Towards the Semantic Desktop

Dr. Øyvind Hanssen
University Library of Tromsø

Agenda

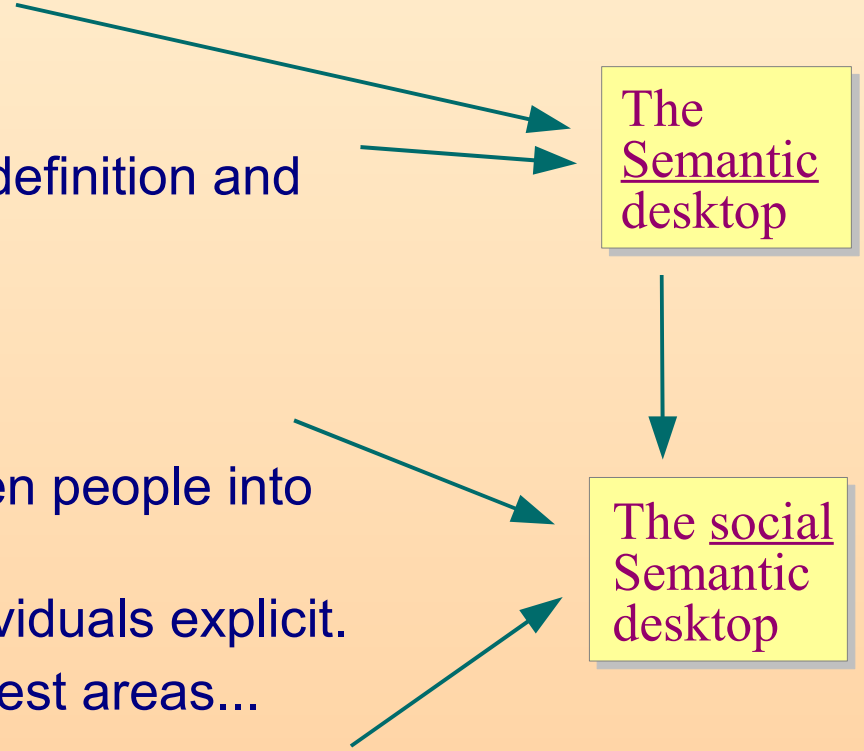
- Background
- Enabling trends and technologies
  - Desktop computing and The Semantic Web
  - Online Social Networking and P2P Computing
- Towards the semantic desktop
  - Definition
  - Challenges
  - Metadata management
- Role of academic libraries

Background

- Information overload
- Information fragmentation
  - Desktops and applications are “data silos”
- Collaboration
  - Blend boundaries between personal and group data
- Information management
  - “Low level” communication
  - Insufficient metadata support
  - Need to interconnect (relate) separate data items.
- Vannevar Bush, Doug Engelbart, T.B. Lee
  - Visions that could not be realised at the time. Now they can!

Towards a social S.D.

- Desktop computing
- The semantic web
  - Standards and technologies for definition and exchange of metadata..
  - Ontologies.
- Online Social networking
  - Map “social” connections between people into technical infrastructure.
  - Make relationships between individuals explicit.
  - Virtual communities around interest areas...
- Distributed and P2P computing
  - Data and computation sharing without centralised infrastructures and centralised control.



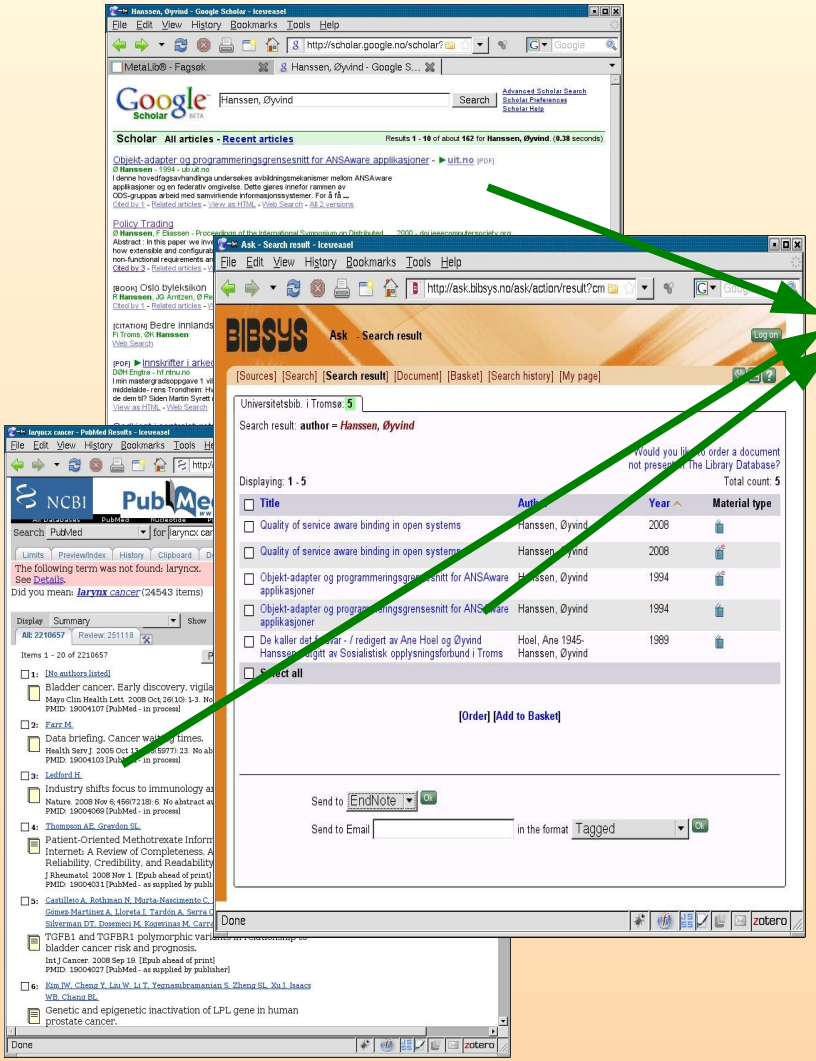
The Semantic Web

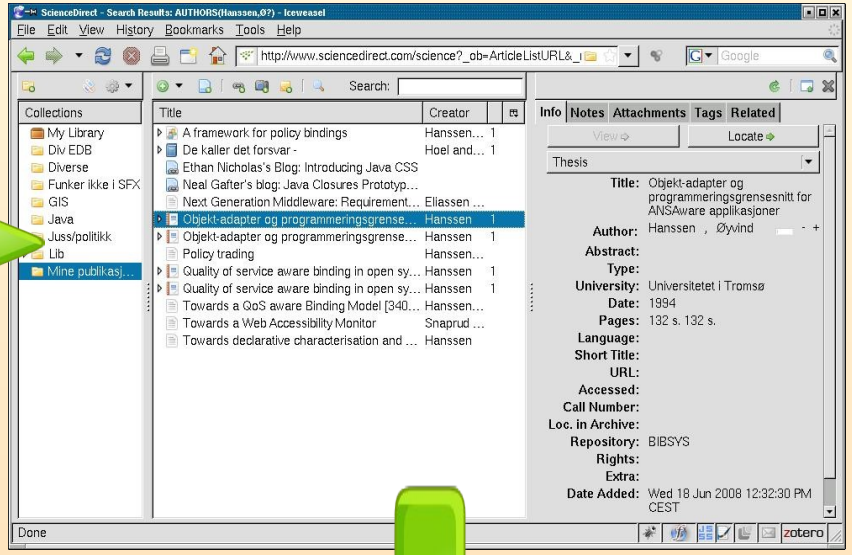
- Envisioned by Tim Berners Lee
- Web content understandable by computers..
  - Computers help find, share, and combine information items.
  - Web pages designed to be readable by humans..
  - Allow automated reasoning.
- Part of Web 3.0
- It is all about metadata
  - HTML markup
  - Resource Description framework (RDF) – a data model for the s.w.
  - Vocabularies, ontologies (conceptual models / schemas)
    - Class hierarchies, relationships, rule-bases

On the left, what browsers see. On the right, what humans see. Can we bridge the gap so ~~browsers~~ computers see more of what we see?

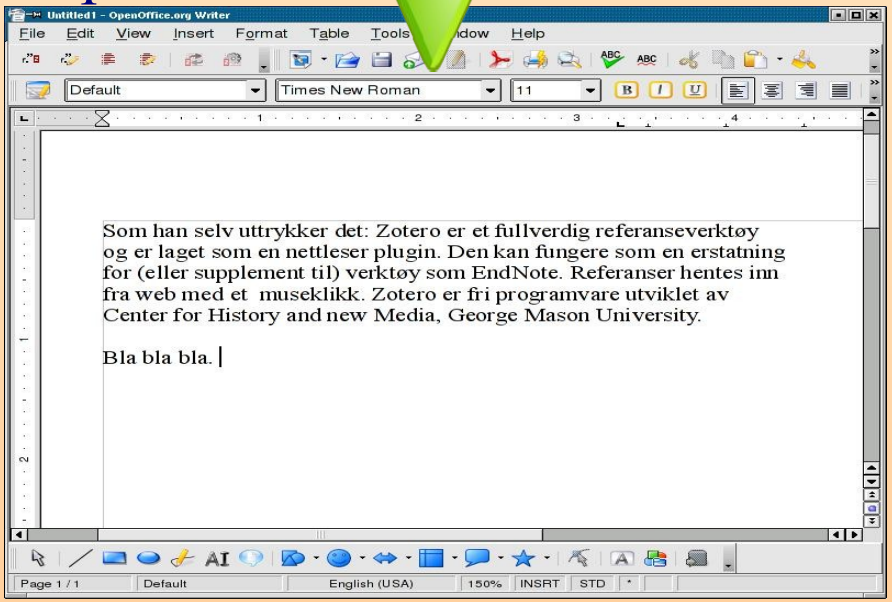
Hentet fra: RDFa Primer – Bridging the Human and Data Webs , *W3C Working group note*, 14 october 2008

Reference management app.





Word processor



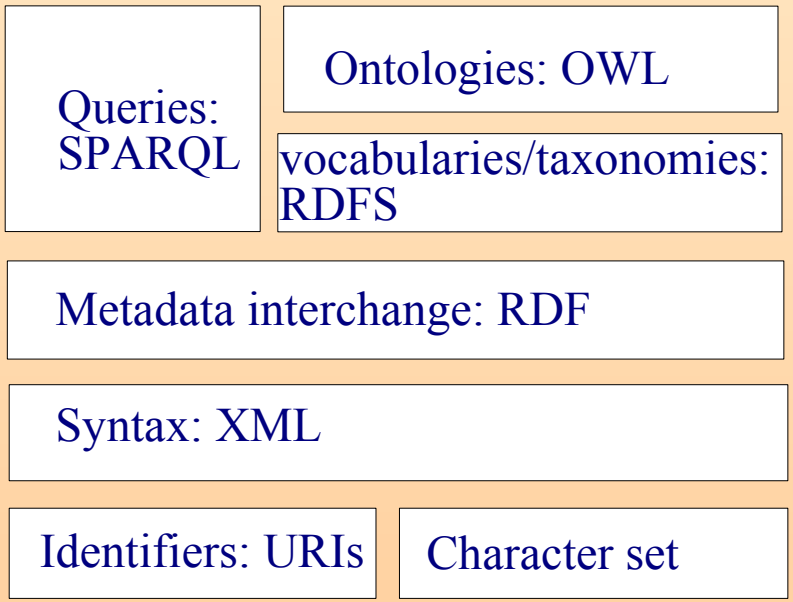
Documents and (bibliographic) metadata on the web

Architecture

User applications

Trust, proof

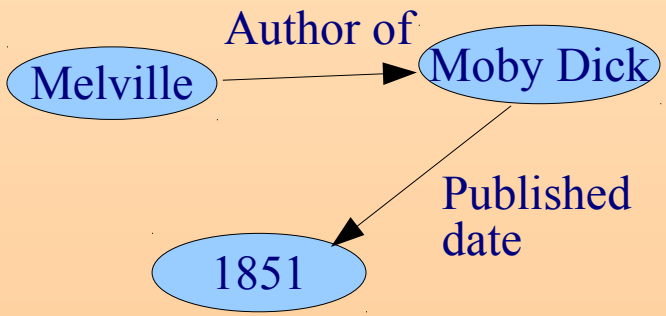
Unifying logic



RDF-metadata are triples:

<subject, predicate, object>

Elements can be appear multiple
Triples → graph ..



P2P computing

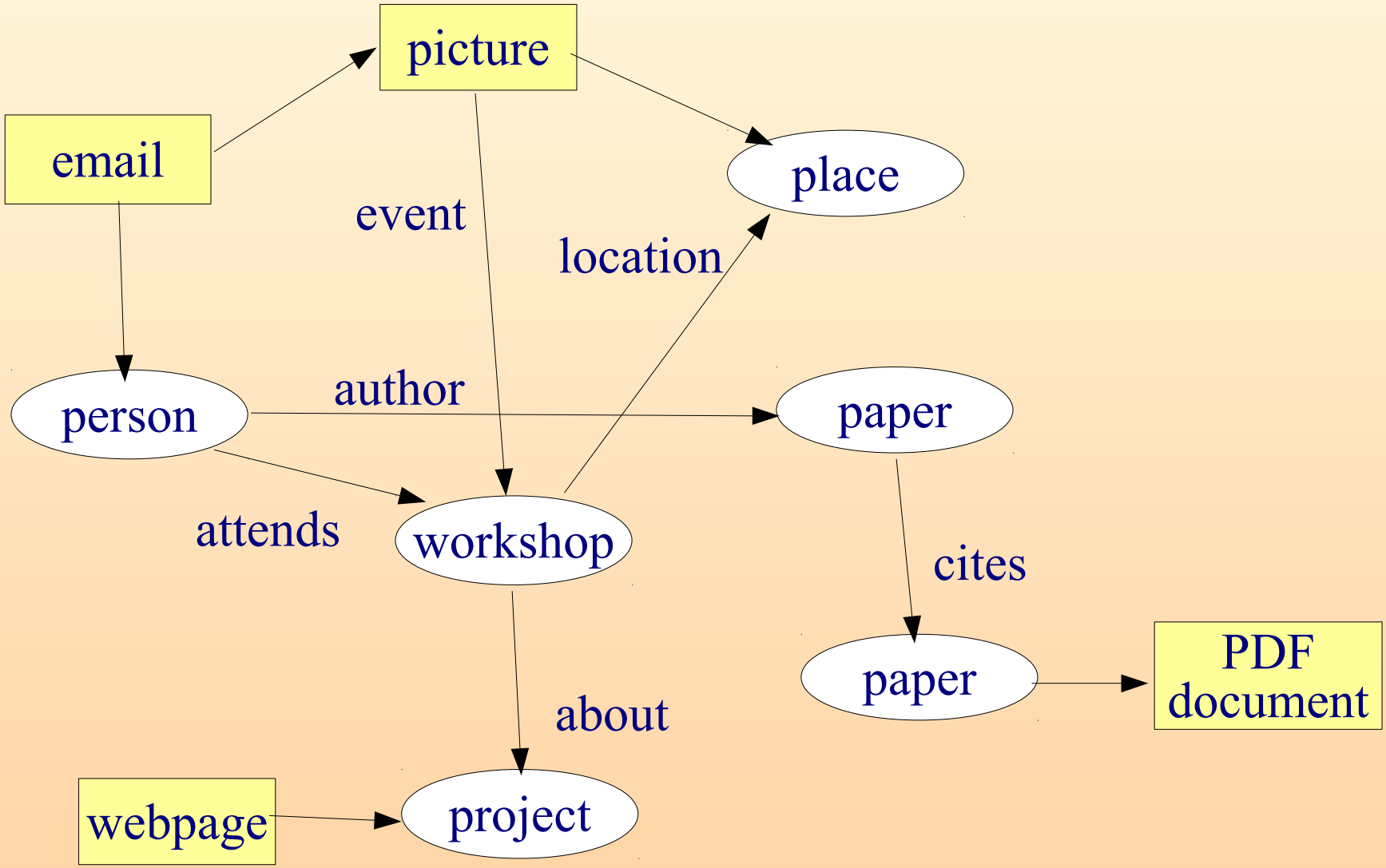
- Different definitions in litterature
  - Strictest: Totally distributed system in which all nodes are completely equivalent
  - "...class of applications that take advantage of resources ... available at the edges of the internet" (Shirky, 2000)
  - "...the sharing of computer resources and services by direct exchange between systems" (Milojicic et.al, 2002)
  - "... interconnected nodes able to self-organize into network topologies with the purpose of sharing resources ... capable of adapting to failures... without requiring the intermediation or support of a global centralized server or authority" (Androutsellis & Spinnellis, 2004)

Online Social Networking

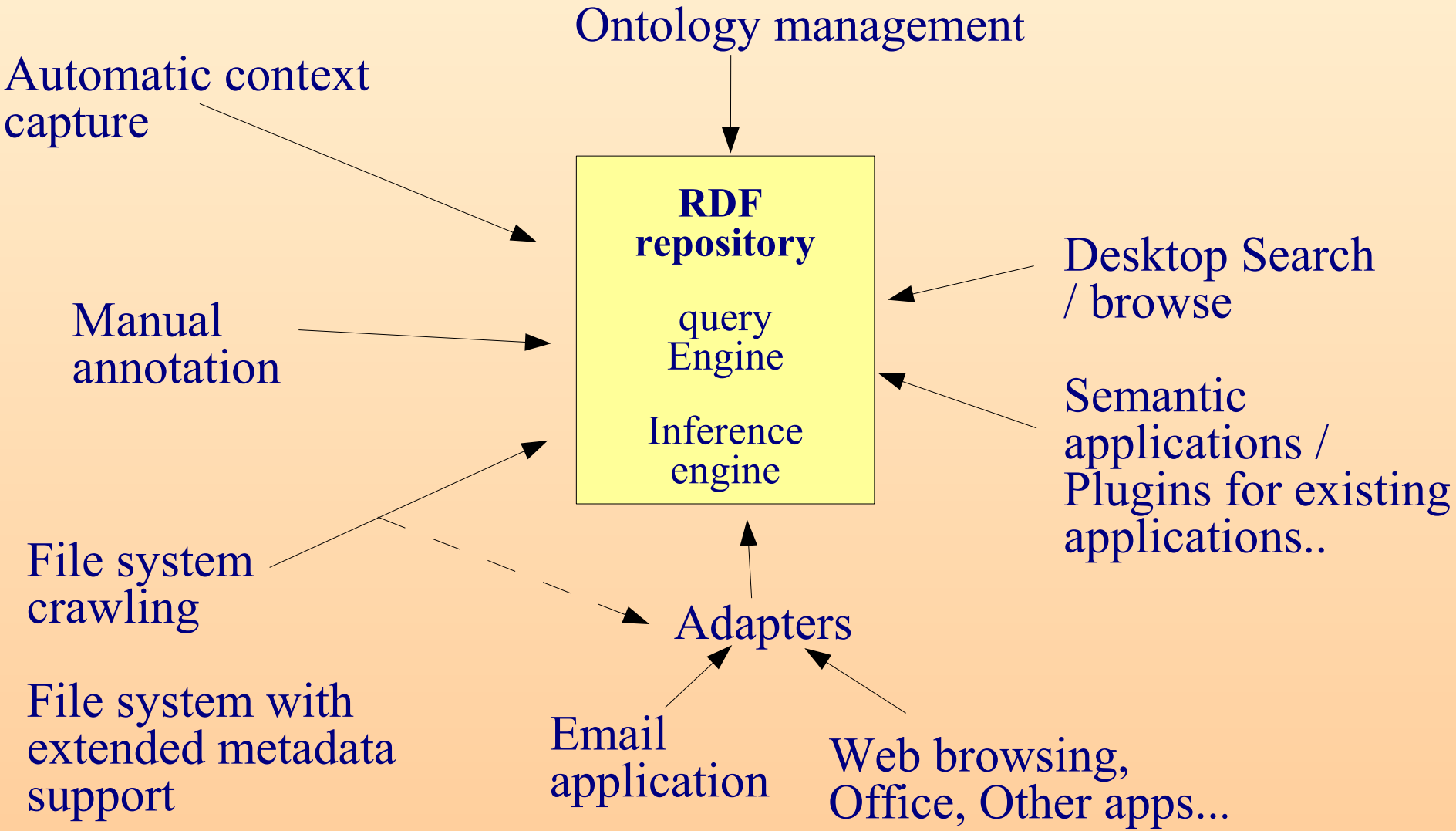
- Support social relationships both for personal and professional use
- Current solutions: Serious issues
  - Limited use
  - Requires major investment, capitalism
  - Sites are unlikely to share information
  - Privacy and information ownership issues
- Build a social networking infrastructure on top of a P2P system, and based on semantic web technology?
  - User information owned by individual users!

Semantic Desktop

- “If the goal is to have a global Semantic Web, one building block is a Semantic Desktop, a Web for a single user”. (Sauermann et al. 2005)
- “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 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. Application 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”. (Sauermann et. Al. 2005)
- Research, and reference architecture:
  - Gnowsis, NEPOMUK


picture
email
place
event
location
person
author
paper
attends
workshop
cites
paper
PDF
document
about
webpage
project

Semantic desktop



Some issues...

- How to capture context of a resource
  - Context of the user. What is the user doing? Context may switch
  - Connect to ontologies. Context ontologies.
  - Application independence
- Ontology mapping and conflict resolution (semantic interoperability)
  - Different persons, institutions, applications, domains
  - Personal vs. domain, etc...
  - Addressed in e.g. (Cruz, 2008).
- Quality of metadata and their ontologies
- Applications and user interface, middleware architecture...

Metadata management

- Ontology
  - Vocabulary and conceptual model
    - Terms, classification-hierarchies, relationships between classes.
    - Rules: Not integrity rules like in databases, but rather to decide what a “thing” is. Define semantics..
- Ontology management
  - Create or extend (define concepts and rules)
  - Tools to check consistency, visualisation etc.
- Ontologies for the desktop – specific to..
  - Domain, application, subject, person, group, institution.
  - One ontology may use another..
  - Layered architecture..

Metadata management

"...but it currently appears that the Semantic Web will rely extensively on human interpretation and judgement to bring metadata into conformance with the ontology, and in fact, to derive and extend the ontology in the first place."

"...establishing trust – that the metadata is a good and consistent representation of content representation of content for the use to which it is put – will be a challenge"

(C.C. Marshall, 2003)

Academic Libraries

- Research librarians
  - Have expertise in research areas.
  - Have bibliographic expertise.
  - Can be important resources in metadata management for semantic desktops.
  - Help in ontology specification and mapping.
  - “Professional counsellor on personal information management” (E. H. Dow, 1987).
- Infrastructure support
  - Institutional and group ontologies?
  - Computational support?

Conclusions

- Trends
  - Personal computer can store your whole “life”
    Emerging infrastructure to manage personal information.
  - More direct (and possibly ad hoc) information sharing between personal computers.
  - Blurred boundaries: Personal, institution, social, ...
- Library role → focus on metadata management?
  - Standardisation, mapping, management of ontologies
  - Metadata quality assurance
  - Provide some of the technical infrastructure?
- This is an active research area
  - Computer science and library science

Litterature

- S. Androutsellis-Theotokis, D. Spinellis, A Survey of Peer-to-Peer Content Distribution Technologies, *ACM Computing Surveys, Vol. 36, No. 4, December 2004, pp. 335-371*
- T. Berners-Lee, J. Hendler, O. Lassila, The Semantic Web. *Scientific American, May 17, 2001,* pp. 34-43.
- V. Bush, As we may think, *The Atlantic Monthly 176(1), 1945,* pp. 101-108.
- I.F. Cruz, H. Xiao, A layered framework supporting personal information integration and application design for the semantic desktop, *The VLDB journal, 17, 2008,* pp. 1385-1406, Springer.
- S. Decker, M. Frank, *The Social Semantic Desktop, DERI Tehnical Report 2004-05-02, 2004.*
- E. Dow, Personal Information Systems: the Library Role, *Library Journal, 112(18),* November 1987.
- C.C. Marshall, F.M. Shipman, Which Semantic Web?, In *proc. HT´03*, ACM, 2003.
- L. Sauermann, A. Bernardi, A. Dengel, Overview and Outlook on the Semantic Desktop, In *proc. ISWC Semantic Desktop Workshop*, 2005.