

Planung: OCRD Modulprojekt 3: „Automatische Nachkorrektur historischer OCR-erfasster Drucke mit integrierter optionaler interaktiver Nachkorrektur“

Inhaltsverzeichnis

1	Stand der Implementierungen	1
2	Meetings	1
2.1	Kick-Off-Treffen 05.3 - 06.3 (Wolfenbüttel)	1
2.2	Videokonferenz 20.4	1
2.3	1. Entwicklertreffen 26.6 - 27.6 (Berlin)	2
2.4	Treffen mit Robert und Lena 07.8 - 08.8 (München)	2
2.5	Videokonferenz 18.9	2

1 Stand der Implementierungen

2 Meetings

2.1 Kick-Off-Treffen 05.3 - 06.3 (Wolfenbüttel)

- [Offizielles Protokoll](#)
- Schwerpunkt: Gemeinsame Datenformate (METS + PAGE) für Eingabe und Ausgabe.

2.2 Videokonferenz 20.4

- [Offizielles Protokoll](#)
- Stand des Modulprojekts erläutert.
- Vorstellung der OCRD-CLI Schnittstelle.

2.3 1. Entwicklertreffen 26.6 - 27.6 (Berlin)

- [Offizielles Protokoll](#)
- Schwerpunkt: Integration des Modulprojekts in OCRD Workflow / Workspace-Umgebung.
- Diskussion mit Leipzig:
 - Wie können Alignierungsergebnisse gemeinsam genutzt werden ?
 - * PAGE.xml als Austauschformat für Alignierungen multipler OCR mit GT.
 - * **Textequiv** zur Speicherung alternativer Alignierungen?
 - Gemeinsame Nutzung des neuronalen Sprachmodell
 - Verständigung auf baldiges Treffen in München zur weiteren gemeinsamen Abstimmung.
- Gespräch mit Kay:
 - Frage nach der Erstellung von (multipler) OCR zu Testzwecken.
 - Zeilenalignierung wird nicht benötigt, man kann davon ausgehen, dass Zeilen korrekt zueinander aligniert sind.

2.4 Treffen mit Robert und Lena 07.8 - 08.8 (München)

- [Protokoll](#)
- Absprachen zum Format des PAGE.XMLs für Alignierung.
- Gemeinsame Nutzung PAGE-Adapters.
- Gemeinsame Nutzung der RNN Sprachmodelle.

2.5 Videokonferenz 18.9

- [Offizielles Protokoll](#)
- Vorstellung Stand + Rückmeldung über positive Erfahrungen mit OCRD-Core Anbindung.
- Ansprechen der Problematik fehlender OCR-Beispieldaten für die Nachkorrektur.