

Anlage A

Themenschwerpunkte für die Zusammenarbeit im Kooperationsprojekt „Transparenzportal“ für die Arbeiten ab 2017 - 2019:

- **Zusammenfassungen von Suchergebnissen**

Die Anzahl der Datensätze im Informationsregister liegt aktuell bei ca. 50000, sie wächst kontinuierlich. Daher liefert das System bei Suchanfragen eine Vielzahl von Treffern. Um den Benutzer bei der Analyse des Suchergebnisses zu unterstützen, könnte eine Übersicht über das Suchergebnis berechnet werden. Diese würde zusammenfassende Merkmale des Ergebnisses neben der Ergebnisliste anzeigen. Beispiele für solche Zusammenfassungen sind örtliche Verteilungen der Ergebnisse (z.B. die involvierten Bezirke) oder betroffene Zeiträume der Ergebnisse (z.B. Jahresbezüge).

- **Ermittlung von themenbezogenen Datensätzen**

Als Ergänzung der erweiterten und der facettenbasierten Suche könnte eine themenbezogene Suche alle Datensätze ermitteln, die ein gemeinsames Thema betreffen, wie z.B. „Haushaltsberichte“ oder „Migranten“. Dazu kann die Verwendung von Data- und Text-Mining Verfahren geprüft werden, um thematische Bezüge von Datensätzen automatisch zu ermitteln. Diese Interpretation der Daten ermöglicht Einsichten in semantische Zusammenhänge, die syntaktischer Suche in Informationsobjekten, wie sie im Moment unterstützt wird, verborgen bleiben.

- **Einbindung von Geografischen Aspekten**

Die Suche kann durch die Verwendung von geografischen Positionen des Suchenden sowie dem geografischen Bezug von Datensätzen erweitert werden. Dazu sollen die Datensätze automatisch durch Text-Mining-Verfahren georeferenziert werden. Weiterhin kann durch eine Kartensuche der Benutzer gezielt Datensätze aus einem gegebenen geografischen Bereich ermitteln.

- **Entwicklung von Semantischen Suchverfahren für das HmbTG**

Aktuell wird eine syntaktische Suche dem Benutzer des Portals zur Verfügung gestellt. Es kann untersucht werden, inwieweit eine semantische Suche den Benutzer (Bürger oder IT-Firmen) unterstützen kann. Eine semantische Suche ermöglicht z.B. die Auflösung von Mehrdeutigkeiten, liefert Erklärungen für Suchergebnisse oder ermöglicht eine hierarchiebezogene Suche, z.B. das Ermitteln von Datensätzen bzgl. *Fahrradwege*, wenn nach *Verkehr* gesucht wurde.

- **Automatische Übersetzungen von Dokumenten**

Die Suche und die Suchergebnisse können durch automatische Übersetzung auch nicht-deutschsprachigen Mitbürgern zugänglich gemacht werden. Dazu könnten vorhandene Werkzeuge für die automatische Übersetzung beim Harvesten und/oder bei der Suche eingebunden werden.

- **Weiterentwicklung des Portals und Anbindung neuer Liefersysteme**

Das bisher entwickelte Portal ist mit den Komponenten Harvester, Zentraler Speicher, Workflow-Prozesse, Portal-Oberfläche ein komplexes System. Speziell die Verarbeitung unterschiedlicher Datenformate und Inhalten verursachen Weiterentwicklungsaufgaben und Administrationsaufgaben im täglichen Betrieb (z.B. Metadatenänderungen, Formatänderungen, Anpassungen der Oberfläche an Browser etc.). Auch die Anbindung neuer Systeme erfordert Integrationsarbeiten und teils neue Harvester. Weiterhin erfordert die Auslieferung der Datensätze Anpassungen an gängige Austauschformate (z.B. neue OpenData-Schemata).

Die geplante externe Evaluierung des HmbTG-Portals wird voraussichtlich Änderungen in allen Komponenten des Portals verursachen sowie eine Ausweitung der veröffentlichungspflichtigen Stellen und damit eine Ausweitung der Liefersysteme zur Folge haben.