

Panel on Interoperability

Interoperability is a goal of many recent language archive initiatives, e.g. AILLA, AIATSIS, E-Meld and DOBES¹; and a number of different practices have been proposed as potential standards. A recent paper “Seven Dimensions of Portability for Language Documentation and Description” (Bird & Simons, 2001) discusses interoperability primarily from the perspective of the Open Language Archives Community (OLAC); and some of the questions below are taken from that paper. But in this panel, we would like the panelists to address interoperability issues from a wider perspective as well, thinking, for example, about the social interactions necessary to achieve interoperability, the role of best practice guidelines within the linguistic community, and the impact of best practice recommendations on the individual fieldworker.

We have listed a number of possible topics below, in the form of suggested questions for the panel. However, we do not want these questions to restrict the discussion so much as to provide a springboard into it. Please, feel free to choose among the topics listed below, and to add other relevant topics of your own.

We will provide the audience with a handout listing the questions to be addressed by the panel. If you have questions or issues you would like to see included in this handout, please send them to hdry@linguistlist.org or peter.wittenburg@mpi.nl.

Social Aspect of Interoperability

- What kinds of interactions are necessary to establish best practice guidelines and to achieve interoperability?
- How can we convince colleagues to follow recommendations?
- How can we create confidence in frameworks?
- What is the scope of best practice guidelines?

Fieldworkers View of Interoperability

- Is there a conflict between best practice guidelines and convenient operational environments, and, if so, what benefits or penalties accrue to the fieldworker from following best practice guidelines?
- How to solve the potential conflict between fieldworkers and archivist requirements?
- Whose task is it to achieve interoperability?

Frameworks and Standards

- What is the value of standards in a continuously changing and migrating world of language resources on the Web?
- Which frameworks are available to establish interoperability in the linguistic domain?
- What is the relevance of common terminology databases?
- What is the relevance of the Resource Description Framework?
- Are there other relevant frameworks?
- Is the ISO TC37/SC4 work of relevance for the domain?

¹ There are a many more which cannot be explicitly mentioned here.

Content and Structure²

- Is it feasible to unify linguistic terminology?
- Which kind of frameworks with a unifying potential for terminology are available?
- What types of mapping do we have to consider?
- What resources are or should be available to define linguistic terminology or “translate” from one terminological system to another?
- How should we describe content relevant structure? Do we need unification? Is structural unification feasible?
- How important is it to include explanatory documentation with the data archived, e.g. to include definitions of the linguistic mark-up along with the marked-up data?
- Is the work involved in careful documentation feasible for the scientists?
- Do we need open repositories where one can easily find definitions? How open do these repositories have to be?
- How important is it to provide the raw video and audio material on which a language description is based?
- Should archives make it a high priority to secure either type of documentation?

Format and Structure³

- Should archives insist or recommend that data be stored in open formats?
- In how far does this decision effect the fieldworker?
- How can we identify open formats which we can rely on?
- Which open formats do we have and which are stable?
- If not, should archives facilitate format conversion in any way?
- How to define and document formal structures of language resources?
- Are there reliable standards which we can use?
- Is encoding of characters with Unicode feasible at this time? What problems are likely to be encountered?
- Do we have to provide character code mapping tables?

Discovery

- Metadata descriptions are important for resource discovery. What are the requirements and trends and how to achieve interoperability on this level?
- How important is language identification and linguistic data type to the linguistic community?
- Which controlled vocabularies can be agreed upon?
- Which suggestions are relevant to improve the codes for language identification? (see <http://linguistlist.org/find-lang.html>)

Access

- What delivery methods (e.g., CD, DVD, web, print) are appropriate for different types of digital resources to interchange data?
- How will access methods change over time dependent on technological developments and new collaboration methods?

Citation

- How dynamic will online linguistic resources be? Which parts can be seen as static? How to document their existence and modifications?

² Content comes along with structure. Some lexical entries get meaning by looking at their structural embedding.

³ The term “format” as used here includes formal structure such as annotation schemes.

- What standards currently exist for identification and citation of digital resources? Are these adequate? For example, how widespread is the use of a persistent Digital Object Identifier?
- How necessary is it to provide fixed versions of resources, with version numbers, or the ability to identify components of resources uniquely? What practical problems, if any, do these pose for the archive?