This document is part of the Network of Excellence "Multilingual Europe Technology Alliance (META-NET)",
co- funded by the 7th Framework Programme of the European Commission through the T4ME grant agreement no.: 249119.

A Network of Excellence forging the
Multilingual Europe Technology Alliance

Vision Document

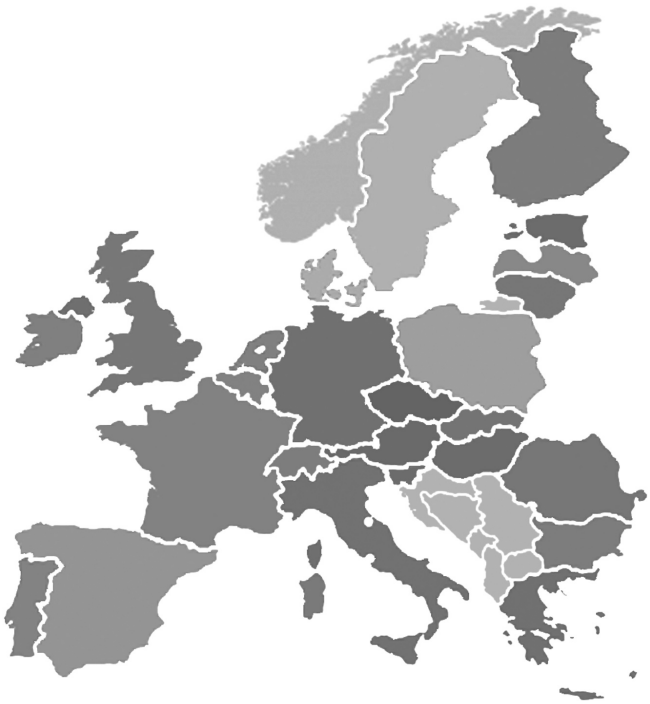
Vision Group Media and Information Services:

Results of first two meetings

Editors: Maria Koutsombogera, Stelios Piperidis

Dissemination Level: Public

Date: 10 November 2010



Grant agreement no.	249119
Project acronym	T4ME Net (META-NET)
Project full title	Technologies for the Multilingual European Information Society
Funding scheme	Network of Excellence
Coordinator	Prof. Hans Uszkoreit (DFKI)
Start date, duration	1 February 2010, 36 months
Distribution	Public
Contractual date of delivery	n.a.
Actual date of delivery	n.a.
Deliverable number	n.a.
Deliverable title	Vision Document- Vision Group "Media and Information Services": Results of first two meetings
Type	Report
Status and version	Draft
Number of pages	16
Contributing partners	ILSP
WP leader	DFKI
Task leader	ILSP
Authors	Maria Koutsombogera, Stelios Piperidis, input from all members of the Vision Group "Media and Information Services" listed in Section 3.2, with Philippe Wacker as Chair and Margaretha Mazura as Rapporteur.
EC project officer	Hanna Klimek
The partners in META-NET are:	Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI), Germany
	Barcelona Media (BM), Spain
	Consiglio Nazionale Ricerche – Istituto di Linguistica Computazionale "Antonio Zampolli" (CNR), Italy
	Institute for Language and Speech Processing, R.C. "Athena" (ILSP), Greece
	Charles University in Prague (CUP), Czech Republic
	Centre National de la Recherche Scientifique – Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (CNRS), France
	Universiteit Utrecht (UU), Netherlands
	Aalto University (AALTO), Finland
	Fondazione Bruno Kessler (FBK), Italy
	Dublin City University (DCU), Ireland
	Rheinisch-Westfälische Technische Hochschule Aachen (RWTH), Germany
	Jozef Stefan Institute (JSI), Slovenia
	Evaluations and Language Resources Distribution Agency (ELDA), France

For copies of reports, updates on project activities and other META-NET-related information, contact:

DFKI GmbH
META-NET
Dr. Georg Rehm
Alt-Moabit 91c
10559 Berlin, Germany

office@meta-net.eu
Phone: +49 (30) 3949-1833
Fax: +49 (30) 3949-1810

Copies of reports and other material can also be accessed via http://www.meta-net.eu

Table of Contents

1 Executive Summary

Language industries will contribute a considerable percentage of Europe's future economic growth. META-NET focuses on understanding the particular goals and constraints of research, user industries, provider industries, administrations and government, surveying the current state and identifying the gaps with the goal to create a shared vision. To this end, it operates three Vision Groups.

The Vision Group Media and Information Services is composed of high-level experts representing the major areas and players in the respective domain. Two Vision Group meetings were organized with the purpose of initiating a discussion and producing draft visions as input to the META Technology Council and the Strategic Research Agenda (SRA). In addition, the Vision Group meetings raised a number of issues which will also be addressed in this document.

The Vision Group Media and Information Services identified a number of horizontal aspects, and addressed domain-specific needs, challenges and opportunities related to the Media and Information Services landscape. A rough listing is presented below:

Needs

- Need #1: Audiovisual search.
- Need #2: Capture of web content at a European level.
- Need #3: Large-scale text modelling.
- Need #4: Improvement of speech technologies.
- Need #5: Industry- Academia synergies in Speech Recognition.
- Need #6: Deeper understanding of text content.
- Need #7: Dialogue and Interaction modelling.
- Need #8: Genre and text-type based layers in language analysis and generation.
- Need #9: New types of GUIs.
- Need #10: Build competitive products for the European market.
- Need #11: Multilingualism should be a priority.
- Need #12: Standardisation.
- Need #13: Semantic search for a unified European market.
- Need #14: Advanced (open source) tools that are available to both large industry players and SMEs.
- Need #15: Deliver highly skilled experts to industry.

- Need #16: Cope with legal/IPR issues, political and economic/market considerations as well as ethical issues.

Visions

- Vision #1: Federated Audiovisual (AV) search.
- Vision #2: Social stream text mining, Information overload handling, Semantic search
- Vision #3: Life logging.
- Vision #4: Automatic minutes production from meetings, automatic indexing for voice search, speech transcription and translation for videos.
- Vision #5: Multimedia multilingual subtitling.
- Vision #6: Improved text mining.
- Vision #7: Personalized task-centered interactive information assistants.
- Vision #8: Genre-aware LT applications.
- Vision #9: Voice Control instead of traditional GUIs.
- Vision #10: Deliver more efficient online advertising.
- Vision #11: Semantic Annotation.
- Vision #12: Boundaries elimination and mass collaboration efforts for making European SMEs more successful.
- Vision #13: Promote the Language Resources ecosystem.
- Vision #14: A common infrastructure to ensure interoperability of the LR production and distribution/sharing processes.
- Vision #15: Focus funding on rapid progress in basic technology for semantic annotation and search.
- Vision #16: Cost-effective porting of LT services and solutions across domains.
- Vision #17: Develop synergies among industry and academia and bring academic excellence to the market.

2 Introduction

2.1 The META-NET Vision Building Process

A central objective of META-NET is the preparation of a major concerted effort geared towards the creation of the needed technological foundation for the European multilingual information society. An essential instrument to this end is the forging of a strategic alliance involving, in addition to the top level R&D centres, the active participation of European LT

and ICT industry and many private and public stakeholders, including the language communities themselves.

The Vision Groups are a central instrument within META-NET. Each of the three groups brings together researchers, developers, integrators and (actual or potential, corporate or professional) users of LT-based products, services and applications. The goal of the groups is to generate domain-specific visions and roadmaps in the form of technology forecasts. This includes ideas for innovative applications of language technology and scenarios for the future knowledge society which can be supported by advanced technology. The visions produced will be gathered by the Technology Council which will consolidate them into a Strategic Research Agenda. The Agenda will contain high-level recommendations and suggestions for joint actions to be presented to the EC and national as well as regional bodies. The three Vision Groups are:

1. Translation and localization
2. Media and Information Services
3. Interactive Systems

The Vision Groups are scheduled to meet twice a year. In 2010, two rounds of meetings have been successfully completed. Their output will be discussed at the 1st Technology Council Meeting (16/11/2010), which, in turn, will prepare the draft SRA. Preliminary findings of the Vision Groups will be presented at META-FORUM 2010 (17/11/2010). Additional meetings will take place in 2011, the goal being to provide input to the SRA draft and to further elaborate the visions.

This document is intended to provide a distillation of ideas, opinions and visions expressed by members of the Vision Group Media and Information Services during its first two meetings. It aims at serving as a basis for discussion on the challenging and innovative LT based scenarios that need to be addressed by 2020 and providing seed ideas for the drafting of the SRA to the Technology Council.

3 Vision Group Media and Information Services

The goal of this Vision Group is to gather distinguished representatives of leading companies and research centres in the areas of (Multi)Media, Information Services as well as Language Technologies (LT) and produce visions to support the development in Europe in this field by establishing a crucial role in this area and by speeding up developments. The focus is on the requirements that Media and Information Services have from language/content processing technologies as communication/information facilitators and the role they foresee for them. A clear view on LT in Media and Information Services will highlight their importance in a sector which is becoming more and more important from a financial point of view and,

thus, trigger financial and strategic support (by the EC and the EP, among others), to give Europe the chance of becoming *the* global player. Specifically, the sectors and players of this group, as well as representative technologies and applications are the following:

Fields: audiovisual sector, news services, digital libraries, portals and vortals, search engines, social networks etc.

Target stakeholders: audiovisual and media industries, web and search engine providers, archives, etc.

Technologies:

- speech processing (recognition, synthesis), subtitling, text simplification, media mining, multilingual data processing, etc.
- topic identification, content classification and structuring, information extraction, summarisation, multidimensional analytics (from fact/event recognition to sentiments/ emotions, etc), authoring, etc.
- LT-enabled digitisation, mono/multilingual/multimedia search, semantic search, etc.
- processing unregulated language and discourse, etc.

Point of Departure – Existing Applications today: Google translate, Voxalead named entity-based access to video news, Hospital report dictation, Business intelligence and Quality Management for call centres, EMM News Explorer – cross-lingual news clustering/NE analysis, Metacarta's NewsMap, ClearForest's OpenCalais, iPhone speech recognition, applications in the job market, real estate, travel, classifieds, logistics, health, financial analysis and research, knowledge management, personal mobile navigation, expert search, social media, etc.

3.1 Recruitment process

Extensive discussions and decisions regarding the composition of the groups, selection criteria for both internal and external members and the invitation process were carried out during META-NET Standing Committee and Management Board meetings. In order to cater for the best topic, sector and geographical coverage, the selection conformed to the criteria of (a) research vs. industry proportion (50% industry, 50% research, at most half of this research percentage from within the META-NET consortium) and (b) even geographic distribution (members from both the big EU member states as well as from smaller countries). It was also decided that the Vision Group starts with initial core members to be extended if needed.

Regarding the population of the Vision Group Media and information Services, a list was compiled after multiple rounds of discussions and suggestions by members of the consortium and the Standing Committee. A target number for participation in the first set of meet-

ings was set to 8-10 recommended external members. The key features concerning the recruitment of the external members are the following:

- Initial member suggestions: 58
- Invitations sent out: 43
- Confirmations: 28
- Refusals: 3
- No answers: 12
- Meeting attendants (cf. table 1): 12

A significant percentage of the invitees had a scheduling conflict between the dates of the meetings and other commitments; some of them rescheduled their agenda so that they can participate in the Vision Group meetings; others suggested substitutes, who were welcome to join the group; but in the end, some of the invitees could not make it to the meetings, despite expressing their interest to participate and contribute. The latter will keep being contacted to participate in the follow-up meetings and in the online discussion forum. The reasons provided by invitees that had to decline the invitation were lack of time and over-involvement in boards and committees.

3.2 Meetings

Two VG meetings were held in 2010, with a total number of 26 participants (cf. table 1):

- 1st meeting: 10/09/2010 at Aero Club de France, Paris,
- 2nd meeting: 15/10/2010 at Barcelona Media, Barcelona.

During the first meeting a mutual understanding was reached regarding the goals of the particular Vision Group and its composition. The participants of the meeting were asked to contribute by expressing their views on challenges, needs, opportunities and gaps regarding future scenarios of Language Technology research and use.

The objective of the second meeting of the Media and Information Services Vision Group was to gather all the visions of the members and organize them in order to produce a draft bullet-point report to be submitted to the META Technology Council. To this end, and in order to achieve a structured input by the members, the following five questions were addressed regarding the role of language technology in the Media and Information Services landscape:

1. What are attractive, plausible, powerful, challenging, innovative language technology-based applications, or combinations of applications with use cases, that in your opinion could be realized through massive concerted research, development and innovation actions?

2. Focusing on media and text analytics, what do you see as the top most important technologies, market challenges and opportunities in the medium (by 2015) and long (by 2020) term?
3. Can you think of novel research advances in language technology that would be needed to support real breakthroughs of the types mentioned in questions 1 and 2 above?
4. What are the expected technological, economic or social developments that have to be considered as prohibitive or enabling factors in the planning of 1, 2 and 3 above?
5. In addition, and in order to have a better idea of the point of departure, we would kindly ask you to provide us with a list of the 3-5 most compelling language technology-based applications in use in the Media and Information Services sectors, at large, today.

The meeting participants prepared their answers which were presented and discussed during the meeting.

	Name	Organisation	City, Country	Sector	1st meeting attendance	2nd meeting attendance
1	Toni Badia	BM	Barcelona, Spain	META-NET	☒	☒
2	Nozha Boujemaa	INRIA	Paris, France	Multimedia content search, Digital libraries	☒	☐
3	Aljoscha Burchardt	DFKI	Berlin, Germany	META-NET	☐	☒
4	Nicoletta Calzolari	CNR	Pisa, Italy	META-NET	☒	☒
5	Roberto Cencioni	EC	Luxembourg	EC	☒	☐
6	Marin Dimitrov	Ontotext	Sofia, Bulgaria	Knowledge management, Semantic Web, Web Services	☒	☒
7	Christoph Dosch	IRT	Munich, Germany	Broadcasting technologies/ multimedia content search	☐	☒
8	René van Erk	Wolters Kluwer	Alphen aan den Rijn, The Netherlands	Global information services	☐	☒
9	Gil Francopoulo	LIMSI	Paris, France	META-NET	☐	☒
10	Robert Gaizauskas	University of Sheffield	Sheffield, UK	IE, Dialogue modelling	☐	☒
11	Gregory Grefenstette	Exalead	Paris, France	NLP/IE/IR	☒	☒
12	Marko Grobelnik	JSI	Ljubljana, Slovenia	META-NET	☒	☐
13	Christopher Kermorvant	A2iA	Paris, France	Document classification, advanced data	☒	☒

				extraction		
14	Hanna Klimek	EC	Luxembourg	EC	☐	☒
15	Maria Koutsombogera	ILSP	Athens, Greece	META-NET	☒	☒
16	Claude de Loupy	Syllabs	Paris, France	Information services, semantic search	☒	☒
17	Joseph Mariani	LIMSI	Paris, France	META-NET	☐	☒
18	Margaretha Mazura	EMF	Brussels, Belgium	META-NET, VG Convenor, VG Rapporteur	☒	☒
19	Alexandre Passant	DERI	Galway, Ireland	Semantic web, social software	☒	☐
20	Stelios Piperidis	ILSP	Athens, Greece	META-NET, VG Convenor	☒	☒
21	Georg Rehm	DFKI	Berlin, Germany	META-NET	☒	☒
22	Sergi Sagàs	MediaPro	Barcelona, Spain	Broadcasting technologies	☐	☒
23	Alessandro Tescari	Pervoice	Trento, Italy	Multilingual Speech	☒	☒
24	Hans Uszkoreit	DFKI	Berlin, Germany	META-NET	☒	☐
25	Philippe Wacker	EMF	Brussels, Belgium	META-NET, VG moderator	☒	☒
26	Jakub Zavrel	Textkernel	Amsterdam, The Netherlands	document understanding, web mining, text matching solutions	☒	☒

Table 1. List of Vision Group Media and Inform. Services meeting participants

3.3 Coverage

The members represent most of the industrial and research sectors addressed by this Vision Group. There is a broad coverage regarding the research and technology issues under discussion. While the group is very strong with regard to content and service providers, it needs to be extended further by information producers, such as broadcasting companies, news agencies, and social network representatives. The conveners of the group have identified those gaps and will focus their efforts on attracting new external members to join the group and thus ensure completeness in terms of sector coverage.

4 Visions on challenging and innovative LT-based scenarios

4.1 Domain-specific needs

1. **Audiovisual search.** Research and performance improvements are still needed regarding intelligent ways of recognizing and identifying objects, persons and actions in videos, as well as OCR and speech recognition of ordinary (non-trained) voices.
2. **Capture of web content at a European level** (like archive.org). The efficient utilization of millions of available Knowledge Bases (structured and linked data sources), linked data/annotated web, social/real time streams etc. would provide a larger number of texts in European languages for mining applications that are needed for domain structuring and would improve performance and scalability of LT. To reach this goal, it is necessary to overcome the constraints imposed by legal/licensing issues, and by lack of data access and availability (especially data derived from social networks).
3. **Large-scale text modelling**. The goal is to gather all the concepts associated to every discipline (medicine, finance, sports etc.), along with the characteristic vocabularies of every field or subspeciality, and a set of related content (videos, pictures, web pages) attached to each concept.
4. **Improvement of speech technologies**. Upgrade the quality of speech recognition and transcription (overlapping speech, named entity recognition, recognition of talk shows, speaker identification), as well as the quality of speech synthesis for speech translation.
5. **Industry- Academia synergies in Speech Recognition**. Text analysis and speech recognition technologies (SR) should be developed in a more integrated way both in the academic and the industrial field. SR researchers should approach the audio-devices industry in order to define new devices expressly designed for high-quality voice acquisition.
6. **Deeper understanding of text content**, in turn requiring: more robust, wider coverage parsing in all community languages – perhaps based on data-driven techniques. To this end, research advances are needed in recognition and interpretation of spatio-temporal expressions in text, in the induction of complex event structures from text, in recognition of entailment phenomena.
7. **Dialogue and Interaction modelling.** Richer understanding of information seeking in task contexts is needed (e.g. knowledge of the user, the user's task and relevant data sources) in order to improve dialogue modelling in various interactions. Also, the ways of interacting with a personal information assistant have to be defined with the goal of exploiting dialogues for efficient information gathering.

8. **Genre and text-type based layers in language analysis and generation**. The knowledge about text types needs to be formalized (also in a Semantic Web context), and be further incorporated into NLP and LT applications as standardized language resources.
9. **New types of GUIs**. Current GUIs are old-fashioned and complex-to-use and thus need to be upgraded with new functionalities.

4.2 Domain-independent needs

10. **Build competitive products for the European market** which is still fragmented along language boundaries:
    a. Identify the industries that would benefit from having a European home market rather than a local home market.
    b. Improve LT research so that the language barrier is eliminated for both the consumer and for SMEs.
11. **Multilingualism should be a priority**. The need for high-quality multilingual LRs for training and testing is evident. To this end, portability of methodologies and best practices across languages should be examined (e.g., LR size and production costs, use of metadata and standards, re-use of development methods and tools etc.).
12. **Standardisation**.
    a. The need of an initiative to reach a consensus on the relevant dimensions for annotation and the formats for representation regarding new methodologies for acquisition and representation (e.g., ISO on Time, Space etc.) has been widely acknowledged.
    b. The involvement of private/industrial actors (besides academic and institutional) in the process of standardization should also be pursued.
13. **Semantic search for a unified European market**
    A single market must be supported by easy searching and matching between supply and demand that so far are given in the form of unstructured or non-alignable sources. Searching and matching should be based on documents enriched with domain-specific semantic structure (concepts, categories, and relationships).
14. Need for **advanced (open source) tools that are available to both large industry players and SMEs**. This would have the result of providing (free) basic NLP infrastructure and resources for all EU languages. The focus should be set on basic technology and fundamental research in order to overcome the current state-of-the-art situation, which is too expensive, not accurate enough, very hard for non-experts to use, and sometimes time-consuming.

15. **Deliver highly skilled experts to industry**. High demand in trained staff has been attested by the industry, with the goal of delivering high quality services.
16. **Legal/IPR issues, political and economic/market considerations as well as ethical issues** need to be solved so that new ideas and opportunities can be effectuated.

4.3 Domain-specific visions

1. **Federated Audiovisual (AV) search**.
   a. Generation of high-level metadata (annotations) to retrieve **specific AV sequences across collections**. Semantic analysis of AV content is needed as well as intelligent search techniques, to **provide intelligent answers to everyday's questions**.
   b. Automatic ontology building based on individual human profiles. This would result in a **consumer-adapted, profiled and secure intelligent way of surfing along one's lines of interest, for business and leisure purposes.**

   (re. Need#1)
2. **Social stream text mining**, including sentiment/opinion mining, brand monitoring, real-time content filtering.

   Information overload handling would be enabled by real time content filtering and real time topic/trend detection.

   Semantic search that is accomplished through the combination of Language Technology and freely available knowledge bases and web annotations, i.e., linked data, and further **integration of linked open data into every information exchange** (Geonames, Musicbrainz, TheMovieDB, all the OpenGov initiatives and others).

   (re. Need#2)
3. a. **Life logging**. The challenge is to capture every utterance and vocal exchange during the day, and semantically structure the individual information. This would result in the creation of a gazetteer of concepts to **gather information on a massive scale and make it useful when someone is looking for something**.

   b. An example that would boost the EU LT industry is to **provide the information produced in the 27 member states of the EU (laws, but also books, movies, newspapers, Radio and TV broadcasts, etc) to the citizens of those 27 member states in their languages.**

   (re. Need#3)

4. **Automatic minutes production from meetings** (speaker identification, accurate speech recognition), **automatic indexing for voice search**, **speech transcription for videos** (subtitling, dubbing) and **live shows** (surtitling), **speech translation for videos**.
   (re. Need#4)
5. **Multimedia multilingual subtitling. This application would foster the access to information for ethnic minorities and new immigrants and facilitate information exchange between EU countries**.
   (re. Need#5)
6. **Improved text mining** in terms of quality, coverage and robustness, to be further exploited in search and in information navigation and presentation. The research advances that will enable this improvement are: recognition and interpretation of spatiotemporal expressions and event anchoring; knowledge base population through entity linking and slot filling; induction of event structures (minimally supervised or unsupervised event extraction); spatial language and reasoning about motion; dynamic logics for modelling motion with qualitative reasoning.
   (re. Need#6)
7. **Personalized task-centered interactive information assistants that know or adapt to: what the user knows or has already asked for, the user's language and the user's education and level of expertise.** The goal is to implement assistants who can interact (through speech and/or text) with the user to probe what they might be seeking, given the knowledge of the user's task.
   (re. Need#7)
8. **Genre-aware LT applications that can use, react and behave in a way that is best suited to a given communication situation.** Such applications would be enabled by the development of tools on text structure recognition and text type detection, as well as language resources that encode text type knowledge.
   (re. Need#8)
9. **Voice Control instead of GUI would be simpler in use and would provide trustful services**.
   (re. Need#9)
10. **Deliver more efficient online advertising** by applying semantic technologies to online advertising solutions.
11. **Semantic Annotation**. Extend research from sentences to discourses, from documents to dialogues, from artificial to natural interaction. Extend research applicability and coverage, including the domains of multimodal, multifunctional, interperson-

al communication, and cross-modal interaction in order to facilitate accessibility, e.g. produce LRs from real situations for processing emotions in Affective Computing, while also taking into account intentions, as well as LRs which include both actual data, and perceived data. To facilitate the production of large quantity of LRs, Web 2.0 or Web 3.0 should be used in a collaborative way, and LR could be continuously growing.

4.4 Domain-independent visions

12. Identification of boundaries to be eliminated and mass collaboration efforts to enrich an international knowledge base and communication code for a specific industry and/or to eliminate a boundary for making European SMEs more successful. Specifically, the focus should be on markets where a hub to provide collaboration and connectivity across a customer's process workflow can be created.
    (re. Need#10)
13. **Promote the Language Resources ecosystem**.
    a. More training, more education for the production (collecting and annotating) and use of LR should be promoted, while LR should also be used more widely in education, and consequently LR producers should be widely acknowledged.
    b. Cater for monolingual, multilingual, cross-lingual Language Technologies.
    (re. Need#11)
14. **A common infrastructure should better ensure interoperability** of the LR production and distribution/sharing processes. The foreseen production of the necessary LR appears as a humongous effort, nevertheless it seems wise to invest in web services to facilitate that task. A unanimous consensus should be reached on the priority of a shared effort, including science, industry and governments, with national and regional funding when appropriate, and EC contribution, ensuring a general agreement for a better share of experience.
    (re. Need#12)
15. Focus **funding on rapid progress in basic technology for semantic annotation and search** to provide the most accurate document understanding models in the shortest time, and using as little labelled data as possible. For that purpose, the next steps are to develop domain specific recognizers and ensure powerful semantic search for structured documents, as well as efficient algorithms for highly complex models (many features, many classes, a lot of structure, background knowledge). Research and technology advances should aim at high quality freely available basic resources and easy to use software toolkits, to ensure easy domain adaptation and transfer of knowledge.

(re. Need#13)

16. **Cost-effective porting of LT services and solutions across domains**. A balance between research and innovation for market products is a priority, so that LT solutions will be provided not by themselves but integrated in greater solutions, e.g., for multimedia content.
    (re. Need#14)
17. **Develop synergies among industry and academia** and bring academic excellence to the market.
    (re. Need#15)

5 Conclusions

This draft report has presented the output of the first two meetings of the Vision Group Media and Information Services. The draft results will serve as input to the road mapping work of the META Technology Council, which in turn will prepare the draft SRA. At the following meetings that will take place in 2011, Vision Group members will provide their feedback to the draft SRA by further refining and enriching their visions. A useful instrument towards this direction is the discussion forum that has been set up on the META-NET website. All Vision Group members and also other interested parties may use the forum to conduct a constructive dialogue that will be triggered by further controversial questions on the future and potential of media, on the role of government and public services etc.