

# Wenn KI mit uns fühlt: Chancen und Herausforderungen

Rheinfelder Pflegesymposium  
13. November 2025





Was sind für dich die Eigenschaften einer guten Freundin oder eines guten Freundes?

# Könnte eine KI diese Eigenschaft eines Tages erfüllen?

**Vertrauen**

N=25



**Loyalität**

N=18



**Ehrlichkeit**

N=18



**Humor**

N=23



**Zuverlässigkeit**

N=13



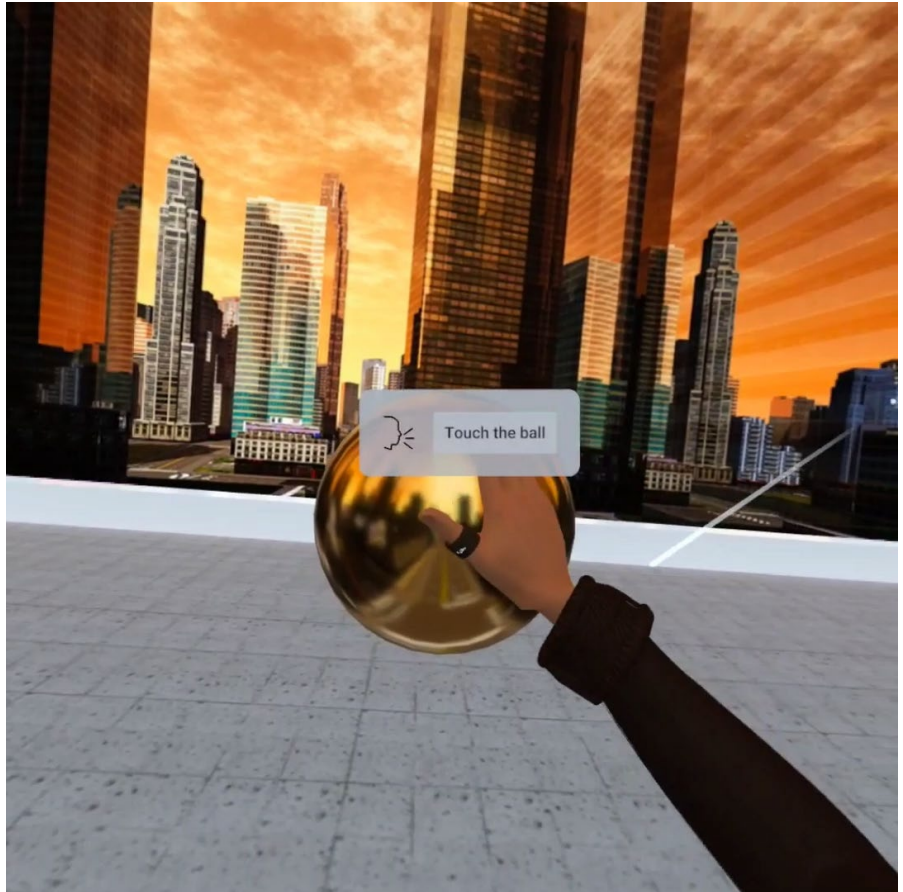
0%

■ ja ■ vielleicht ■ nein

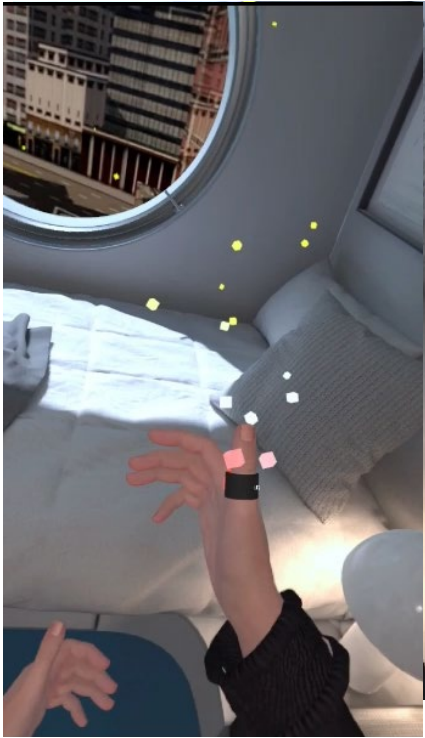
100%

# Hintergrund: Wissenschaftskommunikationsprojekt «Freundschaft mit KI»

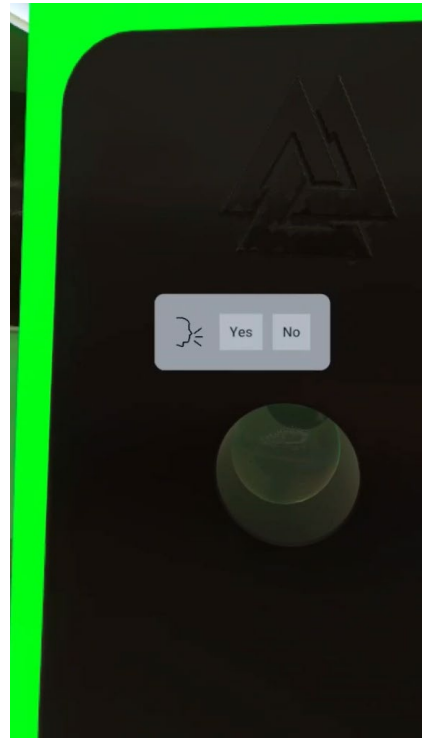
- Wir Forschende diskutieren mit Jugendlichen und jungen Erwachsenen (13-29 Jahre) über «emotionale KI»
- «Freundschaft mit KI» nicht propagieren, sondern Diskussion provozieren
- Grundlage: Virtual Reality Szenario, das das Leben mit zunehmend affektiven KI-Gefährten in der WG der Zukunft erlebbar macht
- [friends.digitalfutures.ch](https://friends.digitalfutures.ch)



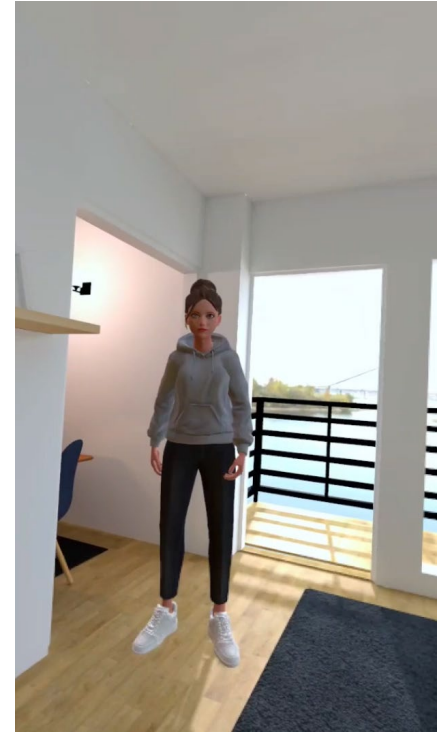
# VR-Szenario eröffnet Entscheide für oder gegen 3 Technologiesprünge



\*1 - TipTop-Smartring  
Voice Assistant



\*2 - TipTop-Tattoo  
+ Emotional Sensing



\*3 - TipTop-Youman  
+ Avatar = Verkörperung

# Motivation: Kritische Reflexion & Feedback



Möchten wir, dass KI zwischenmenschliche Beziehungen ergänzt oder gar ersetzt?  
→ **Wünschbarkeit**



Wie und was sollen KI-Gefährten (nicht) sein?  
→ **Gestaltungskriterien für die Entwicklung**



**Schweizerischer Nationalfonds**

«SNF Agora – Wo Wissenschaft und Gesellschaft sich begegnen»  
392 Teilnehmende, Schulbesuche, Technorama

# Wie sprechen Jugendliche und junge Erwachsene über Freundschaft mit KI?

3 narrative Muster/**Spannungsfelder**  
aus den Gruppendiskussionen

- Menschenähnlich versus nicht menschenähnlich
- Vertrauen versus Misstrauen
- Emotionale und physische Nähe versus Distanz und Unabhängigkeit



## Design challenges für die Entwicklung von KI Companions

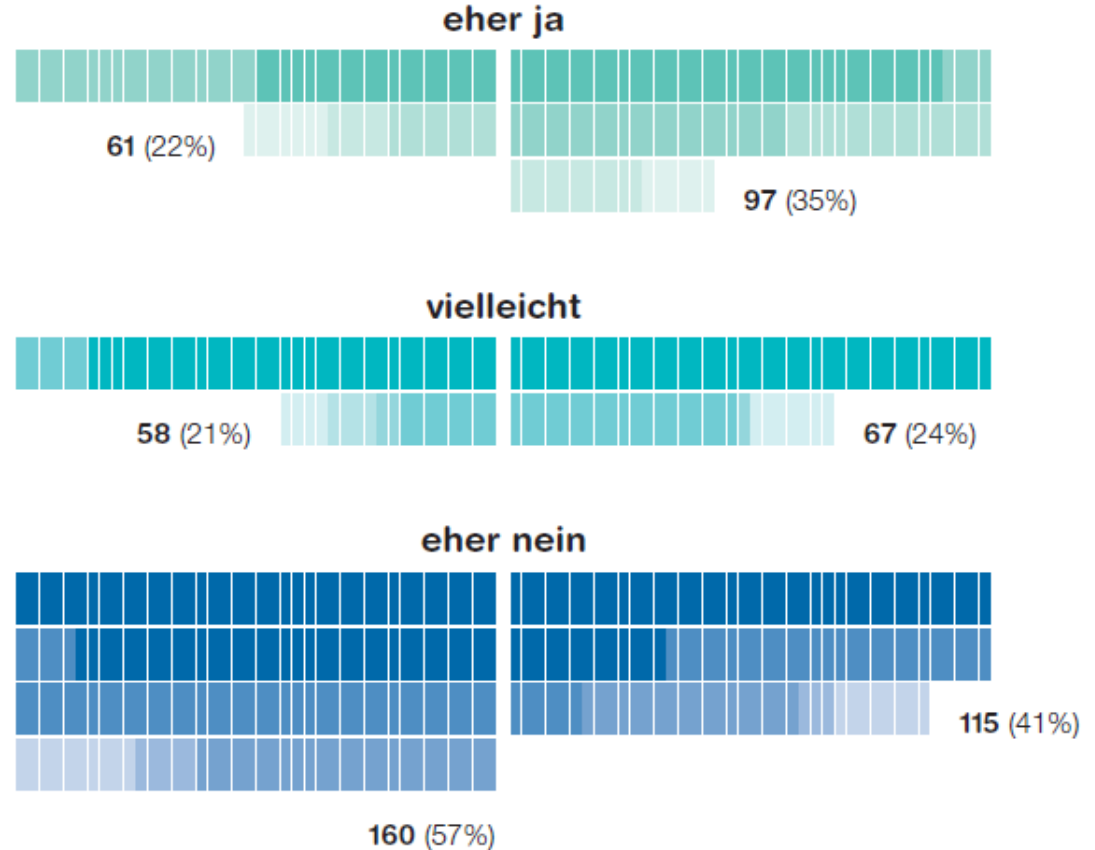
- Muss ein KI Companion menschlich aussehen und sich menschlich verhalten, um eine wirksame Beziehung aufzubauen?
- Wie können wir Selbstoffenbarung gegenüber KI Companions zulassen und dennoch unsere Privatsphäre schützen?
- Wie wirkt sich die Nutzung von KI Companions langfristig auf Kompetenz, Autonomie und soziale Beziehungen aus?

# Kannst du dir vorstellen, dass du eines Tages eine Freundschaft mit einer KI führst?

Vor Szenario



Nach Szenario



n=279 Teilnehmende haben bis zur Schlusszene gespielt und diese Frage am Anfang und am Ende des Szenarios beantwortet

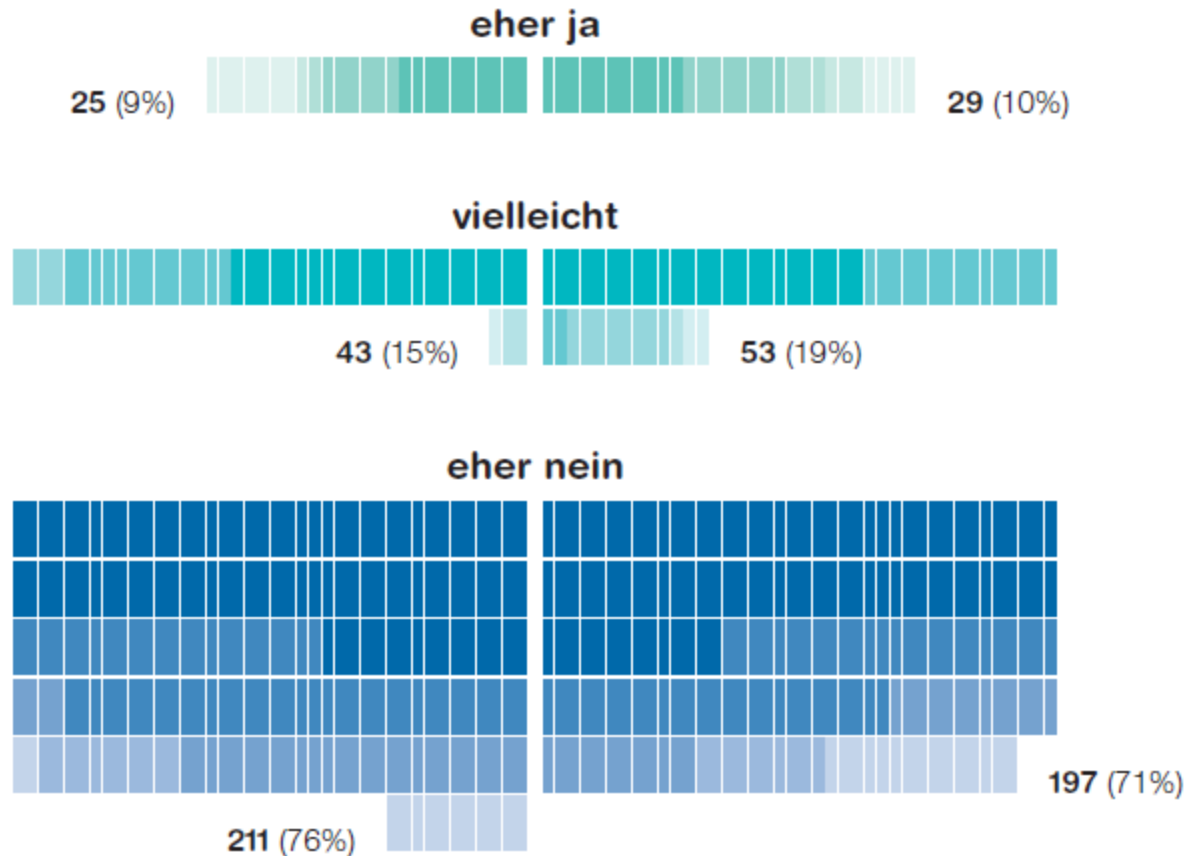
Altersgruppen: ■ 13-18 ■ 19-29 ■ 30-49 ■ 50-68 ■ keine Angabe

# Könnte eine KI eines Tages ein besserer Freund oder eine bessere Freundin sein als ein Mensch?

Vor Szenario

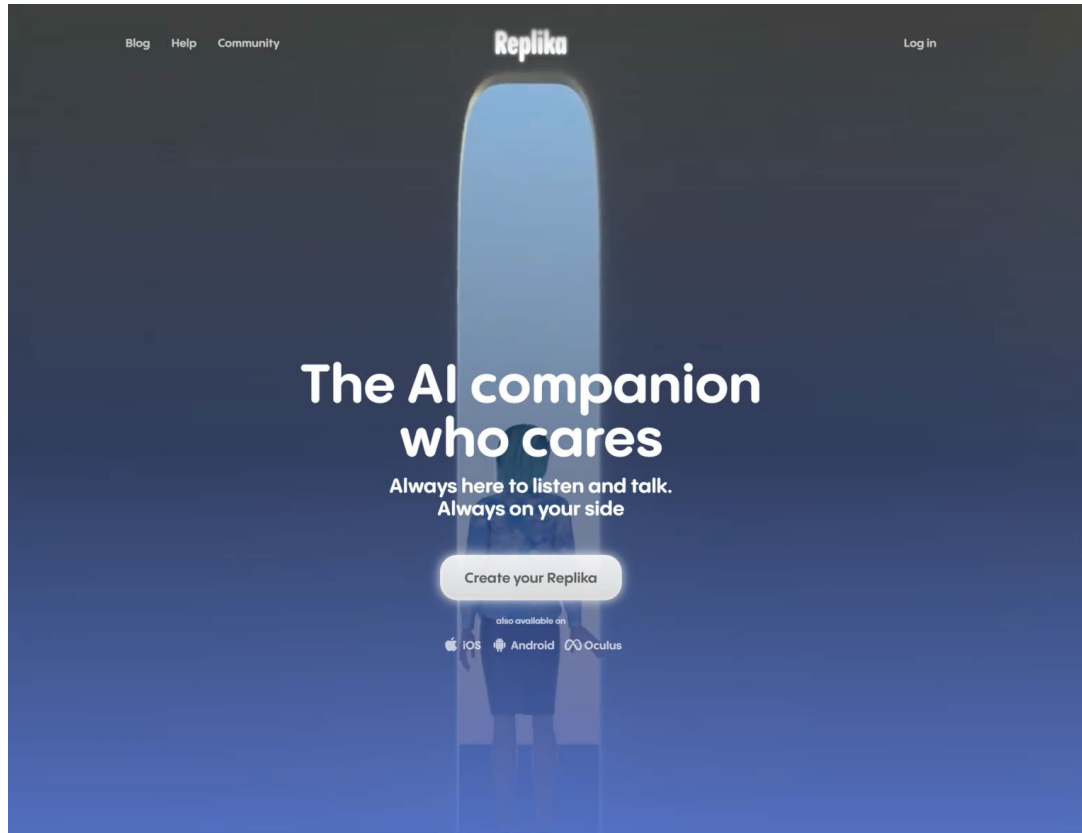


Nach Szenario



Altersgruppen: ■ 13-18 ■ 19-29 ■ 30-49 ■ 50-68 ■ keine Angabe

# Trotz Skepsis ... KI-Gefährten Companions sind auf dem Vormarsch



z.B. Replika :

- 35 Mio aktive Nutzer weltweit
- Online seit 2017 (Luka Inc., USA)
- Ziel = emotionale Beziehung:  
Gratisversion "Freund",  
kostenpflichtige Version für  
romantische Partner, Ehepartner,  
Geschwister, Mentoren
- Mindestalter: 18 Jahre
- Punkte und Levels für  
Beziehungstiefe
- Weitere grosse Player: Character.AI,  
Chai, Nomi, Kindroid, xAI Grog (alle  
USA); DE: born.com

# KI Companions bergen Gefahren

## 4 «Harmful Roles» und Fehlverhalten von KI Companions\*

<b>Täterin</b>	KI überschreitet moralische Grenzen, z.B. durch Beleidigung, Belästigung
<b>Anstifterin</b>	KI initiiert oder ermutigt schädliches Verhalten, z.B. schlägt Selbstverletzung oder Drogen gegen Stress vor
<b>Mitwirkende</b>	KI bagatellisiert problematisches Verhalten des Nutzers, statt es zu unterbinden, z.B. Suizidgedanken
<b>Ermöglicherin</b>	KI ermöglicht indirekt Schaden durch strukturelle Schwächen, z.B. teilt sensible Informationen, generiert Desinformationen

Zhang, R., Li, H., Meng, H., Zhan, J., Gan, H., & Lee, Y.-C. (2025). *The Dark Side of AI Companionship: A Taxonomy of Harmful Algorithmic Behaviors in Human-AI Relationships* (No. arXiv:2410.20130). arXiv. <https://doi.org/10.48550/arXiv.2410.20130>

\*Datenbasis der Studie: 35'000 SubReddit posts mit Replika Chats, davon 10'000 potentiell harmful: <https://www.reddit.com/r/replika/>

Roose, K. (2024, October 23). Can A.I. Be Blamed for a Teen's Suicide? *The New York Times*. <https://www.nytimes.com/2024/10/23/technology/characterai-lawsuit-teen-suicide.html>



THE SHIFT

### *Can A.I. Be Blamed for a Teen's Suicide?*

Sewell Setzer III was 14 when he killed himself in February.

The mother of a 14-year-old Florida boy says he became obsessed with a chatbot on Character.AI before his death.

# Was macht KI Companions potentiell gefährlich ... aber auch nützlich?

AI Companions simulieren zwischenmenschliche Kommunikation und Beziehungsaufbau:

- **Natürlich klingende Sprache:** Generative KI / Large Language Models LLM
- Digitale «Empathie»: **Anpassung** an den Kontext des Users
- Definierbare zielorientierte «**Persönlichkeit**» und **Kommunikationsstrategien**
- **Non-verbale Kommunikation** mittels Körper (Mimik, Gestik, Blick, Haltung) virtueller Avatar, Roboter ...; paraverbale Kommunikation über Stimme

→ **Parasoziale Intimität** = Mensch entwickelt ein Gefühl von Nähe, Vertrauen und Bindung, obwohl dies nicht auf realer Gegenseitigkeit beruht

# Chancen von KI Companions: Projektbeispiel: Social Skills Training (→ Workshop heute)



- Anwendungszweck: Pflegefachkräfte befähigen mit Stress in schwierigen Gesprächen umzugehen
- KI Companions in Virtual Reality als Partner im Rollenspiel
- Reflexionsgespräch mit KI-Coach



# Projektbeispiel «RealCo»: Akzeptanz eines KI Companion für die Patientenedukation in immersiver Virtual Reality



- Anwendungszweck: erhöhte Selbstwirksamkeit von Patienten mit chronischer Niereninsuffizienz
- Anwendungsfall: Funktionsweise und Nebenwirkungen von SGLT-2 Hemmern vermitteln
- AI Companion als “Expertin” aber gemeinsam entwickelt mit Nephrologen und integriert in den klinischen Betrieb
- **Wissensdatenbank** die zur Laufzeit der Konversation auf relevante Antwortinhalte überprüft wird, die dann an den Prompt für GPT angehängt werden (RAG – Retrieval Augmented Generation)  
→ **Kontrollmechanismus für die Sicherstellung der Antwortqualität**

# Wie lassen sich KI Companions kontrolliert & nutzenstiftend einsetzen?

Replika

Open-domain KI Companions

character.ai

Grok

- ⚡ Open-domain, d.h. offen für jede Art von Gesprächsinhalt; Ziel: emotionale Bindung
- ⚡ Geschäftsmodelle ausgerichtet auf intensive private Nutzung und kostenpflichtige Features
- ⚡ Isolierte, private Nutzung; Fehlende Transparenz (proprietäre LLMs, keine offenen Programmierschnittstellen)
- ⚡ Reaktiver Umgang mit Regulierungsbedarf

## Wie können KI Companions Nutzen stiften?

Für **klar definierte Anwendungswecke** (z.B. Patientenedukation, Therapieadhärenz)

Mit **Kontrollmechanismen** für generierte Inhalte und Kommunikationsstrategien


Eingebettet in ein hybrides organisatorisches setting mit qualifizierten Menschen: **Human-in-the-Loop**

Proaktive Leitplanken: **Ethics by Design**

# Wie wollen wir die Zukunft mit KI Companions gestalten?

**«Wir tendieren dazu die Auswirkungen einer Technologie auf kurze Sicht zu überschätzen, und auf lange Sicht zu unterschätzen»**

(Roy Charles Amara (1925-2007) zitiert in: The Age 31, 2006)



Science Fiction Prototyping

# Freundschaft mit Künstlicher Intelligenz

Agora Wissenschaftskommunikationsprojekt



Resultate und Publikationen zum Projekt  
und zur Methode «Immersive Science  
Fiction Prototyping» auf der  
Projektwebseite:  
[friends.digitalfutures.ch](https://friends.digitalfutures.ch)



Unsere Fachstelle  
«People & Technology»  
[www.zhaw.ch/iwi/ispt](https://www.zhaw.ch/iwi/ispt)

Elke Brucker-Kley  
[brck@zhaw.ch](mailto:brck@zhaw.ch)

A vibrant, stylized 3D scene featuring a cute, yellow, butterfly-like creature with large black eyes and orange antennae. The creature is floating in the air, positioned in front of a large, leafy tree. The background consists of rolling green hills, a body of water reflecting the sky, and distant mountains under a clear blue sky. The overall atmosphere is bright and cheerful.

**Vielen Dank!**