



Fostering Language Resources Network

Grant Agreement No. ECP-2007-LANG-617001

<http://www.flarenet.eu>

A FORUM FOR Language Resources (LR) & Language Technologies (LT)

Nicoletta Calzolari (coord.)
glottolo@ilc.cnr.it

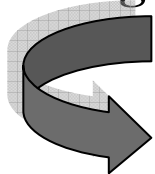


Old Slide: 2004

Political & Strategic Issues

Need of ... International research infrastructures for LRs

- How coordinate national initiatives (ERA)
- No HLT action line → difficulty in designing common global long-term strategies (more opportunistic)



**Need of bodies where to discuss
a broad research agenda & strategic actions
for W&S LRs together**

Technical scientific issues are obviously important

BUT organisational, coordination, political issues play a major role

SO ... should we organise ourselves???



Today

Why a Network of LRs & LTs?

Many dimensions around the notion of language

We need to consider together

- technical,
- organisational,
- strategic,
- economic,
- cultural,
- political issues wrt LRs & LTs

finally

A new EU Networking Initiative: FLaReNet

Sensitive

Political issues

e.g. a commonly agreed list of requirements for “national” LRs

Economic, social issues

- Applications
- Services

Cultural issues

→ Language ... and cultural identity

Multilingualism

Inter- & Multidisciplinarity

Technical, scientific issues

Need of bodies for a broad research agenda & strategic actions for LT&LRs (W/S /MM) based on all the dimensions





Not out of the blue

Today, many vitality & success signs



- In Spoken, Written, Multimodal areas in new emerging areas
 - Different dimensions & layers: Content (Ontologies), Emotion, Time, Cognition, ...
 - Statistical, data-driven approaches
 - Robustness, coverage, ...
 - For Training
 - For Evaluation
 - ...
-
- **LREC** (> 900 submissions; > 1100 participants)
 - Many LRs at **COLING** and even at **ACL**!!
 - **ELRA** (self-sustaining) & **LDC & Jap & Chin**
 - **LRE** (new Journal: N. Ide & NC)
 - **ISO-TC37-SC4/WG4** (International Standards for LRs)
 - **AFNLP...**
 - **New calls or initiatives** in EU, US, ASIA, on LRs, interoperability, cooperation, ...
 - **ESFRI - CLARIN**
 - **→ FLaReNet**



Today

The wealth of data & of basic technologies is such that:

■ We should reflect again at the field as a whole & ask if

□ Standards

← Content interoperability

□ Creation of LRs

← Collaborative creation & Manag.

□ Automatic acquisition

← Dynamic LRs

□ Distribution

← Sharing

are still “the” important components,
or how they have changed/must change

... Which new challenges/initiatives towards a
new & more mature infrastructure of LRs & LTs??





Many LRs & LTs exist, but a global vision, policy & strategy is still missing

Which Communities?

core

- Language Resources
- Language Technologies
- Standardisation
- Content/Ontologies
- System developers
- Integrators
- ...

for

- Many tasks & application domains
 - MT
 - CLIR
 - ...
 - e-government
 - content industry
 - intelligence
 - e-culture
 - ...

**EU
Forum**

with

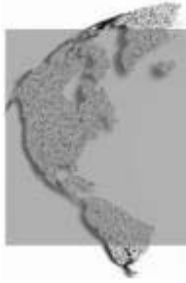
- EC
- National funding agencies
- Industry
- ...

Multilinguality

**FLaReNet
Network**

Focus on cooperation





FLaReNet

Fostering Language Resources Network



<http://www.flarenet.eu>

A European forum

- to facilitate interaction among LR stakeholders

The Network structure considers that LRs & LTs

- present **various dimensions**:
 - technical, but also organisational, economic, legal, political, ...
- must be approached from many perspectives & involve **various communities**

Addresses also **multicultural & multilingual aspects**

- essential when facing access and use of digital content in today's Europe
- **To provide recommendations for future actions** (to the EC, national agencies, industry)
- **To give priorities**
- **→ Need of 'visions'**





LRs, LTs & the Landscape



- **Definition of LRs: very broad**
 - Intertwined with LTs & also with usages, tasks, applications
 - Implications on the involved Community/ies
 - Definition: today & evolving
 - Wikipedia of LRs
- **The landscape today and in the future**
 - **Coverage:** both in the vertical & the horizontal dimensions
 - Portability, Interoperability, Basic Language Resource Kit, Multilinguality, ...
 - Which are the big challenges? The realistic ones? Those with more impact?
 - Driving forces? External factors? Against & in favour?
 - Main objectives?
 - ...





Objectives & expected results



The **largest Network of LR & LT players**, with diverse approaches, to **structure the area of LR & LT of the future** by discussing **new strategies** to:

- Enable progress toward **community consensus**
- **Recast the definition of LRs** in the light of recent scientific, methodological, technological, social developments
- **Consolidate** methods & approaches, common practices, frameworks & architectures
- **Integrate** so far partial solutions into broader infrastructures
- Convert existing LTs related to LRs into useful economic & societal benefits
- ⊗ **Anticipate the needs of new types of LRs / LTs**

A “roadmap”

- Identifying areas where consensus is achieved/emerging vs. areas where more discussion & testing is required, together with an indication of **priorities**
- ➡ Recommendations in the form of a **plan of coherent actions for the EU, national organisations & industry**



Ambitious!



A Layered Structure



Leading experts & groups (national, European, international institutions, SMEs, large companies) for all relevant LR areas (**about 40 members so far**)

- ◆ in collaboration with CLARIN, EuroMatrix, NSF INTEROP SILT & other world-wide initiatives
- ◆ to **ensure coherence of LR-related efforts in Europe**

To be enlarged

- ◆ with **new members &**
- ◆ the involvement of the community, as **supporters**





First FLaReNet Members



- ILC-CNR - Italy (*Coordinator*)
- ELDA - France
- ILSP - Greece
- LIMSI-CNRS & IMMI - France
- Universitat Pompeu Fabra - Spain
- Universität Wien - Austria
- Universiteit Utrecht - The Netherlands
- Max Planck - Germany
- Bielefeld University - Germany
- Morphologic - Hungary
- Prague University - Czech Republic
- Sheffield University - United Kingdom
- DFKI - Germany
- CEA LIST - France
- Copenhagen University - Denmark
- The Open University - United Kingdom
- Bulgarian Academy of Sciences - Bulgaria
- Vrije Universiteit - The Netherlands
- Romanian Academy - Romania
- Manchester University - UK
- Jožef Stefan Institute - Slovenia
- Adam Mickiewicz University - Poland
- Linguattec - Germany
- Langenscheidt - Germany
- Fondazione Bruno Kessler - Italy
- University of Antwerp - Belgium
- Universitat Politècnica de Catalunya - Spain
- University of Helsinki - Finland
- Universität Tübingen - Germany
- European Parliament - Luxembourg
- DGT - EC DG Translation - Luxembourg
- Vassar College - USA
- Kyoto University - Japan
- Brandeis University - USA
- LDC - Linguistic Data Consortium - USA
- National Institute of Advanced Industrial Science and Technology - Japan
- Korea Advanced Institute of Science and Technology - Korea
- Tsinghua University - China





FLaReNet Steering Committee



- ❑ Nicoletta Calzolari, ILC-CNR, Italy (*Coordinator*)
- ❑ Khalid Choukri, ELDA, France
- ❑ Stelios Piperidis, ILSP, Greece
- ❑ Joseph Mariani, LIMSI-CNRS & IMMI, France
- ❑ Nuria Bel, Universitat Pompeu Fabra, Spain
- ❑ Gerhard Budin, Universität Wien, Austria
- ❑ Jan Odijk, Universiteit Utrecht, The Netherlands





First FLaReNet Supportes



- Telefónica Investigación y Desarrollo – Barcelona, Spain
- Neurolingo - Athens, Greece
- Archetypon S.A. Information Technologies - Athens, Greece
- LIBRIS-TECH S.A. – Athens, Greece
- Innoetics Ltd - Athens, Greece
- Vecsys – Paris, France
- The Royal Institute of Technology (KTH), Department of Speech, Music and Hearing – Stockholm, Sweden
- ALPHABET Translation & Interpreting Services EPE - Athens, Greece
- ESTeam AB – Älgårås, Sweden
- Mikro Vaerkstedet – Odense, Denmark
- QUALIA Technologies of understanding - Athens, Greece
- European Bioinformatics Institute (EMBL-EBI) – Cambridge, United Kingdom
- Agence France Press (AFP) Medialab R&D Unit – Paris, France
- Bertin Technologies – Montigny-le-Bretonneux, France
- University of Edinburgh – Edinburgh, United Kingdom
- Translendum SL – Barcelona, Spain
- ELSNET Foundation – Utrecht, The Netherlands
- DAEDALUS-Data, Decisions and Language, S.A. – Madrid, Spain
- SYNTHEMA S.R.L. – Soluzioni Informatiche, Consulenza e Ricerca – Pisa, Italy
- France Telecom R&D - Lannion, France
- European Media Laboratory GmbH – Heidelberg, Germany
- Real Academia Española – Madrid, Spain
- Instituto de Engenharia de Sistemas e Computadores Investigação e Desenvolvimento em Lisboa – Lisbon, Portugal
- Norwegian Language Council – Oslo, Norway
- Siemens AG – Munich, Germany





Organised in Thematic Areas

Working Groups



- The **Chart for the area of LRs** in its different dimensions
- Methods and models for **LR building, reuse**, interlinking and maintenance
- **Harmonisation** of formats and standards
- Definition of **evaluation protocols** and evaluation procedures
- Methods for the **automatic construction and processing** of LRs

To build together:

- **Evolving RoadMap**
- **Blueprint of actions and infrastructures**





Outcomes of FLaReNet



The outcomes will be **of a directive nature**

- ❑ to **help the EC**, and national funding agencies, **identifying priority areas** of LRs/LTs of major interest for the public that need public funding to develop or improve

A **blueprint of actions** will constitute **input to policy development both at EU and national level**

- ❑ for identifying **new language policies** that support **linguistic diversity** in Europe
- ❑ in combination with **strengthening the language product market**, e.g. for **new products & innovative services**, especially for less technologically advanced languages

➡ **A European model for the LRs/LTs of the next years**





International Vocation



- Calls for **international cooperation also outside Europe**

To set up a global

- **Worldwide Forum for Language Resources and Language Technologies**





Join FLaReNet!



- We invite all interested players in the field to express their interest in becoming part of the Network
- **How to join?**
 - To be part of the FLaReNet Network fill the form available on the project website





Come to the FLaReNet Launching Event



The European Language Resources and Technologies Forum:
Shaping the Future of the Multilingual Digital Europe

Vienna, 12-13 February 2009

- See the website for the program (in progress)
 - Occasion for identifying the ground for (→ influence) future directions
 - Large space for discussion and interventions
 - Play a role in the definitions of recommendations for future actions
 - Online **free registration** open
- **A contest for best contribution!**

