

National Library of Finland Digitisation Programme 2021–2024

National Library of Finland 2021

ISBN 978-951-51-7103-0



National Library of Finland Digitization Programme 2021–2024

Authors of the Digitisation Programme: Johanna Lilja, Jussi-Pekka Hakkarainen, Pirjo Karppinen, Pekka Lampinen

Material selected by: Jouni Ahmajärvi, Jussi-Pekka Hakkarainen, Mika Hakkarainen, Anu Haldén, Pirjo Karppinen, Jukka Kervinen, Pasi Koste, Pekka Lampinen, Eeva-Liisa Leppänen, Tarja Lehtinen, Tuula Pääkkönen, Irma Reijonen, Leena Saarinen, Liisa Savolainen, Jaakko Tahkokallio, Marleena Vihakara

1. The Digitisation Programme is part of the Strategic Plan of the National Library of Finland 2021–2030

Digitisation and transparency were significant aspects of the National Library's Strategy for 2016–2020. The National Library of Finland's new Strategy continues in the same direction, but with an even stronger understanding of the significance of the National Library's openly accessible digital services as part of Finnish society as well as of international networks of research and technology. The four sets of targets for the Strategy include: 1. Cultural heritage for the common good; 2. The National Library at the heart of the academic community; 3. The National Library as an education and learning resource; and 4. Creating a powerful hub of expertise through network cooperation. The cornerstones of our work are openness, renewal and *Bildung*.

In the Strategic Plan of the National Library of Finland 2021–2030, the Library sets the objectives of digitising extensive unified sets of resources and supporting open science and the development of methods and tools for digital humanities. Data generated by the Library are managed according to FAIR principles, and their users are instructed to use the data responsibly. With its open digital services, the Library wants to create conditions that can help transform teaching, learning and innovation while supporting the use and significance of cultural heritage resources. Active networking develops competence.

The National Library's Digitisation Programme 2021–2024 lists the resources that the Library wants to digitise during the first performance period of the ten-year strategic period. In addition to the Strategy, the Digitisation Programme is connected to the Library's Collections Policy, Fundraising Plan, Digital Humanities Policy, the Open National Library Policy and the metadata vision.

2. Selection criteria for resources to be digitised and digitisation metrics

The rich and varied collections of the National Library have evoked many requests for digitisation, both from among Library staff and its partners and customers. As resources are limited, these requests must be prioritised. The selection of the material to be digitised has been planned in a workshop of National Library experts and then discussed with supervisors from the Research Library service function, the National Library's leadership in charge of planning as well as the management group along with the National Library's advisory boards.

The criteria used to evaluate the success of digitisation were considered when selecting the material for digitisation. In addition to quantitative metrics, such as download and production figures, the Library has wanted to highlight operative and ethical metrics. The goal is for the digitised material to make library work more efficient and to serve people from a wide variety of language and cultural backgrounds. The Strategy emphasises the needs of research. By serving research, digitisation indirectly benefits all of society..

<p>Current metrics</p> <ol style="list-style-type: none"> 1. Digitisation volume (digitised pages) 2. Use and download figures of the material 3. Customer satisfaction, user and visitor surveys, other feedback 4. External funding for digitisation projects (in euros) 5. To include Partnerships and cooperation in the future 	<p>Criteria based on Library needs</p> <ol style="list-style-type: none"> 1. Use of unique / original items to be discontinued 2. Fewer resources needed for book search and customer service 3. Mobilisation of resources (e.g., the Language Bank of Finland, Europeana) 4. International benchmarking with other cultural heritage institutions
<p>Added value for other social functions</p> <ol style="list-style-type: none"> 1. Research 2. General public and citizen science 3. Art and culture 4. Local institutions 5. Businesses 6. Education 	<p>Ethical criteria</p> <ol style="list-style-type: none"> 1. Making the collections accessible supports the goals of digitisation, crisis preparedness and climate awareness. 2. Sustainable use 3. Equality: presenting material from various language, ethnic and cultural groups

Image 1. Current digitisation metrics and other factors used to measure the success of digitisation

When selecting the material to be digitised, the National Library must also consider whether it can be rendered into a high-quality, user-friendly digital resource. Digitising complete sets of material is particularly important for research. To guarantee the technical quality of the digitised material, sufficient resources must be budgeted for the production process. Reliable metadata ensures the findability and long-term storage of the material. Clear terms of use and an easy-to-use and reliable user interface improve usability.

<p>Content quality</p> <ol style="list-style-type: none"> 1. An extensive and comprehensive set of material 2. No individual pieces (with the exception of exhibition use) 	<p>Technical quality</p> <ol style="list-style-type: none"> 1. High-quality scanning 2. Appropriate post-processing and structuring for different material types 3. High-quality OCR text recognition with support for different languages
<p>Findability</p> <ol style="list-style-type: none"> 1. High-quality metadata, cf. the Metadata vision 2. Uniform metadata for all collections 3. Compatibility, accessibility, visibility on Finna search service 4. A long-term preservation certified final product 	<p>Usability</p> <ol style="list-style-type: none"> 1. Clearly indicated users' rights according to copyright legislation 2. Data packages in a machine-readable format 3. Easy-to-use user interface and user tools (digi.kansalliskirjasto.fi) 4. Secure operation

Image 2. High-quality digitisation consists of many elements.

3. The digitising process and associated resources

Digitisation is conducted with the Library's own budget funding at the Library's Digitisation Services, which provides retroactive digitisation for existing material, as well as Newspaper Services, a service for digitising all newspapers only available in print format as part of the legal deposit process. In the budget, a total of eight person-years have been allocated for retroactive digitisation, of which three person-years are intended for the digitisation of sound recordings. This allocation does not include IT support for digitisation or the development of the digi.kansalliskirjasto.fi service.

The National Library's Centre for Preservation and Digitisation in Mikkeli has a wide and varied selection of scanners, suitable for use with different material types. The scanning of books which are in good condition is almost entirely automated thanks to the automatic scanner intended for this purpose. For fragile paper and loose-leaf material, the Centre has manual scanners with book lifts that require an attendant to change pages by hand. The material may be scanned either with or without a glass plate, including when processing large maps. Newspaper material on microfilm is scanned using a roll film scanner.

All scanned material requires some post-processing before it can be published and entered into long-term preservation. Equipment specifically designed for mass digitisation is used for this purpose. In post-processing, pages are cropped, text areas straightened, images and text sections automatically recognised, and optical character recognition (OCR) conducted. The result is a file which contains all the pertinent information necessary for the publication or long-term preservation of the book or newspaper, including the metadata.

The digitising process requires cooperation between several National Library units and experts. The Helsinki-based units Collection Services and Research Services are in charge of selecting, describing and packing material for digitising. Collection experts also participate in the acquisition of

external funding for digitising the resources they select. Experts at the Mikkeli Centre can also provide conservation services for material that needs it. The Legal Deposit Office is in charge of developing the long-term preservation process. When digitising new material types, the Centre must cooperate with Metadata Services under Library Network Services, the National Metadata Repository Melinda and the developers of Finna.

The implementation of the Digitisation Programme is coordinated by the working group for contracts, fundraising and digitisation, which discusses all proposals requiring external funding and plans the copyright negotiations required for providing open access to certain resources. This group checks all contracts relating to digitisation and user access. External funding projects are budgeted according to the University of Helsinki's pricing principles, and the director of the National Library approves all projects above €20,000.

External funding is very important for the production of digitisation. The main sources of external funding include:

- Partnership digitisation with publishers and media companies
- Funds and foundations
- Research funders
- The National Library's Cultural Heritage Fund

In a partnered digitisation, the publisher or funder of a particular publication agree upon the digitisation of the resource with the National Library. Any use of a resource which is broader than that specified in legislation, typically the publisher's own restricted use or use by the audience, is negotiated with the Kopiosto copyright organisation. The digitised material will also always be accessible through the digi.kansalliskirjasto.fi service. If the resource cannot be released to the public, it will be accessible locally on the legal deposit workstations in legal deposit libraries.

The National Library's Cultural Heritage Fund coordinates donations made to the National Library and annually grants funding to targets specified by the rules of the Fund according to its action plan. Funding is sought from funds and research funders for projects relating to the digitisation of the Library's resources and related research.

Thus far, such external funding has covered 2.8 person-years at Digitisation Services along with a part of the costs associated with equipment, machinery and facilities. The intention is to increase the share of external funding in the future to cover an additional 0.5 person-years of work relating to preparing material for digitisation at the National Library's Collection Services. Digitising non-newspaper material often requires selecting specific items for digitisation and describing them, as the bibliographic description of the oldest sections of the National Collection has major omissions.

The Mikkeli Centre is prepared to increase production if sufficient external funding can be acquired to cover the expanded production capacity and the recruitment of additional staff for at least one year.

4. Material to be digitised during the performance period 2021–2024

A variety of funding options have been considered in the selection of material to be digitised during the performance period 2021–2024. The material has tentatively been categorised into two groups according to whether the digitisation is to be financed by core or external funding. There is always an element of uncertainty to external funding. It is possible that projects hoping for external funding may have to be postponed to the following performance period, if the funding application is unsuccessful.

Image 3 presents the schedule for digitising various sets of resources. Other than newspapers, all material requires some preparation and bibliographic description work. This has been considered in the project schedule. There are many factors to consider in the scheduling of projects. Newspapers with a high use-rate, and for which the digitisation process has effectively been automatised are digitised throughout the performance period with core funding. Resources with fixed funding or a clear plan for acquiring funding have been scheduled towards the beginning of the performance period.

Projects where there is still considerable planning to be done, or where the release of the material through digi.kansalliskirjasto.fi requires further development of the service, are scheduled later in the period. Such material includes sheet music, as the service must be adapted so that parts of a work, not just complete titles, can be searched for and presented. In the case of sheet music, parts of the work could be individual songs from a songbook, for example. Despite eager expectations, the digitisation of the National Library's diverse archive collections and ephemera cannot begin before an information system is acquired for them to enable the administration of the associated metadata and the retrieval of the metadata through the digi.kansalliskirjasto.fi service. Additionally, digi.kansalliskirjasto.fi requires further development to enable search and viewing of items from the archive collections. An information system for archive material will be sought during autumn 2020

Digitisation programme 2021–2024		2021				2022				2023				2024			
Funding	Material	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
core	Newspapers 5 volumes / year																
core	UH doctoral theses																
core	Imperial Russian Fennica																
external	Newspapers and periodicals with publishers																
external	Swedish time Fennica, library collection																
external	Clandestines																
under preparation	Sheet music 1801–1916																
under preparation	Ephemera: gardening																
under preparation	Ostrobothnian mystics																
under preparation	Ptolemaic atlases																
under preparation	Lecture notes																
external – research	Swedish time Fennica, reproductions																
external – research	Broadsheet ballads																
external – research	Inari and Skolt Sámi material																
core	Audio CDs and vinyl pilot																

Image 3. The Digitisation Programme timeline. Projects under the Library's core funding are shown in yellow, projects for which external funding has been secured are in green, projects for which donor and foundation funding is primarily sought are in orange, and projects intended to be covered by research funding are in blue.

Newspapers

Newspapers in the National Collection are a unique resource of topical news communication, and as the collection expands every day, it offers a rich and varied source for anyone interested in people, history, language, culture and society. National, regional, local and town newspapers, dailies, general interest newspapers and various specialist and free newspapers provide many different perspectives on the events of that time. The scope of the collection is approximately 50 million pages, and it grows annually with approximately 300 newspaper titles and 850,000 pages.

All newspapers printed in Finland between 1771 and 1939 have been digitised. According to agreements made with the copyright organisation, the National Library has been able to release all newspapers published before the year 1940 to the public, and all digitised newspapers until 2018 for researchers through its digi.kansalliskirjasto.fi service. All new newspapers have been digitised since the beginning of 2017. The most recent publications can be accessed through the legal deposit workstations.

However, the National Library has only digitised less than 20% of its newspaper collection. Newspapers published between 1940 and 2016 are being digitised with core funding and grants, and in cooperation with publishers and funds.

Doctoral theses from the Imperial Alexander University

More than 1,300 doctoral theses were published at the Imperial Alexander University (1828–1917) in Finland. These digitised theses, along with other academic publications from Finland's time as part of the Russian Empire, constitute a significant resource for the history of science. Together with doctoral theses from the Academy of Turku, which have already been digitised, they make the history of university research available to all. The lecture notes from the Grand Duchy of Finland, which are included in the Digitisation Programme, expand this body of material.

The University of Helsinki has funded this digitisation project, which launched in autumn 2020 and will end in the beginning of 2021.

National Collection from the Grand Duchy of Finland

The National Collection from Finland's time as part of the Russian Empire, i.e., material from 1809/1828–1917, has not been systematically digitised. The Digitisation Programme features an option to launch the digitisation of this material with core funding from 2022, if other projects have been unsuccessful in acquiring external funding and the Library has digitising capacity available. This set of material contains approximately 50,000 books and is primarily free of copyright. Approximately 1,800 titles from this collection have been previously digitised.

Periodicals

The National Collection includes tens of thousands of periodical titles, from the beginning of the 19th century to the present day. Currently, there are approximately 4,000 periodicals in print in Fin-

land. The collection is constantly expanding, and it includes all Finnish periodicals from general-interest magazines to professional, organisational, client and specialist magazines. There are also magazines specifically aimed at children and young people. Magazines provide supplemental, in-depth information on the daily news by highlighting the background stories of current events, and discuss events and themes in their topic areas in a more comprehensive way. People typically keep their old magazines and reread them. Digitised periodicals published before the year 1930 have been released to the public.

Periodicals are being digitised with core funding, grants and in cooperation with publishers and funds.

National Collection from the period of Swedish rule

The collection of Finnish literature from the period of Swedish rule is the oldest section of the National Collection. This collection, Rv Fennica, contains approximately 17,400 publications from 1488–1809/1827 which were printed in Finnish or in Finland, published by Finns or people working in Finland, or published elsewhere but which discuss Finland. The collection includes spiritual literature, various non-fiction books, doctoral theses, prose, official publications and other ephemera.

The digitisation of the collection has previously been conducted as part of various projects, but during 2019 and 2020 the digitisation was carried out systematically thanks to grant funding. A total of 6,850 titles from this collection have been released in digital format. The Rv Fennica features 4,650 publications as reproductions, i.e., as microfilm, photocopies, copy negatives or positive copies. The digitisation of these items should be done from the original publication, in cooperation with other Nordic libraries.

External funding to complete the digitisation of the Rv Fennica collection is being sought in 2020. If no external funding can be acquired, the collection will be digitised with basic funding. Work on the reproductions may be conducted in 2021 and 2022.

Clandestine manuscripts

As part of a donation from the St Petersburg Academy of Sciences in the 1830s, the National Library gained a considerable part of the personal library of Johan Albrecht von Korff, which Catherine the Great had bought from Korff in the 1770s. Korff's library included "clandestine manuscripts", or banned philosophical literature from the 17th and 18th centuries which could not be printed, so they were released as manuscripts and copied for dissemination. These manuscripts are often texts that oppose the religion, church and autocracy of their time.

The clandestine manuscripts are the object of considerable international interest and constitute a unique source material for European cultural history and the history of philosophy. The digitisation of this collection will promote the visibility and use of the National Library's reference collection.

Funding for the digitisation of this material was granted by the National Library's Cultural Heritage Fund in summer 2020, and the digitisation could be carried out in 2021.

Sheet music 1801–1916

The National Library's sheet music collection comprises a total of 45,000 titles. The collection holds many different types of cultural heritage value. Sheet music publications are not just part of the history of musical composition and arrangement, but also that of lyrics, translation and illustration. They shed light on political movements, international attitudes and music publishing. Studying sheet music publications deepens our understanding of Finnish culture and history. Thus far, the sheet music collection has seen very little use.

The digitisation of the extensive collection must be divided into several phases and sub-projects. The development of the digitisation process can continue alongside the projects. The Library will also carefully survey the potential development of technical solutions for releasing sheet music online.

The digitisation of sheet music has been explored and piloted during 2017–2019. This project requires both external funding and further development of the digi.kansalliskirjasto.fi service. The digitisation of the sheet music collection has been scheduled for 2022–2023.

Ephemera: Gardening 1810–1944

Ephemera are among the most popular digitised resources, as they are interesting both visually and in terms of content, and hold interest for both hobbyists and researchers. The material in the National Library's collections is unique in its scope and is not available anywhere else.

The ephemera collection is yet to be catalogued. The intention is to describe them in the same information system as the archive collection. Consequently, this digitisation project has been scheduled for 2024. The digitisation requires external funding.

Ostrobothnian Mystics

The Ostrobothnian Mystics were a religious sect in Ostrobothnia active from the late 18th until the mid-19th centuries. The sect was founded by the German Jakob Böhme, whose thinking spread to Ostrobothnia thanks to the Swedish priest Anders Collin. The sect started near Vaasa and spread from there to central Ostrobothnia.

The texts from the Ostrobothnian Mystics were completely different in terms of content than other printed Finnish literature from that time. The manuscript continued to be a channel for alternative publication for decades. As censorship prevented any public criticism of the authorities and the Lutheran church, critical reformist literature both in the secular and spiritual realms was disseminated as manuscripts and illegal prints.

Between 1780 and 1830, artisans, village scribes and peasants created this nationally and internationally significant body of work: mystical texts written by laypeople as well as other texts, the first Finnish translations of European literature. The Ostrobothnian Mystics collection is an impressive sign of the rise of new independent groups of writers and readers. In the late 18th century, European thought began to cross national borders and spread among the lower social classes.

Once digitised, this collection will supplement the collections of broadsheet literature and periodicals to illustrate how information and ideas were disseminated towards the end of Finland's time under Swedish rule, while being a testament to the rise of independent groups of writers and readers. The National Library's manuscript collections feature 29 items relating to the Ostrobothnian Mystics, and the intention is to digitise them with external funding in 2022.

Ptolemaic atlases

The A. E. Nordenskiöld collection consists of historical maps, geographical literature, travel narratives and rare books collected by researcher and explorer Adolf Erik Nordenskiöld in the late 19th century.

The collection features a great deal of rare early cartographic and geographic literature, approximately 3,870 titles, along with a few manuscripts and some 24,000 maps. The collection also includes a large amount of literature regarding Arctic areas as well as the arrival of Europeans in America. The collection as a whole documents the increase in geographical knowledge in the West. Nordenskiöld's library is among the most significant collections of the history of cartography in the world.

The most valuable section of Nordenskiöld's collection comprises the approximately 40 Ptolemaic world maps published in the 15th and 16th centuries. The collection was added to UNESCO's Memory of the World Register in 1997, largely because it contains nearly all editions of Ptolemaic maps from this period. Digitising this part of the collection serves researchers of old maps around the world and would be a digital calling card for the Nordenskiöld collection. Most of the atlas editions have been digitised elsewhere, but the nearly complete set of Ptolemaic maps would make this a very useful resource for scholars of historical cartography.

External funding is being sought for the digitisation of the Ptolemaic atlases, and the National Library is ready to complete the digitisation and release the material in 2022–2023.

Lectures and lecture notes

The personal archives in the National Library's Manuscript Collection feature a considerable amount of unique lectures from university teachers as well as students' lecture notes from the time of the Royal Academy of Turku to the present day. These lectures comprise a significant resource for the history of ideas. Digitising would make well-known figures in Finnish history (including Henrik Gabriel Porthan, Anders Johan Sjögren, Johan Wilhelm Snellman, Zacharias Topelius, Matthias Calonius) more accessible, while highlighting other, less well-known students and teachers.

This is an important resource for the history of ideas, and it would complement the digitisation of the doctoral theses from the Grand Duchy of Finland. The digitisation of this material would also only begin towards the end of the performance period, as a separate information system must be found for the metadata of the archive material, and the digi.kansalliskirjasto.fi service must be further developed.

Broadsheet ballads

A broadsheet ballad is a song leaflet of one, half or a quarter of a printed broadsheet. Broadsheet ballads were written as song lyrics, and they typically indicated the melody that was intended for them. In the 17th century, broadsheet ballads were hymns. Most of the earliest broadsheet ballads were written by members of the upper classes. Towards the end of the 19th century, the authors and print numbers increased, with more and more broadsheet ballads being secular, focusing on love stories, grisly murders and unusual events.

The broadsheet ballad is an exceptional genre at the intersection of printed literature, vocal music and oral tradition. Its nature and characteristics reveal new information as researchers unlock new concepts, vocabulary and categories. Broadsheet ballads describe the Finnish cultural environment from the 17th until the beginning of the 20th century in a new way. With the help of modern technology, the results of this project can be widely applied, both in research focusing on broadsheet ballads or on media and communications, for example.

Funding for the digitisation of broadsheet ballads was applied for in conjunction with a related project during autumn 2020. If funded, approximately 5,000 of the National Library's broadsheet ballads, mainly from the 19th century, can be digitised between 2021 and 2023. Previous digitisation has focused on the oldest section of the broadsheet ballad collection, from the 17th to the early 19th century.

Resources in Inari Sámi and Skolt Sámi

One of the missions of the National Library is to ensure that all Finns have equal access to information, including linguistic and other minorities. However, Finland's minority languages are still not equal, and the Inari and Skolt Sámi communities, for example, can be considered unequal in comparison with the majority languages. The problem is not the lack of IT solutions or skills, but a typical conundrum of small languages: information on the language has not been collected in one place and made openly available.

This project comprises all printed literature in the Inari and Skolt Sámi languages. This broad scope is intended to enable the creation of extensive corpora, which would make it possible to employ linguistic applications used in the processing of major languages. The material to be digitised would be an excellent resource for researchers, both for ongoing and future projects. In its current formats, the material is not very accessible. Providing it to researchers will promote a more detailed understanding of how the written languages have developed, while allowing the material to be connected to the data created in the course of the documentation of the languages.

The digitisation will also improve the position of minority languages spoken in Finland. By digitising all literature in Inari and Skolt Sámi, the National Library supports research and teaching into the Sámi languages as well as their revitalisation. The digitisation requires copyright agreements and external funding. If funded, this digitisation project could take place in 2023–2024.

Sound recordings

The National Library is digitising extensive sets of material in the National Sound Archive and producing the related metadata, which will be added to the Viola database. In addition, the Library is prepared to digitise donated material to ensure its preservation. Some donations may be in fragile formats, such as reel-to-reel tapes, which require immediate digitisation. In terms of volume, cassette tapes, CDs and vinyl records are the largest material types in the Library's sound recording archives. Most of the Library's unique cassette tapes have already been digitised. Approximately 1,000 cassettes remain to be described and digitised. The Library's collection of more than 55,000 CDs has grown by approximately 2,000 records annually until recently. The Library began systematically digitising its CDs in 2018. At the current production pace, two-thirds of the CDs will be digitised by the end of the current performance period. Towards the end of the period, in 2023–24, the Library can begin preparing for the next large project, digitising LPs and singles.

5. Fundraising

The goal of the National Library's fundraising is to acquire funding to store, digitise and conserve the cultural heritage materials in the Library's collections. Fundraising may also be targeted for other purposes, such as the acquisition of significant source material. Fundraising targets for digitising projects are determined in the Digitisation Programme.

External funding for the digitisation and preservation of cultural heritage materials is sought through cooperation partners, such as publishers, funds, foundations and private individuals.

Donations made to the National Library's Cultural Heritage Fund, which is part of the University of Helsinki Funds, are allocated directly to funding this work.