



# First Ground Truth for historic Sütterlin Reading Primers

**Katharina Krüger**

Leibniz-Institut für Bildungsmedien | Georg-Eckert-Institut (GEI)

**Nadine Towara**

Leibniz-Institut für Bildungsmedien | Georg-Eckert-Institut (GEI)

16.12.2021



7.400  
historic  
german  
textbooks

over 1,7mil  
digital pages

History,  
Geography,  
Politics and  
Alphabet books/  
Reading Primers

from the start of  
textbook production  
until 1918

over 1,4mil  
full texts



divided in  
before 1871  
and

German Empire  
(1871-1918)

about 250  
reading primers  
digitized

before 1871  
circa 500 books  
circa 35.000 pages

German Empire  
circa 300 books  
circa 24.000 pages

oldest digitized from  
1650

Management and  
Indexing  
with Goobi

Convey  
key messages  
of  
Christian teaching

Genre  
children's books

compulsory  
schooling  
around 1850

Reading Primers  
≠  
Alphabet Books

small  
research area

# Sütterlin



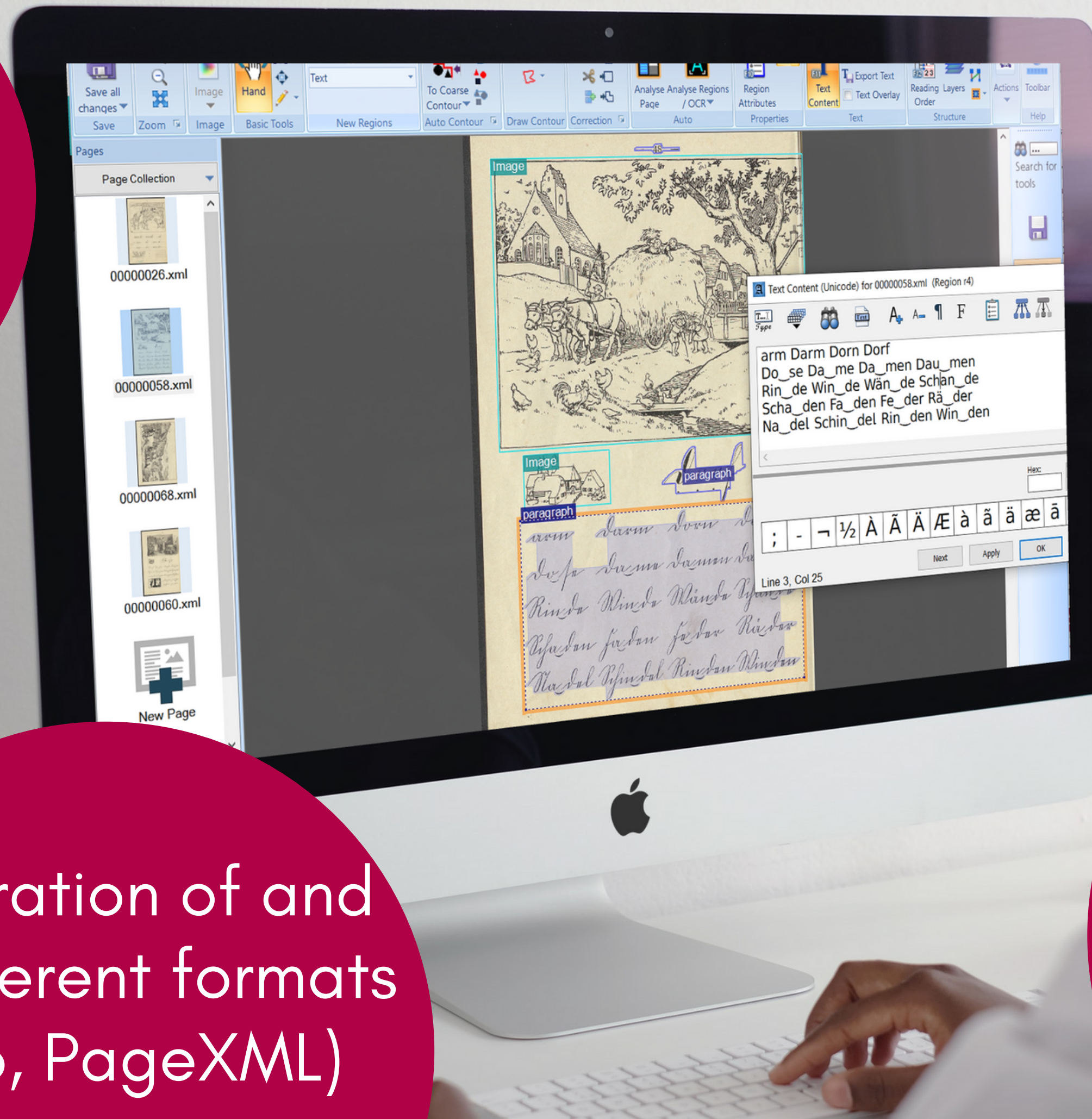
ⱱ	Ɱ	Ⱳ	ⱳ	ⱴ	Ⱶ	ⱶ	ⱷ	ⱸ	ⱹ	ⱺ	ⱻ	ⱼ
a	b	c	d	e	f	g	h	i	j	k	l	m
ⱽ	Ȿ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ
n	o	p	q	r	s	t	u	v	w	x	y	z
Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ
A	B	C	D	E	F	G	H	I	J	K	L	M
Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ	Ɀ
ä	ö	ü	Ä	Ö	Ü	ß	mm	nn	Schluss-s			

AletheiaPro

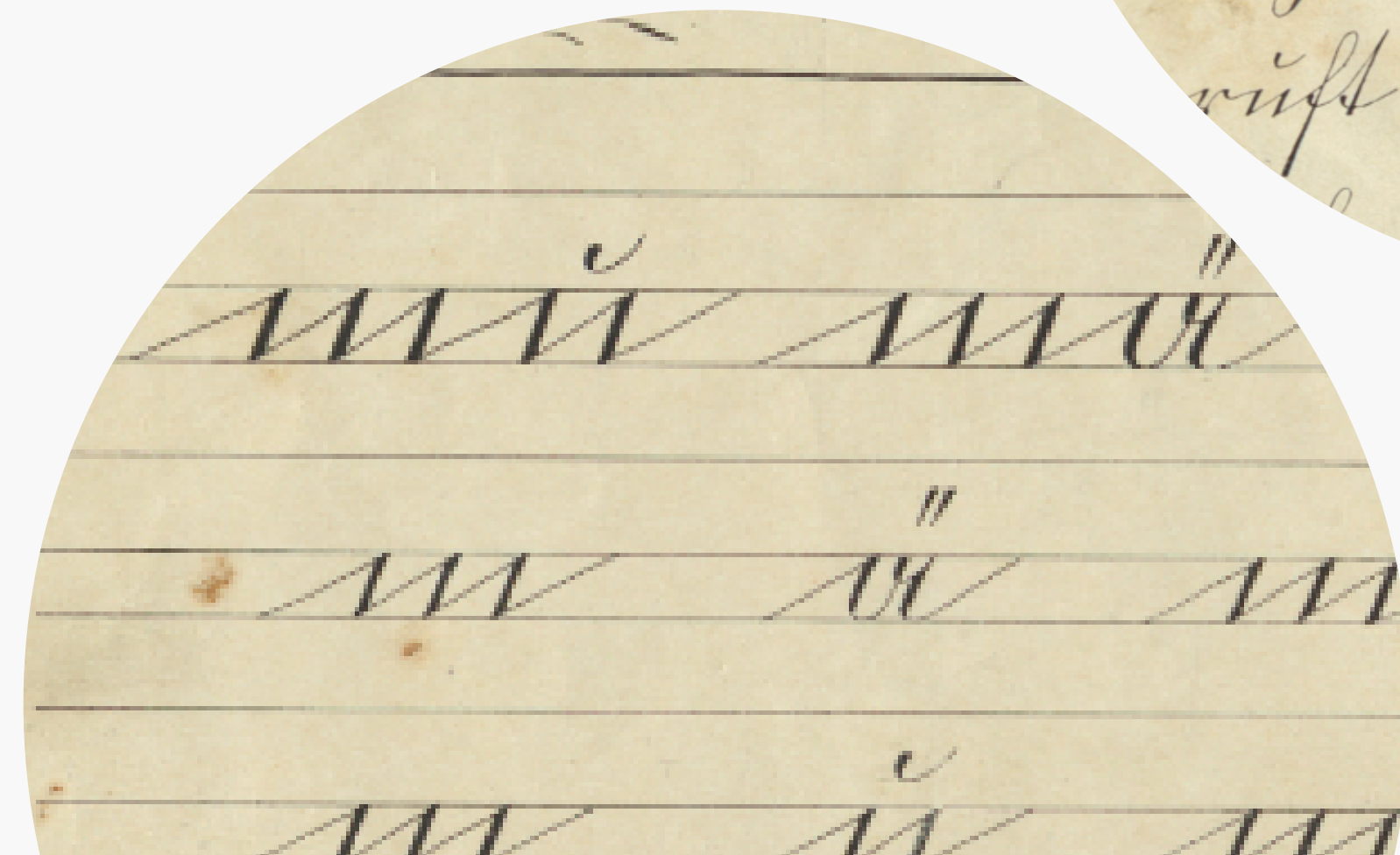
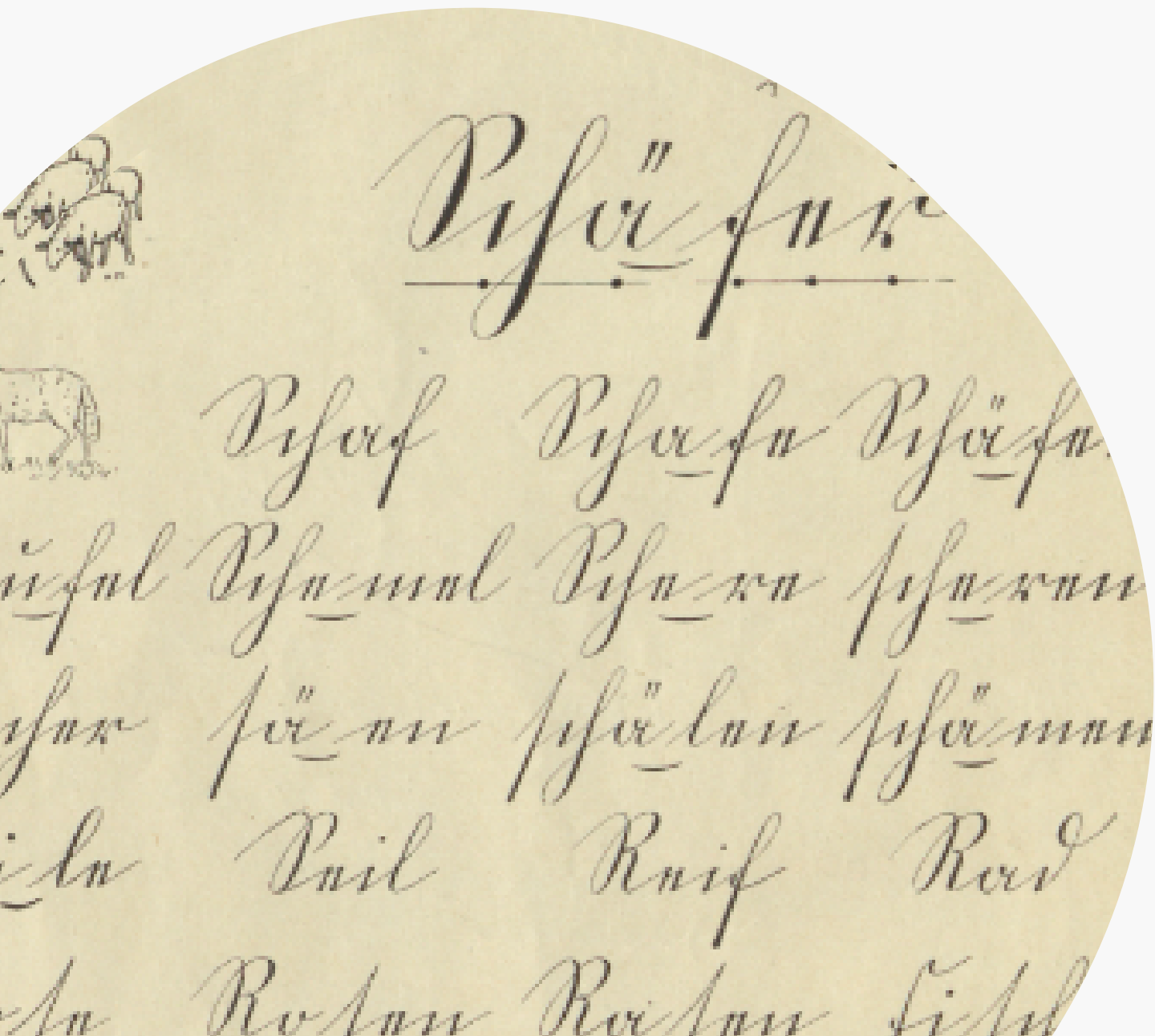
Development  
of new OCR  
models

Integration of and  
to different formats  
(Alto, PageXML)

Integration and  
optimization of  
existing OCR  
models



# Challenges



Ground  
Truth at  
Github

The screenshot shows a GitHub repository page for the 'ground-truth' repository. The main content area displays a list of commits with their commit hashes, messages, and dates. Below this is the README.md file content, which includes the title 'Transkriptionen von Fibeln (19. Jahrhundert)' and a description of the collection. The right sidebar shows repository metadata such as 'Readme', 'CC0-1.0 License', 'Releases', 'Packages', and 'Contributors'.

Commit Hash	Message	Time Ago
PPN1011424150	Fix GT errors (s/l)	4 months ago
PPN1020133104	Fix GT errors (s/l)	4 months ago
PPN1024726142	Fix transcription (f/f confusion)	4 months ago
PPN1024784126	Fix GT errors (s/l)	4 months ago
PPN1025195825	Fix GT errors (s/l)	4 months ago
PPN643815198	Create README.md	2 years ago
.gitignore	Add .gitignore	11 months ago
LICENSE	Initial commit	2 years ago
README.md	Update README.md	2 years ago

**README.md**

## Transkriptionen von Fibeln (19. Jahrhundert)

Diese Sammlung enthält Transkriptionen von Fibeln des 19. Jahrhunderts, die für Training oder Validierung von OCR-Modellen verwendet werden können.

Bezugsquellen:

- [Fibeln Kaiserreich](#) (Georg-Eckert-Institut Braunschweig)
- [Fibeln Kaiserreich](#) (Digitalisierungszentrum der SUB Göttingen)

**Contributors** 4

- stweil Stefan Weil
- Ma-Nuechter Maria Nüchter
- FVoigtschild
- kruegeka

in Cooperation  
with UB  
Mannheim







Contact:  
[gei-digital@gei.de](mailto:gei-digital@gei.de)

Merry Christmas!