

## Erklärungen gemeinsam entwickeln | 02.2026

Im Newsletter des Transregio (TRR) 318 „Constructing Explainability“ präsentieren wir Forschungsprojekte, Workshops zu Künstlicher Intelligenz sowie aktuelle Nachrichten und Vorträge. Sie sind herzlich eingeladen, mit uns auf [LinkedIn](#) und [Instagram](#) zu interagieren und uns eine E-Mail mit Ihrer Frage zu KI zu schreiben. Lassen Sie uns gemeinsam Erklärungen entwickeln!

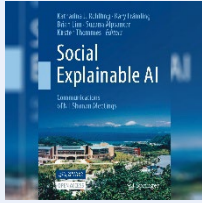
[english version below: click here ▼](#)



## Neue Gesichter, neue Perspektiven

Mit Beginn der zweiten Förderphase hat der TRR 318 nicht nur neue Forschungsprojekte gestartet, sondern auch zahlreiche neue Doktorand\*innen willkommen geheißen. Während viele Promovierende der ersten Förderphase ihre Promotion bereits erfolgreich abgeschlossen haben oder kurz vor der Disputation stehen, sind in den vergangenen Monaten 15 neue Wissenschaftler\*innen zum Forschungsverbund gestoßen.

So kommen in der zweiten Förderphase rund 75 Forschende aus Informatik, Linguistik, Medienwissenschaft, Philosophie, Psychologie, Soziologie und Wirtschaftswissenschaft zusammen – mit unterschiedlichen Erfahrungen und frischen Ideen für die weitere Arbeit rund um #SocialXAI.



## Sammelband zu sozialen Konzepten in der KI-Forschung

Erklärbarkeit ist mehr als eine technische Eigenschaft von KI – sie entsteht im Zusammenspiel zwischen Menschen und Systemen. Das neue Open-Access-Buch „Social Explainable AI“, an dem zahlreiche Forschende des TRR 318 mitgewirkt haben, zeigt, wie Erklärungen als sozialer und interaktiver Prozess gestaltet werden können und stellt Nutzer\*innen, Kontext und Kommunikation in den Mittelpunkt.

[Weiterlesen](#)



## TRR 318 bei einer der weltweit größten Robotik-Konferenzen

Der TRR 318 war auf der diesjährigen IEEE International Conference on Robotics and Automation (ICRA) in Wien vertreten, die mit mehr als 7.600 Teilnehmenden zu den weltweit führenden Konferenzen im Bereich Robotik und Automatisierung zählt. In einem Workshop zur Roboter-Ethik präsentierten Sinem Görmez ([A05](#)) ihre Forschung zu verantwortungsvoller verbaler Unterstützung in der Mensch-Roboter-Interaktion sowie Carolin Klute ([A03](#)) ein ethisches Framework für die Entwicklung emotionsadaptiver KI-Systeme.

[Weiterlesen](#)



## Zwischen Erregung und Gelassenheit

Wann vertrauen Menschen Empfehlungen von KI – und wann hinterfragen sie diese kritisch? Forschende von Projekt [A03](#), Prof. Dr. Kirsten Thommes, Dr. Olesja Lammert und Maurice Pape, haben untersucht, wie der emotionale Zustand den Umgang mit KI-Empfehlungen beeinflusst. Die Ergebnisse zeigen: Besonders Menschen in einem mittleren Erregungszustand prüfen KI-Empfehlungen kritischer als Personen mit hoher oder niedriger emotionaler Erregung. Langfristig könnten die Erkenntnisse dazu beitragen, erklärbare KI-Systeme zu entwickeln, die den emotionalen Zustand ihrer Nutzer\*innen stärker berücksichtigen. Maurice Pape hat die Ergebnisse auf der XAI Konferenz in Brasilien vorgestellt.

[Weiterlesen](#)



## Forschung im Alltag: TRR-Promovierende in der Video-Reihe „Science irl“

In der LinkedIn-Video-Reihe „Science irl“ der Universität Bielefeld geben unsere neuen Doktorandinnen Helena Stahlschmidt ([B07](#)) und Sinem Görmez ([A05](#)) Einblicke in ihre Forschung und ihren Arbeitsalltag. Helena erklärt, wie Missverständnisse zwischen Menschen und KI entstehen – und wie Chatbots damit umgehen können. Sinem zeigt, wie Roboter Anweisungen so formulieren sollten,

dass sie für Menschen klar und gut verständlich sind.

[Helenas Video ansehen](#)

[Sinems Video ansehen](#)



### TRR 318 auf der GENIALE

Der TRR 318 ist am 28. Oktober bei der diesjährigen GENIALE in der Wissenswerkstadt Bielefeld mit dem Format Wissenshappen vertreten. Unter dem Titel „KI als Assistenz-Arzt – Künstliche Intelligenz in der Medizin verstehen und hinterfragen“ geben Prof. Dr. Hanna Drimalla, Prof. Dr. Stefan Kopp und Dr. Helen Beierling Einblicke in ihre aktuelle Forschung. Besucher\*innen erfahren, wie erklärbare KI medizinische Entscheidungen transparenter machen und Ärzt\*innen sowie Patient\*innen im Alltag unterstützen kann.

[Weiterlesen](#)



### Save the Date

Die 4. TRR 318-Konferenz findet am 16. und 17. März 2027 in Paderborn statt. Das Thema der Konferenz lautet „Scaffolding Social XAI“.

[Weiterlesen](#)

Weitere News

## Im Kontext

„Wie schaust du, wenn du verwirrt bist? Oder wenn du etwas verstanden hast?“

Unsere Forschung fragt genau danach. Sie beschäftigt sich damit, wie Menschen ihr Verständnis nonverbal ausdrücken, z.B. indem sie nicken oder die Augenbrauen heben, und wie diese Signale am besten interpretiert werden können. Konkret untersuchen wir, wie Stress und psychische Erkrankungen solche nonverbalen Signale verändern. So möchten wir herausfinden, wie Missverständnisse in der Kommunikation entstehen. Dafür führen wir unsere Studie zurzeit sowohl in der Psychotherapie-Ambulanz der Universität Bielefeld (PAdUB) als auch in der Universitätsklinik für Psychiatrie und Psychotherapie Bethel durch. Die Teilnehmer\*innen bekommen die Regeln von zwei Brettspielen erklärt. Währenddessen zeichnen wir ihre Gesichtsausdrücke per Video auf.

Erste Ergebnisse zeigen, dass Menschen unter Stress tatsächlich ihr Verständnis manchmal anders



ausdrücken. Langfristig möchten wir dieses Wissen nutzen, um Erklärsituationen inklusiver zu gestalten. Außerdem soll es dabei helfen, KI-Systeme zu entwickeln, die nonverbale Rückmeldungen richtig interpretieren und angemessen darauf reagieren können.”

*Vivien Pantel für Projekt [A06](#): Erklärungen des multimodalen Ausdrucks von Stress in klinischen Erklärungen*

## Veröffentlicht

### **Forschungsartikel: How are Gaze Aversions and Mutual Gaze Related to the Topical Development of Dyadic Explanatory Interactions?**

Die Studie erforscht, wie Blickabwendungen und gegenseitiger Blickkontakt mit Themenwechseln und Themenfortführungen in erklärenden Gesprächen zusammenhängen. Die Ergebnisse zeigen, dass Blickabwendungen, insbesondere gemeinsame Blickabwendungen beider Gesprächspartner\*innen, häufig mit dem Abschluss eines Themas und dem Übergang zu einem neuen Thema verbunden sind.

[Weiterlesen](#)

### **Forschungsartikel: Navigating the Dual Nature: Do Explainers Adapt to Explainee Interests When Explaining Technical Artifacts**

Die Studie untersucht, ob Menschen ihre Erklärungen technischer Artefakte an die Interessen ihres Gegenübers anpassen. Im Mittelpunkt steht die sogenannte „duale Natur“ technischer Artefakte und die Frage, wie Erklärende beide Seiten berücksichtigen: zum einen die Funktionsweise eines technischen Geräts, zum anderen seinen Zweck und seine Bedeutung. Die Ergebnisse zeigen, dass Erklärende zwar auf die Interessen ihres Gegenübers eingehen, ihre Erklärungen jedoch weiterhin überwiegend auf die Funktionsweise des Artefakts ausrichten.

[Weiterlesen](#)

## Medientipps

### **Blog Post: The Internet Felt Like This in 1994.**

Douglas Rushkoff berichtet in seinem Blog von seinen ersten praktischen Erfahrungen mit KI und wie sie ihm half, seine Website neu aufzubauen. Dabei diskutiert er, ob KI trotz ihrer Risiken auch für kreative und gesellschaftlich nützliche Zwecke eingesetzt werden kann.

[Weiterlesen](#)

### **Podcast Folge: Wissen2Wirtschaft ... und zurück**

In der fünften Folge von „Wissen2Wirtschaft ... und zurück“ der Universität und Wirtschaftsförderung Paderborn ist TRR-Forscher Prof. Dr. Axel Ngonga zu Gast. Er spricht darüber, warum erklärbares KI immer wichtiger wird, welche Rolle Wissensgraphen dabei

spielen und wie Unternehmen KI sinnvoll einsetzen können.

[Podcast anhören](#)

## Was habe ich gelernt?

„Im Arbeitsalltag bewege ich mich meist im Umfeld meiner Arbeitsgruppe – mit Kolleg\*innen, die an ähnlichen Themen arbeiten und einen ähnlichen fachlichen Hintergrund haben wie ich. Genau deshalb war das Training zum interdisziplinären Arbeiten, das speziell für den TRR 318 angeboten wurde, für mich besonders spannend: Es hat mir vor Augen geführt, wie wichtig es ist, bewusst über den eigenen disziplinären Tellerrand hinauszuschauen.

Was mich dabei am meisten überrascht hat: Obwohl wir im TRR oft an verwandten Fragestellungen arbeiten, unterscheiden sich die Perspektiven, Prioritäten und Herangehensweisen der verschiedenen Disziplinen erheblich. Besonders nachhaltig hat mich die Diskussion darüber beeindruckt, was in den jeweiligen Disziplinen überhaupt als Wahrheit oder wissenschaftliches Ergebnis gilt. Darauf aufbauend haben wir uns gemeinsam überlegt, welche Schwierigkeiten interdisziplinäres Arbeiten mit sich bringen kann, etwa unterschiedliche Fachbegriffe, abweichende Methodenstandards und Erwartungen oder schlicht das Risiko, aneinander vorbeizureden.

Ebenso wichtig war die Frage, wie sich solche Hürden überwinden lassen: durch das frühzeitige Klären gemeinsamer Begriffe, durch gegenseitige Offenheit für andere Herangehensweisen und durch die Bereitschaft, eigene Selbstverständlichkeiten zu hinterfragen.“

*Carolin Klute*

*Doktorandin im Projekt [A03](#)*



---

TRR digital



Oder [direkt per Mail](#) mit Fragen oder Feedback an uns.

Newsletter [abonnieren](#).

## Developing explanations together | 02.2026

In the newsletter of Transregio (TRR) 318 "Constructing Explainability" we present our research projects, workshops on artificial intelligence as well as new publications and upcoming talks. You are invited to interact with us on [LinkedIn](#) and [Instagram](#) and email us with your questions about AI. Let's develop explanations together!

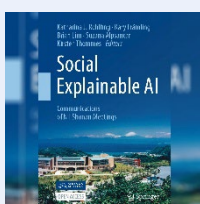


## New Faces, New Perspectives

With the start of the second funding phase, TRR 318 has not only launched new research projects but also welcomed numerous new doctoral candidates. While many doctoral candidates from the first funding phase have already successfully completed their doctorates or are about to defend their dissertations, 15 new researchers have joined the research network in recent months.

In the second funding phase, around 75 researchers from various fields, including computer science, linguistics, media studies, philosophy, psychology, sociology and economics, will come together, each bringing a range of experiences and ideas to the next phase of the project about #SocialXAI.

## News



### Handbook on Social Concepts in AI Research

Explainability is more than a technical feature of AI—it emerges through the interaction between people and intelligent systems. The new open-access book "Social Explainable AI," featuring contributions from numerous TRR 318 researchers, explores how explanations can be designed as a social and interactive process, placing users,

context, and communication at the center.

[Read more](#)



## TRR 318 at One of the World's Largest Robotics Conferences

TRR 318 was represented at this year's IEEE International Conference on Robotics and Automation (ICRA) in Vienna, one of the world's leading conferences in robotics and automation, with more than 7,600 participants. At a workshop on robot ethics, Sinem Görmez ([A05](#)) presented her research on responsible verbal support in human-robot interaction, while Carolin Klute ([A03](#)) introduced an ethical framework for the development of emotion-adaptive AI systems.

[Read more](#)



## Between Excitement and Composure

When do people trust AI recommendations—and when do they question them? Researchers from Project [A03](#), Professor Kirsten Thommes, Dr. Olesja Lammert, and Maurice Pape investigated how users' emotional state influences the way they respond to AI recommendations. Their findings show that people in a moderate state of emotional arousal evaluate AI recommendations more critically than those with either low or high arousal. In the long term, these insights could help develop explainable AI systems that better adapt to users' emotional states. Maurice Pape presented the results at the XAI 2026 Conference in Brazil.

[Read more](#)



## Save the Date

The 4th TRR 318 Conference will take place on 16 and 17 March 2027 in Paderborn. The theme of the upcoming conference is "Scaffolding Social XAI."

[Read more](#)

[More News](#)

## In Context

"What does your face look like when you're confused? Or when you've understood something?"

Our research focuses on exactly that. Specifically, our study examines how people express understanding nonverbally, such as nodding or raising their eyebrows, and how these signals can be best interpreted.



Specifically, we're investigating how stress and mental illness alter these signals. Our goal is to determine how misunderstandings arise in communication. To this end, we are conducting a study at the Bielefeld University's Psychotherapy Clinic (PAdUB) and the Bethel University Clinic for Psychiatry and Psychotherapy. Participants are given instructions on how to play two board games, while we record their facial expressions on video.

Preliminary findings suggest that people under stress sometimes express their understanding differently. In the long term, we hope to use this knowledge to help create explanations that are clearer and more inclusive for a wider audience. This knowledge should also help us develop AI systems that can correctly interpret nonverbal cues and respond appropriately."

*Vivien Pantel for project [A06](#): Explaining the multimodal display of stress in clinical explanations*

## Published

### **Research Paper: How are Gaze Aversions and Mutual Gaze Related to the Topical Development of Dyadic Explanatory Interactions?**

This study investigates how gaze aversions and mutual gaze are related to topic changes and topic continuations in explanatory interactions. The results show that gaze aversions, particularly joint gaze aversions by both interlocutors, are frequently associated with the closure of a topic and the transition to a new one.

[Read more](#)

### **Research Paper: Navigating the Dual Nature: Do Explainers Adapt to Explainee Interests When Explaining Technical Artifacts**

The study examines whether people adapt their explanations of technical artifacts to the interests of their conversation partner. The focus is on the so-called 'dual nature' of technical artefacts, examining how explainers consider both the way a technical device works and its purpose and significance. The findings show that while explainers do respond to their conversation partner's interests, their explanations remain predominantly focused on how the artifact works.

[Read more](#)

## Media Tips

### **Blog Post: The Internet Felt Like This in 1994.**

Douglas Rushkoff reflects in his blog on his first hands-on experience with AI while rebuilding his website. He explores whether AI, despite its risks and drawbacks, can be used for creative projects and positive social change.

[Read more](#)

## What Have I Learned?

“In my day-to-day work, I usually interact within my research group with colleagues who have similar academic backgrounds and work on similar topics. That’s why the interdisciplinary training offered specifically for TRR 318 was particularly exciting for me. It made me realize how important it is to consciously look beyond the boundaries of my own discipline.

What surprised me most was that, although we at the TRR often work on related issues, the various disciplines have different perspectives, priorities, and approaches. I was particularly struck by the discussion about what constitutes truth or a scientific finding in each discipline. Building on that, we discussed the challenges that interdisciplinary work can pose, such as different technical terms, varying methodological standards and expectations, and the risk of misunderstanding one another.

It was equally important to consider how such obstacles could be overcome. This could be by clarifying common terminology early on, being open to other approaches, and being willing to question one’s own assumptions.”



*Carolin Klute*  
PhD student **A03**

---

TRR digital



Or **message us directly** for questions or feedback.

**Subscribe** to the newsletter for free.

**TRR 318 „Constructing Explainability“**

Teilprojekt WIKO „Fragen zu erklärbaren Technologien“  
Universität Bielefeld  
Universitätsstraße 25  
33615 Bielefeld

[communication@trr318.uni-paderborn.de](mailto:communication@trr318.uni-paderborn.de)