Ein Bild, das Diagramm enthält.

Automatisch generierte Beschreibung

Version 29.03.2023

# Aufgabenstellung Team GeoContentAnalyser

* Entwicklung eines Softwaretools, um quantitative Daten zum Thema „Geo Content Management“ zu ermitteln und statistisch auswertbar abzuspeichern
* Agile Entwicklung
* Als ausschließlich englischsprachiges Open Source Project unter Account Prof. Weinkauf auf GIT HUB
* [Eine Open-Source-Lizenz](https://help.github.com/articles/open-source-licensing/#where-does-the-license-live-on-my-repository)
* [Eine README](https://help.github.com/articles/create-a-repo/#commit-your-first-change)
* [Beitragsrichtlinien](https://help.github.com/articles/setting-guidelines-for-repository-contributors/)
* [einen Verhaltenskodex](https://opensource.guide/de/code-of-conduct/)
* Als ausschließlich englischsprachige JAVA-Konsolenanwendung ohne spezielle Bibliotheken
* Data Storage als Files im Format Markdown Table in einem konfigurierbaren Datenverzeichnis
* Recherche vorhandener Tools und Bibliotheken
  + Im Internet nach Tool, Bibliotheken und Quellcode
  + In wissenschaftlichen Papers nach Beschreibungen von Tools und ähnlichen wiss. Untersuchungen (siehe wiss. Suchmaschinen / Literatursuche unter Organisatorisches in meinem ILIAS-Bereich)
* Algorithm
  + Importiere zu einem definierten, dokumentierten Stichtag die URL’s der Landkreisliste laut Wikipedia
  + Erzeuge im Datenverzeichnis einen Unterordner mit Datum und Uhrzeitangabe im Ordnernamen, z.B. 20230322-10-02
  + Lege eine Logdatei „Landkreiswebsitegesamtstatistik“ mit Datum und Uhrzeit an für die Ergebniszusammenfassung. Die Datei wird eine Tabelle aller Landkreise mit den aggregierten Ergebnissen je Landkreis speichern.
  + Gehe auf jede Landkreisseite mit einem zeitgesteuerten Crawler, der (möglichst) nicht als BOT identifiziert wird (Alternativ kann auch ein Standarddownloader für temporäre, lokale Daten in Betracht gezogen werden.)
    - Lege eine Logdatei „Landkreisseitenstatistik“ mit Datum und Uhrzeit an für die Analysestatistik pro einzelner Landkreisinhaltsseite
    - Gehe auf jede einzelne Seite der Landkreiswebsite und analysiere deren Inhalt
      * Anzahl von „Informationsobjekten“ mit wahrscheinlichem Raumbezug
      * Anzahl interaktiver, eingebetteter Karten
      * Anzahl extern verlinkter Karten
      * Anzahl von Adressangaben (indirekter Raumbezug)
      * Anzahl von Koordinaten (Paaren …)
      * Anzahl von E-Services direkt auf dieser Seite
      * Anzahl von Links auf E-Services auf dieser Seite
      * Schreibe für jede einzelne Seite eine Logdatei mit den Beweisen (File mit Markdown Table), beweise jede Zählung
      * Aktualisiere die Datei Landkreisseitenstatistik mit den auf dieser Seite gefundenen Infos.
    - Aktualisiere nach Abschluss der Analyse der Landkreiswebsite die Gesamtstatistik, d.h. die zusammengefassten Ergebnisse für den hier betrachteten Landkreis in der Datei „Landkreiswebsitegesamtstatistik“.
  + Anreicherung der Landkreiswebsitegesamtstatistik um weitere Metadatenfelder pro Datensatz = Tupel = Zeile = Landkreis:
    - Bundesland
    - Flächengröße in qkm
    - Anzahl Einwohner
* Test 1: Konvertierung der Tabelle in Excel (<https://tableconvert.com/markdown-to-excel>) und erste Statistische Auswertung mit Pivot Tabellen in Excel
* Test 2: Konvertierung der Tabelle (<https://tableconvert.com/markdown-to-excel>) und Import der Daten in R und oder einer anderen wissenschaftlichen Statistiksoftware auf Pool-PC der HoMe
* Ausblick: Anwendung auf Wikipedia-Liste der Städte und Gemeinden

* Vergleichen Sie ihre erste Statistik mit den in der wissenschaftlichen Literatur gemachten Aussagen zum Geo-Content Management!

* Um die in der Literatur gemachten Aussagen zum Geo-Content Management zu finden, führen Sie bitte vorher eine Systematische Literaturanalyse (SLR Systematic Literature Review) nach der Methode von Kitchenham durch, inkl. Planung, protokollierter Durchführung und wissenschaftlichem Report …
  + Suchstring: Geocms OR "geo-content management system" OR "geo content management system" OR "geo-content-management-system"
  + 173 Treffer bei Scholar am 22.3.23
  + Idee 1: (Geocms OR "geo-content management system" OR "geo content management system" OR "geo-content-management-system") AND government AND Website
  + Idee 2: Zeitraum einschränken, z.B. 2012-2022 wäre geeignet
  + Idee 3: Sprache einschränken: Nur englisch
  + RQ vorschlagen und abstimmen …
  + …
* *Gedanke: Beispiele von Tina Balke und den Studentenprojekten raussuchen, die manuell erhoben wurden …*
* **Zusätzlich kann gern nach am Markt vorhandenen GeoCMS Lösungen im Internet recherchiert werden, wobei diese Systeme nicht immer als GeoCMS bezeichnet werden, z.B. Drupal ….**