

**Language Resources:
a pillar of Language Technology**

Nicoletta Calzolari

Istituto di Linguistica Computazionale – CNR
Pisa, Italy
glottolo@ilc.cnr.it

Even if language resources have a relatively short history, they are recognised as one of the pillars of language technology.

Their traditional production process is too costly. It is urgent to create a framework that enables effective cooperation of many groups on common tasks, adopting the paradigm of accumulation of knowledge so successful in more mature disciplines, such as biology, astronomy and physics. This requires a change in the paradigm, and the design of a new generation of language resources, based on open content interoperability standards. The semantic web notion may help in determining the shape of the language resources of the future, consistent with the vision of an open distributed space of sharable knowledge available on the web for processing. This enables building on each other achievements, integrating results, and having them accessible to various systems and applications. This is the only way to make a great leap forward.

A new generation of multilingual language resources – based on these principles – will serve better the needs of today and tomorrow language technology and language applications, semantic web. This is true in particular for coping with the need of more and more ‘knowledge intensive’ large-size language resources for effective multilingual content processing.

A new paradigm of R&D is emerging, pushing towards the creation of open and distributed language infrastructures for language resources and language technology, based on sharing language resources and tools. I will highlight current European infrastructural and networking initiatives in this area, such as *CLARIN*, *FLaReNet* (*Fostering Language Resources Network*), and *META-NET / META-SHARE*, which may have influence in how we shape the future of our field. I will also highlight the relevance of the new *LRE* (*Language Resources and Evaluation*) *Map*, a collaborative means of highlighting the current language resource landscape and its trends.

Technical scientific issues are obviously important, but organisational, coordination, political issues play a major role in our field as in every other. Technologies exist and develop fast, but the infrastructure that puts them together and sustains them must be created and sustained.

The Multilingual Web is crucial to the success of such an infrastructure, critically based on interoperability, aimed at improving sharing of language resources and accessibility to multilingual content, thus stimulating automation and innovation, especially in the translation industry.