

idea_lab

Das Interdisziplinäre Digitale
Labor der Universität Graz



IDEa Lab

Das Interdisziplinäre
Digitale Labor
der Universität Graz

IDea_Lab

Mission

IDea_Lab
Das Interdisziplinäre Digitale
Labor der Universität Graz



Wir
verbinden
Wissenschaft
, Wirtschaft
und
Gesellschaft

Wir arbeiten
am Puls der
Digitalen
Transformatio
n.

Wir bringen
die Zukunft
ins Jetzt.

Wir
schaffen
Raum für
innovative
Forschung.

FORSCHUNGSZENTRUM

- Data Analysis
- Ethik & KI
- Didaktik & KI
- Machine Learning

DATA LAB

- uniGPT & studiGPT¹
- Forschungsdaten Governance²
- Workshops & Events³

METHODS LAB

- Micro-Degree⁴
KI & Gesellschaft
- Konferenz:
Technology Impact Summit (TIS)
- Events³

Digitale Transformation: KI, Big Data

Interdisziplinarität

Missbrauch von Sprachmodellen zur Desinformation

DI (FH) Michael Freidl, MA

1.7.2025



Nutzungsverhalten und Informationsquellen

Social Media als primäre Nachrichtenquelle

37% der Europäer:innen nutzen hauptsächlich soziale Medien, um sich über aktuelle Nachrichten zu informieren. Diese hohe Abhängigkeit von sozialen Plattformen als Informationsquelle macht die Bevölkerung besonders anfällig für die Verbreitung von Desinformation und wenig vertrauenswürdigen Informationen.



Quelle: Eurobarometer 2023

🔍 Verbreitung wenig vertrauenswürdiger Informationen

💬 Telegram: Hotspot für Verschwörungstheorien

- **Extrem hohe Desinformationsrate:** Über 40% der geteilten Links in Gruppen, die sich zu Verschwörungstheorien austauschen, enthalten wenig vertrauenswürdige Informationen
- Telegram fungiert als besonders problematische Plattform aufgrund der geschlossenen Gruppenstrukturen und geringen Moderation

🔍 Verbreitung wenig vertrauenswürdiger Informationen

🌿 Twitter/X: Politisches Verhalten im Fokus

Unterschiede nach Regionen:

- **USA:** Republikaner:innen teilen 9 x so viele unseriöse Links als Demokraten
- **Deutschland:** CDU/CSU-Mitglieder teilen 6 x so viele unseriöse Links als SPD-Mitglieder
- **United Kingdom:** Tories teilen 5 x so viele unseriöse Links wie Labour Party

🏢 Parteienpolitische Unterschiede: Parteien rechts der politischen Mitte verbreiten deutlich mehr wenig vertrauenswürdige Informationen als andere politische Gruppierungen (Deutschland)

□ wachsende Bedrohung durch KI-generierte Inhalte

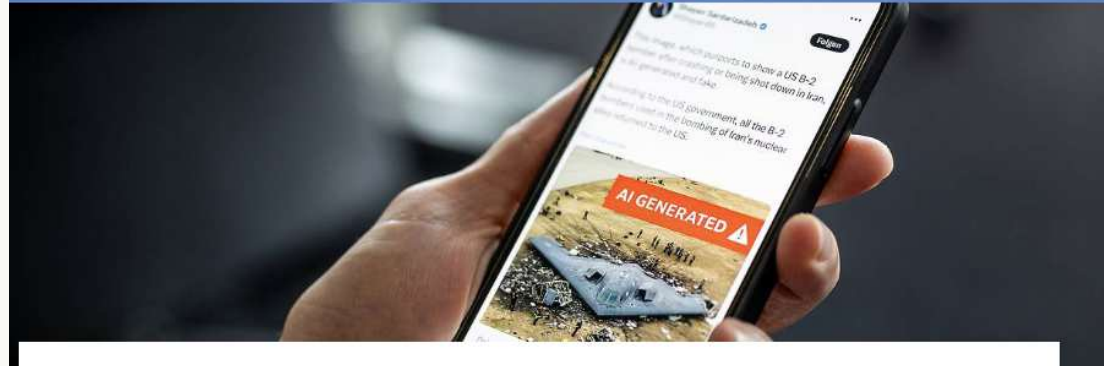
🔧 Herausforderungen bei der Quantifizierung

- Die Messung KI-generierter Desinformation gestaltet sich äußerst schwierig aufgrund von:
 - **Technische Limitationen:** Tools zur automatischen Erkennung von KI-generiertem Content funktionieren nicht zuverlässig
 - **Mangelnder Datenzugang:** Social Media-Plattformen gewähren Forschenden kaum noch Zugriff auf ihre Daten

□ wachsende Bedrohung durch KI-generierte Inhalte

📈 Konkrete Befunde

- **NewsGuard-Bericht:** 1.271 komplett KI-generierte "Nachrichtenseiten" wurden manuell identifiziert
<https://www.newsguardtech.com/special-reports/ai-tracking-center/>
- **📺 Fact-Checking-Organisationen berichten:** Zunehmende Häufigkeit von KI-generierten Videos und Bildern mit "Nachrichten" auf TikTok, YouTube und Instagram – insbesondere mit Bezug zu den Kriegen in Iran, Gaza und der Ukraine.



ISRAEL – IRAN

Krieg der KI-Bilder

Der Alltag vieler Redaktionen wird gerade von der Berichterstattung über die Eskalation des Konflikts zwischen Israel und Iran geprägt. Eine Flut von mit Hilfe künstlicher Intelligenz generierten Bildern überflutet die Nachrichtenportale. Ein Verifikationsexperte Derek Bowler von der European Broadcasting Union (EBU) beschreibt gegenüber ORF defacto, dass diese Bilder verdienen, während Userinnen und User i

24. Juni 2025, 18.15 Uhr

Der Krieg zwischen Israel und dem Iran



KI-generiert

Designing Social Media Recommendation Algorithms for Societal Good (DeSiRe)

- AIM: research alternative content recommendation algorithms for social media platforms.
- https://janalasser.at/talks/ERC_StG/

Univ.-Prof. B.Sc. Dr.rer.nat. M.Sc.

Jana Lasser

Principal Investigator

✉ [jana.lasser\(at\)uni-graz.at](mailto:jana.lasser(at)uni-graz.at)

☎ +43 316 380 - 1633

@ ORCID: [0000-0002-4274-4580](https://orcid.org/0000-0002-4274-4580)

🌐 <https://www.janalasser.at/>



Deepfakes als Cybersecurity-Bedrohung

Explosionsartiger Anstieg von Deepfake-Angriffen

- **19% mehr Deepfake-Vorfälle im ersten Quartal 2025 als im gesamten Jahr 2024** (TechAdv, AI Cyber Attack Statistics 2025)
- **Deepfakes sind nun für 6,5% aller Betrugsangriffe verantwortlich - ein Anstieg von 2.137% seit 2022** (TechAdv, AI Cyber Attack Statistics 2025)
- **Deepfake-Phishing verzeichnete einen 3.000% Anstieg in 2023, angetrieben durch die Fortschritte und Zugänglichkeit generativer KI** (Integrity360, How is AI Changing Social Engineering Attacks 2024)

Medien-wiedergabe

KI & UNIGPT IM ARBEITSALLTAG: ORIENTIERUNG UND ANWENDUNGSHILFEN

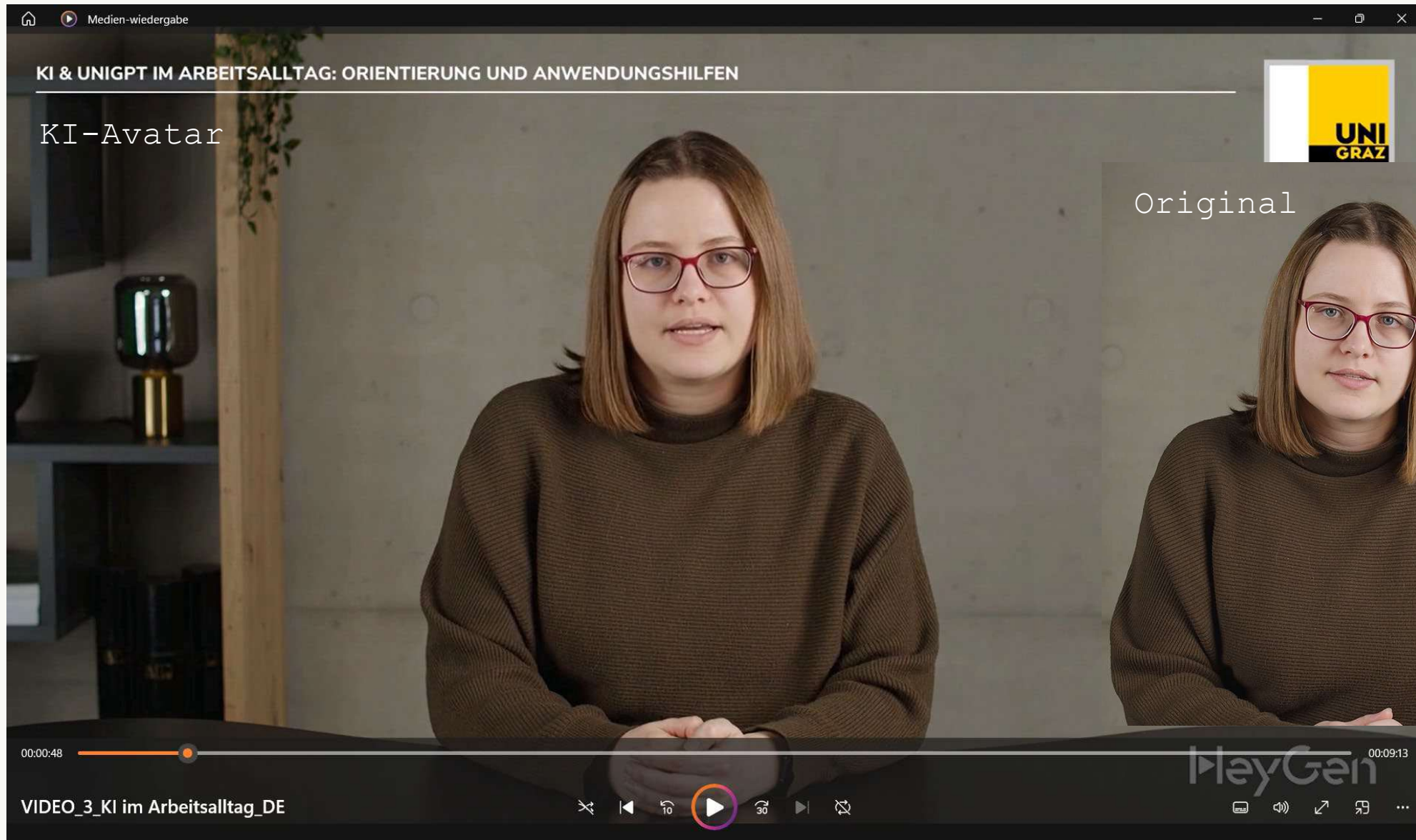
KI-Avatar

Original

00:00:48 00:09:13

VIDEO_3_KI im Arbeitsalltag_DE

PlayGen



Let's stay in touch



Kontakt

Idea_Lab – Das
Interdisziplinäre Digitale
Labor der Uni Graz



Leechgasse 34, 8010 Graz



idealab@uni-graz.at



+43 316 380 – 1640



idea-lab.uni-graz.at

