

#### Grammatikka I oikiakirjutos.

oisess kluassas annettu sanomustiloin kertaamia. Uu sanomuksen soderzanja ja sanoin yhtehyia sano- ja Upraznenjoja sanomuksiin ladimises eri teemoille kohittämistöön yhtehyes.

rostoi sanomus, lauhilu sannoist, kummat vastajaat kysy- e: **ken? mikä?** (subjekta) i sannoist, mit vastajaat kysy- e: **ken tekkiöö?** mitä tekkiöö? (predikaattia). Subjektian kaatan keskinäiset ottozenjat.

rajennettu sanomus. Eroittammin sanomuksest sanoja, alitaa subjektian i predikaattia.

es kluassas mäntylin predmettasannoin kertaamiin. nimio nimisanan kera, ku.npa osoittaa predmettoja i lamliissilä, javienjoja, tapahtumia).

misannoin muutamiin sanomuksees. (Kala ono mukuisa kalan pyytö tapalituu ympäar vootta. Keskiakesäl vällä ää kallaa j. n. e.) Sanan pohjan ja sanan lopua eroit- sanast.

misena subjektina i toimitussanana. Nimsanan lopun nniin, kons se on subjektina ja kons se ono toimitus-

sikkö ja monikko. Nominetiva, ku forma, kons möö mä predmettaa. Toiset padezat. Padezoiin niinet. Pade-

toimitusnoin teemoittom kera. Tekosanan ponnates, tees- mukkaa. Tekosannoin aikaformat tapalituva aika (presens tapalituva aika (imperfekta), sloinoi tuleva aika (noisen teki- mä). Tekosanan infinitiva format: a) 1-n infinitiva mitä tehdä, kytää, pyytää b) 2-in infinitiva, mitä tekemää? tekemä kytäämä, pyytämä. Tekosanan loput monikoo i yksiköös.

11. Yhissety sanat, niiden oikakirjutos i lauttuun.

#### NELJÄS KLAASSA.

##### Lukemiin.

1. Präntätyin ja käsikirjutosleksten äänee lukemiin (tänne voojen töön jatkuimiin vähä rokkoman materlalanka). Tuntu merkiiä luetust materialast meellin, ajatuksiin eri o.tenkej. Oppiissa lukemaa hyväst äänee hudozestvennoita kirjaa. Tuntu valmistus itseksee lukemaa äänne väelle.

2. Totumus tööää kirjanka. Lukemiin itseksee opeitaja annettua tekstaa täll itseen valittua tekstaa. Tuntu tehi mu stoppanoja luetust tekstast.

3. Kirjaa juko osii ja glovoihe. Tuntu antaa omija zag- violta luetule tekstalle. Tuntu jakkaa osiitsee luettu tekstaa antaa omat zaglaviat eri osille. Plaanaa laittimiin.

Tuntu löyää lätkäämyksest glavnoi paikat ja vtorostepe- sät paikat. Tuntu sitta neet kostonää. Täällä niinet keskiä-

## **ABSTRACT**

This brochure describes the equipment, processes and organisation of the pilot project for digitising Finno-Ugrian materials in 2012-13. The pilot project is part of a project group under the Kone Foundation Language Programme 2012-16, and it is referred to in the Programme section titled “The Digitization of the Finno-Ugrian materials at St. Petersburg, Pilot Project (2012-13)”.

The material to be digitised for the pilot project will be published through the open source publication system maintained by the National Library of Finland. This allows the digital material to be available for open use by the public while ensuring access to the academic community to use the language materials in research. In addition, crowdsourcing tools and resources will be created for the material in conjunction with its publication.

The report suggests that a specially appointed project manager, located in the National Library of Finland and answerable to the pilot project steering group, be in charge of the implementation and coordination of the project. This project manager will make sure the digitisation undertaken at the Russian National Library (St Petersburg) complies with the quality standards. The manager will supervise the process of making the material available in the National Library of Finland publication system, make sure that relevant copyright issues are addressed and coordinate the work to be undertaken at the National Library of Finland. The project manager is also in charge of communications regarding the pilot project in cooperation with the National Library communications unit. The National Library’s Library Network Services unit will be responsible for the maintenance and technological development of the publication system, while the Research Library Services unit will conduct the description and formatting work. All expenses associated with the project, including salaries, will be covered by a grant from the Kone Foundation. The National Library of Finland provides tools, work stations and other immediate infrastructure needed by the project employees.

Based on a report on the pilot project, a grant of EUR 356,040 was applied from the Kone Foundation to implement this pilot project between July 2012 and December 2013. The pilot project will generate practices and technological implementations, which shall benefit the digitisation of other materials in the follow-up project in 2013-2016.

## **PARTICIPANTS AND ORGANISATION**

The National Library of Finland will produce the research material and infrastructure required by the Kone Foundation Language Programme using the grant from the Kone Foundation. The National Library of Finland will also oversee the cooperation with the Russian National Library (St Petersburg) and the Helsinki University Library. The international cooperation will be guided by an agreement drafted between the National Libraries of Finland and Russia. Based on this, a separate agreement between the libraries will be negotiated during spring 2012, specifying the respective duties and responsibilities for each party regarding the pilot project, with relevant sanctions. Copyright issues and fees will be separately agreed between the National Library of Finland and the Russian National Library Resource (NLR, Национальный библиотечный ресурс, НБР), an umbrella organisation providing copyright services for the Russian library sector.

The pilot project will have a steering group comprising representatives from the Finnish participant organisations and academics, making sure to have expert representation from the library sector. As the applicant for the grant, the National Library of Finland will be responsible for reporting on the pilot project and presenting its results to the Kone Foundation.

The main duty of the steering group is to support the project manager, who will report on the progress of the project to the steering group. Locating the project manager within the organisation of the National Library of Finland provides him or her with the best possible opportunities to implement the pilot project and to supervise the progress of different stages of the project. At the launch stage, starting in July 2012, the project manager is responsible for the overall planning of the pilot project, budgeting, agreements with the different parties and other relevant bodies, as well as agreeing on early information system training and cooperation practices with both Finnish and international organisations. At the later stages, the project manager's duties will include conducting quality control of the material to be digitised, planning work processes, supervising the accessibility of the material, managing the community of the crowdsourcing services and monitoring copyright issues. The project manager is also in charge of cooperation with the academic community. Media relations and organising communications in cooperation with the communications of the National Library of Finland are also part of the project manager's duties. In regards to billing and all monetary transactions, the project manager will co-operate with the financial administration of the National Library of Finland.

In addition to the tasks and duties of the project manager, the different stages of the project will require a range of library technical skills to enable the maintenance and development of the publication systems. The pilot project also requires staff skilled in library description and metadata, particularly for the exceptional meta-data practices and formats associated with the project.

## **MATERIAL**

In the autumn of 2011, Kone Foundation representatives selected the Finno-Ugrian materials to be digitised from the collections of the National Library of Russia (St Petersburg). This material will be digitised at the National Library of Russia in St Petersburg and delivered to the National Library of Finland in a digital format. The selected material includes approximately 7,000 pages of publications in the Ingrian and Veps languages, comprising 66 monographs, most of which are textbooks and dictionaries. In addition to these monographs, nearly 20,000 pages of newspapers in the Mari and Mordvinic languages, published mainly in the 1920s and 30s, will be digitised.

The technical requirements for the material to be digitised have been set in cooperation with the National Library of Russia. As the materials cannot be sent to Finland for digitisation, the National Library of Russia will digitise and send them to the National Library of Finland. The monographs will be digitised in grayscale at 300 dpi resolution, while the resolution for the newspapers has been set at 400 dpi. The primary scanning will produce TIFF images, and the elements required for optical character recognition (OCR) will be added. A single PDF file will be compiled from the OCR digitised text and the image files. PDF is also the format in which the material will be made accessible in the publication system. The original TIFF image files will also be handed over to the National Library of Finland for any necessary processing later.

The National Library of Russia will also provide the bibliographic data of the material to be digitised to enable the description of the material at the National Library of Finland and its publication system. The bibliographic data entries cannot be utilised as such in the publication system because Russian libraries utilise their own metadata format, RusMarc, while the pilot project uses DublinCore, a metadata format specifically developed for the description of digital publications. As the digitised materials are catalogued, their bibliographic metadata will be edited to be compatible with DublinCore descriptions of digital resources. Material

digitised in the pilot project will not be catalogued in local databases. Instead, access to the material is facilitated by ensuring its availability in the National Digital Library and possibly the Europeana portal, as far as copyright restrictions allow.

## **SYSTEM AND INTERFACES**

The material digitised in the pilot project will be made available through the publication system maintained by the National Library of Finland, based on the open source DSpace software. For the purposes of the pilot project, a separate DSpace environment will be established for other digital collections of the National Library of Finland, to address the needs of both the academic community and crowdsourcing. Tools for crowdsourcing and improved conditions for generating new functions for the project will be developed through this customised DSpace environment. One tool for crowdsourcing created during the pilot project is the OCR text editor, which allows visitors to fix and edit the OCR text of the digitised material in the publication system. The tool will be developed in cooperation between the National Library of Finland and researchers.

The customised DSpace environment also promotes the accessibility and use of the digitised material internationally, among both the international academic community and speakers of Finno-Ugrian languages living in Russia. This DSpace environment will also be available in Russian, and its user interface can be made available in additional languages. To enhance communications and publicity, an easily recognisable visual style will be designed for this separate environment of the pilot project. In this way the publication system complies with the target of advancing a culture of openness and interaction, as specified in the Kone Foundation Language Programme.

Publishing the material in the DSpace publication system also makes searching for content easier. The search function of the system supports both searches in the metadata of the publications and full-text searches of their content. The publications entered into the system will be appointed permanent URN addresses, which make the material searchable even if the location of the publications were to change. The material in the pilot project will be provided for open access also through the coming public interface of the National Digital Library and equivalent international search services. The materials will also be searchable through Google and other common search engines.

## **COPYRIGHTS**

Both Finnish and Russian copyright legislation must be considered in the pilot project, since the digitising will be conducted in Russia and the material made available in Finland. As most of the material to be digitised is within copyright, the conditions for digitising and publishing the material must be further examined together with the National Library of Russia and the National Library Resource.

## **COMMUNICATIONS**

The project manager is in charge of communications regarding the pilot project in cooperation with the National Library communications unit. The pilot project will have its own page on the National Library of Finland website. The page will include information and news about the project in Finnish, Swedish, English, Russian and other languages if possible. The project can also utilise other communication channels of the National Library of Finland, such as client magazines and social media, enabling communications with specific user groups. The project manager is in charge of the communication content and of making sure the information is up to date.

The pilot project will be presented at a range of Finnish (such as Kirjastoverkopäivät) and international events (such as IFLA 2012) jointly with the Kone Foundation and other organisations. The pilot project will also be presented in further detail at the kick-off event organised later in the autumn, intended to introduce the project to both the academic community and the general public.

## **FOLLOW-UP PROJECT**

The pilot project (July 2012-December 2013) will develop practices and technological implementations which will ease the digitisation of other materials under the Language Programme, to be conducted in the follow-up project in 2014-2016. The intention is to also enable new forms of international cooperation and to widen the cooperation, if possible, to cover other memory institutions managing and digitising Finno-Ugrian language materials in Russia and Finland. Developing cooperation during the pilot project eases the implementation of the follow-up project and will facilitate issues with the metadata formats in particular (METS/Alto packages and the RusMarc conversion).

**\* \* \***

**CONTACT DETAILS AND ADDITIONAL INFORMATION**

Jussi-Pekka Hakkarainen  
Project Manager  
Research Library

National Library of Finland  
P.O. Box 15 (Unioninkatu 36)  
00014 UNIVERSITY OF HELSINKI  
+358 9 191 40793  
+ 358 50 363 9223  
[jussi-pekka.hakkarainen@helsinki.fi](mailto:jussi-pekka.hakkarainen@helsinki.fi)

