# Sitzungsprotokoll

Sitzung am: 25.06.2015 um: 16:00 Uhr

Ort: Büro Manuel Burghardt

Teilnehmer: Manuel Burghardt, Miriam Nickl, Katia Buchhop, Thomas Schmidt, Florian Fuchs

Thema: Zwischenbericht & weitere Planung

Fragen:

- Wann und mit wem können wir uns wegen der Informationen zu den Metadaten und ihrer Verwendung treffen?

Verlauf:

Öffnen der Scan-Dateien:

* Mac: Gehe zu: Pfad eingeben. Mounten…
* Windows: Netzlaufwerke verbinden …

Ergebnis: „Standard OCR-Workflow vom Ergebnis her ok aber bei Weitem nicht optimal.“

* 758 Einzelne JPEG-Dateien (einschließlich leerer Rückseiten 🡪 müssten anhand der leeren Eintragungen in XML aussortiert werden)
* Noten werden falsch als Text erkannt
* Stringtag für jedes Wort
* Es wird angegeben wo genau das Wort steht (in der XML-Variante)
* Zu untersuchen: Kann man anhand dieser Angaben den Trash (automatisch) herausfiltern?
* Strophenblöcke werden erkannt (in der XML-Variante)
* Handschriftliches wird (fast) nicht erkann

Bericht über bisher betrachtete OCR-Tools:

* Handschriftentool (Transcriptorium.de) unbedingt ausprobieren.

ToDo:

* Mounten des Laufwerks für die Ergebnisse der Bib-Scans
* Sonst via Festplatte von Burghardts Rechner beziehen

🡪 Miriam holt sie im Laufe des Abends ab und verteilt sie an die anderen Gruppenmittglieder.

Anforderungen:

* Anforderungen werden hauptsächlich von Herrn Burghardt gestellt
* aber auch von Profs der Kulturwissenschaft (werden heute/ morgen angeschrieben und nach Terminvereinbarung von Burghardt mit Requirements-Analysefragen interviewt)
* Musikwissenschaftler Herr XY [//…Name entfallen] hat Ideen geliefert, was ihn an diesen Liedern interessieren könnte
* 🡪 Bis Ende nächster Woche bekommen wir die gesammelten Anforderungen von Burghardt

Nächste Schritte:

* Planung der Tests
* Beginn der OCR-Evaluation
  + Softwareanschaffung über Bib möglich
  + 🡪 Testen mit Kostenpflichtigen Tools. Möglich solange Trail kostenfrei ist
* Report der Metadatenanalyse

Festlegung: Wir können mit den Daten, so wie wir sie haben arbeiten. (Scanvorgang in der Bib muss nicht weiter evaluiert/ angepasst werden.)

Debrifing:

Weitere Planung:

* Nächstes Treffen Di, 30.06.2015 14:00 Uhr:
  + Planung der Tool-Tests
* 30.06.2015 14:45 Uhr: Projektleitertreffen: Besprechung des Plans und der Anforderungen
* Mi, 01.07.2015: Durchführung der OCR/OMR-Tool-Tests

Aufgaben:

* Miriam:
  + Reservierung des Testraums (und ggf. Testrechner)
  + Verteilung der Scanergebnisse
* Katia:
  + Vorbereiten des Projektplans
  + Erstellen des Metadatenanalysedokuments
* Thomas, Florian:
  + Geeigneten, einheitlichen Testablauf für die Tool-Tests überlegen
  + vertraut machen mit Handschriftentool (Transcriptorium.de)