Consortium CORLI
COrpus, LAnguages and Interactions

PRESENTATION

CORLI (CORpus, Languages and Interaction) is a consortium of Huma-Num (https://huma-num.fr) dedicated to the sharing of methodological approaches, tools and software, best practices and training within the community of linguists building and investigating corpora.
Organising such a community is particularly complex because people have different theoretical and methodological approaches, practices and needs. The development of digital humanities and widespread access to data have introduced significant changes in linguistics, both on the methodological and theoretical levels.

OBJECTIVES

The goals of CORLI are to promote corpus visibility, reusability , easy access, good practices in corpus linguistics, as well as to help developing corpora, tools and formats. CORLI helps researchers carry out complex tasks they would not have the means to achieve without this support. Researchers can ask CORLI to help them, especially in domains in which they do not have enough expertise and knowledge. The goal of CORLI is then to provide information and support, and help researchers get into working groups to examine collectively such questions. CORLI assists them in disseminating the recommendations and standards that they developed. CORLI has general and specific actions according to the needs of the community.

ORGANISATION

Governance

Christophe PARISSE, Modyco, Inserm & Nanterre University
Céline POUDAT, BCL, Nice Cote d'Azur University

Direction

33 researchers and engineers from as many as possible French linguistic laboratories (more than 22)

Steering committee

In all working groups, more than 180 participants representing various research fields

Working process

CORLI has a bottom-up approach. The practices that researchers most use are the ones which are promoted and disseminated. If different or even opposite practices coexist, CORLI will help researchers exchange and discuss to find, when appropriate, common solutions.
Consensus helps to promote recommendations on the use of common standards and norms, making ressources available for the largest audience, and following common guidelines and practices.

Working groups
Thematic actions and priorities

- **WG 1 – Interoperability / Corpus Exploration Tools and practices**
- **WG 2 – Multimodality and new forms of communication**
- **WG 3 – Multilingual and plurilingual corpora**
- **WG 4 – Legal advice and corpus rights**

General actions targeting the whole community

- **GA 1 – Finalisation of corpora (financial support – about 50K € each year)**
- **GA 2 – Training sessions on tools and practices**
- **GA 3 – Corpus evaluation, FAIR principles**
- **GA 4 – Corpus repositories and standards**

Some examples of CORLI achievements in the last few years

- participation in standards and recommendations for the encoding of corpora and metadata
- creation of software to help conversion of format and metadata edition
- creation of white papers and recommendations in various domains (legal issues, CMC corpora, corpus exploration)
- dissemination of good practices in workshops and meetings
- community awareness of CLARIN

Mailing list: corli@groupes.renater.fr