epistemological as well as ethical and political awareness among its users about the impact of - possibly implicit or minor - programming decisions on the information they obtain and on epistemic justice. I will conclude by showing how such widgets can enhance critical awareness and reflection among users while empowering them to make informed, context-dependent epistemic choices.
Keywords: MyChoice, ethics, epistemology, trust, recommender systems, epistemic injustice, feminist epistemology, trust propagation, transparency, reflexivity

1. Introduction: Epistemic Trust - Trusting to Know

1.1. Trust in Science

Trust has been a topic in ethics (McLeod 2006, Baier 1992) for quite some time before receiving attention in epistemology, i.e. the philosophical discipline concerned with the process of knowing and criteria for knowledge. In his seminal paper "The role of trust in knowledge" (1991), John Hardwig exposes the function of trust for knowledge creation in science. He states that in classical epistemology, knowledge and trust are conceptualized antithetically: we trust when we do not know; when we know, we do not have to trust. However, "[m]odern knowers", he argues "[…] cannot be independent and self-reliant, not even in their own fields of specialization" (Hardwig 1991, p.693). His analysis departs from the observation that the majority of research is nowadays conducted in teams and he presents two examples of major scientific achievements in physics and mathematics as case studies in support of his claims. Co-operation in science is supposedly needed to overcome time pressure and to handle rising specialization. As a consequence, in scientific co-operations scientists have to trust the competency and the honesty of their colleagues, because they do not only lack the time to perform every subtask of their research on their own, but mostly they also lack the necessary expertise in the respective area of research. Thus, in order to successfully operate in science, scientists need to assess their colleagues not only *epistemically* but also *morally.*

Unfortunately, such assessment is not immune against injustice and the attribution of trustworthiness is frequently influenced by social categories, such as gender, class or race as feminist epistemologists have shown (e.g. Scheman 2001, Alcoff 2001).

Accordingly, the task for a sound epistemological analysis of the attribution of trustworthiness should be twofold. First, reasonable analyses of epistemic trust have to address the empirical question of whether trustworthiness actually is attributed adequately and fairly. However, the assessment of adequacy and fairness depend upon a prior discussion of different possible ideals of rational assessment and attribution of trustworthiness. Secondly, based on these analyses, epistemologists have to develop normative standards of how epistemologically just attributions of trustworthiness can be achieved. Since even an empirically informed epistemology proper can only show *differences* in how epistemic trustworthiness is applied, ethical considerations have to be taken into account to decide upon what epistemic justice is supposed to be. Thus, the development of normative standards and procedures to enforce them depends on a) the epistemological, ethical and political discussion about different possible forms of epistemic justice *and* b) empirical data about forms of epistemic injustice rooted in current epistemic practices. Elizabeth Fricker has offered quite recently an interesting analysis of the ethical dimension of epistemic practices. In her book "Epistemic Injustice" she differentiates two forms of epistemic injustice, testimonial and hermeneutical injustice, and develops normative standards to ensure that these epistemic injustices can be countered (Fricker 2007).

1.2. Trust in Everyday Life and the World Wide Web

If we compare those insights about scientific reasoning to reasoning in everyday life, we soon realize that similar processes of knowing, trusting and attributing trustworthiness take place there as well. Taking into account how many decisions you have to make in your daily life, where you lack the necessary expertise, it should become quite obvious that trusting others for epistemic purposes, trusting others to know is an extremely prevalent phenomenon pervading all possible aspects of our lifes: we ask our doctors about possible illnesses, mechanics about our cars, our insurance brokers about necessary insurances - and we will probably cross-check with other offers, different agents and maybe some independent agencies. The effort that we put into cross-checking will depend on what is at stakes: potential loss of money, potential danger to life or health, etc.

Similar processes occur when you search for information on the World Wide Web. You might have a default to trust resources on the web, but depending on the topic, the stakes, the spare time you have at the moment, you will spend more or less effort on finding supporting or contradiction information elsewhere. For instance, if you are interested in the weather conditions in Vienna at the moment, you might use a search engine. Maybe you will simply trust the first information you get shown on the top of your search results. If you are a bit more suspicious, you might check, whether the information was provided by some agency trustworthy with respect to weather forecast.[i]

1.3. Trust and Recommender Systems

Let’s turn to a specific aspect of trust and social software. There are two very distinct approaches towards concetualizing the relationship between trust and

software in web science. The field that I am not going to deal with is *trust in software*, which focuses especially on security and privacy issues. In the following I will focus exclusively on *trust in other people via software* and use trust-aware recommender systems (RS) to exemplify my claims. RSs in general are systems that suggest new items to users, which he or she might like (e.g. books, music, etc). Classical RS techniques have several shortcomings, including the so-called *cold start problem*, i.e. the difficulty to generate recommendations for new users.

When a new user enters a system, the system does not "know" anything about this new user and this ignorance makes it difficult to generate appropriate recommendations for her. To counteract this problem, traditionally, new users have been asked to rate a few items so that the system can "learn" something about the user in order to provide personalized information on interesting items for her. However, especially in large databases necessary correlations are scarce and thus, this procedure often turns out to be quite ineffective. In consequence, Massa & Bhattachasjee (2004) have developed an algorithm for "Trust-aware Recommender Systems", arguing that this problem can be solved by implementing a notion of trust between users into the system (Massa & Bhattachasjee 2004). The difference between traditional RSs and trust-aware RSs is quite simple: "[w]hile traditional RSs exploit only ratings provided by users about items, Trust-aware Recommender Systems let the user express also trust statements, i.e. their subjective opinions about the usefulness of other users" (Massa & Avesani 2006). This seemingly minor change proves to be highly effective to remedy the cold start problem because "it is able to exploit trust propagation over the trust network by means of a trust metric" (Massa & Avesani 2006).

It becomes especially obvious that the work of Massa and his colleagues is also interesting from an ethical point of view in the case of "controversial users". It is here that the underlying values and possible consequences of different trust metrics become visible. In an empirical study, Massa & Avesani (2007) analyzed data from Epinions.com, a web site, where people can publish reviews about a variety of products and rate reviews of others. The goal of Epinions.com according to its self-description is to help "[…] people make informed buying decisions. It is a […] reliable source for valuable consumer insight, unbiased advice, in-depth product evaluations and personalized recommendations".[ii] Users of Epinions.com can assign binary trust statement to other users, indicating whether they in principle trust or distrust their reviews. This process leads to webs of trust. Controversial users are users that receive diverging trust statements from other users of the community, i.e. many users trust them while many others express their distrust in them. Trust metrics are techniques for answering questions such as "Should I trust this person?" in virtual communities and in this inquiry they tackle the philosophical question of whether trust is warranted in a certain situation.

In their analyses, Massa and Avesani distinguish between global and local trust metrics and define them as follows: "[while] [g]lobal trust metrics assign to a given user a unique trust score, the same independently of the user that is evaluating the other user's trustworthiness […], a local trust metric provides a personalized trust score that depends on the point of view of the evaluating user." (Massa & Avesani 2007, p. 40).

It is of epistemological, ethical and political interest that controversial users are valued very differently in these two different metrics. Local trust metrics explicitly stress and appreciate the individuality and situatedness of every trust statement and state that controversial users by definition do not have a global trust value for the whole community. By contrast, global trust metrics suggest a fictitious consensus between users by calculating an averaged trust value for each user. Through this process, the controversial user is rendered “unreliable” and gets statistically eliminated.

These different types of trust metrics do not only have different underlying assumptions about the value of those users and about deviation from the mean - or norm - more generally. They also have an impact on which information you receive and whose opinions are included. And they might even retroact on cultural and societal values on how to deal with minority views. Averaging out controversial users by means of statistics has a similar effect as other mechanisms of sorting out (Bowker & Star 1999) and silencing: they exclude those from participation that deviate too much from the norms or do not fit in ready-made categories.

1.4. Trust and Wikipedia

The second major source of inspiration for this paper and the development of my own widget in the next section have been the works of Ed Chi and his colleagues of the Socially Augmented Cognition Group at the Palo Alto Research Center (PARC). In the following, I will give a brief description of one of their tools and the related experiments they have conducted concerning the relationship between social transparency, accountability and trustworthiness on the web.

Ed H. Chi, Aniket Kittur, Brian A. Pendleton and Bongwon Suh from PARC have developed a tool called WikiDashboard[iii] "[...] that visualizes the social dynamics and editing patterns of every article and editor of Wikipedia" (Chi, Suh & Kittur 2008). By unveiling the evolvement of articles and the role and amount of conflict, this tool is aimed at raising the social transparency and accountability and by this the trustworthiness of Wikipedia (Suh, Chi, Kittur & Pendleton 2008). The basic idea behind the development of this tool was that the fact that anyone can edit any Wikipedia article does not necessarily have to be regarded as a threat to reliability, but also as a source for it. Suh et al. (2008) argue that it is precisely the possibility to put ideas into discussion, to examine and challenge each others' claims, that is crucial for knowledge generation in science and that similar processes also occur on the web. Reliability and growth of knowledge might thus be advanced by discussions and mutual criticism in combination with practices increasing social transparency, such as attribution, indication of past performance and provision of sources. How relevant the revealing of sources is for an epistemically valid attribution of trust, the assessment of the quality and the identification of the potential bias of information was shown by the WikiScanner (wikiscanner.virgil.gr). By tracking the IP addresses of anonymous editors, this tool unveiled that numerous organizations were editing a diversity of Wikipedia articles anonymously in a way that served their particular interests.

WikiDashboards exist for users and for articles. Thus the editing activity of a specific user or a specific article is visualized and can be used as a cue for assessing the trustworthiness of an article at a given moment or as a proxy for

the trustworthiness of a user. The WikiDashboard embedded within each article of Wikipedia is intended to make to user aware of interesting editing patterns, he might otherwise not notice. Examples would be sudden bursts of edits due to recent events in the case of articles. WikiDashboards on user sites might indicate the user's specific editing habits as well as the range and variety of topics she has contributed to.

Referring to theories of social translucence (Erickson, et al. 2002, quoted from Suh et al. 2008), they consider three things essential for effective communication and collaboration: "[...] making socially significant information visible and salient; supporting awareness of the rules and constraints governing the system; and supporting accountability for actions." (Suh et al. 2008, p. 1039). They conclude that the WikiDashboard might be a useful tool for supporting social translucence and that it might not only benefit the readers in trying to assess the trustworthiness of a Wikipedia article or user, but that it might also have an impact on the behavior of editors and authors.

What I am particularly interested in is the role, visualization can play for rational - and just - attribution of trustworthiness and its benefit for informed decisions about the value of information on the web. It is not some fancy visualizations that I have in mind, but rather the basic *process of rendering things visible.* And this is often a matter of degree and location: *what* is *how* visible *for whom*?

For instance, the discussion and history pages of Wikipedia, which serve as input data for Chi et al.'s (2008) tool are in principle easily accessible to all users of Wikipedia. However, many people do not look at these pages. Maybe because it is too much effort; maybe because they do not want to be bothered; maybe because they do not understand the interface; maybe because they are overwhelmed by the sheer number of revisions or discussion entries; and maybe also because they do not understand how this information might be useful to them. It is especially for this very last fraction of users that the WikiDashboard as well as the application that I have conceptualized might be useful (cf. section 2). However, even some of those from the 'I don't care'-category might be turned into 'Well, now-I'm-interested', if the indication of trustworthiness was a *salient* feature of the website and if it had proven *useful* to them in their first few trials.

So the tool developed at PARC should be regarded not simply as just another web2.0 widget. Rather it should be seen as a tool for empowerment, a tool that raises awareness about the functioning - and possible malfunctioning - of a system that is widely used for information gathering. The skeptic, and especially a Wikipedia skeptic, might still ask why such a tool should be useful. Please allow me to take you on a brief excursus.

1.5. Excursus: A personal note on using Wikipedia and self-observation as tools for reasoning

It is quite often very useful and informative to take a look at one's very own methods and practices of information searching, knowledge acquisition and trust attribution in order to get a clearer view on epistemological problems. When reading epistemological papers, articles on knowledge and trust on the web as well as apocalyptic prognoses about the future of human knowledge, I frequently get the impression that authors assume that they are among the only

ones who reason - at least approximately - accurately. I think there is no reason to be that pessimistic; the epistemic situation of humankind is in all probability not quite as disastrous as critics and worriers often want to make us believe.

Wikipedia for that matter has many critics, among them philosophical, pedagogical, economic and political thinkers and controverters (e.g., Keen 2008, Sanger 2009, Waters 2007). Every once in a while, there are even political cries to ban Wikipedia and often it seems that some critics tend to throw the baby out with the bathwater. Despite the fact that I do not think efforts, such as banning Wikipedia from academia are enforceable, I do also think that these reactions are neither reasonable nor are they based on adequate empirical knowledge about the actual usages of Wikipedia. I use Wikipedia frequently myself and I think it is a very valuable tool. So, the issue should not be *whether* to use it, but rather *how* to use it.

According to some of its critics, Wikipedia is predominantly used with the default of blind trust. I do not agree with this speculation unless someone has provided convincing evidence that indeed the majority of users trusts Wikipedia more blindly than other, similar resources. Rather, I consider it to be paternalistic and arrogant to assume that all other users of Wikipedia are less skilled than I am myself in assessing the quality of information. I am convinced that most users assess - however tentatively and provisionally - the reliability and trustworthiness of the information they receive on a certain topic, at a certain time from Wikipedia (or any other source if you wish). We all have to rely on proxies and indicators of reliability (such as overlap with aspects of the topic we already knew before, etc.).

However, depending on what is at stake, I will raise my epistemic standards of accepting information (cf. also Origgi 2008). If I have a lot to lose, I will probably look for supporting or contradicting information elsewhere and if I am still not convinced I will reject the information or - if possible - suspend judgment. For instance, if I want to have a first idea about what might be causes of a minor skin irritation, I might search Wikipedia or just type the symptoms into Google. It might actually really be the case - as an example of ‘epistemic luck’ (cf. Pritchard 2005) - that I find the correct information and maybe even a solution to relieve me from the itching with this random search. However, except from rare occasions, (e.g. I am alone in the middle of nowhere or the last human being on earth, who has only access to the internet, but not to other people or experts) will I rely solely on the output of Google. And I will even less so be satisfied with this information, when I experience more serious symptoms. I have oversimplified things in this example. I might for instance trust information on an illness that is provided by a medical department of a university more than some random website without any institutional 'voucher'. When it comes to trusting information about effects and especially side effects of medication, however, I might actually end up putting more trust into the newsgroup of a patients' association than on the medical department's website or even the supposed expertise of a pharmaceutical company selling this medication.

Since I assume that I am not a rare example of a rationality and that other people deliberate what to believe as well, there seems to be hope and no need to become desperate about the future of human knowledge and knowing. However, there is of course, room for improvement of epistemic practices as

well as of software. There is a lot of work to do to raise transparency and accountability, to support education, reflection and empowerment. And hope the widget depicted in the next section will make a contribution to this effort.

2. Developing Tools for Empowerment and Reflexivity

The goal of this second part will be to conceptualize tools that accommodate for all the insights gained from the Part 1: the relationship between knowledge and trust in science, everyday-life and the web in particular; the danger of unjust attribution of trustworthiness; the epistemological duty to empower people to make informed decisions about which information they want to trust how much in which context.

The fact that one of the definite books on usability design is entitled "Don't make me think" (Krug 2005) is certainly not very encouraging for an epistemologist trying to improve the epistemological awareness among users by implementing a new widget. However, as noted before, I prefer to stay optimistic about the willingness of people to think. If they can see a clear advantage of a new feature that is easy and intuitive to use and does not cause them much extra hassle, most people I know appreciate epistemic support. Thus, instead of being pessimistic, the lesson learned from the success of this book should rather be that when developing a new tool, simplicity and ease of use as well as perceived usefulness for the user have to be taken into account.

2.1. Should I Stay or Should I Go? Traffic Lights of Trustworthiness

So what are the conclusions that can be drawn from my previous analyses for the development of a tool to improve the information seeking behaviour on the web both epistemologically and ethically? I think that the WikiDashboard is a good example of a tool that empowers its users and enables them to make more informed decisions about the information they are being provided (Suh et al. 2008). However, I also think that the interface of the WikiDashboard is quite complex and that a majority of users might not use it, because they might feel still overwhelmed by the huge amount of data provided even in this aggregated format. So what I am interested in is whether there might be even simpler tools that still raise the epistemological and ethical reflexivity of many differently skilled and interested users.

Let's for instance assume that you see a traffic light either in red, yellow or green on the top of each Wikipedia article. Wouldn't this make you stop and wonder for a second what this signal is trying to tell you? Whether you should treat the information of the article differently depending on whether the lights are on red or green? The reason why I chose the traffic light is that it is an almost universally recognized warning system. Given how widely used Wikipedia is around the globe, the only alternative would be to localize the symbol you intend to use as a cue for trustworthiness depending on the language of Wikipedia. However, if you only take into account in how many countries English is spoken and in how many more the English Wikipedia site is used, e.g. because the version in their own language is much smaller, the potential problems around localization become clearer.[iv]

The traffic lights are simply a much more condensed and intuitive format than the rather complex interface of the WikiDashboard. The traffic light signal would be a dynamic, automatically generated indicator of controversy which preceded the temporal fixation of the article of Wikipedia at the moment you

read it. This indicator can then be used as a proxy for the trustworthiness of the current state of the article[v]. That this indicator is dynamic and automatically generated is crucial for any website as dynamic as Wikipedia. If the traffic lights cannot potentially change with each revision of the article, they would soon be rendered unreliable and thus useless as indicators of trustworthiness of a potentially constantly changing article. But this is exactly what the user is interested in: a quick assessment of the quality of the article they see *at the moment* they see it.[vi]

2.2. MyChoice: Empowering users, raising reflexivity of users and developers

The second tool I have conceptualized in more detail is as simple as the traffic light example, but it is meant to amend searches for recommendations. I have labelled my widget **MyChoice** for two reasons. First it is supposed to indicate that epistemologically and ethically relevant choices are constantly being made in the course of developing software. Thus, by the label **MyChoice**, users should be made aware that these decisions are built into software and have certain effects. Secondly, **MyChoice** is a tool that empowers users to make informed choices of their own where it is normally the programmer who has decided for them.

MyChoice comprises of a dual search-function and a visualization option and can be implemented in principle into all websites using recommender systems[vii]. In the following, I will explain the features of **MyChoice** and show how they correspond to the previous epistemological and ethical considerations of this paper. I have also sketched a user interface to illustrate this tool. For reasons of exemplification, I have used the epinions.com-website as a background into which I have inserted the new features of **MyChoice**. In figure 1 you can see the normal starting page of Epinions.com.

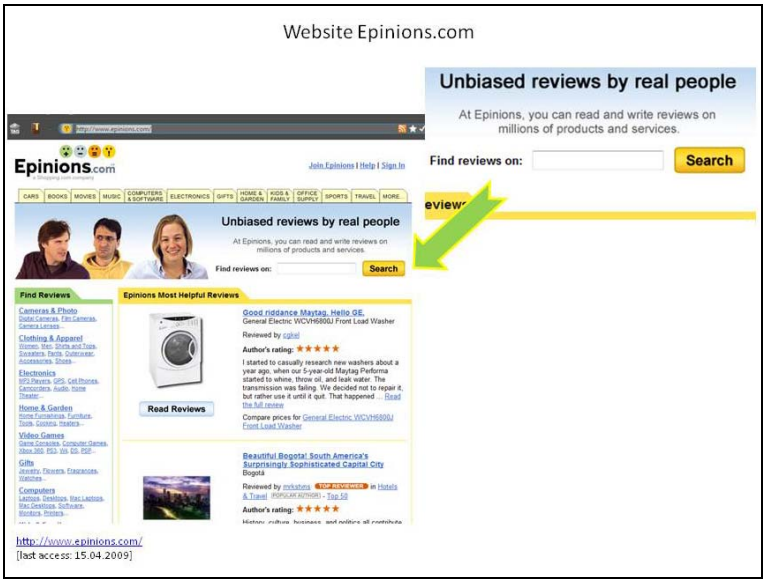

Fig. 1: Screenshot taken from the website Epinions.com [http://www.epinions.com/; Date of access: 16.04.2009]

MyChoice is depicted in figure 2. For the purpose of illustration, I have superimposed the features of **MyChoice** on the background of the Epinions.com starting page.

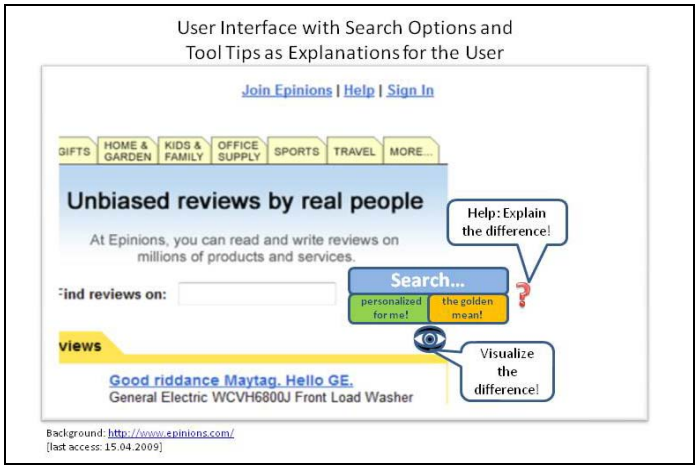

Fig. 2: Sketch of the user interface of **MyChoice** with search options and tool tips as explanations for the users]

Basically, **MyChoice** has two distinct features. One is a dual search button, by which you can chose between two different trust metrics to generate recommendations for your search query. The labels that I have chosen are "Search... personalized for me" versus "Search... the golden mean". They correspond to the local and global trust metrics as described in Massa & Avesani (2007), but other metrics are also possible. You can decide which trust metric you want to set as your default, however, by clicking on one of the two buttons, you can change this for each new query.

The second feature is an "open-eye"-button. If you click on this button, the differences between the two trust metrics are visualized. There are various options on how to visualize this difference. The simplest version would be to just show the different search results next to each user (similar to the function of comparing versions in Wikipedia). However, I would opt for a graphic display that is more visually stimulating and intuitive. It should be possible to display – at least in parts - the different trust propagation patterns. The resulting displays should resemble social networks graphics. Since those graphics are by now embedded into many popular websites, such as Facebook.com, etc., I would assume that many users will be familiar which this type of graphic display. The crucial point will be to show the differences that result from using the local as opposed to the global trust metric. The simplest option might be to show both graphics next to each other. However, it might be more instructive to superimpose the two graphics. Or it would be possible to switch slowly between the two metrics, so that differences become dynamically visible and the users can follow the changes.

In the end, there are different visualization options and it would be necessary to develop different prototypes and test them for their respective usability. It might be even possible to let the user decide upon her preferred mode of visualization. However, I think one should avoid providing too many choices,

because this might eventually alienate some users, especially the less experienced ones.

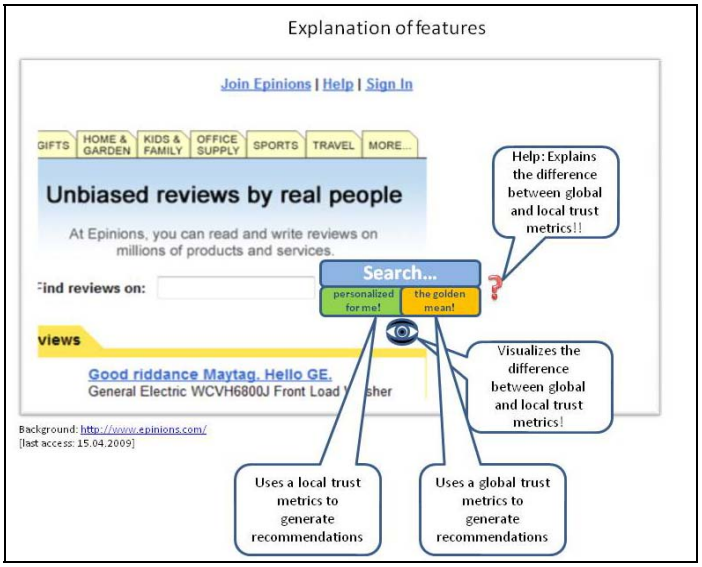

Fig. 3: **MyChoice**: explanations of its features

Another feature that should be included into the graphic displays is the use different colour markers for different types and groups of users. It might be informative to see how different user types change in their positions and their impact depending on the metrics used. The fate of the controversial user in global trust metrics has been depicted before, but there are other possible examples. What happens to the user who always rates each product she evaluates with the highest possible value? What about her picky counterpart? How are these biases accounted for? What happens to the user that very rarely writes reviews, but those he writes are highly valued but many users? What about his prolific neighbour, who writes five reviews a day, one as useless as the other? Highlighting those groups of users who are most affected by changes in trust metrics will render the discriminative consequences of different metrics visible. Moreover, this information about different users and user types might not only be illuminative for the users, but also for developers trying to improve their metrics and algorithms.

For the moment, I can only hint at different visualization options, which have to be developed and tested in the future. The guiding principle should be that the resulting visualizations should be as instructive and informative as possible while remaining usable and intuitive. Thus, over-complex as well as under-complex graphic displays should be avoided.

So what would be the utility of such a tool? I assume that once people start using this tool and become aware of the differences in results that are caused by a simple click on a differently coloured search button, they will start to think about this. They will realize that *different* metrics and algorithms can be used and *are* used in different applications. And that these different metrics have an impact on which information they receive, whose voice is heard and who is automatically silenced or sorted out. The presence of the two search buttons

and the visualization of differences will raise awareness about the huge consequences of seemingly minor programming decisions on information retrieval *and* epistemic justice.

MyChoice is also a tool for empowerment, enabling the user to make more informed decisions about the information she wants to receive. For instance, it would be possible to switch between the global and the local metric depending on the context. Users might opt for the more situated, local option when looking for movie recommendations, but for the rather universalist, global one when they want to learn something basic about computing, statistics or gardening. In fields where users are novices or for certain reasons more interested in mainstream recommendations, they might press the "Golden Mean"-button. And for other questions, they might prefer the "Personalized For Me"-search. In the end this decision is up to the user. But I think that such a simple tool as a button to decide which trust metric to use and have the differences displayed would be extremely valuable. By using **MyChoice**, users will learn about the functioning *and* the consequences of different metrics and algorithms. And this will have positive epistemological and ethical consequences.

Let me summarize the possible benefits of **MyChoice**. Besides enabling the user to decide on a case-to-case basis which metric he prefers depending on the context, **MyChoice** also has several pedagogic functions. People using this system would on the one hand be empowered to decide upon which metric they want to use, i.e. they have a *choice* of which information they want to be displayed. Moreover, people using this system would become much more aware of how implicit assumptions and values of the programmers are inscribed into technology, i.e. they would realize the impact of certain programming decisions of the retrieval of information. This effect would already be triggered by the dual search option, but it would be amended by the visualization of the different search mechanism and differences of retrieved information. Knowing what happens to controversial users once you start averaging people out that deviate too far from the norm, might also raise critical awareness of processes of silencing minorities more generally. Just remember that depending on the context, controversial users might just as well be labelled "pointed users".

MyChoice renders visible the possibly far-reaching consequences of seemingly minor programming decisions implemented into systems. Further, if the use of tools like **MyChoice** becomes more widespread; if people use it, because they like this option and see its benefit for their usage, this would possibly lead to a re-thinking and more awareness among programmers and software developers as well. It should be an epistemological goal to empower people to make rational decisions about how much trust to put on information they receive in general and on the web in particular. And it is an epistemological and ethical duty to raise awareness about epistemic justice and to provide tools that support it. I think **MyChoice** is a starting point for this endeavour and I hope it will ignite more and possibly different reflection about decisions made in software development and their epistemological, ethical and political consequences.

3. Conclusions

I hope to have shown that the notion of epistemic trust is a topic where epistemology has to meet ethics. Assessing the quality of information, deciding

whom to trust and whom to distrust is not limited to information obtained on the WWW. However, it becomes all the more obvious in an environment in which information can be exchanged with high speed over long distances, enormously increasing the amount of interactions with people we do not know personally but whom we have to trust – or decide to distrust – nonetheless. Taking these developments into account, a thorough ethical, epistemological and political analysis of the relationship between trust and knowledge should be indispensable – for philosophy just as much as for web science and software development. Simple widgets, small changes in programming can have huge epistemological, ethical and political consequences. To give inspiration and concrete ideas for the development of normative standards as well as tools that improve epistemological as well as social and political justice on the web and in society more generally should be a central task for contemporary epistemologists. I hope this paper convinced you both of the necessity and feasibility of such an endeavour and that it serves itself as a first small contribution to it.

Acknowledgements

I would like to thank the three anonymous reviewers as well as Noah Holtwiesche for their very encouraging and constructive feedback on earlier drafts of this paper. The research for this paper was enabled by two grants: the ANR2008 grant (Agence Nationale de la Recherche, France) CSOSG-CAHORS for a Project on "Information Evaluation, Analysis, Organization and Ontologies for Intelligence and Security" and a research scholarship from the University of Vienna, Austria for a project on notions of knowledge, sociality and trust in social epistemology and social software (Project number: F-405).

References

Alcoff, L. M. (2001). On Judging Epistemic Credibility: Is Social Identity Relevant? *Engendering Rationalities*. N. Tuana and S. Morgen. Albany, SUNY Press: 53-80.

Baier, A. C. (1992).*Trust.* The Tanner Lectures on Human Values, Princeton University 6-8.3.1991. available at: http://www.tannerlectures.utah.edu/lectures/documents/baier92.pdf. Last access: 15.04.2009

Bowker, G. C. and S. L. Star (1999). *Sorting Things Out: Classification and Its Consequences.* Cambridge, MIT Press.

Chi, E. H., B. Suh, et al. (2008). Providing social transparency through visualizations in Wikipedia. *Social Data Analysis Workshop at CHI 2008*, Florence, Italy.

Fricker, M. (2007). *Epistemic Injustice. Power and the Ethics of Knowing*. Oxford, Oxford University Press.

Hardwig, J. (1991). The role of trust in knowledge. The *Journal of Philosophy 88*(12): 693-708.

Keen, A. (2008). *The Cult of the Amateur*. New York, Doubleday.

Krug, S. (2005). Don't *Make Me Think!: A Common Sense Approach to Web Usability*, New Riders.

Massa,P. & Bhattacharjee, B. (2004). Using Trust in Recommender Systems: an Experimental Analysis. *Proceedings of iTrust2004*, International Conference.

Massa, P. and P. Avesani (2006). Trust-aware Bootstrapping of Recommender Systems. *Proceedings of ECAI*, Riva del Garda, Italy.

Massa, P. and P. Avesani (2007). "Trust metrics on controversial users: balancing between tyranny of the majority and echo chambers." *International Journal on Semantic Web and Information Systems 3(1)*: 39-64.

McLeod, C. (2006). *Trust*. Stanford Encyclopedia of Philosophy. Edward N. Zalta (ed.), URL = < http://plato.stanford.edu/entries/trust/ > [last access: 15.4.2009]

Origgi, G. (2008). "Trust, authority and epistemic responsibility." *Theoria 61*: 35-44.

Pritchard, D. (2005). *Epistemic Luck*. Oxford, Clarendon Press.

Sanger, L. M. (2009). "The Fate of Expertise after Wikipedia." *Episteme 6*(1): 52-73.

Scheman, N. (2001). Epistemology Resuscitated: Objectivity as Trustworthiness. *Engendering Rationalities*. N. Tuana and S. Morgen. Albany, SUNY Press: 23-52.

Suh, B., E. H. Chi, et al. (2008). Lifting the Veil: Improving Accountability and Social Transparency in Wikipedia with WikiDashboard. *26th Annual ACM Conference on Human Factors in Computing Systems*, Florence, Italy, NY: ACM.

Waters, N. L. (2007). "Why You Can't Cite Wikipedia in My Class." *Communications of the ACM 50(9):* 15-17.

Endnotes

[i] An aspect I want to stress here is that trust concerning competency is of course context-specific. You might trust your mechanic on his opinion about the brakes of your car, but probably not about the best treatment for your liver disease.

[ii] Please confer: http://www1.epinions.com/about [last access: 16.4.2009]

[iii] Please confer http://wikidashboard.parc.com/ [last access: 16.4.2009]

[iv] Deciding on a localized indicator or an almost universal one are two options you can choose from. Even though I prefer the universal traffic light in this example, I do not want to suggest that this is always the better option. The best alternative might depend on the context and has to be decided case by case.

[v] I am aware that the level of controversy is only one possible cue for assessing trustworthiness. Thus it would be possible and plausible to aggregate different algorithms and merge them into the ternary symbol of the traffic light.

[vi] Others who are more interested and would check the history and discussion pages anyway are not the target group of such an application and also not those who do not even care about the traffic lights, because they either blindly trust or distrust or randomly trust the information provided on Wikipedia. But I would assume that remaining group of users is quite large, so that the development of such tools has a good chance of getting used.

[vii] This is true at least for those RSs that allow for different trust metrics.

Biography

Judith Simon is a fellow at the Institut Jean Nicod in Paris since April 2009. Previously, she has worked at the Department of Philosophy, University of Vienna for four years and spent two years in technology assessment of biomedicine in Berlin. Her current areas of interest are concepts of knowledge, trust and sociality in epistemology and social software. She

holds a Master in Psychology from the Free University Berlin and has been a visiting scholar at Stanford University's Forum on Contemporary Europe for half a year in 2008.