



Liquid Democracy

# Gemeinwohlorientierte KI für demokratische Öffentlichkeiten

Beispiele aus der Beteiligungs-Praxis

Maria Lederer

Liquid Democracy e.V.

04. OKTOBER. 2024

# Agenda

1. Liquid Democracy & Adhocracy
2. KI & Online-Beteiligung
3. Praxis-Einblick KOSMO
4. Praxis-Einblick DeFaktS
5. Diskussion

1

# Liquid Democracy & Adhocracy

Unsere Vision ist eine demokratische Kultur, in der Mitgestaltung für jede:n selbstverständlich ist.





adhocracy+

adhocracy.plus



meinBerlin

mein.berlin.de



Deutscher Bundestag

enquetebeteiligung.de



# Ideen einreichen

<https://adhocracy.plus/werder-havel>

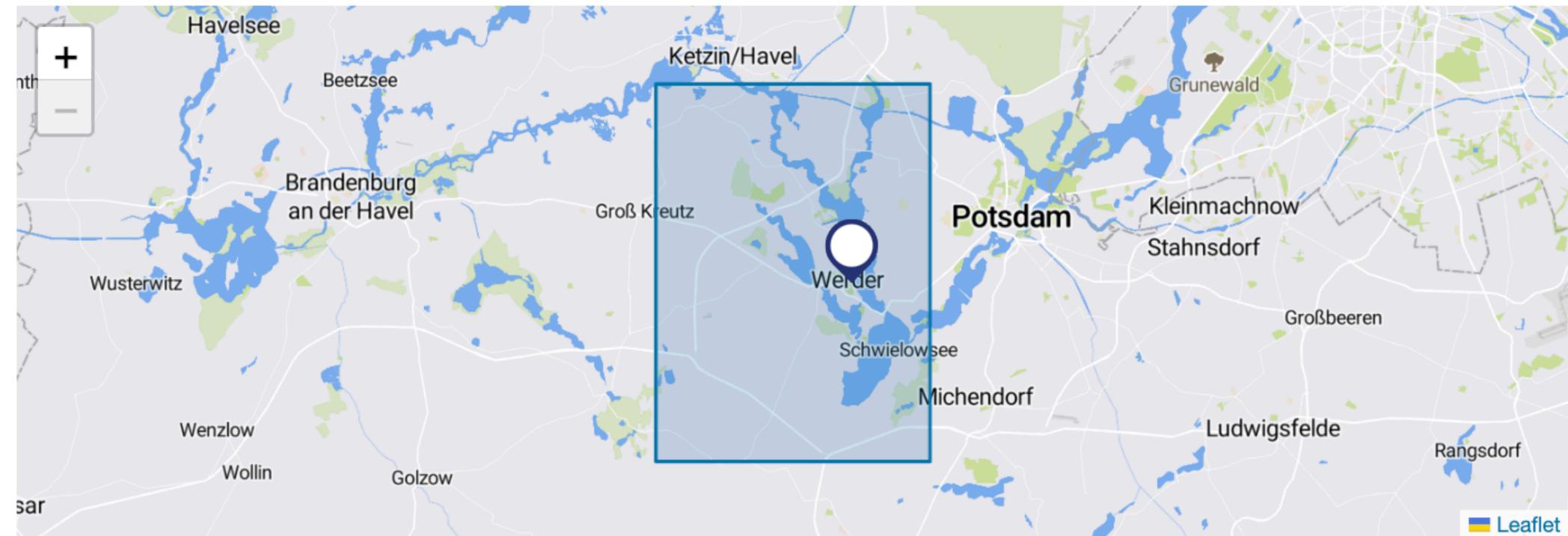
← ZURÜCK

## Einbahnstraße und Radweg in der Potsdamer Straße

Fahrradweg

📍 Potsdamer Straße

Die Potsdamer Straße ist für Fahrradfahrer besonders unangenehm, der Verkehr ist sehr stark und es gibt nur wenig Platz für Autos zum Überholen. Ein Einbahnstraßensystem für Autos in Kombination mit einem breit ausgebauten Radweg würde die Situation sehr erleichtern.



AnnaF 14. Mai 2020

Referenznr.: 2020-01430

# Ideen diskutieren

<https://adhocracy.plus/werder-havel>



Blüte

4. Juni 2020, 12:50



Die Mehrheit der Autofahrer würde sich stark gegängelt fühlen, wenn Einbahnstraßenverkehr droht und sie nicht befragt werden. Sie sind es, die die Radwege mitfinanzieren und sind keine schlechten Menschen, mitunter beides. Keine Autofahrer = keine Radwege! Am besten ist es wie bisher, wenn sich alle den Verkehrsraum teilen - und zwar gleichberechtigt. Die Verschwendung von Verkehrsraum durch...

Weiterlesen...

^ 0 v 0

[Antworten verbergen](#) [Teilen](#) [Melden](#)



Nonita

Letzte Änderung am 4. Juni 2020, 14:50



Nicht zeitgemäß ist, das Auto als einzige Lösung zu sehen! Wer sagt denn, dass die Radfahrer keine Autofahrer sind und nicht selbst die Straßen mitfinanzieren? Die Radfahren tragen außerdem dazu bei, dass die Straßen leerer sind und weniger CO2 ausgestoßen wird, als wenn sie auch immer mit dem Auto führen. Weiterhin ist es wichtig mehr an die Zukunft zu denken. Wenn Kinder bereits lernen, dass ma...

Weiterlesen...

^ 0 v 0

[Teilen](#) [Melden](#)



Ungerath\_SMG

4. Juni 2020, 15:55



2

# KI & Online-Beteiligung

Vorstellung + Fragerunde

**Was versteht ihr/verstehen Sie unter  
„gemeinwohlorientierter KI“?**

Vorstellung + Fragerunde

**Was versteht ihr/verstehen Sie unter  
„gemeinwohlorientierter KI“?**

**Wie unterscheidet sie sich von „for-profit-AI“ ?**

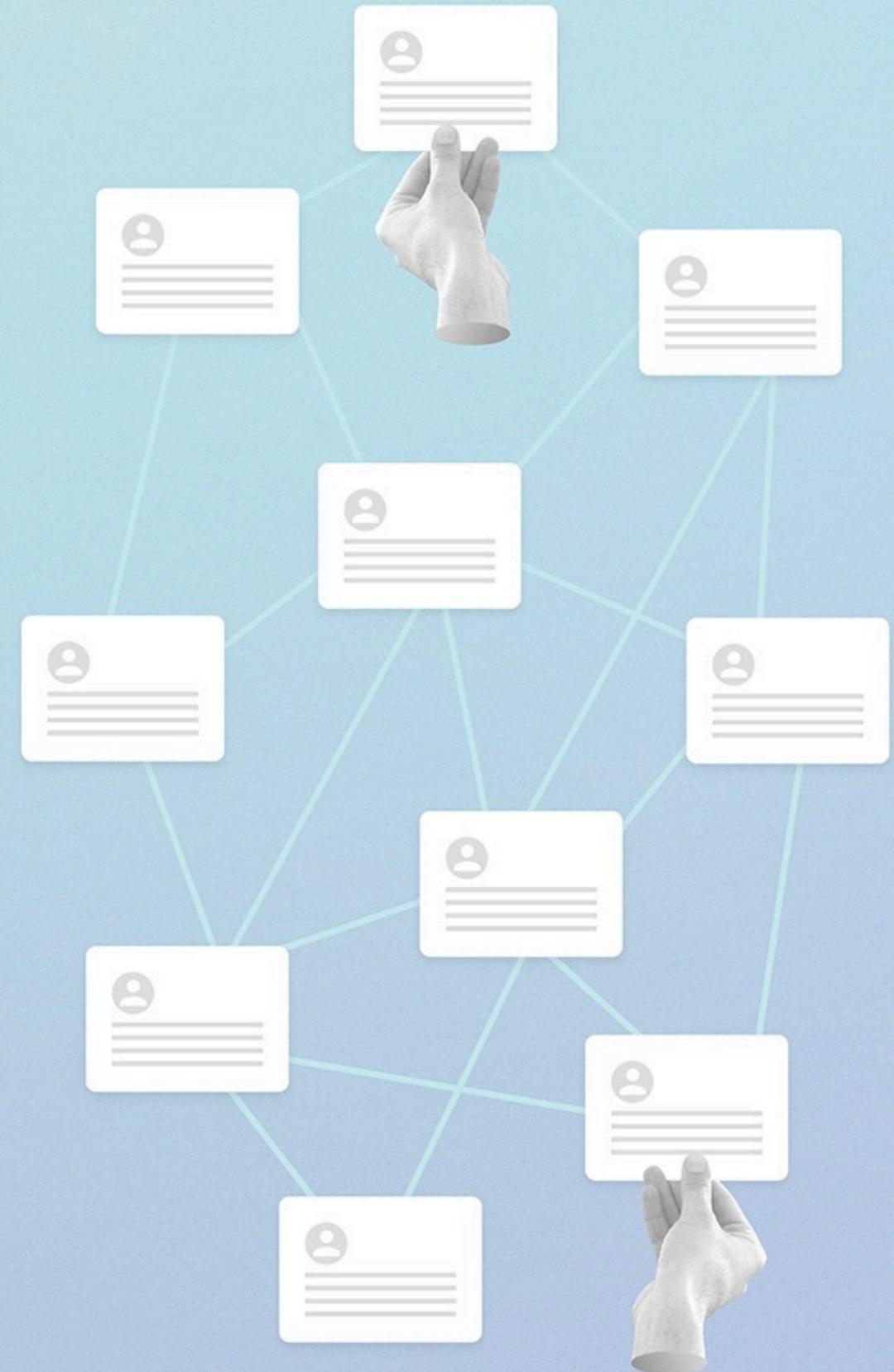
# Gemeinwohlorientierte KI

Idee

KI-Systeme sollen so zum Einsatz kommen, dass sie einem kollektiven gesellschaftlichen Wohl dienen.

Voraussetzungen

- Bewusstsein für Kontext und Öffentlichkeit
- Problem- und Bedarfszentriertes Denken als Grundhaltung
- Öffentliche, NICHT profitorientierte Rechtfertigung für die Nutzung von KI
- Grundsatz der Gleichberechtigung von Menschen



# KI & Liquid Democracy

Wie?

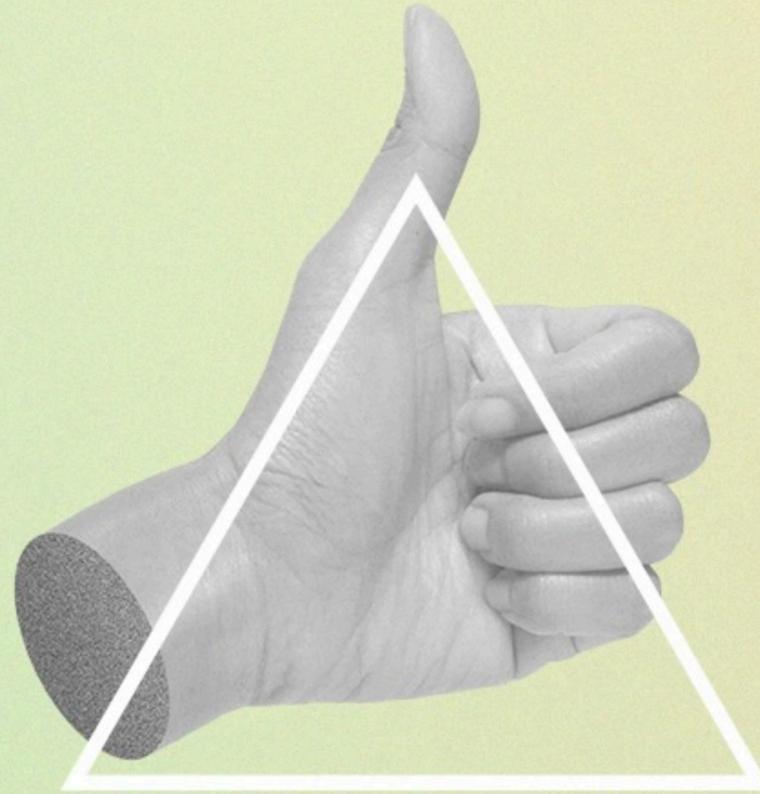
- Erproben & Erforschen von KI-Einsatz im Beteiligungskontext mit Projektpartnern
- ständiger Lernprozess für uns als Verein
- Einsatz von KI birgt sowohl Gefahr als auch Potential für Demokratie
- Besondere Anforderungen an gemeinwohlorientierte KI-Entwicklung

Was?

Forschungsprojekte: KOSMO & DeFaktS

Zentrale Frage:

**Kann KI unsere Diskurse verbessern?**



3

# Praxis-Einblick KOSMO

# KOSMO

KI-gestützte, kollektiv-soziale Moderation von Onlinediskussionen

Partner

Liquid Democracy, Heinrich Heine Universität Düsseldorf (HHU) & Institut für partizipatives Gestalten (IPG)

Ziel

Verbesserung von Online-Diskussionen als relevanten Moment der demokratischen Meinungsbildung

Ergebnis

Praxisnah entwickelter Prototyp, basierend auf der OS-Software Adhocracy



KOSMO

# KOSMO-KI

KI erkennt schädliche und konstruktive  
Diskussionsbeiträge sowie Tatsachenbehauptungen

Moderation kann entsprechend auf Beiträge  
reagieren

Grundsatz: KI für die Moderation & human-in-the-  
loop

- KI soll bei der Moderation unterstützen
- KI soll keine autonomen  
Moderationsentscheidungen treffen
- KI soll Menschen in Moderation nicht ersetzen



# Prototyp

Meldungen

Schädlich

Tatsachenbehauptung

Konstruktiv

Moderationsoptionen

Kommentar blockieren

Feedback geben

Kommentar hervorheben

The screenshot displays the KOSMO moderation interface. At the top left is the KOSMO logo with a 'BETA' tag. The top right shows a language selector set to 'de' and a user profile for 'Ricky Admin' with the text 'Guten Tag Ricky Admin'. Below the header are two tabs: 'ÜBERSICHT' and 'MODERATION', with 'MODERATION' being the active tab. The main content area is split into two columns. The left column features a user profile for 'Ricky Admin', including a circular profile picture, the name 'Ricky Admin', and the text 'Seit 23 Apr 2021'. Below the profile is a 'Mehr Info' button. The right column shows a navigation arrow '← Zurück' and the title of a post: 'KI IN DER VERWALTUNG (EINES DEMOKRATISCHEN STAATES)'. Below the title is the text 'von Praxis-Tests'. There are three filter and sorting controls: 'Filter: Alle Kategorien', 'Filter: Ausstehend', and 'Sortierung: Neueste Meldung'. A notification box indicates that 'Karl Kritik' has reported a post, with a 'SCHÄDLICH: 1' badge and a timestamp of '21. März 2023, 10:36'. The reported post text reads: 'Warum brauchen wir hier eine technische Unterstützung? Unsere Demokratie funktioniert seit jeher gut und ist eine eingespielte Sache. Da fände ich es vollkommen bescheuert so etwas unerprobtes wie eine KI einzusetzen. Die Forschung ist noch viel zu rudimentär.' Below the text are three moderation options: 'Feedback bearbeiten', 'Hervorgehoben', and 'Blockieren'. At the bottom, there is a 'Feedback der Moderation' section with the text 'Zuletzt bearbeitet am 20. März 2023, 14:24' and a question: 'Haben Sie ein Beispiel für eine Forschungsarbeit, die Sie "rudimentär" finden?'.

# Entwicklungsphasen

## Trainingsdaten

- Beispielkommentare durch Clickworker oder Research Assistents annotieren
- Moderierte Kommentare exportieren

## Vortrainierte Sprachmodelle

- Aufbauen auf vortrainierten Sprachmodellen
- Fine-tuning mit den Trainingsdaten, um relevante Kommentare zu erkennen

## Implementierung

- Loop zw. KI und Back-End
- Entscheidung der KI wird im Dashboard der Moderation angezeigt
- Next Steps: Feedback-Loop durch Moderation

# Herausforderungen

1. Was „gute“, bzw. „schlechte“ Beiträge sind ist subjektiv, durch KI nur Annäherung möglich
2. Qualitative KI braucht große Mengen an Daten (Ressourcenfrage)
3. Qualitative KI braucht passgenaue Daten (ähnliche Diskussionsformate)
4. KI-Weiterentwicklung nach Projektablauf sichern

4

# Praxis-Einblick DeFaktS

# DeFaktS

Desinformationskampagnen beheben durch  
Offenlegung der Faktoren und Stilmittel

Partner

Liquid Democracy, Forschungszentrum  
Informatik (FZI), Murmuras GmbH,  
Universität Marburg

Ziel

Desinformation mit KI erkennen & vor ihr  
warnen, um Online-Debatten zu verbessern

Ergebnis

Praxisnah entwickelter Prototyp,  
basierend auf der Open-Source-Software  
Adhocracy



# KI in DeFaktS

KI erkennt für Desinformation charakteristische Faktoren & Stilmittel

Diese KI bildet Komponente für Explainable Artificial Intelligence (XAI)

Nutzer:innen werden informiert & aufgefordert, entsprechend zu reagieren (Call to Action)

Grundsatz: KI für die Nutzer:innen, XAI-Ansatz (kein human-in-the-loop)



Ich, wenn ich online auf Desinformation treffe



Ich, wenn KI mir hilft, Desinformation zu identifizieren

# Prototyp

## Meldungen

Desinformation nach definierten Stilmitteln (z.B. sensationalistisch, emotionalisierend, pseudowissenschaftlich) erkennen

## Usertestings für XAI-Komponente

- Warnhinweis unter/über dem klassifizierten Post
- Ausführlicher Erklärungstext
- Confidence Score in Prozentzahlen
- Hervorhebung relevanter Textstellen



NebelWanderer

1/9/2020, 1:47 PM

Diese 5G-TECHNOLOGIE IST EIN KOMPLETTER BETRUG!!! Ich habe GEHÖRT und GELESEN, dass diese Strahlen direkt das HIRN UNSERER KINDER manipulieren können!! Wer weiß, welche GEHEIMEN BOTSCHAFTEN sie da reinstrahlen und unsere Jugend dazu bringen, gegen ihre eigenen Familien zu handeln! Und von wem kommt diese Technologie? AUSLÄNDISCHEN Unternehmen, die nur darauf aus sind, UNSERE KULTUR ZU ZERSTÖREN!!! Wieso lassen wir zu, dass sie diese GEFAHR direkt in unseren Schulen installieren? Das ist doch ABSICHT, um unsere nationale Identität zu untergraben und uns ALLE KONTROLLIEREN zu können! WACHT AUF, LEUTE!!! Wir sollten zurück zu den Wurzeln, zu Kreidetafeln und Büchern, und diese DESTRUKTIVE und GEFÄHRLICHE "Fortschrittstechnologie" ABLEHNEN!!!



Die defakts KI hat diesen Kommentar als Desinformation eingestuft.



Die KI ist zu **67 Prozent sicher**, dass es sich bei dem Kommentar um Desinformation handelt. Er wurde als **pseudowissenschaftlich** und als **gefälschter Inhalt** eingestuft, basierend auf den folgenden Aspekten: **emotionaler Inhalt, geringe Komplexität des Inhalts**

Eine gute Diskussion lebt von konstruktiven Beiträgen. Bitte fügen Sie Ihre Ideen hinzu. Dies können Quellen, Argumente oder Fragen zum Thema sein. [Weniger anzeigen](#)



6



2



Antworten



Teilen

# Herausforderungen

1. Was „wahre“, bzw. „unwahre“ Beiträge sind ist subjektiv, durch KI nur Annäherung möglich (kein fact checking)
2. Ansatz „no-human-in-the-loop“ braucht XAI und damit besondere Aufmerksamkeit auf UX/UI Design
3. Desinformationen werden nicht auf allen Plattformen gleichermaßen geteilt (Datengrundlage)
4. KI-Weiterentwicklung nach Projektablauf sichern

Fragen?

5

# Diskussion

Diskussion in Kleingruppen (2-3 Personen)

**Bevorzugt ihr/Bevorzugen Sie den Ansatz „human-in-the-loop“ oder reicht eine gute „XAI“ im Kontext von demokratischen Öffentlichkeiten?**

**Welche Anwendungsbeispiele seht ihr/sehen Sie in der Zivilgesellschaft?**

Liquid Democracy

# Vielen Dank!

Zum KOSMO-  
Abschlussbericht



 [liqd.net](https://liqd.net)

 @liquid-democracy

 @liquiddemocracy

 @liqd\_de

 [hallo@liqd.net](mailto:hallo@liqd.net)