

Vorlesung 24: Ethik, Bias & Verantwortung in der KI

Gruppendiskussion - Kontroverse Thesen & Standpunkte

Ablauf der Diskussion

1. **Gruppenbildung** (3-4 Gruppen à 4-6 Personen)
 2. **Vorbereitung** (15 min): Gruppen lesen ihre Ressourcen, sammeln Argumente
 3. **Diskussionsrunden** (je 10 min pro These): Strukturierte Debatte zwischen den Standpunkten
 4. **Reflexion** (5 min): Was haben wir gelernt? Wo liegen die echten Herausforderungen?
-

These 1: KI-Bias vs. Effizienz

"Fairness-Anforderungen machen KI-Systeme ineffizienter und sollten nur in besonders sensiblen Bereichen gelten."

Gruppe A: "Effizienz First"

Standpunkt: Wirtschaftlichkeit und technische Performance haben Priorität

- KI soll optimal funktionieren, Fairness-Constraints bremsen Innovation
- Markt regelt Diskriminierung langfristig selbst
- Zu viele Regulierungen verhindern technologischen Fortschritt

Gruppe B: "Ethics by Design"

Standpunkt: Fairness ist eine Grundanforderung, nicht optional

- Diskriminierung verursacht gesellschaftliche Kosten
 - Ethische KI ist nachhaltiger und vertrauensvoller
 - Kurzfristige Effizienz vs. langfristige Akzeptanz
-

These 2: Transparenz vs. Geschäftsgeheimnisse

"Unternehmen sollten ihre KI-Algorithmen vollständig offenlegen müssen, um Vertrauen und Überprüfbarkeit zu gewährleisten."

Gruppe A: "Open Source Everything"

Standpunkt: Vollständige Transparenz ist notwendig für Vertrauen

- Gesellschaft hat Recht auf Nachvollziehbarkeit
- Open Source fördert Innovation und Sicherheit
- Intransparenz ermöglicht versteckte Diskriminierung

Gruppe B: "Competitive Advantage"

Standpunkt: Geschäftsgeheimnisse sind notwendig für Innovation

- R&D-Investitionen müssen geschützt werden
 - Zu viel Transparenz hilft Angreifern (Security durch Obscurity)
 - Branchenstandards statt Volloffenlegung
-

🌀 These 3: EU AI Act - Innovation vs. Regulierung

"Der EU AI Act behindert europäische KI-Innovation und macht Europa im globalen Wettbewerb weniger konkurrenzfähig."

Gruppe A: "Regulation Kills Innovation" 🚫

Standpunkt: Überregulierung schadet dem Technologiestandort Europa

- USA und China haben weniger Restriktionen → Wettbewerbsnachteil
- Bürokratie verlangsamt Entwicklung und Deployment
- Startups können sich komplexe Compliance nicht leisten

Gruppe B: "Standards Create Trust" 🛡️

Standpunkt: Regulierung schafft Vertrauen und langfristige Wettbewerbsvorteile

- "Privacy by Design" war auch erfolgreich (DSGVO-Vorteil)
 - Standards schaffen Rechtssicherheit für Unternehmen
 - Europäische Werte als Exportschlager
-

🌀 These 4: Menschliche Entscheidung vs. KI-Automatisierung

"In kritischen Bereichen wie Kreditvergabe und Personalauswahl sollte immer ein Mensch die finale Entscheidung treffen."

Gruppe A: "Human in the Loop" 👤

Standpunkt: Menschliche Aufsicht ist unverzichtbar

- Menschen können Kontext und Ausnahmen besser bewerten
- Verantwortung braucht menschlichen Entscheider
- KI als Unterstützung, nicht als Ersatz

Gruppe B: "Algorithms Are More Fair" 🤖

Standpunkt: KI-Systeme können objektiver sein als Menschen

- Menschen haben unbewusste Vorurteile (Implicit Bias)
 - KI-Entscheidungen sind konsistenter und nachprüfbarer
 - Skalierung: Menschen können nicht alle Entscheidungen überprüfen
-

🌀 These 5: Globaler KI-Wettlauf und Ethik

"Im globalen KI-Wettlauf mit China und den USA kann sich Europa ethische Bedenken nicht leisten - Hauptsache, wir bleiben technologisch konkurrenzfähig."

Gruppe A: "Win at All Costs" 🎲

Standpunkt: Geopolitische Realität erfordert Pragmatismus

- Technologische Souveränität wichtiger als ethische Ideale
- China und USA nehmen keine Rücksicht auf europäische Werte
- Wirtschaftliche und sicherheitspolitische Abhängigkeit vermeiden

Gruppe B: "Values as Competitive Advantage" 🌱

Standpunkt: Europäische Werte sind ein Alleinstellungsmerkmal

- Vertrauen ist ein Wettbewerbsvorteil
- Nachhaltige KI ist zukunftsfähiger
- Europa als Regulierungsexporteur (Brussels Effect)

Zentrale Ressourcen für alle Thesen

Die wichtigsten Quellen (alle getestet und funktionsfähig):

1. **EU AI Act - Europäische Kommission**
 - Offizielle EU-Regulierung zu KI, alle Thesen relevant
2. **UNESCO Ethics of AI Recommendation**
 - Globaler Standard für KI-Ethik, 194 Länder
3. **Microsoft Responsible AI**
 - Industriepraxis und konkrete Tools
4. **Stanford HAI - Human-Centered AI**
 - Führende Forschung zu KI und Gesellschaft
5. **Brookings Institution - AI Policy**
 - Unabhängige Politikforschung und Analysen
6. **Fairness and Machine Learning (Online-Buch)**
 - Kostenlos, technisch fundiert, gut verständlich
7. **Plattform Lernende Systeme (acatech)**
 - Deutsche KI-Strategie und Ethik-Leitlinien

Diskussionsregeln

Für alle Gruppen:

1. **Argumente statt Emotion:** Sachlich bleiben, Quellen nutzen
2. **Stahlmann-Prinzip:** Gegnerische Position in bester Form darstellen, dann widerlegen
3. **Beispiele nutzen:** Konkrete Fälle sind überzeugender als abstrakte Theorie
4. **Trade-offs anerkennen:** Es gibt selten einfache Antworten

Moderationshinweise:

- **Zeitlimit einhalten:** Prägnante Argumente fördern
 - **Alle einbeziehen:** Leisere Stimmen aktiv einbinden
 - **Synthese suchen:** Wo gibt es gemeinsamen Boden?
 - **Realitätscheck:** Welche Position ist in der Praxis umsetzbar?
-

Lernziele der Diskussion

Nach der Diskussion sollten die Studierenden:

- **Komplexität verstehen:** Es gibt selten einfache Schwarz-Weiß-Antworten
- **Trade-offs erkennen:** Verschiedene Werte und Ziele stehen oft in Konflikt
- **Stakeholder-Perspektiven:** Verschiedene Akteure haben berechnigte, aber unterschiedliche Interessen
- **Lösungsansätze entwickeln:** Pragmatische Kompromisse zwischen idealen und machbaren Lösungen finden

Take-Home-Message: Verantwortliche KI-Entwicklung erfordert kontinuierliche Abwägung zwischen verschiedenen Werten und Zielen - es ist ein Prozess, keine einmalige Entscheidung!