



FlareNet: International Cooperation

A Data Center Perspective

Khalid Choukri, Christopher Cieri

Profiles of ELRA and LDC

- Two dimensions
 - Activities carried out
 - Identification (collection, dissemination of information)
 - Distribution & Licensing of LRs but also sharing
 - Curation
 - Production of LRs from scratch or enhancement
 - Evaluation of HLT, rankings, evaluation packages
 - Population being served
 - principally HLT
 - some social science



PreRequisites for Cooperation

What are the pre-requisites to international cooperation in our communities?

common need

challenge

respect, trust, flexibility

favorable cost/benefit ratio

ELRA (EU) ↔ LDC Connections

- Joint Distribution
 - TED, also involving ELSNET
- Special Distribution Arrangements
 - LDC Data in TC-Star, MEDAR
 - ELRA EMILLE data in REFLEX LCTL
- Joint Development & Distribution
 - Net-DC Arabic Broadcast News Corpus
- LDC Participation in LREC
 - 14 employees, 18 papers (not counting these workshops, panels, chairpersonships)

ELRA (EU) ↔ LDC Connections

- Subcontracts
 - Med-LTC on GALE
 - Tool & Specification Exchange
- Zampolli Prize to Liberman
- Joint Missions
 - Amazigh in Morocco
- Annual Exchange Program

Multilingualism

- Multilingualism as a unifying issue?
 - need to work with multiple languages should, sometimes does, bring together, contributors from traditionally separate communities
 - However, the same factors that lead to the divergence of language may create barriers to cooperation.
 - isolation, bias
 - not created here syndrome
 - ...

Interoperability

- How to boost the Interoperability issue?
 - Why/how does practice vary
 - purpose, target user (community)
 - types of LR that need to interoperate
 - same: increases volume, diversity
 - different: enables new insight
 - source material
 - lack of specification, human variability
 - annotation
 - nature, expertise required, inherent variability
 - producer
 - epoch

Interoperability

- How to boost the Interoperability issue?
 - Challenges to Interoperability
 - scope of planning
 - consistency versus innovation
 - tactical versus strategic development

Interoperability

- Other Approaches
 - Two Existing LRs
 - Linking
 - Merging
 - Converting, revising, re-designing to be interoperable
 - One Existing, one New LR
 - Build new LR to imitate old
 - Change annotation of old
 - Add annotation upon old
 - New LRs
 - Build with interoperability in mind
 - Specifications (Henry's comments on value of LR based upon quality of description)
 - Use of standards
 - Use of data category registry
 - Build multipurpose LR
 - Design suites of interoperable resources

Priorities

- How to draw priorities per language? How to conduct a BLARK/ELARK for each?
 - centralized the knowledge of
 - what constitutes a BLARK/ELARK
 - regional priorities for target languages
 - EU priorities for EU and neighbor languages
 - other regions naturally have
 - other languages
 - other immediate priorities
 - different parameters which affect the priorities

Common Strategies

- Common Strategies for the next years? In which framework: local, regional, international frameworks
 - How is “international” defined?
 - Within EU, where frameworks may help encourage sharing?
 - Among more geographically diffuse groups lacking shared history and regulation?
 - What are the prospects for creating broadly international frameworks like those that affect the EU?

Who/What Cooperates?

- Cooperation through the initiatives or by people (consensus/ cooperation through discussion or also through actions)?
 - Yes.
 - Net-DC Example

Modalities of Cooperation

- Realistically, there is no way to avoid multiple (instances of) LRs addressing the same purpose, data centers.
- However, those that cooperate effectively increase their chances of
 - excelling
 - persevering
 - attracting funding
 - influencing research policy
- However, in order to succeed, such cooperation requires
 - public support
 - formal structure
 - infrastructure (fora)