Towards a Data-Centric Notion of Trust in the Semantic Web
(A Position Statement)

Olaf Hartig

Department of Computer Science
Humboldt-Universität zu Berlin
hartig@informatik.hu-berlin.de

Abstract. Existing research on trust in the Semantic Web extensively studies trustworthiness and trust in the context of active entities such as persons and agents. However, few work exist that focus on the content in the Semantic Web and that study trustworthiness as an information quality criterion. Hence, computer systems that use the trustworthiness of Semantic Web data for filtering or decision making usually apply a very simple assessment approach: each data object is related to some kind of a source for which a trust score can be determined using one of the methods that exist for active entities; this score is then adopted for the trustworthiness of the data object. In this position paper we argue that such a simple notion of trustworthiness for data is insufficient and we propose to adjust the focus of trust research for the Semantic Web from an actor-centric view to a data-centric perspective.

1 Introduction

Today, a large amount of RDF data is published on the Web; large datasets are interlinked; new applications emerge that utilize this data in novel and innovative ways. However, the openness of the Web and the ease to combine Linked Data from different sources creates new challenges. Unreliable data could dominate the result of queries, taint inferred data, affect local knowledge bases, or may have negative or misleading impact on software agents. Hence, questions of reliability and trustworthiness must be addressed.

A great many approaches exist that allow for a calculation of trust values for active entities such as persons, software agents, or peers in a P2P scenario [1]. While several of these approaches can be applied to consider trustworthiness of data providers in the Semantic Web (e.g. [2,3,4]), little has been done considering the data itself. Existing work applies a very simple assessment approach: each data object is related to some kind of a source for which a trust score can be determined using one of the methods that exist for active entities; this score is then adopted as the trustworthiness of the data object. However, simply adopting the trustworthiness of a source for its data does not consider cases where statements have multiple sources, where providers (re)publish data aggregated from the original sources, or where inference engines discover implicit facts from statements of other sources. Hence, source-level approaches are too coarse-grained and, thus, insufficient for the Web of data. Furthermore, our knowledge of the provenance of a data object is not the only criterion that can be applied

to assess the trustworthiness of the object. Other factors such as the correctness of the data or the opinion of another data consumer may affect our decision.

In this paper we argue for making data the central subject of research on trust in the Semantic Web. Therefore, we propose to reconsider the actor-centric trust research for the Semantic Web and conceive trust in the Semantic Web more as an effort that fits in the wider area of information quality (IQ) research. IQ reflects the fitness for use of information [5]. Since this fitness for use may depend on various factors, IQ is a multi-dimensional concept which includes different IQ criteria such as accuracy, completeness, and timeliness [6]. Consequently, we consider the trustworthiness of Linked Data as another such IQ criterion.

Our fundamental understanding of the trustworthiness of data is the subjective belief or disbelief in the truth of the information represented by this data [7]. The decision to believe or to disbelieve is affected by a broad variety of influences. Notice, this complexity renders the actor-centric idea of simply representing the trustworthiness of data by adopting the trust value of an actor as insufficient. We propose to classify the influences in three categories: i) information quality, ii) provenance, and iii) others' opinions. In the remainder of this paper we discuss these categories in more detail (cf. Sections 3 to 5). As a basis for this discussion we review existing approaches that focus on the trustworthiness of data or on content in general (cf. Section 2).

2 Existing Research

In this paper we propose to focus trust research in the Semantic Web on the trustworthiness of data. Similarily, Gil and Artz [8] identify that the majority of existing work on trust "focuses on entity-centered issues such as authentication and reputation and does not take into account the content." Therefore, the authors propose to study *content trust* which "is a trust judgment on a particular piece of information in a given context." As the units of content that are being judged Gil and Artz identify Web resources in general.

To the best of our knowledge, there is still only very few work on content trust in the Semantic Web community. With IWTrust and FilmTrust, two systems have been proposed that consider the trustworthiness of statements during processing tasks and for decisions. IWTrust [9], the trust component of the Inference Web answering engine, understands trust in answers as the trust in sources and in users. Similarily, FilmTrust [10] represents the trustworthiness of movie reviews by a user's trust in the reviewer and in other users' competence to recommend movies. A similar understanding of the trustworthiness of statements published on the Semantic Web has been presented by Rowe and Butters [11]. Their approach adopts a contextual trust value determined for the person who asserted a statement as the trustworthiness of the statement itself. Hence, even if these approaches take the trustworthiness of statements into account they still apply an actor-centric view.

Systems that explicitly focus on trust assessments for statements are TRELLIS and QUATRO Plus. The TRELLIS [12] system assesses the truth of statements by considering their provenance and related statements. Users can rate

information sources and follow the assessments that are presented with the corresponding analysis and the influencing facts. The QUATRO Plus [13] system enables trust assessments for descriptions of Web resources; trust assessment is based on user ratings of these descriptions. Both approaches, however, do not provide a trust model that explicity represents the trustworthiness of content.

Mazzieri [14] and Richardson et al. [15] propose such trust models; they represent content trust for RDF data on the level of RDF statements. Mazzieri introduces fuzzy RDF; a *membership value* associated with each statement represents the likelyhood that the statement belongs to the RDF graph. By equating those membership values with trustworthiness of statements Mazzieri inappropriately mixes two different concepts; trustworthiness is not the same as a fuzzy notion of truth nor is trustworthiness of RDF statements tied to a specific RDF graph. Richardson et al. [15] represent a user's personal belief in a statement by a value in the interval [0,1]. Besides the vague explanation that a "high value means [...] the statement is accurate, credible, and/or relevant" the approach lacks a more formal definition of those values. Thus, what is missing in all cases is a well-founded definition of the meaning of trustworthiness of RDF data.

Another related system is the WIQA framework [5] that permits quality based filtering of data aggregated from the Web. Filtering is based on policies; these policies are constraints that are enforced during query evaluation and that restrict the resultset of queries. Furthermore, the system explaines why data should be trusted, more precisely, why results passed the filters. The WIQA approach does not use explicit scores for IQ criteria. However, missing scores prevent comparisons of the trustworthiness of different pieces of data; moreover, without explicit ratings it is impossible to compare the opinions of multiple data consumers regarding the trustworthiness of the same data. Instead of a filtering approach other work focuses on the ranking of Linked Data [16,17].

Other relevant research is provided by the IQ community where trustworthiness of data is often considered synonymous to believability [6]. Lee et al. [18] decompose believability into three sub-dimensions: trustworthiness of source, reasonableness of data, and temporality of data. Following this differentiation, Prat and Madnick [19] propose a provenance based approach to measure believability by aggregating quality scores for the sub-dimensions. Another provenance based approach has been proposed by Dai et al. [20]. Their main idea is to determine the trustworthiness of a data item by considering source data from which the item has been derived. Furthermore, the approach compares data items to other, similar, but also to conflicting data items. In [21] we present a generic approach for methods that assess IQ of Web data and we apply this approach for the IQ criterion timeliness. Similarily, this generic approach can also be used as the base for a method to assess the trustworthiness of Web data.

3 Influence Category: Provenance

The decision to believe that a data object represents the truth includes considering questions such as the following:

– How was the creation of the data conducted?
– Who or what participated in the creation of the data and how much do I trust this participant?
– To what extend does the input from which the data was produced represents the truth?
– What happened to the data since its creation; how likely is a manipulation?

These questions refer to the provenance of the data object. We understand the provenance of a data object as everything that is related to how the object in its current state came to be. Hence, provenance information about a data object is information about the whole history of this object. This history may start long before the object has been created itself because the provenance of source artifacts used for the creation is also a relevant part of this history. Hence, this history includes multiple actors that participated in various, different roles. All of these actors had a certain influence on the data object and the current state of the object in which it is available to us.

Traditionally, there are two main areas in which researchers study provenance of data: workflows and databases [22]. Research in these areas usually focuses on the creation of data, be it a data product generated by a workflow [23] or the query results created by a database query engine [24]. This focus is reasonable given that workflows and databases are self-contained systems. The Web, in contrast, is a much more open environment. A data object on the Web may have passed through many (virtual) hands before it is finally available in the current application. Hence, the history of a data object includes more aspects than the creation. This additional information is of interest when it comes to assessing the trustworthiness of data objects from the Web as the last of the aforementioned, provenance-related questions illustrates. For this reason we propose a new model for Web data provenance in [25]; this model considers the Web based access to data and the creation of this data equally important. Based on this model we present concepts and tools to integrate provenance information into the Web of data in [26]. This information can then be used to apply provenance based assessment approaches as we introduce in [21].

4 Influence Category: Information Quality

Even if we consider trustworthiness as a criterion of information quality other IQ criteria are likely to affect our trustworthiness assessment. Knowing about a lack of correctness, accuracy, or consistency in the data reduces our belief in the truth of the information represented by that data. Apart from these obvious influences it often depends on the context if a specific criterion is relevant for the trustworthiness assessment. An example is completeness: knowing that a dataset is incomplete may give rise to doubt because missing data may change the information in the dataset. On the other hand, the part of the data that is available might still be trustworthy and, thus, be usable in some application. Similarily, the relevancy of time-related criteria depends on context and application: old data might not be believed to be true anymore in some context; in another application the low currency might not cause a reduced trustworthiness score for

the same data, for instance, when the development of certain data values over time should be analyzed.

As can be seen from these examples, trustworthiness can be understood as a more abstract kind of IQ criteria. Compared to other, independent criteria, trustworthiness comprises multiple other criteria. This characteristic means that scores for other, relevant IQ criteria have to be determined as a prerequisite to assess trustworthiness. These scores, then, have to be weighted during the actual trustworthiness assessment so that the context-dependent relevancy of the corresponding IQ criteria is reflected.

5 Influence Category: Other Opinions

An additional factor that can be used to assess a user's belief in the truth of a data object is the opinion of other consumers of this data. This approach is similar to the idea of determining the trustworthiness of actors using trust assertions in a Web of trust. In addition to the trust assertions about the actor in question, these Web of trust approaches also take the trustworthiness of the actors into account that provided the assertions. This principle must be adopted for opinion based assessment of the trustworthiness of data: the assessment is either based on the opinion of trusted consumers only or opinions must be weigthed by the trust in the corresponding consumer to provide reliable trustworthiness scores for data. Furthermore, we note that the development of trustworthiness assessment approaches that take other consumers' opinion into account can benefit from existing work on recommendation systems.

6 Conclusions

In this paper we argue to apply a data-centric view for further research on trust for the Semantic Web. We understand trustworthiness of Semantic Web data as a criterion of information quality and identify main categories of factors that affect the assessment of this criterion. Even if we discuss these categories separately we suggest that an actual assessment approach should take factors from all categories into account. As an additional requirement for such trustworthiness assessment approaches we note that the assessment system should be able to explain a determined trustworthiness score to end users.

References

1. Artz, D., Gil, Y.: A Survey of Trust in Computer Science and the Semantic Web. Journal of Web Semantics **5**(2) (2007)
2. Golbeck, J., Parsia, B., Hendler, J.A.: Trust Networks on the Semantic Web. In: Proc. of the 7th Int. Workshop on Cooperative Information Agents. (2003)
3. Ziegler, C.N., Lausen, G.: Spreading Activation Models for Trust Propagation. In: Proc. of the Int. Conference on e-Technology, e-Commerce, and e-Service (EEE)

4. Brondsema, D., Schamp, A.: Konfidi: Trust Networks Using PGP and RDF. In: Proc. of the Workshop on Models of Trust for the Web at WWW. (2006)
5. Bizer, C., Cyganiak, R.: Quality-Driven Information Filtering using the WIQA Policy Framework. Journal of Web Semantics **7**(1) (2009)
6. Naumann, F.: Quality-Driven Query Answering for Integrated Information Systems. Springer Verlag (2002)
7. Hartig, O.: Querying Trust in RDF Data with tSPARQL. In: Proc. of the 6th European Semantic Web Conference (ESWC). (2009)
8. Gil, Y., Artz, D.: Towards Content Trust of Web Resources. Journal of Web Semantics **5**(4) (2007)
9. Zaihrayeu, I., da Silva, P.P., McGuinness, D.L.: IWTrust: Improving User Trust in Answers from the Web. In: Proc. of the 3rd International Conference on Trust Management (iTrust). (2005)
10. Golbeck, J., Hendler, J.: FilmTrust: Movie Recommendations using Trust in Web-based Social Networks. In: Proc. of CCNC. (2006)
11. Rowe, M., Butters, J.: Assessing Trust: Contextual Accountability. In: Proc. of the 1st Workshop Trust and Privacy on the Social and Sem. Web at ESWC. (2009)
12. Gil, Y., Ratnakar, V.: Trusting Information Sources One Citizen at a Time. In: Proc. of the 1st International Semantic Web Conference (ISWC). (2002)
13. Archer, P., Ferrari, E., Karkaletsis, V., Konstantopoulos, S., Koukourikos, A., Perego, A.: QUATRO Plus: Quality You Can Trust? In: Proc. of the 1st Workshop on Trust and Privacy on the Social and Semantic Web at ESWC. (2009)
14. Mazzieri, M.: A Fuzzy RDF Semantics to Represent Trust Metadata. In: Proc. of the Italian Workshop on Semantic Web Applications and Perspectives. (2004)
15. Richardson, M., Agrawal, R., Domingos, P.: Trust Management for the Semantic Web. In: Proc. of the 2nd International Semantic Web Conference (ISWC). (2003)
16. Harth, A., Kinsella, S., Decker, S.: Using Naming Authority to Rank Data and Ontologies for Web Search. In: Proc. of the Int. Semantic Web Conference. (2009)
17. Toupikov, N., Umbrich, J., Delbru, R., Hausenblas, M., Tummarello, G.: DING! Dataset Ranking using Formal Descriptions. In: Proc. of the Linked Data on the Web Workshop at WWW. (2009)
18. Lee, Y., Pipino, L., Funk, J., Wang, R.: Journey to Data Quality. MIT Press, Cambridge, MA, USA (2006)
19. Prat, N., Madnick, S.: Measuring Data Believability: A Provenance Approach. In: Proc. of the 41st Hawaii Int. Conference on System Sciences (HICSS). (2008)
20. Dai, C., Lin, D., Bertino, E., Kantarcioglu, M.: An Approach to Evaluate Data Trustworthiness Based on Data Provenance. In: Proc. of the 5th VLDB Workshop on Secure Data Management. (2008)
21. Hartig, O., Zhao, J.: Using Web Data Provenance for Quality Assessment. In: Proc. of the Role of Sem. Web in Provenance Management at ISWC. (2009)
22. Tan, W.C.: Provenance in Databases: Past, Current, and Future. IEEE Data Engineering Bulletin **30**(4) (2007)
23. Davidson, S.B., Boulakia, S.C., Eyal, A., Ludäscher, B., McPhillips, T.M., Bowers, S., Anand, M.K., Freire, J.: Provenance in Scientific Workflow Systems. IEEE Data Engineering Bulletin **30**(4) (2007)
24. Cheney, J., Chiticariu, L., Tan, W.C.: Provenance in Databases: Why, How, and Where. Foundations and Trends in Databases **1**(4) (2009)
25. Hartig, O.: Provenance Information in the Web of Data. In: Proc. of the Linked Data on the Web Workshop at WWW. (2009)
26. Hartig, O., Zhao, J.: Publishing and Consuming Provenance Metadata on the Web of Linked Data. In: Proc. of 3rd Int. Provenance and Annotation Workshop. (2010)