Using the Semantic Web

Leo Sauermann
leo.sauermann@dfki.de
http://www.dfki.de/~sauermann/
God morgen!
Leo Sauermann

- Researcher on Semantic Desktop since 2002, Vienna
- At DFKI since 2004
- Currently at the Nepomuk EU project
- Scientist: workshops, lectures, students, diploma theses
- Community: open source, developer, blogger, SWEO W3C
- Training: workshops, consulting, talks

leo.sauermann@dfki.de
http://www.dfki.de/~sauermann/
The Structure of DFKI

Key Directors
Prof. W. Wahlster (CEO)
Dr. W. Olthoff (CFO)

Administration and Services
Dr. H.-J. Bürckert

PR and Corporate Communication
R. Karger, M.A.

Research Lab
Image Understanding and Pattern Recognition
Prof. Th. Breuel

Research Lab
Knowledge Management
Prof. A. Dengel

Research Lab
Deduction and Multiagent Systems
Prof. H. Uszkoreit

Research Lab
Intelligent User Interfaces
Prof. W. Wahlster

Institute for Information Systems (IWI)
Prof. P. Loos

Project Office
Prof. N. Reithinger

Robotics and Intelligent Technical Systems (RITS)
Prof. B. Krieg-Brüggen

Safe and Secure Cognitive Systems (SKS)
Prof. D. Zuehlke

Center for Human-Machine-Interaction (ZMMI)
Prof. D. Zuehlke

Transfer Center

Living Labs

Competence Centers

CCSW

KL
SB
HB
B
DFKI is based in Kaiserslautern, Saarbrücken, Bremen and Berlin

Saarbrücken Site
DFKI GmbH
Campus D32
Stuhlsatzenhausweg 3
D-66123 Saarbrücken
Tel.: +49 681 302 5252
Fax.: +49 681 302 5341
email: info@dfki.de
www.dfki.de

Kaiserslautern Site
DFKI GmbH
Trippstadter Straße 122
D-67663 Kaiserslautern
Tel.: +49 631 20575 0
Fax.: +49 631 20575 503
email: info@dfki.de
www.dfki.de

Bremen Labs
DFKI GmbH
Robert-Hooke-Straße 5
D-28359 Bremen
Tel.: +49 421 218 64100
Fax.: +49 421 218 64150
email: info@dfki.de
www.dfki.de

Project Office Berlin
DFKI GmbH
Rosenstraße 2
D-10178 Berlin
Tel.: +49 30 243102 101
Fax.: +49 30 243102 22
email: info@dfki.de
www.dfki.de
in a changing world...

You are happy driving a steam wheeler. but what when others start driving cars... ?
So why do people use Semantic Web?

- I did some research on SWEO case studies

Key Benefits of Using Semantic Web Technology

- The ability to integrate diverse source of data
- More flexibility in being able to incorporate additional unanticipated data sets
- The ability to be able to perform targeted queries or to be able to navigate through the relationships in the data
- The capability of viewing all data relating to entities of interest, even if the information available for each entity differs
## The Key Benefits of Semantic Web

<table>
<thead>
<tr>
<th>benefit</th>
<th>description</th>
<th>priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>profit</td>
<td>customer satisfaction, shareholder value, user work support</td>
<td>21</td>
</tr>
<tr>
<td>data integration across sources</td>
<td>A consistent data model to build upon. Integrate content from different organizations/providers/departments, disparate data sources, legacy data</td>
<td>13</td>
</tr>
<tr>
<td>expressive queries</td>
<td>i.e. sparql</td>
<td>11</td>
</tr>
<tr>
<td>taxonomy multilingual</td>
<td>Software uses the terminology of the user, vocabulary, synonyms, multilinguality</td>
<td>9</td>
</tr>
<tr>
<td>accessible for users</td>
<td>Generate website make data accessible to users</td>
<td>8</td>
</tr>
<tr>
<td>navigation, related items, <strong>serendipity</strong></td>
<td>Discover hidden information, navigate across data boundaries, &quot;what works with what&quot;</td>
<td>8</td>
</tr>
<tr>
<td>extensible metadata</td>
<td>Incorporate unanticipated data sets, new requirements do not change code, dynamic adaptability</td>
<td>5</td>
</tr>
</tbody>
</table>
# The Key Benefits of Semantic Web

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Export knowledge: existing data can be repurposed, unprecedented</td>
<td>4</td>
</tr>
<tr>
<td>use cases, conserve culture standards</td>
<td></td>
</tr>
<tr>
<td>Semantics: inference owl reasoning formal</td>
<td>4</td>
</tr>
<tr>
<td>Simple: cheap implementation lightweight</td>
<td>2</td>
</tr>
<tr>
<td>Natural:</td>
<td>2</td>
</tr>
<tr>
<td>Personalization:</td>
<td>2</td>
</tr>
<tr>
<td>Automation:</td>
<td>2</td>
</tr>
<tr>
<td>Explanation:</td>
<td>1</td>
</tr>
<tr>
<td>Improve content quality consistency</td>
<td>1</td>
</tr>
<tr>
<td>Web of linked open data</td>
<td>1</td>
</tr>
<tr>
<td>Unique id</td>
<td>1</td>
</tr>
</tbody>
</table>

This will grow
Open linked data

~ 2,000,000,000 facts
~ 680,000 links

• http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData

... more people with similar idea

Google Social API + FOAF

unique id
March 13, 2008
The Yahoo! Search Open Ecosystem

In the coming weeks, we'll be releasing more detailed specifications that will describe our support of semantic web standards. Initially, we plan to support a number of microformats, including hCard, hCalendar, hReview, hAtom, and XFN. Yahoo! Search will work with the web community to evolve the vocabulary framework for embedding structured data. For starters, we plan to support vocabulary components from Dublin Core, Creative Commons, FOAF, GeoRSS, MediaRSS, and others based on feedback. And, we will support RDFa and eRDF markup to embed these into existing HTML pages. Finally, we are announcing support for the OpenSearch specification, with extensions for structured queries to deep web data sources.

applications
from SWEO case studies
and foafs
Vodafone live!

Problem:
• Vodafone is one of the largest mobile carriers
• **Content** for portal *Vodafone live!*
• Multiple *content* providers

Solution
• *Vodafone live!* publishes RDF vocabulary
• Content providers conform
• Content metadata is delivered in RDF

Project:
• **Content** ringtones, games, pictures
• **Metadata** phone platforms, ratings adult|violence, availability
• exact documentation, validation of XML
• key is involvement of content providers

*Courtesy of Kevin Smith, Vodafone Group R & D (SWEO Use Case)*
Vodafone live!

vocabularies
- Dublin Core
- **PRISM** for embargos and availability
- **Internet Content Ratings Association (ICRA)** for violence, erotic, gambling
- extensions

- user enters profile at portal, appropriate content delivered
- significant increase in content download after the introduction

*Courtesy of Kevin Smith, Vodafone Group R & D ([SWEO Use Case](#))*
Semantic Desktop: Problem of Personal Information Management

- Problem: **Data** about single ideas are stored in several applications and without context

- People have **mental models** of their environment
NEPOMUK KDE 4.0

~ 1-8 million users
~ 1k developers
integrate the desktop data

11. January 2008: release v4.0

?! use RDF

• convert all data to RDF, index it
• allow user to annotate with tags, ratings, etc
• allow applications to annotate
NEPOMUK KDE 4.0

- Soprano indexing, Dolphin file explorer

http://nepomuk.kde.org/
Convert Leo from Address-Book to RDF

Microsoft Outlook

<table>
<thead>
<tr>
<th>Name</th>
<th>Leo Sauermann</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>Scientist</td>
</tr>
<tr>
<td>Eirma</td>
<td>DFKI</td>
</tr>
<tr>
<td>Speichern unter</td>
<td>Sauermann, Leo</td>
</tr>
</tbody>
</table>

PersonContact

is a

desktop://outlook/contact/ECD4B993

fullname

“Leo Sauermann”

<rdf:RDF
 xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
><nco:PersonContact rdf:about="desktop://outlook/contact/ECD4B993">
 <nco:fullname>Leo Sauermann</nco:fullname>
</nco:PersonContact>
</rdf:RDF>

done with http://aperture.sourceforge.net
since ~2000 RDF is used by plugins to configure how they extend the mozilla platform

http://developer.mozilla.org/en/docs/Deploying_a_Plugin_as_an_Extension

<RDF xmlns="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
<Description about="urn:mozilla:install-manifest">  
<em:id>RhapsodyPlayerEngine@rhapsody.com</em:id>  
<em:name>Rhapsody Player Engine</em:name>  
<em:version>1.0.0.487</em:version>  
<em:targetApplication>  
<Description>  
<em:id>{ec8030f7-c20a-464f-9b0e-13a3a9e97384}</em:id>
<em:minVersion>1.5</em:minVersion>
<em:maxVersion>1.5.0.*</em:maxVersion>
</Description>
</em:targetApplication>
</Description>
</RDF>
Mozilla?
Semantic Web has been in use for so long?
Many applications

Changes
• Version 1: http://.../1
• Version 2: http://.../2

Index of pages
• Page 1: http://.../1
• Page 2: http://.../2

• list of resources
• title
• author
• date
• content

http://web-2-0.gemzies.com/show/entry_3402/RSS_logo.html
Linda T
key learnings from W3C SWEO case studies

Vodafone

- Use an open, flexible model
- Research existing vocabularies and reuse industry metadata standards where possible.
- Provide full documentation and a schema (RDF or XML) for validation and IDE support.
- Ensure appropriate stakeholder support

Sun and Zepheira

- Start with, and maintain, an investment in data integrity and process automation
- Provide flexible and extensible administration tools
- Ensure appropriate stakeholder involvement and ownership
- Consider program name selection – choose a name that is simple to remember and spell, is not similar to other program or project names within that organization, and does not require explanation
- Produce an accurate definition and wide communication of the technology's success metrics
make your own semantic web sites
Its easy!

“Planning a Semantic Web site”
Prepare your site for structured data
Rob Crowther
04 Mar 2008
http://www.ibm.com/developerworks/
  xml/library/x-plansemantic/index.html

look on del.icio.us for
“SWEO tutorial”
DFKI Competence Center Semantic Web

• Contact us to get connected to researchers within DFKI
  – Ontology creation and maintenance
  – Support for annotating and analysing data
  – Research on demand
  – Consulting, teaching, training

http://semanticweb.dfki.de

Michael Sintek
michael.sintek@dfki.de

Paul Buitelaar
paulb@dfki.de

+ Leo
mash it up!

</SemanticWeb>
</rdf:RDF>

Leo Sauermann

leo.sauermann@dfki.de
http://www.dfki.de/~sauermann/

Thanks to Jesus, Ingrid Brunner-Sauermann, Heiko Maus, Sven Schwarz, Ansgar Bernardi, Andreas Dengel, Frank Osterfeld, Dominik Heim, Man Luo, Jeen Broekstra, Giovanni Tumarello, Michael Zeltner, Stephan Baumann, Gunnar A Grimnes, Ludger Van Elst, Harald Holz, Stefan Decker, Malte Kiesel, und alle anderen auf #swig
Semantic Web is part of a changing world

(as is art)

Perception
Business Models

- Texts
- Databases
  - Mashups
  - Taxonomies
- Web sites
  - Search
  - Navigation
- Social networks
- B2B
  - Exchange
  - Knowledge
- Cultural data

Vocabularies
- Search
- Data integration
- Serendipity

Group Me!

GoPubMed

ReVu!.com

Vodafone Live!
The New York Times and the Los Angeles Times tried charging for some online content, then abandoned the practice. Not so Consumer Reports, which has charged Internet readers the same price as print subscribers for a decade, currently $26 a year (or $5.99 for a month’s online access or $45 a year to get the magazine both in print and on the Web). While the rest of the industry sees print readers as more valuable—because advertisers do—Consumer Reports actually makes more money from readers on its Web site, because it avoids printing, trucking, and mailing costs.

Contextually relevant suggestions of links, pictures, related content and tags will make your blogging fun again.

Similar thing as Reuters did launch January 2008 with OpenCalais.com but with pictures, links, tags and users