Richtlinie zum „Software-Sprint“

Crossfoam - Browser-Plugin zur Analyse und Visualisierung von Filterbubbles mit Hilfe von maschinellem Lernen

Schlussbericht

Zuwendungsempfänger:

ULTRAPOP - Sebastian Meier und Daniel Lommes Gbr

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01IS18S47 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor.

Kurze Darstellung der Aufgabenstellung und Motivation

Im Gegensatz zum Begriff der Öffentlichkeit, in dem ideal-typisch alle sprechen und von allen gehört werden können, werden in Filterblasen homogene Meinungen einem homogenen Publikum vorgetragen. Die fortschreitende Verschiebung des Nachrichtenkonsum in Soziale Medien, gepaart mit deren algorithmischer Kuratierung, führt zum weitgehenden Verblassen der ersten, während tausende Filterblasen blühen. Im Gegensatz zu Ansätzen der Zensur und/oder der Priorisierung “klassischer” Medienunternehmen versuchen wir Menschen zu sensibilisieren, und ihnen ein Werkzeug an die Hand zu geben, um Filterblasen zu erkennen und Inhalte in diesen verorten zu können. Dadurch wird Data Literacy aktiv gefördert.

Aus Datenschutzgründen wollen wir versuchen die Anwendung soweit wie möglich ohne Cloud-Infrastrukturen aufzubauen. Ziel ist es eine Client-seitige Anwendung zu entwickeln, welche im Browser der User*in läuft. Die Wahl eines Browserplugins ist nicht nur technologisch motiviert, sondern zielt darauf ab die Nutzung für möglichst viele Menschen so niederschwellig wie möglich zu machen. Das Plugin selber nutzt JavaScript als primäre Programmiersprache. Das Projekt wird sich in der ersten Phase auf Twitter konzentrieren und darauf aufbauend für andere Netzwerke erweitert werden. Für Twitter nutzen wir die API um Nutzer*innen-Netzwerke und deren Inhalte zu erschließen. Ein datenbank-ähnliches System auf Seiten der Nutzer*innen dient als Cache der Informationen. Diese können dann über Visualisierungen und Erklärungende Darstellungen erkundet werden.

Beitrag des Projektes zu den Zielen der Förderinitiative „Software-Sprint“

Die Zielgruppe unserer Anwendung sind Nutzer*innen sozialer Medien. Speziell Nutzer*innen welche soziale Medien für den Konsum von Nachrichten nutzen. In Bezug auf diese Zielgruppen, versucht unser Projekt nicht nur ein Werkzeug zur Verfügung zu stellen, sondern durch die Sensibilisierung für das Thema der Filterblasen, die Nutzer*innen zur digitale Selbstermächtigung zu befähigen. Wir nehmen durch diesen Fokus direkt Bezug auf das Themenspektrum B der Data-Literacy. Das Tool selber nutzt verschiedene Methoden der Datenvisualisierung und des daten-gestützten Storytellings und nimmt somit auch Bezug auf das Themenspektrum A.

Ausführliche Darstellung der Ergebnisse

Vorgehen:

Im Rahmen der Förderung konnten wir erfolgreich die anvisierte Browser-Erweiterung umsetzen. Im ersten Projektabschnitt haben wir die Basisinfrastruktur der Erweiterung entwickelt. Hierbei lag die Herausforderung darauf die Erweiterung von Anfang an für mehrere Browser-Engines zu entwickeln. Hilfreich hierbei war der von Mozilla bereitgestellte Polyfill für Browser-Erweiterungen. Nichts desto trotz stellte uns die Cross-Browser-Compatibility vor große Probleme, z.B. in Hinblick auf das Styling oder die Datenspeicherung. Bei der Datenspeicherung mussten wir auch von der vorgesehenen SQLite Datenbank abweichen und statt dessen einen simplen Key-Value-Storage nutzen. Um diesen speicher- und performanzeffizient nutzen zu können, mussten wir unsere Informationsarchitektur entsprechend anpassen. In der zweiten Projektphase wurde das client-seitige Daten-Harvesting entwickelt. In dieser ersten Stufe der Erweiterung wurde dies für Twitter entwickelt.
Die dritte Phase konzentrierte sich auf die Visualisierung der in Phase zwei gesammelten Daten. Mit Hilfe der JavaScript Bibliothek D3 wurden eine Reihe an Visualisierungen entwickelt, welche den Nutzer*innen erlauben die gesammelten Daten visuell zu explorieren.
Die vierte Phase konzentrierte sich, anstatt des vorgesehenen Austausches über ein offenes Repository, auf die Optimierung der Visualisierungen. Im offenen Repository sahen wir zu viele potentielle Gefahren, gerade für nicht technisch versierte Nutzer*innen. Statt dessen konzentrierten wir uns auf die Optimierung der Visualisierungen. Speziell große Netzwerke führten die Performanz der Browser an seine Grenzen. Deshalb entwickelt wir in der vierten Phase ein auf Web-GL aufgebautes System, welches die D3-Visualisierungen optimierte.
Den letzten 5. Milestone konnten wir zeitlich nicht ganz erreichen, da das Veröffentlichen der Erweiterungen in den jeweiligen Browser-Shops doch länger dauerte als angenommen. Wir hoffen, dass dies in den nächsten Wochen abgeschlossen werden kann und das Projekt somit seinen vorerst letzten Milestone erreicht.

Ethisches Design & Entwickeln und Nutzertests:

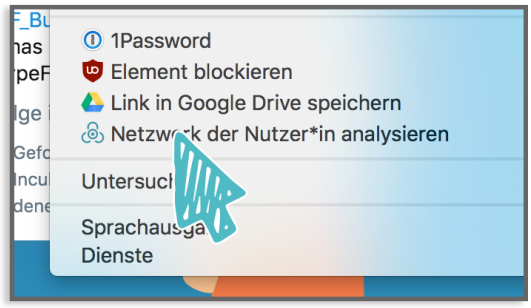
Über alle Phasen hinweg, haben wir auf mehreren Metaebenen das Projekt überwacht. Zum einen haben wir mehrere Evaluationen mit Nutzer*innen durchgeführt, um frühzeitig Probleme des Systems zu erkennen und eine agile und nachhaltige Umsetzung zu gewährleisten. Hierüber haben wir z.B. festgestellt, dass die lange initiale Wartezeit von Nutzer*innen nicht besonders gut aufgenommen wurden. Deshalb wird die Erweiterung nun mit einem voroptimierten Datensatz ausgeliefert, welcher von den Nutzer*innen sofort exploriert werden kann. Zusätzlich gibt es eine progressive Visualisierung, welche den Benutzer*innen einen visuellen Status gibt, während die Daten weiter gesammelt werden. Eine weitere Erkenntnis aus den Tests war auch, dass Nutzer*innen

das Werkzeug nicht nur zur Erforschung ihrer eigenen Netzwerke in Bezug auf Filterblasen nutzen, sondern, dass sich das Werkzeug auch eignet, um das Netzwerk zu optimieren und aufzuräumen. Auf einer weiteren Ebene haben wir den Entwicklungsprozess aus einer ethischen Perspektive überwacht. Mit der Entwicklung jedes neuen Moduls haben wir versucht, zu analysieren in wie weit durch dieses neue Module nachhaltige Effekte für potentielle Nutzer*innen entstehen können.

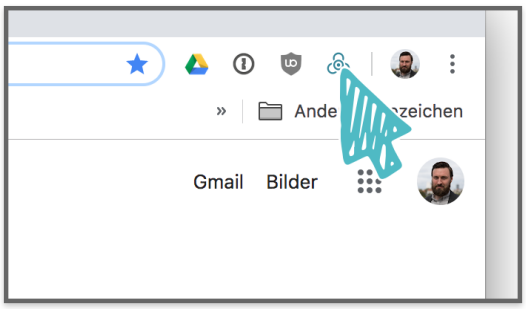
Wie funktioniert Crossfoam:



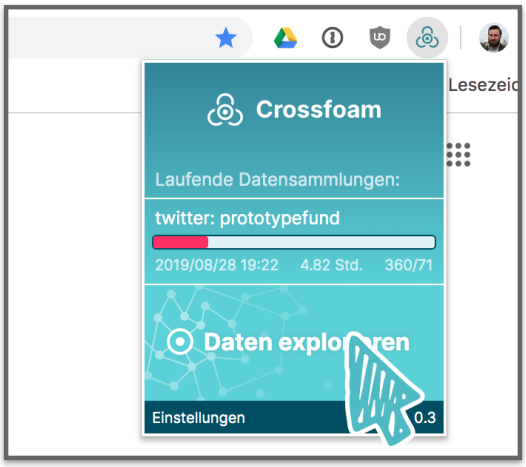
Nutzer*innen können eine beliebige Nutzer*in auf Twitter auswählen…



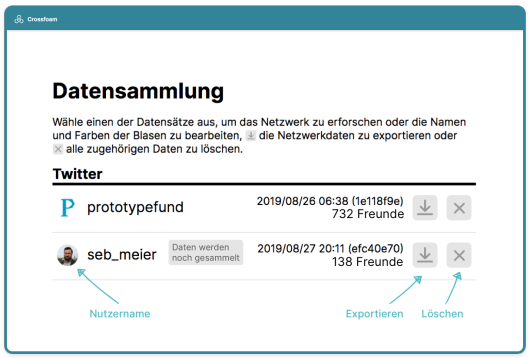
.. und direkt im Browser über die Erweiterung Crossfoam analysieren lassen.



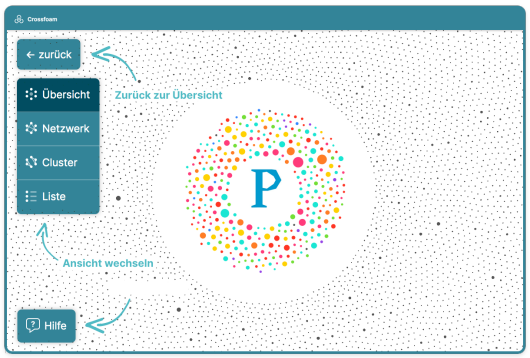
Damit Nutzer*innen sich über die Filterblasen informieren können. Müssen die zugehörigen Daten erst gesammelt werden.



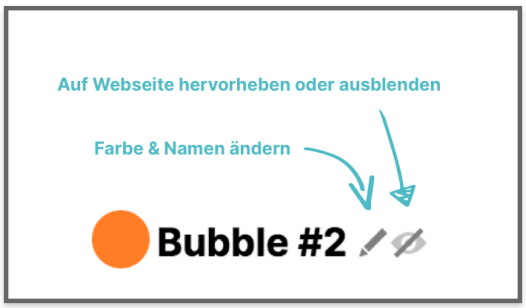
Den Status dieser Datensammlung kann von der Nutzer*in jeder Zeit bequem über das Erweiterungsmenü abgerufen werden.



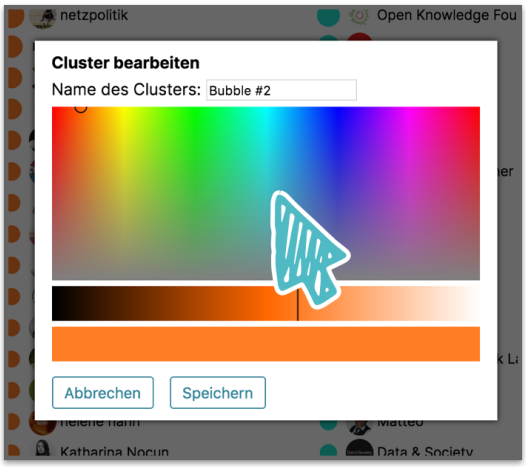
Auf der Übersichtsseite der Erweiterung werden alle Datensammlungen angezeigt, sowie deren Status. Von hier können die Daten exportiert werden oder direkt in der Erweiterung visualisiert werden.



Alle Visualisierungen haben eine Hilfefunktion. Über diese werden Nutzer*innen durch die Funktionalität der Visualisierung geleitet.



Die vom System erkannten Blasen können von der Nutzer*in angepasst werden um den persönlichen Wünschen zu entsprechen.

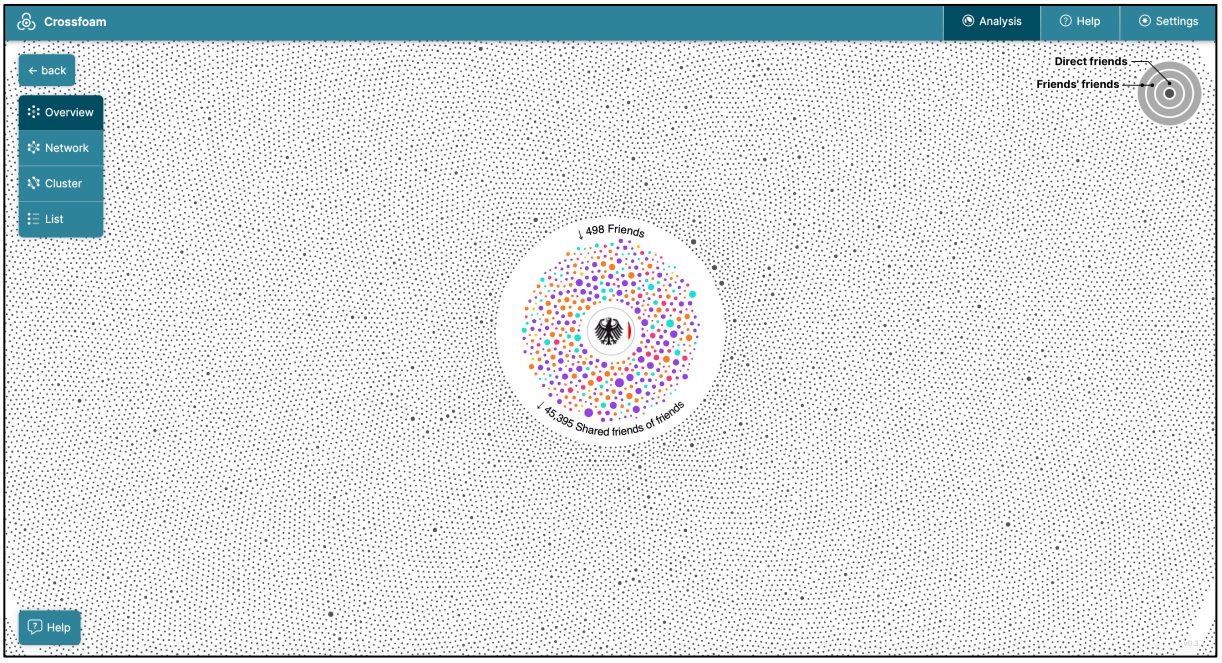


Nicht nur die Beschriftung auch die Farbe kann entsprechend angepasst werden.

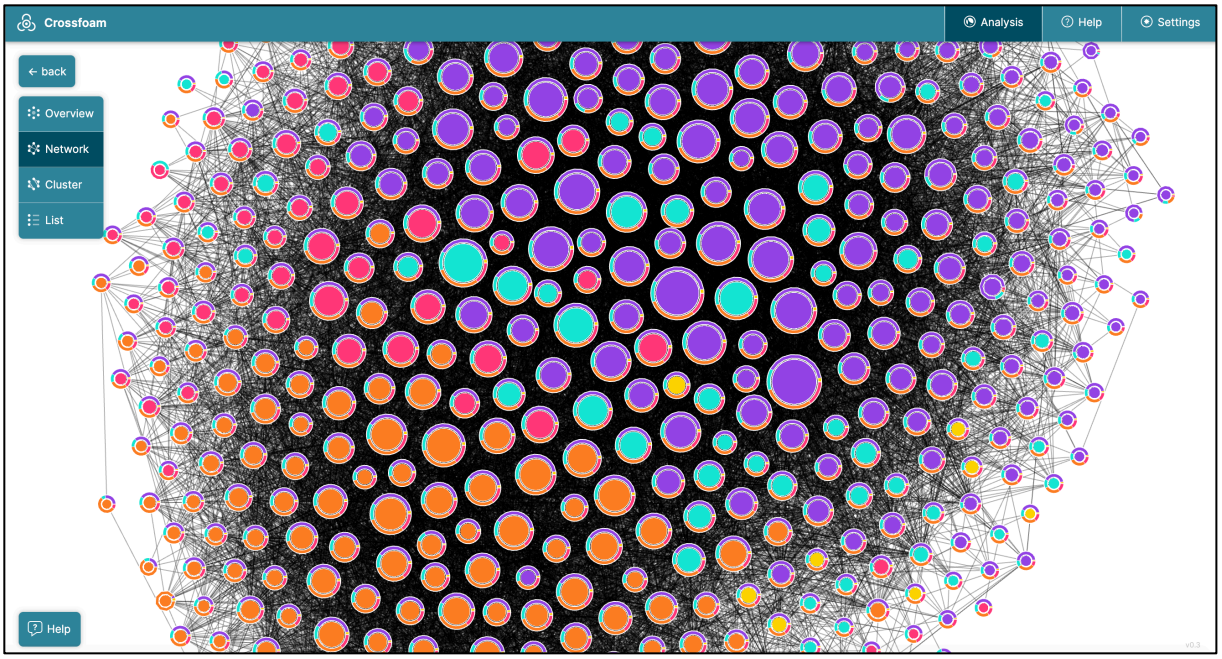


Wenn die Nutzer*in durch das soziale Netzwerk oder andere Seiten surft, werden erkannte Personen, mit entsprechenden Labels versehen. So lässt sich schnell erkennen zu welchen Blasen eine bestimmte Nutzer*in gehört.

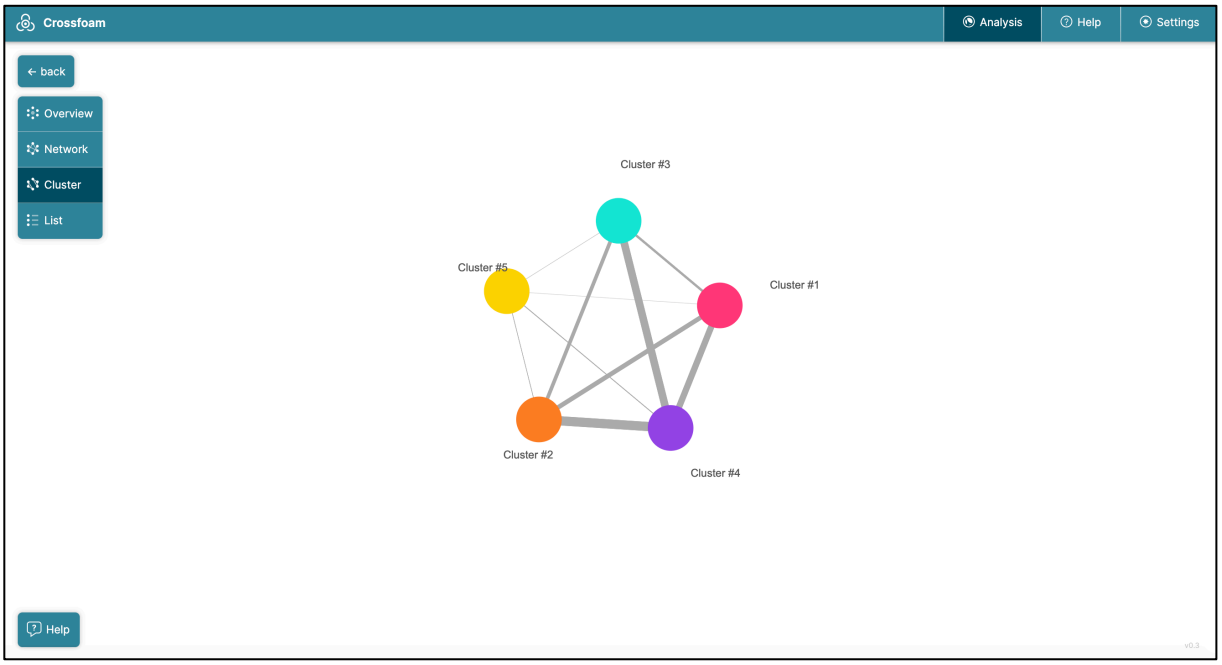
Visualisierungen:



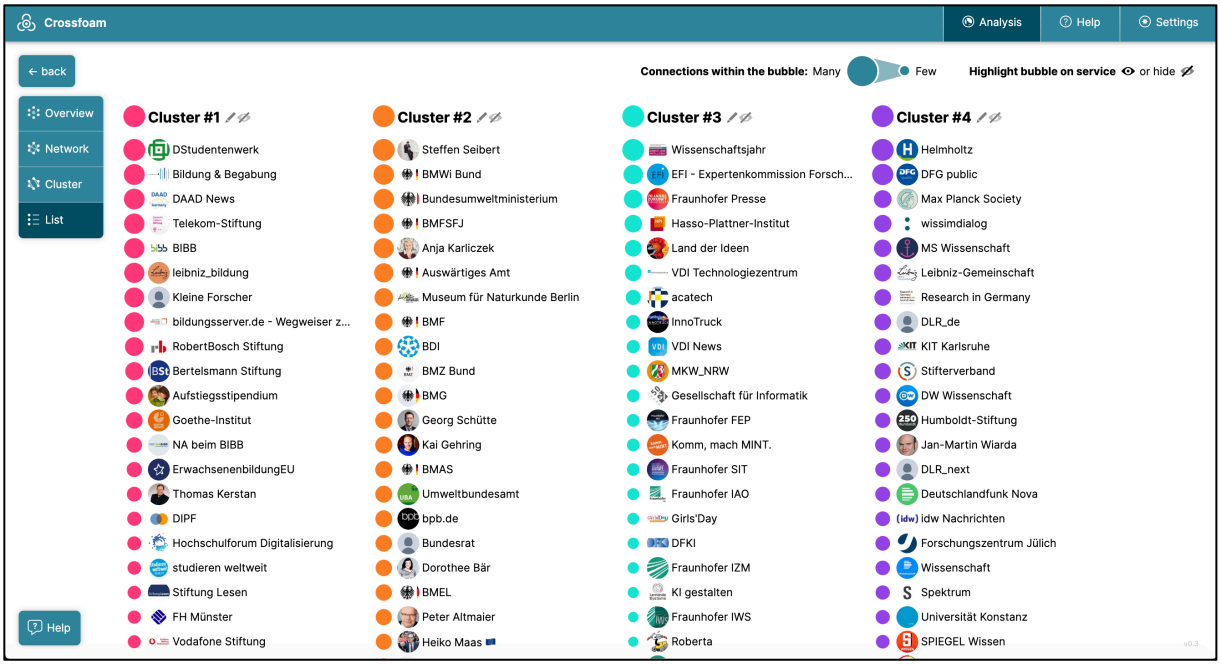
Die Übersichtsvisualisierung vermittelt den Nutzer*innen einen Eindruck über die Größe ihres pirmären und sekundären Netzwerks. Die Visualisierung verschafft einen ersten Eindruck, welche Personen einen besonders großen Einfluss auf die Informationen haben, welche der Nutzer*in in ihrem Netzwerk angezeigt werden.



Die Netzwerkansicht gibt einen Einblick in die Filterblasen. In dieser Ansicht werden nicht nur die zentralen Akteur*innen der Blasen sichtbar sondern auch, dass Nutzer*innen zwar einer Blase primär zugeordnet werden, aber durchaus Überlappungen zwischen einzelnen Blasen bestehen.



In der Filterblasenansicht liegt der Fokus auf den Blasen selber. Neben den Überlappungen, welche schon in der Netzweransicht sichtbar wurden, kann die Nutzer*in in die einzelnen Netze eintauchen und dort tiefergehende Analysen durchführen.



Die Listenansicht gibt der Nutzer*in einen schnellen Überblick über alle Personen in Ihrem Netzwerk und deren Blasenzugehörigkeit. Von hier aus können Nutzer*innen auch die Blasen anpassen.

Zielgruppe, Nutzen und mögliche Weiterentwicklungen

Mit Hilfe der Browser-Erweiterung können Nutzer*innen sich über ihre persönlichen Sozialen Netzwerke informieren. Den Nutzer*innen werden so Einblicke in die Hetero- bzw. Homogenität ihrer Informationssphären gegeben. Nutzer*innen sollen durch die Interaktion mit den eigenen Daten für die Thematik der Filterblasen sensibilisiert werden. Aktuell ist die Erweiterung nur auf Twitter anwendbar. Bei der Entwicklung wurde aber bereits darauf geachtet, dass alles modular und flexibel programmiert wurde, sodass weitere Services zukünftig ergänzt werden können. Im nächsten Schritt soll die offene Plattform Mastodon ergänzt werden.

Die unter offenen Lizenzen veröffentlichten Module sollen andere Entwickler*innen dazu befähigen, ebenfalls privatsphären-zentrierte Anwendungen zu entwickeln, welche auch komplexe und langwierige Prozesse auf den Systemen der Nutzer*innen durchführen. Zu diesem Zweck haben wir z.B. ein spezielles Queue-Management Modul entwickelt, welches auch über den Verlust der Internetverbindung oder den Neustart des Rechners hinweg einen Ablauf an Aufgaben abarbeiten kann.

Während wir in der Vergangenheit schon diverse web-basierte Anwendungen konzipiert und umgesetzt haben, war die Anwendung dieser Kenntnisse und Fähigkeiten auf die Entwicklung einer Browser-Erweiterung eine neue Herausforderung.

Kurze Darstellung der Arbeiten, die zu keiner Lösung geführt haben

In der initialen Projektbeschreibung war vorgesehen, dass Methoden des Natural-Language-Processing (NLP) in Kombination mit Ansätzen des maschinellen Lernen eingesetzt werden. Gleichzeitig haben wir versucht, uns an den Prinzipien des ethischen Designs zu orientieren und ethische Konsequenzen der entwickelten Technologien zu antizipieren. Während der Implementierung von algorithmischen Vergleichen von Textähnlichkeiten, innerhalb nutzer*innen-

spezifischer sozialer Netzwerke, wurde deutlich, dass dies negative Effekte haben könnte. Ein Problem sozialer Netzwerke ist die Verfolgung von einzelnen Individuen. Um dieser Verfolung entkommen zu können, löschen Nutzer*innen ihre Accounts und erstellen unter einem neuen Pseudonym einen neuen Account. Über die automatisierte Analyse von Textähnlichkeiten über eine große Anzahl an Accounts hinweg, wäre es mit großer Wahrscheinlichkeit möglich die neuen Pseudonyme zu identifizieren. Als Alternative, um weitere Erkenntnisse über soziale Netzwerke zu ermöglichen, wurden weitere Netzwerkanalysefunktionen implementiert.

Außerdem haben wir schlussendlich die Verwaltung der Service-Anbindungen nicht über ein öffentliches Repository geregelt, sondern die Services direkt in das Plugin verbundlet. Da die Module für die Service-Anbindung direkt mit Internet kommunizieren müssen, war uns die Gefahr zu groß, das Nutzer*innen sich Schadcode installieren. Deshalb werden etwaige Service-Erweiterungen von uns nun erst evaluiert und dann in das Bundle integriert.

Kurze Angabe von Präsentationsmöglichkeiten für mögliche Nutzer*innen

- Es gibt eine Projektwebsite: http://www.ultrapop.de
- Der Code befindet sich in einem GitHub-Repository: https://github.com/ultrapop-de
- Die Browser-Erweiterung wurde in den Erweiterungs-Store für Firefox & Chrome geladen und wird dort aktuell noch überprüft und wird bald verfügbar sein.
- Außerdem schreiben wir aktuell einen Artikel für den Launch

Kurze Erläuterung zur Einhaltung der Arbeits- und Kostenplanung

Die initial aufgestellte Kostenplanung konnte eingehalten werden. In der Arbeitsplanung wurden die Pakete des NLP in die Netzwerkanalyse überführt.

Kurze Darstellung von etwaigen Ergebnissen bei anderen Stellen

Seid der Beantragung des Projektes hat Microsoft angekündigt, dass der Edge Browser zukünftig die Chromium Engine nutzen wird. Dank dieser Entwicklung, konnten wir die Browser-Erweiterung auch für den Edge Browser verfügbar machen.