



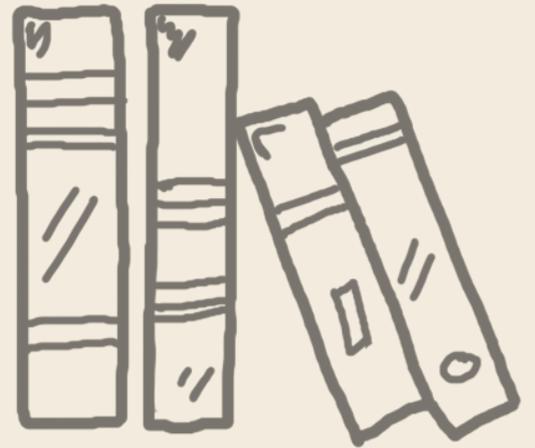
Transkribus®

23. Februar 2023 | DigEdTnT: Workshop - Meet the Tools and Developers

Sebastian Colutto & Matthias Sorg, READ-COOP SCE



Können Sie das lesen?



16 Feb. Ao 1610.

Ehrentveste, achtbare vnd hochgelarte,
Großgönstige Herrn, Nehist meiner
alzeit willig gefliessener Dienste er-
bietung, kan E. E. g. Ich vnvormeldet
nit lassen, Das vorschienere Jare Marx

16 Feb. Ao 1610

Ehrentveste, Achtbare vnd hochgelarte,
Großgönstige Herrn, Nehist meiner
alzeit willig gefliessener Dienste er-
bietung, kan E. E. g. Ich vnvormeldet
nit lassen, Das vorschienere Jare Marx

Das Transkribus-Ökosystem:

Die digitale Verfügbarkeit ermöglichen

Bilderfassung

Bearbeitung, Datenerfassung +
Annotation und KI-Training

KI-gestützte
Texterkennung

Veröffentlichung



citizen&science

Transkribus
metagrapho <api>



Was ermöglicht **Transkribus**?



Transkription des Textes



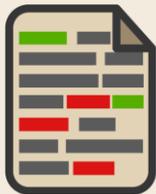
KI-Training



Layout-Erkennung



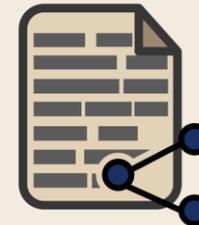
Texterkennung



Anreicherung des Textes



Durchsuchen von Dokumenten



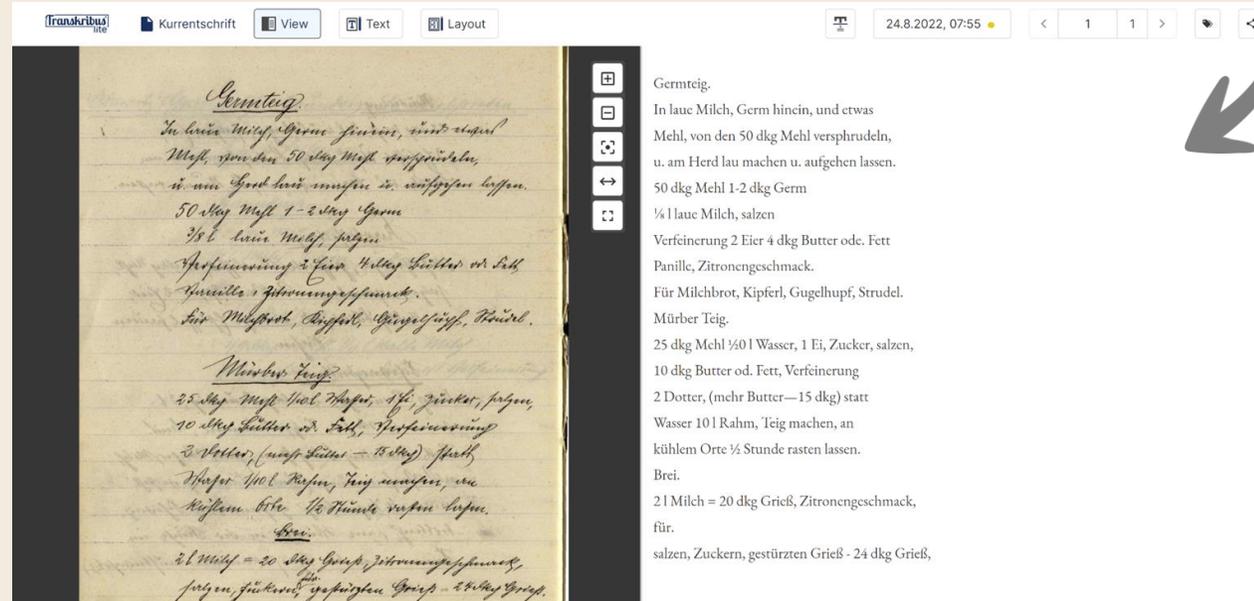
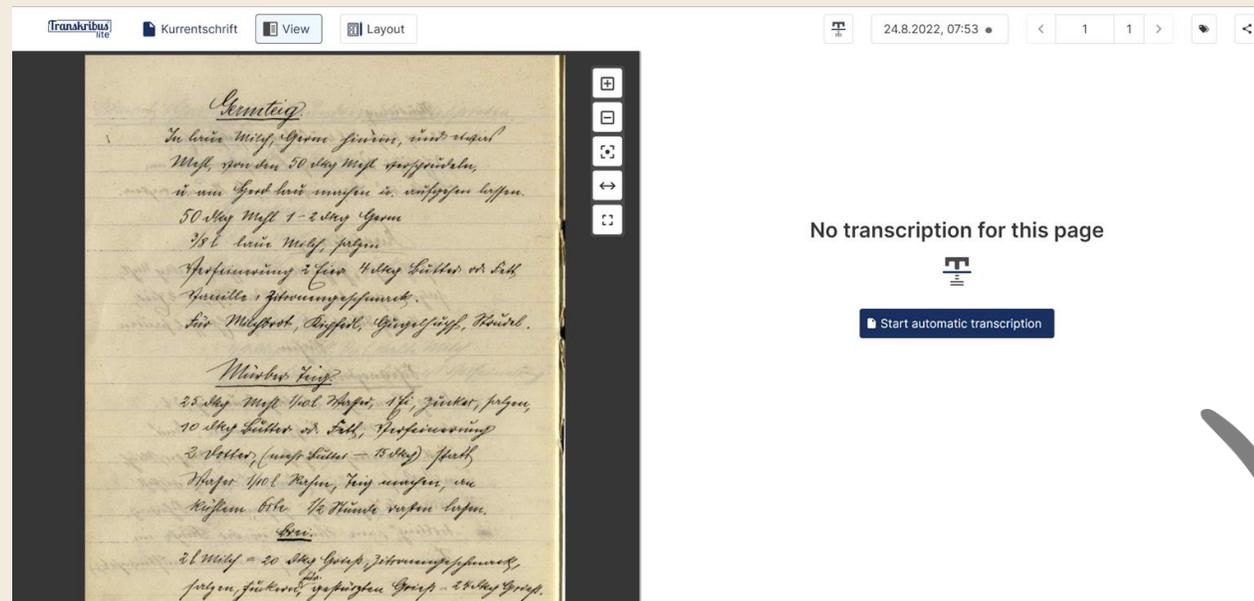
Export/Veröffentlichung



Zusammenarbeit

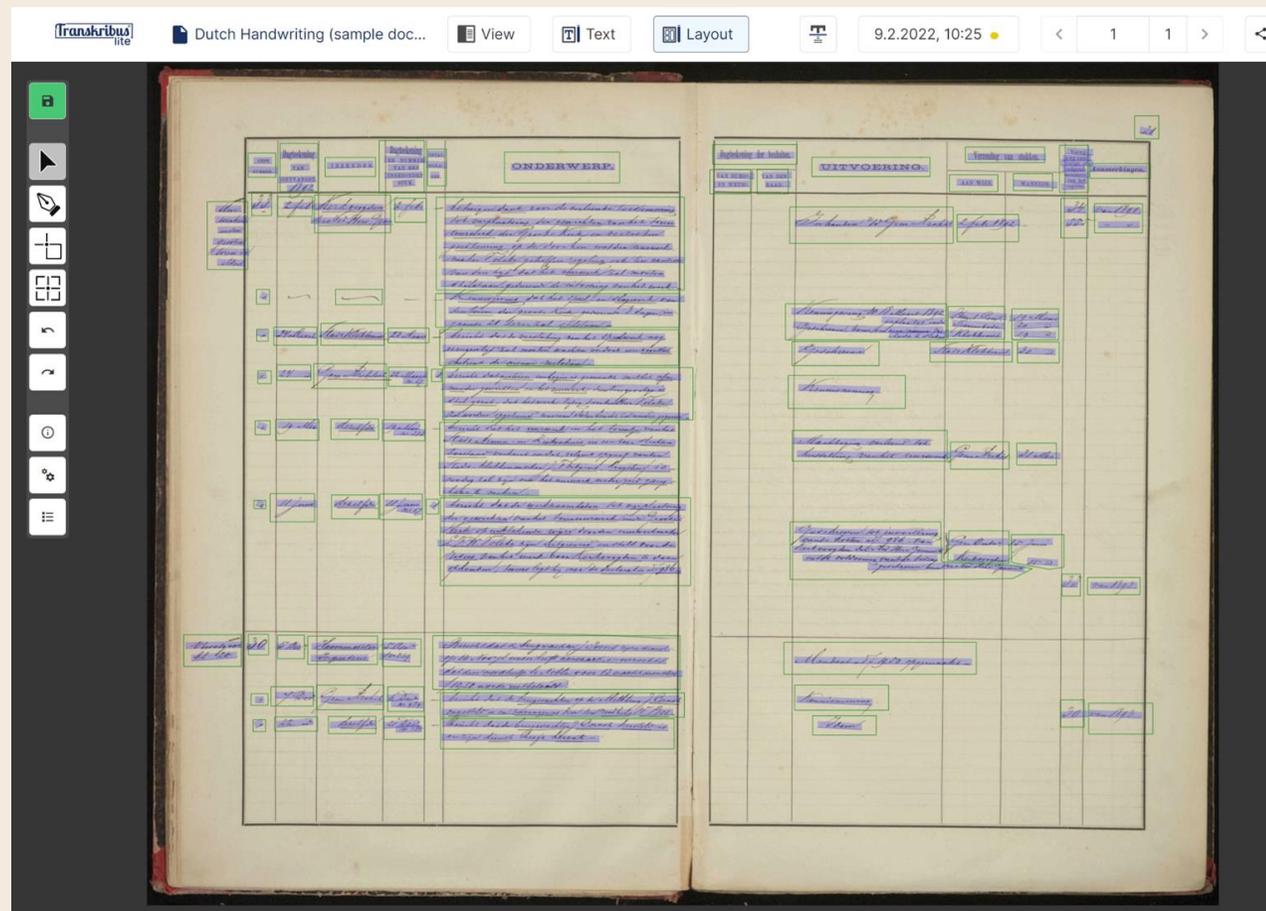
Texterkennung

- Automatische Erkennung von Text auf Scans
- Handschriftliche und gedruckte Texte
- Möglich durch KI-Modelle



Layout-Erkennung

- Trainierbare Layout-Erkennung
- Basis einer guten Texterkennung
- Möglichkeit der strukturierten Erkennung komplexer Dokumente



Eigene KI-Modelle trainieren

- KI für jeden zugänglich
- Erstellung eigener Modelle möglich
- Weiterentwicklung bestehender Modelle möglich

The screenshot shows a web interface for training a text model. At the top, there are navigation tabs for 'Workdesk' and 'Training', a search bar with the placeholder 'Enter search term ...', and a user profile icon with the number '899'. Below the navigation, there is a 'Text Model' button and a collection name 'learn@readcoop.eu Collection'. The main section is titled 'Text Training' and features a progress bar with four steps: 'Model Setup', 'Training data', 'Validation data', and 'Start'. The 'Model Setup' step is currently active. Below the progress bar, there are several input fields: 'Model Name' with the value 'Kurrento - My Custom AI Model', 'Description' with the text 'This is my own model to recognise German Kurrentschrift', and 'Language' with 'German' selected and an 'Add a language' button. There is also a 'Centuries' slider with markers for '17' and '19'. At the bottom, there is a 'Transcript version' section with two buttons: 'Latest Transcript' (selected) and 'Ground truth only'.

Zusammenarbeit

- Kollaboration in Teams
- User Management
- Zentrale Creditverwaltung möglich

The screenshot displays the Transkribus interface. On the left is a sidebar with 'TOOLS' (Text Recognition, Layout Recognition, Upload, Download, Delete, Users-Manager) and 'TRANSKRIBUS ORGANIZER' (Collections, Jobs, Tag Manager, Recent Documents). The main area shows a 'Collection Overview' for 'learn@readcoop.eu' with a grid of document thumbnails. A 'Users' dialog box is open in the foreground, featuring a 'Manage' tab and an 'Add' button. It includes a search field with the email 'shop@readcoop.eu' and a 'Search' button. Below the search, it lists 'Found following users:' with one entry: 'User: shop@readcoop.eu', 'Email: shop@readcoop.eu', and a 'Transcribe' dropdown with a '+ Add' button. A 'Feedback' button is visible on the right edge of the interface.

Durchsuchen von Dokumenten

- Volltextsuche
- Smart Search, um schwierige Suchbegriffe zu finden sowie Rechtschreib- und Erkennungsfehler oder Schreibvarianten zu überbrücken

The screenshot displays the Transkribus interface. On the left is the 'TRANSKRIBUS ORGANIZER' sidebar with options: Collections, Jobs, Tag Manager, and Recent Documents. The main area shows a search for 'Mehl' with 6 results. The first result is from a document titled '/ Kurrentschrift / Page 1'. The text in the image is a recipe: '... Germteig. In laue Milch, Germ hinein, und etwas Mehl, von den 50 dkg Mehl versphrudeln, u. am Herd ... lau machen u. aufgehen lassen. 50 dkg Mehl 1-2 dkg Germ 1/2 l laue Milch, salzen Verfeinerung 2 Eier 4 ... Teig. 25 dkg Mehl 1/2 l Wasser, 1 Ei, Zucker, salzen, 10 dkg Butter od. Fett, Verfeinerung 2 Dotter ...'. Below the text are four horizontal strips of the original handwritten manuscript, with the word 'Mehl' highlighted in red in each strip. The interface also includes a search bar at the top right with 'Mehl' entered, a 'Search' button, and a 'Feedback' button on the far right.

Transkription des Textes

- Integrierter Transkriptions-Editor
- Dokumente manuell transkribieren
- Automatisch generierte Transkriptionen bearbeiten

The screenshot displays the Transkribus web interface. At the top, there is a navigation bar with the logo 'Transkribus', a document icon labeled 'Kurrentschrift', and buttons for 'View', 'Text', and 'Layout'. The date '24.8.2022, 07:55' and page numbers '1 1' are also visible. The main area shows a scanned page of handwritten text in cursive. A section of the text is highlighted with a grey box, and a corresponding transcribed text block is shown below it. The transcribed text is as follows:

25 dkg Mehl 1/20 l Wasser, 1 Ei, Zucker, salzen,
10 dkg Butter od. Fett, Verfeinerung
2 Dotter, (mehr Butter—15 dkg) statt
Wasser 10 l Rahm, Teig machen, an
kühlem Orte 1/2 Stunde rasten lassen.
Brei.
2 l Milch = 20 dkg Grieß, Zitronengeschmack,
für.
salzen, Zuckern, gestürzten Grieß - 24 dkg Grieß,

Anreicherung des Textes

- Anreicherung mit Metadaten durch Markieren, Taggen und Annotieren
- Mit oder ohne Unterstützung der KI
- Aufbau einer Grundlage für eine ausgefeilte Informationsextraktion

The screenshot displays the Transkribus interface. On the left, a handwritten manuscript page is shown with the title 'Gernteig' and the following text: 'In laue Milch, Germ hinein, und etwas Mehl, von den 50 dkg Mehl verschrudeln, u. am Herd I au machen u. aufgehen lassen. 50 dkg Mehl 1-2 dkg Germ 3/8 l laue Milch, folgen Verfeinerung 2 Eier 4 dkg Zucker od. Salz Vanille, Zitronengeschmack. Für Milchbrot, Kipferl, Gugelhupf, Strudel. Mürber Teig. 25 dkg Mehl 1/2 l Wasser, 1 Ei, Zucker, salzen, 10 dkg Butter od. Fett, Verfeinerung 2 Dotter, (mehr Butter—15 dkg) statt Wasser 10 l Rahm, Teig machen, an kühl em Orte 1/2 Stunde rasten lassen. Brei. 2 l Milch = 20 dkg Grieß, Zitronengeschmack, für sa len, Zuckern, gestürzten Grieß - 24 dkg Grieß.'

On the right, the digital transcription is shown with various annotations. The title 'Gernteig.' is followed by the text: 'In laue Milch, Germ hinein, und etwas Mehl I, von den 50 dkg Mehl verschrudeln, u. am Herd I au machen u. aufgehen lassen.' Below this, ingredients are listed: '50 dkg Mehl' (highlighted in blue), '1-2 dkg Germ', and '3/8 l laue Milch' (highlighted in blue), followed by ', salzen'. A 'Verfeinerung' section includes '2 Eier', '4 dkg Zucker', and 'Zutaten'. The text continues: 'Pani I le, Zitronengeschmack. Für Milch brot, Kipferl, Gugelhupf, Strudel. Mürber Teig. 25 dkg Mehl I 1/2 l Wasser, 1 Ei, Zucker, salzen, 10 dkg Butter od. Fett, Verfeinerung 2 Dotter, (mehr Butter—15 dkg) statt Wasser 10 l Rahm, Teig machen, an kühl em Orte 1/2 Stunde rasten lassen. Brei. 2 l Milch = 20 dkg Grieß, Zitronengeschmack, für sa len, Zuckern, gestürzten Grieß - 24 dkg Grieß.'

Export von Dokumenten

- Verschiedenste Formate verfügbar
- z.B. XML, PDF, TEI oder Docx
- Export der Bilder möglich

The screenshot displays the Transkribus web interface. On the left, a sidebar lists various tools and organizer functions. The main area shows a collection of handwritten document pages. A 'Download' dialog box is open in the foreground, offering options to download as an image or as a file in various formats. The 'Files' section includes checkboxes for Page XML, Alto, PDF, TEI, and Docx. Below this, it indicates that one document is selected and provides a green 'Start' button to initiate the download process.

Transkribus lite

Workdesk Training

Enter search term ... 898

Home Collection Overview learn@readcoop.eu Collection

Filter by title or ID

TOOLS

- Text Recognition
- Layout Recognition
- Upload
- Download
- Delete
- Users-Manager

TRANSKRIBUS ORGANIZER

- Collections
- Jobs
- Tag Manager
- Recent Documents

Download

Image

- Images

Files

- Page XML
- Alto
- PDF
- TEI
- Docx

Selected documents

- Kurrentschrift Pages: 1

Start

Kurrent-Rezept

- Select

Test

- Select

TRAINING_VALIDATION_SET_Test

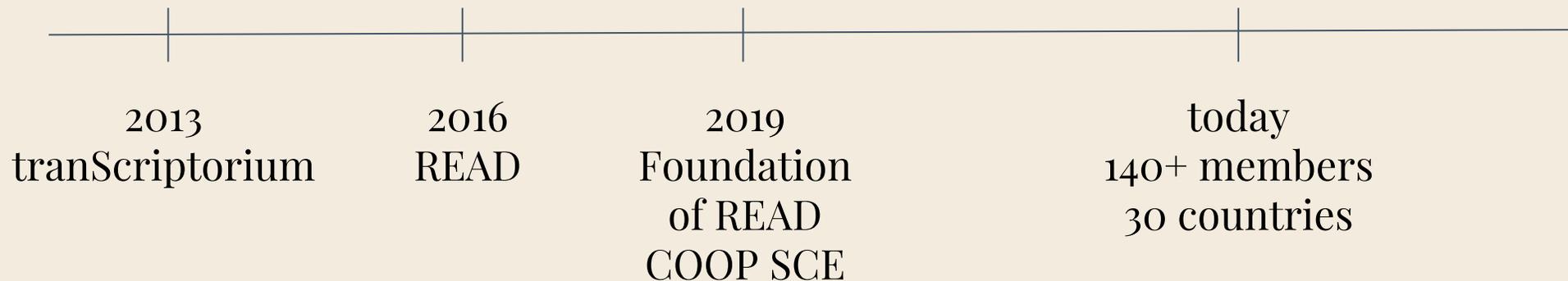
- Select

Feedback

Wer steckt hinter **Transkribus**?

READ-COOP SCE

Eine Europäische Genossenschaft zur Aufrechterhaltung und Weiterentwicklung der Transkribus-Plattform.

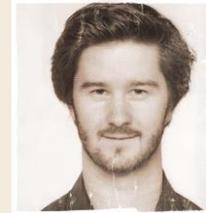


Wer steckt hinter Transkribus?



Das KI-Team

30+ Mitarbeiter mit Sitz in Österreich, Deutschland, Italien und UK



* Einige
Gesichter
fehlen noch :-)

READ-COOP in Zahlen

~ 110.000 Nutzer

30 Mio. Seiten texterkannt

**50.000.000
Hochgeladene Scans
insgesamt**

**150 +
Länder mit aktiven
Nutzer*innen**

**~ 30
read&searches online**

**~ 20.000
trainierte Modelle**

**40.000 +
Monatliche Website-
Besucher*innen**

**1000 +
ScanTents in Verwendung**

Das Transkribus-Ökosystem:

Die digitale Verfügbarkeit ermöglichen

Bilderfassung

Bearbeitung, Datenerfassung +
Annotation und KI-Training

KI-gestützte
Texterkennung

Veröffentlichung



citizen&science

Transkribus
metagrapho <api>



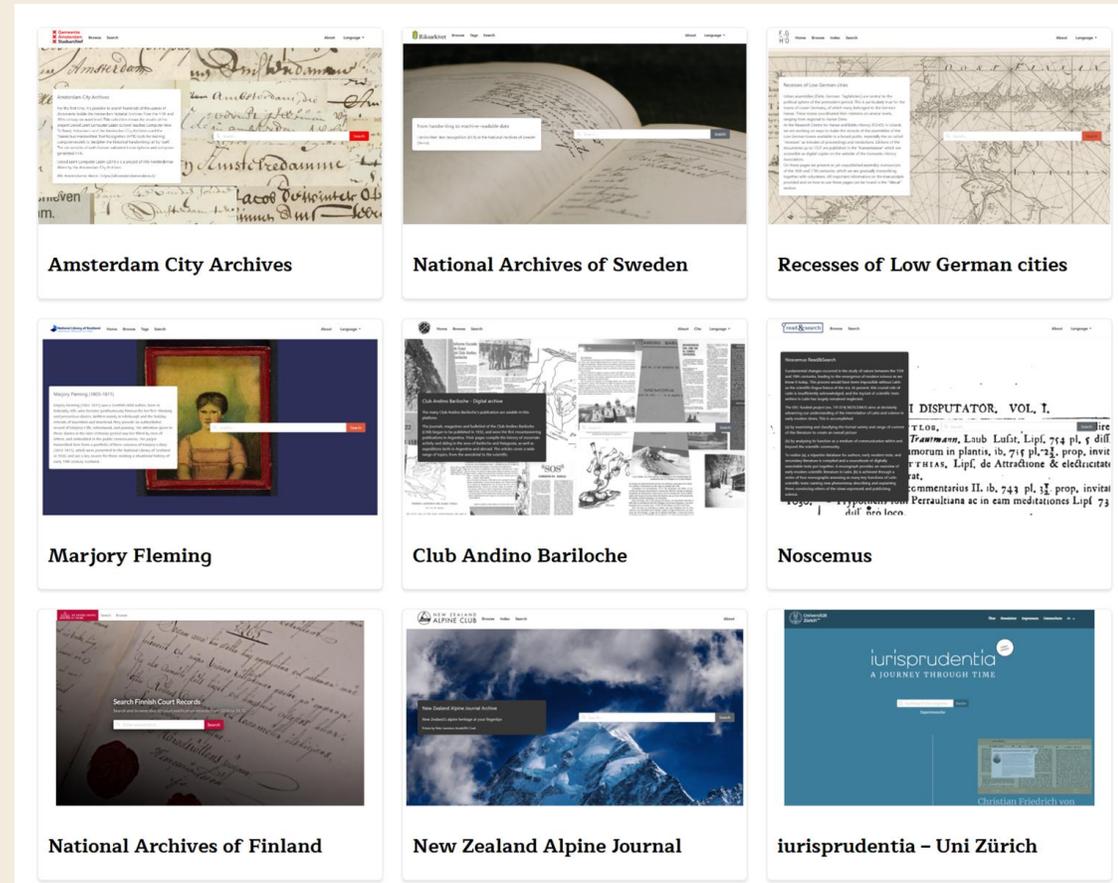
Das ScanTent

- Hochwertiges Scannen für unterwegs
- Nutzung des eigenen Smartphones
- Eigens entwickelte DocScan App
- Leichtes, faltbares Zelt
- Schnell auf- und abgebaut
- Standardisierter Hintergrund
- Ideale Lichtbedingungen durch integrierte LED Beleuchtung



read&search

- Einfache und effiziente Publikationsmöglichkeit
- Ohne Programmierkenntnisse oder IT-Ressourcen nutzbar
- Teilen mit Öffentlichkeit oder privater Gruppe möglich
- Volltext- und Fuzzy-Suchfunktionen
- Smart Search integrierbar
- Tag-Suche verfügbar
- Überlagerung des Scans mit einer Text-Layer oder Gegenüberstellung möglich
- Beispiel: Amsterdamer Stadtarchiv



metagrapho api

- Extraktion von Text- und Layout-Informationen
- Integration von KI-basierter Texterkennung in Ihren bestehenden Workflow oder Ihre Toolchain
- Erstellung völlig neuer Anwendungen zur Texterkennung
- Zugriff auf alle KI-Modelle, die Sie in Transkribus trainiert haben

```
{
  "processId": 47725,
  "status": "FINISHED",
  "content": {
    "text": "Eisenschrott",
    "regions": [
      {
        "id": "region_1",
        "coords": {
          "points": "0,1 282,1 282,107 0,107"
        },
        "lines": [
          {
            "coords": {
              "points": "5,106 37,109 59,92 80,91 97,105 156,108 195,76 220,69 254,82
283,66 283,11 247,9 228,26 205,5 173,15 146,3 98,13 71,0 4,6"
            },
            "baseline": {
              "points": "9,68 278,67"
            },
            "text": "Eisenschrott"
          }
        ]
      }
    ]
  }
}
```

Transkribus On-Prem

- Ermöglicht das Hosting des Transkribus Workflows in Ihren Räumlichkeiten
- Nutzung Ihrer eigenen Infrastruktur
- Daten verlassen nie Ihre Server
- Lokale Verarbeitung der Scans
- Senden von Daten an Transkribus-Server ausschließlich zu Lizenzzwecken und Statistik (ohne Internetzugang möglich)

Transkribus[®]
On-Prem

Laufende Projekte



Das Landesarchiv Berlin

Einwohnermeldekartei Berlin

~3,6 Millionen Karteikarten

Zeitraum: mehr als 50 Jahre (erste Hälfte des 20. Jahrhunderts)

Handschriftliche Texterkennung

Layoutanalyse für **Formulare**

Web-Interface für rasche Korrekturen

Veröffentlichung der Ergebnisse über **read&search**

Männl. | Weibl.

Eltern:

Eltern des Ehegatten: M. _____

Ehemann: _____

Staatsangehörigkeit: _____

Beruf: _____

Letzter Dienstgrad bei der Wehrmacht: _____

Lfd. Nr.	Vornamen Rufnamen unterstreichen Ehefrau mit Geburtsnamen	Fam.-stand	Geburts-			Rel.	Aus-zugs-vern.	Personenausweis:
			tag	mo-nat	jahr			

... auf Hausnr. 1948

zeit.punkt – Zeitungsportal

Zeitungsportal (Nordrhein-Westfalen)

15 Millionen Zeitungsblätter

Universitäts- und Landesbibliothek
Bonn

Verwendung des **NewsEye-Modells**
(EU-Projekt koordiniert von der
Universität La Rochelle)

Texterkennung

Automatische experimentelle
Artikel-Separierung (AAS)



Das Staatsarchiv Amsterdam - VeleHanden

“Crowd bringt dem Computer das Lesen bei”

Citizen-Science-Plattform **VeleHanden**

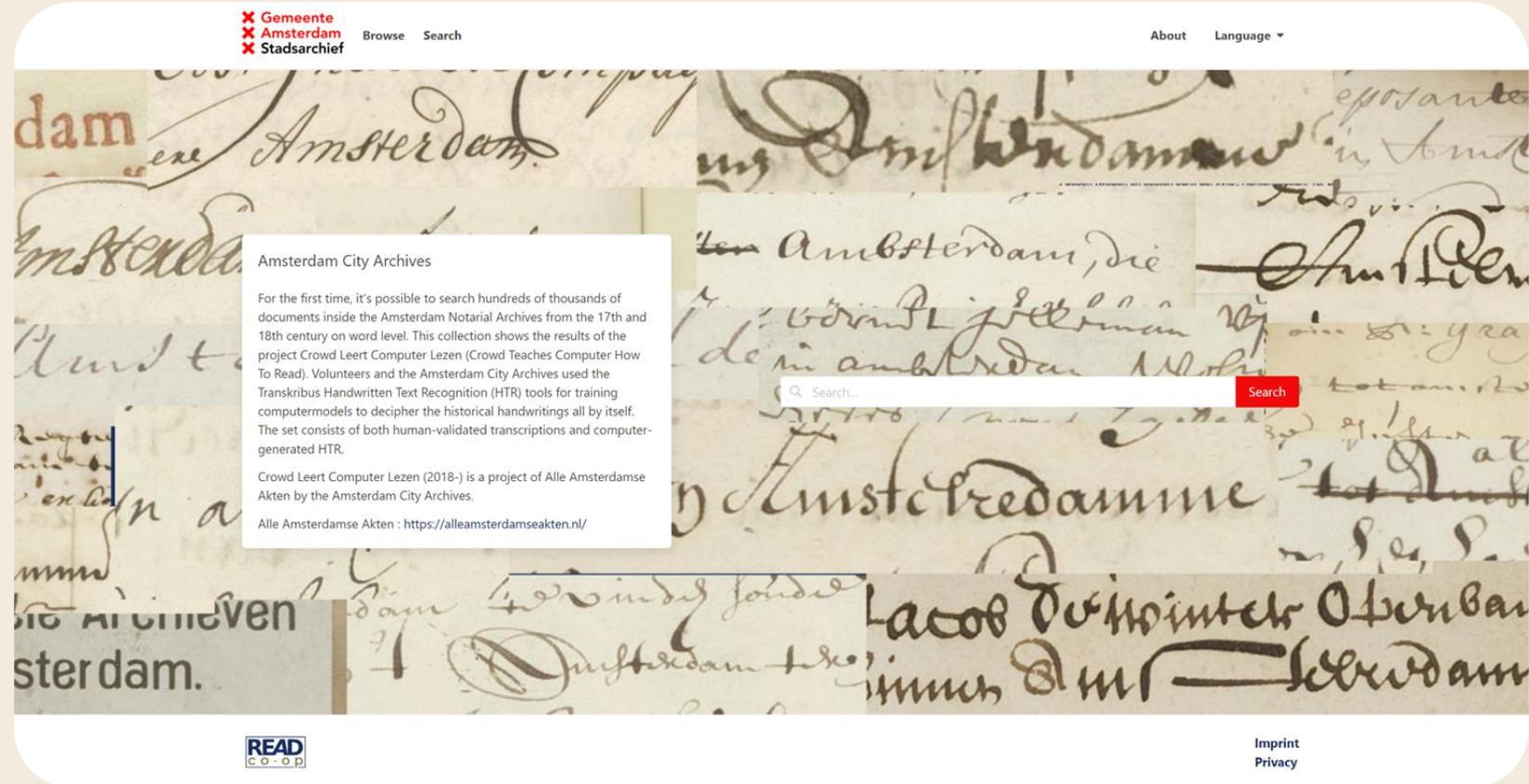
Geplant: 4.000 Seiten

Aktuell: **42.000 Seiten**

von Freiwilligen mithilfe der
Transkribus-Plattform transkribiert

300.000 Seiten über **read&search**
zugänglich gemacht

Interessante Geschichte:
Ein **neues Rembrandt-Gemälde** wurde
entdeckt



Gemeente Amsterdam Stadsarchief Browse Search About Language

Amsterdam City Archives

For the first time, it's possible to search hundreds of thousands of documents inside the Amsterdam Notarial Archives from the 17th and 18th century on word level. This collection shows the results of the project Crowd Leert Computer Lezen (Crowd Teaches Computer How To Read). Volunteers and the Amsterdam City Archives used the Transkribus Handwritten Text Recognition (HTR) tools for training computer models to decipher the historical handwritings all by itself. The set consists of both human-validated transcriptions and computer-generated HTR.

Crowd Leert Computer Lezen (2018-) is a project of Alle Amsterdamse Akten by the Amsterdam City Archives.

Alle Amsterdamse Akten : <https://alleamsterdamseakten.nl/>

SEARCH

READ co-op Imprint Privacy

**Gemeente
Amsterdam
Stadsarchief**

VeleHanden.nl

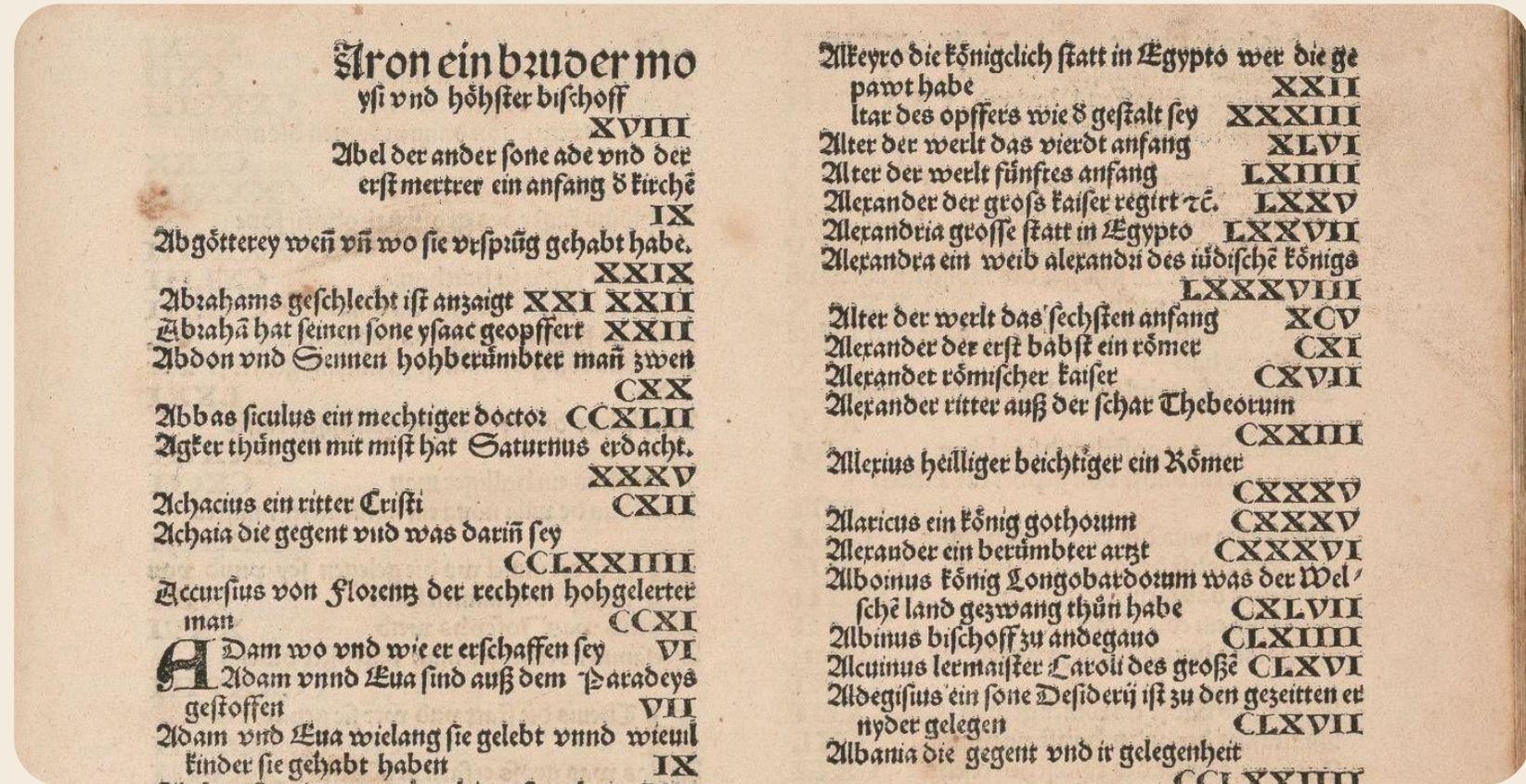
Bibliotheca Hertziana/Max-Planck Institut für Kunstgeschichte

Humanities Connect

read&search-Website

Texterkennung und Strukturanalyse

Änderung der **TEI Export-Funktion** + **Tagging, doppelte Erkennung von Abkürzungen** (abgekürzt und erweitert)



KaTi: Universität Innsbruck / Land Tirol

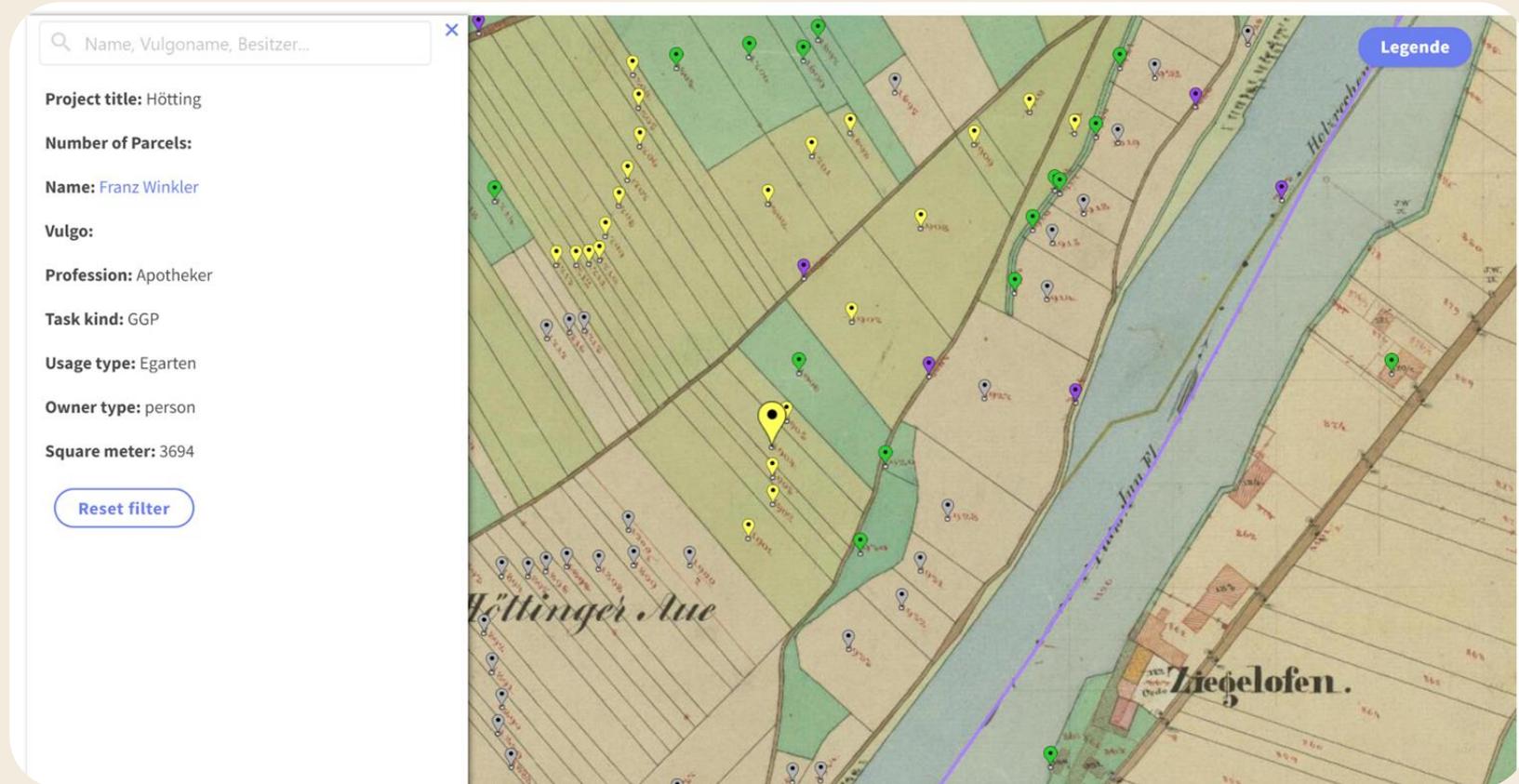
Citizen-Science / Crowdsourcing

Erster österreichischer **Kataster**
in Tirol - **1860-1865**

Citizen Scientists arbeiten mit
Grundbüchern (Tabellen) und
Karten

mehr als 500.000 Einträge
bereits in der Datenbank (=ca.
30%)

Ziel – Dashboard, das der
Öffentlichkeit die **Suche und**
Visualisierung aller Daten
ermöglicht



EnrichEuropeana+

**Unterstützung der Crowd-
Arbeit für das europäische
Kulturerbe**

Weiterentwicklung der
Transcribathon-Crowd-Plattform

Transkriptions-Editor

Zugang zu HTR über die
metagrapho API



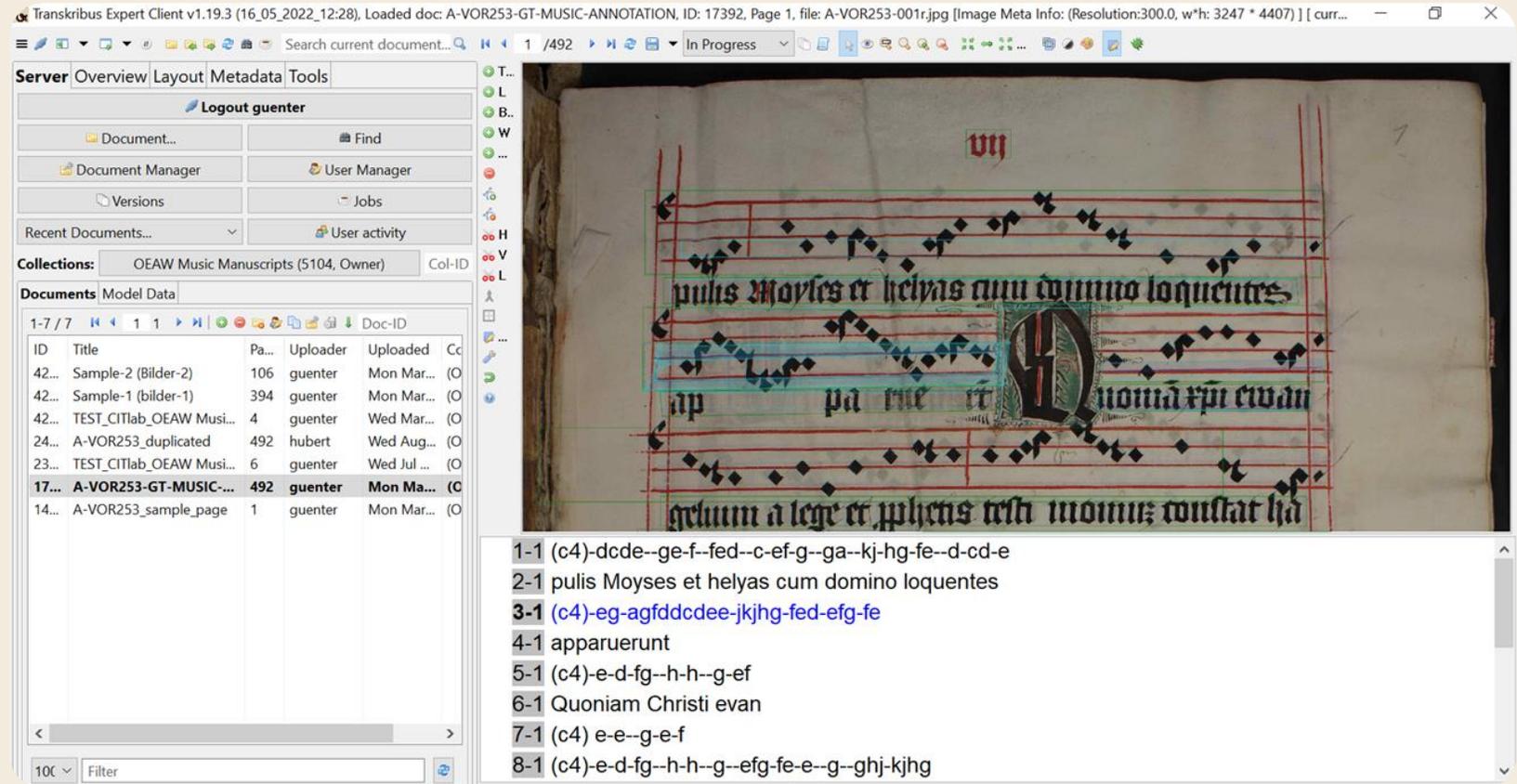
Österreichische Akademie der Wissenschaften + READ-COOP

Optical Music Recognition

Entwicklung von Musikererkennung
für mittelalterliche Dokumente

wird **Open Source** gestellt

Zusammenarbeit mit der
Universität von Alicante (Jorge
Calvo)



Transkribus Expert Client v1.19.3 (16_05_2022_12:28). Loaded doc: A-VOR253-GT-MUSIC-ANNOTATION, ID: 17392, Page 1, file: A-VOR253-001r.jpg [Image Meta Info: (Resolution:300.0, w'h: 3247 * 4407)] [curr...]

Server: Overview | Layout | Metadata | Tools

Logout guenter

Document... Find

Document Manager User Manager

Versions Jobs

Recent Documents... User activity

Collections: OEAU Music Manuscripts (5104, Owner) Col-ID

Documents | Model Data

ID	Title	Pa...	Uploader	Uploaded	Cc
42...	Sample-2 (Bilder-2)	106	guenter	Mon Mar...	(O)
42...	Sample-1 (bilder-1)	394	guenter	Mon Mar...	(O)
42...	TEST_CITlab_OEAU Musi...	4	guenter	Wed Mar...	(O)
24...	A-VOR253_duplicated	492	hubert	Wed Aug...	(O)
23...	TEST_CITlab_OEAU Musi...	6	guenter	Wed Jul ...	(O)
17...	A-VOR253-GT-MUSIC-...	492	guenter	Mon Ma...	(O)
14...	A-VOR253_sample_page	1	guenter	Mon Mar...	(O)

1-7 / 7

100 Filter

1 / 492 In Progress

W

pulis Moyses et helyas cum domino loquentes

ap pa tue et Quonia xpi ewan

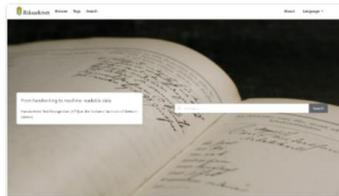
gelum a lege et pphens testi monuz constat ha

- 1-1 (c4)-dcde--ge-f--fed--c-ef-g--ga--kj-hg-fe--d-cd-e
- 2-1 pulis Moyses et helyas cum domino loquentes
- 3-1 (c4)-eg-agfdccdee-jkjhg-fed-efg-fe
- 4-1 apparuerunt
- 5-1 (c4)-e-d-fg--h-h--g-ef
- 6-1 Quoniam Christi evan
- 7-1 (c4) e-e--g-e-f
- 8-1 (c4)-e-d-fg--h-h--g--efg-fe-e--g--ghj-kjhg

Viele, viele weitere Projekte



Amsterdam City Archives



National Archives of Sweden



Recesses of Low German cities



Czech and Slovak Fracture
(Schwabacher)



Grenzregelungsausschuss



Mining Hub



Marjory Fleming



Club Andino Bariloche



Noscemus



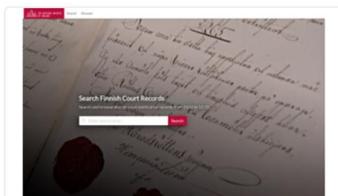
Mountain Hub



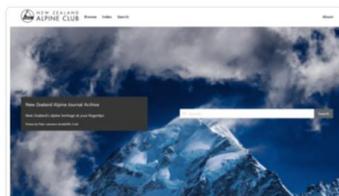
Regesta Pomeraniae Monastica



Deutsche Ratsprotokolle 17. bis
19. Jh. (Bautzen)



National Archives of Finland



New Zealand Alpine Journal



iurisprudētia – Uni Zürich



Tiroler Grundbuchblätter



Historischer Roman



Transkriberade handskrifter

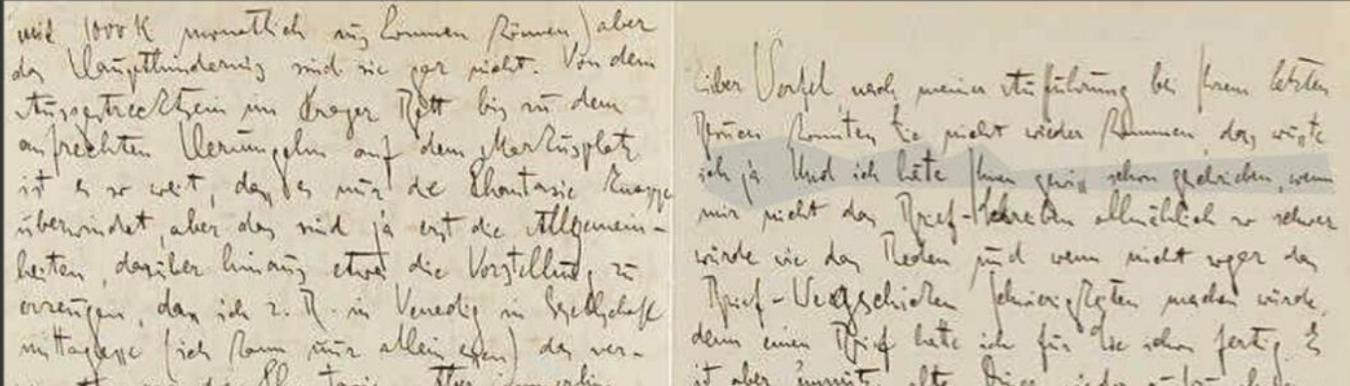


The Next Generation Of

Transkribus®



Literatur > Kafka



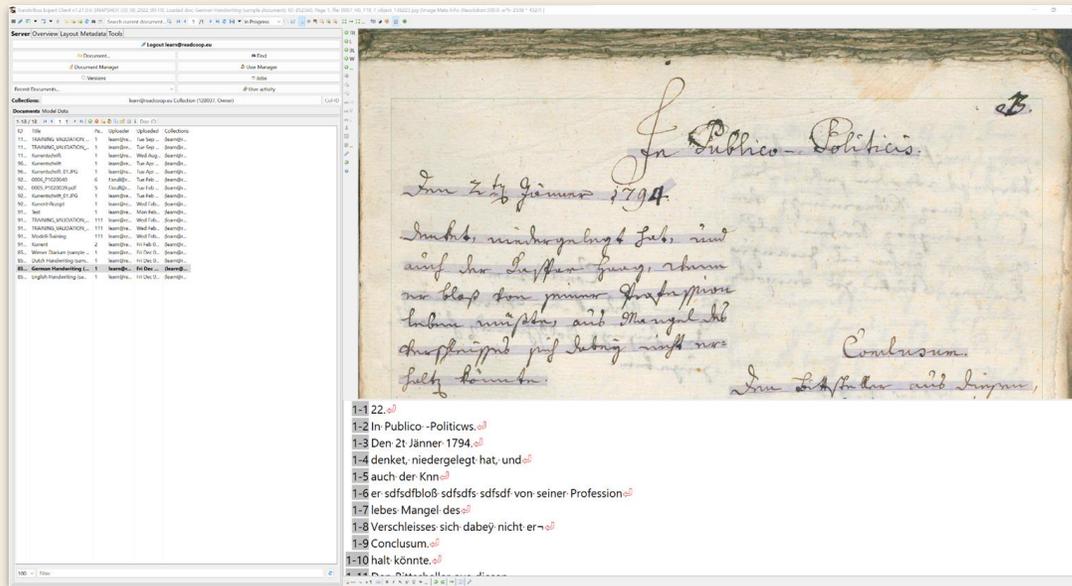
sz uurictlich aun konnen sönnen, als
nit
des Caupthinkerug mich sie ger nicht. Von den

SEARCH
comment
type name
number

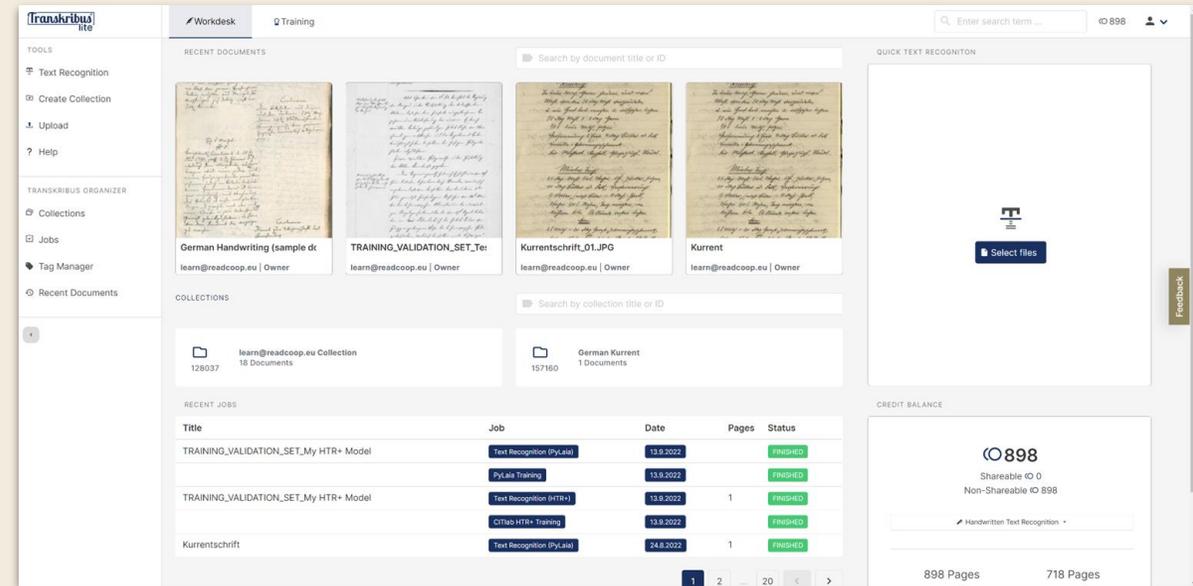
Navigation icons: +, -, ↺, ↻, ↵, ←, →, A+, A+, ☰

Moving Transkribus to the web

So we went from

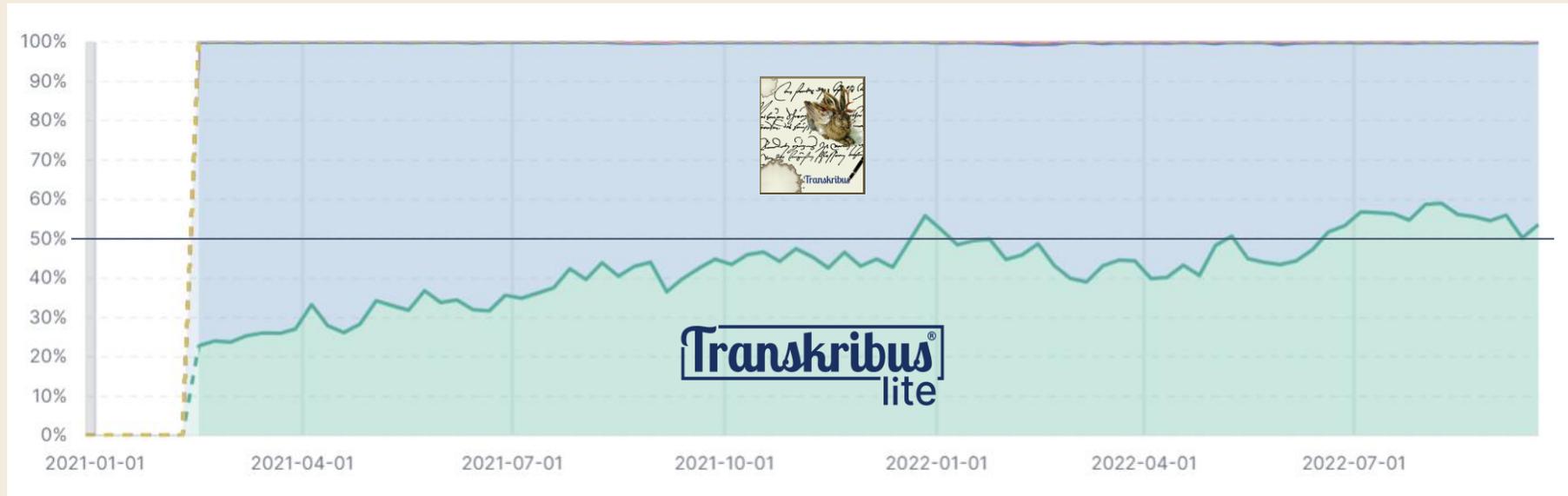


From this



To this

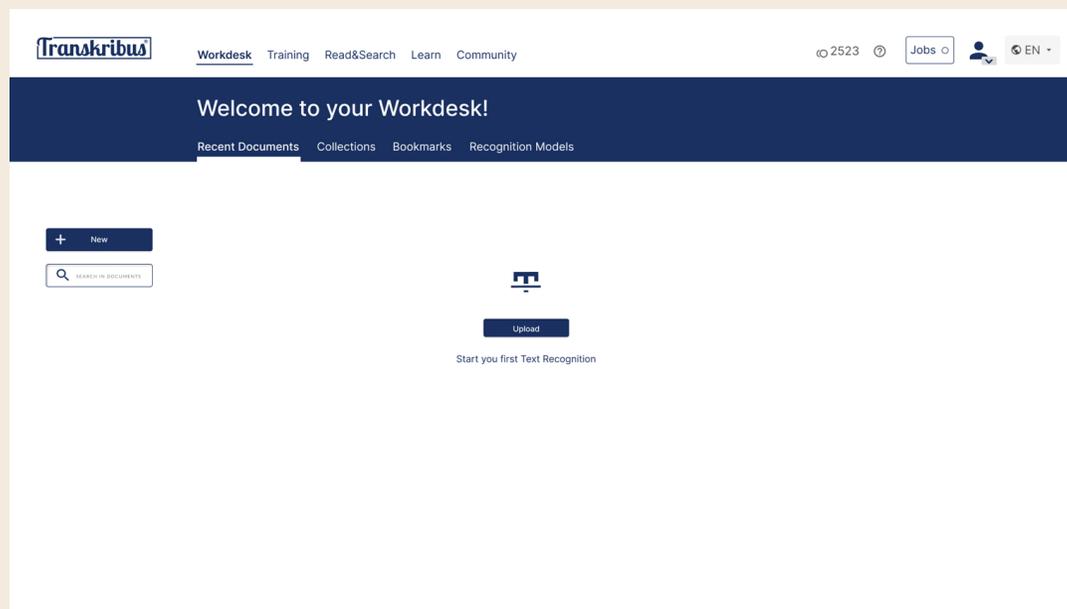
Increasing importance of our web application



More than 50% of Transkribus
Users use **Transkribus** Lite

It's just

Transkribus®



beta.transkribus.eu

Making Transkribus simple *and* powerful

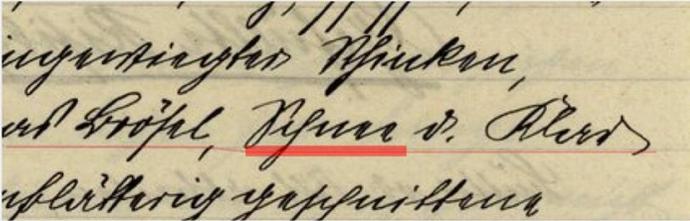
Transkribus Learn

← Back to collection

Cooking book (Sütterlin), 1926

Page 1

< 1 / 5 >



Text Text Text

Check word 

Score: 0/5

Learn how to read old handwriting with Transkribus

- Set and edit tasks
- Choose words or select randomly
- Statistics for tasks
- Learn groups

read&search CMS

The screenshot displays the read&search CMS interface. On the left, a handwritten document is shown with a search bar at the top containing the text "In Publico- Politicis." and a page number "5." The document text is in cursive and includes the date "Actum Bozen den 5 ten Jänner 1792." and lists several names: "Herr Oberlieut. Karl v. Kuglern, Herr Dr. Leopold Spielmann, Herr Franz v. Mayrl, und Herr Dr. Andree di Pauli." The right side of the interface shows the transcription of the document, starting with "5 In Publico- Politicis. 74 Actum Bozen den 5 ten Jänner 1792. Vor Titl Herrn Bürgermeister Jos. v. Remich et. In Beyseyn der Magistrats Rätthe: Herr Oberlieut. Karl v. Kuglern, Herr Dr. Leopold Spielmann, Herr Franz v. Mayrl, und Herr Dr. Andree di Pauli. Titl Herr Bürgermeister V Remich. No 5. Kreisämt. Reskript d. d. 31 ten Xbr 1791, praes. 3 ten Jänner d. J. daß die von der Gemeinde Lei= fers angebrachte Bitte um An= stellung eines andern Schullehrers anstatt des Jos. Obwurzer der malen nicht Platzhabe, übrige= gens aber derselbe sich vor dem Kreisamte sich persönlich zustellen. verhalten, und zugleich die Anzei= gen erstattet werden solle, ob er von der hohen Landesstelleschon Conclusu. das Bestätigungsdekret empfangen 66 Ist dem Anwald zu Leifers"

Publish your material online with read&search

- Lightweight CMS System
- Publish your Transkribus Collections
- Powerful search features

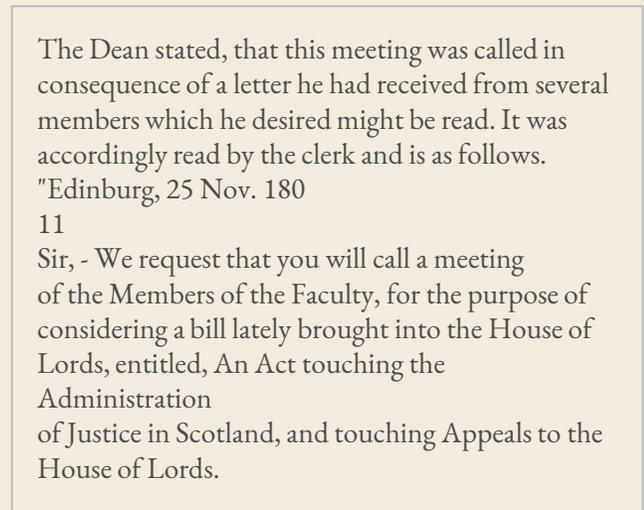
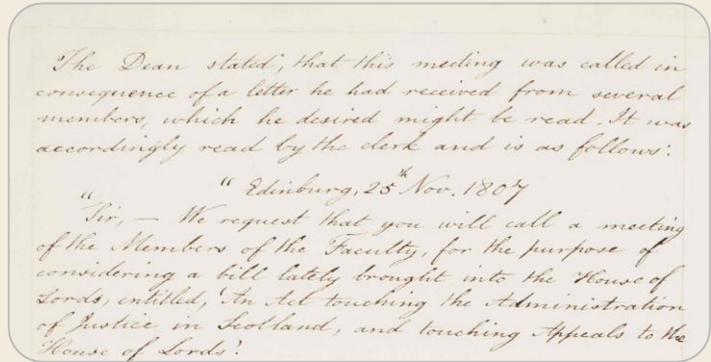
Transkribus[®]
metagrapho <api>

Processing API for text recognition

Processing API



- First use case: EnrichEuropeana+
 - Integration of Transkribus text recognition in the Transcribathon platform
- Simplifies HTR processing for single images
 - “Image in, layout+text out”
 - No persistent storage of image data
 - No document/collection management
- Technical documentation according to OpenApi Specification v3 (Swagger UI: <https://transkribus.eu/processing>)
- Many text recognition models available
 - Your private models trained in Transkribus
 - All public models <https://readcoop.eu/transkribus/public-models/>
- Already featured in various applications
 - <https://transkribus.ai>
 - CCS docWizz
 - Goobi
 - Europeana Transcribathon

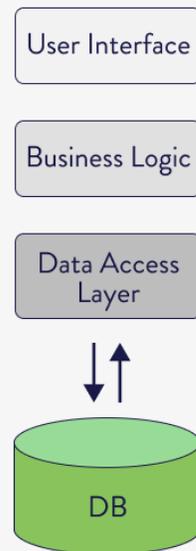


Transkribus[®]
On-Prem

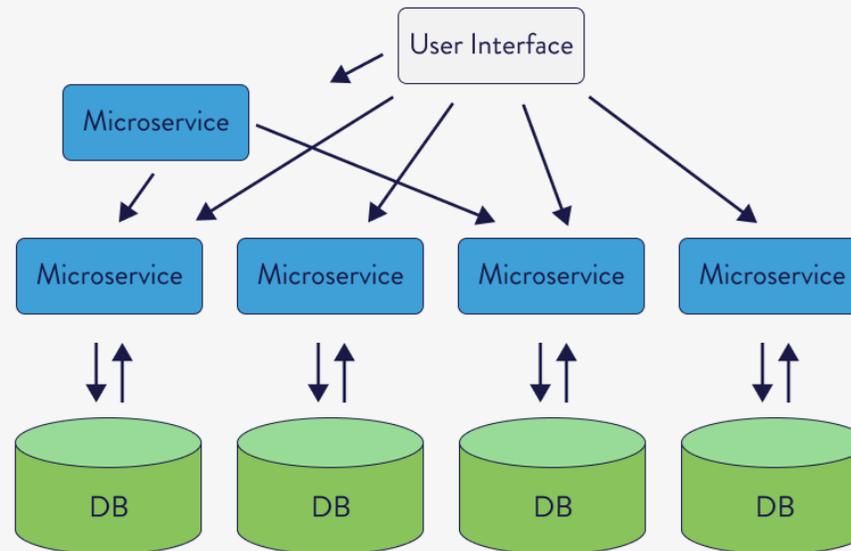
Current work on architecture: Microservices

Monolith and Microservice Architecture Difference

MONOLITH ARCHITECTURE

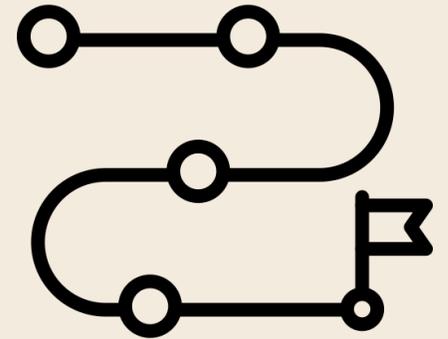


MICROSERVICES ARCHITECTURE

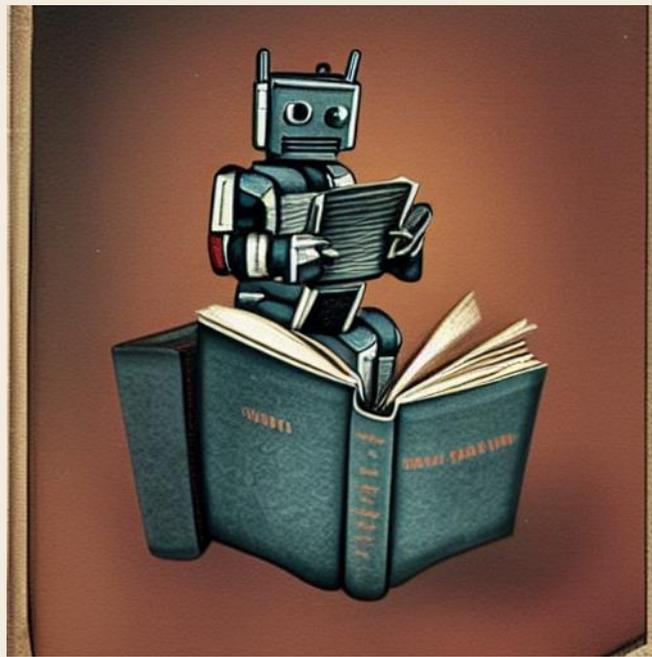


Rough Roadmap

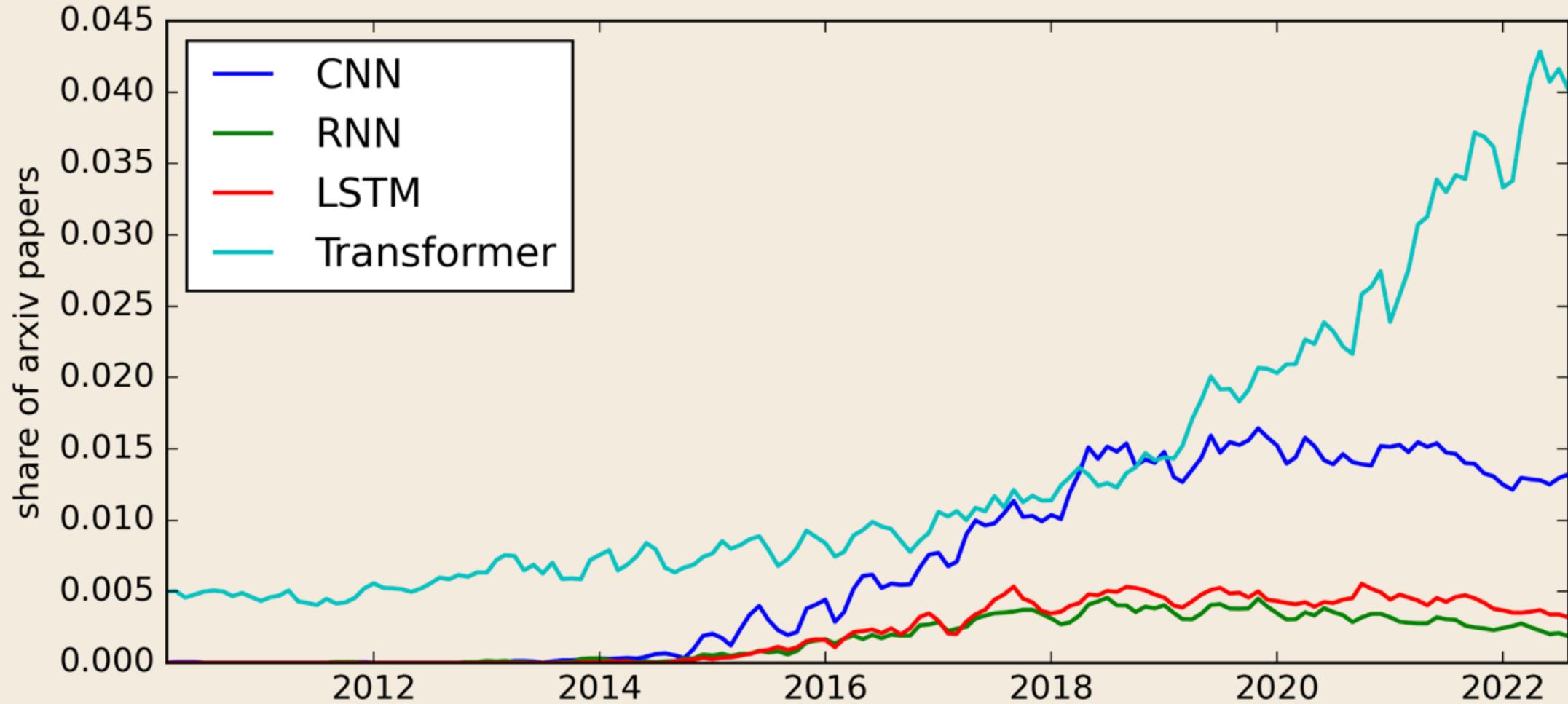
- First phase (completed)
 - Client implementation as production-grade collection of scripts + local java GUI
- Second phase (ongoing, Q4 2022/Q1 2023)
 - Docker images for core services
 - Install in Kubernetes environment
- Third phase (sometime during 2023):
 - Integration of additional GUIs: Transkribus (web version) and r&s



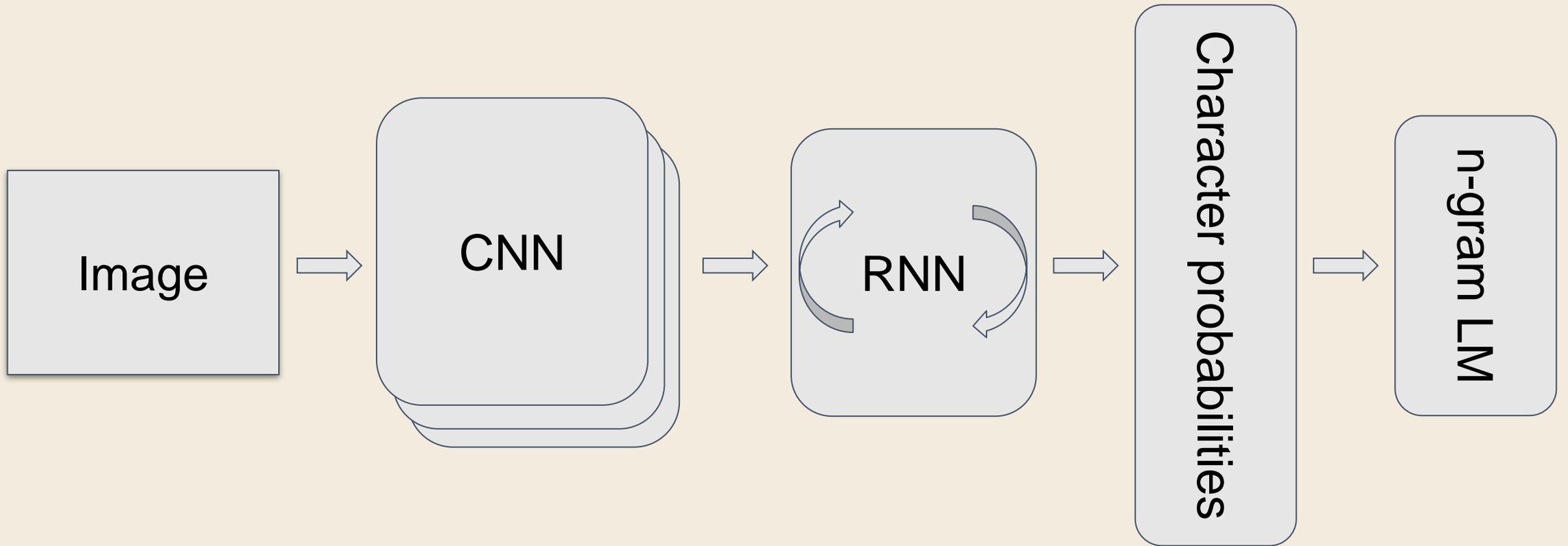
Transformer Models



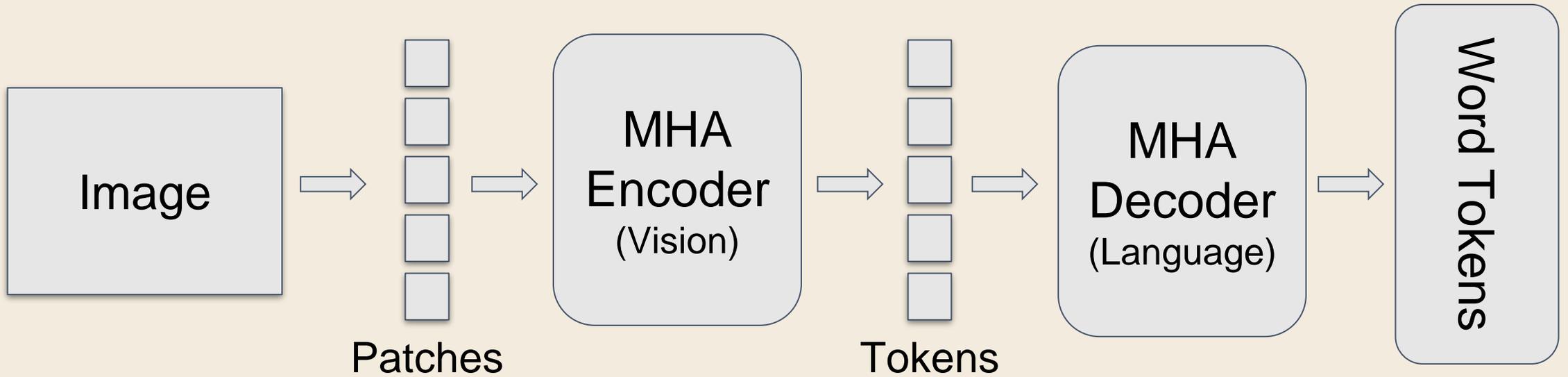
State of research



Classic HTR architecture (e.g. PyLaia)



Transformer HTR



First trHTR model now available in Transkribus!

- One big general-purpose model
- Trained on gigantic dataset
- 6+ Languages
 - English, French, German, Finnish, Swedish, Dutch
- Ideal for trying out recognition or creating GT
- Rollout has started

Trainable Layout Analysis & Tables & Forms Recognition

ORDE NUMMER.	Dagteekening VAN ONTVANGST.	INZENDER.	Dagteekening EN NUMMER VAN HET INGEZONDEN STUK.	DIETAL BIJLA- GEN.	ONDERWERP.
111.	2 feb.	Kerkvoorsden der Herv. Geme.	2 feb.		betreft de dank voor de verleende toestemming tot verplaatsing der gewichten van het terrein verreken der Grootse Kerk en verzoeken goedkeuring op de door hen medegeleverde maken welke geroefte regeling ook ten aanzien van den tijd dat het terrein dat moeten aankomen, gedurende de uitvoering van het werk.

Dagteekening der besturen.		UITVOERING.	Verzending van stukken.		Verrij- ding naar verige of volgende nummers van het register.	Aanmerkingen.
VAN BURO's EN WETH.	VAN DEN RAAD.		AAN WIEN.	WANNEER.		
		De handen van Geme. Raad 2 feb. 1891.			34	van 1891
					55	" "

Trainable LA + Tables & Forms Recognition

- New Layout Analysis (LA) coming to Transkribus
 - Not general purpose, but trained on certain collections / documents
- Based on the open-source Detectron2 framework
- Allows for different kinds of segmentation
 - Semantic
 - Instance



Detectron2

Form segmentation

- Trainable LA based on existing PageXML's
- Neighboring instances can be segmented

21831	Halmis	T
2		
3		
4		
5	Pole 7. d	U
6		
7	Rum 7. d	W
8	Pole	
9	Rum	
21840	Pole	



NUMMER welke ieder STUK Ontvangt.	DATUM der Registratie.	NUMMER der ontvangene STUKKEN.	AANWIJZING van het OBJEKT van IEDER STUK.	Autoriteit of Particulier die de ontvangene Stukken heeft geschreven of aan wien men schrijft.	ZAKELIJKE INHOUD
					VAN IEDER STUK.
11	1811	111	Handwritten description	Handwritten authority	Handwritten content
12	1811	112	Handwritten description	Handwritten authority	Handwritten content
13	1811	113	Handwritten description	Handwritten authority	Handwritten content
14	1811	114	Handwritten description	Handwritten authority	Handwritten content
15	1811	115	Handwritten description	Handwritten authority	Handwritten content

Aan wien de zaak is Gedemandeerd.	DISPOSITIE.	Bemerkingen.
	INHOUD.	
Handwritten name	Handwritten disposition	

...giovani
...Cappuccini, e
...S. Barbara
...1763
...ti suddetti Carlo
...Bassano
...S. Maria
...condizione
...Cappuccini
...S. Maria
...S. Antonio

...infraz tutto confermo
...Caracci e Cape

...et
...in fidei
...manus
...in mal

...fidei
...de
...die

...Joseph
...Caraccano

Barbara di Pietro
...Con solenne
...solenne, e speciale
...in casa propria della
...stato, con
...a
...Parola
...valuta di Milano
...alla vota
...per
...ca il suo
...e titolo del

- 1.°
- 2.°
- 3.° Che entro il termine
a loro spesa in
- 4.° La Spese del
Sub. Conomator

Antonij de Aleb
...habitat
...erit
...Vxon
...De Martinio
...Borghij
...quod
...ad tertius
...tanti
...per
...ad
...Tusino
...tanta
...Antonio Aleb
...Stada
...Antony Vergo
...et
...Francij
...Cectio: ab ipso erecto

Unlock our written past.

