META-NET

The META-NET Strategic Research Agenda and Linked Open Data

Georg Rehm

DFKI, Germany
georg.rehm@dfki.de

Multilingual Web Workshop – Dublin, Ireland
June 11, 2012

Outline

- Introduction

- Strategic Research Agenda (SRA): Overview

- The SRA and Open Data

Multilingual Europe

META-NET

□ **Challenge:** Providing each language community with the most advanced technologies for communication and information so that maintaining their mother tongue does not turn into a disadvantage.

□ While research has made considerable progress in recent years, the pace of progress is not fast enough to meet the challenge within the next 10-20 years.

□ All stakeholders – researchers, LT user and provider industries, language communities, funding programmes, policy makers – should **team up for a major dedicated push**.

Objectives

META-NET is a network of excellence dedicated to fostering the technological foundations of the European multilingual information society.
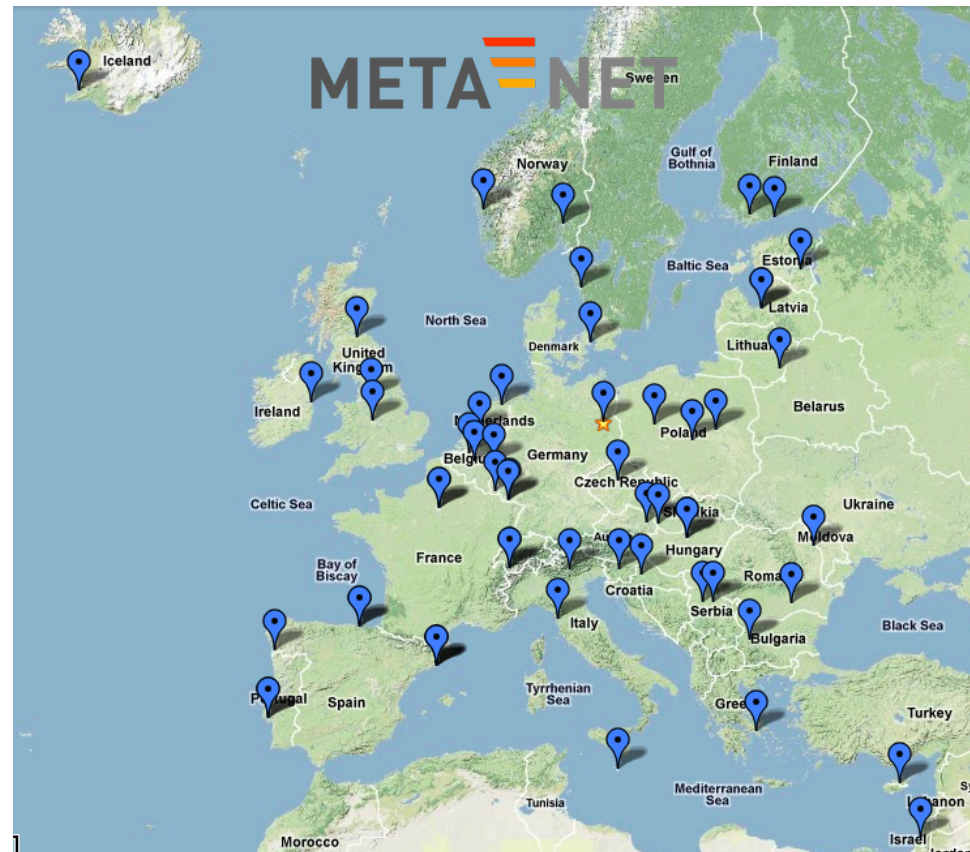
META-VISION: Building a community with a shared vision and strategic research agenda

META-SHARE: Building an open resource exchange infrastructure

META-RESEARCH: Building bridges to neighbouring technology fields

Four Funded Projects

- Initial project: T4ME (FP7; 13 partners, 10 countries)

- Three ICT-PSP consortia since Feb. 2011: CESAR, METANET4U, META-NORD

- All EU member states and several non-member states covered.

- META-NET in June 2012: **57** members in **33** countries.
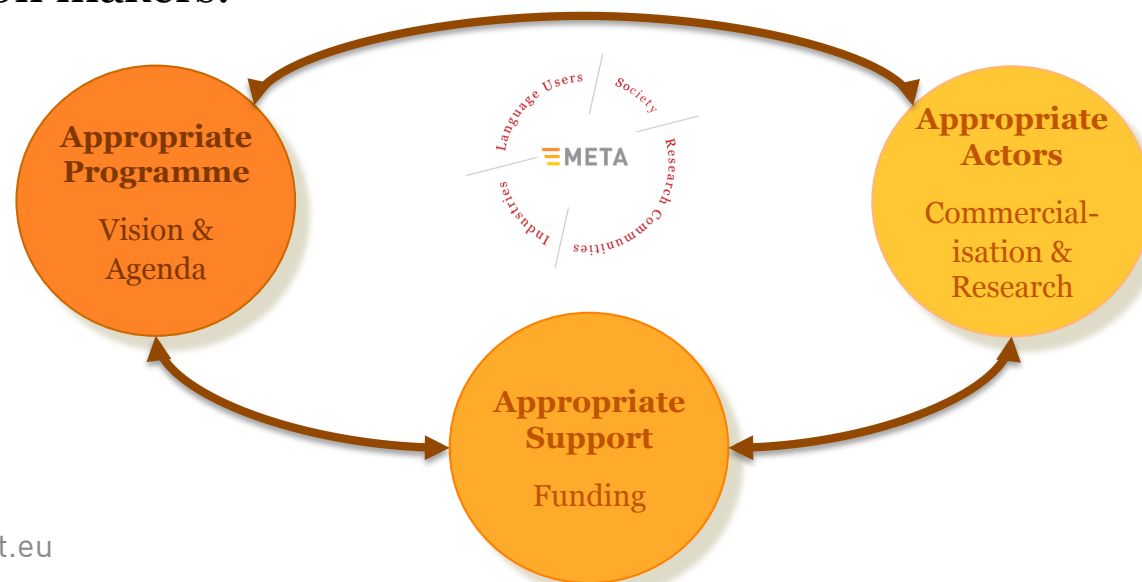

http://www.meta-net.eu/members

META-VISION

Strategic Research Agenda

Shared Vision and SRA

❑ Mobilize researchers, users and providers of LT for cooperation and collaboration to initiate and support a large-scale joint action to

  ▪ Building a **community** around Language Technology in Europe (META),

  ▪ Creating a **shared vision**,

  ▪ Preparing a **Strategic Research Agenda** for Multilingual Europe 2020, to be presented to national/international politicians, funding agencies, decision makers.
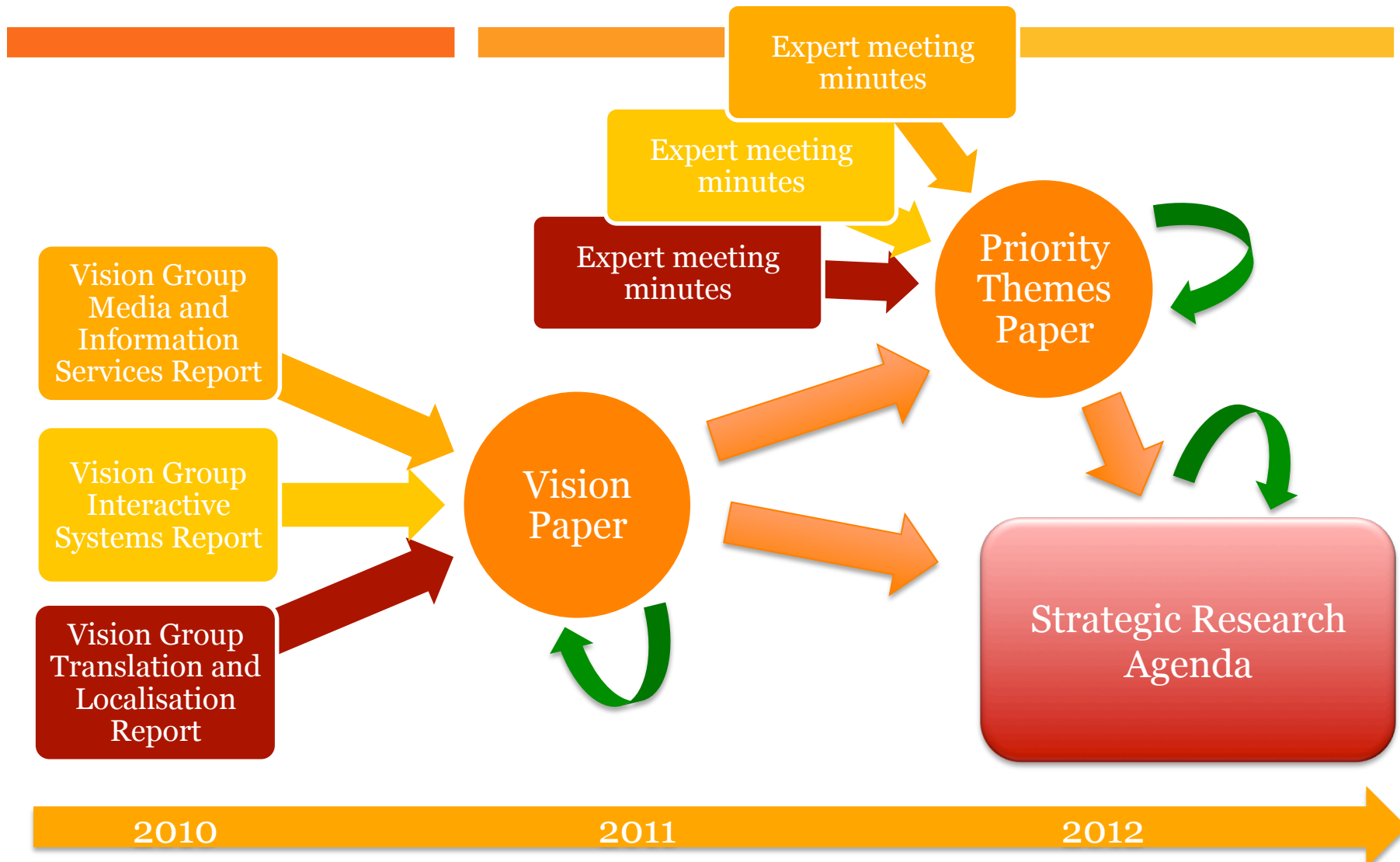
From Visions to the SRA

- Three **Vision Groups** brought together researchers, developers, integrators and (corporate or professional) users of LT-based products, services and applications (ca. 25 members each).

- Collected domain-specific visions and prepared individual reports.

- Vision Group
  Translation and Localisation
  - July 23, 2010 — Berlin, Germany
  - September 28, 2010 — Brussels, Belgium
  - April 7/8, 2011 — Prague, Czech Republic

- Vision Group
  Media and Information Services
  - September 10, 2010 — Paris, France
  - October 15, 2010 — Barcelona, Spain
  - April 1, 2011 — Vienna, Austria

- Vision Group
  Interactive Systems
  - September 10, 2010 — Paris, France
  - October 5, 2010 — Prague, Czech Republic
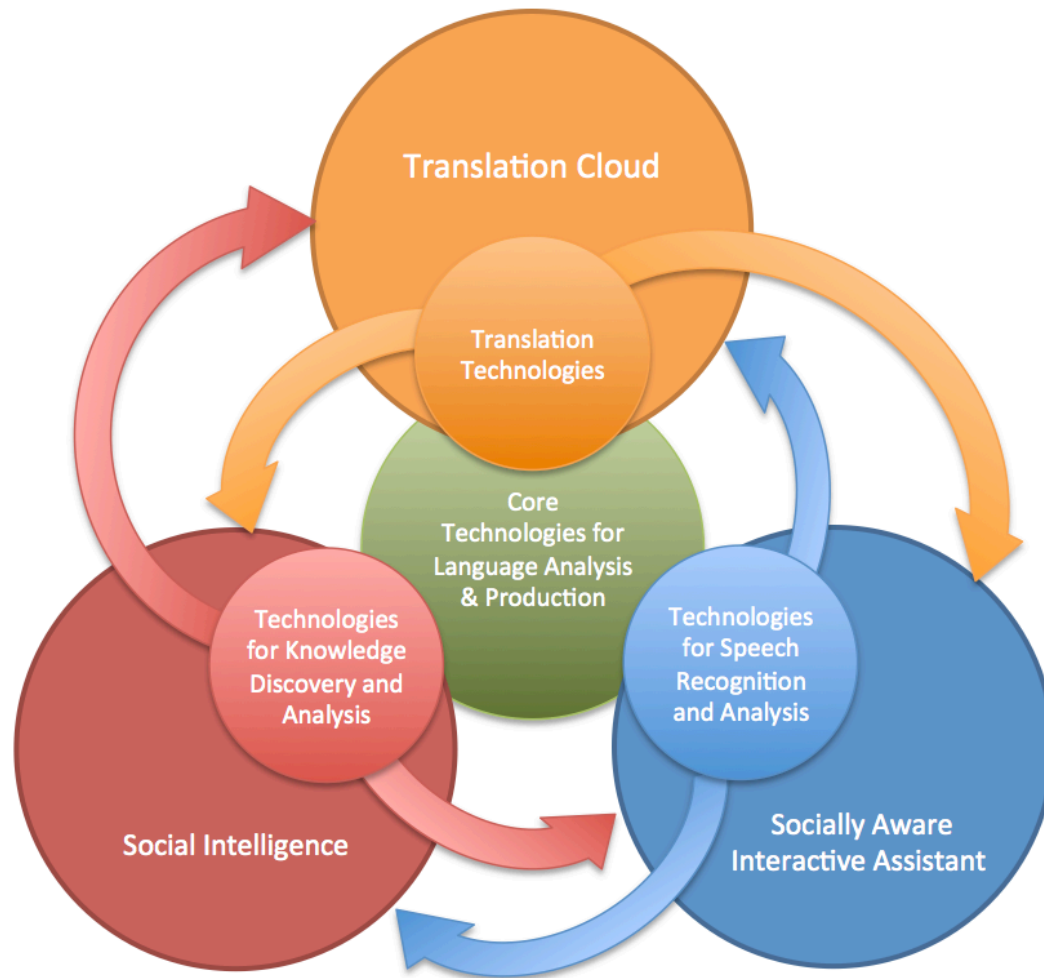 - March 28, 2011 — Utrecht, The Netherlands

Steps taken towards the SRA

- Many suggestions by: Vision Group members; META-NET; META.
- Additional input in meetings, workshops, discussions etc.
- Language White Papers: research gaps; funding situation etc.
- Hundreds of people participated in the vision building process.
- We screened the Strategic Research Agendas of other initiatives.
- We discussed procedures, input and structure of the SRA in three meetings of the META Technology Council.
  - Brussels, Belgium, November 16, 2010
  - Venice, Italy, May 25, 2011
  - Berlin, Germany, September 30, 2011
  - *Upcoming:* Brussels, Belgium, June 19, 2012

From Visions to the SRA

META-NET

Expert meeting minutes

Expert meeting minutes

Expert meeting minutes

Priority Themes Paper

Vision Group Media and Information Services Report

Vision Group Interactive Systems Report

Vision Group Translation and Localisation Report

Vision Paper

Strategic Research Agenda

2010

2011

2012

Priority Research Themes

Open Data in Horizon 2020

META=NET

Jose L. Marin
@jluismarin

Great news for #opendata companies RT @sclopit: #edf2012 data will be one of five main areas addressed in Horizon 2020 program (2014-2020)

2:55 AM - 6 Jun 12 via Twitter for Android · Embed this Tweet

stefano bertolo
@sclopit

#edf2012 data will be one of five main areas addressed in Horizon 2020 program (2014-2020)

2:44 AM - 6 Jun 12 via Twitter for iPhone · Embed this Tweet

Sergio Fernández
@wikier
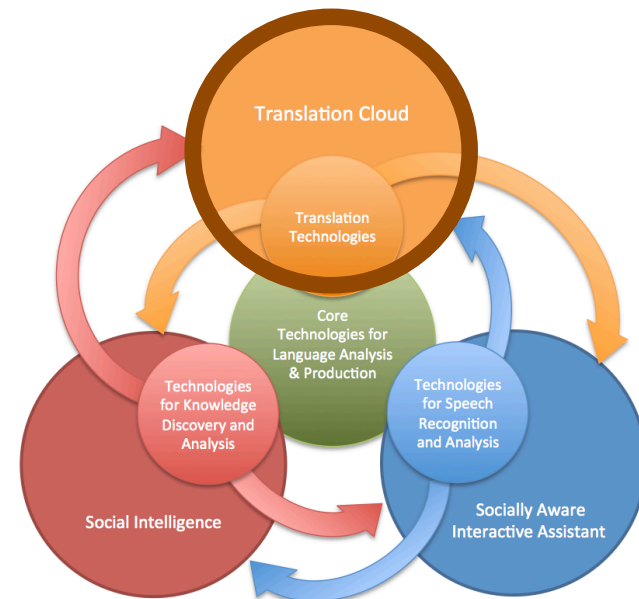
EC is going to reorganize some things, welcome the data value chain #edf2012

2:25 AM - 6 Jun 12 via TweetDeck · Embed this Tweet

European Data Forum 2012: Data Challenge will be one of five main topics in H2020!

Data and Language Technology

META=NET

- Data challenge is concerned with:
  - Big data
  - Open data
  - Linked data
  - Data value chain

- Strong relation between our priority themes and the data challenge:
  - *As beneficiaries:* using data as data or language resources – big, open, and linked open data sets can be conceptualized as *resources* for LT.
  - *As contributors:* produce, extract, generate, connect data.
  - Also: interlinked language resources as an inherent part of the linked open data cloud (*Linguistic Linked Data,* as presented by, among others, Sebastian Hellmann and John McCrae).
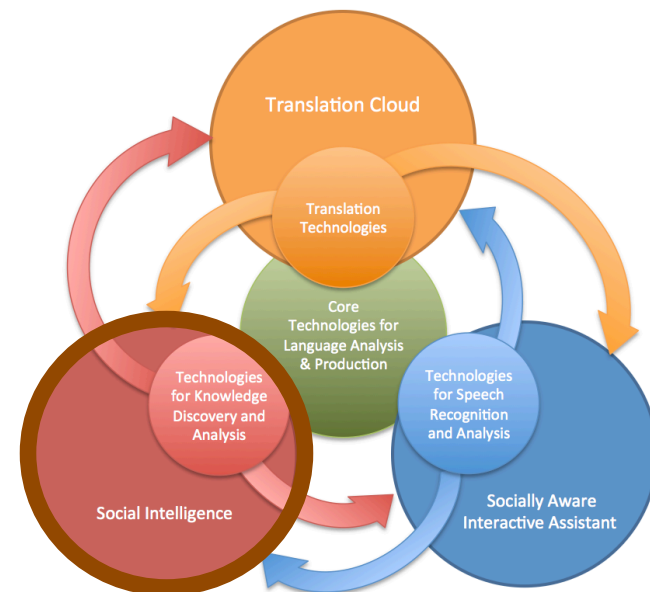
Translation Cloud

META NET

- As beneficiary:
  - Exploit data sets to get access to data across languages – can be used to, for example, improve machine translation models.

- As contributor:
  - Provide multilingual access to data sets.
  - Deal with big data and language barriers between data sets.
  - Build links and connections in data sets between and within languages.
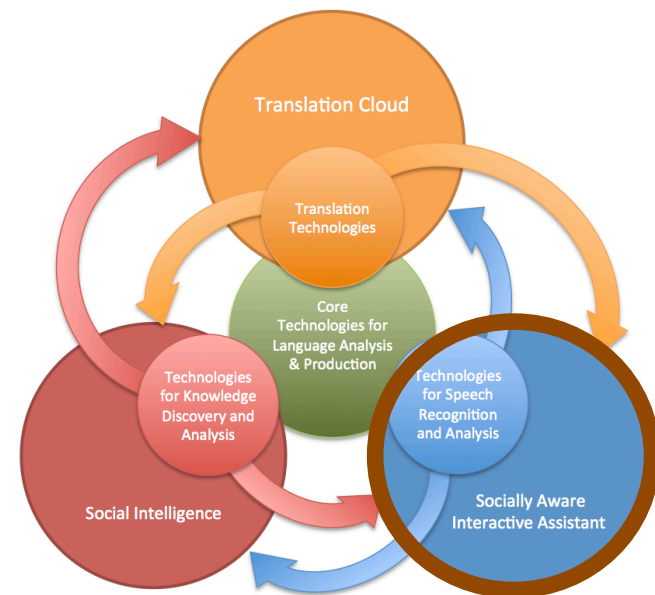 - Example: generate cross-lingual references between named entities.

Social Intelligence and e-Participation

- ❑ As beneficiary:
  - ▪ Make use of an exploit open, governmental data (for example, data.gov.uk or the upcoming European data portal).
  - ▪ Crucial for the e-Participation topic!

- ❑ As contributor:
  - ▪ Provide novel methods of accessing data sets in Europe.
  - ▪ Help to clean data sets.
  - ▪ Technological means: machine translation, information extraction, text mining, sentiment detection, relation extraction, cross-lingual information access, ...
  - ▪ *Language is the biggest data of them all!*

Interactive Assistant

META-NET

- As beneficiary:
  - Data sets about organizations, individuals, settings of interaction etc. can help for different interaction settings, virtual meetings etc.

- As contributor:
  - Provide novel methods of access to data sets (voice interaction)
  - Data aware-infrastructure for multilingual virtual meetings and conferences

Crucial Aspects

- How will data sets be distributed? Do we need private or public data infrastructures? Both? (Important point for re-use, privacy, licensing schemes etc.)

- Information about the origin, provenance, trustworthiness, quality, coverage etc. of data sets are crucial for a solid data value chain.

- Publish terminological, lexical, translation memory, word nets, named entities and other general LT resources as linked (open) data.

- Make LT generally applicable for various language-specific or country-specific data sets (e.g., financial records).

META

- **META-NET** is a network of excellence.

- **META** is an open and growing strategic technology alliance: **Multilingual Europe Technology Alliance**.

  - 600+ members, including W3C, Google, Microsoft, GALA, research centres, LT companies etc.

  - Main goal: to support our Strategic Research Agenda.

  - **Join us!** **http://www.meta-net.eu/join**

Conclusions – Next Steps

- ❑ Meet with national research planners, funders, policy makers and inform them about the Strategic Research Agenda.
- ❑ Present and discuss SRA and roadmap in the community and gather as many supporters (through membership in META) as possible.
- ❑ FP7 and others have already built bridges between open data and LT (DBpedia, Monnet, Wikidata, META-SHARE etc.)
  - ▪ Now is the time to strengthen these relations!
  - ▪ *This will ensure long-term availability (including easy, natural language access) of data for the European multilingual information society!*
- ❑ *Upcoming opportunities can provide sufficient resources to make our visions for Europe's citizens and economy, as described in the SRA, a reality (Horizon 2020 and Connecting Europe Facility, CEF).*

Q/A

META NET

Thank you very much!

See you in Brussels on June 20/21 for

META FORUM 2012

http://www.meta-forum.eu

office@meta-net.eu

http://www.meta-net.eu
http://www.facebook.com/META.Alliance