

# Tekoäly sukututkimuksessa - käytännön opastusta

19.2.2025

Pohjois-Kymen sukututkijat

Tuomas Kunttu

# Viimeaikaisia koulutuksia

- 9.12.2024 Juha Vuorela: **Sukututkijan maanantaitunti: Tekoäly**  
<https://jvuorela.wixsite.com/juhansuku/tietoboksi7>
- Katsottavissa vielä ainakin 18.2.2025
- 20.1.2025 Ari Vesalainen: **Tekoälyn hyödyntäminen sukututkimuksessa** (ei enää katsottavissa)

# Nyt käsittelyssä

- Automaattinen tekstintunnistus
- Generatiivinen tekoäly
  - ChatGTP

# Automaattinen tekstintunnistus

- Tekoälyä on opetettu isoilla tekstimassoilla
  - Kansallisarkiston (KA) Multicentury-HTR-Demo  
<https://huggingface.co/spaces/Kansallisarkisto/Multicentury-HTR-Demo>
  - Transkribus <https://www.transkribus.org/>
- Automaattista tekstintunnistusta hyödyntävä valmis sovellus:  
Tuomiokirjahaku <https://tuomiokirjat.kansallisarkisto.fi>

# Omia havaintoja KA:n tekstintunnistustyökalusta

## **Onnistuu paremmin**

- 1700-1800-luvun teksti
- Ruotsin kieli
- Koko sivun teksti
- Ei-rakenteinen teksti

## **Onnistuu huonommin**

- 1900-luvun teksti
- Suomen kieli (muita ei lainkaan)
- Sivulta rajattu teksti
- Rakenteinen teksti
  
- Yliviivaukset, alleviivaukset, läpikuultavat sivut jne.

# Generatiivinen tekoäly

- **Generatiivinen tekoäly** on tekoälyn osa-alue, joka pystyy luomaan uutta dataa sen sijaan, että se vain analysoi olemassa olevaa dataa.
- Yleisimpiä sovelluksia ovat tekstin, kuvan, äänen ja videon generointi.

# Tekoälysovelluksien hyötykäyttö

- Kääntäminen vanhasta ruotsista suomeksi
- Hakemiston tekeminen
- Kertovan tekstin tuottaminen
  
- Generatiivisen tekoälyn sovelluksia
- <https://chatgpt.com/>
- <https://copilot.microsoft.com/>
- <https://gemini.google.com/>
- <https://tekoälysovellukset.fi/>

# Miten auttaa tekoälyä tekemään paremman käännöksen?

## 1. Anna konteksti

- Sanojen ja ilmausten merkitys riippuu ajasta, paikasta ja asiayhteydestä -> kerro mihin teksti liittyy (esim. käräjäpöytäkirja, kuntakokous, sotapäiväkirja), paikka ja ajankohta
- Tämä voi auttaa tekoälyä valitsemaan oikeat termit.

## 2. Teksti mahdollisimman selkeänä

- Jos tekstissä on epäselviä kohtia (esim. käsinkirjoitettu vanha teksti), voit yrittää tulkita niitä ensin itse tai kertoa, mitkä kohdat ovat epäselviä.

## 3. Kerro toivotusta tyylistä

- Haluatko käännöksen mahdollisimman tarkasti vanhaa kieltä mukaillen vai sujuvaksi nykysuomeksi?

## 4. Pyydä tarkennuksia

- Jos jokin kohta käännöksessä tuntuu epäselvältä, pyydä tekoälyä tarkentamaan sitä tai antamaan vaihtoehtoisia käännöksiä.

## 5. Anna valmiita käännöksiä termeille ja paikannimille

- Jos tiedät, että jokin sana tarkoittaa tiettyä asiaa jossain vanhassa tekstissä, voit mainita sen.
- Esim. ”käytä käännöksessä seuraavia suomennoksia: nämndeman=lautamies, Eufraåminne=Eurajoki”

## 6. Kokeile pyytää tekstin ensin nykyruotsiksi ja vasta sen jälkeen suomeksi

## 7. Tee käännös kirjautuneena käyttäjänä



# ChatGPT-käännökset

## Ei kirjautunut

Mutta koska Henrich ei ole saapunut itse paikalle eikä ole lähettänyt edustajaa esteelleen, niin hänet tuomitaan väliin jääneeksi ja määrätty maksamaan 3 lakkia viljaa, 1 Dr 28/-, 5 Dr 24/- rahaa, sekä 1 Stuth, 7 C, joka yhteensä tekee 14 d:r 20/-.

Hänen on myös palautettava kärjääntoimittajan tekemiä oikeudenkäynnin kuluja 24 sölviä.

## Kirjautunut käyttäjä

### Käännös suoraan suomeksi

Koska mainittu Henrich ei ole saapunut paikalle eikä myöskään lähettänyt ketään ilmoittamaan poissaolonsa syytä, hänet julistetaan poissaolevaksi (contumax) ja tuomitaan maksamaan edellä mainitut 3 kuormaa viljaa, joiden arvo on 1 taalari 28 äyriä, lisäksi 5 taalaria 24 äyriä rahassa sekä härkä, jonka arvo on 7 taalaria. Kaikkiaan summa tekee 14 taalaria ja 20 äyriä.

Hänen tulee myös korvata kantajan oikeudenkäyntikulut 24 äyrillä hopearahaa.

## Kirjautunut käyttäjä

### Muunnos ensin nykyruotsiksi ja sitten suomeksi

Koska mainittu Henrik ei saapunut paikalle eikä lähettänyt mitään ilmoitusta poissaolonsa syystä, hänet tuomitaan poissaolosta ja määrätään maksamaan edellä mainitut 3 kuormaa viljaa arvoltaan 1 taalari ja 28 äyriä, 5 taalaria ja 24 äyriä rahana sekä härkä arvoltaan 7 taalaria. Tämä tekee yhteensä 14 taalaria ja 20 äyriä.

Hänen on myös korvattava kantajan oikeudenkäyntikulut, jotka ovat 24 äyriä hopearahaa.

# Hakemiston tekeminen

## Copilot

Tässä tulee pitkä usean sivun teksti Eurajoen (Euraåminne) käräjäpöytäkirjasta vuodelta 1726. Tee lista henkilöistä, jotka siinä mainitaan. Mainitse henkilön koko nimi sekä sivunumero, jolla henkilö mainitaan. Jos nimeen liittyy titteli tai paikan- tai talon nimi, merkitse se suluissa henkilön nimen perään. Järjestä nimet aakkosjärjestykseen etunimen mukaan. Sivunumero mainitaan aina sivun alussa suomeksi, muuten teksti on vanhaa ruotsia.

Tässä on lista henkilöistä, jotka mainitaan Eurajoen (Euraåminne) käräjäpöytäkirjassa vuodelta 1726. Lista on järjestetty etunimen mukaan aakkosjärjestykseen ja sisältää henkilön koko nimen, mahdolliset tittelit tai paikan- tai talon nimet suluissa sekä sivunumeron, jolla henkilö mainitaan:

1. Abraham Lufven (Inspectoren, Wälbetrodd) - sivu 149, 154
2. Anders Carlström (Capitain, Wälborne Hr Niels) - sivu 162
3. Anders Thomas (Anders Thomasson) - sivu 178
4. Brita Henrik (Brita Hendricks dr, Ryttare Enkan) - sivu 143
5. Brita Johans dr - sivu 144, 157, 160
6. Erik Anderson (Frälsebonden) - sivu 145
7. Erik Anderson Teppo (Borgaren, utj Raumo) - sivu 170
8. Erik Erson (Erik Ersons Resolins Enka Lisa Matz dr) - sivu 168
9. Eskil Johansson (Bonden och Nembdeman) - sivu 157
10. Fredrik Jöransson (Pakola) - sivu 174

## ChatGPT

1. Abraham Lufven (Inspector, Wälbetrodd) – Sivun 149
2. Anders Hillendahl (Lijfdragoun, förafskierad) – Sivun 149
3. Anders Matsson (Augments Bonden, Kaufonmäki) – Sivun 171
4. Anders Anderson (Poika, Zachris Bollstedtin palveluksessa) – Sivun 163
5. Anna Leistenia (Hustru, Johan Simolins maka) – Sivun 166
6. Beata Hendriksdotter (Dreng hustrun, Kautomäki) – Sivun 162
7. Brita Hendriksdotter (Enka, Löyty) – Sivun 144
8. Brita Jacobsdotter (Rusthållare Enka, Sampanala) – Sivun 144
9. Brita Johansdotter (Dreng hustrun, Kukola, Lufwia, Eskill Johansonin sisar) – Sivun 144, 157-160
10. Carl Henrikson (Bonde, Attestant) – Sivun 161

# Kertovan tekstin tuottaminen faktoista

Tässä on tietoja Johan Hermaninpoika Kuntun elämästä. Kirjoita näiden tietojen sekä muualta saatavien lisätietojen perusteella suorasanainen elämäkerta Johanista. Lisää Suomen historiaan liittyviä asioita sopiviin kohtiin esim. perustelemaan muuttoa.

Johan Hermaninpoika Kunttu

s. 25.1.1848 Viipurin mlk, Rakkolanjoki

Vanhemmat itsellinen Herman Matinpoka Kunttu 1824-1886 ja

Maria Pekantytär Töllikkö 1824-1898.

renkinä Rakkolanjoen Sollolla 1866 ja 1867-1872 Joenpolven tilalla.

vihitty 29.6.1879 Maria Hämäläisen kanssa Rakkolanjoella

lapsi Alma Wilhelmiina 15.7.1880-19.8.1881

Vaimo kuoli 28.11.1880 Rakkolanjoella

uusiin naimisiin 18.12.1881 Rakkolanjoella Anna Sofia Naukkarisen kanssa.

Lapset:

Antti s. 2.12.1884, lähti ja hävisi Amerikkaan

Herman Johan s. 4.4.1893 Rakkolanjoki

August s. 15.2.1896 Valkeala

Perhe muutti Valkealan Ruotsulaan 18.4.1894

Johan Hermaninpoika Kunttu työskenteli työmiehenä Kymen

## Johan Hermaninpoika Kunttu (1848–1918)

Johan Hermaninpoika Kunttu syntyi 25. tammikuuta 1848 Viipurin maalaiskunnan Rakkolanjoella itsellinen Herman Matinpoka Kuntun ja Maria Pekantytär Töllikön perheeseen. Hän kasvoi aikana, jolloin Suomi oli osa Venäjän keisarikuntaa, mutta vanhoja ruotsalaisia hallintotapoja ja perinteitä noudatettiin yhä.

Nuorena Johan työskenteli renkinä Rakkolanjoen Sollolla vuonna 1866 ja myöhemmin vuosina 1867–1872 Joenpolven tilalla. Rengintyö oli tavallista talonpoikaisnuorille, jotka eivät perineet omaa maata. 1800-luvun jälkipuoliskolla Suomi eli suurten yhteiskunnallisten muutosten aikaa, ja maatalous kehittyi, mutta samalla teollistuminen alkoi vaikuttaa työväestön elämään.

Johan avioitui 29. kesäkuuta 1879 Maria Hämäläisen kanssa Rakkolanjoella. Parin esikoistytär Alma Wilhelmiina syntyi 15. heinäkuuta 1880, mutta menehtyi vain vuoden ikäisenä 19. elokuuta 1881. Jo ennen lapsen kuolemaa, 28. marraskuuta 1880, Johan jäi leskeksi vaimonsa Marian menehdyttyä.

Elämän oli kuitenkin jatkuttava, ja Johan meni uusiin naimisiin 18. joulukuuta 1881 Anna Sofia Naukkarisen kanssa. Perheeseen syntyi kolme lasta: Antti (1884), Herman Johan (1893) ja August (1896). Antti lähti aikuisena Amerikkaan, mutta hänen kohtalonsa jäi perheelle tuntemattomaksi.

Vuonna 1894 perhe muutti Viipurin maalaiskunnasta Valkealan Ruotsulaan. Muutto osui aikaan, jolloin maassamuutto yleistyi, ja yhä useampi suomalainen etsi työtä maatalouden ulkopuolelta. Johan siirtyi työskentelemään työmiehenä

# Muita

- Tekstin parantaminen
- Yhteenvetojen tekeminen
- Käännökset

# Luotettavuus ja tietoturva

- Tarkista aina kaikki tekoälyn tuotokset
- Ilmaispalveluita
  - Muuttuvat jatkuvasti
  - Ei velvoitteita ilmaiskäyttäjiä kohtaan
  - Jostain kuitenkin saavat tulot (mainostajat, annettujen tietojen myynti jne.)
- Älä kerro tekoälylle arkaluontoisia tietoja

*Ibidem*  
*Tobidem*

*Michel AnderBon*

*Hr. Carin Josephs dr*

*Mag:n Henr MattzBon.*

*Sr Sophia Michels dr.*

*dott: Caisa Michels dr*

*Mattz PhilipBon*

*Hr Anna Christurs dr*

*Son Christer MattBon*

*Erick: Lisa Mattzdr*

Uotidem

Tobidem

Michel AnderBon

Hr. Carin Josephs dr

Mag:n Henr MattzBon.

Sr Sophia Michels dr.

dott: Caisa Michels dr

Mattz PhilipBon

Hr Anna Christurs dr

Son Christer MattBon

Erick: Lisa Mattzdr

Den 3 Septemb: Hendrich LahrBon Ryttere för Niman,  
tago och pigan Maria Zachariadåtter i Parskylä.  
Den 9 Dito Erich ThomBon en äre dreen i Suomenkylä och  
pigan Maria Thomasdotter ibid.  
Den 26 Dito Rytteren Erich MichelBon i Yljoki och Anna  
Andersdåtter piga i Wärla.  
Den 30 Dito Hendrich MichelBon Smedh i Lemagård och Kir-  
stin Jöransdåtter NyckeEpigan ibid.

Den 3 Septemb: Hendrich LahrBon Ryttere för Niman  
tago och pigan Maria Zachariadåtter i Parskylä.  
den 9 Dito Erich ThomBon en äre dreen i Suomenkylä och  
pigan Maria Thomasdotter ibid:  
den 26 Dito Rytteren Erich MichelBon i Yljoki och Anna  
Andersdåtter piga i Wärla.  
den 30 dito Hendrich MichelBon Smedh i Lemagård och Kir-  
stin Jöransdåtter NyckeEpigan ibid.

Restart this Space

This Space is sleeping due to inactivity.




# Multicentury HTR Demo

The HTR pipeline contains three components: text region detection, textline detection and handwritten text recognition. The components run machine learning models that have been trained at the National Archives of Finland using mostly handwritten documents from 17th, 18th, 19th and 20th centuries.




Input image can be uploaded using the *Input image* window in the *Text content* tab, and the predicted text content will appear to the window on the right side of the image. Results of text region and text line detection can be viewed in the *Text regions* and *Text lines* tabs. Best results are obtained when using high quality scans of documents with a regular layout.

Text content Text regions Text lines

**Input image**



Drop Image Here  
- or -  
Click to Upload

**Predicted text content**

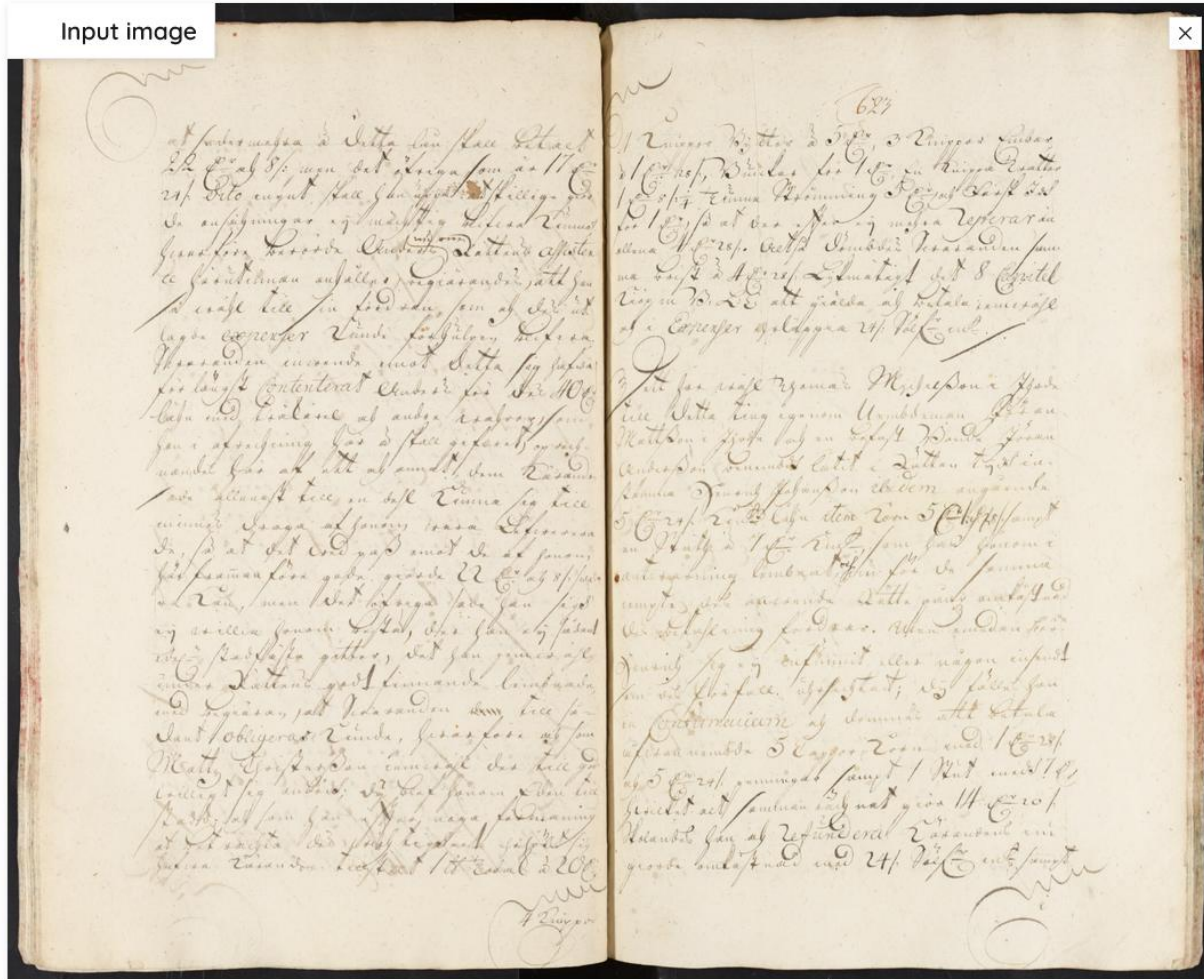
**Process image**

Text content

Text regions

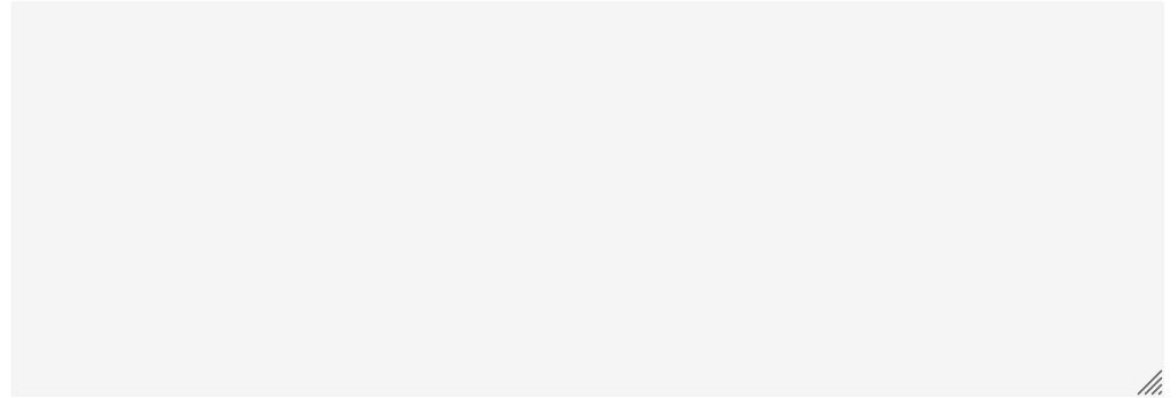
Text lines

Input image

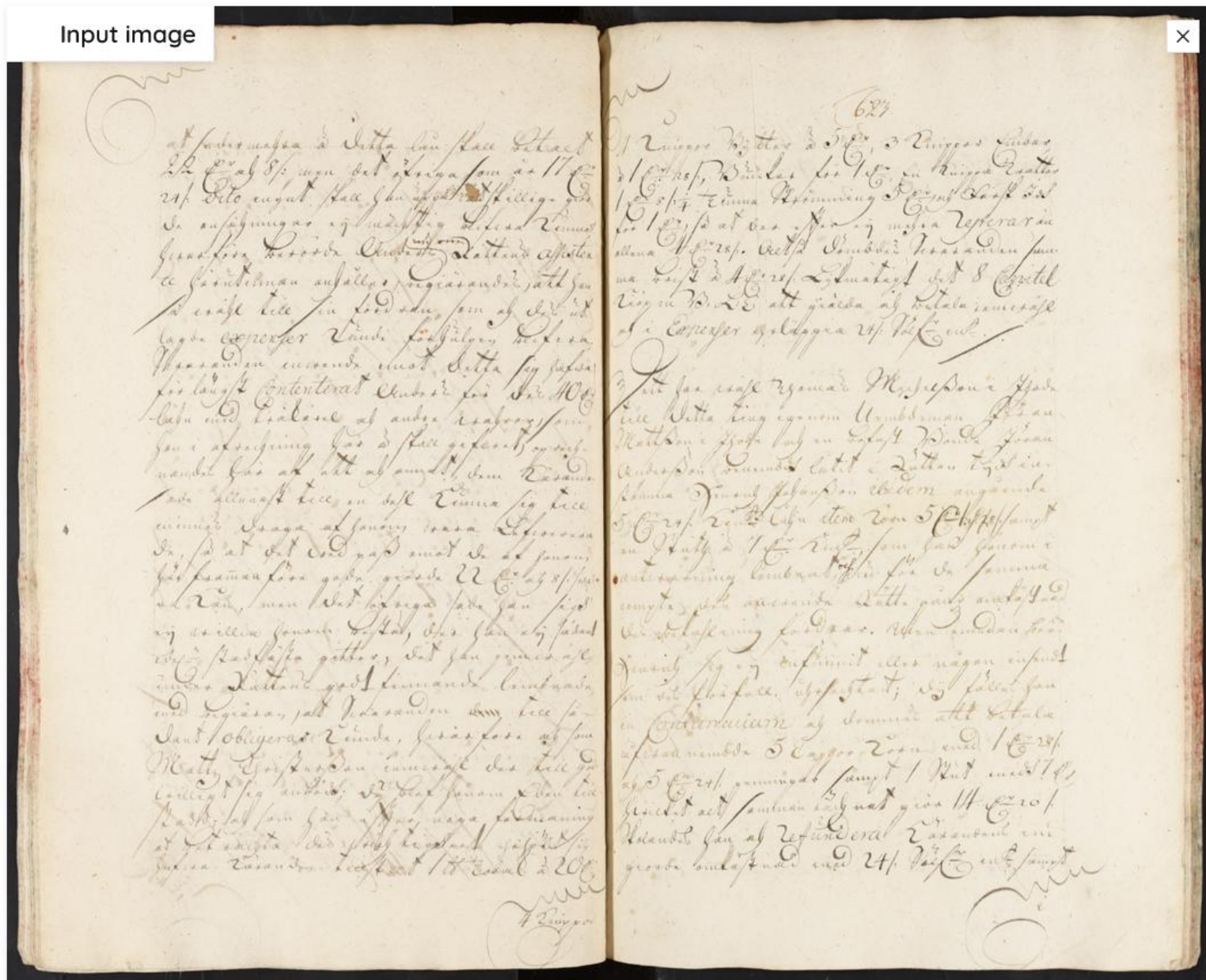


Process image

Predicted text content



## Input image



## Predicted text content

at sedermera å detta lån skall betalt  
 2 / 2 C:r och 8 /: men det öfriga som är 17 d:r  
 24 /: Dito mynt skall han öppåt åtskillige gör  
 de ansågningar ej mächtig blifwa Tunna  
 nu om  
 hwarföre berörde Anders Rättens assisten  
 ce härutinnan anhåller begiärandes, att han  
 så wähl till sin fordran, som och dess ut  
 lagde expenser kunde förhulgen blifwa  
 Swaranden inwende emot detta sig hafwa  
 för långst Contenterat Anders för des 40 c:r  
 lähn med tråkarel och andre wahror, som  
 han i afreching här å skall gifwet, op och  
 nandes här af ett och annat, dem Kärande  
 sade allenast till en dehl kunna sig till  
 minnes draga af honom wara Lefwerera  
 de, så at det wed pass emot de af honom  
 här framman före gode gjorde 22 d:r och 8 /: swär  
 en kan, men det öfriga sade han sikh  
 ej willia honom bestå, der han ej sådant  
 edel:n stadfästa getter, det han jemwähl  
 under Rättens godt finnande lembnade  
 med begiäran, at Swaranden by till så  
 dant obligeras kunde, hwarföre och som  
 Matts Christersson jemwähl der till god  
 willigt sig anhädt; du blef honom Eden till