

Chancen/Risiken-Bewertung des KI-Projekts *Demokratieförderquote*

Chancen

Bias/Fairness: Das Team von **AWO digital** hat eine klare, selbst entwickelte Definition von Demokratieförderung genutzt. Dadurch konnten sie den Algorithmus gezielt trainieren und nachvollziehbare Kriterien anwenden. Die transparente Offenlegung des Datensatzes und des Klassifizierers schafft Nachvollziehbarkeit und ermöglicht Dritten, die Definition zu prüfen oder zu verändern. Das Projekt trägt zu einer faktenbasierten Debatte bei und verhindert rein subjektive Einschätzungen.

Datenschutz/Sensible Daten: Fokus auf öffentliche Daten: Die genutzte Förderdatenbank enthält keine personenbezogenen Daten, sondern Förderprogramme und Budgets. Das senkt das Risiko von Datenschutzverletzungen erheblich.

Nachhaltigkeit: Effizienter Ressourceneinsatz: Das Projekt setzt keine großen KI-Modelle ein, sondern nutzt ein maßgeschneidertes NLP-Modell, das deutlich weniger Rechenleistung, Strom und Speicher benötigt als große generative Sprachmodelle. Durch den Open-Data-Ansatz und frei zugängliche Tools können Ergebnisse mehrfach genutzt werden, ohne erneut Rechenressourcen für dieselbe Aufgabe aufzuwenden.

Risiken

Bias/Fairness: Die Arbeitsdefinition von Demokratieförderung ist subjektiv. Sie basiert auf menschlichen Einschätzungen, andere Akteure könnten sie als zu eng oder zu weit gefasst kritisieren. Wenn das NLP-Modell mit Beispielen trainiert wird, die eine Schieflage enthalten, kann es systematisch falsche Zuordnungen vornehmen. Programme werden binär als Demokratieförderung oder nicht eingestuft (Ja/Nein-Logik). Komplexe Inhalte könnten falsch kategorisiert werden.

Datenschutz/Sensible Daten: Auch wenn hier keine personenbezogenen Daten betroffen sind, birgt Webscraping generell rechtliche Unsicherheiten. Falls ähnliche Projekte auf detailliertere, personenbezogene oder organisationsbezogene Daten zugreifen, könnten Datenschutzrisiken steigen.

Nachhaltigkeit: Falls das Projekt stark erweitert wird, kann der Ressourcenbedarf deutlich steigen. Häufiges Neuaufsetzen des Modells ohne klare Notwendigkeit erhöht Energieverbrauch und CO₂-Bilanz.

Maßnahmen zur Risikominimierung

Gegen Bias/Diskriminierung

1. Diverse Perspektiven einbeziehen: Definition und Trainingsdaten gemeinsam mit verschiedenen gesellschaftlichen Gruppen entwickeln.
2. Mehrstufige Validierung: Ergebnisse stets durch menschliche Prüfung absichern.
3. Mehrdimensionale Klassifizierung: Graubereiche statt reiner Ja/Nein-Logik zulassen.
4. Dokumentation und Transparenz: Kriterien, Daten-Auswahl und Modellgrenzen offenlegen.

Gegen Datenschutzrisiken

1. Rechtsprüfung bei Webscraping: sicherstellen, dass das Vorgehen rechtskonform ist.
2. Minimierung von Datensammlung: nur notwendige Daten erheben.
3. Sichere Speicherung: auch bei öffentlichen Daten Zugriffsrechte regeln.
4. Klare Lizenzierung: Open Data mit eindeutigen Nutzungsbedingungen versehen.

Für Nachhaltigkeit

1. Kleine, spezialisierte Modelle bevorzugen, ressourcensparend statt groß.
2. Modell-Updates gezielt planen, nur bei relevanten Änderungen neu trainieren.
3. Rechenressourcen grün betreiben, z. B. mit erneuerbaren Energien.
4. Wiederverwendung fördern: offene Ergebnisse und Tools so gestalten, dass Dritte sie nutzen können.

Fazit



Das Projekt **Demokratieförderquote** ist ressourcenschonend im Vergleich zu vielen anderen KI-Anwendungen, bietet hohe Transparenz und eine gute Basis für öffentliche Diskussionen. Risiken bestehen vor allem in der Subjektivität der Definition und der möglichen Ausweitung ohne Effizienzstrategie. Partizipative Definitionen, kontrollierte Modell-Updates und Open-Data-Prinzipien können Fairness und Nachhaltigkeit stärken.