Call for interest: A Swedish BLARK
Workshop proposal for GSLT retreat 2006

Organisers

- Anna Sågvall Hein, anna@lingfil.uu.se (chair)

- Eva Forsbom, evafo@stp.lingfil.uu.se

Motivation

At the closing seminar of the VINNOVA language technology program, Nationell konferens i språkteknologi, 20-21 Oktober, 2005 Uppsala[1], the subject of a Swedish BLARK (Basic LAnguage Resource Kit (Krauwer, 1998)) was brought up for discussion, and Anna was asked to be convenor for a first meeting with interested parties, so here we go.

The idea of a BLARK is to define a minimal set of language resources (LRs) to be made available for as many languages as possible, Swedish in our case, and map the actual gaps which should be filled in order to meet the needs of the HLT field (Mapelli and Choukri, 2003). A BLARK is meant to be used for education and research, and pre-competitive industry needs. BLARK initiatives for at least two languages are on the way: Dutch/Flemish (Strik et al., 2002) and Arabic (Krauwer et al., 2004).

ELRA/ELDA provides four matrices for LRs online where potential BLARK definers and instantiators can grade resources for various applications or modules to get an overview of what LRs are needed, and set up a priority list for developing the LRs that are missing (with the Arabic experience as example).

Topics

After a short introduction by the organisers, and any submitted contributions, the floor is open for discussion on the following topics:

- Organisation

- Funding

- Definition

- Grading

- Inventory of LRs and tools

---

[1] http://www.lingfil.uu.se/info/nkst05.html

- Priorities

- Development of missing LRS and tools

- Distribution

- Availability and open source aspect

Participants who would like to contribute with, for example, a talk during the first part should contact the organisers as soon as possible.

Equipment needs

- Projector

- Internet connection

References

OLAC (Open Language Archives Community). URL `http://www.language-archives.org/`.

ELDA (Evaluations and Language resources Distribution Agency). URL `http://www.elda.org/blark/`.

Eva Forsbom. Återanvändbarhet för språkvara, 2004. URL `http://stp.lingfil.uu.se/~evafo/gslt/java/paper.ps.gz`. Literature review for GSLT Java course.

Steven Krauwer. ELSNET and ELRA: A common past and a common future. *ELRA Newsletter*, 3(2), 1998. URL `http://www.elda.org/blark/fichiers/elsnet&elra.doc`.

Steven Krauwer, Bente Maegaard, Khalid Choukri, and Lise Damsgaard Jørgensen. Report on BLARK for Arabic, 2004. URL `http://www.nemlar.org/Publications/index.htm`.

Valérie Mapelli and Khalid Choukri. Report on a (minimal) set of LRs to be made available for as many languages as possible, and map of the actual gaps. ENABLER Deliverable 5.1, 2003. URL `http://www.enabler-network.org/reports.htm`.

H. Strik, W. Daelemans, D. Binnenpoorte, J. Sturm, F. de Vriend, and C. Cucchiarini. Dutch HLT resources: From BLARK to priority lists. In *Proceedings of ICSLP, Denver, USA, pp. 1549-1552*, pages 1549–1552, Denver, USA, 2002. URL `http://lands.let.kun.nl/literature/strik.2002.2.pdf`.