Elias Lönnrot Letters online

The *Elias Lönnrot Letters* publishes the facsimile images and transcriptions of the approximately 1800 private letters sent by Lönnrot. The edition is built on the open source digital publishing platform *Omeka* and is designed for user groups like cultural historians and Folklore Studies researchers.

The guiding principle of the *Elias Lönnrot Letters* online edition is open access, not only in terms of the open source platform but also the published source material. The authors did not try to predict what scholars and other likely users of the edition would find useful, but rather to make the data available for free use and processing according to each use case's needs. Thus, while the edition is designed for easy viewing online, arranged to collections according to each recipient and with the facsimile images and transcriptions viewable side by side, emphasis is on shareability. The data is made available for download in several formats. While providing easy access to the data, this shareability opens new avenues for cooperation with other actors in digital humanities. The SKS is already working with The language Bank of Finland (*Kielipankki*), who will receive the XML/TEI5 files of all the letters contained in the edition; the international READ project, which develops tools for machine reading of hand-written text; and the STRATAS project, which develops linguistic tools and uses Lönnrot's letters as part of their data.

The edition makes the transcriptions of all the letters available as plain text for processing with qualitative data analysis tools like *Atlas*. Similarly, the statistical data produced by the open text search, which uses Apache's *Solr*, can be exported for easier processing. This kind of statistical data may reveal aspects that might remain hidden when examining a large mass of source material document by document.

Most importantly, the XML/TEI5 files of all the letters are made available for download. The entire edition can be obtained at once, or parts of it separately - for example individual recipients' collections or the XML/TEI5 files of letters found by the open text search. Due to the design, it is not possible to incorporate tools needed by linguistic researchers in the edition. However, many such tools are already provided by *Kielipankki*: for example *Korp*, a concordance tool for corpus queries based on morphosyntactic analysis, and a dependency parser for Finnish text. The free availability of the XML/TEI5 files makes Lönnrot's letters accessible for research with these as well as various other linguistic tools (topic modeling, morphological analysis, entity identification etc.). Likewise, the data is readily available for visualization tools such as map applications.

Maria Niku, PhD (History), BA (Computer Science), Software programmer

Biographical Centre, Finnish Literature Society

Hallituskatu 2B, PL 259, FI-00171 Helsinki

tel +358 40 7791655, maria.niku@finlit.fi