Using the Semantic Web to generate Trust Indicators

Marc Sel

PwC - Enterprise Advisory - Belgium
marc.sel@be.pwc.com

Abstract

This article discusses how Semantic Web technology such as RDF and SPARQL can define and compute trust indicators related to Trust Service Providers (TSPs) using independent public domain information. Such Trust Indicators can complement the purely cryptographic trust evaluations that are common today.

1 Introduction

These Trust Indicators are set in the context of the new Regulation of the European Commission on electronic identification and trust services for electronic transactions in the internal market (COM 2012 238). The new regulation introduces amongst others a more stringent approach for supervision; however it still focuses on trust which is mainly based on cryptographic verification. While such verification should obviously be mandatory, it does not take into account relevant, independent and freely available information from public domain sources such as central banks.

A short discussion on trust is provided, followed by an introduction to the Semantic Web components used (RDF and SPARQL). Then the public domain information relied on is presented, such as the various EU and Member State Trust lists, as well as other sources. Then it is demonstrated how such information can be transformed in a semantic model. Subsequently it is shown how different triples can be merged. Finally, it is demonstrated how 'semantic' trust indicators can be defined, and computed using SPARQL queries. Such trust indicators can be used by Supervisors, users of Trust Services and Relying Parties to evaluate the overall trustworthiness of such service providers.

2 Trust

Trust can be seen as a factor that contributes to the taking of a decision. In cases as different as the ordering from a website, making a payment to a specific counterparty, or starting a medical treatment, trust will play a role in the decision taking. Such trust is based on elements as the existence (or lack) of positive (negative) outcomes related to similar decisions taken in the past. We may have obtained these outcomes ourselves, or we may have learned about them from other sources we rely on. Other elements include the extent to which some form of transaction-reversal is possible. Furthermore, the product or service provider may offer some form of guarantee or refund. Finally, a regulator may force the provider to take liability.

N. Pohlmann, H. Reimer, W. Schneider (Editors): Securing Electronic Business Processes, Springer Vieweg (2014), 1-14

For the remainder of this article, we will use trust as a factor in a decision process, unless explicitly indicated otherwise. A trust indicator is information that is relevant in a decision process, and that contributes to the taking of the decision in a positive or negative way.

3 Basic Semantic Web concepts

According to the Encyclopaedia Britannica, semantics is the philosophical and scientific study of meaning in natural and artificial languages. Semantics refers to the meaning of languages (and artefacts created through them), as opposed to their form (syntax).

The term Semantic Web was proposed by Tim Berners-Lee[1]. We now briefly introduce its key concepts according to the loosely defined Semantic Web Stack [SWS1]. It can be observed that cryptography is situated across the different layers of the stack, since it offers different functionality at different layers.

As these concepts are well documented on the W3C website and in literature such as [SWB1], we limit ourselves to a short overview, introducing the concepts required for understanding of the rest of this article. For a more detailed description, please refer to the W3C website, or to [FSWT]. At the top of the stack, user interfaces and applications are situated. In descending order, the next layers are trust, proof, and unifying logic. These layers contain technologies that are not yet standardized or contain just ideas that should be implemented in order to realize Semantic Web. In the middle of the stack, in order of increasing reasoning power, we encounter:

- Vocabulary: a set of URIs to which a particular meaning is attached;
- Taxonomy: a hierarchical classification of a vocabulary (i.e., subClassOf relations).
- Ontology: a more complete set of inference rules.

The information expressing the above can be formalised in schemes and triples, and saved in a triplestore. It can be processed for example with the query language SPARQL (a recursive acronym for SPARQL Protocol and RDF Query Language) which is, as indicated by its name, an RDF query language. It allows a query to consist of triple patterns, conjunctions, disjunctions, and optional patterns. Controlled vocabulary schemes mandate the use of predefined, authorised terms that have been preselected by the designer of the vocabulary, in contrast to natural language vocabularies, where there is no restriction on the vocabulary. Rules allow encoding knowledge, and inference engines can combine such rules with information and draw conclusions.

The Dublin Core Metadata Initiative (DCMI) is a membership organisation that supports shared innovation in metadata design and best practices. The term “Core” is used because its elements are broad and generic, usable for describing a wide range of resources. The DCMI publishes the DCMI Metadata Terms, represented in RDF schema language. DCMI identifies metadata terms with Uniform Resource Identifiers (URIs) in different DCMI namespace URIs such as terms, types, etcetera. These include the classic Dublin Core Metadata Element Set (DC MES), the DCMI Type Vocabulary, and resource classes used as formal domains and ranges. The DC MES [DCterms] is the original set of 15 classic metadata terms for use in resource description. They are considered a de-facto standard for description of properties of resources in the Semantic Web. They are endorsed in the IETF RFC 5013 and ISO 15836:2009 documents.

1 Tim Berners-Lee, homepage http://www.w3.org/People/Berners-Lee/

At the bottom of the stack we find RDF, XML, URI and Unicode. The RDF data model is graph-oriented. An RDF document describes a directed graph, i.e. a set of nodes that are linked by directed edges. Both nodes and edges are labelled with identifiers (nodes can also be literals). URIs (strictly speaking URIRefs in RDF 1.0 and IRIs in RDF 1.1) are used as identifiers.

In contrast, XML documents are structured in tree-mode. Trees are suitable for structuring information which is often hierarchical in nature, and allow efficient processing of such information. However, RDF was conceived to describe relationships between objects (RDF refers to these as resources). RDF allows to make subject-predicate-object statements ("triples") about web resources e.g. "the sky/has the colour/blue". The RDF data model is similar to conceptual modelling approaches such as entity-relationship or class diagrams. It is based upon the idea of making statements about resources (in particular web resources) in the form of subject-predicate-object expressions. These expressions are known as triples in RDF terminology. The subject denotes the resource, and the predicate denotes traits or aspects of the resource and expresses a relationship between the subject and the object.

For example, one way to represent the notion "The sky has the colour blue" in RDF is as the triple: a subject denoting "the sky", a predicate denoting "has the colour", and an object denoting "blue". We do not address the other concepts (XML, URI and Unicode) as they are generally well-known. For a more detailed overview of Semantic Web standards, see [SWS2].

4 Problem statement

We situate our trust indicators in the context of a market (demand/supply), where Business Service Consumers call upon services from Business Service Providers. We use the term 'business' to refer to public and private sector organisations. Such services can include anything that can be delivered in electronic form. Trust services assist in the delivery of such services. We use the term 'trust services' to refer to any services created to establish, maintain or improve trust between different actors engaging in an electronic transaction. A more formal definition can be found e.g. in Menezes, van Oorschot and Vanstone [HAC], or COM (2012) 238 [EIDAS2012].

Qualifying information in an electronic form or a service as 'trusted' is not trivial since:

- Many contributing factors are under control of different stakeholders. Examples include the different components of client and server hardware and software. Software can be in-house developed, obtained from a vendor or from the open source community. This may include also the implementations of algorithms, key stores and certificates. Different types of assurance may be given at component, system or service level, by parties as diverse as an independent laboratory, an auditor or a regulator.
- Many terms used to describe such a qualification are overloaded (i.e. bear multiple meanings that are not necessarily identical or similar). This is illustrated in the previous discussion of the meaning of trust, or in the many competing interpretations of the term 'identity' and 'electronic identity'.

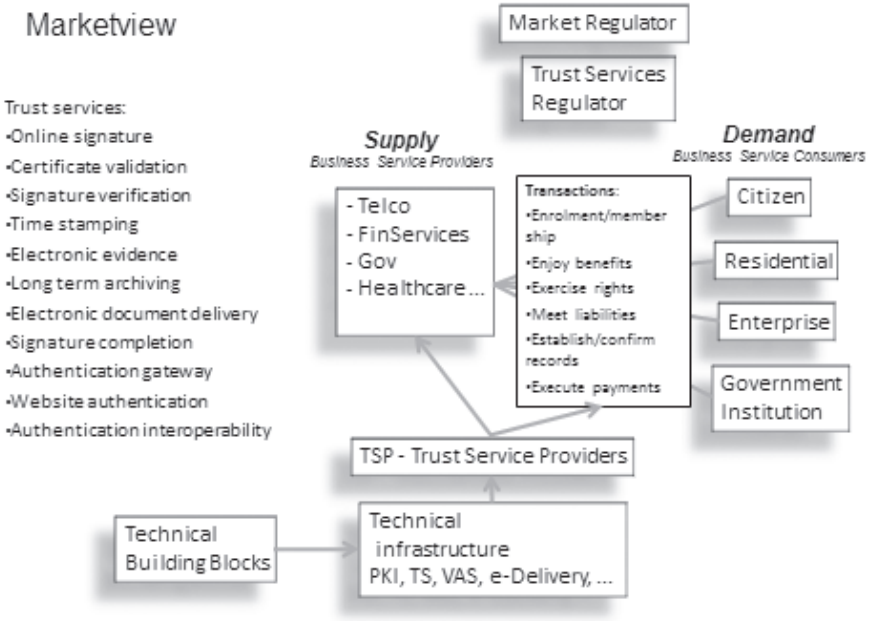

Fig 1: Market view

There are many services in use in the field of electronic identity, authentication and trust, and their number will most likely only further increase. Unfortunately, the vocabularies used by these services are neither necessarily transparent nor clear in the actual semantics of the outcome of the invocation of such service. Hence multiple challenges remain, including:

- How can an identity or another piece of electronic information be described in a way that is semantically unambiguous in a set of Use Cases, seen from the perspectives of different participants? Use Cases can be as diverse as:
  - Enrolment for a service or membership, demonstrating eligibility;
  - Participation in electronic business transactions to exercise rights, meet liabilities, establish or confirm records and any other business logic;
  - Execution of payments related to the services consumed.
- Participants include Business Service Provider, Citizen, Enterprise, Government Institution, as well as Trust Service Providers and Regulators.
- Which attributes are required? How can they be modelled?
- How can the different sources of these attributes be trusted?
- What attributes or other relevant information can be logically deduced, how, and
- Why would such information be trusted, and how can it be used in trust decisions?

While the 'value adding' transactions in the above Use Cases are often conducted between two parties (Business Service Consumer/Business Service Supplier), it is common to find a network of related actors that help preparing, executing or analysing these transactions. Related actors can include Network Provider, Identity Provider, Enterprise Register, Attribute Provider, Certifying

Party, Regulator, Supervisor and many more. Prior to, during and after the execution of a transaction, each actor may be alternatively acting in a claiming, asserting and relying role.

While there are many sector and transaction specific models to establish the appropriate trust between those actors, the semantics used are not necessarily transparent. Neither is there a model that can easily be shared by all actors alike. As a consequence, it is difficult for actors to understand what security is actually covered, and to trust one another. This is even more the case in a volatile setting (e.g. a mobile or cloud-based eco-system). Furthermore, both for electronic identity/authentication and for trust services, the relationship between a service that is invoked and the legal effect that is obtained is not well designed; this makes establishing a network of trusted services across the EU (e.g. a trusted STORK2-backbone) both more difficult and less transparent. The Supervision landscape is complex, within a nation, at the level of EU cross-border transaction, and even more so at a global level. Regulatory improvements envisioned e.g. by the eIDAS proposal [EIDAS2012] do currently not include an approach to close the semantic gaps between the stakeholders (relying parties, subscribers, service providers, regulators, supervisors, ...). In the context of this broad range of challenges, we limit ourselves to trust indicators only.

4.1 Description of the proposed solution

A proposal for a vocabulary should build on existing concepts where relevant, but should offer new possibilities for decision support in the value chain of trust services.

We propose a domain model, from which we derive selected information and map it into a semantic model. The initial semantic model is enriched by linking data to it. On the enriched model, we calculate a trust indicator.

4.1.1 Domain model

We start by creating the following informal 'real world' domain model:

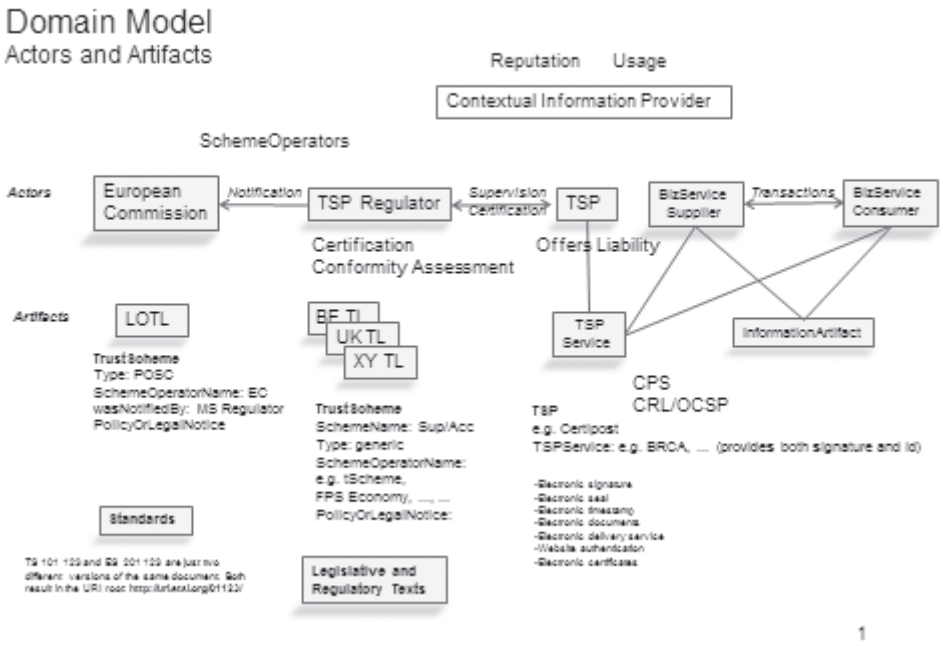

Fig 2: Informal domain model of Actors and Artefacts

The domain model includes actors such as the European Commission, Trust Service Provider regulators (per Member State), TSPs, as well as Suppliers and Consumers of Business Services. Furthermore, there are many Contextual Information Providers such as e.g. Central Banks, the DNSSEC Trust Anchor, the CAB Forum, Conformity Assessment Bodies, and information provided by accountants and auditors.

The domain model also contains artefacts. We now briefly describe the main artefacts, the LOTL and TLs. The List of Trusted Lists (LOTL) is an EU-wide mechanism for electronic trust, publicly on-line available in signed XML and PDF formats. It identifies the TLSO (the Trusted List Scheme Operator, which is the European Commission), and provides information about the scheme. It makes use of various existing namespaces such as *xmlns:tsl*. The LOTL contains both technical and legal scheme information as well as references to Member State Trust Lists (TLs). These TL references are pointers to the TLSOs (national regulatory authorities), competent for electronic trust.

The domain model also contains relationships between actors and artefacts.

Technically, the XML version of the LOTL starts with the *tsl:TrustServiceStatusList* element. Next, technical information is provided such as version, type and scheme operator name (the European Commission). Furthermore, the element *tsl:PointersToOtherTSL* contains a set of *tsl:OtherTSLPointer* elements, which contains amongst much other information, the various *tsl:TSLLocation* elements.

The structure of a Member State's TL is similar to that of the LOTL. However, the TSLO is a national regulatory authority, rather than the European Commission. We find *tsl:Distribution-Points*, containing the *uri* of the Member State TL, both in human readable and XML formats. We also find the *tsl:TrustServiceProviderList* element which contains the *tsl:TrustServiceProvider* element. This points to the actual national root CA's and in Belgium also to e.g. SWIFT's root CA.

1.1.1 Common Vocabulary

We propose elements for a Common Vocabulary for Trust Services, based on Actors and Artefacts. This Common Vocabulary represents selected information from the 'real world', in a format that can be processed by Semantic Web Technologies. We use an RDF graph rather than a more traditional Entity-Relationship or Object-Oriented model since this allows:

- Greater flexibility to capture the various types of information that different stakeholders may find relevant, and which is available in different forms in the public domain;
- Possibilities to deduct conclusions based on the use of queries via languages such as SPARQL, and via 'Reasoners'. Such Reasoners use built-in inference capabilities over the information present in the graph model.

Much relevant information is published today in XML, where relations are implicit and hierarchical. Such XML can be electronically signed. To construct a richer information set by linking more data to one-another, or to calculate or reason over this information using Semantic Web technologies, we need to map this information in a Common Vocabulary, expressed in RDF.

Trust in an Actor can come from contributors such as past use of the Actor's services without problems (own usage), reputation of the Actor (opinion of other parties – reputation), certifications of the Actor (quality indicators by independent parties), the Actor's capability to handle his liability in case something goes wrong (liability).

Trust in an Artefact can come from contributors such as an indication that reliance on the Artefact is supported by one or more Actors, stand-alone safeguards that demonstrate required qualities such as integrity or authenticity.

We create our Common Vocabulary for Trust Service by modelling properties and relationships both for Actors and Artefacts that will allow quering/inferencing Trust Indicators. We limit ourselves to trust indicators only.

4.2 Technical deployment model

In the current implementation, a Trust Indicator is simply a variable that can take an integer value. To make a Trust Indicator available to applications, we create it through the layers depicted below. At the lowest layer, data acquisition takes place from machine readable public domain information. We then apply XSL transformation to create RDF triples from this data. We further elaborate these triples, so they can be processed by a SPARQL endpoint. A SPARQL query calculates a Trust Indicator for a particular Trust Service Provider. The Trust Indicator can be used in transactions between a Supplier and a Consumer from Business Services that rely on the TSP.

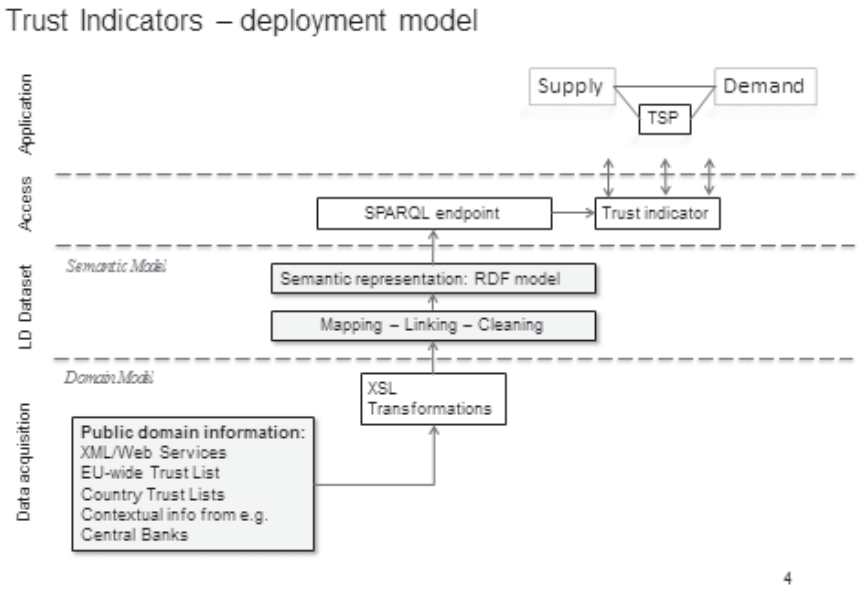

Fig. 3: Deployment model to derive a Trust Indicator

5 Design of the core vocabulary and RDF model

5.1 Reuse of existing terms

When transforming information into an RDF model, the first decision required is whether to reuse existing terms, or to invent new ones.

5.1.1 XML Schema

RDF by default re-uses XML Schema datatypes where possible (prefix *xsd*).

5.1.2 Dublin Core

The DCMI already proposes a wide range of terms to choose from. However these are not particularly suitable for describing information or knowledge in the field of eIDAS. The term eIDAS is used to refer to the combination of electronic identification, authentication, signatures, and related trust services.

Whatever terms the triple-based RDF model contains, these should be encoded in XML Node and Property elements. The LOTL itself can be represented as a Node element, using an attribute as URI reference to its location. For the representation of the LOTL components that will form part of the RDF model, DCMI terms can be used or new terms can be proposed. Candidates from the DCMI Metadata Terms [DCterms] include *isReferencedBy, conformsTo, hasPart, description, references, publisher* and many more.

The term *dcterms:publisher* is described as an entity responsible for making the resource available. Examples of a *dcterms:publisher* include a person, an organization, or a service. Hence *dcterms:-publisher* could potentially be used to indicate the LOTL is the publisher for the references to the Member State TLSOs. However, mere publication does not reflect the implied trust for which the LOTL was established, so we drop this term as a candidate.

The term *dcterms:isReferencedBy* is described as a related resource that references, cites or otherwise points to the described resource. Hence *dcterms:isReferencedBy* could potentially be used by resources to link back to the LOTL or other TLs. However, such linking referencing is not strong enough to reflect the implied trust for which the LOTL was established, so we drop this term as a candidate. The term *dcterms:conformsTo* is described as a reference to an established standard. Hence *dcterms:conformsTo* could be used to indicate the LOTL conforms to ISO or ETSI standards. So we keep this term as a candidate to describe a relevant feature of the LOTL and other TLs.

The term *dcterms:hasPart* is described as a related resource that is included either physically or logically in the described resource. Hence *dcterms:hasPart* could be used to indicate the LOTL contains other TLs as parts. So we keep this term as a candidate to describe a relevant feature of the LOTL and other TLs.

5.2 Transformation of the LOTL into RDF

The LOTL XML version from https://ec.europa.eu/digital-agenda/en/eu-trusted-lists-certification-service-providers was used. The LOTL has 2 major elements:

- The *tsl:SchemeInformation*, indicating that the TLSO was the European Commission, and containing the pointers to the various Member State TLSOs,
- The *ds:Signature* containing the signature of the TLSO over the information provided

The XSL transformation selects *tsl:SchemeInformation* and extracts *tsl:TSLLocation* elements by visiting all *tsl:PointersToOtherTSL* elements and their children. The transformation produces an RDF graph as output, which is serialised into RDF/XML. Our transformation uses an RDF-node element to represent the LOTL. This node element has multiple property elements that contain the TL locations. The node element also has an attribute that refers to the LOTL's location.

We use two XSL transformations on the LOTL to extract relevant properties. We call these transformations *ArtLotlPt01* and *ArtLotlPt02*. The information extracted and transformed provides information about the *SchemeName*, the *SchemeOperatorName*, the *SchemeTerritory* and the Scheme's *LegalNotice*. All this information relates to the single EU-wide Scheme, officialised through the EU List Of Trusted Lists.

They yield RDF/XML, which we merge into a single graph, *RDFMM02.rdf*. The result can be validated with the W3C RDF validator [W3CRV] and graphically represented. A small fraction of the graph is displayed below.

Fig 4: Partial graph of RDFMM02, the result of two transformations on the LOTL.

The transformations are available at http://www.marcsel.eu/SemWebSrc/. The resulting RDF/XML and graph are available at http://www.marcsel.eu/3Store/, as RDFMM02.rdf and RDFMM02.png.

As a consequence of the transformations (*ArtLotlPt01* and *ArtLotlPt02)*, the triplestore contains a graph *RDFMM02.rdf* with information on the European List Of Trusted Lists.

5.3 Transformation of a TL into RDF – Part 1 Scheme information

We then transformed the Trusted Lists from Belgium, Spain, Estonia, The Netherlands, France and Germany. For brevity, we only describe how we processed the one from Belgium; all others were handled in a similar fashion. The XML version of the Belgian Trusted List was used from http://tsl.belgium.be/tsl-be.xml. This TL has 3 major elements:

- The *tsl:SchemeInformation*, indicating that the TLSO was FPS Economy, SMEs, Self-employed and Energy - Quality and Security - Information Management, where FPS stands for Federal Public Service,
- The *tsl:TrustServiceProviderList* identifying the different TSPs, and
- The *ds:Signature* containing the signature of the TLSO over the information provided

The XSL transformation *ArtMstlPt01* selects *tsl:TrustServiceProviderList* and extracts TSP information from all *tsl:TrustServiceProvider* elements. The information extracted and transformed provides information about the *SchemeName*, the *SchemeOperatorName*, the *SchemeTerritory* and the Scheme's *LegalNotice*. All this information relates to the various Schemes that are officialised through the country's Trust List.

It can be observed that we used the TSP's public *uri* as unique identifier for each TSP and its services. Many alternatives are possible e.g. his root certificate.

For every country, the transformation produces an RDF graph which is serialised into RDF/XML. The transformations and resulting RDF/XML and graphs are available at the same location

as for the LOTL, under their respective filenames. Furthermore, the results from the individual countries are available in a single merged file, *RDFMM11.rdf*.

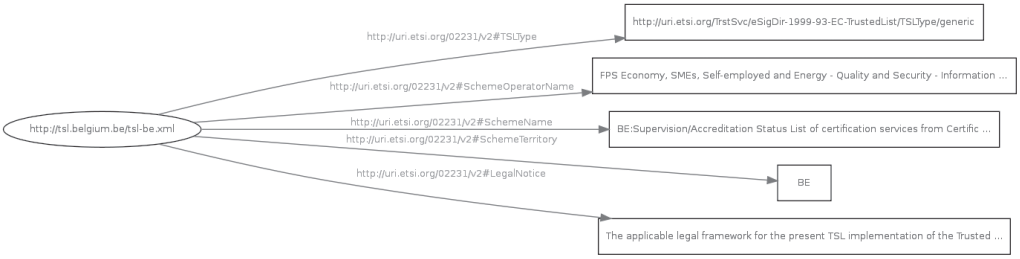

Fig 5: Partial graph of RDFMM11, the result of the transformed Belgian TL

As a consequence of these transformations (*ArtMstlPt01 and 11-15)*, the triplestore contains a graph *RDFMM11.rdf* with information on the Trusted Lists of Belgium, Spain, Estonia, The Netherlands, France and Germany.

5.4 Transformation of a TL into RDF – Part 2 TSP information

Furthermore, from the same TLs, information can be extracted and transformed with regard to the actual Trusted Service Providers (TSPs). We then processed the same Trusted Lists again, with a focus on individual TSP information. For brevity, we only describe how we processed the one from Belgium; all others were handled in a similar fashion.

The XSL transformation *Act2ActhasPartRTBE* selects the various TSPs active under the territory's scheme from the TL. For every selected country, the transformation produces an RDF graph which is serialised into RDF/XML. The transformations and resulting RDF/XML and graphs are available at the same location as for the LOTL, under their respective filenames. Furthermore, the results from the individual countries are available in a single merged file, *RDFMM22.rdf*.

Fig 6: Partial graph of RDFMM22, the second result of the transformed Belgian TL

We continue by merging the graphs *RDFMM02, RDFMM11 and RDFMM22* into the single graph *RDFMM021122.rdf*. The resulting RDF/XML and graphs are available at the same location as for the LOTL, under their respective filenames. The complete resulting graph is available at *http://www.marcsel.eu/3Store/RDFMM021122.rdf*, and its graphical representation in *RDFMM021122.png*.

As a consequence of the transformations (*Act2ActhasPartRT** with ** as 2-letter country symbol, BE, EE, ES, NL, ...)*, and the merging, the triplestore contains a consolidated graph *RDFMM021122.rdf* with information on the Trusted Lists of Belgium, Spain, Estonia, The Netherlands, France and Germany.

5.5 Transformation of a TL into RDF – Part 3 TSP offered services

Furthermore, from the same TLs, also information can be extracted and transformed with regard to the actual services offered such as Root Certification Authorities etc. The XSL transformation ActOffersServicesRTBE2 selects the various services listed for a TSP, active under the territory's scheme from the TL. For Belgium, this yields:

```
- <rdf:RDF>
  - <rdf:Description rdf:about="http://www.certipost.be/">
      <tsl:X509SubjectName>CN=Belgium Root CA, C=BE</tsl:X509SubjectName>
      <tsl:X509SubjectName>CN=Belgium Root CA2, C=BE</tsl:X509SubjectName>
      <tsl:X509SubjectName>CN=Belgium Root CA3, C=BE</tsl:X509SubjectName>
      <tsl:X509SubjectName>CN=Belgium Root CA4, C=BE</tsl:X509SubjectName>
    - <tsl:X509SubjectName>
        CN=Certipost E-Trust Primary Qualified CA, O=Certipost s.a./n.v., C=BE
      </tsl:X509SubjectName>
    - <tsl:X509SubjectName>
        CN=Certipost Public CA for Qualified Signatures, O=Certipost n.v./s.a., C=BE
      </tsl:X509SubjectName>
    </rdf:Description>
  - <rdf:Description rdf:about="http://www.swift.com/">
      <tsl:X509SubjectName>SWIFTNet PKI Certification Authority</tsl:X509SubjectName>
    </rdf:Description>
  </rdf:RDF>
```

Fig 7: RDF generated from the Belgian TL, illustrating root CA's of Certipost and SWIFT

This graph is merged with RDFMM021122.rdf from the previous step, yielding *RDFMM021122BE2.rdf*. As a consequence the triplestore now contains triples that link the TSPs (for Belgium: Certipost, responsible for the eID PKI, and SWIFT) to the Certification Authority services they offer.

5.6 Data enrichment from contextual information sources

We then further enriched this RDF graph with public domain information from the Balance Sheet department of Belgian Central Bank. Here information about public companies such as TSPs contains records that specify they submitted an annual report in good order, and whether this was supported by the opinion of an independent auditor, and if so, which auditor. The statutory accounts of Belgian public companies subject to publication are made available on-line at no cost, in XBRL format. We did not yet investigate the official publication channels in the other countries, but we assume comparable information will be available. XBRL is used to define and exchange financial information, such as balance sheets and financial statements. As XBRL is XML-based and uses the XML syntax and related XML technologies such as XML Schema, XLink, XPath, and Namespaces, it can easily be integrated in our processing. The XBRL of the TSP Certipost contains information about its auditor:

```
- <rdf:RDF>
  - <rdf:Description rdf:about="http://www.certipost.be">
      <EntityCurrentLegalName>CERTIPOST</EntityCurrentLegalName>
      <EntityIdentifier>BE0475396406</EntityIdentifier>
    - <EntityAdministrators>
        <ParticipantIndividualName rdf:resource="MEUNIER"/>
        <ParticipantIndividualFirstName rdf:resource="BAUDOUIN"/>
        <ParticipantIndividualName rdf:resource="COOLS"/>
        <ParticipantIndividualFirstName rdf:resource="NICO"/>
        <ParticipantIndividualName rdf:resource="WINAND"/>
        <ParticipantIndividualFirstName rdf:resource="PIERRE"/>
      </EntityAdministrators>
    - <Auditor>
        <AccountantsEntity>ERNST & YOUNG BEDRIJFSREVISOREN</AccountantsEntity>
        <ParticipantEntityIdentifier>BE0446334711</ParticipantEntityIdentifier>
        <ParticipantRepresentativeName rdf:resource="De Luyck"/>
        <ParticipantRepresentativeFirstName rdf:resource="Jan"/>
      </Auditor>
    - <LegalParent>
        <ParentEntity>Bpost NV van publiek recht</ParentEntity>
        <ParentEntityIdentifier>BE0214596464</ParentEntityIdentifier>
      </LegalParent>
    </rdf:Description>
  </rdf:RDF>
```

Fig 8: RDF generated from XBRL of the Balance Sheet department of Belgian Central Bank

We used transformation ActXbrlPt01.xsl to generate the above RDF, and once again, we merge this new RDF with the already existing. As a consequence, there are now triples that link the TSP to its auditor.

5.7 SPARQL query

Once all data is merged, the query below calculates a sample Trust Indicator.

```
PREFIX j.0: <http://uri.etsi.org/02231/v2#>
PREFIX j.1: <http://purl.org/dc/>
PREFIX xsd: <http://www.w3.org/2001/XSDSchema#>
SELECT ?MemberStateSchemeOperatorName ?TSPuri
(IF(BOUND(?MemberStateSchemeOperatorName), 1, -1) AS ?TE1)?TSPAuditorInfo
(IF(BOUND(?TSPAuditorInfo), 1, -1) AS ?TE2) ((?TE1 +?TE2) as ?TESum)
WHERE  { ?EUTrustedList j.0:TSLLocation ?MemberStateTrustedList .
       ?MemberStateTrustedList j.1:termshasPart ?TSPuri .
       ?MemberStateTrustedList j.0:SchemeOperatorName
       ?MemberStateSchemeOperatorName .
OPTIONAL {?TSPuri <http://example.com/foo#Auditor> ?TSPAuditorInfo}. }
```

This will yield a numerical value of e.g. '2' in case the TSP both operates under a Member State scheme (in which case there is a matching MemberStateSchemeOperatorName), and an Auditor can be identified (in which case there is TSPAuditorInfo). Lack of such elements will yield a lower value such as zero or '1'. Obviously this is only a crude selection of two possible elements to

consider. More work needs to be done to look into other elements and into building up a better articulated vocabulary.

6 Conclusion

It was demonstrated how for transactions relying on Trust Service Providers, additional Trust Indicators can be defined and calculated that go beyond the pure cryptographic checks. Such Trust Indicators can rely on public domain information and can be created using Semantic Web techniques such as RDF and SPARQL.

References

[DCMI] DCMI Usage Board. Dcmi metadata terms 2012-06-14, downloaded from http://dublincore.org/documents/dcmi-terms/.

[DCterms] DCMI Usage Board. Dublin core metadata element set. http://purl.org/dc/terms/.

[HAC] The Handbook of Applied Cryptography, CRC Press, 1997, Section 1.11.3 on trusted third parties and public-key certificates

[EIDAS2012] DG Connect, "Proposal for a Regulation of the European Parliament and of the Council on electronic identification and trust services for electronic transactions in the internal market", European Commission, 2012, COM(2012) 238.

[FSWT] Foundations of Semantic Web technologies, by Pascal Hitzler, Markus Kroetzsch and Sebastian Rudolph, CRC Press, ISBN 978-1-4200-9050-5

[LOV] Linked Open Vocabularies, Barnard Vatant and Pierre-Yves Vandenbussche, http://lov.okfn.org

[PROV2012] Jun Zhao and Olaf Hartig: Towards Interoperable Provenance Publication on the Linked Data Web. In Proceedings of the 5th Linked Data on the Web (LDOW) Workshop at WWW, Lyon, France, April 2012, downloaded from http://www.dbis.informatik.hu-berlin.de/fileadmin/research/papers/conferences/2012_ldow_hartig.pdf

[SWB1] Foundations of Semantic Web technologies, by Pascal Hitzler, Markus Kroetzsch and Sebastian Rudolph, CRC Press, ISBN 978-1-4200-9050-5

[SWS1] W3C, Semantic web stack, downloaded from http://www.w3.org/2000/Talks/1206-xml2k-tbl/slide10-0.html

[SWS2] W3C, Semantic web standards, http://www.w3.org/standards/semanticweb/

[TRDF] Trust in RDF Graphs, Dominik Tomaszuk, Karol Pak, and Henryk Rybinski, downloaded from http://ii.uwb.edu.pl/~dtomaszuk/papers/trust_in_rdf_graphs.pdf

[W3CRV] W3C RDF Validator, accessible at http://www.w3.org/RDF/Validator/