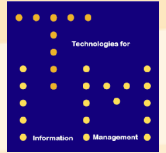


KHRESMOI

Kick-off meeting



Large data

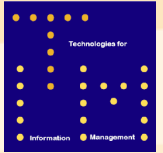
Challenge and opportunities

12 October, 2010
Wien

Werner Janusch

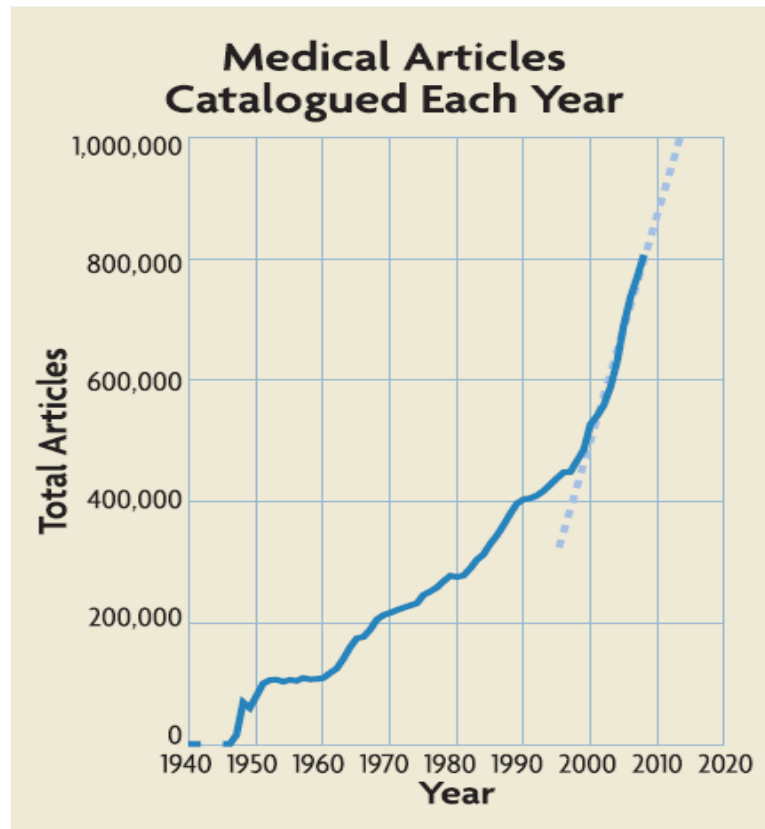
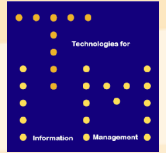
DG Information Society and Media
Unit E2 – Technologies for Information Management

Agenda



- Data growth
- Large data
- FP7 ICT Call5 Intelligent Information Management
- Other relevant projects for KHRESMOI
- Outlook
 - Digital agenda for Europe
 - New FP7 ICT workprogramme 2011-2012
 - Next calls

Information Explosion in Medicine

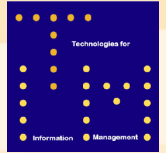


Ref.

- The Fourth Paradigm
- Data Intensive Scientific Discovery
- Microsoft Research
- 2009

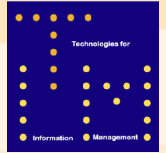


Data growth trends



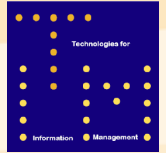
- IDC March 2008 report: by 2010 less than 50% of digital info will be authored documents
- The rest will come from
 - automated experiments (e.g. DNA sequencing)
 - ambient sources (surveillance, sensors, RFID)
 - interaction logs (web searches/services, telecom, financial transactions)
 - ...
- Need to make sense of it (**knowledge based economy**)

Data grows faster than intelligence



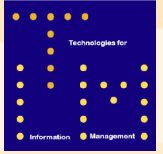
Large data

/what



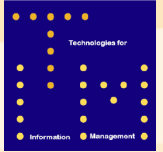
- Text
- Images
- Video
- 3D models
- Engineering Designs
- Statistical Data
- Sensor Data (remote sensing satellites, in-situ sensors, medical sensors, ...)
- Biological Data
- Medical Data
- Environmental Data
- Scientific data
- RFID data streams
- Formal Knowledge Representation
- Databases
- Logs of internet searches, telephone calls, financial transactions, etc...

Large data /challenge



- Available Large Data Processing Technologies
 - Cloud Computing, Map reduce, Hadoop, parallel processing, columns oriented databases...
- What's next?

Large data



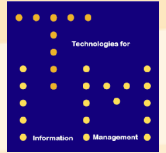
A Problem,

an Opportunity,

perhaps a Solution?



Intelligent Information Management ICT Call 5 (closed)

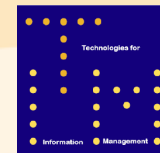


Key consultation themes:

- **Capturing tractable Information**
- **Delivering pertinent information**
- **Collaboration and decision support**
- **Personal sphere**
- **Impact and S&T leadership**

Focus: Large datasets
(aka big data, the petabyte age...)

Intelligent Information Management ICT Call 5 Results



- **Selection for funding by instruments*:**

- 9 STREP
- 5 IP
- 2 NoE
- 1 CSA

} First projects started
September 2010

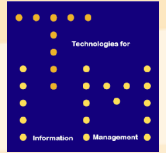
+ 3 enlargements of existing projects

* Negotiations partially not yet finally concluded



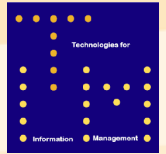
Intelligent Information Management

Call5 – main advancements



- **more scalable processing methods** for multimodal and multimedia information
- **novel and better scaling reasoning methods**, in particular for streams and very large scale RDF stores
- **advanced algorithms for data and knowledge mining**, in particular – relationship mining, entity identification, linking open data
- **novel real-time techniques** for processing and semantically fusing sensor data streams and extracting complex events and actions from them.

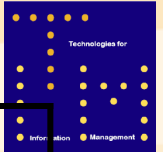
Intelligent Information Management Call5 – domains covered by projects



- Medical sciences data management
- Product engineering/design
- Earth sciences
- Financial and business intelligence
- Management of large on-line communities
- Improvement of linked-data resources
- Large scale public governmental data



Closed FP7 Calls in figures

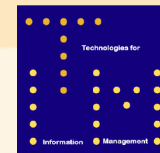


Inputs:	Call 1	Call 3	Call 5
Proposals	148	252	169
Participants	210	2017	1387
Countries	50	49	51
Request M€	473	817	611
Available M€	51	50	70
Outputs:	Call 1	Call 3	Call 5
Projects	15	13	17*
Participants	128	106	148
Countries	21	21	22
Total: 45 projects, 392 contractors, 171 M€			
* additional 3 enlargements of existing projects (10p)			



Closed FP7 Calls

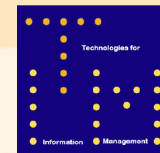
Selected Projects / Call & Instrument



Instruments	call1	call3	call5	Total
IP	4	4	5	13
STRP	9	7	9	25
NOE		1	2	3
CSA	2	1	1	4
Total	15	13	17	45



Intelligent Information Management ICT Call 5 Projects



2 project examples out of 17; started 9/2010

Relevance for your project?

Creating Knowledge out of Interlinked Data.

Will develop:

1. tools and methodologies for ..managing very large amounts of structured information on the Data Web,
2. ... high-quality multi-domain, multi-lingual ontologies from sources such as Wikipedia and OpenStreetMap.
5. adaptive tools for searching, browsing, and authoring of Linked Data.

LOD2

IP

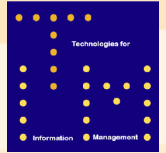
PLANETDATA

NoE

large-scale data management; community building; training; data provisioning, processing and Infrastructure

http://cordis.europa.eu/fp7/ict/projects/home_en.html

Intelligent Content and Semantics ICT Call 3 Projects

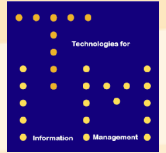


2 project examples out of 13; started begin of 2009

Relevance for your project?

E-LICO	STR	e-Laboratory for Interdisciplinary Collaborative Research in Data Mining and Data-Intensive Sciences Pilot application on diseases of kidney and urinary pathways (KUP).
CALBC	SA	Collaborative Annotation of a Large Biomedical Corpus

Intelligent Content and Semantics ICT Call 1 Projects



2 project examples out of 15; started begin of 2008

Relevance for your project?

**FOKUS-
K3D**

CA

Knowledge intensive technologies in analysis and production of 3D media

4 areas: Medicine and Bioinformatics; Gaming and Simulation; CAD/CAE and Virtual Product Modelling; Archaeology and Cultural Heritage.

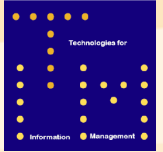
LARKC

IP

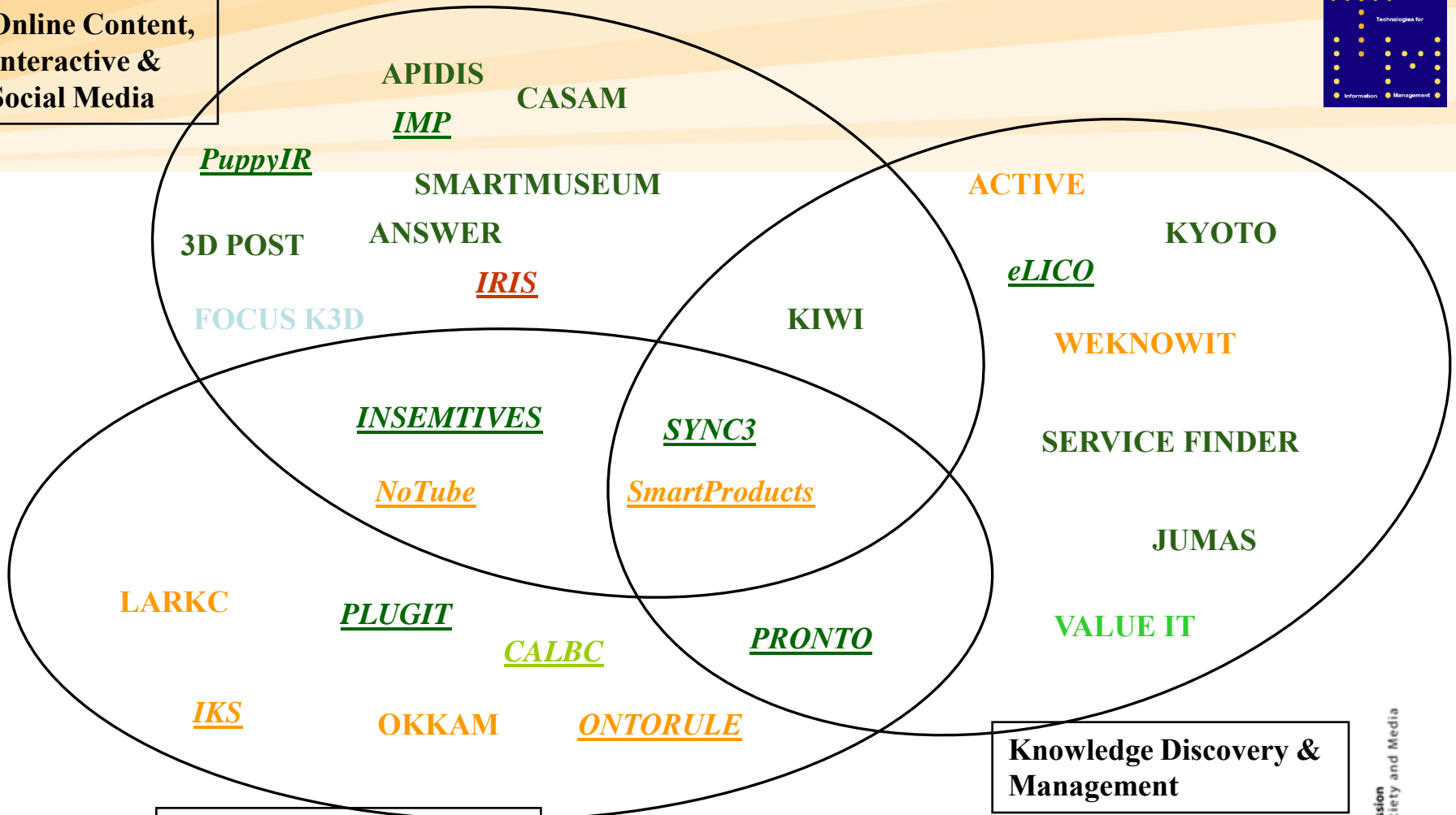
A pluggable distributed infrastructure for Web-reasoning and search

3 use cases: telecommunication services, bio-medical research, and drug-discovery

FP7 Intelligent Content - project portfolio



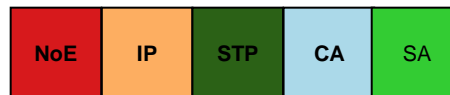
Online Content,
Interactive &
Social Media



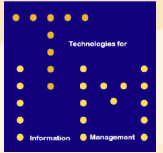
Reasoning & Information
Exploitation

Knowledge Discovery &
Management

CALL 1 CALL3



Agenda



- Data growth
- Large data
- FP7 ICT Call5 Intelligent Information Management
- Other relevant projects for KHRESMOI

Outlook

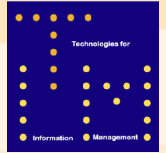
- Digital agenda for Europe
- New FP7 ICT workprogramme 2011-2012
- Next calls

AN OVERALL STRATEGIC FRAMEWORK FOR COMMUNITY ACTION

- First pillar of “Europe 2020” to devise and implement adequate European policies
- <http://ec.europa.eu/digital-agenda>



Challenges of ICT - Workprogramme 2011-2012



1. Pervasive and Trusted Network and Service Infrastructures
2. Cognitive Systems and Robotics
3. Alternative Paths to Components and Systems
- 4. Technologies for Digital Content and Languages**
5. ICT for Health, Ageing Well, Inclusion and Governance
6. ICT for a low carbon economy
7. ICT for the Enterprise and Manufacturing
8. ICT for Learning and Access to Cultural Resources

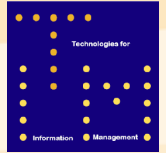
FET: Future and Emerging Technologies

International Cooperation

Horizontal Actions

Publication of WP: 23 July 2010

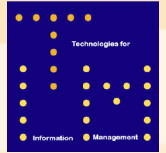
SME initiative on Digital content and language technologies (SO 4.1)



- **Reuse** of digital content resources
- To **exploit and contribute to large digital resource pools.**
- **User-centred experimentation** with the aim of demonstrating the integration of data-intensive technologies within innovative solutions and processes.



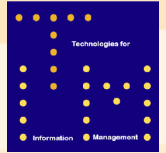
SME initiative facts and figures



- **Call Opening:** 01 February 2011
- **Short proposals:** 28 April 2011
- **Full proposals:** 28 September 2011
- **Indicative Budget:** 35 M€ in total (for Digital content and Language technologies)
- **Specific participation and evaluation rules** (min. 2 SMEs having >30% of max 2M€ budget; specific evaluation criteria, 2 steps)



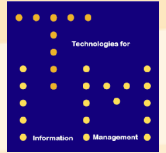
SME initiative Information Days



- **Monday, 22 November 2010 in Frankfurt**
- **Thursday, 9 December 2010 in Budapest**
- **http://cordis.europa.eu/fp7/ict/content-knowledge/fp7-callsme_en.html**



Intelligent Information Management in Call 8 (SO 4.4)



- **Focus on large data**
- **Call 8 Opening:** 26 July 2011
- **Deadline:** 17 January 2012
- **Indicative Budget:** 50M€
43M€ (>30% IPs, >50% STREPs)
7M€ (other instruments)

-



ICT Proposers' Day 2011
19 - 20 May, Budapest
Networking for European ICT R&D



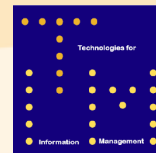
- Aim of the event:
to prepare for Calls 8 and 9 (together >1 billion €)
 - by networking and partnerships building
 - by first-hand information from >100 EC officials

- Structure:
 - thematic sessions with presentations of proposal ideas
 - information stands & meeting points

- Registration:
free of charge, open from January 2011



Further info



- **ICT under FP7**

<http://cordis.europa.eu/fp7/ict/>

- **Experts data base:**

<https://cordis.europa.eu/emmfp7/>

- **Unit E2 – Technologies for Information Management**

URL: <http://cordis.europa.eu/info-management/>

eMail to: info-e2@ec.europa.eu

Thank You!

