Welcome to the Semantic Desktop Hack

Leo Sauermann
DFKI GmbH
April 2007
Berlin
Leo Sauermann

• Work on Semantic Desktop since 2002
• Researcher at DFKI since 2004
• Currently at the **Nepomuk EU** project
• University Lectures on Semantic Web, Students and Diploma theses
• Open Source developer, blogger
• Sci-Workshops, Trainer, Consultant

Leo Sauermann

[Contact Information]

leo.sauermann@dfki.de
http://www.dfki.de/~sauermann/
How did we manage?

photos from flickr pool everyday informationarchitecture, users: Ariel van Spronsen, typedown, ludita, el shack del hombre
Das Problem ist zu schwierig für einen einzelnen Mann. Ich werde mein Elektronengehirn befragen.

“What did they offer us?”

“This problem is too big for one lone man. I have to consult the computer brain”
Personal Information Management

• Definition by Richard Boardman (PhD, 2004)
  PIM can be defined as the management of personal information (information owned by an individual, and under their direct control) as performed by the owning individual.

• It's all about appointments, e-mails and contacts
• massive amounts of data – disks are cheap
• across different applications
• each application offers a separate way of categorization
People have **mental models** of their environment.

It's all about Projects, People, Goals, Structure.

Documents are related to concepts from the model.
Offered Solutions

PIM Software
• def: PIM = Microsoft Outlook®
• other applications connect with it
• or with Lotus Notes®, …® …®

Information Management Software
• Mindmanager by Mindjet
• Wikis

Standards
• WinFS – data in the filesystem
• XTM, RDF, OWL – meta-data standards
Semantic Web

- Extension of the current Web
- The Web is a database
- Common Framework
  - RDF
    Resource Description Framework
  - to describe resources
  - to describe my files?
- Across Applications
  - that could mean: across all my applications?
- RDFS, OWL
  - to describe how to describe resources
- Ontologies
  - to describe my mental model?
- Inference Engines deduct Information
  - suddenly we know more about the described resources

http://www.w3.org/2001/sw/

The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries. It is a collaborative effort led by W3C with participation from a large number of researchers and industrial partners. It is based on the Resource Description Framework (RDF), which integrates a variety of applications using XML for syntax and URIs for naming.
Break: Real problems leading to a Semantic Desktop

- PIM is **not** e-mails/address book/calendar
- A desktop is **not** a bunch of applications
- We have **mental models**
- Our concepts show in every application we use, are managed across the whole desktop

- If PIM is the main use of a PC – why is it not defined in the operating system?
- If the Semantic Web could do that on a global scale, it may be useful for one desktop?
The leap we make

• Semantic Web as enabling technology

• The data swamp on a single PC

• Semantic Desktop
Goals of a Semantic Desktop

- Managing all personal information across application borders
- Based on Semantic Web standards
  - URI
  - RDF
  - RDF/S + Ontologies
  - HTTP
- Represent the mental model of the user
- Semantic communication and collaboration
Definition

A Semantic Desktop is a device in which an individual stores all her digital information like documents, multimedia and messages. These are interpreted as Semantic Web resources, each is identified by a Uniform Resource Identifier (URI) and all data is accessible and queryable as RDF graph. Resources from the web can be stored and authored content can be shared with others. Ontologies allow the user to express personal mental models and form the semantic glue interconnecting information and systems. Applications respect this and store, read and communicate via ontologies and Semantic Web protocols. The Semantic Desktop is an enlarged supplement to the user's memory.

The semantic desktop is an integrated, coherent whole, build on many parts

- Personal Information Management - PIM
- Personal Information Model – PIMO
- Document classification, analysis, NLP
- Data Integration
- Semantic Search
- Personal Semantic Wikis
- Task Management, Work Context Detection
- Organizational Memories
- Artificial Intelligence assistance systems
- Collaboration, Peer-To-Peer, semantic e-mail
- Semantic Web 2.0 applications
- great user interfaces

- that’s a lot. but its better than doing all that separated.
Phases towards the Social Semantic Desktop

Today necessary technologies & communities exist:
Standardized metadata: Semantic Web
Scalable distributed infrastructure: P2P Computing
Knowledge articulation and interaction: Desktop Technology
Human centric information exchange: Online Social Networks

Challenge: Extension & merging of research streams

Phase 1
- Desktop/Wiki
- Semantic Web
- P2P networks

Phase 2
- Semantic Desktop
- Semantic P2P

Phase 3
- Social Semantic Desktop
- Ontology driven distributed Social Networking
- Social Networking

Stefan Decker and Martin Frank 2004: The Social Semantic Desktop
Personal Information Model PIMO

- represents **mental model** of an individual person
- formal: classes, individuals, properties using RDF/S
- upper classes are predefined: *person, project, organization, location, time*
- own classes and things can be created *my visit to Dublin, Dublin, Claudia,…*
- documents (e-mails, websites) are “tagged” using defined *things*
- all data should be integrated into the PIMO, facts and relations stored here for cross-application reuse
Example from the PIMO

@prefix leo: <gnowsis://sauermann.dfki.de/pimo>.
@prefix pimo: <http://ontologies.opendfki.de/repos/ontologies/pim/pimo#>.

leo:iswc2006 a pimo:Event;
    rdfs:label "ISWC2006";
    pimo:occurrence
        <imap://sauermann@mail/mail%2fConferences%2fISWC2006%2fSemDeskWS2006;TYPE=LIST>
    pimo:hasPart leo:semdeskws2006.
Middleware to integrate Data

- Extraction of Data from source applications to be integrated into a central RDF database
- Middleware transforms application data to RDF
- Search and data-integration works then on the crawled data
- Personal data-warehouse
- Target is the PIMO
All is RDF

• Example: me

<http://leo.gnowsis.com/outlook/contact/0000000ECD4B99358B9814B9DAFE2255CD8AE9A048F2900>
a foaf:Person;
  foaf:currentProject <http://www.gnowsis.org/>;
  foaf:family_name "Sauermann";
  foaf:firstName "Leo";
  foaf:workInfoHomepage <http://www.dfki.de/~sauermann>;

vcard:address
[ a vcard:AddressPostal ;
  vcard:streetAddress "Pirmasenser Strasse 18"
  vcard:locality "Kaiserslautern" ;
  vcard:postalcode "67655" ;
  vcard:country "Deutschland" ;
]
  kiss:kissed

<http://leo.gnowsis.com/outlook/contact/0000000ECD4B99358B9814B9DAFE2255CD8AE9AC4612000>;
Rules and Artificial Intelligence

• Extract the contents of an invitation e-mail and infer the location and attendees of the meeting

• Dynamic ToDo-Lists: What I have to do is determined by rules that analyze e-mails, contacts, projects, personal settings, …

• Query expansion in search

• Document classification, pattern recognition, training of assistance systems

Semantic Meeting Annotation, Man Luo, Diploma Thesis 2006
Example for a personal semantic web
Wiki: From text to information

• Free text is easy: to understand, to write
• Free text is everywhere: comments, notes, documents, websites, blogs, e-mails, …
• Free text is the basis for semantic wikis

Nepomuk
Nepomuk is an EU project. It’s topic is the Social Semantic Desktop. DFKI coordinates.
Personal Semantic Wiki is the key for note-taking

- Terms
- Types
- Relation types
- Relations
- RDF

We can use a personal semantic wiki everywhere: comments, notes, documents, websites, blogs, e-mails, …
Personal Semantic Desktop Wikis

• Malte Kiesel, Leo Sauermann
  *Towards Semantic Desktop Wikis*

• David Aumueller
  *Towards a Semantic Wiki Experience – Desktop Integration and Interactivity in WikSAR*

• Eyal Oren
  *SemperWiki: a semantic personal Wiki*

• Jörg Richter, Max Völkel, Heiko Haller
  *DeepaMehta – A Semantic Desktop*

• (and more…*)
Document understanding for the Personal Information Model

- Documents can be analyzed to get related topics, people, projects, etc.
- The user trains this analysis over time, personalizing the system
- The PIMO is used and trained during normal document work

→ like your SPAM filter

1. training with existing data
2. incoming e-mail
3. suggestions based on training
4. classification & training

Andreas R. Dengel: Six Thousand Words about Multi-Perspective Personal Document Management, EDM 2006
Theory

Prototype

http://flickr.com/photos/eti-eti/114554777/
SeMouse

• Annotation of documents using the middle mouse button
• Works with Word, Browser, PDF, …
• Can be integrated with other frameworks (gnowsis)
• Sergio F. Anzuola
Open IRIS

Re: irisdock (11/3/05 11:45 AM)

Subject: Re: irisdock
From: Jim Carpenter <james.carpenter@iri.com>
Date: 11:45 AM
To: Rich Schofield <rich@schofieldnetworks.com>
Cc: Adam Cheyer <adam.cheyer@iri.com>

Rich,

I think that looks great. I can’t wait to get the docking framework into Iris. I am not sure if we want the Search Panel/Toolbar areas to be a dockable window, should also break apart the context service panel into summary, connections, suggestions and history. Is the status bar the same as the existing notifications area? Should we move the notifications below the app tree on the left side?

Jim
What connects Jerome Euzenat with Ireland

Multiple Data Entry Options
Prototype

Community

http://flickr.com/photos/eti-eti/114554777/
Web resources and community

• Community resources
  – www.semanticdesktop.org community website with wiki and many projects
  – people@semanticdesktop.org mailinglist

• Scientific events
  – ISWC workshop 2005
  – ISWC workshop 2006
  – i-Semantics Special Track 2007
  – ESWC workshop 2007
Two more articles – shameless self advertising

• read a book
  Kapitel Semantic Desktop - Der Arbeitsplatz der Zukunft

• one good overview
  – Overview and Outlook on Semantic Desktop, Leo Sauermann, Ansgar Bernardi, Andreas Dengel

• Nepomuk publication list (many)
  – http://nepomuk.semanticdesktop.org
Social Semantic Desktop Science Fiction

UI group
fancy browser/editor

Ontology matching scientists

standard interfaces

EAI group adapters

p2p cscw approach

PIM support algorithm
Summary

- PIM is **not** e-mails/address book/calendar
- The Semantic Web provides a common framework that allows data to be shared and reused across applications
- The Semantic Desktop manages all my information based on standards
- Everything has a URI, is transformed to RDF, is interlinked within a Personal Information Model
- This needs research and development in many areas, from user interfaces down to databases
- You are all interested doing this
mash it up!
questions?

</SemanticDesktop>
</rdf:RDF>

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

persona created using www.sp-studio.de
Thanks to Ingrid Brunner-Sauermann, Jean Rohmer, Martin Memmel, Heiko Maus, Sven Schwarz, Ansgar Bernardi, Andreas Dengel, Frank Osterfeld, Dominik Heim, Man Luo, Jeen Broekstra, Giovanni Tummarello, Michael Zeltner, Stephan Baumann, Gunnar A Grimnes, Ludger Van Elst, Harald Holz, Stefan Decker, Malte Kiesel, and the others