mediactor
Media Impact Monitor – Medienwirkungsanalyse für zivilgesellschaftliche Akteure

Schlussbericht

Zuwendungsempfänger:

David Pomerenke

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01IS24S15 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor.

Kurze Darstellung der Aufgabenstellung und Motivation

Was war Deine Motivation? Welches Problem wolltest Du mit Deinem Projekt lösen? Wie war die geplante Vorgehensweise zur Problemlösung (auch Angabe der wichtigsten Meilensteine)?

Der Media Impact Monitor soll ein neuartiges Tool sein, das zivilgesellschaftlichen Akteuren ermöglicht, ihre Wirkung auf den gesellschaftlichen Diskurs zu messen und zu visualisieren. Durch die Analyse verschiedener Medienquellen, einschließlich lokaler und nationaler Zeitungen, sozialer Medien und parlamentarischer Debatten, kann ermittelt werden, wie sich Aktionen von beispielsweise Protestbewegungen auf den Diskurs auswirken. Dadurch können diese ihre Aktionen quantitativ und qualitativ evaluieren und zielgerichtet anpassen, um Aufmerksamkeit auf ihre gesellschaftlichen Anliegen zu lenken.

Zivilgesellschaftliche Akteure möchten oft mehr öffentliche Aufmerksamkeit auf wichtige Themen lenken, um gesellschaftliche Veränderungen anzustoßen. Dafür veranstalten sie z. B. Podiumsdiskussionen, veröffentlichen Beiträge auf Social Media, verschicken Pressemeldungen, organisieren Demonstrationen, üben zivilen Ungehorsam oder setzen Kunstaktionen um. Zur Evaluation ihrer Arbeit benutzen die Organisationen oft "press hits" als Metrik – aber das ist ein schlechter Ersatz für die breitere Wirkung auf die Themensetzung im Diskurs, die sie eigentlich gerne wissen möchten.

Diese Wirkung auf die Themensetzung im Diskurs ist statistisch schwer zu fassen, da der Diskurs, die Aktionen und externe Ereignisse in einem komplexen kausalen Zusammenhang stehen. Eine korrelationale und selbst eine einfache Regressionsanalyse würden daher zu stark verfälschten Ergebnissen führen.

Hier soll der Media Impact Monitor ansetzen: Er soll mittels innovativer Causal Inference-Methoden quantifizieren und visualisieren, wie sich verschiedenste Aktionen auf die Themensetzung im öffentlichen Diskurs auswirken. Konkrete Beispiele: Die "Letzte Generation" kann damit zum Beispiel sehen, ob ihre Aktionen vom Thema ablenken oder zur konstruktiven Diskussion beitragen; und NGOs wie Greenpeace können evaluieren, welche ihrer Formate medial am effektivsten sind.

Technische Umsetzung:

App: Die Daten sollen in einem Python-Backend mit joblib-Pipelines verarbeitet und mit Flask und gunicorn per API zugänglich gemacht werden. Das Frontend soll auf React basieren und die Daten mittels D3 visualisieren. Daten von Nutzenden sollen lokal in deren LocalStorage gespeichert werden.

Daten: Als Datengrundlage für den öffentlichen Diskurs soll in erster Linie das Open Data Projekt MediaCloud dienen, das eine API für den Inhalt sämtlicher Online-Zeitungen bereitstellt. Daten zu Print-Zeitungen sind über das deutsche Referenzkorpus DeReKo verfügbar. Protestereignisse sollen standardmäßig aus dem laufend kuratierten ACLED-Datensatz übernommen werden.

Methoden: Die Causal Inference Methoden sind in den Bibliotheken DoWhy/EconML und tfcausalimpact implementiert. Die Anwendung der Methoden für die Abschätzung der Wirkung von Ereignissen habe ich in meiner Masterarbeit detailliert beschrieben und evaluiert; auf diesem Code möchte ich aufbauen.

Zahlreiche kommerzielle Tools helfen Unternehmen, die Reichweite ihrer Marke und assoziierter Schlagworte zu evaluieren (bspw. Meltwater Media Intelligence, Talkwalker Media Monitoring, Onclusive Media Impact Analysis). NGOs möchten aber ihre Wirkung auf den generellen gesellschaftlichen Diskurs (z. B. Klimadiskurs, Rassismusdiskurs) messen. Für kommerzielle Anbieter ist damit nicht viel Geld zu machen. Daher soll mit dem Projekt eine freie und offene Lösung entwickelt werden.

Ursprünglich geplante Meilensteine:

Phase 1: MVP

- Statische Visualisierung der schon gesammelten Daten: zeitliche Darstellung des Diskurses; Wirkung von Ereignissen auf den Diskurs; geografische Verteilung. (~1 Monat)
- Pipelines für laufende Einbindung aktueller Daten von ACLED und MediaCloud. (~0,5 Monate)
- Interface, mit dem Nutzende eigene Aktionen und Kampagnen definieren und anhand eigener Kriterien evaluieren können. (~1 Monat)

Phase 2: Erweiterungen

- Einbindung weiterer Quellen: Nationale Zeitungen, soziale Medien, parlamentarische Debatten. (~1 Monat)
- Erweiterte Analysemöglichkeiten: Topic Models, kontextuelle Sentimentanalyse. (~1 Monat)

- Verbesserung der Causal Inference Techniken: doubly robust estimation, double machine learning. (~0,5 Monate)
- Weitere nutzerzentrierte Iterationen und Puffer. (~1 Monat)

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

Wer ist die Zielgruppe für Deine Lösung? Wie profitiert sie von den Ergebnissen Deines Projekts? Welche Bezüge gibt es zu den Themenfeldern und Zielen des Software Sprints?

Zielgruppen sind vielfältige zivilgesellschaftliche Akteure, vor allem NGOs, soziale Bewegungen, Protestbewegungen, Bürgerinitiativen, und Stiftungen. Durch die Forschung zu Klimaprotesten besteht Kontakt mit dem Social Change Lab, Future Matters, GermanZero, Umweltinstitut München, Fridays for Future, Letzte Generation, Extinction Rebellion, und Ende Gelände. Mit diesen soll das Produkt agil und nutzerzentriert entwickelt werden.

Es sollen möglichst auch Akteuere außerhalb der Klimabewegung einbezogen werden, beispielsweise Organisationen aus der Antirassismusbewegung oder der Tierschutzbewegung.

Das Tool soll vom Social Change Lab (socialchangelab.org) auf deren Website gehostet werden und über deren Newsletter, Twitter-Kanal, etc. beworben werden. Das Social Change Lab verfügt über ein Netzwerk mit Organisationen, die großes Interesse an Forschung und Evaluation ihrer Arbeit haben, also das Tool gut gebrauchen können.

Der Media Impact Monitor passt daher insbesondere in das Themenfeld "Civic Tech" des Software Sprint, sowie ebenfalls zum Thema "Data Literacy", indem er komplexe Datenanalysen für zivilgesellschaftliche Akteure aufbereitet.

Ausführliche Darstellung der Ergebnisse

Welche konkreten Ergebnisse hast Du erzielt? Konnten alle Meilensteine erreicht werden? Welche zusätzlichen Erkenntnisse hast Du aus der Projektarbeit gewonnen, auch im Hinblick auf die Begleitung durch die Open Knowledge Foundation?

Es wurde wie geplant eine interaktive Datenanalyse-Website umgesetzt und unter https://mediaimpactmonitor.app veröffentlicht.

Umsetzung der Meilensteine:

Phase 1: MVP

- Statische Visualisierung der schon gesammelten Daten: umgesetzt; auf die geografische Visualisierung wurde verzichtet, da sich in den User-Interviews kein Bedarf dafür gezeigt hat.
- Pipelines für laufende Einbindung aktueller Daten von ACLED und MediaCloud: umgesetzt.
- Interface, mit dem Nutzende eigene Aktionen und Kampagnen definieren und anhand eigener Kriterien evaluieren können: nicht umgesetzt; dies wäre sehr

aufwendig (kein vorheriges Caching möglich; ggf. User-Accounts notwendig), und hat sich in den User-Interviews als relevant aber nicht vorrangig herausgestellt.

Phase 2: Erweiterungen

- Einbindung weiterer Quellen: Print- und Online-Zeitungen wurden eingebunden (bei den Print-Zeitungen fehlen allerdings noch die Volltexte), sowie Google Trends Daten. Social Media hat sich aufgrund starker Restriktionen bei den API-Zugängen als zunächst nicht umsetzbar herausgestellt. Parlamentarische Daten wurden exploriert, aber bisher nicht eingebunden.
- Erweiterte Analysemöglichkeiten: Topic Models, kontextuelle Sentimentanalyse: umgesetzt; statt traditionellem Topic Modelling wurde ein LLM-basierter Ansatz gewählt.
- Verbesserung der Causal Inference Techniken: nicht umgesetzt wegen Zeitmangels.
- Weitere nutzerzentrierte Iterationen und Puffer: umgesetzt; so wurde bspw. als zusätzliches Feature ein Pressespiegel zu einzelnen Events; und die Charts und dazugehörigen Erklärtexte wurden teils in mehreren Iterationen verbessert.

Zusätzliche Erkenntnisse:

- Bisher ist nur die einfachste Causal Inference Methode implementiert, nämlich Time Series Regression. Es hat sich dabei gezeigt, dass diese weniger verlässlich ist als gedacht. Als nächstbessere Alternative sollte Interrupted Time Series implementiert werden, aber die dafür angedachte Bibliothek hat sich als ungeeignet herausgestellt, und eine eigenen Implementierung bedingt Hyperparameter Tuning, welches wir zunächst zurückgestellt haben. Beste absehbare Lösung ist die Implementierung von Synthetic Control oder eines spezielleren Modells, wurde wegen Zeitdruck aber noch nicht umgesetzt.
- Es gibt ein erhebliches Interesse an Social Media, insbesondere TikTok; gerade weil es sehr schwierig ist, dazu Daten zu bekommen.
- Durch Superbloom haben wir gelernt, schwierige Entscheidungen teils über User-Tests statt über Diskussionen anzugehen.
- Die OKF / das PTF Team hat uns insbesondere durch Hinweise auf Anschlussförderung und relevante Netzwerk-Veranstaltungen geholfen, sowie durch den Austausch mit den anderen Projekten und den Demo Day.

Zielgruppe, Nutzen und mögliche Weiterentwicklungen

Welcher Nutzen ergibt sich für die Zielgruppe aus den Ergebnissen Deines Projekts? Welche weitergehenden Effekte ergeben sich aus der Open-Source-Stellung der Ergebnisse? Gibt es Ideen für die Weiterentwicklung Deiner Lösung und Pläne zu deren Umsetzung?

Zivilgesellschaftliche Organisationen (NGOs und Protestgruppen) können durch unser Tool ihre Arbeit besser evaluieren, kontraproduktive Strategien identifizieren und korrigieren, und effektive Strategien priorisieren. Für die interessierte Öffentlichkeit werden das Protestgeschehen in Deutschland und dessen Wirkung vermittelt.

Durch die Open-Source-Stellung ergibt sich die Eingliederung in ein offenes Ökosystem (ACLED, Media Cloud) und die Wahrnehmung als gemeinnütziges Projekt. Prinzipiell können

Aktivistinnen sich auch direkt an der Weiterentwicklung beteiligen (einige Organisationen haben eigenes Personal für das Media Monitoring, teils mit Coding-Skills).

Das Projekt soll als kleines Startup weitergeführt werden. Zwei Funding-Bewerbungen waren nicht erfolgreich, es gibt noch mindestens zwei weitere mögliche Funding-Partner, und andernfalls die direkte finanzielle Kooperation mit den zivilgesellschaftlichen Organisationen. Sollte alles scheitern, kann das Projekt immerhin weiter laufen und immer aktuelle Daten aufbereiten; für Maintenance sind auch ehrenamtliche Ressourcen vorhanden.

Hat die Arbeit in dem Projekt Dich in Deiner persönlichen, fachlichen Weiterentwicklung unterstützt?

Durch die Arbeit haben wir gelernt, als unabhängiges Team gut zusammenzuarbeiten, und dadurch grundlegende Skills für ein Startup erworben. Dabei sind insbesondere die Skills zur Planung und Kooperation und Strategie hervorzuheben, wohingegen wir auf technischen Skills ja schon aufbauen konnten.

Kurze Darstellung der Arbeiten, die zu keiner Lösung geführt haben

Gab es Arbeiten bzw. Lösungsansätze, die nicht weiter verfolgt wurden? Was waren die Hintergründe, und wie bist Du alternativ vorgegangen?

Es wurde an der Einbindung parlamentarischer Daten zu Debatten, sowie zu politischen Prozessen gearbeitet. Insbesondere die Daten zu politischen Prozessen sind sehr interessant, da das Ziel der Akteure oft ein politisches ist. Allerdings hat es sich dabei um ein zusätzliches Datenformat gehandelt. das separat hätte visualisiert werden müssen, daher wurde das nicht weiter integriert.

Kurze Angabe von Präsentationsmöglichkeiten für mögliche Nutzer

Wo können sich Interessenten detailliert über Deine Projektergebnisse informieren (z.B. Webseite, GitHub, Veröffentlichungen)?

https://github.com/SocialChangeLab/media-impact-monitor

https://mediaimpactmonitor.app

Kurze Erläuterung zur Einhaltung der Arbeits- und Kostenplanung

Gab es im Projektverlauf Ereignisse, die eine Anpassung der Planung erforderlich machten – z.B. Mehr- oder Minderaufwand bei der Bearbeitung von Teilaufgaben?

Das Projekt wurde nicht alleine, sondern im Team mit Lucas Vogel (Frontend-Design und -Entwicklung) und Lennart Klein (Data Science) umgesetzt. Dadurch hat sich eine neue Aufgabenverteilung ergeben. Auch der Scope wurde nach den anfänglichen User-Interviews nochmals neu erarbeitet und priorisiert, es haben sich aber (außer den bei den Meilensteinen beschriebenen Anpassungen) keine größeren Änderungen ergeben. Negativ ist anzumerken, dass zu Projektende keine verlässliche Causal Inference Methode implementiert ist, aufgrund oben beschriebenen Mehraufwands, sowie einem größeren Fokus auf Frontend-Entwicklung. Das soll bald behoben werden.

Kurze Darstellung von etwaigen Ergebnissen bei anderen Stellen

Gab es Entwicklungen anderer Personen oder Institutionen, die Einfluss auf Deine Arbeiten und die Zielsetzung hatten? Wenn ja, worin bestand dieser und wie bist Du damit umgegangen?

Es wurden 2024 mehrere relevante wissenschaftliche Publikationen zur Wirkung von Protesten veröffentlicht, teils auch mit Fallstudien zur deutschen Klimabewegung, was Fokus dieses Projekts war. Deren Methoden und Daten verbessern unseren Ansatz teilweise leicht, und wir werden das in Zukunft noch berücksichtigen. Es hat sich auch sowohl ein Interesse als auch das Vorliegen von Daten zur öffentlichen Meinung und politischen Präferenzen herausgestellt, auch das wäre eine spannende Ergänzung.