Trustable B2C Markets on the Semantic Web

Robert Tolksdorf,[1] Christian Bizer,[2] Rainer Eckstein[3] and Ralf Heese[3]

[1] Freie Universität Berlin, Institut für Informatik, AG Netzbasierte Informationssysteme,
Takustr. 9, D-14195 Berlin, Germany,
research@robert-tolksdorf.de, *http://www.robert-tolksdorf.de*

[2] Freie Universität Berlin, Institut für Produktion, Wirtschaftsinformatik und OR
Garystrasse 21, D-14195 Berlin, Germany
bizer@wiwiss.fu-berlin.de, *http://www.bizer.de*

[3] Humboldt-Universität zu Berlin, Institut für Informatik, Datenbanken und
Informationssysteme, Unter den Linden 6, D-10099 Berlin, Germany,
{*Rainer.Eckstein* | *rheese*}*@dbis.informatik.hu-berlin.de*

Abstract. Semantic Web technologies will deeply influence the further development of the Internet Economy. A major challenge is, however, to find a practical solution for trust problems arising from their deployment in real-world scenarios. In this paper we develop a concrete application scenario based on Semantic Web technologies for the domain of business to consumer electronic commerce to illustrate these problems. A possible underlying technical architecture can well be realized building on todays standards. The resulting scenario is used for a preliminary analysis of the potential impacts of Semantic Web applications on market participants.

Specific to the scenario we describe trust requirements and outline a Semantic Web trust architecture which fulfils them. The architecture focuses on the allowing subjective and task-specific trust policies as a combination of reputation-, context- and content-based trust mechanisms.

1 Introduction

The past 10 years of Web evolution have established electronic markets and led to the rise and fall of the "new economy". The next 10 years may be characterised by the transformation of the Web from a document publication medium intended for human consumption into a medium for intelligent knowledge exchange [19]. This development is led by the W3C Semantic Web initiative with a joined effort of scientific (MIT, Stanford, ILRT etc.) and business institutions (HP, IBM, Nokia etc.). The basic idea of the Semantic Web is to publish—in addition to classic HTML pages—data directly on the Web. The vision is to use the Web as a global distributed database, which could be queried like a local database today. The W3C Semantic Web architecture stack is defining the reference architecture in the ongoing standardisation process. The standardisation of the basic layers of the architecture is already at a very advanced state. The standardisation of the higher layers (logic, proof, trust) is just starting.

It is still unclear what economic effects these new technologies have on markets and enterprises:

- How do semantic technologies affect market-transparency in electronic markets?
- How does this affect the business models of market participants and intermediaries?
- How does an increased efficiency of information exchange affect industry specific value chains?

Even if the final vision of a global distributed database maintained on a peer-to-peer basis does not become reality in the midterm, Semantic Web technologies provide a huge longterm potential in several application domains [13]:

- *Enterprise Information Discovery and Integration.* Ontology based data integration frameworks will significantly reduce integration costs. Seen from a system architecture perspective, a stronger decoupling of data and applications will become possible. Data could become an independent resource, used by several applications. Semantic Web technologies could also play a major role in the Web Service description, discovery, composition context [3].
- *Knowledge Management.* Ontology based search engines will offer a direct access to information sources in corporate intranets and on the global Web, which will reduce search costs. In addition, adaptive Web sites could enable a dynamic reconfiguration according to user profiles, and make precisely the information needed available in the personally desired format and on the preferred device. Because semantic knowledge networks are based on language independent ontological concepts, it could be even possible to render a large amount of Web content in the user's preferred language.
- *E-Commerce.* The development of XML-based e-commerce standards has led to the same problems EDI initiatives ran into in the past: A wide range of competing standards has been developed and is used by different trading communities. Today suppliers often have to maintain their product catalogs in several XML formats for different market places and customers. Semantic technologies could help to solve these problems by offering a framework for standard mapping and to identify entities like products and traders across different standards. Using RDF's URI identification mechanism, the integration of different product descriptions, ratings, and offers from multiple independent sources by a shopping portal or single user agent might become possible. This would enlarge the data basis of shopping and comparison agents and enhance their usefulness. Seen from the market perspective, this could lead to an increase in market transparency and make the business models of a range of trading intermediaries and market places obsolete.

Before these potentials can be realised, there is still a range of technical and organisational issues to be solved. There has been an agreement on basic data formats and protocols. In order to integrate data from different sources, there has to be consensus about a set of domain ontologies and mappings between them. Mechanisms to decide which data is trustworthy have to be developed and a relevant amount of data has to be published according to the RDF data model and the ontological vocabularies.

The access to distributed, machine readable and semantically annotated information could widely influence the further development of the Internet economy, if these problems can be solved.

Taking up ideas from [20], we use scenarios as a basis to explore the effects of the deployment of Semantic Web technologies. A scenario is defined in the cited work as "a possible set of events that might reasonably take place" in the future. Scenarios should stimulate thinking in order to enable a management of change.

As shown in Figure 1, we approach the potential impact of semantic technologies from the business and the technical viewpoint in order to make predictions about the influence of the new technologies on markets, enterprises and individuals. Who will benefit? Who will lose market positions or will have to change his business model? This analysis of the participant's roles combined with the analysis of technical restrictions allows us to project if a scenario has chances to be realised.

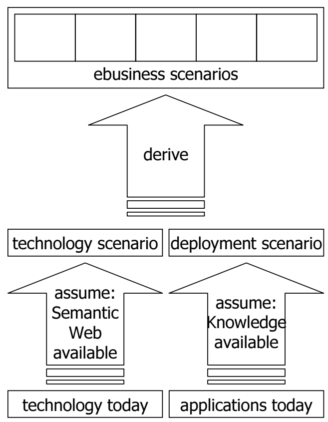

Fig. 1. The scenario based approach

In section 2 we develop a concrete application scenario based on Semantic Web technologies for the domain of business to consumer electronic commerce. First we explain the current state of technological development and highlight the major shortcomings. Afterwards we describe the potentials of Semantic Web technologies to solve these problems. In section 2.6 we analyse the potential impacts on market transparency and the business models of the market participants. In section 3 we describe the scenario-specific trust requirements and outline a Semantic Web trust architecture which fulfils them. The architecture focuses on allowing subjective and task-specific trust policies as a combination of reputation-, context- and content-based trust mechanisms.

2 Business to Consumer Markets on the Semantic Web

Semantic Web technologies will have different impacts depending on the characteristics of a specific markets. We think that they are likely to have the greatest impact if a market fulfils the following set of criteria:

- Suppliers must have similar influence on the market and equal access to it, since in a market without sufficient competition no evolutionary changes can be expected.

- The products and services should already be traded via the Internet, so that customers are already used to this distribution channel.
- The products and services traded must be uniquely and precisely identifiable, because the Semantic Web relies on the unique identification of things about which statements are made.
- There should be no negotiations involved in a usual business process and the products should be traded at fixed prices.

Following these criteria, consumer markets where medium complex, high involvement products, like video cameras, HiFi, computer or other electronic equipment or car accessories are traded, could benefit most from the usage of Semantic Web technologies. Examples of service markets which could benefit are the travel and the job market.

2.1 Current State of Development

In order to identify potentials for improving electronic commerce by the use of Semantic Web technologies, this section describes the current state of development and highlights weaknesses of the technologies currently in use. We use the example of buying car tires to illustrate the different aspects.

The Web sites relevant for this purchasing decision can be classified into the following groups based on assortment strategy, business models and kinds of information provided (sample web sites in parentheses):

- *Online stores* offering a wide range of products, including a shallow assortment of car accessories, like child safety seats, cleaning supplies or roof racks. *(www.quelle.de, www.neckermann.de)*
- *Online stores* with specialised, deep assortments, offering a wide range of tires or just high performance tires as part of a sports car assortment. *(www.tirerack.com, www.discounttire.com)*
- *Electronic marketplaces and online auctions*, where individuals and companies offer new and used tires. *(www.ebay.com, auctions.yahoo.com)*
- *Specialised search services* for car accessories which lead to both online and offline stores.
- Web sites of *manufacturers* describing their products. *(www.goodyear.com, www.dunloptire.com)*
- Web sites of *car magazines* offering product test and comparisons of different tire types. *(www.caranddriver.com, www.europeancarweb.com)*
- *Rating sites* like epinions.com, where consumers report experiences with a product.

These sources together offer a huge amount of information about a specific tire and a wide range of possibilities to purchase it. The problem for the customer is to find all relevant information sources and to compare the information fragments offered. This problem is aggravated by the fact, that most content of the Web sites is generated dynamically from databases and therefore not indexed by search engines like Google. Examining the product descriptions on the various stores we noticed that the structure does not diverge much.

In conclusion, someone searching online for a product is confronted with a wide range of information sources offering separate pieces of information about an item and a variety of shops, marketplaces and actions offering the item itself. The main problems using current Web technologies are:

– *Finding* all relevant information sources and online stores for a specific product.
– *Integrating* all information available on the Web, for comparing products and vendors.

Thus buying decisions today are based only on parts of the relevant information available on the Web.

2.2 Deploying the Semantic Web

Semantic Web technologies address the above two issues directly and improve market transparency. The following sections describe how the technologies could be utilized. After an architectural overview we present two typical use cases which show the benefits of the architecture. Next we describe the ontology development, the information provision and the information usage in more detail.

2.3 Architectural Overview

The Semantic Web technologies offer three important building blocks for our e-commerce scenario [22]:

– The use of *URIs as a global identification mechanism* for products and traders.
– The *RDF data model* together with its XML-based serialisation syntax for the direct publication of data on the Web.
– The *Web Ontology Language (OWL)* for the definition of common terms and concepts needed to understand RDF data.

In an e-commerce scenario based on Semantic Web technologies, the market participants will have the following roles.

A manufacturer of a product will define a URI for his product which identifies it globally, e.g. http://goodyear.com/eagle_f1/gs_emt/p245/45Zr-17. Based on this URI he will publish product descriptions and additional technical information about the product using the RDF data format. A merchant offering the product then only has to publish his price for the product together with shipment and payment details, referring to the product with the URI. The use of a common identification schema and a common data model will allow shopping portals to integrate all information available on the Web about the item.

The portal will not collect links to Web pages or online shops, it will directly collect the data from the sites. This allows the integration and direct comparison of the content of different sites. A car accessories portal would collect all available information about an item from manufacturers, merchants, test and rating services and integrate them into a personalised offer for customers. Buyers then can use one central portal instead of collecting information fragments from different sites. Under the assumptions that all relevant information providers participate and that mappings between different description schemata are possible, large parts of the information in the market would be available in an open and machine processable manner. Figure 2 illustrates the role model for the scenario and gives an overview about the technical architecture.

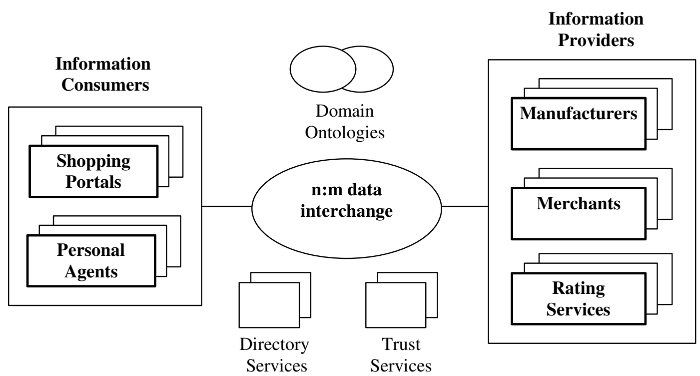

Fig. 2. Architecture Overview

2.4 Ontology Development

The communication between the market participants is based on a set of ontologies which provide shared terms to describe products, traders, shipment and payment options. Using the tire example again, the market ontology will include the concept of a tire and define several properties like size or rain behavior together with the range of possible property values. The market ontology is the result from the merging of existing ontologies by defining mappings between them.

To create the ontologies needed for the market, a lot of standardisation efforts from the EDI community can be reused. Standards like EDIFACT or ebXML define already many of the necessary business concepts and can be converted into ontologies. An overview of the relevant standards can be found in [12].

The existing standards provide the general framework, but have to be complemented with more fine grained domain concepts. Because the manual creation of fine grained ontologies is very costly, different semi-automatic approaches for ontology creation using text mining and language processing tools are being researched. In our example domain, many concepts and relations between them could be extracted from the existing Web sites and product descriptions. An overview of the different approaches and ontology creation tools can be found in [15].

The experiences from the EDI community show that it is impossible to reach agreement on a single standard for a domain. The Semantic Web approaches this problem by allowing the co-existence and co-usage of multiple ontologies for the same domain. To integrate them, mappings amongst similar or equal concepts in separate ontologies can be defined. In perspective, these mappings lead to ontology convergence. For example the technical term “wet traction” is related to the colloquial superconcept “rain behavior”. There are different manual and semi-automatic approaches being researched [1, 24] to generate mappings between different ontologies.

2.5 Information Provision and Usage

To participate in an electronic market based on Semantic Web technologies, the information providers (producers, merchants and rating services) will map their local data models and identification schemata to an ontology used in the market. Collecting these mappings a network of ontologies is established.

Most product and pricing data is stored today in relational databases and can be easily reused on the Semantic Web. There are different approaches to map relational data into RDF [18, 6]. After the mapping, the RDF data is published using standard Web servers like Apache or specialised query interfaces as in [16, 28].

All the published information will be accessed by potential buyers using either personal agents who collect information for them or using semantic shopping portals which provide the access to the information through a standard HTML interface. A shopping portal which is presenting the published data to the customers will:

- Use a directory service to locate information providers for the car accessories domain.
- Use a robot to collect the data from the different providers.
- Use a mapping engine to integrate data published using different ontologies.
- Render the data according to the user’s preferences (level of detail, device, language, trust policy, . . .).
- Provide semantic search services for the customers based on the knowledge included in the ontologies and the available market data [17].

Following the above architecture, a portal would have to store huge amounts of collected RDF data and provide easy access to this data for his customers. An overview of specialised databases and query languages to accomplish this task can be found in [4] and [26].

To allow searches on a semantic level [17], the shopping portal will use a reasoning engine which combines the knowledge included in the ontologies with the instance data collected from the Web. Thus vague, similar, or synonym concepts can be matched against the existing information. For example the vague concept “fast” could be matched to a special tire property.

2.6 Scenario Analysis

Our scenario is based on the new technological foundation defined by Semantic Web standards and raises a couple of interesting economic questions. This section will discuss the possible implications on the business models of the participants and point to open issues, which have to be solved before the technologies can be deployed with success in real markets.

The communication processes between manufacturer, merchant and shopping portal are changed in the Semantic Web scenario. Formerly there were controlled 1:n communication links between them. The manufacturers supplied a known set of merchants with product information. The merchants published this information on their Web sites and forwarded it to the set of marketplaces and auctions in which they participate. Semantic Web technology changes this communication structure. An offer published by a merchant can be used by an unknown number of market places. Formerly controlled 1:n communication links change to an open n:m communication situation.

There are also changes in the kind of information the different parties have to provide. Classically, product description were provided and maintained redundant by every single shop or market place. In the Semantic Web Scenario a merchant would not have to replicate product descriptions to his site, because they are already available from the manufacturer and can be integrated with his offer by the shopping portal.

The application architecture and the possibility of higher market transparency arise new strategic questions for the market participants which will ultimately determine their willingness to participate: Is the architecture in the economic interest of all potential participants? Is it possible to construct an economic win-win situation which would motivate all parties to participate?

For the *customers*, these questions can easily be answered positively. Customers would benefit from the higher market transparency and could make their buying decisions based on a solid, computable information basis. For the *manufacturers* the answers are also positively. They are interested in informing as many potential buyers as possible about the existence of their products. A second advantage for the manufacturers is that they gain more control on the information presented to the customers about their products.

For merchants and shopping portal operators, it is unclear whether the questions can be answered positively. On the one hand it would be much easier to find *merchants* for a specific product if the product is globally identified by an URI. Merchants would also save costs, because publishing data on the Semantic Web is cheaper then maintaining high quality human-readable online stores. But what is likely to weigh more is the fact that merchants are not interested in high market transparency, because in a perfect market their profit margins fall close to zero. Their information advantage would decrease, allowing them to differentiate themselves from other merchants only by additional services.

Shopping portal operators face similar problems. They would all operate on the same information basis. This would allow them to differentiate only in the way they present the information to potential buyers and by additional services like insurances and bonus programs.

Another question is the business model of the *rating services* in the scenario. Today rating services in consumer markets like epinions.com make their profits and are able to pay consumers for their ratings by the revenues they are marking from advertisements on their Web sites and by provisions they are getting from the merchants when a buyer is directed to a shop by the rating service. If the rating data would be published on the Semantic Web and presented to the buyers together with the product information by the shopping portals, the business model of the rating services would have to change to a pure pay-per-view model.

Before Semantic Web technologies can be successfully used in an e-business context and the scenario described above could be realised, numerous technical and organisational obstacles have to be overcome.

- *Missing ontologies*. Most of the RDF ontologies available for the domain of e-commerce are just research prototypes and hardly fulfil the requirements of real electronic markets. There are a lot of standards, like ebXML or EDIFACT, which could be used on semantic networks. But there is no awareness for potentials of Semantic Web technologies in the communities developing these standards.
- *Missing Identification Schemes*. There are also no commonly accepted identification schema for products, companies, places and people within the Semantic Web community. For integrating information from different sources commonly accepted URI schemata or at least URI mapping mechanisms would be necessary. The identification problem has also been addressed in the EDI community for a long time. So existing approaches like EAN numbers for products, Dun and Bradstreet numbers for companies or UNCOCODES for places could be reused. What would have to be achieved is a consensus about a set of identification schema and ways to publicise product and organisational URIs. If the the Semantic Web grows in the future, it could be common practice to have URIs on business paper and product labels, like it is common practice today with URLs.
- *Diffusion strategies*. In order to show the potentials of the Semantic Web and to involve larger communities, seeding application and business cases with clear economic benefits are needed. The Semantic Web is facing a chicken/egg problem today. Because there is hardly any real world RDF data online it is difficult to demonstrate the benefits of the technology. Because the benefits are unclear, major potential users remain in a waiting position and do not publish RDF data online.

3 A Trust Architecture for Semantic Web Markets

Not everything found on the Web is true and the Semantic Web does not change that in any way. Seen from a trust perspective, the participants of the scenario form an open, dynamic network of independent information providers all having

- *different views of the world and different levels of knowledge*. Thus they will publish information which seams to be wrong to an information consumer having a different view or a different level of knowledge.
- *different intentions*. Information providers may intentionally publish wrong information in order to gain economic advantages.

Thus all information published on the Web has to be seen as a *claim* by its author rather than as a fact. The central enabling factor for our scenario is the question, whether it is possible to build a trust architecture which allows customers to decide which *claims* are trustworthy and therefore should be used in buying decisions. In the following section we give an overview of the different types of trust relevant information and different trust mechanisms which can be combined in trust evaluations. Subsequently we describe the scenario-specific trust requirements and outline a trust architecture which could fulfil these requirements.

3.1 Trust Relevant Information

Most trust approaches proposed for the Semantic Web are based on reputational information in the form of explicit information or information source ratings [14, 2, 7]. These approaches imply that a user is willing to provide up-to-date and topic-specific ratings. However, experience from existing electronic markets shows that users tend to show free-riding behaviour in many situations and only provide quality ratings if they get direct compensations, for example offered by sites like epinion.com. In our scenario there is a dense mesh of other, trust relevant information publicly available: Products are described by different sources. In order to be found, participants publish information about their roles in the scenario. Merchants publish their address due to legal requirements. Thus a trust architecture for our scenario should not be based exclusively on explicit ratings but also use all other trust relevant information available.

Figure 3 shows an abstract view on the trust situation in our scenario. All information which could be used in trust evaluations is shaded grey.

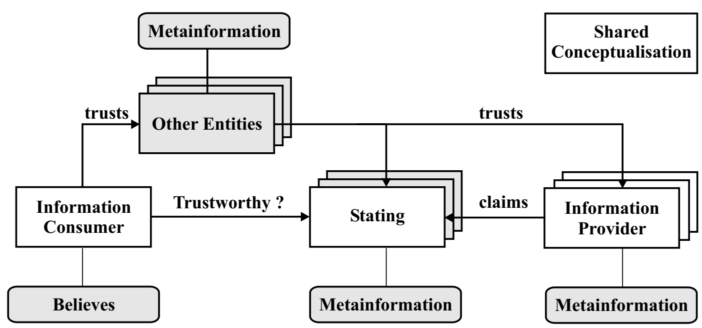

Fig. 3. Abstract View on the Trust Situation

The trust relevant information can be classified into the following three categories:

- *Reputational Metainformation* in the form of information- or information source ratings provided by other users or specialised rating services (Box: Other Entities) or known from past interactions (Box: Believes).
- The *Information Content* itself together with related information about the same topic provided by other information providers (Gray statings in the middle).
- *Context Metainformation* about the circumstances in which a statement has been asserted, e.g. who (Box: Metainformation about the information provider) said what, when and why (Box: Metainformation about the statings). Another kind of context metainformation is information created in the gathering process like the retrieval

date and the retrieval URL or the information whether an attached digital signature is verifiable or not.

We think that it is essential, that the architecture allows users to formulate trust policies based on information from all three categories.

3.2 Trust Policies and Trust Mechanisms

A trust policy is a subjective procedure used for evaluating the trustworthiness of information in a specific situation. In everyday life, we use a wide range of different trust policies. These policies depend on the specific situation, our subjective preferences, our past experiences and the trust relevant information accessible: We might trust Andy on restaurants but not on computers, trust a professor about the topics he is researching, distrust all information provided by a political party that we do not like and buy only from sellers on eBay who have more than 100 positive ratings.

Trust policies combine elements from the following general trust mechanisms:

- *Reputation-based Trust Mechanisms* include rating systems like the one used by eBay [27] and Web-Of-Trust mechanisms [2]. The general trust policy behind these approaches is "Trust someone because you know someone you trust who also trusts him or because you have heard about someone you do not directly know who trusts him." There has been a lot of research about different trust metrics for reputation-based trust mechanisms [23]. The general problem of all approaches of this category is that they require explicit trust statements and that defining such statements and keeping them up-to-data places a big burden on information consumers.
- *Context-based Trust Mechanisms* use metainformation about the information provider or metainformation about the published information for trust evaluations. An example from the first category are role-based trust mechanisms, using the information provider's role or his membership of a specific group for trust decisions. Example policies are: "Prefer product descriptions published by the manufacturer over descriptions published by a vendor" or "Distrust everything a vendor says about its competitor." An example policy using the statement context is "Distrust all product ratings, that are older than a year."
- *Content-based Trust Mechanisms:* These approaches do not use metadata about information but rules and axioms together with the information content itself and related information content published by other information providers [21]. Example policies following this approach are "Believe information which has been stated by at least 5 independent sources." or "Distrust product prices that are more than 50% below the average price."

Context- and content-based trust mechanisms do not require explicit ratings, but rely on the availability of a dense mesh of background information. We argue that in situations like the one in our e-commerce scenario such a mesh of background information will be available and should be used for trust decisions. Thus a trust architecture should allow combinations of all three mechanisms.

3.3 Requirements

A trust architecture, that might be able to handle the trust situation arising from our scenario, should fulfil the following requirements:

- *Openness and Decentralisation:* Trustworthiness should be subjectively evaluated by each application that processes information found on the Web. The architecture should not require central trusted third parties. The principle of openness also implies, that the architecture should be tolerant to incomplete information. The system should not exclude information providers that have not been rated or do not publish trust relevant information in specific way, e.g. sign the information. On the other hand, the system should be able to use all trust relevant information (signatures, context information, related information and ratings) published or generated during the information gathering process (source URL, date).
- *Support for Subjective and Task-Specific Trust Policies:* The users in our scenario have different trust requirements depending on their purchasing intentions. When searching for a low priced product, a buyer might be willing to take a risk in exchange for a better price. Buying other higher priced products, he will require maximal information reliability. Different users have different trust requirements in the same situation and subjective preferences for specific trust mechanisms. As a consequence an architecture should support a mixture of different mechanisms depending on the level of trust required and the context and background information available. The architecture should allow users to formulate subjective and task-specific trust policies.
- *Ability to Explain Trust Decisions:* The key factor for building trust based on context and content is the user's understanding of the information used in trust decisions. Thus an architecture should have the ability to explain its trust decisions and support something like Tim Berners-Lee's "Oh, yeah?"-button [5], meaning that a user can click on every piece of information within an application and get explanations why he should trust the information.
- *Information Ranking:* Instead of eliminating distrusted information, the architecture should use a more Google-like approach and rank information according to its trustworthiness.

3.4 Building Blocks of the Trust Architecture

In this section we describe the main building blocks of our trust architecture. The architecture can be logically divided into four layers:

- *The Information Integration Layer* which handles the aggregation of information from different sources and adds provenance metadata, like source URL and crawling date, to the information. The second task of the integration layer is the verification of information origin. If published information is signed [9] and the public key of the signer is available from a trusted source, information can be marked descending from a verified information provider. In an open environment it can not be expected that all information will be signed. Thus unsigned information should not be rejected by the system but marked descending from a unverified information provider.

– *The Repository Layer* for storing the aggregated information. The trustworthiness of information is not evaluated when it is stored on the repository layer. Trustworthiness is evaluated later when the information is retrieved.
– *The Query and Trust Evaluation Layer* handles the actual trust decisions depending on the user-specific trust policy, the gathered data and the available metainformation about this data.
– *The Application and Explanation Layer* on which the retrieved 'trusted' data is used and which provides the user with functionality to browse through explanations why he should trust this data.

For storing aggregated data, context and reputational information together as an integrated knowledge model we use *Named Graphs* [10], an extension to RDF which allows avoiding the usage of reification when attaching information to graphs. For querying the aggregated data we use *TriQL.P* [8], a query language extending *TriQL*. TriQL is similar to RDQL but uses graph patterns instead of triple patterns for querying named graphs. TriQL.P allows the expression of trust-policies within queries and returns justification trees together with the query results. It supports set operations and different ranking mechanisms like Web-of-Trusts. The result of a *TriQL.P* query are sets of variable bindings together with a *justification tree* for each set of bindings. *Justification trees* contain information about why each set matches the query goals. Applications can use justification trees to explain why the retrieved information fulfils the trust requirements formulated within the query. A justification tree attached to information returned by a query, which uses a reputation-based trust mechanism, includes all known ratings for the selected object. A justification tree for a role-based query includes the relevant authorship information. Similar approaches have been researched within data warehousing community [11]. [25] proposed an alternative web explanation architecture. Their approach focuses on explaining distributed proof traces. Our concept of justification trees focuses on explaining the primary data which has been used in trust decisions based on different trust policies.

4 Conclusion

In this paper we have developed an e-commerce scenario, showing how semantic technologies could influence electronic markets. We described the potentials of the new technologies and the open issues which have to be solved before the scenario can be realised. Afterwards we derived the scenario-specific trust requirements and outlined the building blocks of a Semantic Web trust architecture which fulfilled these requirements. When the Semantic Web begins to have an economic impact it is likely that a lot of intended misinformation will be published. Trust and security mechanisms are part of the W3C Semantic Web layer cake, but have to be matured by more research before they can be used in an open e-commerce context like the one described in this paper.

Our main contributions to the current discussion about a Semantic Web trust architecture are:

– *Context-based and Content-based Trust Mechanisms:* The Semantic Web will be a dense mesh of inter-related statements and a big portion of these statements can be

relevant for trust decisions. Thus a trust architecture the Semantic Web should not be based exclusively on explicit ratings but use all trust relevant information available. In addition to explicit ratings, it should include context metainformation and related information published by other information providers into trust evaluations.
- *Subjective and Task-Specific Trust Policies:* We stress the need that a general trust architecture should not be restricted to a single trust mechanism, but should allow users to formulate subjective and task-specific trust policies as a combination of different mechanisms.

We think that the usage of context- and content-based trust policies within Semantic Web applications presents an interesting path for future research. Our next steps are implementing the trust architecture described above and simulating our e-commerce scenario within this architecture.

Our work on application scenarios and trust mechanism for the Semantic Web is performed within the project *Wissensnetze* funded by the German Ministry of Research BMBF as part of the Berlin Research Centre for the Internet Economy InterVal.

References

1. K. Aberer, P. Cudré-Mauroux, and M. Hauswirth. The Chatty Web: Emergent Semantics Through Gossiping. In *Proceedings of the 12th International World Wide Web Conference, WWW2003*, 2003.
2. R. Agrawal, P. Domingos, and M. Richardson. Trust Management for the Semantic Web. In *Proceedings of the 2nd International Semantic Web Conference, ISWC2003*, 2003.
3. A. Ankolekar et al. DAML-S: Web Service Description for the Semantic Web. In *Proceedings of the 1st International Semantic Web Conference, ISWC2002*, 2002.
4. D. Beckett. Scalability and Storage: Survey of Free Software / Open Source RDF storage systems, 2003. http://www.w3.org/2001/sw/Europe/reports/rdf_scalable_storage_report/.
5. T. Berners-Lee. Cleaning up the User Interface, 2003. http://www.w3.org/DesignIssues/UI.html.
6. C. Bizer. D2R MAP - A Database to RDF Mapping Language. In *Proceedings of the 12th International World Wide Web Conference, WWW2003*, 2003.
7. C. Bizer. Semantic Web Trust and Security Resource Guide, 2003. http://www.wiwiss.fu-berlin.de/suhl/bizer/SWTSGuide.
8. C. Bizer. TriQL.P - A Query Language for Querying Named Graphs Published by Untrustworthy Sources, 2004. http://www.wiwiss.fu-berlin.de/suhl/bizer/TriQLP/Spec/.
9. J. Carroll. Signing RDF Graphs. In *Proceedings of the 2nd International Semantic Web Conference, ISWC2003*, 2003.
10. J. Carroll, C. Bizer, P. Hayes, and P. Strickler. Named Graphs, Provenance and Trust, 2004. Hewlett Packard Labs, Technical Report HPL-2004-57.
11. Y. Cui and J. Widom. Practical Lineage Tracing in Data Warehouses. In *Proceedings of the 16th International Conference on Data Engineering, ICDE2000*, pages 367–378, 2000.
12. Diffuse IST Project. Standards and Specifications List, 2003. http://www.diffuse.org/standards.html.
13. Y. Ding, D. Fensel, and H.-G. Stork. The Semantic Web: from Concept to Percept. *Austrian Artificial Intelligence Journal*, 21(4), 2003.
14. J. Golbeck, B. Parsia, and J. Hendler. Trust Networks on the Semantic Web. In *Proceedings of the 7th International Workshop on Cooperative Intelligent Agents, CIA2003*, 2003.

15. A. Gómez-Pérez et al. A survey on ontology tools, 2002. http://www.aifb.uni-karlsruhe.de/WBS/ysu/publications/OntoWeb_Del_1-3.pdf.
16. R. Guha and R. McCool. TAP: A Semantic Web Platform. *Computer Networks: The International Journal of Computer and Telecommunications Networking*, 42(5):557–577, 2003.
17. R. Guha, R. McCool, and E. Miller. Semantic Search. In *Proceedings of the 12th International World Wide Web Conference, WWW2003*, 2003.
18. S. Handschuh, S. Staab, and R. Volz. On Deep Annotation. In *Proceedings of the 12th International World Wide Web Conference, WWW2003*, 2003.
19. J. Hendler, T. Berners-Lee, and E. Miller. Integrating Applications on the Semantic Web. *Journal of the Institute of Electronic Engineers of Japan*, 122(10):676–680, 2002.
20. M. Jarke, X. T. Bui, and J. M. Carroll. Scenario Management: An Interdisciplinary Approach. *Requirements Engineering Journal*, 3(3–4):155–173, 1998.
21. Joseph M. Reagle Jr. Finding Bacon's Key - Does Google Show How the Semantic Web Could Replace Public Key Infrastructure?, 2002. http://www.w3.org/2002/03/key-free-trust.html.
22. M.-R. Koivunen and E. Miller. W3C Semantic Web Activity. In *Semantic Web Kick-Off in Finland*, pages 27–45, 2001.
23. R. L. Levien. Attack resistant trust metrics, 2002. PhD Thesis, UC Berkeley.
24. A. Maedche, B. Motik, L. Stojanovic, R. Studer, and R. Volz. An Infrastructure for Searching, Reusing and Evolving Distributed Ontologies. In *Proceedings of the 12th International World Wide Web Conference, WWW2003*, 2003.
25. D. McGuinness and P. Pinheiro da Silva. Infrastructure for Web Explanations. In *Proceedings of the 2nd International Semantic Web Conference, ISWC2003*, 2003.
26. L. Miller. Databases, Query, API, Interfaces report on Query languages, 2003. http://www.w3.org/2001/sw/Europe/reports/rdf_ql_comparison_report/.
27. P. Resnick, K. Kuwabara, R. Zeckhauser, and E. Friedman. Reputation Systems. *Communications of the ACM*, 43(12):45–48, dec 2000.
28. A. Seaborne. Joseki - Project Homepage, 2003. http://www.joseki.org.