The TriQL.P Trust Policies Enabled Semantic Web Browser

Christian Bizer, Richard Cyganiak, Tobias Gauss, Oliver Maresch
Freie Universität Berlin, Germany
chris@bizer.de, richard@cyganiak.de, tobias.gauss@web.de, oliver-maresch@gmx.de

Abstract

The TriQL.P browser is a general purpose RDF browser that supports users in exploring RDF datasets aggregated from multiple sources. Information can be filtered using a wide range of user-definable trust policies which can be based on the content or context of the information, ratings of the information or its source, and on digital signatures. To help users understand filtering decisions, the browser can explain why a piece of information fulfils the selected trust policy.

1 Trust Policies for the Semantic Web

The Semantic Web is a dynamic network of independent information providers all having different views, levels of knowledge, and intentions. Thus, it contains claims rather than facts. Before using these claims, an information consumer has to evaluate their trustworthiness and determine the subset which he wants to trust for a specific task.

In everyday life, we use a wide range of policies to evaluate the trustworthiness of information: We might trust Andy on restaurants but not on computers, trust professors on their research field, believe foreign news only when they are reported by several independent sources and buy only from sellers on eBay who have more than 100 positive ratings.

The choice of policy depends on the task, our subjective preferences, past experiences and the trust relevant information available. A trust framework for the Semantic Web can and should support a similarly wide range of policies as used in the offline world [Bizer and Oldakowski, 2004].

Every trust policy employs one or more trust assessment methods which can be classified into three categories:

Rating-Based Methods include systems like the one used by eBay and Web-Of-Trust mechanisms. Most trust architectures proposed for the Semantic Web so far fall into this category. They require information consumers to maintain explicit and topic-specific trust ratings.

Context-Based Methods do not require explicit ratings, but rely on the availability of rich background meta-information, like who said what, when and why. They include methods that use the author's role or group membership for trust decisions. Examples are: 'Prefer product descriptions published by the manufacturer over those published by a vendor' or 'Distrust everything a vendor says about its competitor.'

Content-Based Methods use rules and axioms together with the information content itself. Examples 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.'

2 The TriQL.P Browser

The TriQL.P browser prototype shows how Semantic Web content can be filtered using a wide range of trust policies that combine the methods described above.

It is based on the Piggy Bank extension for the Firefox browser [Huynh *et al.*, 2005]. Piggy Bank extracts Semantic Web content from Web pages as users browse the Web. Extracted information can be sorted and searched through a comfortable user interface and saved into a local repository.

The TriQL.P browser gives users the additional ability to filter the local repository according to a trust policy. The screenshot on the next page shows its user interface. When the user selects a trust policy from the right-hand box, the left-hand view is updated to show only matching information.

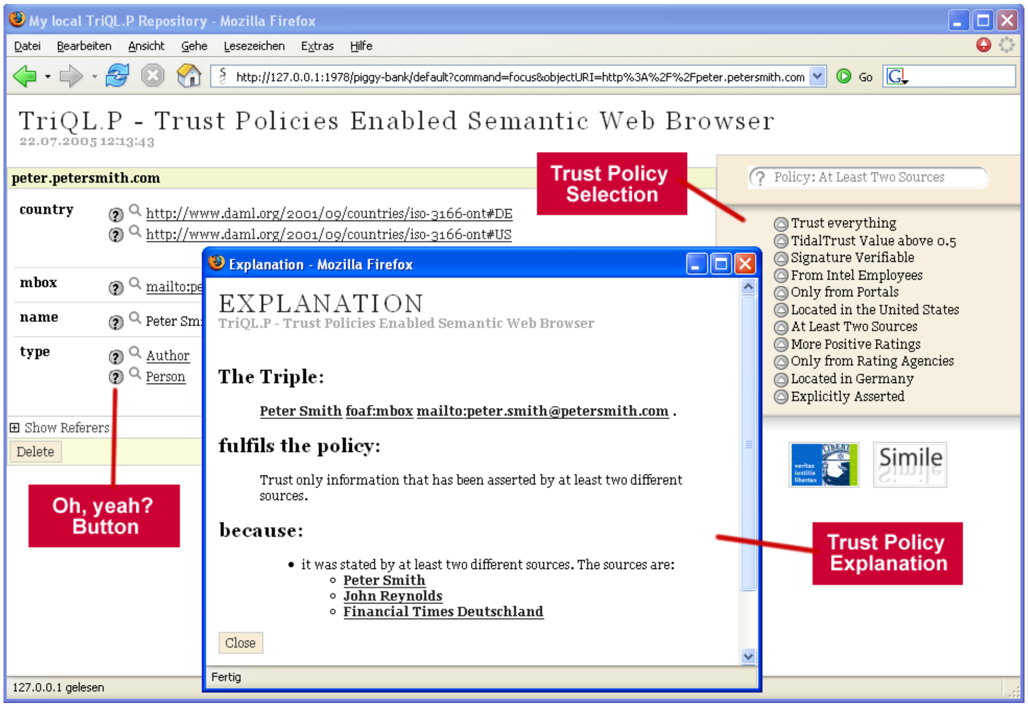
There is an 'Oh, yeah?' button [Berners-Lee, 1997] next to each piece of information. Pressing one of these buttons displays an explanation why the selected policy is fulfilled. The explanation in the screenshot establishes that information about Peter Smith's email address was 'Asserted by at least two different sources,' as required by the policy in effect.

In addition to saving information directly from the Web, users can also import information aggregated by a third party into the local repository using the RDF/XML, TriX and TriG syntaxes.

3 Representing Information

The TriQL.P browser uses Named Graphs [Carroll *et al.*, 2005] as the internal data model. This slight extension of the RDF abstract syntax provides well-defined semantics for the attachment of provenance information to RDF graphs.

Whenever the browser saves information from a web page into the local repository, it creates a new named graph and stores the current timestamp, the URL of the page and the authority (website URL) together with the actual information.


My local TriQL.P Repository - Mozilla Firefox
Datei Bearbeiten Ansicht Gehe Lesezeichen Extras Hilfe
http://127.0.0.1:1978/piggy-bank/default?command=focus&objectURI=http%3A%2F%2Fpeter.petersmith.com
Go
Trust Policy Selection
Oh, yeah? Button
Trust Policy Explanation
Explanation - Mozilla Firefox
Close
Fertig
127.0.0.1 gelesen
Simile

TriQL.P - Trust Policies Enabled Semantic Web Browser
22.07.2005 12:13:43

peter.petersmith.com

country http://www.daml.org/2001/09/countries/iso-3166-ont#DE
http://www.daml.org/2001/09/countries/iso-3166-ont#US

mbox mailto:pe

name Peter Sm

type Author
Person

⊞ Show Referers

Delete

Policy: At Least Two Sources

- Trust everything
- TidalTrust Value above 0.5
- Signature Verifiable
- From Intel Employees
- Only from Portals
- Located in the United States
- At Least Two Sources
- More Positive Ratings
- Only from Rating Agencies
- Located in Germany
- Explicitly Asserted

EXPLANATION
TriQL.P - Trust Policies Enabled Semantic Web Browser

The Triple:

Peter Smith foaf:mbox mailto:peter.smith@petersmith.com .

fulfils the policy:

Trust only information that has been asserted by at least two different sources.

because:

- it was stated by at least two different sources. The sources are:
  - Peter Smith
  - John Reynolds
  - Financial Times Deutschland

4 Expressing and Applying Policies

Our trust policies translate queries against an imagined trusted graph into queries against the entire untrusted repository. Policies are expressed as query templates containing explanation templates. The following policy accepts only information that has been asserted by at least two different sources.

```
:Policy6 a tpl:TrustPolicy;
tpl:policyName "Asserted by at least two
    sources";
tpl:textExplanation "it was stated by at
    least two different sources.
    The sources are:";
tpl:graphPattern [ tpl:pattern
   "(?GRAPH swp:assertedBy ?warrant .
    ?warrant swp:authority ?authority)";
  tpl:textExplanation "@@?authority@@"];
tpl:constraint "COUNT(?authority) >= 2".
```

The graph patterns and constraints are combined with a find(spo) query triple from the browser into a complete TriQL.P query, using the special variables ?GRAPH, ?SUBJ, ?PRED, ?OBJ and ?USER. TriQL.P is a query language similar to but predating SPARQL. In addition to basic graph pattern matching, TriQL.P offers a COUNT() language construct for formulating quantity conditions and METRIC() as an open interface to different rating metrics. Up till now, we have implemented four metrics plug-ins: eBay, TidalTrust, Appleseed and PageRank.

Each 'Oh, yeah?' button corresponds to one trusted RDF triple. The engine generates the corresponding explanation by inserting the variable bindings collected during query execution into the tpl:textExplanation templates included in the policy definition. METRIC() plug-ins generate their own explanations about their internal calculation process.

5 Conclusions

The TriQL.P browser integrates a trust policy framework into a general-purpose Semantic Web browser. We have shown a flexible way to express trust policies, to filter untrusted information based on these policies, and to explain the filtering decisions. We don't rely solely on explicit ratings but also use information context and content for trust assessments.

We hope that our prototype facilitates further thinking about pragmatic ways to incorporate trust policy frameworks into Semantic Web applications, as this essential topic is often ignored by current applications.

The TriQL.P browser is available under BSD licence. More information about the browser, example RDF datasets and example policy suits are found at: http://www.wiwiss.fu-berlin.de/suhl/bizer/TriQLP/browser/

References

[Berners-Lee, 1997] Tim Berners-Lee. Cleaning up the user interface, section - the "oh, yeah?"-button, 1997. http://www.w3.org/DesignIssues/UI.html.

[Bizer and Oldakowski, 2004] Christian Bizer and Radoslaw Oldakowski. Using context- and content-based trust policies on the semantic web. In *13th World Wide Web Conference (Poster)*, 2004.

[Carroll *et al.*, 2005] Jeremy Carroll, Christian Bizer, Patrick Hayes, and Patrick Stickler. Named graphs, provenance and trust. In *14th International World Wide Web Conference*, 2005.

[Huynh *et al.*, 2005] David Huynh, Stefano Mazzocchi, and David Karger. Piggy bank. Submitted to the International Semantic Web Conference, 2005.

6 Demo Description

In our demonstration, we want to show how a trust policy enabled Semantic Web browser can be used in a real-world scenario.

Our running example is an investor who explores information about the financial world. He is subscribed to a financial information service. This service collects business news, postings from business related news groups, stock quotes and company ratings from various information sources and sends the aggregated information to subscribers as a set of Named Graphs. The service includes provenance information about the original sources into the dataset, thus giving our investor the ability to make his own trust decisions. The aggregated dataset is a file in TriG syntax. Our investor imports the file into his TriQL.P browser and starts to explore the news. While browsing, he uses different trust policies to filter the information according to his trust requirements:

6.1 Using Context-based Trust Policies

Our investor has found an article saying that Intel ran into serious problems. In order to decide if he should trust the article, he investigates the article's background. He queries for further information about the author, Peter Smith, using a very liberal trust policy. The investor notices that Peter Smith is claimed to be American and German at the same time. He thus asks the browser to explain the provenance of both pieces of information.

The statement that Peter Smith is German has been asserted by the Financial Times Deutschland. The statement that he is American has been stated by himself. As our investor thinks that people know their own nationality best, he decides to belief that Peter is American.

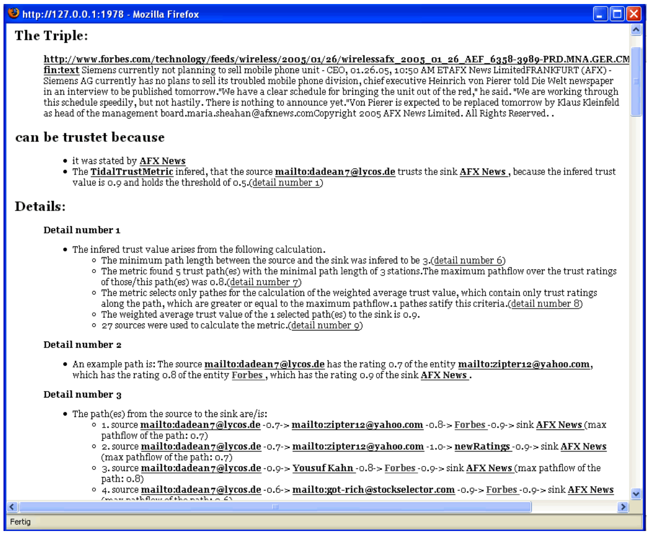
The investor then decides to check if the news have already affected Intel's rating. As the investor prefers American rating agencies like Standard & Poor or Moody's to European ones, he changes the trust policy to "Trust information only from sources in the United States", and thereby removes all other rating agencies from the view.

After looking at the ratings, the investor checks the postings of two financial news groups for related information. As he is especially interested in the opinion of Intel employees about the problem, he decides to filter the postings using the policy "Use only information from information providers that work for Intel. The statement that an information provider works for the company has to be asserted by Intel itself."

6.2 Using Reputation-based Trust Policies

Our investor decides to use a simple eBay-style reputation metric for filtering news articles about Intel. Sources that have more negative than positive ratings are removed from the view. The browser's explanations tell him how many positive and negative ratings an article's source has, and give him a list of the parties that have cast positive and negative votes.

The source of an interesting article has been rated only a few times; not enough to convince our investor. He thus changes the policy to the more sophisticated Tidal Trust reputation metric, which also uses another (hopefully bigger) set of ratings. Using this metric, he gets the more detailed explanation shown in figure 1.



The Triple:

http://www.forbes.com/technology/feeds/wireless/2005/01/26/wirelessafx_2005_01_26_AEF_6358-3989-PRD.MNA.GER.CM
fin:text Siemens currently not planning to sell mobile phone unit - CEO, 01.26.05, 10:50 AM ETAFX News LimitedFRANKFURT (AFX) - Siemens AG currently has no plans to sell its troubled mobile phone division, chief executive Heinrich von Pierer told Die Welt newspaper in an interview to be published tomorrow."We have a clear schedule for bringing the unit out of the red," he said. "We are working through this schedule speedily, but not hastily. There is nothing to announce yet."Von Pierer is expected to be replaced tomorrow by Klaus Kleinfeld as head of the management board.maria.sheahan@afxnews.comCopyright 2005 AFX News Limited. All Rights Reserved. .

can be trustet because

- it was stated by **AFX News**
- The **TidalTrustMetric** infered, that the source **mailto:dadean7@lycos.de** trusts the sink **AFX News** , because the infered trust value is 0.9 and holds the threshold of 0.5.(detail number 1)

Details:

Detail number 1

- The infered trust value arises from the following calculation.
  - The minimum path length between the source and the sink was infered to be 3.(detail number 6)
  - The metric found 5 trust path(es) with the minimal path length of 3 stations.The maximum pathflow over the trust ratings of those/this path(es) was 0.8.(detail number 7)
  - The metric selects only pathes for the calculation of the weighted average trust value, which contain only trust ratings along the path, which are greater or equal to the maximum pathflow.1 pathes satify this criteria.(detail number 8)
  - The weighted average trust value of the 1 selected path(es) to the sink is 0.9.
  - 27 sources were used to calculate the metric.(detail number 9)

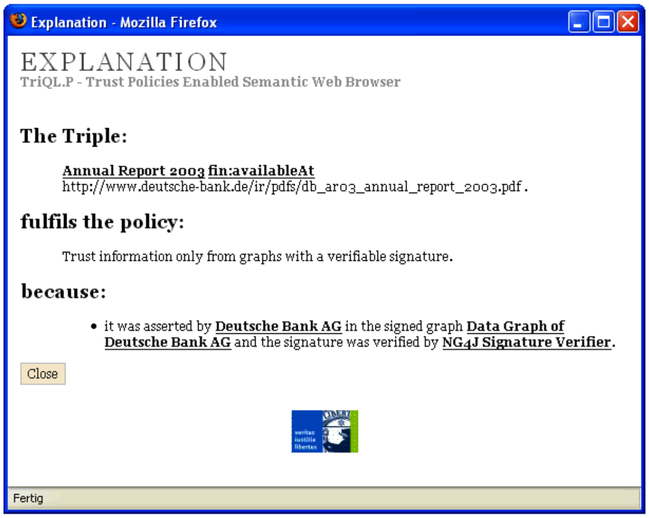
Detail number 2

- An example path is: The source **mailto:dadean7@lycos.de** has the rating 0.7 of the entity **mailto:zipter12@yahoo.com**, which has the rating 0.8 of the entity Forbes , which has the rating 0.9 of the sink **AFX News** .

Detail number 3

- The path(es) from the source to the sink are/is:
  - 1. source **mailto:dadean7@lycos.de** -0.7-> **mailto:zipter12@yahoo.com** -0.8-> Forbes -0.9-> sink **AFX News** (max pathflow of the path: 0.7)
  - 2. source **mailto:dadean7@lycos.de** -0.7-> **mailto:zipter12@yahoo.com** -1.0-> **newRatings** -0.9-> sink **AFX News** (max pathflow of the path: 0.7)
  - 3. source **mailto:dadean7@lycos.de** -0.9-> **Yousuf Kahn** -0.8-> Forbes -0.9-> sink **AFX News** (max pathflow of the path: 0.8)
  - 4. source **mailto:dadean7@lycos.de** -0.6-> **mailto:got-rich@stockselector.com** -0.9-> Forbes -0.9-> sink **AFX News** (max pathflow of the path: 0.6)

Figure 1: Explaining a policy that uses the Tidal Trust metric



EXPLANATION
TriQL.P - Trust Policies Enabled Semantic Web Browser

The Triple:

Annual Report 2003 **fin:availableAt**
http://www.deutsche-bank.de/ir/pdfs/db_ar03_annual_report_2003.pdf .

fulfils the policy:

Trust information only from graphs with a verifiable signature.

because:

- it was asserted by **Deutsche Bank AG** in the signed graph **Data Graph of Deutsche Bank AG** and the signature was verified by **NG4J Signature Verifier**.

Figure 2: Explanation for the signature verification process

6.3 Using Digital Signatures

Some of the graphs in the dataset might be signed by the original information providers. Our investor verifies the signatures against his set of trusted digital certificates. He uses the policy "Trust information only from graphs with a verifiable signature" to verify the download URL of an annual report. The explanation in figure 2 shows that the report has indeed been published by the company.

6.4 Summary

We hope that the demonstration shows how a trust architecture for the Semantic Web supporting a wide range of different trust policies could be used to establish the trustworthiness of claims and thus support users in decision-making.

An extended version of the scenario, along with screenshots of many explanations, are available on the TriQL.P website at http://www.wiwiss.fu-berlin.de/suhl/bizer/TriQLP/browser/ .