

Introduction to Handwritten Text Recognition

u^b

b
UNIVERSITÄT
BERN

Tobias Hodel (Walter Benjamin Kolleg, Digital Humanities)

Michigan, 13th May 2026

Bild: Stable Diffusion, Prompt:
«A deep neural network
in the style of a Dali painting»



Goals

We want to...

- get an Introduction to HTR/ATR
- Try out an Integrated Transcription Environment (ITE)
- Get to Know the Technology
- Compare it to Visual Language Models
- Discuss Best-Practices, Issues, etc.
- Develop our own approaches
- Get Access to Resources

Overview (Part I)

- 09:45 ATR: Where are we at the moment?
- 10:10 Round of Introductions: ATR — Where are you?
- 10:25 Key Concepts & Epistemological Decisions
- 10:40 Integrated Transcription Environments
- 10:55 How to use eScriptorium
- 11:20 eScriptorium vs. Transkribus
- 11:25 Business models, open science, and HTR-United

Overview (Part II)

- 14:15 Test your own material
- 14:45 Show your results
- 15:05 Introduction to VLMs and command-line tools
- 15:25 Developing Workflows and Identifying Issues
- 15:45 Open Questions and Inputs on Demand

Preparation and Resources

You will need:

- Access to escriptorium.flow-project.net
 - User account: mail address
 - Password: Kzoo2026
- Documentation for today: <https://thodel.github.io/atr/>
- Slides: XXX

MICHAEL SCHONHARDT,
TIM GEELHAAR, TOBIAS HODEL,
JAN ODSTRČILÍK
.....
**AUTOMATED TEXT
RECOGNITION**
.....
**THEORY, PLATFORMS,
BEST-PRACTICES**

DIGITAL HUMANITIES RESEARCH
UNIVERSITY
BIELEFELD PRESS

A very short overview: From Optical Character Recognition to (Automatic/Handwritten) Text Recognition

1990ies: Commercial Optical Character Recognition

2000thes: Recognition of «Fraktur» with OCR
(EU Project IMPACT, ABBYY XIX)

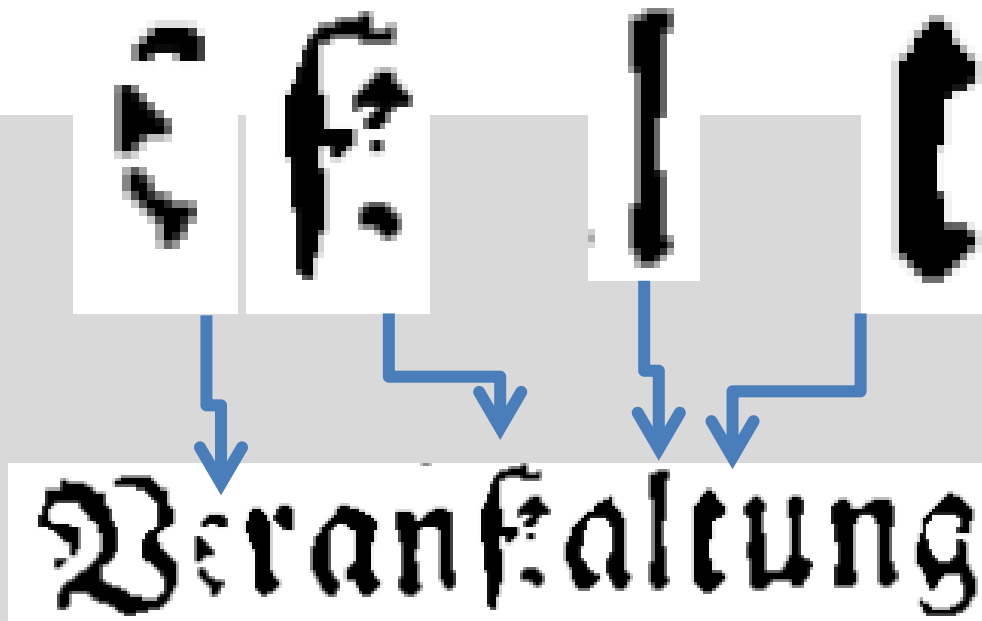
2010ns: Simple Language Model / Introduction of Neural Networks
(to the field)

After 2016: Dominance of *deep learning*
(high-dimensional language model)

2021: Transformer-based Models for text recognition (trOCR)

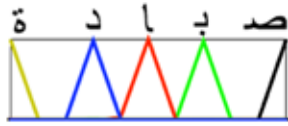
2024: Visual/Multimodal Language Models, based on LLMs

Automated Text Recognition?

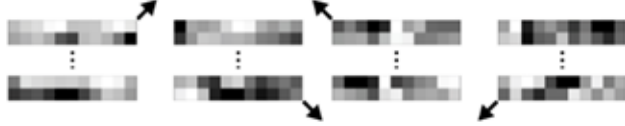


Automated Text Recognition

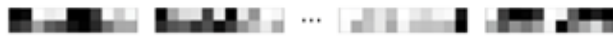
- Machine learning using neural networks
- Processes writing by line, rather than by character
- Needs to be trained by being shown document images and transcripts
- More training data → more accurate recognition
- Create a model to transcribe and search a collection of documents



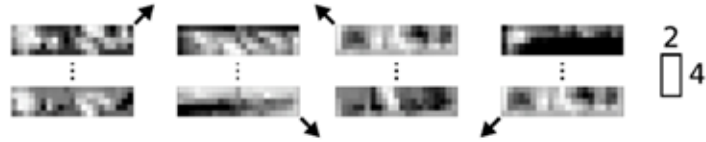
Output
121 x CTC



MDLSTM
4 x 50 cells



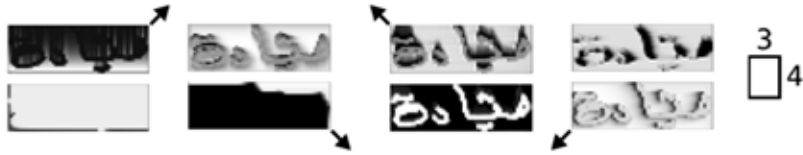
Feedforward
20 x *tanh*



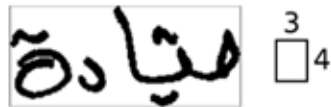
MDLSTM
4 x 10 cells



Feedforward
6 x *tanh*



MDLSTM
4 x 2 cells



Input

u^b

^b
UNIVERSITÄT
BERN

Carolingian Minuscles

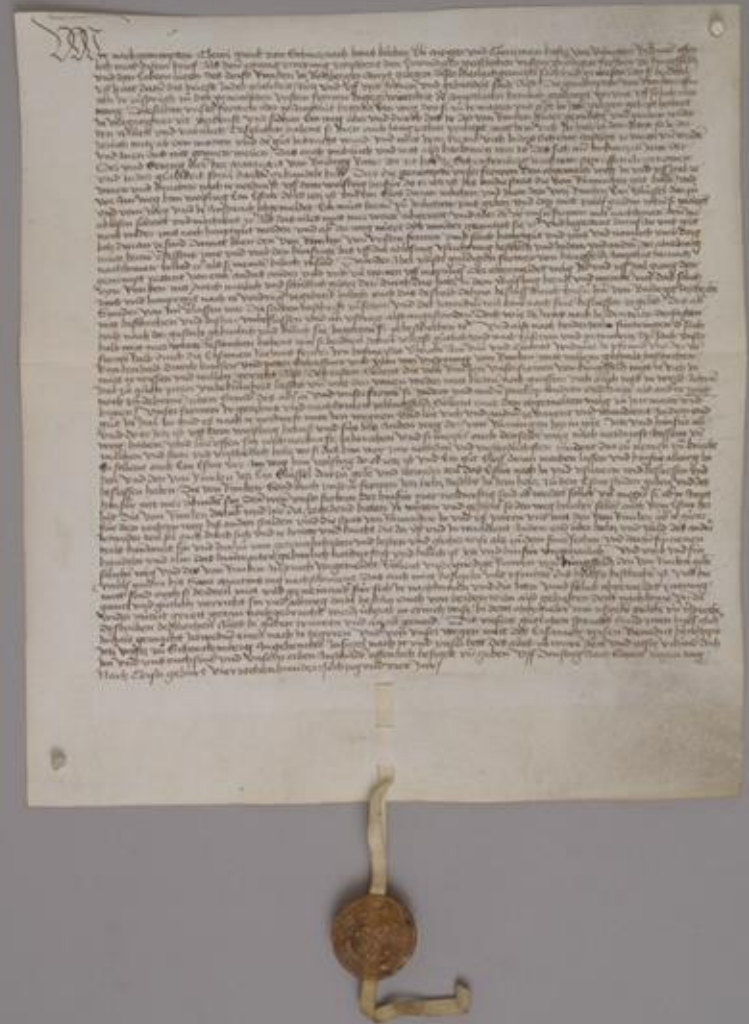
5'000 words, 1'000 lines:

Validation pages: 7% CER

24
illem: que p. fec. deo fo. mit. at
tino rem per quec um uer se que
prius non sine formidine: obser
uabec absq. ullo labore uel ut
nec curcliter & consudine
in capi & custodire: Non icem tino
re gehennę: sed amore xpi: & con
sudine ipse bonę & delecta
tione uirtutum: quae dñs icem
in operetum suum mundum
auertit & peccatis spūs scō dignabitur
demonstrare: **De officio diuini uiri**
Hiemis tempore **in hoc uis**
id est: a kalendis nō uembris
usque in pascha iuxta considerat
tionem rectionis oē uahore noc
tis surgen dum est: ut mo dice
amplius de media nocte peccat &

Charter 15th century

77'000 words, 3'500 lines:
Validation pages: 4,8% CER



Council Minutes 19th Century

148'000 words, 31'000 lines:
2,5% CER

4. Januar 1851. 1

Actum Samstag den 4. Januar 1851.
In Gegenwart des gesammten Regierungsraths.

Am 30. d. M. wurde im Rath der Bescheid über die Anträge der Bauunternehmer, die einen Zuschuss zur Herstellung der öffentlichen Gebäude beantragt haben, beschlossen. Die Bauunternehmer sind verpflichtet, die Kosten der Herstellung der öffentlichen Gebäude zu tragen. Der Regierungsrath hat beschlossen, die Bauunternehmer zu verpflichten, die Kosten der Herstellung der öffentlichen Gebäude zu tragen.

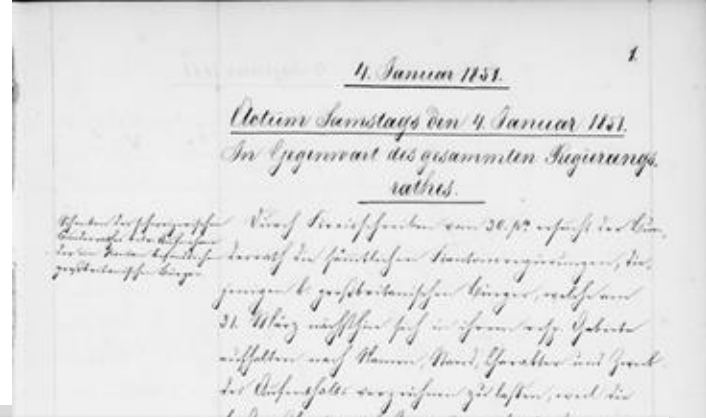
Am 31. d. M. wurde im Rath der Bescheid über die Anträge der Bauunternehmer, die einen Zuschuss zur Herstellung der öffentlichen Gebäude beantragt haben, beschlossen. Die Bauunternehmer sind verpflichtet, die Kosten der Herstellung der öffentlichen Gebäude zu tragen. Der Regierungsrath hat beschlossen, die Bauunternehmer zu verpflichten, die Kosten der Herstellung der öffentlichen Gebäude zu tragen.

Am 1. d. M. wurde im Rath der Bescheid über die Anträge der Bauunternehmer, die einen Zuschuss zur Herstellung der öffentlichen Gebäude beantragt haben, beschlossen. Die Bauunternehmer sind verpflichtet, die Kosten der Herstellung der öffentlichen Gebäude zu tragen. Der Regierungsrath hat beschlossen, die Bauunternehmer zu verpflichten, die Kosten der Herstellung der öffentlichen Gebäude zu tragen.

Am 2. d. M. wurde im Rath der Bescheid über die Anträge der Bauunternehmer, die einen Zuschuss zur Herstellung der öffentlichen Gebäude beantragt haben, beschlossen. Die Bauunternehmer sind verpflichtet, die Kosten der Herstellung der öffentlichen Gebäude zu tragen. Der Regierungsrath hat beschlossen, die Bauunternehmer zu verpflichten, die Kosten der Herstellung der öffentlichen Gebäude zu tragen.

Am 3. d. M. wurde im Rath der Bescheid über die Anträge der Bauunternehmer, die einen Zuschuss zur Herstellung der öffentlichen Gebäude beantragt haben, beschlossen. Die Bauunternehmer sind verpflichtet, die Kosten der Herstellung der öffentlichen Gebäude zu tragen. Der Regierungsrath hat beschlossen, die Bauunternehmer zu verpflichten, die Kosten der Herstellung der öffentlichen Gebäude zu tragen.

Generalizing of Writing Types Current/19th century



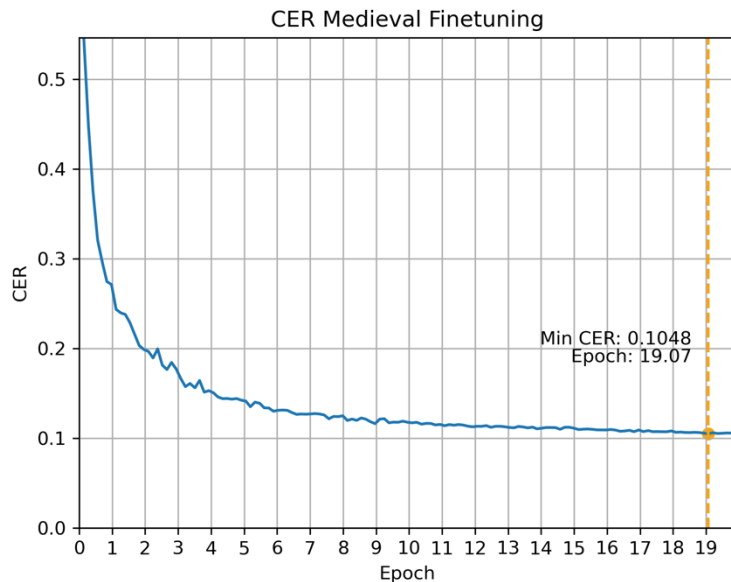
HTR MODEL	HTR ENGINE	CER MEAN %	CER MEDIAN %	CER UPPER BOUND (WORST)
German Kurrent M2	HTR+	3.43	2.76	9.13
	PyLaia	18.77	13.30	51.05
Transkribus German Kurrent	HTR+	5.90	4.85	10.20
RRB	HTR+	9.15	8.13	16.28



Generalizing of Writing Types

Medieval Scripts (TrOCR)

→ Transformer-based Models



What matters

Keep in mind...

- uniformity is key
- it's about your goals for a text/document
- there's not such a thing as a *perfect* transcription
- most machine learning algorithms start with random states; meaning: it doesn't hurt to re-run training processes

Transkribus, eScriptorium & Co.

Integrated Transcription Environments (ITE), Schonhardt et al. 2026, p. 37ff.

- Processing Platform (User, Login, Server Space)
- Sharing of Documents
- Creation of Training Material
- Run Training jobs (if infrastructure is available)

Visual Language Models (as an alternative)

Based on latent space vectorization as in Large Language Models

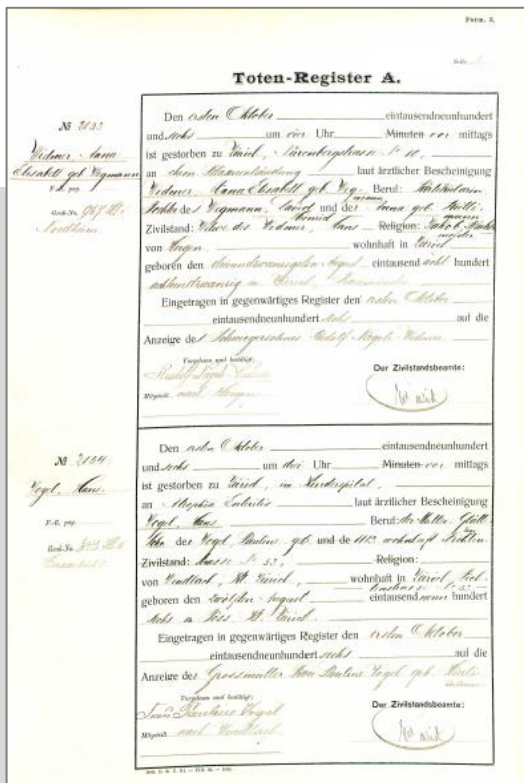
Run recognition process on entire pages

- no segmentation of layout
- no training of specific handwritings
- relying on large closed- / open-providers (Mistral, QWEN, Google)

Test out with Gemini (most successful for historical handwriting):

<https://gemini.google.com/app>

VLM: Text



Preprocessing (Segmenting in two entries) with YOLO v7 (specific model)

Text Recognition with QWEN 3 VL (8B)

→ for prints in GE after 1850 (handwritten and printed)

Information Extraction with GPT-OSS (120 B)

→ Segmenting in information pieces

VLM: Tables

13

1838 Spitz Aufstimm. 184077 72 Gesamtübersichten nach Gemeinden. 184077 72 Kanton Zürich.

Gemeinde Fueschlenen (Bevölg. 1992)				Gemeinde Fimach (Bevölg. 852)				Gemeinde Furlingen (Bevölg. 502)				Gf. Grossandelfingen (Bevölg. 855)				Zürcher Stad- t. Be- zirkel								
ge- mei- nde	ver- we- sen	Be- schäf- tigte	ge- mei- nde	ver- we- sen	Be- schäf- tigte	ge- mei- nde	ver- we- sen	Be- schäf- tigte	ge- mei- nde	ver- we- sen	Be- schäf- tigte	ge- mei- nde	ver- we- sen	Be- schäf- tigte										
in 1838	in 1840	in 1840	in 1838	in 1840	in 1840	in 1838	in 1840	in 1840	in 1838	in 1840	in 1840	in 1838	in 1840	in 1840										
140	37	103	266	145	411	355	57	198	272	325	507	99	23	76	229	908	437	168	36	142	235	176	501	
41.	4	37	64	42	106	163	14	149	193	198	291	56	-	56	66	79	145	90	-	90	114	137	251	
35	4	34	59	8	3	1	42	101	162	14	145	191	198	290	55	1	55	64	79	143	89	111	137	248
5	2	1	2	2	2	2	1	2	1	2	1	1	2	1	2	1	2	1	2	1	2	1	2	1
45	22	167	157	40	197	60	35	25	68	105	178	28	15	10	156	114	270	50	10	31	188	14	172	
5	-	5	11	4	15	8	3	5	8	8	8	8	2	1	4	4	10	-	10	22	4	4	28	
21	12	9	9	25	34	17	11	6	7	15	22	12	11	1	2	11	13	9	5	4	10	5	15	
12	7	5	69	24	69	25	16	9	33	33	10	4	4	6	74	5	80	18	0	9	76	1	77	
8	3	3	12	4	16	5	2	14	10	5	14	90	104	3	74	97	171	2	2	8	8	8	20	
4	4	4	56	7	68	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
35	31	24	53	77	20	4	16	3	17	20	9	5	7	2	13	15	19	2	17	16	34	40	0	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
7	7	10	6	22	7	7	8	3	11	5	3	3	5	1	6	5	1	4	34	1	4	34	34	0
6	4	9	5	4	9	5	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1																				

VLM: Table Recognition



Kategorie	A_col.1	A_col.2	A_col.3	A_col.4	A_col.5	A_col.6	A_col.7	A_col.8	A_col.9	A_col.10	A_col.11	A_col.12	B_col.1	B_col.2	B_col.3	B_col.4	B_col.5	B_col.6	B_col.7	B_col.8	B_col.9	B_col.10	B_col.11	B_col.12	B_col.13	B_col.14	B_col.15	B_col.16	B_col.17	B_col.18
Gesamtsum	154	52	102	503	183	386	457	140	317	993	529																			
darunter hi-																														
A. Gewinn_93	5	84	166	54	260		184	14	170	589	171																			
Aa. Bergba_2		2	1	1	1		1		1																					
Ab. Landw_91	5	82	165	54	259		180	14	166	282	171																			
Ac. Forstw_							3		3	107																				
B. Veredlur_44	36	8	25	74	99	174	99	75	435	627		76	52	24	31	68	99	133	96	37	152	243	595	93	63	30	65	135	200	
B.a. Herst_2		2	5	5	5	22	2	20	70	31		2	2	2	3	5	6		6	11	4	15	7			7	7	1	8	
darunter hi-																														
B.b. Herst_7	6	1	4	5	7	51	54	17	50	47		13	5	4	6	6	15	25	20	5	12	14	26	22	14	6	5	17	22	
darunter hi-																														
B.c. Herst_6	5	5	4	1	5	52	6	23	150	2		8	3	5	15		15	30	15	15	43	1	44	15	5	10	21	5	24	
darunter hi-																														
B.d. Herst_27	25	2	12	70	62	52	50	2	136	680		48	59	9	1	56	57	62	55	7	71	223	294	42	41	1	28	114	142	
darunter hi-																														
B.e. Herst_						1			7	1																				
B.f. Chem_										8																				
B.g. Bearb_5						11	5	6	18			4	1	5	5	5	9	5	4	15			15	4	5	5	2	4		4
darunter hi-																														
B.h. Vervi_						1			18	2																				
B.i. Wass_						1			1	6																				
B.k. Gewer_										6																				
C. Handl_14	4	10	10	14	24	76	15	61	69	90		11	1	10	5	12	21	36	10	26	50	44	74	25	6	19	10	5	19	
Ca. Lebend-						6	5	8		8		1	1		1	1	4	4					2	1	1	1	1		1	1
C.b. Rohpr_						2		2	4														2	2			1			1
C.c. Landw_3	1	2	6		6																									
C.d. Steir_						4	3	4	5	4		1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
C.e. Holzwa-						10	90	104																						
C.f. Buch_1																														
C.g. Spinnst-						7		7	7	6		1		1	1	1	5	1	1	4	6	7	13							
C.h. Ausstei-																														
C.i. Waren-																														
C.k. Bazars						2		2	4	3		5		3	3	2	5	9	4	2	5	2	8	10	6	1	8	2	4	6
C.l. Lebens_4	3	1		2	2	21	6	15	13	19												1	1	1	1	1	1	1	1	1
C.m. Geträ_						6	1	5	15																					
C.n. Tabak_						1		1	5	2																				
C.o. Drogw_						1		1	5	2																				
C.p. Abfälle																														
C.q. Tröddel-																														
C.r. Bankwe-									1	5	5																			
C.s. Versich-						1		1	2	5																				
C.t. Vermitt-																														
C.u. Gastw_7	7	4	12	16	21	21	11	52																						
C.v. Autom_																														
D. Verkehr_ triebe	100	triebe	1000	betri	lich	lich	Stimmen	100	triebe	1000	betri	lich	lich	Stimmen																
E. Offentl_1	5	100	500	100	200																									
Ea. Offentl_2	2				2	8	2	6	77	700																				
Eb. Rechts_1					1	15	10	5	23	51																				
Ec. Gesund-						1		1																						
Ed. Unterri-																														
Ee. Übrige_1	1				1	10	6	4	20	53																				
Ef. Künste						2		2	1																					
darunter hi-																														

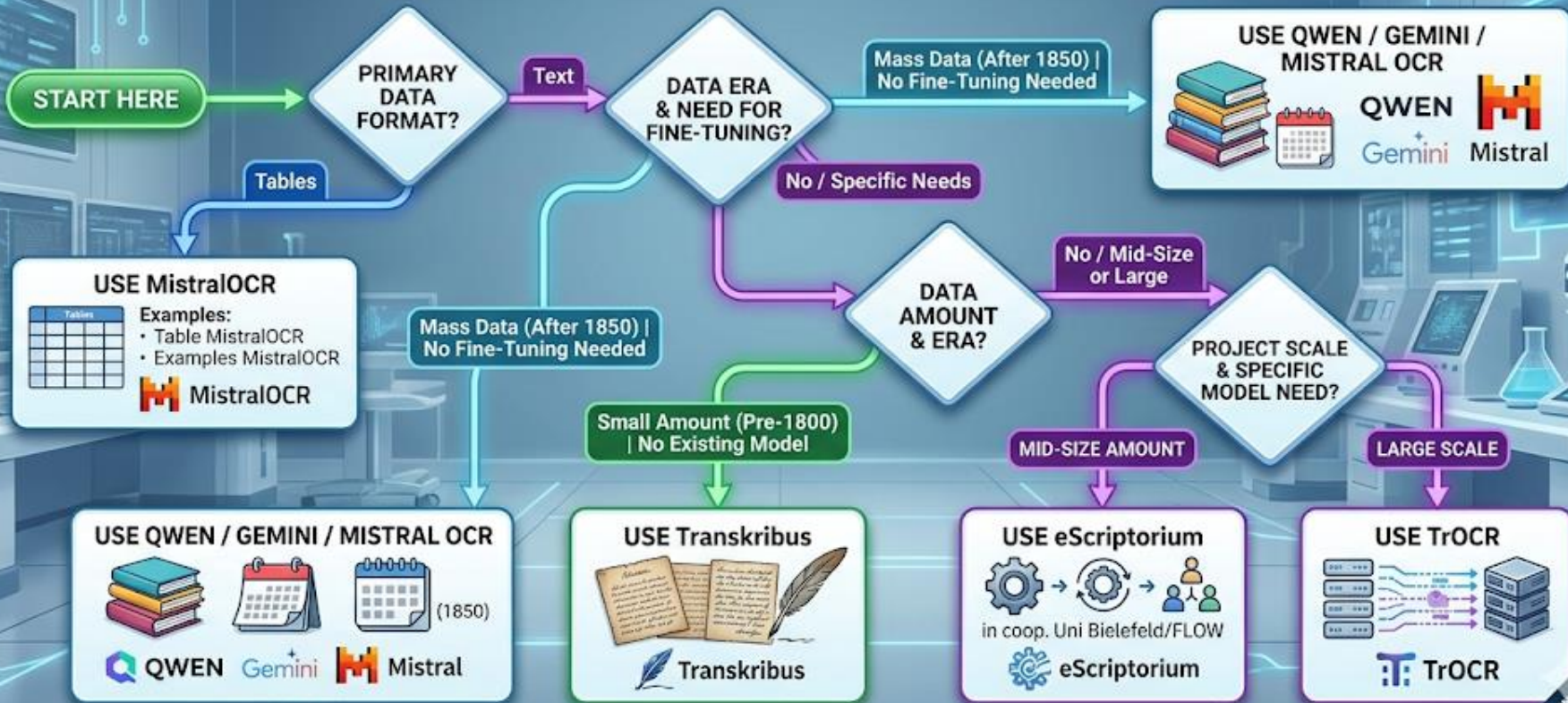
FARBLEGENDE:
 Gelb Zelle war leer/Null (0 eingefügt)
 Orange Vertikaler Summenfehler (Spaltensumme passt nicht)
 Lila Horizontaler Summenfehler (Zeilensumme passt nicht)

Automatic Text Recognition: Strategy

- Small Amount of Data without Model (pre-1800)
→ Transkribus
- Mid-Size Amount of Data with need for specific model
→ eScriptorium (in coop. Uni Bielefeld/FLOW)
- Large Scale projects with specific model
→ TrOCR
- Mass data after 1850, without need for fine-tuning →
QWEN/Gemini/Mistral OCR
- Tables → MistralOCR



DHLab TEXT RECOGNITION DECISION TREE | 2026



Round of Introduction

Where are you in regard to HTR

- Your name, background, pronouns
- Material you are working with
- Goals for this day
- Long-term application
- Experience so far

Key Concepts

... and why they matter

- Machine Learning
- Supervised Deep Learning
- Ground Truth
- Character Error Rate
- Word Error Rate
- Accuracy
- Character Set

Integrated Transcription Environments

Preparing, Training, Recognizing

- Graphical User Interface
- Set-up Once, Runs for Many
- Cooperate

Not necessarily

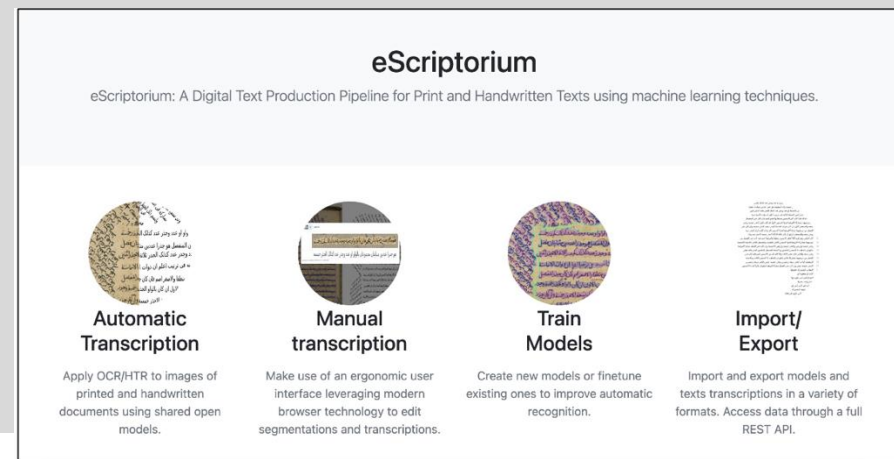
- Open Source
- Reusability of Models

The logo for Transkribus, featuring the word 'Transkribus' in a blue, serif font with a registered trademark symbol, enclosed in a blue rectangular border.

eScriptorium

An FOSS ITE

- Graphical User Interface
- Set-up Once, Runs for Many
- Cooperate
- Open Source
- Reusability of Models



eScriptorium

What is eScriptorium

- **Web-based tool** for transcription (no local installation necessary)
- **Free and Open Source** (FOSS, Code available on GitLab)
- **Developed by:** Université Paris Sciences et Lettres (PSL) as part of the projects *Scripta* and *RESILIENCE*.
- **Alternative to Transkribus:** Offers different features and is non-commercial.

eScriptorium

What is eScriptorium

- **Digital Text Production Pipeline:** Uses machine learning for print and handwritten texts.
- **Automatic Transcription:** Apply OCR/HTR to images using shared open models.
- **Manual Transcription:** Ergonomic user interface to edit segmentations and transcriptions.
- **Train Models:** Create new models or fine-tune existing ones to improve recognition.
- **Import/Export:** Support for various formats and access through a full REST API.

eScriptorium

What is eScriptorium

Compatible with (can be imported):

- Page-XML
- Alto-XML
- IIIF-Manifest
- PDF
- Images

```
1 <?xml version="1.0" encoding="UTF-8" standalone="yes"?>
2 <PcGts xmlns="http://schema.primaresearch.org/PAGE/gts/pagecontent/2019-07-15" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLo
3 <Metadata>
4   <Creator>escriptorium</Creator>
5   <Created>2024-01-22T15:21:00.961404+01:00</Created>
6   <LastChange>2024-01-24T12:23:49.148807+01:00</LastChange>
7 </Metadata>
8 <Page imageFilename="typewriter.jpeg" imageWidth="1400" imageHeight="1980">
9   <TextRegion id="eSc_textblock_da7af2e3" custom="structure {type:Main;}">
10    <Coords points="47,300 47,652 1231,652 1229,307"/>
11    <TextLine id="eSc_line_59a6603c" custom="structure {type:default;}">
12      <Coords points="92,204 92,216 1144,222 1145,204 1145,179 92,178 92,204"/>
13      <Baseline points="94,206 1147,206"/>
14      <TextEquiv>
15        <Unicode>Why Writing On A Typewriter Feels So Strangely Modern</Unicode>
16      </TextEquiv>
17    </TextLine>
18    <TextLine id="eSc_line_7d9fec4c" custom="structure {type:default;}">
19      <Coords points="92,237 92,250 236,260 255,252 279,260 374,252 399,262 562,262 574,252 648,260 733,252 744,260 787,252 815,260 846,252 870,
20      <Baseline points="93,239 1107,239"/>
21      <TextEquiv>
22        <Unicode>Years of voice dictation trained me to think before I write</Unicode>
23      </TextEquiv>
24    </TextLine>
```

eScriptorium

is a tool for

- **Optical Character Recognition (OCR) and Handwritten Text Recognition (HTR)**
- **Automatic segmentation** and transcription
- **Manual transcription:** Used, for example, to produce **Ground Truth** (material for training models)
- **Training and Fine-tuning:** Powered by the free-and-open-source **Kraken** OCR Python engine in the background
- **Data Management:** Import, export, and cooperation features

eScriptorium

Demo

The eScriptorium instance can be reached at:
<https://escriptorium.flow-project.net>

eScriptorium

Demo

The eScriptorium instance can be reached at:
<https://escriptorium.flow-project.net>

eScriptorium

Strengths (+)	Weaknesses (-)
Free and Open Source	Software development can be slow and buggy
Opportunity for further development	Web-based tools can provoke restrictions (e.g., browser compatibility)
Great line mask recognition thanks to Kraken	No tagging/abbreviation features (requires external tools like Inception)
Can be self-hosted/operated	Few segmentation models available; prone to errors with complex layouts
Community models available (via Zenodo.org)	

Training

Create your own model

Necessities:

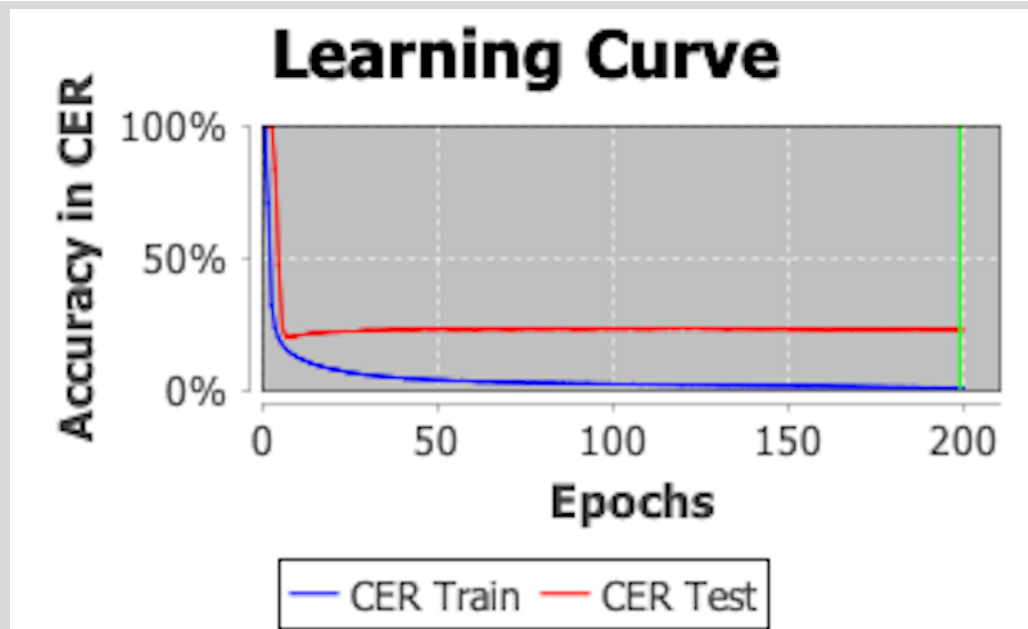
- Training Material
- Infrastructure
- Idea of ML approaches

Experience

- Interpretation of Training Curves

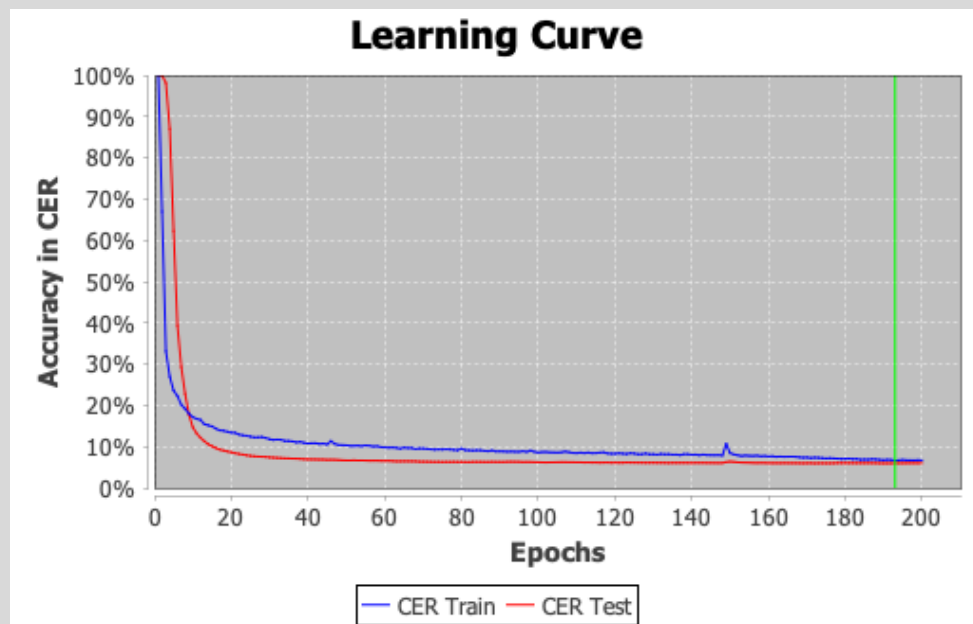
Interpretation of Graphs

Diverging curves: overfitting



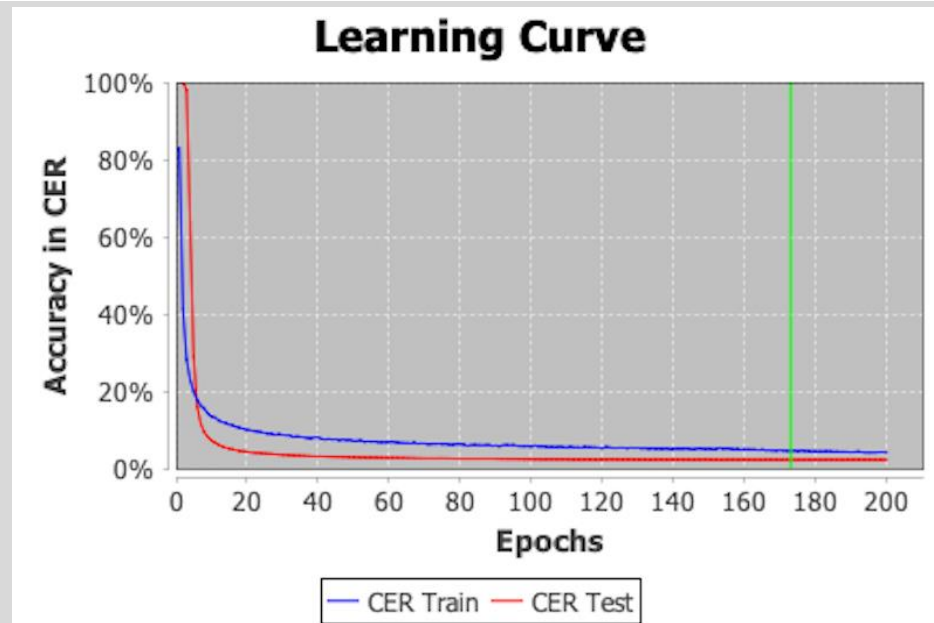
Interpretation of Graphs

Fully trained/optimal network



Interpretation of Graphs

Keep on training...



How to achieve good results

- define goals beforehand...
- ... adapt your style of transcription accordingly
 - Adherence to community guidelines

How to achieve good results

Transcription Guidelines:

- CATMUS: <https://hal.science/hal-03697382>

Segmentation Ontology:

- <https://jdmdh.episciences.org/14953>
- <https://segmonto.github.io/>

Cooperation

Data Sharing: HTRUnited

HTR-United

Home Browse the Catalog Record New Data Models Tools Github Automation The Team

HTR-United

HTR-United is a Github organization without any other form of legal personality. It aims at gathering HTR/OCR models for and transcriptions of all periods and style of writing, mostly but not exclusively in French. It was born from the mere necessity -for projects- to possess potential ground truth to rapidly train models on smaller corpora.

Use the data Provide data

htr-unity.github.io

Filters

Language:

Script:

Script type:

Project:

Dates: Not before: Not after:

Number of results: 9

Statistics about the results

	Unit	Amount	Projects with this unit
Characters		866.354	6
Lines		89.930	7

- Options
- Show Transcription Guidelines
 - Show Citation Informations

CREMMA Early Modern Books

CREMMA 1500 - 1779

Language: Script: Script Type: Hands: Volume: 84726 characters Volume: 148 files Volume: 17033 lines Known characters (PDF): 148 License: CC-BY 4.0 Software: eScriptorium + Kraken

Collection of book samples in early print forms, 16th to 17th century, in Latin and pre-orthographic French.

Authors: Clérice, Thibault

Complete record Tweet

CREMMA Medieval Latin Manuscripts

CREMMA 1000 - 1059

Language: Script: Script Type: Hands: Volume: 240291 characters Volume: 100 files Volume: 6148 lines Known characters (PDF): 718 License: CC-BY 4.0 Software: eScriptorium + Kraken

Ground truth for medieval latin manuscripts

Authors: CfxErice, Thibault and ChaguXE9, Alix and Vlachou Efstathiou, Matamateria

Complete record Tweet

Caroline Minuscule by Rscribe

Rscribe 800 - 1100

Language: Script: Script Type: Hands: Volume: 457 lines Volume: 17 files Volume: 45 regions Known characters: 16309 License: CC-BY 4.0 Software: eScriptorium + Kraken

Charters and Records of Königsfelden Abbey and Bailiwick (1308-1662)

1302 - 1670

Language: Script: Script Type: Hands: Volume: 60000 lines License: CC-BY 4.0 Software: Transkribus

The data set is the publication of the data of the scholarly edition "Urkunden und Akten des Klosters und der Hofmeisterei Königsfelden".

Cooperate

Romein et al. 2024.
“Exploring Data Provenance in Handwritten Text Recognition Infrastructure: Sharing and Reusing Ground Truth Data, Referencing Models, and Acknowledging Contributions. Starting the Conversation on How We Could Get It Done.”
Journal of Data Mining & Digital Humanities, 2024.
<https://doi.org/10.46298/jdmdh.10403>.

The screenshot shows the Zenodo article page for the paper "Exploring Data Provenance in Handwritten Text Recognition Infrastructure: Sharing and Reusing Ground Truth Data, Referencing Models, and Acknowledging Contributions. Starting the Conversation on How We Could Get It Done" by Romein et al. (2024). The page features a blue header with the Zenodo logo, a search bar, and navigation links for "Upload" and "Communities". The article title is prominently displayed, along with the authors' names and a list of ORCID iDs. The article is categorized as a "Journal article" and is "Open Access". On the right side, there are statistics showing 1,421 views and 795 downloads, with a link to "See more details...". Below the statistics, it indicates the article is indexed in "OpenAIRE". The "Publication date" is listed as March 24, 2023, and the DOI is 10.5281/zenodo.7765903. The "Keyword(s)" section includes "Automatic Text Recognition", "Handwritten Text Recognition", "Data Publication", "Open Data", "Data Provenance", "Data Curation", "Ground Truth", and "Sharing". The article is published in the "Journal of Data Mining and Digital Humanities".

Overview (Part II)

- 14:15 Test your own material
- 14:45 Show your results
- 15:05 Introduction to VLMs and command-line tools
- 15:25 Developing Workflows and Identifying Issues
- 15:45 Open Questions and Inputs on Demand

Test your own material
and share it with us

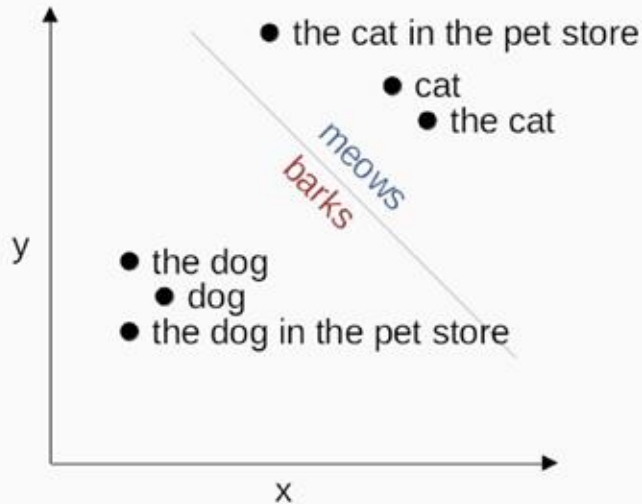
u^b

b
UNIVERSITÄT
BERN

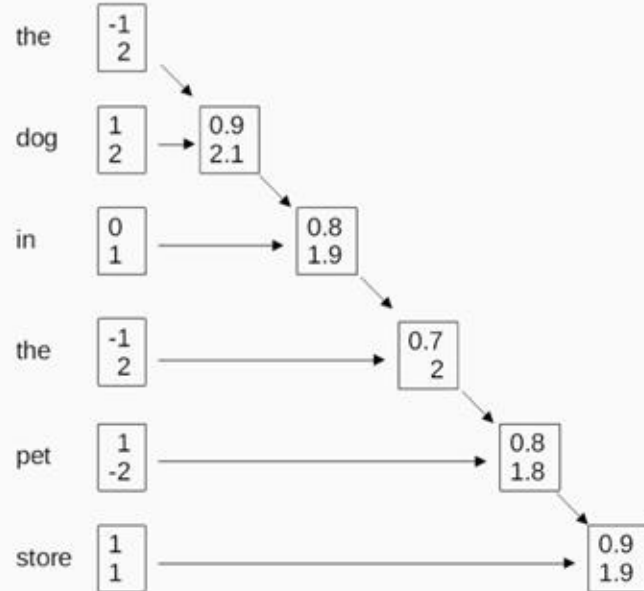
Introduction to VLMs

- How to LLMs work (way too short)
- Adding more dimensions (= multi-modal)

Vector Space



Representation as mathematical function/neural network



Quelle: Rico Sennrich, Interaktive Sprachmodelle: Ein Blick hinter die Kulissen.

https://kilof.unibe.ch/veranstaltungen/vergangene_veranstaltungen/kilof_tagung/index_ger.html

VLMs: Brainstorming potential challenges

Test out *commercial* VLMs

<https://www.hodelweb.ch/vlm/>

User: medievalist

Pw: kzoo2026!

Opportunity to compare

- Gemini
- QWEN-VL-Max
- Pixtral Large

Command line Tools

If you feel comfortable with the Command Line (and some Python)

- TrOCR:

<https://github.com/microsoft/unilm/tree/master/trocr>

- Kraken:

<https://kraken.re/main/index.html>

Workflows? Workflows!

Define:

- Goal / task
- Quality (threshold), be pragmatic
- A desired output

Identify:

- A suitable infrastructure
- Preexisting models
- An evaluation strategy