



Leverage from  
the EU  
2014–2020

## Download Tool User Guide

To find the data you wish to download using the Tool, go to [digi.kansalliskirjasto.fi](https://digi.kansalliskirjasto.fi) and search for the documents. You can use the Tool to download data from newspapers and magazines, ephemera, books, manuscripts and sheet music to the extent that their content is considered to be largely free of copyright. Newspapers printed before the end of 1918 and ephemera printed before the end of 1910 are available for download. You can download other documents as indicated by their copyright metadata, the time limit depending on the data and work. Watch the video guide: <https://www.youtube.com/watch?v=o-ykh9zTOdI>

To open to the Tool, select the Download Tool shortcut icon. You will find the icon in the Download Tool directory where this Guide is also located.

In the Digin hakutulokset sisältävän sivun osoite (Address of the page containing the Digi search results), enter the address of the page containing your search results.

You can then choose whether to download text as ALTO or text files and whether to download the data as jpg page images or pdf files. To download all pages of the volumes contained in the search results, select Lataa kaikki niteen sivut (Download all pages of the volume). If no selection is made, the Tool will only load pages that are included in your search results.

You can also select the storage location on your computer for downloading the data. By default, the data is downloaded to the same directory where the Download Tool is located. To change the location, select the Valitse sijainti (Select location) button. To change the directory structure of your storage location, select an option from the Hakemistopuun rakenne (Directory tree structure) drop-down menu. By default, BindingTitle\_BindingID is selected and the data is stored in directories whose names consist of the volume name and the volume unique identifier.

To begin the download, select the Lataa aineisto (Download data) button. You will see the progress and estimated duration of your download. You can pause and resume the download by selecting Keskeytä lataus (Pause download) and Jatka latausta (Resume download). To cancel your download completely, select Peruuta lataus (Cancel download). You can close the Tool and then resume a download you have paused earlier. When the download begins, you will also see the number of pages included in the search and, if you used search terms in the search criteria, the occurrence of search terms on the pages found in the search. To download the number of pages in your search and the download metadata as csv files, select Tallenna hakutulosten sivujen määrät (Save the number pages included in the search) and Tallenna ladattujen aineistojen tiedot (Save metadata for the downloaded data). You can merge csv files for future new downloads by selecting Yhdistä tulevien latausten csv-tiedostot (Merge csv files for future downloads). In this case, the number of pages of future download searches and the metadata for the downloads will be added to the same csv files. The download metadata, which also includes the copyright information and references in the data, is stored in a csv file beginning with ladatut aineistot (downloaded data), automatically created in the storage location when the download begins.

The Download Tool has been developed as part of the Digital Open Memory project, funded by the European Union Regional Development Fund and its Leverage from the EU 2014-2020 programme. The project develops the National Library's services for researchers and the use of its digital collections. The project has collected information in a user-oriented manner from researchers who use the National

Library's digital collections, and interviewed and observed parties who access the international researcher services developed around digital collections.

## Detailed descriptions of Tool buttons

### Tallenna hakutulosten sivujen määrät (Save the number pages included in the search)

Saves a csv file that contains the number of pages that a Digi search returned. If the search included search terms, the csv file also stores the number of pages where the terms appeared. For example, the following search returns a csv file, the results of which are shown below:

<https://digi.kansalliskirjasto.fi/search?query=mikkeli&startDate=1870-01-01&endDate=1871-12-31&title=1457-4721&formats=NEWSPAPER>

Year	Pages included in the search	Pages containing the search terms
1870	437	89
1871	633	143

### Tallenna ladattujen aineistojen tiedot (Save the metadata for the downloaded data)

Saves a csv file that contains general information about the downloaded data, copyright information and the references in the data. This csv file is also automatically created in the storage location when the download begins.

## Descriptions of downloadable file formats

Below you will find descriptions of the file formats which the Tool uses for downloads. Each format is suitable for a variety of uses.

### ALTO

An XML schema that describes the appearance of a page and contains the OCR-recognised raw text of the page. The pages may be from printed documents, such as newspapers or books. Read more: <https://www.loc.gov/standards/alto/description.html>

### Text file

Contains unformatted text. Text files downloaded via the Download Tool only contain the OCR-recognised raw text of the page.

### JPG

An image format that uses lossy compression and is supported by many programmes. JPG is used for page images, in other words, images of the pages that make up newspapers, books and other printed documents.

### PDF

A file format in which the appearance of the data always remains the same, independent of the programmes or operating systems used to view the file. PDF files downloaded via the Tool contain all page images of a single volume.



Leverage from  
the EU  
2014–2020

Address of the Digi page with the search results

<https://digi.kansalliskirjasto.fi/search?startDa>

Text format

☒ ALTO xml ☐ text

Pictures

☒ jpg pictures ☐ pdf

Download all pages of the volume

☐

Material storage location

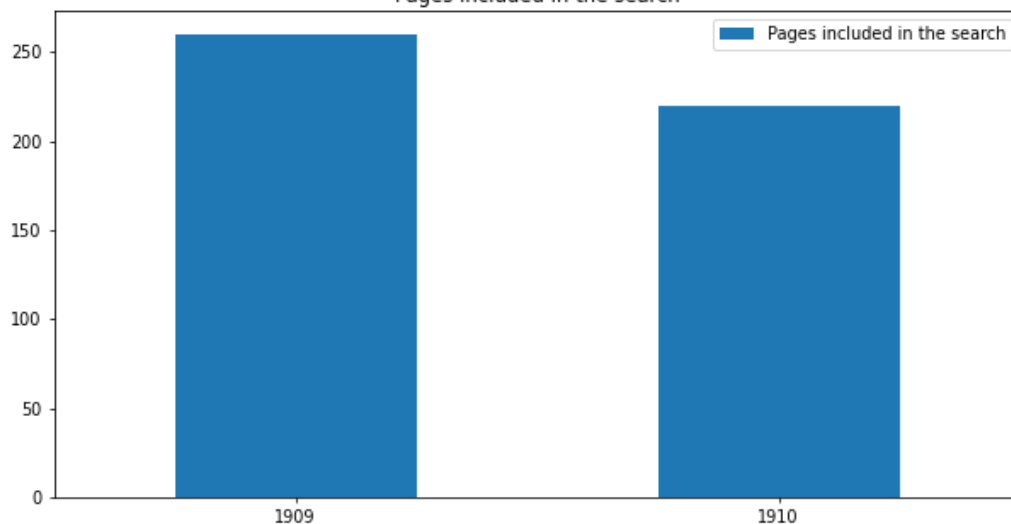
C:\Users\esliukko\Documents\NetBeansProjects\digi\_aineistolataaja\_eng

Choose location

Directory tree structure

year/Bindinglssn

Pages included in the search



Save the number of search result pages

Save the information on downloaded materials

☐ Merge the CSV files of future downloads

Download status

Downloading search result: 5/50

Downloaded: 15.54 MB

Free disk space: 46.11 GB

Estimated download duration: 1.4 min

Pause download

Cancel download