POWDER and QUATRO Plus:
Integrating the Semantic Web with Social Networking to Enhance the Acces to Online Resources

Barbara Carminati Elena Ferrari Andrea Perego

DICOM, Università degli Studi dell'Insubria, Varese, Italy

DICOM, 18 April 2008, Varese

Outline

What is a content/quality label

Definition A machine-readable description of the content/characteristics of online resources

Purpose To make an end user able to take decisions on how to use a resource, depending on his/her requirements and/or preferences

Current applications Child protection from possible harmful online content

Existing standars Platform for Internet Content Selection (PICS), released by the W3C in 1996
[Resnick and Miller, 1996]

An Example of PICS Label

```
(PICS-1.1 "http://www.gcf.org/v2.5"
  by "John Doe"
  labels on "1994.11.05T08:15-0500"
  until "1995.12.31T23:59-0000"
  for "http://w3.org/PICS/Overview.html"
  ratings (density 0 color/hue 1))
```

How to associate a PICS label with a resource:

- By embedding the label in the source code of an HTML page, using the `META` tag
- By configuring the web server in order to deliver the PICS label through the `PICS-Label` HTTP header

Usage and Diffusion

So far, resource labelling has not gained success:

- The effort needed to describe resources can be justified only if labels bring real benefits to a content provider
- Web resources' content and characteristics may frequently change, it is necessary to update content labels accordingly, to be sure that they actually describe the resources they refer to
- There exist other alternative solutions (at least, as far as child protection is concerned)

Are Content/Quality Labels Still Useful?

The situation has changed:

- Web metadata are currently seen by content providers as a means to personalise Web access and to assure the quality of online information
- There currently exist several Web-based social networks (WBSNs) providing their members the ability of specifying and sharing metadata (referred to as *tags*) concerning online resources
- Examples of WBSNs supporting *collaborative/social tagging*:
  - del.icio.us: `http://del.icio.us`
  - RawSugar: `http://rawsugar.com`
  - Flickr: `http://flickr.com`
  - Last.fm: `http://last.fm`

POWDER and QUATRO Plus

POWDER and QUATRO Plus build on this new attitude:

- **POWDER** (Protocol for Web Description Resources) is a W3C Working Group, which started its activity on March 2007
- It aims at the definition of a new generation of general purpose content/quality labels based on Semantic Web technologies, namely, RDF/OWL [Klyne and Carroll, 2004, Bechhofer et al., 2004], referred to as *Description Resources*.
- **QUATRO Plus** is a European project, funded in the framework of the Safer Internet *plus* Programme, which started its activity in October 2007
- Follow-up of the former EU project QUATRO ("Content Labels for Quality Assurance"), from which POWDER has sprung
- It aims at developing POWDER-based platform to be used by both labelling authorities and end users for creating, sharing, and distributing labels

POWDER: http://www.w3.org/2007/powder/

Active W3C members:

- Family Online Safety Institute (FOSI), UK (WG Chair)
- America On Line (AOL.com), USA
- Deutsche Telecom, UK
- National Council for the Scientific Research (NCSR) "Demokritos", Greece
- Opera Software, Norway
- Università degli Studi dell'Insubria, Italy
- Vodafone, UK

Published specifications:

- POWDER: Description Resources. W3C Public Working Draft, 17 March 2008
- POWDER: Grouping of Resources. W3C Public Working Draft, 24 March 2008
- POWDER: Use Cases and Requirements. W3C Working Group Note, 30 October 2007
- POWDER: Web Description Resources (WDR) Vocabulary. W3C First Public Working Draft, 25 September 2007
- POWDER: Web Description Resources Datatypes (WDRD). W3C First Public Working Draft, 25 September 2007

Description Resource Requirements

- A Description Resource (DR) must be able to describe one or multiple resources
- A DR must be able to denote any property/characteristics of a set of resources. Therefore:
  - DRs are independent from their possible uses
  - there is no restriction on the vocabularies used to describe a set of resources
- A DR must provide information which can be used to verify the DR's provenance and, possibly, when it has been published
- Multiple DRs may apply to the same (set of) resource(s), and the same (set of) resource(s) may refer to multiple DRs
- It must be possible to modify a DR without the need of modifying the (set of) resources applying to it. This means that DRs must be stored separately from the resources they apply to

The Structure of a DR

A POWDER DR consists of the following components:

Attribution It denotes the individual/organisation who created the DR, plus other information, such as the DR's issue date

Scope It denotes the set of resource a DR applies to, in terms of their URIs

Description It describes the content/characteristics of the resources denoted by the DR scope

Summary An optional component, providing a NL summary of the claims expressed in the DR

How to associate a DR with a resource:

- By linking the file containing the DR from an HTML page, using the `LINK` tag
- By configuring the web server in order to deliver the DR through the `Link` HTTP header

An Example of DR

Attribution	*Alice claims that*
Scope	*the resources hosted by example.org*
Description	*are safe for children*

DRs in RDF/OWL

The previous DR can be expressed in RDF/OWL as follows:

- Define the class S of resources hosted by example.org (the DR's scope)
- Define the class D of resources which are safe for children (the DR's description)
- State that class S is a subclass of class D

But:

- How can we define the class of resources matching a given URI?
- And what about the DR's attribution?

Grouping Resources by URI

Two alternative ways:

- enumerating the set of resources in the DR's scope
- use pattern matching

The former solution is feasible only if

- we know in advance the URIs of the corresponding resources
- they are not so many

The latter solution is the most suitable, *but RDF does not support pattern matching*

Grouping Resources by URI

POWDER defines the following semantic extension to RDF:

- a set of RDF properties are defined, in order to denote *a set of URIs* (not of resources) matching a given pattern
- an RDF property is defined (namely, wdr:hasIRI), which maps a resource to its URI (actually, its IRI)

The RDF/OWL representation of the DR in the previous example is then as follows:

- Define the class S of URIs having a `host` component ending with string `example.org`
- Define the class D of resources which are safe for children
- State that the class of resources having one of the URIs in class S is a subclass of class D

DR's Attribution

- The DR's attribution states who is the DR's author and when the DR has been issued
- This mean that we need to express *statements on statements*, which in RDF can be achieved only by using *reification* [Hayes, 2004], which does not work well with OWL
- The alternative (and adopted) solution is to use an rdf:Description applying to the whole document
- Consequences: an RDF/OWL can contain only DRs sharing the same attribution

An Example of RDF/OWL Representation of a DR

```
1  <?xml version="1.0"?>
2  <rdf:RDF xmlns:wdr="http://www.w3.org/2007/05/powder#"
3           xmlns:foaf="http://xmlns.com/foaf/0.1/"
4           xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
5           xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
6           xmlns:owl="http://www.w3.org/2002/07/owl#"
7           xmlns:dc="http://purl.org/dc/elements/1.1/"
8           xmlns:dcterms="http://purl.org/dc/terms/0.1/"
9           xmlns:ex="http://example.org/vocab#">

11   <rdf:Description rdf:about="">
12     <foaf:maker rdf:resource="http://authority.example.org/foaf.rdf#alice" />
13     <dcterms:issued>2007-12-14</dcterms:issued>
14   </rdf:Description>

16   <wdr:iriset rdf:nodeID="iriset_1">
17     <owl:intersectionOf rdf:parseType="Collection">
18       <owl:Restriction>
19         <owl:onProperty
                 rdf:resource="http://www.w3.org/2007/05/powder#includehosts" />
20         <owl:hasValue>example.com</owl:hasValue>
21       </owl:Restriction>
22     </owl:intersectionOf>
23   </wdr:iriset>
```

continues in the next slide...

An Example of RDF/OWL Representation of a DR (continued)

```
24    <owl:Class rdf:nodeID="descriptorset_1">
25      <owl:intersectionOf rdf:parseType="Collection">
26        <owl:Restriction>
27          <owl:onProperty rdf:resource="http://example.org/vocab#childsafe" />
28          <owl:hasValue>true</owl:hasValue>
29        </owl:Restriction>
30      </owl:intersectionOf>
31    </owl:Class>

33    <owl:Restriction>
34      <owl:onProperty rdf:resource="http://www.w3.org/2007/05/powder#hasIRI"/>
35      <owl:someValuesFrom rdf:nodeID="iriset_1"/>
36      <rdfs:subClassOf rdf:nodeID="descriptorset_1"/>
37    </owl:Restriction>

39  </rdf:RDF>
```

Using Reification for DR's Attribution

```
...

  <owl:Restriction>
    <owl:onProperty rdf:resource="http://www.w3.org/2007/05/powder#hasIRI"/>
    <owl:someValuesFrom rdf:nodeID="iriset_1"/>
    <rdfs:subClassOf rdf:ID="dr_1" rdf:nodeID="descriptorset_1"/>
  </owl:Restriction>

  <rdf:Description rdf:about="#dr_1">
    <foaf:maker rdf:resource="http://authority.example.org/foaf.rdf#alice" />
    <dcterms:issued>2007-12-14</dcterms:issued>
  </rdf:Description>

...
```

Problems & Considerations

- RDF/OWL makes very difficult to represent the semantics of DRs: OWL is both too much and not enough expressive for DRs
- The RDF/OWL representation of DRs is extremely verbose and error prone
- There exist Semantic Web technologies which perfectly suits DRs, namely, *rule languages*, such as N3Logic [Berners-Lee et al., 2006] and SWRL (Semantic Web Rule Language) [Horrocks et al., 2004]:
  - The semantics of DR is equivalent to a *if*... *then* statement (e.g., "*if* a resource is hosted by example.org, *then* it is safe for children")
  - Rule languages implement *if*... *then* statement is form of Horn-like clauses
- **But**: at the moment there does not exist any *standard* Semantic Web rule language

N3Logic

Relevant features:

- it makes use of the compact and readable notation provided by N3 (Notation 3)
- *quoted formulae*: thanks to this it is possible to specify statements where the subject and/or the object are RDF graphs (each quoted formula is an RDF graph)
- support for rules
- support for a set of predicates for manipulating and comparing strings and numbers

N3Logic supports negation in the form of *scoped negation as failure*, but it does not support disjunction.
However:

$$(p \wedge q) \vee (r \wedge s) \rightarrow t \equiv (p \wedge q \rightarrow t) \wedge (r \wedge s \rightarrow t)$$

Example of DR in N3Logic

```
@prefix      wdr: <http://www.w3.org/2007/05/powder> .
@prefix     foaf: <http://xmlns.com/foaf/0.1/> .
@prefix dcterms: <http://purl.org/dc/terms/0.1/> .
@prefix       ex: <http://example.org/vocab#> .

{
  {@forall X .
    {X wdr:hasIRI [wdr:includehosts "example.org"]}
      =>
    {X ex:childsafe "true"} .
  } dcterms:issued "2007-12-14"

} foaf:maker <http://authority.example.org/foaf.rdf#alice> .
```

Resolution

- The basic representation of POWDER DRs is made by using an XML dialect, enforcing their *operational* semantics
- XML DRs can be converted into the corresponding RDF/OWL representation, enforcing their *formal* semantics, by using transformation rules
- The W3C GRDDL technology (Gleaning Resource Descriptions from Dialects of Languages) [Connolly, 2007] is used for this purpose, possibly in combination with XSL transform rules [Clark, 1999]

An Example of XML Representation of a DR

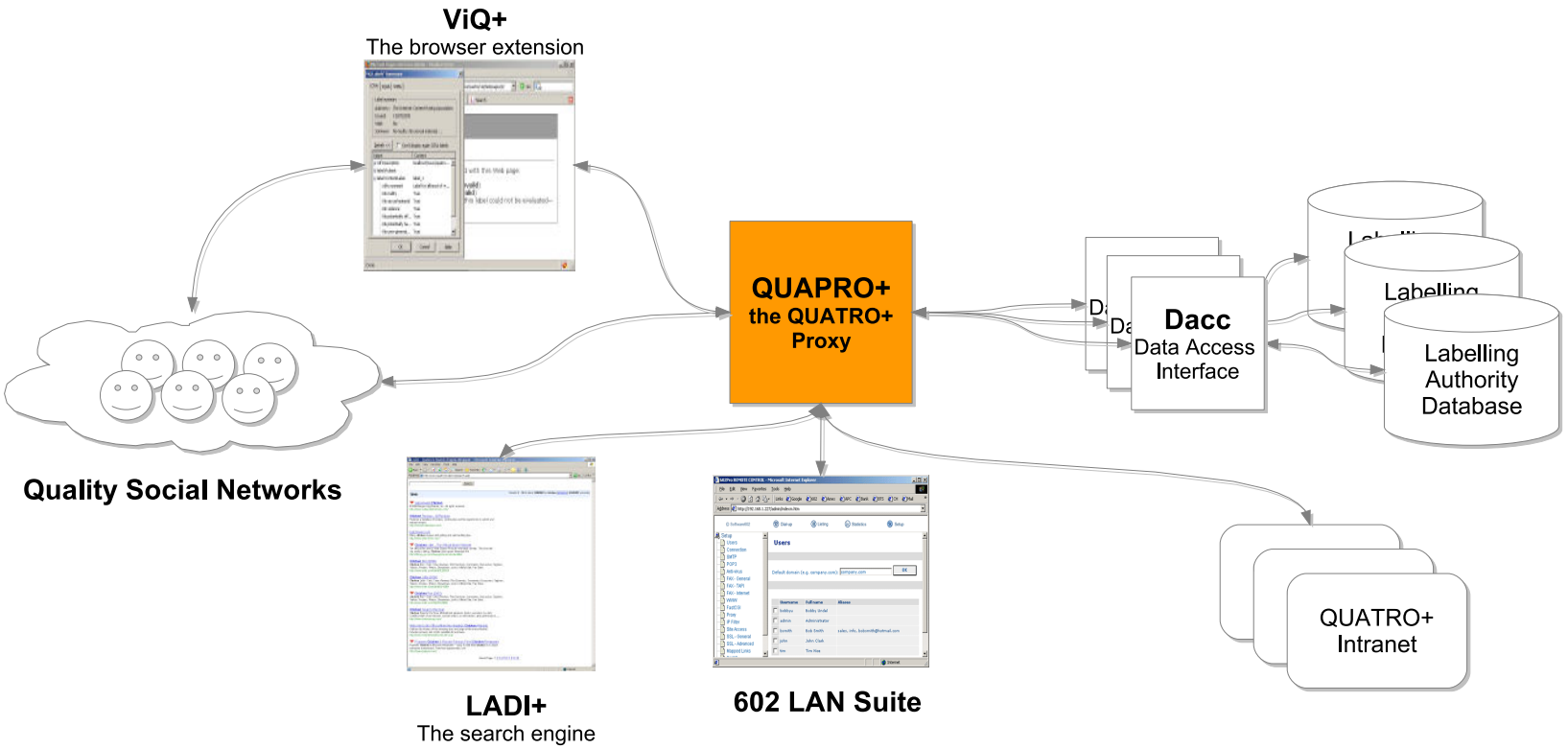
```
<?xml version="1.0"?>
<powder xmlns="http://www.w3.org/2007/05/powder#"
        xmlns:ex="http://example.org/vocab#">

  <attribution>
    <maker>http://authority.example.org/foaf.rdf#alice</maker>
    <issued>2007-12-14</issued>
  </attribution>

  <dr>
    <iriset>
      <includehosts>example.org</includehosts>
    </iriset>
    <descriptorset>
      <ex:childsafe>true</ex:childsafe>
    </descriptorset>
  </dr>

</powder>
```

QUATRO Plus: http://www.quatroplus.org/

Project partners:

- Labelling authorities:
  - FOSI, UK
  - Web Mèdica Acreditada (WMA), Spain
  - The Internet Quality Agency (IQUA), Spain
- Universities:
  - NCSR "Demokritos", Greece (Technical Coordinator)
  - Università degli Studi dell'Insubria, Italy
  - Università degli Studi di Milano, Italy
- Industry:
  - MD Partners, UK (Project Manager)
  - CoolWave, UK
  - Software 602, Czech Republic
  - ECP.NL, The Netherlands

The QUATRO Plus Platform



What is a Web-based Social Network

- Services which allow Web users to join communities established in order to address given interests or purposes
- Provide an information space where Web users can
  - establish relationships
  - publish and share resources (personal data, blogs, photos, etc.)
- Due to the work we have carried out in this field (see, e.g., [Carminati et al., 2006, Carminati et al., 2007b, Carminati et al., 2007a, Carminati et al., 2008]), in QUATRO Plus we are responsible of the design and implementation of the social networking component of the QUATRO Plus platform, i.e., the *Quality Social Network*

The Quality Social Network (QSN)

- Besides supporting the basic social networking services, it allows its members to specify labels, and to express their dis/agreement about existing labels, by associating *ratings* with them
- The collected user-defined labels and ratings are *aggregated* and statistically analysed in order to assess the trustworthiness of the labels associated with a given (set of) resource(s)
- What is this information used for and by whom?
  - it is made available to the other components of the QUATRO Plus platform, and to their users
  - it is used to enforce *user preferences*, thanks to which end users can state which action must be performed by the user agent (in our case, ViQ+) on a given resource, upon detection of labels containing given descriptors and having a given trustworthiness

Role of the QSN

It addresses some of the issues which have limited the success of resource labelling, by establishing a convergence between 'official' and collaborative labelling.
More precisely, the QSN aims to:

- increase the number of labelled resources
- give a measure of how much I can trust the claimed content/characteristics of a resource
- make easier the detection of labels which no longer describe the resource(s) they are associated with

Trust Computation

Relevant information:

- number of occurrences of a descriptor in the labels associated with a given (set of) resource(s)
- number of ratings associated with (the labels containing) that descriptor

Possible rating values:

- positive: "I agree"
- negative: "I disagree"
- neutral: "I don't know"

Global or Local Trust?

- Occurrences of labels and ratings allow the computation of a *global* trust score (what is called *reputation score*), i.e., a trust score computed considering all the QSN members equally trustworthy
- It might be useful to provide also a *local* trust score, which may vary depending on the QSN member
- The principle is that I can consider more trustworthy the labels and/or ratings of member m_1 than those of member m_2
- This can be achieved by refining trust computation by taking into account the *trust relationship* existing between QSN members

Trust Relationships

Two different types:

Explicit Trust relationships which are established explicitly by QSN members. E.g., Alice my state that she consider Bob to be trustworthy, but that she does not trust Eve

Implicit Trust relationships derived from *trust policies*, denoting the characteristics of the QSN members that I consider to be trustworthy

Trust relationships can be:

- *binary* or *scalar*: in the former case, I can just state whether I consider a given user to be trustworthy or not, whereas in the latter I can also specify how much I trust a given user
- *topical* or *absolute*: i.e., I can state that I consider trustworthy a given user for specific topics or for any topics

Trust Algorithm

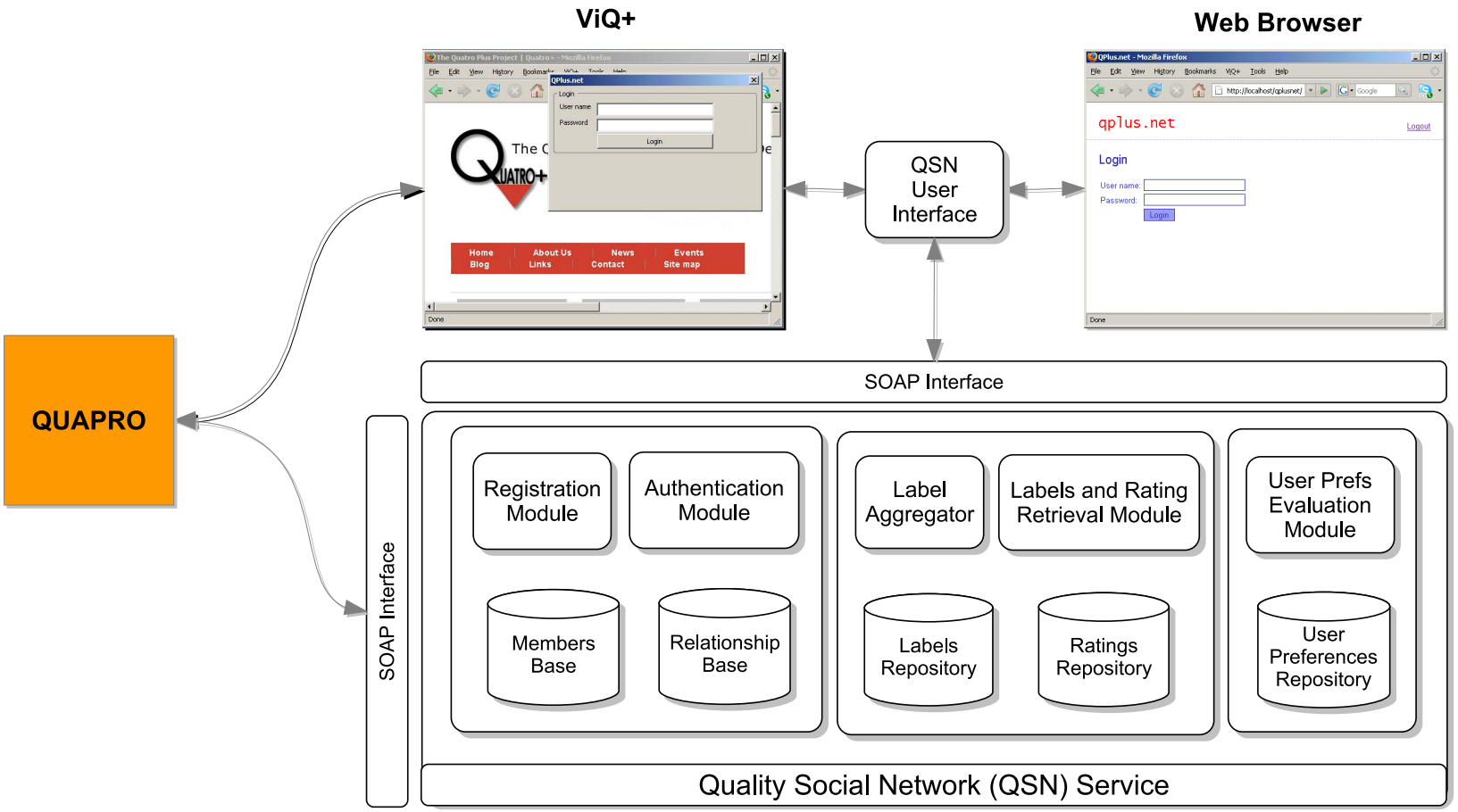
The algorithm to be adopted depends on the following factors:

- whether the trust score is global or local
- whether trust relationships are binary or scalar
- whether trust relationships are topical or absolute

Possible options:

- EigenTrust [Kamvar et al., 2003], for global trust scores
- TidalTrust [Golbeck, 2005], for local trust scores
- Designing a new algorithm

QSN Architecture



The QSN User Interface (QUI)

Issues to be addressed:

- QSN services must be accessed both via a Web interface ((X)HTML) and ViQ+ (XUL) [Goodger et al., 2001]
- Localised versions must be available

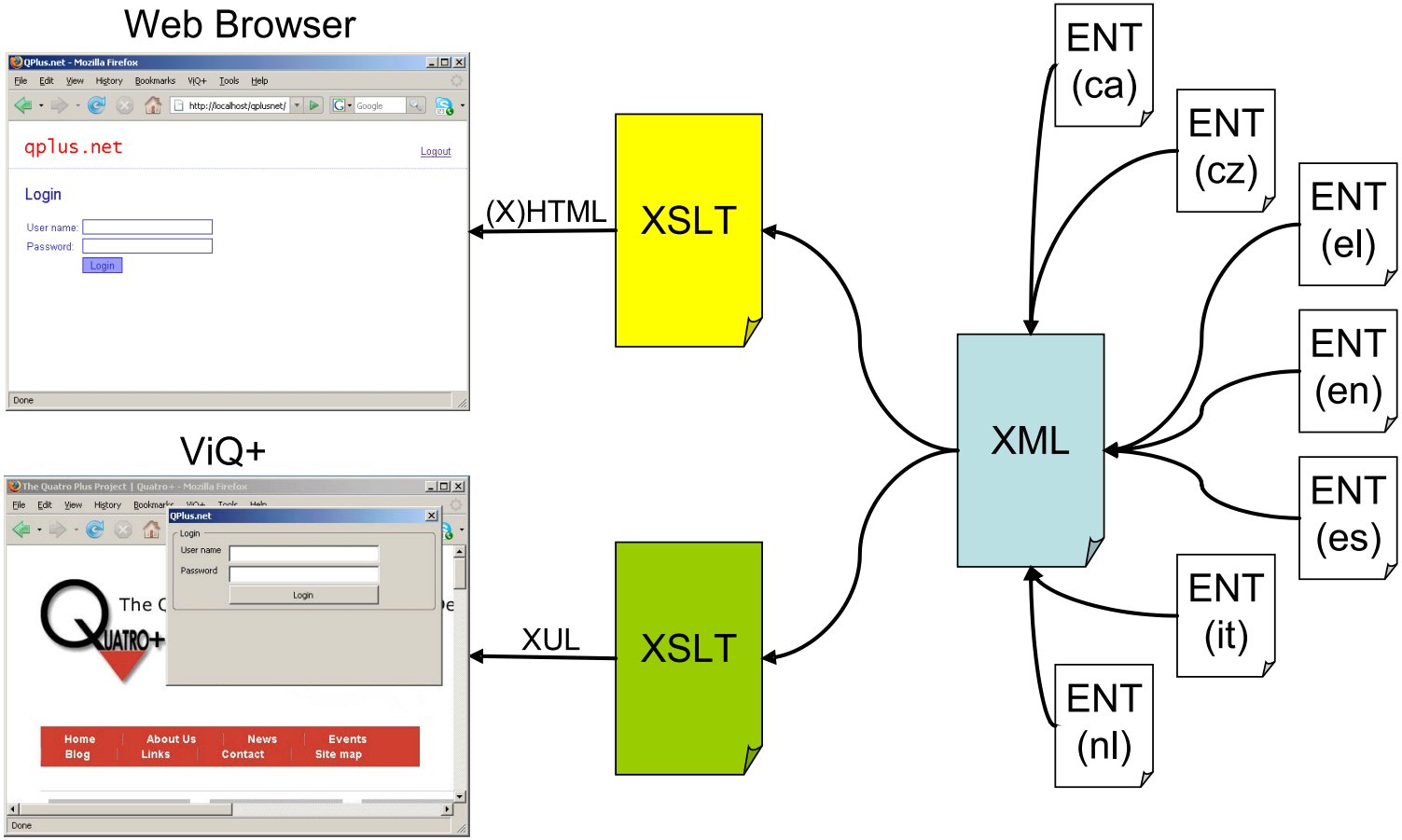
Adopted solution: using XML, in combination with XSLT and *entity references*

- XML-based interface + XSLT(s) = (X)HTML or XUL interface
- XML-based interface + XML entity references = localised versions

Advantages:

- Just one version of the QSN interface
- Language-independent interface

QUI Flow Diagram



References I

Bechhofer, S., van Harmelen, F., Hendler, J., Horrocks, I., McGuinness, D. L., Patel-Schneider, P. F., and Stein, L. A. (2004).
OWL Web Ontology Language: Reference.
W3C Recommendation, W3C.
Available at: http://www.w3.org/TR/owl-ref.

Berners-Lee, T., Connolly, D., Kagal, L., Scharf, Y., and Hendler, J. (2006).
N3Logic: A logic for the Web.
Technical report, MIT CSAIL DIG.
Available at: http://dig.csail.mit.edu/2006/Papers/TPLP/n3logic-tplp.pdf.

Carminati, B., Ferrari, E., and Perego, A. (2006).
Rule-based access control for social networks.
In *OTM 2006 Workshops*, volume 2 of *LNCS 4278*, pages 1734–1744. Springer Verlag.

Carminati, B., Ferrari, E., and Perego, A. (2007a).
Enforcing access control in web-based social networks.
Submitted for journal publication.

Carminati, B., Ferrari, E., and Perego, A. (2007b).
Private relationships in social networks.
In *ICDE 2007 Workshops*, pages 163–171. IEEE CS Press.

References II

Carminati, B., Ferrari, E., and Perego, A. (2008).
Security and privacy in social networks.
In Khosrow-Pour, M., editor, *Encyclopedia of Information Science and Technology, Second ed.* IGI Publishing, Hershey, PA.
To appear.

Clark, J. (1999).
XSL Transformations (XSLT) — Version 1.0.
W3C Recommendation, W3C.
Available at: http://www.w3.org/TR/xslt.

Connolly, D. (2007).
Gleaning resource descriptions from dialects of languages (GRDDL).
W3C Recommendation, W3C.
Available at: http://www.w3.org/TR/grddl/.

Golbeck, J. A. (2005).
Computing and Applying Trust in Web-based Social Networks.
PhD thesis, Graduate School of the University of Maryland, College Park.
Available at: http://trust.mindswap.org/papers/GolbeckDissertation.pdf.

Goodger, B., Hickson, I., Hyatt, D., and Waterson, C. (2001).
XML User Interface Language (XUL) 1.0.
Technical specification, Mozilla.org.
Available at: http://www.mozilla.org/projects/xul/xul.html.

References III

Hayes, P. (2004).
RDF semantics.
W3C Recommendation, W3C.
Available at: http://www.w3.org/TR/rdf-mt/.

Horrocks, I., Patel-Schneider, P. F., Boley, H., Tabet, S., Grosof, B., and Dean, M. (2004).
SWRL: A Semantic Web rule language combining OWL and RuleML.
W3C Member Submission, W3C.
Available at: http://www.w3.org/Submission/SWRL.

Kamvar, S. D., Schlosser, M. T., and Garcia-Molina, H. (2003).
The Eigentrust algorithm for reputation management in P2P networks.
In *WWW 2003*, pages 640–651.

Klyne, G. and Carroll, J. J. (2004).
Resource Description Framework (RDF): Concepts and abstract syntax.
W3C Recommendation, W3C.
Available at: http://www.w3.org/TR/rdf-concepts.

Resnick, P. and Miller, J. (1996).
PICS: Internet access controls without censorship.
Commun. ACM, 39(10):87–93.