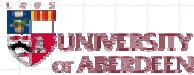# The Semantic Web: A Short, Practical Introduction

Alun Preece

Department of Computing Science
University of Aberdeen, Scotland, UK
apreece@csd.abdn.ac.uk
http://www.csd.abdn.ac.uk/~apreece

UNIVERSITY of ABERDEEN

Or…

```
<rss:item
  rdf:about="http://www.csd.abdn.ac.uk/~apreece/
  talks/TheSemanticWebAShortPracticalIntroduction/">
  <rdf:type rdf:resource=
  "http://xmlns.com/foaf/0.1/Document" />
  <rss:title>The Semantic Web: A Short, Practical
  Introduction</rss:title>
  <rss:description>Slides for plenary talk at Aberdeen
  University's Conference for Computing Teachers,
  September 10 2003</rss:description>
  <rss:link>http://www.csd.abdn.ac.uk/~apreece/
  talks/TheSemanticWebAShortPracticalIntroduction/
  TheSemanticWebAShortPracticalIntroduction.pdf
  </rss:link>
  <dc:creator>Alun Preece</dc:creator>
</rss:item>
```

UNIVERSITY of ABERDEEN

Scope of this talk

- What is the Semantic Web, and why do we need it now?
- How does the Semantic Web relate to the "traditional" Web?
- What are the main components of the Semantic Web information architecture?
- What does Semantic Web technology buy us in terms of applications...
  - ... that are "lightweight" and easy to build?
  - ... that are more "heavyweight" and more challenging to build?

UNIVERSITY OF ABERDEEN

What is the Semantic Web?

W3C "The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation."
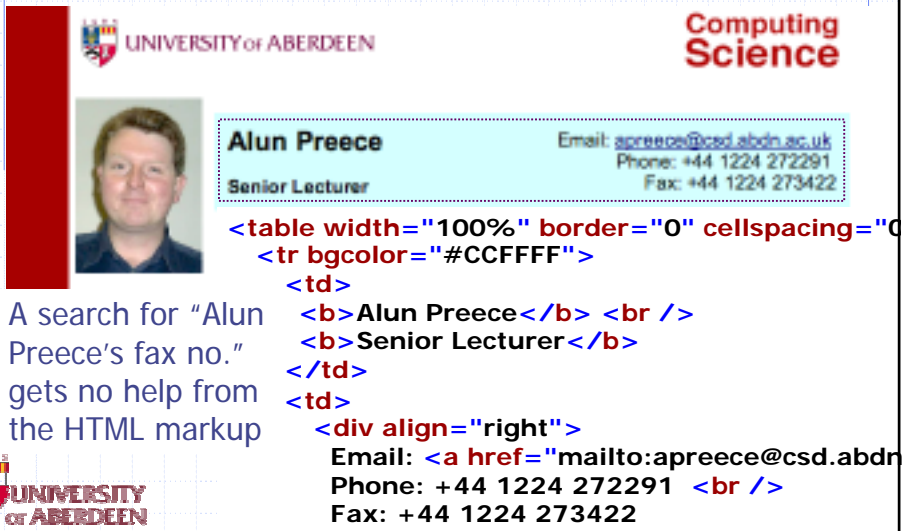
> Tim Berners-Lee, James Hendler, Ora Lassila,
> *The Semantic Web*, Scientific American, May 2001

- Aim: to create a network of machine-processable resources
- Existing in parallel with the current World Wide Web
- Enables software to carry out tasks on users' behalf
- Moving from a Web of "finding things" to a Web of "doing things"

UNIVERSITY OF ABERDEEN

The problem with the HTML Web

◆ A typical homepage, today:

UNIVERSITY of ABERDEEN

Computing Science
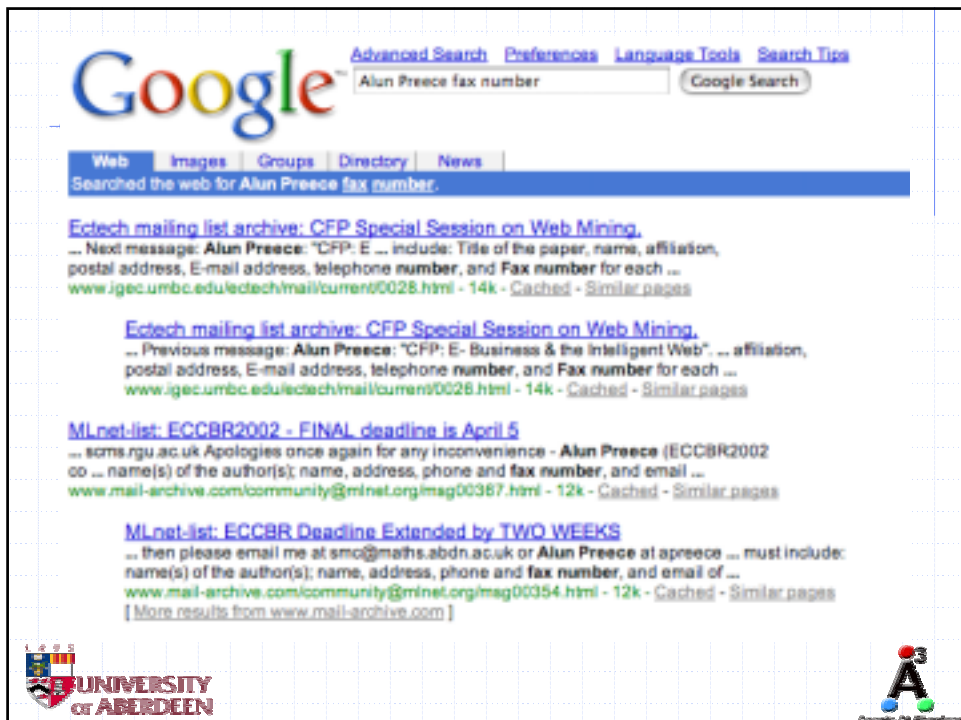
Alun Preece

Senior Lecturer

Email: apreece@csd.abdn.ac.uk
Phone: +44 1224 272291
Fax: +44 1224 273422

A search for "Alun Preece's fax no." gets no help from the HTML markup

```
<table width="100%" border="0" cellspacing="0
  <tr bgcolor="#CCFFFF">
    <td>
      <b>Alun Preece</b> <br />
      <b>Senior Lecturer</b>
    </td>
  <td>
    <div align="right">
      Email: <a href="mailto:apreece@csd.abdn
      Phone: +44 1224 272291  <br />
      Fax: +44 1224 273422
```

UNIVERSITY of ABERDEEN

Google

Advanced Search   Preferences   Language Tools   Search Tips

Alun Preece fax number   Google Search

| Web | Images | Groups | Directory | News |

Searched the web for **Alun Preece** fax number.

Ectech mailing list archive: CFP Special Session on Web Mining.
... Next message: **Alun Preece**: "CFP: E ... include: Title of the paper, name, affiliation, postal address, E-mail address, telephone **number**, and **Fax number** for each ...
www.igec.umbc.edu/ectech/mail/current/0028.html - 14k - Cached - Similar pages

Ectech mailing list archive: CFP Special Session on Web Mining.
... Previous message: **Alun Preece**: "CFP: E- Business & the Intelligent Web". ... affiliation, postal address, E-mail address, telephone **number**, and **Fax number** for each ...
www.igec.umbc.edu/ectech/mail/current/0028.html - 14k - Cached - Similar pages

MLnet-list: ECCBR2002 - FINAL deadline is April 5
... scms.rgu.ac.uk Apologies once again for any inconvenience - **Alun Preece** (ECCBR2002 co ... name(s) of the author(s); name, address, phone and **fax number**, and email ...
www.mail-archive.com/community@mlnet.org/msg00387.html - 12k - Cached - Similar pages

MLnet-list: ECCBR Deadline Extended by TWO WEEKS
... then please email me at smo@maths.abdn.ac.uk or **Alun Preece** at apreece ... must include: name(s) of the author(s); name, address, phone and **fax number**, and email of ...
www.mail-archive.com/community@mlnet.org/msg00354.html - 12k - Cached - Similar pages
[More results from www.mail-archive.com]

UNIVERSITY of ABERDEEN

Doesn't XML solve this problem?

◆ With XML, custom sets of tags can be defined, which can be:

- ▪ "styled" into conventional HTML
- ▪ processed directly by software - so can be queried...

Alun Preece

Email: apreece@csd.abdn.ac.uk
Phone: +44 1224 272291
Fax: +44 1224 273422

```
<Xfgpds ID="apreece">
 <ghew>Alun Preece</ghew>
 <ngre resource="mailto:apreece@csd.abdn.ac.uk" />
 <hytcx resource="tel:+44-1224-272291" />
 <hytcx resource="fax:+44-1224-273422" />
</Xfgpds>
```

But this is just more syntax - the tags don't <u>mean</u> anything!

UNIVERSITY of ABERDEEN

The <u>Semantic</u> Web

◆ XML alone is not enough - the XML tags need a defined semantics, to make them <u>meaningful</u>

◆ Formally, to relate the tag symbols to the things they <u>represent</u> in the real world

◆ So the markup becomes a <u>model</u> of the real world

```
<Person ID="apreece">
 <name>Alun Preece</foaf:name>
 <mbox resource="mailto:apreece@c
 <phone resource="tel:+44-1224-272
 <phone resource="fax:+44-1224-27
</Person>
```

FAX

UNIVERSITY of ABERDEEN

Two Webs in parallel

- ◆ The Semantic Web is <u>not</u> a replacement for the current Web
- ◆ <u>Semantic markup</u> is designed to exist alongside HTML markup - often in the form of <u>metadata</u> (data that describes other data)
  - Humans will continue to view the HTML
  - Software can process the Semantic markup
  - Hence the W3C's aim: "better enabling computers and people to work in cooperation"
- ◆ (Because the Semantic markup is in XML, it is possible to generate HTML from it...)
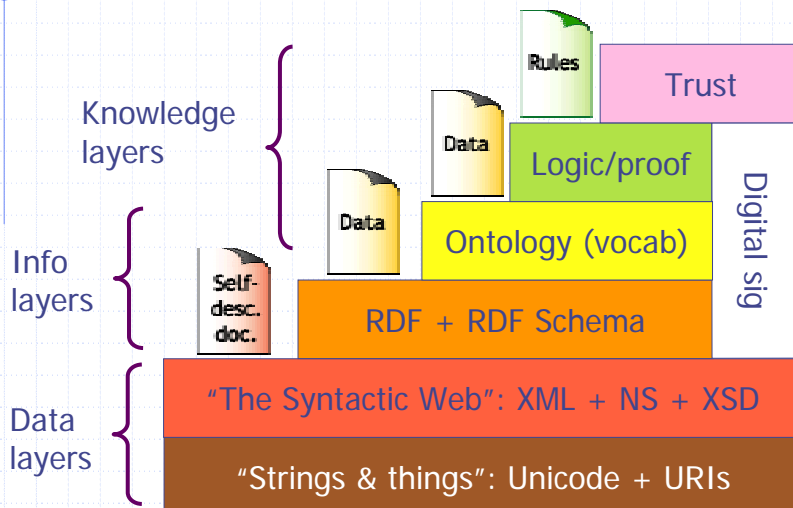

Web/Semantic Web example

http://www.csd.abdn.ac.uk/~apreece/index.html

http://www.csd.abdn.ac.uk/~apreece/index.rdf

Semantic Web architecture

Knowledge layers

Info layers

Data layers

Trust

Logic/proof

Ontology (vocab)

RDF + RDF Schema

"The Syntactic Web": XML + NS + XSD

"Strings & things": Unicode + URIs

Digital sig

[Adapted from Semantic Web "layer cake" slide due to Tim Berners-Lee]

Data layers: Unicode, URIs, XML

Like the "traditional Web" (in recent years), Semantic Web data is based on W3C-recommended standards:

◆ Unicode for strings (in all languages)

◆ URIs - Uniform Resource Identifiers - to name "things"

◆ XML as the standard extensible markup language

◆ XML Schema for a variety of primitive datatypes (integer, real number, string, date, URI, ...)

◆ XML namespaces to give global scope for tag names

```
<rdf:RDF
    xmlns:vc="http://www.w3.org/2001/vcard-rdf/3.0#">
  <rdf:Description
    rdf:about="http://www.csd.abdn.ac.uk/~apreece">
    <vc:FN>Alun Preece</vc:FN>
    <vc:EMAIL>apreece@csd.abdn.ac.uk</vc:EMAIL> ...
```

Information layer: RDF + RDFS

- RDF - Resource Description Framework - is the foundation of the Semantic Web standards
- RDF provides:
    - a simple semantic data model with classes (entities) & properties (relationships)
    - schema definition constructs (RDF Schema) to define simple vocabularies of terms
    - an XML syntax for marking-up RDF data
- RDF is the best-developed aspect of the Semantic Web:
    - many RDFS vocabularies are currently available
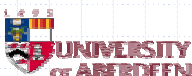   - a suite of software tools exists to process RDF

Sample RDF Schema fragment

- An RDFS definition for a class called Person:

```
<rdfs:Class rdf:ID="http://xmlns.com/foaf/0.1/Person" />
```

- RDFS definitions for three RDF properties - name, mbox (email address), and phone:

```
<rdf:Property rdf:ID="http://xmlns.com/foaf/0.1/name">
 <rdfs:range rdf:resource=
  "http://www.w3.org/2000/01/rdf-schema#Literal" />
</rdf:Property>
<rdf:Property rdf:ID="http://xmlns.com/foaf/0.1/mbox" />
<rdf:Property rdf:ID="http://xmlns.com/foaf/0.1/phone" />
```

- Note how all these terms are named globally with URIs

Sample RDF data fragment

```
<rdf:RDF
  xmlns:rdf=
      "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xml:base=
      "http://www.csd.abdn.ac.uk/~apreece/foaf.rdf#">

  <foaf:Person rdf:ID="apreece">
    <foaf:name>Alun Preece</foaf:name>
    <foaf:mbox rdf:resource=
        "mailto:apreece@csd.abdn.ac.uk" />
    <foaf:phone rdf:resource="tel:+44-1224-272291" />
    <foaf:phone rdf:resource="fax:+44-1224-273422" />
    …
  </foaf:Person>

<rdf:RDF />
```
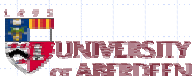
UNIVERSITY
of ABERDEEN

Querying RDF

◆ RDF statements can be parsed from the XML format into an RDF model, for example using Hewlett Packard's Jena toolkit

◆ Jena RDF models can be queried directly using the RDQL query language

◆ Example, "retrieve the phone number(s) of the person whose name is 'Alun Preece'":

```
SELECT ?y
WHERE ( ?x, <foaf:name>, "Alun Preece" )
AND ( ?x, <foaf:phone>, ?y > )
USING foaf FOR <http://xmlns.com/foaf/0.1/>
```

◆ (RDF can also be queried in XML RDF syntax using University of Aberdeen's RDF Query-By-Example)

UNIVERSITY
of ABERDEEN

More than just a smarter Google

- ◆ In addition to offering accurate searching, by querying, Semantic Web data enables many kinds of applications
- ◆ Examples
  - charting communities of friends and colleagues
  - building collaborative community-oriented apps
  - information integration based on standard vocab
  - Web "push": publish & subscribe
  - automated Web services
- ◆ A lot can often be done with small amounts of semantic markup!
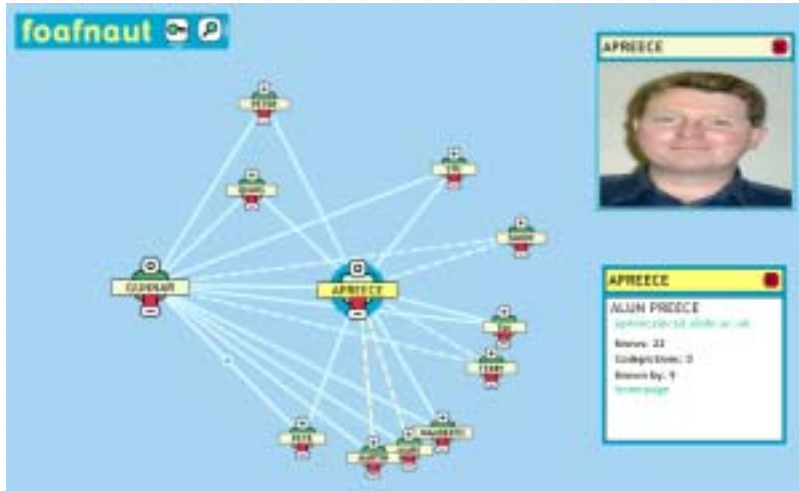

Six degrees of separation: FOAF

- ◆ The Friend-Of-A-Friend (FOAF) vocabulary covers
  - entities: people, organisations, projects, documents
  - "identifying" details: mbox, homepage, phone, depiction
  - relationships between people: who knows who
- ◆ We've already seen FOAF definitions for the Person class, and the name, mbox, & phone properties
- ◆ The FOAF knows property:

```
<rdf:Property
  rdf:about="http://xmlns.com/foaf/0.1/knows">
 <rdfs:domain
  rdf:resource="http://xmlns.com/foaf/0.1/Person" />
 <rdfs:range
  rdf:resource="http://xmlns.com/foaf/0.1/Person" />
</rdf:Property>
```

FOAFnaut

◆ FOAFnaut is a tool that browses knows links:

The need for ontologies

◆ RDF is designed to be simple
◆ To define more sophisticated vocabulary, we need to go one layer higher: to the ontology layer
◆ The Semantic Web ontology language, OWL, extends RDF with some additional functionality
◆ Concrete examples:
  ▪ a Person must have at least one name
  ▪ a Person must have exactly one age
  ▪ the class Person is the disjoint union of the classes Man and Woman
  ▪ an email address (mbox) belongs to only one Person
◆ This last example is crucial to FOAF...

A little ontology goes a long way

◆ FOAF uses OWL to define the mbox property:

```
<rdf:Property
  rdf:about="http://xmlns.com/foaf/0.1/mbox".">
  <rdf:type rdf:resource=
   "http://www.w3.org/2002/07/owl#
      InverseFunctionalProperty" />
</rdf:Property>
```

◆ This definition means: "mbox is a personal mailbox, i.e. an Internet mailbox associated with exactly one owner"

◆ This means, in database terms, the value of mbox acts as a primary key for Persons in the FOAF world - a unique ID

Ontology mapping

◆ OWL can also be used to map one vocabulary to another

◆ Example: the vCard EMAIL property is the same as FOAF's mbox:

```
<rdf:Property
  rdf:about=
    "http://www.w3.org/2001/vcard-rdf/3.0#EMAIL">
    <owl:equivalentProperty
     rdf:resource="http://xmlns.com/foaf/0.1/mbox" />
</rdf:Property>
```
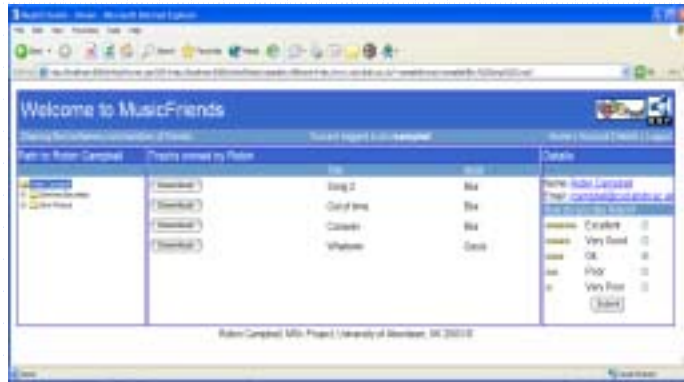
◆ An OWL reasoner could use this equivalence to derive a value for some resource's vcard:EMAIL if it can find a value for foaf:mbox

Application snapshot: MusicFriends

◆ MP3 file sharing among a community of friends
◆ Uses FOAF RDF vocab for friend-to-friend links
◆ Uses MusicBrainz RDF vocab for MP3 collections

Application snapshot: AKT 3store

◆ Repository of over <u>7 million</u> RDF statements
◆ Covers entire UK computing science community

Application snapshot: RSS

◆ RDF Site Summary (RSS) is an open framework for "publish and subscribe" applications, using RDF

◆ Many news sites (and other sites with frequent updates) now provide RSS channels

◆ By using a "newsfeed" tool, one can subscribe to RSS channels of one's choosing

◆ When new items are published in RSS/RDF format, subscribers are notified

◆ Items can be anything with a URI
  ▪ news stories
  ▪ published documents
  ▪ slides of talks...

"Eating our own dogfood"

◆ Here's the RSS metadata for this talk:

```
<rss:item
  rdf:about="http://www.csd.abdn.ac.uk/~apreece/
  talks/TheSemanticWebAShortPracticalIntroduction/">
  <rdf:type rdf:resource=
  "http://xmlns.com/foaf/0.1/Document" />
  <rss:title>The Semantic Web: A Short, Practical
  Introduction</rss:title>
  <rss:description>Slides for plenary talk at Aberdeen
  University's Conference for Computing Teachers,
  September 10 2003</rss:description>
  <rss:link>http://www.csd.abdn.ac.uk/~apreece/
  talks/TheSemanticWebAShortPracticalIntroduction/
  TheSemanticWebAShortPracticalIntroduction.pdf
  </rss:link>
  <dc:creator>Alun Preece</dc:creator>
</rss:item>
```
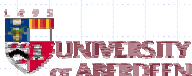
Semantic Web services

◆ The key features of the Semantic Web...
- machine-processable data
- standard vocabularies
- compatibility with the "Web family" of standards

◆ ... makes the technology very appealing for automated Web services in all sectors:
- E-business
- E-science
- E-health
- E-governance

Web services use Web standards to allow client software to call upon Web servers to carry out <u>tasks</u> - far more than just information retrieval...

UNIVERSITY of ABERDEEN

Granite Nights service

◆ Semantic Web service: helps a user to schedule a night out in Aberdeen!

◆ Sources of information, all in RDF:
- Restaurants (uses standard ontology)
- Cinema shows (uses standard ontology)
- Pubs (uses a home-grown ontology)

◆ Remembers and recalls user preferences - semantic profiling

◆ AI-based scheduler maps RDF data to constraints and produces valid schedules

◆ Part of EU-funded Agentcities.NET project
(Worldwide network of intelligent Web services)

UNIVERSITY of ABERDEEN

Granite Nights – input page

Granite Nights – output page

Static info source (restaurants)

```
<res:Restaurant rdf:about="#lalombarda">
  <res:name>La Lombarda</res:name>
  <res:averageMealDuration>2</res:averageMealDuration>
  <res:address>
    <add:Address rdf:about="rest#lombardaaddr" />
  </res:address>
  <res:atmospheres rdf:resource="res#CasualAtmosphere" />
  <res:atmospheres rdf:resource="res#RelaxedAtmosphere"/>
  <res:caterings rdf:resource="res#ALaCarte" />
  <res:caterings rdf:resource="res#HomeDelivery" />
  <res:facilities rdf:resource="res#SmokingFacility" />
  <res:typeOfCuisine rdf:resource="res#ItalianCuisine" />
</res:Restaurant>
```

UNIVERSITY of ABERDEEN

Dynamic info source (cinemas)

SCOOT® Busine

Cinema details
You searched for U G C Ci
Aberdeenshire

U G C CINEMA ABERDEE
QUEENS LINK LEISURE PAR
LINKS ROAD, ABERDEEN
AB24 5EN

Tel: 0870 1550502

View website

Films showing

B Mile (15) Film Revie
FRI,SAT,SUN,MON,TUES,WE
9:00PM,

Pianist, The (15)
FRI,SAT,SUN,MON,TUES,WE

```
<s:Shows rdf:ID="ugc_PianistThe">
 <s:time>
  <s:ShowScheduleCollection>
   <s:consistsOf>
    <s:ShowSchedule>
     <s:startTime>
      <c:Calendar>
       <c:calendarDate>
        <c:Date>
         <c:dateDayOfWeek rdf:resource="cal#Thursday" />
         <c:year>2003</c:year>
         <c:month>1</c:month>
        </c:Date>
       </c:calendarDate>
       <c:calendarTime>
       <c:Time>
        <c:format rdf:resource="cal#24h" />
        <c:timeHour>20</c:timeHour>
        <c:timeMinute>20</c:timeMinute>
       </c:Time>
       ...
 <s:location rdf:resource="cinemas#ugc" />
 <s:description>Certificate: 15</s:description>
 <s:show>
 <s:CinemaPerformance rdf:ID="PianistThe">
   <s:title>Pianist, The</s:title>
 </s:CinemaPerformance>
```

UNIVERSITY of ABERDEEN

Summary

- The Semantic Web is exciting from several perspectives:
  - as a piece of computing science technology
  - as a "new generation" for the Web
  - as a platform for diverse kinds of applications
- It's still the Web we know and love:
  - it co-exists with all our messy HTML, etc data
  - it's a global system: URIs are <u>universal</u>!
  - it's extremely open
  - it's not <u>too</u> hard to get started...

- IT'S WORTH GETTING INVOLVED!

Credits & Links

- Work done at Aberdeen in collaboration with
  - Agentcities & Granite Nights: Gunnar Grimnes, Pete Edwards, Stuart Chalmers
  - MusicFriends: Robin Campbell
- URIs:
  - W3C Semantic Web: http://www.w3.org/2001/sw/
  - RDF & apps: http://www.w3.org/RDF/
  - FOAF: http://www.foaf-project.org/
  - FOAFnaut: http://jibbering.com/foaf/
  - RDF-vCard: http://www.w3.org/TR/2001/NOTE-vcard-rdf-20010222/
  - MusicBrainz: http://www.musicbrainz.org/MM/
  - Jena toolkit: http://www.hpl.hp.com/semweb/
  - Granite Nights: http://www.csd.abdn.ac.uk/research/AgentCities/compo/
  - AKT 3store: http://triplestore.aktors.org/demo/AKTiveSpace/