

**WRITE**

**&**

**FLaReNet:**

**a new European Thematic Network  
for Language Resources and Language Technologies**

**Nicoletta Calzolari & Steven Krauwer**

Istituto di Linguistica Computazionale del CNR, Pisa, Italy

[glottolo@ilc.cnr.it](mailto:glottolo@ilc.cnr.it)



# WRITE Committee

## *Written Resources Infrastructure, Technology, & Evaluation*

Another initiative of Antonio Zampolli:

### ■ International forum

- to support international coordination & cooperation in the field of written LRs and technology assessment methodologies
- to formulate a common agenda of medium- and long-term research priorities
- to discuss innovative research needs

@ Involvement of various 'continents', topics, distribution centers

@ On the model of & in cooperation with **COCOSDA** for Speech



@ **Common meetings (Spoken & Written) at LRECs**

# A common Roadmap for S&W LRs and HLT

## A first list of priorities for S&W LRs:

- provide basic LRs for a larger set of languages
- increase multilingual LRs
- reduce development time of LRs
- enhance LR content interoperability
- foster synergies between S&W areas & with neighbouring areas (e.g. terminology, Semantic Web)
- develop new methodologies & tools for LR management, quick domain and application adaptation, data-driven tuning, etc.
- **Maintain Roadmap for Written & Spoken LRs & LTs (see LREC Roadmap Workshop)**
  - Evaluate Roadmap wrt state-of-art
  - Methodological issues
  - Users of Roadmap: e.g. extraction of 'hot topics'
  - ...

# Dissemination of info on standards & dissemination of standards

- Reporting on practical experience of standard use
- Look at models/guidelines for adapting standards to different purposes
- Look at good practice scenarios
- How to maintain standards
- How to evaluate standards
- **Promote the definition & adoption of shared standards for Written & Spoken LRs (e.g. through ELRA)**
- ...
- Advisory group of experts to help community

# Political & Strategic Issues

## Need of ... International research infrastructures for S&W LRs

- How coordinate national initiatives (ERA)
- How cope with a changing scene in funding scheme
- No HLT action line → **difficulty in designing common global long-term strategies (more opportunistic)**



ERA-Net  
&  
WRITE

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

# Many dimensions around the notion of language

We need to put together

- ✿ **technical,**
- ✿ **organisational,**
- ✿ **strategic,**
- ✿ **economic,**
- ✿ **political issues of LR**s



**Two new European Infrastructural & Networking Initiatives**  
**CLARIN & FLaReNet**

## Political issues

e.g. a commonly agreed list of requirements for “

## Economic, social issues

- ➔ Applications
- ➔ Services

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

- ➔ Language ... and **cultural identity**
- ➔ Language ... and **the Humanities**

**Multilingualism**

**Interdisciplinarity & Multidisciplinarity**

**Technical issues**

# We have still a long path ...

## & also a “new project”

in an *e-Contentplus* Call for a:

- “Thematic Network on Language Resources”:

### **FLaReNet**

- To provide common recommendations (to the EC) for future actions
- To give priorities
- Need of ‘visions’

In a global context, in cooperation with *CLARIN*  
& also with *non-EU members*

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

## Which Communities?

core

- Language Resources
- Language Technologies
- Standardisation
- Ontologists
- Content

EU  
Forum

for

- Humanities
- Social Sciences
- Digital Libraries
- Cultural Heritage
- ...

CLARIN  
ResInf

FLaReNet  
Network

Multilinguality

Focus on cooperation

for

- EC
- Funding agencies
- ...

for

- Many **application** domains  
(eculture, egovernment, ehealth, intelligence, domotics, content industry, ...)



# FLaReNet

## Fostering Language Resources Network

<http://www.ilc.cnr.it/flarenet/>

### A European forum

- to facilitate interaction among LR stakeholders

The Network structure considers that LRs present **various dimensions** and must be approached from many perspectives:

- technical, but also
- organisational
- economic
- legal
- political

Addresses also

- multicultural and multilingual aspects**, essential when facing access and use of digital content in today's Europe

# Organised in Thematic Working Groups

- A **layered structure**, with **leading experts & groups** (national and European institutions, SMEs, large companies) for all relevant LR areas (**about 40 partners**)
- ◆ in collaboration with CLARIN
  - ◆ to **ensure coherence of LR-related efforts in Europe**

FLaReNet will

- @ **consolidate** existing knowledge, presenting it analytically and visibly
- @ contribute to **structuring the area of LRs of the future** by discussing **new strategies** to:
  - ◆ convert existing and experimental technologies related to LRs into useful economic and societal benefits
  - ◆ integrate so far partial solutions into broader infrastructures
  - ◆ consolidate areas mature enough for recommendation of best practices
  - ◆ **anticipate the needs of new types of LRs**

# Thematic Areas

- 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**

# Objectives & expected results

The **largest Network of LR and HLT players**, with diverse approaches, efforts and technologies

- ④ Enable progress toward **community consensus**
- ④ Give an extended picture of LRs & **recast its definition** in the light of recent scientific, methodological, technological, social developments
- ④ **Consolidate** methods & approaches, common practices, frameworks and architectures
- ④ A **“roadmap”** identifying areas where consensus has been achieved or is emerging vs. areas where additional discussion and testing is required, together with an indication of **priorities**
- ④ Recommendations in the form of a **plan of coherent actions for the EU and national organizations**
- ④ A **European model for the LRs of the next years**

Ambitious!

# Outcomes of FLaReNet

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

- to help the EC, and national funding agencies, **identifying priority areas** of LRs 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

# These Initiatives, ... together

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

and will be relevant for

- setting up a global **worldwide Forum of Language Resources and Language Technologies**