

MLT Language technology resources (LT2304), Fall 2011, take-home exam

Below, you will find nine (9) take-home exam questions. Out of these, you should individually choose five (5) to which you will write essay-format answers (1–2 A4 pages per question, plus references, in 12-point Times [New Roman]). The answers are to be sent by email to me <lars.borin@svenska.gu.se> as pdf-format attachments by 31st October 2011.

In the answers, you should illustrate your reasoning and your claims by concrete examples and support them by references to relevant literature. For a passing grade, references to course readings will suffice, but in order to qualify for a pass with distinction you will need to consult other work as well.

1. Explain the notion of BLARK. How is the BLARK defined and motivated?
2. LT resources can be acquired/collected or created completely automatically, (mainly) manually, or in a process having both automatic and manual elements. Discuss the pros and cons of the two extremes – automatic and manual creation – in terms of prerequisites and likely outcomes, for *data resources* (in general or a specific kind of data resource).
3. LT resources can be acquired/collected or created completely automatically, (mainly) manually, or in a process having both automatic and manual elements. Discuss the pros and cons of the two extremes – automatic and manual creation – in terms of prerequisites and likely outcomes, for *tool resources* (in general or a specific kind of tool).
4. For automatic evaluation of LT resources – tool resources in particular – there is a fairly standard method found in the literature. Describe this method and discuss its pros and cons.
5. Motivate the need for (text) corpus resources.
6. Motivate the need for lexical resources.
7. Motivate the need for speech resources.
8. Standardization of data formats and content models is increasingly felt to be important for the effective use of LT resources. Elaborate on this by discussing the main issues involved in both kinds of standardization and factors that stand in the way of the adoption of standards for LT resources.
9. Describe how LT resources can be used to build e-science applications. What are the key characteristics of disciplines where this is feasible and what are the hoped-for outcomes of such an endeavor?