Using ELAN for annotating sign language corpora in a team setting

Onno Crasborn
Centre for Language Studies
Radboud University Nijmegen
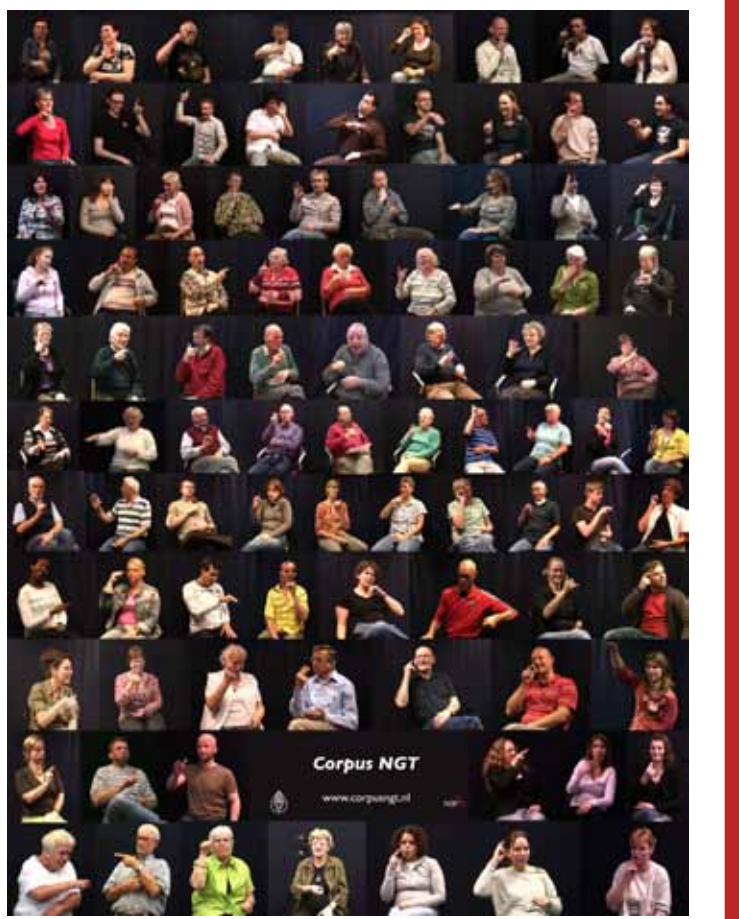
Han Sloetjes
Max Planck Institute
for Psycholinguistics, Nijmegen

Background

ELAN was originally created with the idea of a single user working on a single document. Within the Corpus NGT project and subsequent linguistic research projects, the functionality to use ELAN to work on **large numbers of documents** in a **research team with various interests** has been slowly expanding. This poster shows some of these functions within ELAN version 3.8.1, and it shows how a complementary solution that has been developed to share files wihin a research group.
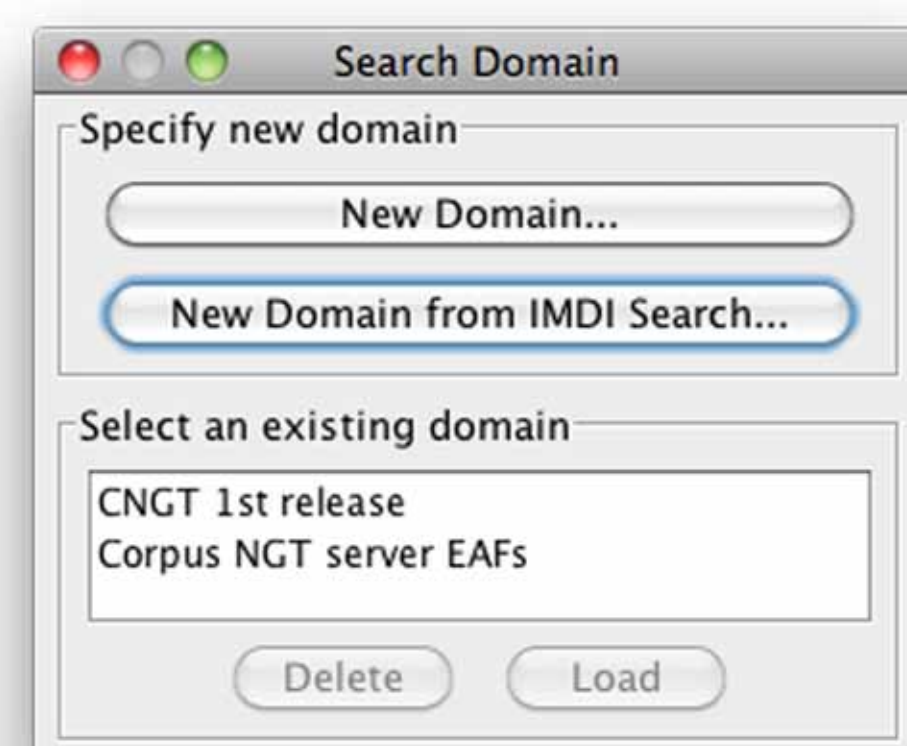
Context

Corpus NGT: 2375 files, used by 14 researchers and assistants simultaneously for different research projects. Total of >70 tiers.
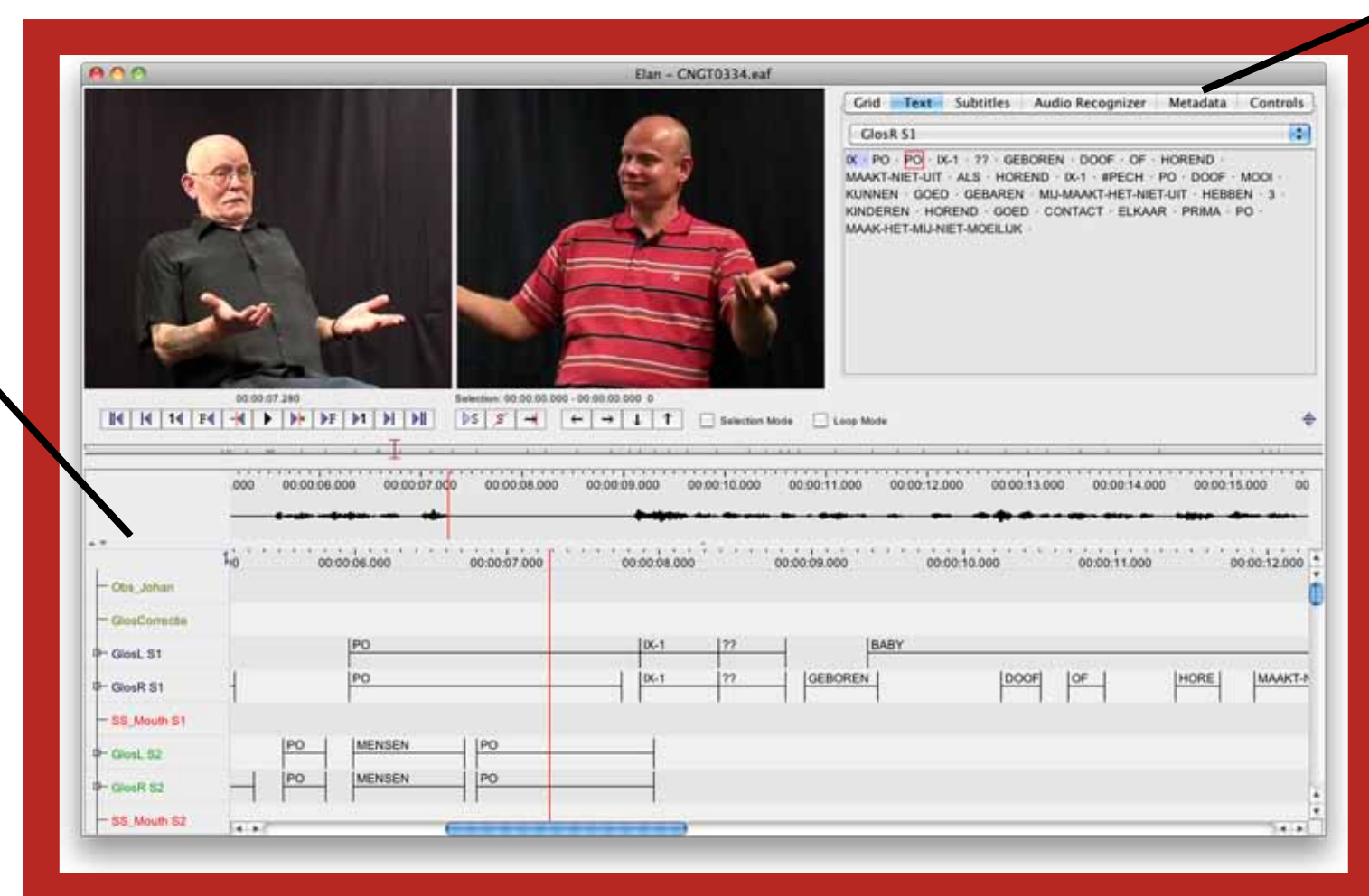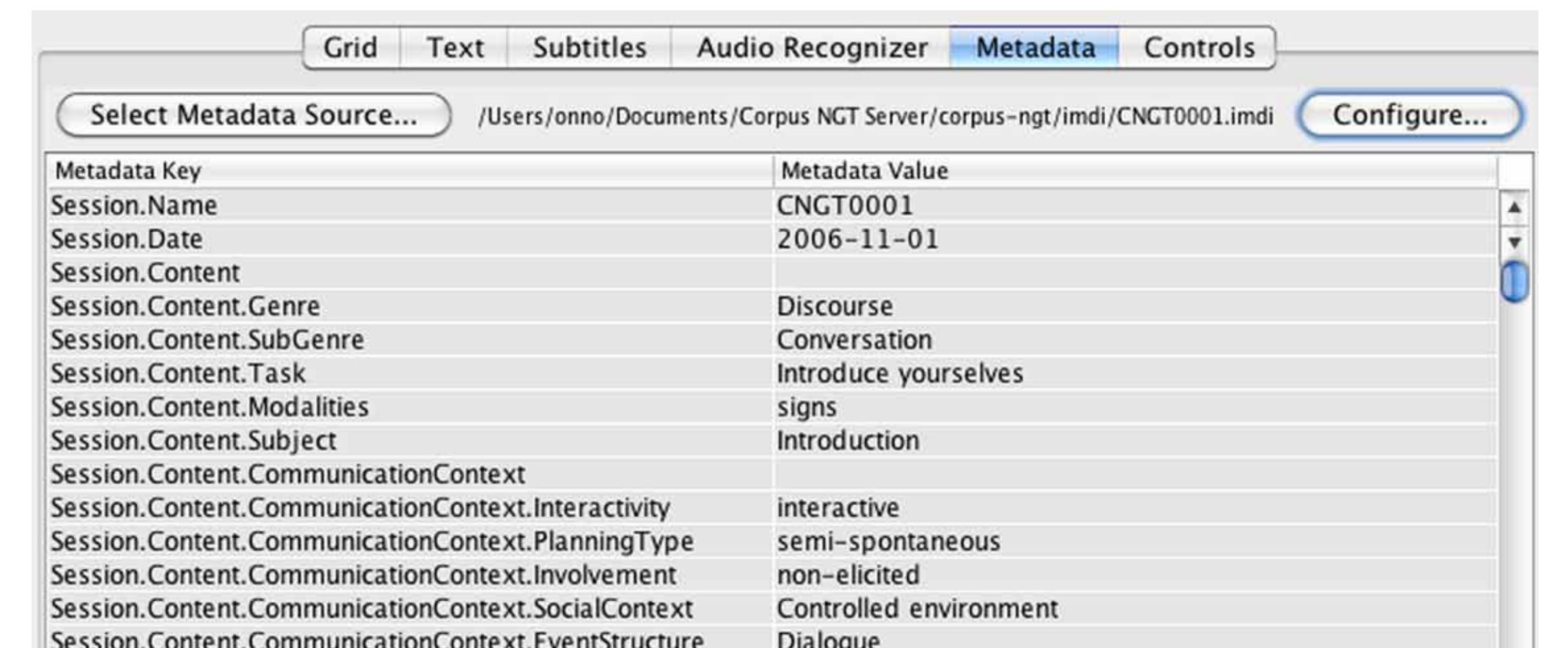
Colours of tier labels help to visually group information in a large set of tiers.
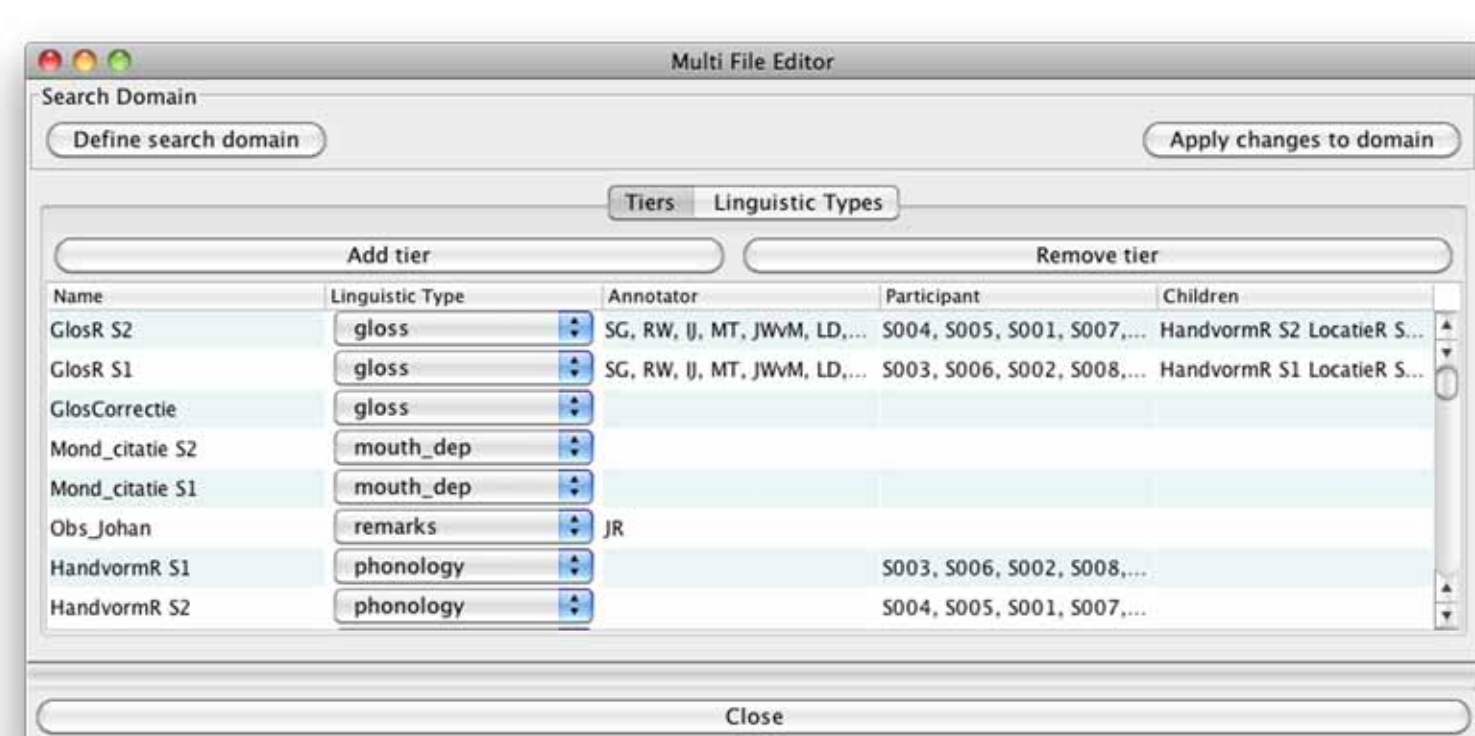
Search results from the IMDI metadata browser ('all men from Amsterdam') can be exported as a domain for ELAN searches.
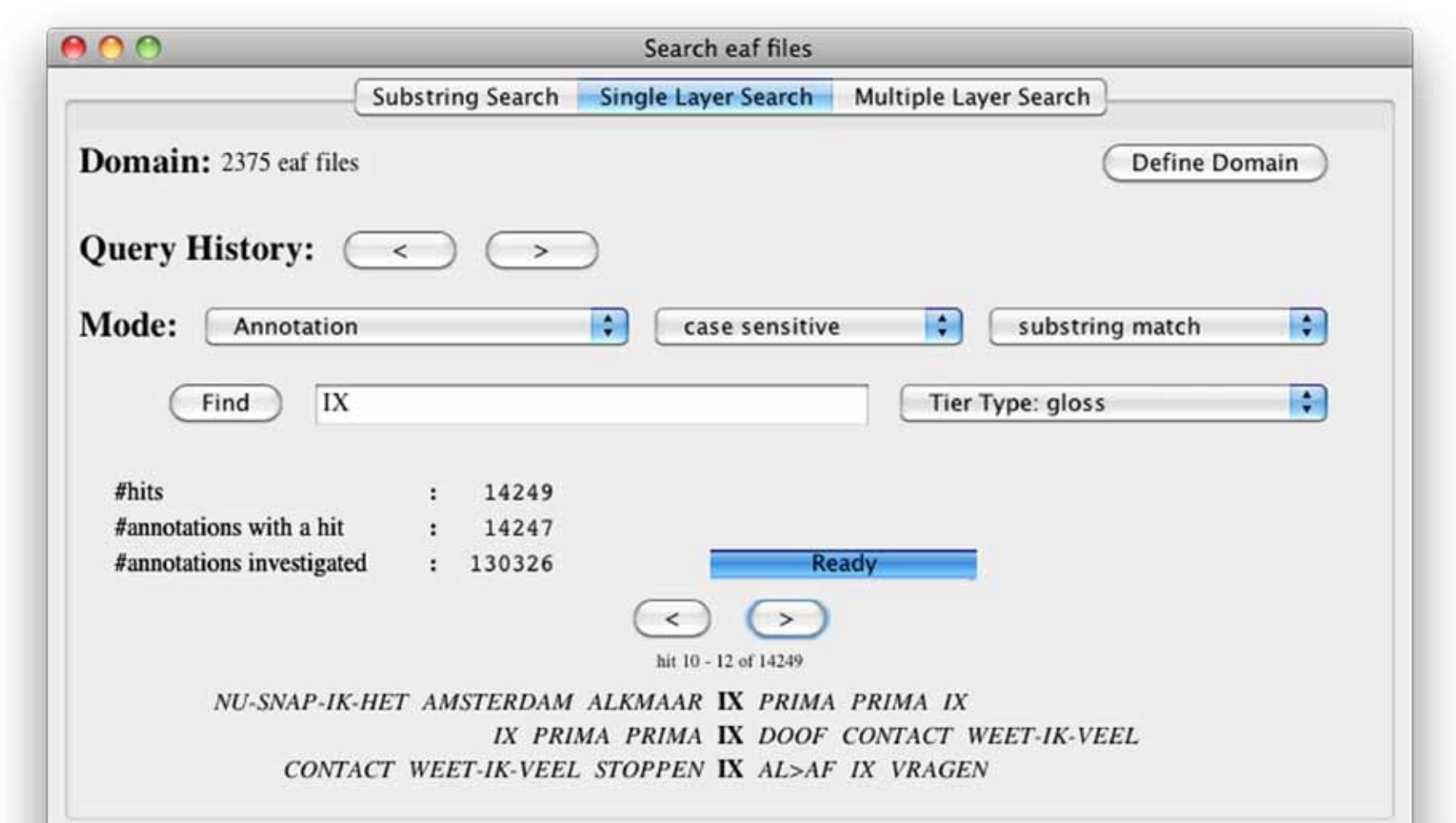
Metadata from related IMDI documents can be visualised in ELAN.

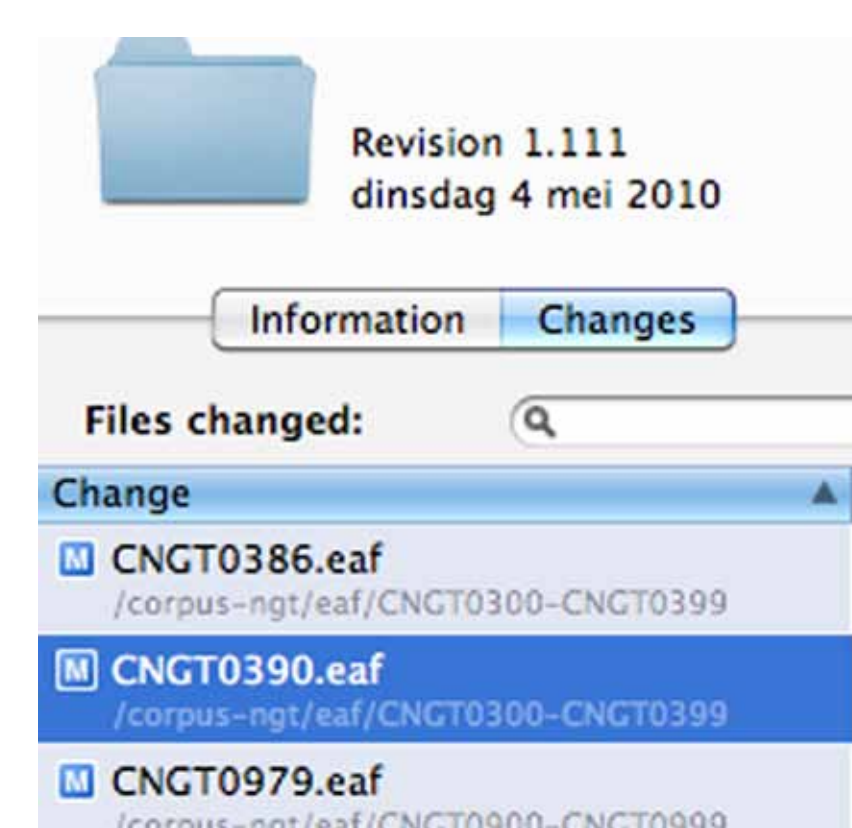Various actions can be performed on a selection of multiple files, including three types of searches.

The Multiple File Editor allows for adding and deleting tiers in a selected domain. By displaying all the properties that tiers have in this whole domain (Linguistic Type, Annotator, Participant), the function can help to keep large corpora well-organised.

Other actions on multiple files:
- find and replace
- generation of statistics
- create new documents based on template and sets of media files
- annotation 'scrubbing' (removal of superfluous spaces, tabs, and newline characters)
- export as word list for a selection of tiers

All annotation and IMDI documents are stored on a Subversion (SVN) server. Users work on a local copy of the archive, and publish their modifications to the server. This setup requires that users 'lock' the file they want to work on, and always ensure that they keep their local copies up to date.

Different versions of preferences files have been created so that users see a selection of tiers that is relevant fo their work. The folder to use can be set in the ELAN preferences.

The SVN server keeps log info about who did what to which file. It stores a copy of every change ever submitted to the archive, and provides tools for seeing changes that have been made and resolving conflicts.