

# Handwritten text recognition & search platform for historical court records

Sampo Viiri  
[sampo.viiri@arkisto.fi](mailto:sampo.viiri@arkisto.fi)

Ville-Pekka Kääriäinen  
[ville-pekka.kaariainen@arkisto.fi](mailto:ville-pekka.kaariainen@arkisto.fi)



THE NATIONAL ARCHIVES  
OF FINLAND

**READ**  
RECOGNITION & ENRICHMENT  
OF ARCHIVAL DOCUMENTS



# READ project

- **R**ecognition and **E**nrichment of **A**rchival **D**ocuments
- 2016-2019
- 14 partners – 3 archives, 11 universities and research centres
- Goal: revolutionise automated handwritten text recognition and access to digitised historical material
- Work continues in READ COOP

# New research platform – Transkribus



- Automated text recognition for both printed and handwritten material
- Search functions for digital material
- Free download: [transkribus.eu](https://transkribus.eu)
- Over 20 000 registered users



# THE NATIONAL ARCHIVES OF FINLAND

Transkribus v1.6.0 (28\_02\_2019\_14:59). Loaded doc: Ahvenanmaan tuomiokunnan renovoidut tuomiokirjat (HMA) Ilmoitusasioiden pöytäkirjat 1881-1882 (KO b22), ID: 163017, Page 1, file: Ilmoitus\_224.jpg [Image Meta Info: (Resolution:300.0, w\*h: 5650 \* 4506)]

Search current document... 1 / 10 Ground Truth

Server Overview Layout Metadata Tools

Logout sampo.viiri@arkisto.fi

Document Manager User Manager

Versions Jobs

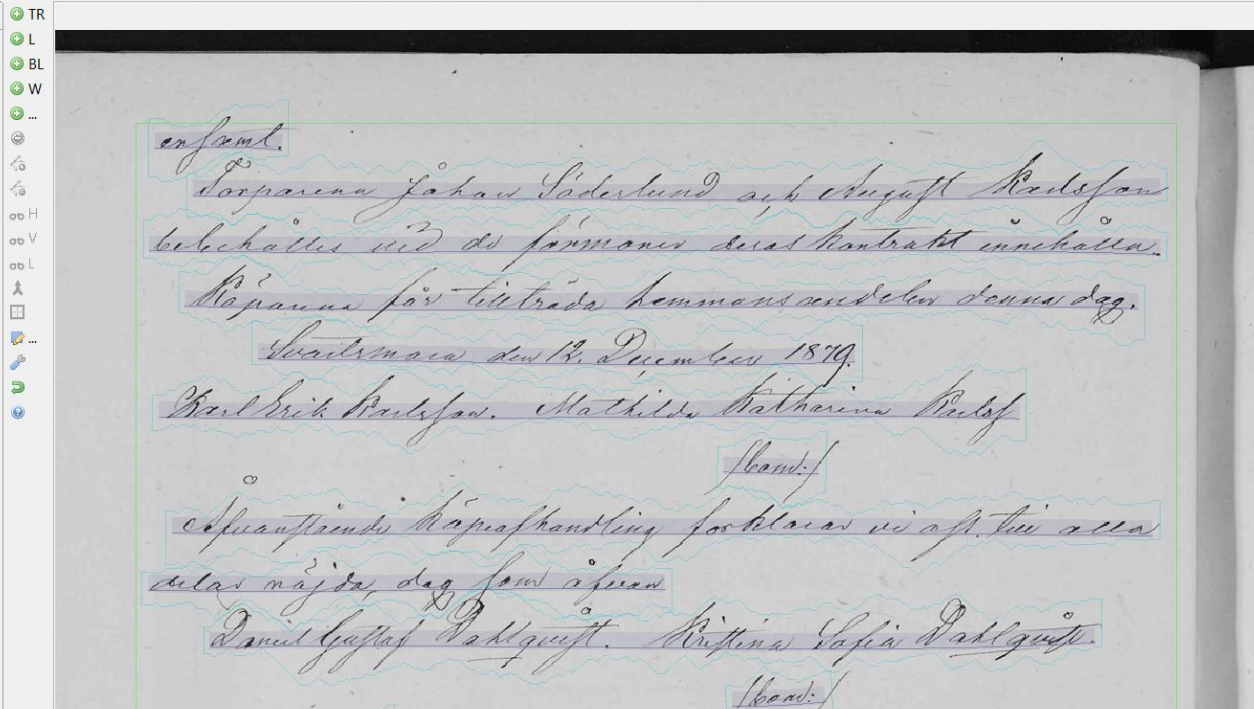
Recent Documents... User activity

Collections: NAF GT Court Records (12568, Owner)

1-88 / 88 1 1 Doc-ID

ID	Title	Pag...	Uf
163054	Tornion tuomiokunnan renovoidut tuomiokirjat Il...	20	sa
163053	Sääksmäen tuomiokunnan renovoidut tuomiokirj...	20	sa
163052	Sääksmäen yliisen tuomiokunnan renovoidut tuo...	20	sa
163051	Sääksmäen alisen tuomiokunnan renovoidut tuo...	20	sa
163049	Suur-Savon tuomiokunnan renovoidut tuomiokirj...	20	sa
163048	Salon tuomiokunnan renovoidut tuomiokirjat Ilm...	20	sa
163045	Rautalammin tuomiokunnan renovoidut tuomio...	20	sa
163043	Raaseporin läntisen tuomiokunnan renovoidut tu...	20	sa
163041	Pohjanmaan itäisen tuomiokunnan renovoidut tu...	20	sa
163039	Piikkiön ja Halikon yliisen tuomiokunnan renovo...	20	sa
163038	Piikkiön ja Halikon alisen tuomiokunnan renovo...	20	sa
163037	Pien-Savon alisen tuomiokunnan renovoidut tuo...	20	sa
163036	Oulun tuomiokunnan renovoidut tuomiokirjat Ilm...	20	sa
163033	Maskun tuomiokunnan renovoidut tuomiokirjat Il...	20	sa
163032	Kymin tuomiokunnan renovoidut tuomiokirjat Il...	20	sa
163030	Kymin tuomiokunnan renovoidut tuomiokirjat (U...	20	sa
163029	Korsholman eteläisen tuomiokunnan renovoidut ...	20	sa
163027	Keski-Pohjanmaan yliisen tuomiokunnan renovo...	20	sa
163026	Kemin tuomiokunnan renovoidut tuomiokirjat Il...	20	sa
163025	Karjalan yliisen tuomiokunnan renovoidut tuomio...	20	sa
163024	Iisalmen tuomiokunnan renovoidut tuomiokirjat L...	10	sa
163022	Iisalmen tuomiokunnan renovoidut tuomiokirjat L...	20	sa
163021	Euran tuomiokunnan renovoidut tuomiokirjat Ilm...	10	sa
163020	Alavuden tuomiokunnan renovoidut tuomiokirjat...	10	sa
163019	Ala-Satakunnan tuomiokunnan renovoidut tuomi...	20	sa
163018	Ala-Satakunnan tuomiokunnan renovoidut tuomi...	10	sa
<b>163017</b>	<b>Ahvenanmaan tuomiokunnan renovoidut tuo...</b>	<b>10</b>	<b>sa</b>
162757	TRAINING_TESTSET_NAF court records test 7	73	ka
162731	TRAINING_TESTSET_NAF court records test7	50	ka
153402	Hollola1810	9	jul

100 Filter



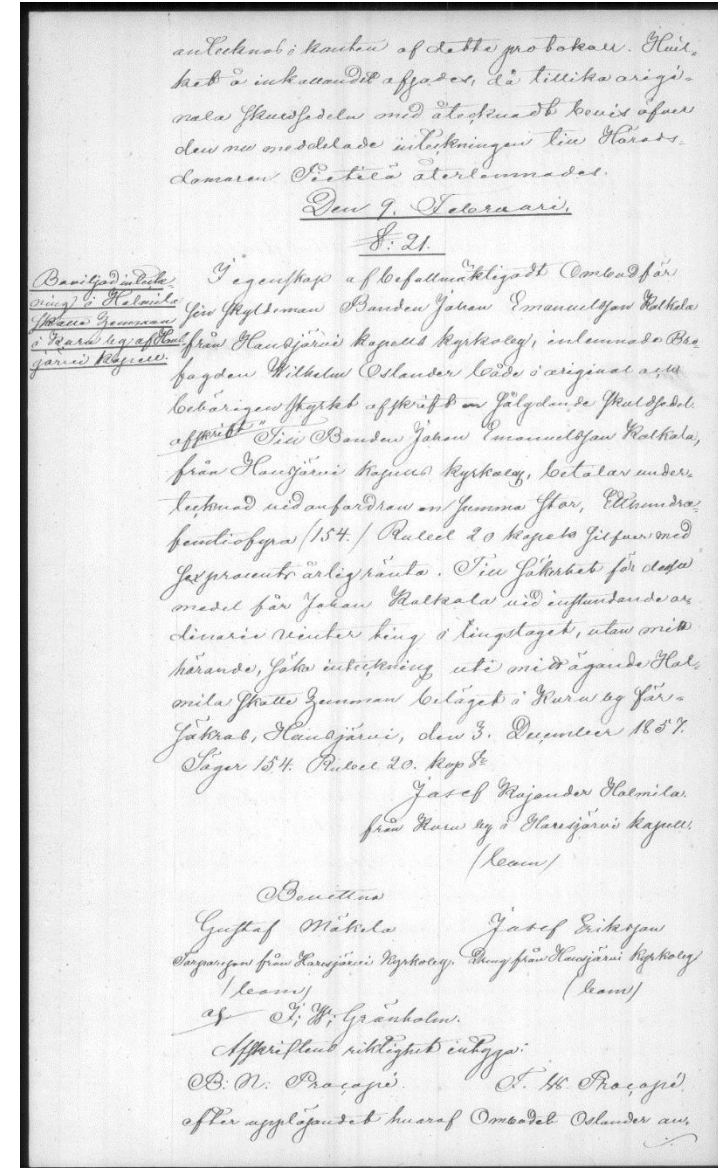
- 1-1 ensamt.
- 1-2 Torparena Johan Söderlund och August Karlsson
- 1-3 bebehålles vid de förmoner deras kontrakt innehålla
- 1-4 Köparena får tilltråda hemmans andelen denna dag.
- 1-5 Svartsmara den 12. December 1879.
- 1-6 Karl Erik Karlsson. Mathilda Katharina Karlss
- 1-7 /bom/
- 1-8 Äfvanstående köpeafhandling förklarar vi oss, till alla
- 1-9 delas nöjda, dag som äfvan
- 1-10 Daniel Gustaf Dahlqvist. Kristina Sofia Dahlqvist.

100 Filter



# 19th century renovated court records

- Pilot project
- Pro:
  - Lots of material, good quality
  - Uniform structure
- Con:
  - Long period, different handwritings





# Achievements during the project

- Text recognition model, character error rate 4-5 %
- 600 000 pages processed with the model
- Court records search interface
  - <https://transkribus.eu/r/kansallisarkisto/>



# Links & information

National Archives project blog  
<https://makingamodernarchive.blogspot.com/>

Court records search platform  
<https://transkribus.eu/r/kansallisarkisto/>

<https://transkribus.eu/>

<https://read.transkribus.eu/>

<https://transkribus.eu/wiki>

Twitter @Transkribus