

SemRA Proposal

**Portal Coordination Meeting
Amsterdam, 14. March 2002**

Peter Gietz

Peter.gietz@DAASI.de



Agenda



Current WWW

- **Mere publishing medium**
- **Huge amount of information**
- **Design for human access only**
- **Lack of structure and organization**
- **Insufficient access methods**
- **Ambiguous:**
 - bank (finance institute) the same as
 - Bank (river bank)

DAASI
International

Directory Applications for
Advanced Security and
Information Management



Visions for the future

- **Web Services**
- **Accessed by humans *and* programs**
- **Quality content**
- **Better structured**
- **Knowlegde enhanced**
- **Disambigued:**
 - **Bank (finance institute) is not the same as**
 - **Bank (river bank)**

DAASI
International

Directory Applications for
Advanced Security and
Information Management



Buzzwords for the new visions

- „Semantic Web“ (Tim Berners-Lee)
- Grid
 - Computational Grid (Foster/Kesselman)
 - Computing power out of the wall
 - Information Grid
 - Information about resources, data and the rest
 - Knowledge Grid
 - Knowledge is relations between concepts and information



How to achieve knowledge

➤ Metadata

- Data about information

➤ Ontologies

- Concepts and relations between them
- Computer knows more than inputed

Input: Parents have children

Input: Mother = female parent

Output: Mothers have children

DAASI
International

Directory Applications for
Advanced Security and
Information Management



SemRA aims

- **Semantic Relationship Architecture for the Internet**
- **Set up and experiment with an architecture that includes:**
 - Metadata crawlers
 - Multiple Ontologies
 - Automatic classification
 - Cooperation Tools
 - Intelligent Access Portal



Metadata crawlers

- **Z39.50 Crawler (legacy)**
- **Webcrawlers using different search strategies**
 - Starting at subject gateways
 - Using automatic classification for intelligent decisions which links to follow
- **OAI Harvester**
 - Open Archive Initiative
 - Promising technology



Ontologies

- **Ontology description languages**
 - DAML+OIL
 - WebOnt
- **Multiple ontologies with links between different ontologies**
- **General ontologies (e.g. GermaNet)**
- **Special ontologies (e.g. for IT and political sciences)**

Ontologie Storage

- **Common repository for metadata and ontologies based on LDAP technology**
- **Large scalability by setting up an Indexing system based on Common Indexing Protocol (CIP) and respective work of TF-LSD**
- **Ontologie data model based on CIM which provides a model for associations that can be used for the relations**

DAASI
International

Directory Applications for
Advanced Security and
Information Management



Automatic Classification

- **Based on Ontologies**
- **Statistical Decision making**
 - Simple reliable algorithms as a start
 - Experiments with advanced high-precision techniques
- **Neural Networks**
- **Machine Learning**
- **Automatic Categorisation**

DAASI
International

Directory Applications for
Advanced Security and
Information Management



Cooperation Tools

- **Tools for net communities for online content publishing including:**
 - Publications
 - Discussions
 - Metadata creation
 - Ontology creation and extension
- **Platform for communication**
- **Online editorial boards**
- **XML based: DocBook, TEI**

Intelligent Access Portal

- **Web site offering cross-server access to a set of integrated applications directed towards a specific user community**
- **Web Services the new technology**
 - SOAP
 - WSDL
 - UDDI / LDAP
- **AA**
 - SAML, PAPI



The Proposal I

- **1. Submission in FW5, 7th Call „Semantic Web Technologies“ did not succeed**
 - More or less expected since time was too short for a complete proposal
 - Evaluation said technically very good worthy of a resubmission
 - Lack of good consortium structure
 - Lack of exploitation plans

The Proposal 2

- **FW 5, 8th call „Pioneering Research“**
 - **2 new consortium members**
 - **Ontology experts**
 - **Exploitation experts**
 - **Better Structure, technical advisory board**
 - **Complete proposal**
 - **Part of the exploitation should be a follow up project that builds on the experiences from experiments with SemRA**



Future Strategies

- **Evaluation will be finished some time in April**
- **If positive SemRA will take place something like September 2002 – August 2004**
- **The call does not have a lot of money**
- **If negative there is a lot of good technology evaluation for future (=FW6) activities**

Possible cooperations for a 6th FW Proposal

**Portal Coordination Meeting
Amsterdam, 14. March 2002**

Peter Gietz

Peter.gietz@DAASI.de



FW 6

- **FW 6 will most probably be focussed on**
 - Grid computing
 - Knowledge technologies
- **FW 6 will want very big projects**

Proposal

- **Renardus Project did a very good job in the field of Subject Gateways**
- **SemRA people learned a lot about newest technologies for knowledge technologies and portal technologies.**
- **Together a good and big enough consortium can be set up**
- **Portal Coord could be the platform for the discussions**

