DIGITIZATION PROJECT OF FINNO-UGRIC LANGUAGES

Jussi-Pekka Hakkarainen
Project Manager
National Library of Finland
Research Library
28 May 2013, Göttingen
PILOT PROJECT, 2012-2013

The National Library of Finland is executing a pilot project (July 2012-October 2013) that aims for digitizing of Finno-Ugric material for the benefit of the linguistic research within the Language Programme of Kone Foundation (2012-2016).

During the pilot project the National Library of Finland will produce the research infrastructure (repository and OCR editor) and takes care of cooperation with the Finnish (Helsinki University Library) and Russian partners (National Library of Russia and National Library Resources).
The researchers made the selection of materials for the digitization plan in autumn 2011.

The selection consists of 17,000 pages of monographs in Veps, Ingrain, Mari and Mordvinic languages and around 20,000 pages of newspapers in Mari and Mordvinic languages.

Monographs are mainly school and text books from the 1920’s and the 1930’s and in many cases they are translations from Russia to the local languages.
COPYRIGHTS

Since Language Programme of Kone Foundation has an objective of free access to the materials, National Library of Finland will publish the material as its own repository and provides an open access without geographical or IP restrictions.

In order to publish the material as public domain, the copyrights regarding material was needed to be cleared. The research on copyrights was conducted by Moscow-based National Library Resource during winter 2013.
OPEN REPOSITORY (FENNO-UGRICA)

FENNO-UGRICA

Фенно-угрика

Фенно-угрика – однородная финно-угорская коллекция Национальной библиотеки Финляндии. Коллекция содержит публикации на ижорском, вепском, марииском (горномарийский и угонаройский) и мордовских (марийский и мокшанский) языках, а также газеты на марийских и мордовских языках, опубликованных в основном в 1920-е и 1930-е годы. В целом коллекция содержит более 120 монографий и почти 20 000 страниц текста газет.

Представленные здесь электронные ресурсы созданы в рамках Проекта оцифровки публикаций на финно-угорских языках в фондах Российской национальной библиотеки в Санкт-Петербурге. Проект является частью Грантов Фонда Кена. Материалы оцифрованы в Российской национальной библиотеке и публикуются на основании исследования, выполненного Национальным Библиотечным ресурсо, по проверке наличия или отсутствия исключительных прав на включение в Проект издания.

В рамках проекта для лингвистов также разработан OCR-редактор, основанный на открытом исходном коде, который позволяет редактировать OCR-тексты на финно-угорских языках. Право редактировать ресурсы данной коллекции предоставляется, при этом, финноугроведам. Права пользователя и редактирования можно получить у администрации Проекта по оцифровке финно-угорских языков. Дополнительная информация и контакты по электронному адресу: kk.fennougrica@helsinki.fi

Разделы и коллекции

- Монография (34)
- Газеты (0)


THE NATIONAL LIBRARY OF FINLAND - Research Library
The material is catalogued directly to the repository in Dublin Core format, but the metadata will be linked to the local library catalogues too.

In order to ease the access to the material, the material will be linked to the National Digital Library of Finland and it browsed through its interface, Finna.

The material will be linked to Europeana too.
In the conjunction with the publishing the digitized material, National Library of Finland has developed an OCR editor for correcting the alphabets that cannot be recognized upon digitization or will be misread by the OCR programme.

The editor has web interface and edited material will be re-uploaded back to Dspace (Fenno-Ugrica) once it has been corrected.
SUB-PROJECT FOR LINKED METADATA

National Library of Finland was awarded by the Finnish Ministry of Education to run a pilot project that would aim to generating a shared and open information system infrastructure for the Finno-Ugrian language materials digitised in different libraries. The sub-project will active until the end of 2013.

This sub-project of the infrastructure project will be implemented by the Research Library Services of the National Library of Finland in cooperation with the national libraries of Russia, Karelia and the Udmurt Republic. Copyright cooperation will be conducted with both the abovementioned libraries and the National Library Resource.
Material linked to the information system will not be provided with new metadata. The intention is for all project participants to provide the bibliographic information of the material to be digitised into the publication system in the form that they have been entered into the local databases.

When the material is transferred to the publication system, its bibliographic metadata will be converted and enriched to correspond to the conventions of digital materials. After the metadata have been harmonised, the material may be used through the user interfaces of third parties.
The linked metadata will be made available in [VuFind](#), that benefits to Dspace in its flexibility.

The public interface will be published in several language, depending on the availability of language packages.

The issues with Cyrillic and Latin alphabets need to be solved.
The project manager began the work relating to the communication between the project partners and general planning in the beginning of March 2013. The work related to agreements, copyrights, metadata management and information systems will be conducted in May-September 2013.

The testing on the material to be linked to the information system should begin in October 2013, and the information system and its features should be ready for presentation to other organisations by the end of the year. The intention is to cover most of the costs of the project period proper in 2014–2016 by external funding as a part of the Language Programme umbrella by Kone Foundation.
CONTACT DETAILS AND FURTHER INFORMATION

Digitization Project of Kindred Languages
www.nationallibrary.fi/services/digitaalisetkokoelmat/finnougric_en_ru.html

Fenno-Ugrica collection
fennougrica.kansalliskirjasto.fi/

Fennio-Ugrica Blog
blogs.helsinki.fi/fennougrica/

National Library of Finland
www.nationallibrary.fi/
Jussi-Pekka Hakkarainen
Project Manager
National Library of Finland
Research Library

The National Library of Finland
P.O. Box 26 (Teollisuuskatu 23)
00014 University of Helsinki

kk-fennouglica@helsinki.fi