From Intelligent Content to Actionable Knowledge Research Directions and Opportunities

Roberto Cencioni

DG Information Society and Media
Unit INFSO.E2
Content & Knowledge

infso-e2@cec.eu.int

SAMT 2006, Athens, 8 Dec 2006

Agenda

- **Context & drivers**

- **FP6 state of play & FP7 outlook**

- **"Intelligent Content & Semantics"**

- Upcoming calls & tips

- Info sources & events

Content: a changing landscape

file sharing, many-to-many, social media, Web 2.0, long tail, non-market economy ...

- explosion in the availability of multimedia content
- produced & remixed by non-professionals
- accessed & consumed on a broad range of devices

- emergence & deployment of distributed (eg peer to peer) and socially enhanced content management applications

- growing cognitive load & diversity of content sources & types
- more and more data produced by devices as opposed to humans

Hidden intelligence

According to **Dr Mike Lynch**, CEO and Founder of **Autonomy**:

« **Meaning-based computing** is the way of the future as 80 per cent of information within enterprises is unstructured and it is understanding this 'hidden' intelligence that is at the heart to improving the way we interact with information. »

(not so) Hidden costs

Content management & business search
(source: IDC, March 2005)

- Information workers spend more than **27 hours a week searching, gathering and analyzing information**

- Information workers waste **3.5 hours a week searching for information that is never found and 3 hours a week recreating content**

- Rapid access to information drives effective business processes, but today's computing environment is still largely composed of **standalone, 'unaware' collaborative applications and services**

Enterprise

Gartner identified the **Corporate Semantic Web** as part of the Hype Cycle for Emerging Technologies:

« The corporate semantic web will reduce costs and improve the quality of **content management, information access, system interoperability, database integration** and data quality »

FP6 portfolio

- **62 projects**, ~700 contractors, ~270 M€ EU funding

- designed to address the intersection of Web, multimedia and KR&R technologies

- main features:
  - **networked** information, ontologies, users
  - **multimedia**, well beyond text (image, video, 3D…)
  - **automation** of the content / knowledge lifecycle

- emphasis on
  - generic enabling technologies
  - flexible application platforms

  URL: http://cordis.europa.eu/ist/kct/projects.htm

FP7 & Cooperation Pgme

(draft) Specific Programme text:

"**ICT for content, creativity and personal development**: novel forms of interactive, non-linear and self-adaptive content; creativity and enriched user experience; cross-media content customisation and delivery; combining all-digital content production and management with emerging semantic technologies; user oriented use, access to and creation of content."

"**Knowledge systems:** methods and techniques to acquire and interpret, represent and personalise, navigate and retrieve, share and deliver knowledge recognizing the semantic relationships in information for use by humans and machines."

"**ICT supporting businesses and industry:** dynamic, network-oriented business systems for product and service creation and delivery; … collaboration services … group management and sharing support."

Directions & priorities

- **extensive consultations:**
  - c 250 field experts
  - face-to-face meetings & written submissions
  - conferences & working visits
  - over a period of 6 months
- **other sources:**
  - ISTAG (high-level IST advisory group)
  - NEM (media) technology platform
  - NESSI (software) technology platform

A lasting research agenda, beyond 2008

URL: http://cordis.europa.eu/ist/kct/fp7.htm

In a nutshell:

- boost creativity

- master content

- dig out « hidden » information

"**Digital Libraries and Content**"

Make content and knowledge abundant, accessible, interactive and usable over time by humans and machines alike.

– Content must be made **available through digital libraries** and its **long term** usability, accessibility and preservation must be ensured

– Effective technologies need to be developed for intelligent **content creation and management**, and for supporting the **capture of knowledge and its sharing and reuse**

– Individuals, organisations and communities must find new ways to **acquire and exploit knowledge, and thereby learn**

European Commission
Information Society and Media

Political framework: « i2010 – Digital Libraries »

Intelligent Content & Semantics

Make digital resources that embody **creativity and semantics** easier and more cost effective to produce, organize, search, personalise, distribute and use across the value chain.

- **CREATORS**: Design more communicative and participative forms of content (media professionals, enterprise designers, talented amateurs)

- **PUBLISHERS**: Increase productivity in creative industries, enterprises and professional sectors (eg health, law)

- **SCIENTISTS**: Automate link between data analysis, theory and experimental validation

- **ORGANISATIONS & COMMUNITIES**: Automate collection and distribution of digital content and machine-tractable knowledge, and their sharing in collaborative environments

European Commission
Information Society and Media

Target socio-economic sectors

- **key features**
  - ICT based, high growth & innovation potential
  - pronounced international character
  - sophisticated users
  - very large data volumes
  - well defined flows & protocols
- **obvious candidates** (in addition to *ICT*!)
  - *creative industries* (film, TV, games, advertising …)
  - *enterprises in information bound industries*
    - utilities eg energy
    - manufacturing & process industries
    - construction & engineering, financial services …
  - *eScience* eg life sciences

(intertwined) Themes

RTD:

– authoring

– workflow

– personalisation

– semantics

– knowledge

Coordination & Support

– community building

– technology assessment

– exploitation channels, take-up

(a) Advanced Authoring

create / capture content

- "creativity": explore new media paradigms & novel forms of content; support creative process & experimentation

  – generate metadata as new content is created

  – annotate & categorize legacy content to ease reuse

  – find reference & inspirational material

- enable user "experience" & control; interactivity, highly visual & non-linear content; building on gaming, 3D, simulation, animation …

- "democratisation": low-cost, scalable-functionality personal tools - editing/sharing/remixing; usability

(b) Collaborative Workflow

- from "film" through files to objects:
  - integrated, metadata & object based postproduction flows
  - management & reuse of content assets; versioning, packaging & repurposing
  - where relevant, adaptation to different target markets & groups including cultural/linguistic elements
- multimedia segmentation, summarisation, aggregation, (scalable) (trans)coding according to distribution channels …

(c) Personalised Presentation & Consumption

- (re)active, self-aware, adaptive … content
- user, context & device adaptation
- visualisation, immersive rendering, multimodal interaction
- privacy preserving logging / feedback datamining

European Commission
Information Society and Media

SEVENTH FRAMEWORK PROGRAMME

(e) Semantic Foundations

objective driven research

- beyond current knowledge models & formalisms

  - approximate reasoning & induction

  - temporal, probabilistic & modal modelling …

- reference implementations esp. web integration of heterogeneous data sources

  - multimedia resources

  - (real-time) data streams

  showing the practical value & power of semantics

(f) Knowledge Systems

problem oriented

- systems, architectures & technologies for information bound organisations & communities
  - ▪ <u>core S&T task</u>: extract "meaning" from information, social interaction & work patterns; make it computer tractable
- knowledge intensive tasks & domains
  - ▪ dynamic data & application integration
  - ▪ business search, knowledge (eg scientific) discovery
  - ▪ cross-organisational interoperability: business processes, dynamic collaboration & knowledge sharing
  - ▪ … identity management, audit trails, privacy …

multimedia, data & process semantics

European Commission
Information Society and Media

Approach & key features

- problem & objective driven

- principled approach, no quick-and-dirty fixes

- scope encompasses (actual mix depends on project types & objectives):

  – **formal** (KE) + **social** (Web 2.0) + **ambient** (Internet of things) resources & approaches

  – **foundational** + **component** + **system** level research

- centred around **<u>real</u> users, data & flows**

- integrated demonstrator(s), field validation & assessment

- usability, scalability, replicability; legacy data/systems

- active promotion & dissemination of results … beyond scientific circles

multi-sector potential, innovative areas, beyond SoA

What we don't do

In 2007-08 we do not intend to support <u>research</u> into:

- **basic research** with no identifiable by-products within 10 years

- **domain specific applications** - not portable/replicable in other socio-economic sectors

- developments addressing **immediate commercial imperatives** eg content protection & monetisation

- issues covered by **other Challenges and Objectives** eg media networking, peer to peer, technology enabled learning …

- topics well covered by **on-going FP6 projects & networks** (see our website)

individual projects can however address one or the other of the above issues, and integrate existing & emerging technologies

European Commission
Information Society and Media

SEVENTH FRAMEWORK PROGRAMME

Evolution, no disruption:

- hard research problems remain invariant

- automation, networked information, multimedia … stay as well

- efficiency & cost effectiveness are key drivers
  - clear application potential & exploitation channels

- broader scope, no longer purely « document » centric
  - task & user centric
  - more attention paid to human & organisational factors
  - so as to encourage new programme entrants

- « systems » approach: reference architectures; reusable tools; interworking with legacy systems

- pro-active promotion & dissemination
  - so as to leave behind resources others can build on

European Commission
Information Society and Media

Calls in 2007-08

- **~2 Beuro in 2007-08, 3 calls in total**

  - Call 1 *   big   Q1, Dec/Jan?

  - Call 2   smaller   Q2, Jun?

  - Call 3 *   smallest   Q4

- **Challenge 4 called twice (calls 1 & 3)**

  - c 100 Meuro for Intelligent Content & Semantics (2 x 50 M)

  - c 100 Meuro for Digital Libraries & Learning (2 x 50 M)

  - same scope for both calls, budget flexibility vav research themes

- 250+ Meuro in total for content & knowledge related research, across Challenges & objectives

- more calls from Q4 2008 on … (~9 Beuro in 2007-13)

24

Schedule of 1st call (provisional)

- 51 Meuro in total of which:

  – 46 Meuro for IP & STR projects

  – 5 Meuro for NoEs & CSA's

- first call expected to close late April (?)

- evaluation/selection mid-May – late Jun (?)

- negotiations until Nov

- contract awarding in Dec

- projects due to start Q1 2008

  highly demanding process …

EVALUATORS WANTED …!

Use of instruments

- IPs   **➔ impact**
    - up to 4 years, 5-9 Meuro (EU funding)
- NoEs   **➔ integration**
    - up to 3 years, 1-4 Meuro
- STRs "research"   **➔ S&T innovation**
    - up to 3 years, 2-4 Meuro
- STRs "demonstration"   **➔ uptake**
    - up to 2 years, 1-2 Meuro
- CSAs (coordination & support actions)
    - up to 3 years, 1-3 Meuro

- **STREPs especially geared towards use cases & field experimentation ("first use")**
- centred around existing, promising but untried technology
- designed to go one step forward towards
  - packaging, configuring … and testing
- assess viability
  - functionality
  - technical performance & flexibility
  - usability (hide complexity!)

  within a well defined domain / user context
- rigorous evaluation plans & metrics
- active user involvement & feedback
- adequate documentation of results (positive/negative)

- main mission: integration & critical mass
  - integrative <u>research</u> & shared facilities
  - convergence, bridges between institutions
  - spreading knowledge, training
- sizeable effort in FP6
  - 6 NoEs underway (representation, reasoning, learning, multimedia semantics, 3D & shapes)
  - adequate coverage at this point in time
- no obvious candidate topics under Call-1; better opportunities under Call-3
- … but let us know …

European Commission
Information Society and Media

SEVENTH FRAMEWORK PROGRAMME

- **theme (d):** horizontal & prospective activities cutting across the Objective; linking research to the broader environment

- either <u>embedded</u> in projects (esp. IPs & NoEs) or via <u>dedicated</u> CSA actions

- main axes for Call-1:

  - **community building**, interdisciplinary exchanges, shared visions

  - **user-supplier dialogue,** awareness, exploitation channels, technology transfer/access to market

  - **technology assessment,** benchmarking

European Commission
Information Society and Media

What kind of project?

	Foundational research	Component technology	System integration & validation
IP	Poss.	Yes	**Yes**
NoE	**Yes**	Poss.	NR
STR - R	**Yes**	**Yes**	Poss.
STR - D	NR	Yes	**Yes**

NR   =   Not recommended
Poss. =   Possible
Yes   =   Highly recommended

Partnerships

- compact consortia:

  - **IPs**   7-12 partners

  - **NoEs**   3-4   "core" partners

  - **STRs**   4-8   partners

- cohesive agenda; competent, committed & reliable partners

- complementarity: cover all areas you need while keeping consortium <u>manageable</u>

- duplication of competence:

  - necessary for NoEs

  - acceptable for IP/STR's where dictated by project needs

- industry, SME, academia … participation:

  - as dictated by project needs

Exp. outcome of 2007 calls

- more focused IPs & NoEs wrt. FP6

- more ambitious STR-R's

- STR-D's to encourage integration, customisation & validation

- c 100 Meuro available (2 x 50 Meuro):
  - 8-9   IPs
  - 1-2   NoEs
  - 14-15 STRPs
  - 2   CSAs

- ~26 proposals likely to be retained for funding … highly selective process!

Guidance & feedback

- ## inquiries & candidate evaluators
  - mail to infso-e2@cec.eu.int
  - 'FP7-ICT Call' in the Subject field

- ## FP6 projects

  cordis.europa.eu/ist/kct/projects.htm

- ## pre-proposal checking
  - available from mid-January onward
  - along with Call 1 technical background notes
  - discont'ed 4 weeks before call closure

  cordis.europa.eu/ist/kct/fp7.htm

34

Pre-proposals

3 pages max, emailed to infso-e2@cec.eu.int

- **rationale & problem area**
  - task & user profile
  - actual/prospective application(s)
  - data sets: source(s), typology, volume
- **contribution to WP objectives**
  - key S&T innovation(s)
  - main concrete results
  - impact (scientific, technical, socio-economic)
  - public outputs
- **consortium**
  - names or profiles, skills mix
  - intended instrument (if known)
- **scale of ambition**
  - est. effort, duration, EU funding

- ICT under FP7   **cordis.europa.eu/fp7/ict/**

- Unit E2 – Content & Knowledge
  URL: cordis.europa.eu/ist/kct/fp7.htm
  mailto: infso-e2@cec.eu.int
  c/o Dr Stefano Bertolo

- AXMEDIS 2006, Leeds   15 Dec 2006
  www.axmedis.org/axmedis2006

- Dir E info session, Luxembourg   24-25 Jan 2007
  cordis.europa.eu/ist/directorate_e/fp7_infodays.htm

- ICT Proposers' day, Cologne   1 Feb 2007
  ec.europa.eu/information_society/events/koln_2007/index_en.htm

European Commission
Information Society and Media

SEVENTH FRAMEWORK PROGRAMME

Thank you

QUESTIONS?

infso-e2@cec.eu.int

Draft ICT Workprogramme:
http://cordis.europa.eu/fp7/ict/

Experts database:
https://cordis.europa.eu/emmfp7