

# Anforderungskatalog – Erste Version

## Funktional

1. Die Software muss eingegebene Textdokumente auslesen und nach Themenbereichen clustern können.
  - a. Die Software muss PDFs als Eingabeformate akzeptieren.
  - b. Die Software muss fähig sein, Texte auf Webseiten auszulesen.
  - c. Die Software muss sowohl englische als auch deutsche Texte verarbeiten können.
2. Die Software muss dem User die Möglichkeit bieten, den Suchlauf über einen Knopf im Programm starten zu können.
3. Die Software muss die erstellten Cluster visualisieren können.
  - i. Subcluster müssen ebenfalls grafisch dargestellt werden.
  - ii. Der User muss einsehen können, welche Dokumente zu einem Cluster zugeteilt wurden.
- b. Wenn ein Suchlauf abgeschlossen ist, muss die Software die erstellte Clustervisualisierung als Grafik exportieren können.
4. Die Software muss relevante Begriffe in einer Wortwolke visualisieren.
  - a. Wenn ein Suchlauf abgeschlossen ist, muss die Software die erstellte Wortwolke als Grafik exportieren können.
5. Wenn ein Suchlauf abgeschlossen ist, muss die Software CSV-Dateien exportieren können.
  - a. Eine CSV-Datei muss eine Auflistung von Clusternummern und zugehörigen Clusternamen beinhalten.
  - b. Eine weitere CSV-Datei muss eine Auflistung der Quellennummern und den zugehörigen Quellenbezeichnungen, Daten und die Anzahl der Cluster enthalten, in denen die jeweilige Quelle vorkommt.
    - i. In dieser CSV-Datei muss außerdem aufgelistet sein, ob es bei der Verarbeitung einzelner Quellen jeweils einen Fehler gab.
  - c. In einer weiteren CSV-Datei muss eine Auflistung der gezählten Wörter, ihrer jeweiligen Nummer und Anzahl in der Dokumentensammlung enthalten sein.
6. Die Software muss fähig sein, eingegebene Suchkonfigurationen als XML-Datei zu exportieren.
7. Die Software muss fähig sein, Suchkonfigurationen in Form von XML-Dateien zu importieren.
8. Die Software muss fähig sein, auf einen lokalen Ordner zuzugreifen, um darin vorhandene Textdokumente zu importieren.
  - a. Die Software muss dem User die Möglichkeit bieten, den Pfad des lokalen Ordners angeben zu können.
  - b. Die Software muss die Textdokumente aus dem lokalen Ordner in einer Listenansicht aufbereiten.
    - i. Die Liste muss für jede Quelle eine Checkbox anzeigen, mit der die Quelle in den Suchlauf einbezogen oder aus diesem ausgeschlossen werden kann.
    - ii. Die Liste muss außerdem für jede Quelle deren Nummer, Namen und Datum auflisten.
9. Die Software muss dem User die Möglichkeit bieten, den Speicherort für Exportdateien angeben zu können.
10. Die Software muss dem User die Möglichkeit bieten, Textdokumente nach Datum filtern zu können.

- a. Die Software muss dem User die Möglichkeit bieten, eine Zeitspanne angeben zu können, um Dokumente nach dieser einzugrenzen.
- 11. Die Software soll dem User die Möglichkeit bieten, eine maximale Anzahl an Clustern, die bei einem Suchlauf erstellt werden können, festlegen zu können.
- 12. Die Software muss Logfiles anlegen können.
  - a. Falls bei einem Suchlauf Fehler auftreten, müssen diese durch die Software in einem Logfile dokumentiert werden.
  - b. Die erstellten Logfiles müssen über einen Knopfdruck im Programm aufgerufen werden können.

### **Nicht-funktional**

- 1. Die Software muss eine Webapplikation sein.
  - a. Die Software muss auf dem aktuellen Windows-Betriebssystem (Windows 10) funktionieren.
- 2. Die Software muss gesetzeskonform sein.
- 3. Der Code der Software muss englisch kommentiert sein.
- 4. Die Software muss eine Hilfefunktion bieten, die über Knopfdruck im Programm aufgerufen werden kann.
  - a. Die Hilfefunktion muss eine kurze Dokumentation darüber beinhalten, wie der User einen Suchlauf durchführen kann.
  - b. Die Hilfefunktion muss darüber informieren, was die Visualisierung beschreibt.
- 5. Die Software muss abteilungsübergreifend genutzt werden können.
- 6. Die Software muss von Nutzern ohne besondere IT-Kenntnisse genutzt werden können.