

# Handwritten Text Recognition in the Archives

Maria Kallio  
Senior Research Officer  
[maria.kallio@arkisto.fi](mailto:maria.kallio@arkisto.fi)

Lauri Hirvonen  
Text Recognition  
Specialist  
[lauri.hirvonen@arkisto.fi](mailto:lauri.hirvonen@arkisto.fi)



THE NATIONAL ARCHIVES  
OF FINLAND

**READ**  
RECOGNITION & ENRICHMENT  
OF ARCHIVAL DOCUMENTS



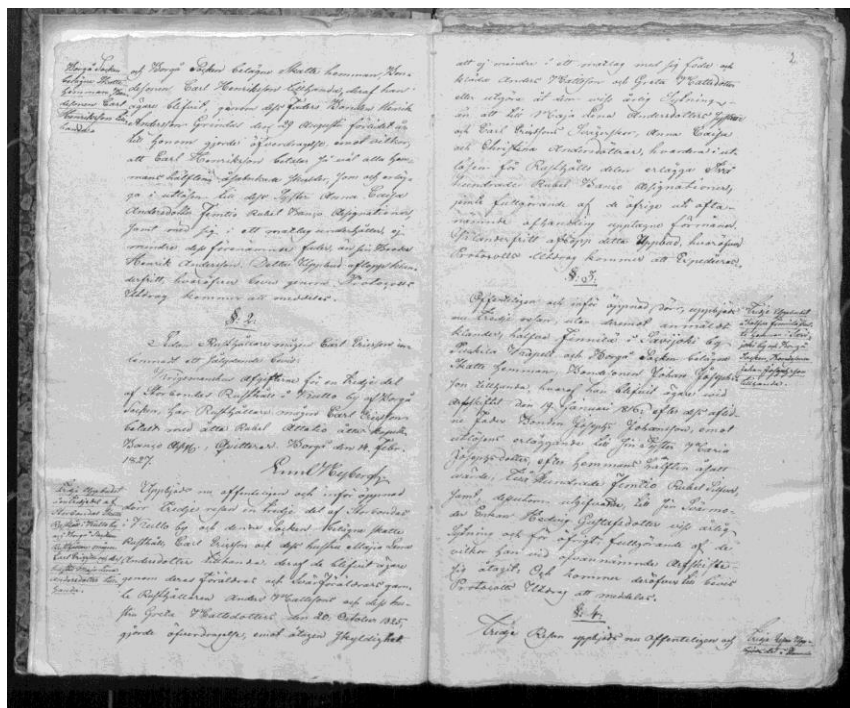


# The National Archives of Finland

- 210 linear kilometres of records
- 76 million digitized records
- 32 000 visits annually
- >1 million visits in the Digital Archives



# Renovated Court Records Collection



- One of the largest collections at NAF
- 1623-1970
- 800 000 pages already digitized
- Records from 1809-1917 will be digitized by FamilySearch in 2018-2019

# Making the Court Records available through HTR

- First batch of GT produced by volunteers in 2017
  - HTR model with CER 9,1%
- New data set of 680 images for training, validation and testing produced during 2018
  - HTR+ model with CER below 5%



Med förestående afhandling förklarar jag  
mig till alla delar nöjd; samt förbinder  
mig äfven att uppfylla densamma uti  
alla delar; som hämnad försäkres. Ohkon som  
äfvän.

Erik Johanson Suojala  
Vander son.  
(barn.)

Till vittnen.

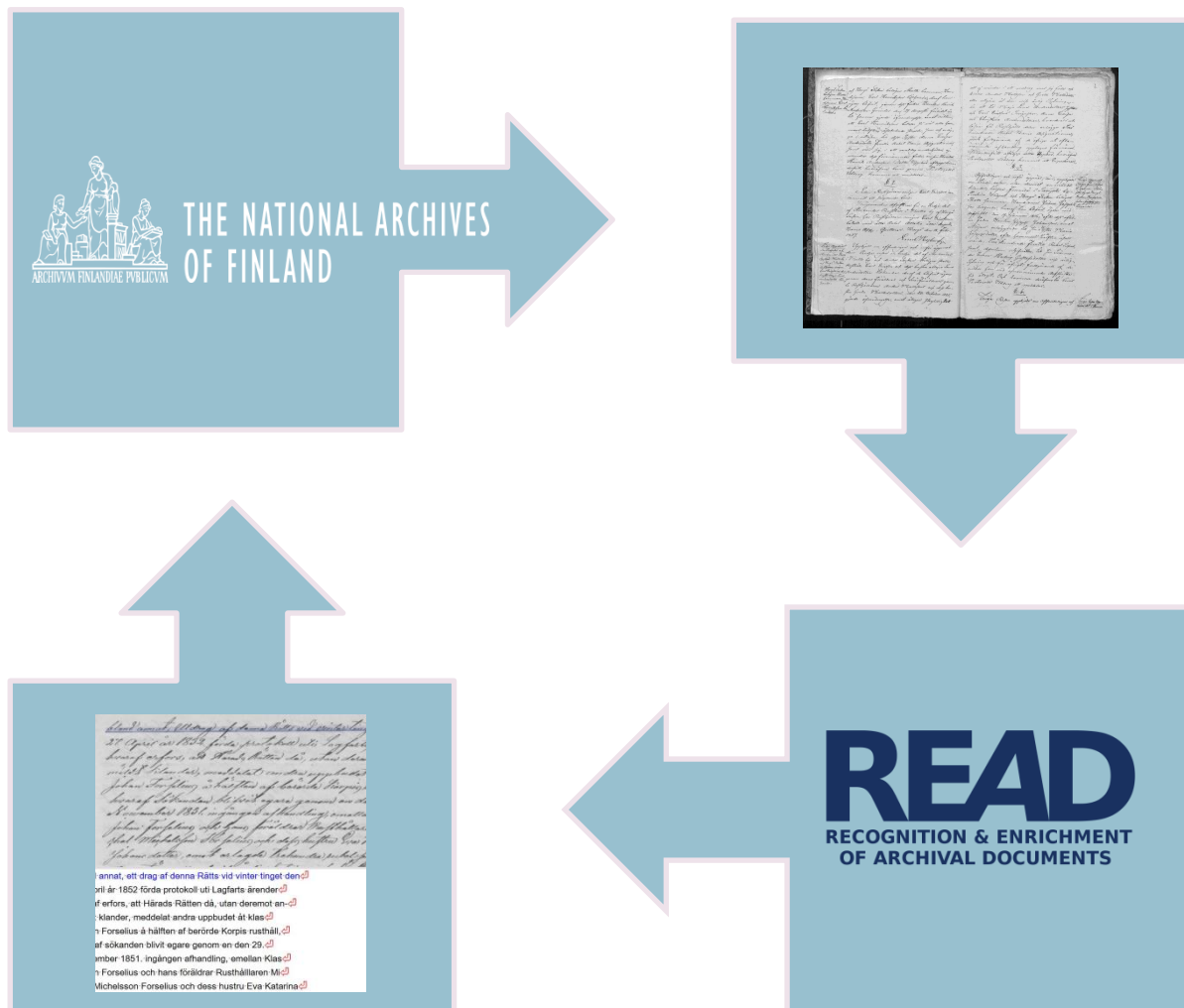
Erik Mattsson Oulila. Fredrik Carlsson Mikkola  
(barn.) Vander isän Ohkon (barn.)

- 23 Med förestående afhandling förklarar jag  
24 mig till alla delar nöjd, samt förbinder  
25 mig äfven att uppfylla densamma uti  
26 alla delar, som hämnad försäkres. Ohkon som  
27 gfon

och hans hustru Kristina Elisabet Bergqvist måtte  
meddelas andra uppbudet å Tolivis skatte hemman  
nr 1 om ett halft mantal i sagde by, hvilken lä-  
genhet sökanderne i ego bekommit förmedelst köpe-  
bref den 26 Januari siffrukne år affslutadt med  
Bonden Karl August Juslin och dennes hustru An-  
rona Karolina Juslin, född Lindahl, för utfärte  
ettusen rubel silfver, utom andra vilkor och be-  
ting, i hvilket affseende sökanden Gustaf Lindström  
inlemnade expedieradt utdrag ur denne Härad-  
rätts den 11 September siffrukne år under S: 18 i lag.  
parts ärendes förda protokoll, hvaraf inhemtades att  
Häradsrätten då meddelat sökanderne, å förbe-  
rörde hemman, det första klanderfria uppbudet, mist  
antecknandet hvaraf och som sökanderne redanta  
förene fullgjort allt hvad till lagfartens vinnande  
Häradsrätten sådt sökanderne i

- 2-1 och hans hustru Kristina Elisabet Bergqvist måtte ↵
- 2-2 meddelas andra uppbudet å Tolivis skatte hemman ↵
- 2-3 No 1 om ett halft mantal i sagde by, hvilken lä ↵
- 2-4 genhet sökanderne i ego bekommit förmedelst köp ↵
- 2-5 bref den 26 Januari sistvikne år afslutadt med ↵
- 2-6 Bonden Karl August Juslin och dennes hustru An ↵
- 2-7 rora Karolina Juslin, född Lindahl, för utfärte ↵
- 2-8 ettusen rubel silfver, utom andra vilkor och be ↵
- 2-9 ting, i hvilket afseende sökanden Gustaf Lindström ↵
- 2-10 inlemnade expedieradt utdrag ur denne Härad ↵
- 2-11 rätts den 11 September sistlidne år under SS: 18 i lag ↵
- 2-12 farts ärendes förda protokoll, hvaraf inhemtades att ↵
- 2-13 Häradsrätten då meddelat sökanderne, å förbe ↵
- 2-14 rörde hemman, det första klanderfria uppbudet, mist ↵
- 2-15 antecknandet hvaraf och som sökanderne redantill ↵
- 2-16 förene fullgjort allt hvad till lagfartens vinnande ↵
- 2-17 erfordrats fann Häradsrätten godt sökandernes ifrå ↵
- 2-18 gavarande anhållan bifalla; derå Taliais skatte ↵
- 2-19 hemman i Wirkby by och Löjs socken andra resan ↵
- 2-20 offentligen uppbjöds Bonden Gustaf Lindström och ↵







# Thank you for listening!

[maria.kallio@arkisto.fi](mailto:maria.kallio@arkisto.fi)

[lauri.hirvonen@arkisto.fi](mailto:lauri.hirvonen@arkisto.fi)