# Sitzungsprotokoll

Sitzung am: 28.05.2015 um: 15:00 Uhr

Ort: Büro Manuel Burghardt

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

Thema: Vorstellen der ersten Analyseergebnisse; Besprechung des weiteren Vorgehens

Fragen:

* Nach welchen Kriterien sollen die Dokumente für den Testkorpus ausgewählt werden? (Zusammenhängend? Vollständigkeit? Wichtigste Merkmale?)
* Wie gehen wir mit fehlerhaften Datensätzen um?
* Wie normalisieren wir die Daten für das Info Retrival?
* BitBucket nur 5 User möglich? Wie mit Stakeholder und Projektbetreuer?
* Excel: Sollen wir eine Kriterienliste erstellen, damit in Zukunft die Aufnahme weiterer Daten einheitlich durchgeführt werden kann?

Verlauf:

* Vorstellen der Excel-Tabellen; Probleme mit den Daten ansprechen (Vollständigkeit; Titel)
* 🡪Einführen: Schritt zur Vereinheitlichung
* Frage: Welche Metadaten sind für uns wichtig?
* Aufgabe: Report: Informationstypen, Wie viele Sind da? Welche Ausprägungen gibt es?
* 🡪 GND Abgleich um aus der Sangesortspalte die Orte zu extrahieren (Ortsnamenlexikon) 🡪 Normalisierte Spalte einfügen
* Vorschlag: Treffen mit Kulturwissenschaftlern zur Besprechung der Daten 🡪 Burghardt meldet sich wegen Termin. Bis dahin Analyse der Tabellen zusammenfassen
* Vorschlag zur Digitalisierung der Noten: Musiker mit Midi Keyboard 🡪 Wie lange braucht eine Musiker um das Notenblatt nachzuspielen? 🡪 Evaluieren?
* Auswahl der Testdocs.?
  + Aufeinanderfolgend (wenn sie geordnet vorliegen)
  + 🡪 1000 digitalisieren und geeignete auswählen
* Sollen wir eine Kriterienliste erstellen, damit in Zukunft die Aufnahme weiterer Daten einheitlich durchgeführt werden kann? 🡪 Wäre für die Erstellung eines Workflows sinnvoll!

ToDo:

* Überlegen: Welche Spalten wählen? Wie normalisieren?
* Mail mit bisherigen Excel-Dokumenten an Stakeholder
* Dokumentieren der Interessantesten Spalten und Besonderheiten
* Präsentationstermin weiterleiten

Debrifing:

* Fertigstellung der Excel-Analyse