

## Erklärungen gemeinsam entwickeln | 01.2022

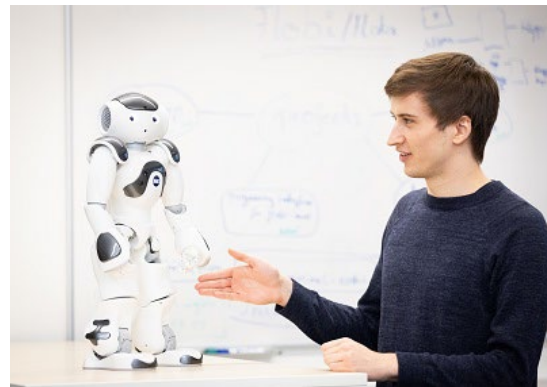
Im Newsletter des Transregios (TRR) 318 „Constructing Explainability“ präsentieren wir Forschungsprojekte und ihre Ergebnisse, Workshops zu künstlicher Intelligenz (KI) sowie Veröffentlichungen und Vorträge. Sie sind herzlich eingeladen, mit uns auf Twitter zu interagieren und uns eine E-Mail mit Ihrer Frage zu KI zu schreiben. Lassen Sie uns gemeinsam Erklärungen entwickeln!

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



### Studien zur Erklärbarkeit von künstlicher Intelligenz laufen an

Im TRR 318 arbeitet ein interdisziplinäres Team an nachvollziehbaren Erklärungen. Die Erkenntnisse sollen unter anderem dazu beitragen, KI begreifbar zu machen. In ihre Studien bezieht die Forschungsgruppe die Nutzer\*innen von KI ein. Interessierte können Workshops besuchen, in denen sie KI erkunden oder an Studien als Versuchspersonen teilnehmen und so einen Einblick in die Forschung des Transregios 318 erhalten. [Weiterlesen](#)



### Erster Workshop zu künstlicher Intelligenz

Anfassen, erforschen und manipulieren: Rund 15 Schüler\*innen der 11. Jahrgangsstufe des ESG Gütersloh haben den ersten Workshop zu künstlicher Intelligenz (KI) vom TRR 318 an der Universität Paderborn besucht. Zuerst entschlüsselten die Schüler\*innen die Funktionen eines mit KI ausgestatteten Systems. Anschließend folgte ein dreistündiges praktisch angelegtes Programm rund um KI-Themen.



## Sonderausgabe von KI-Magazin

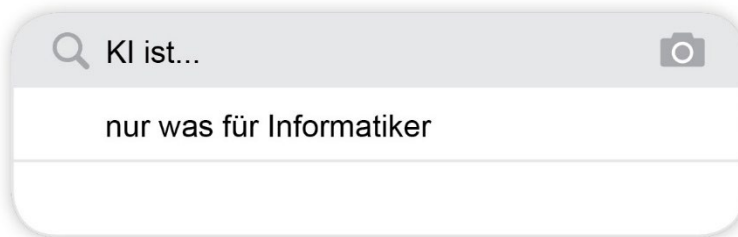
Der Springer-Verlag veröffentlicht im Sommer eine Sonderausgabe der Zeitschrift für KI. Sie beschäftigt sich mit dem Forschungsfeld „Erklärbare künstliche Intelligenz“ (XAI). Forschende aus dem TRR 318 sind mit Artikeln vertreten. [Weiterlesen](#)



## Veranstaltung zu lernenden Maschinen

Bei der Eventreihe „50 Mosaik“ zum 50-jährigen Jubiläum der Universität Paderborn bietet ein TRR-Team einen offenen Workshop für Familien an. In dem Mitmach-Angebot arbeiten die Teilnehmenden mit lernenden Maschinen (Teachable Machines). [Weiterlesen](#)

### Weitere News



„KI-Systeme werden von Informatiker\*innen programmiert, das ist richtig. Genutzt wird KI aber in vielen Bereichen, wie zum Beispiel der Medizin. Für viele Anwender\*innen sind KI-Systeme oft komplex und undurchsichtig. Die Erkenntnisse aus unserer Forschung tragen dazu bei, Erklärprozesse zu durchschauen und damit KI interaktiver und begreifbarer zu machen. Dafür untersuchen wir zunächst, was nachvollziehbare Erklärungen zwischen Menschen, aber auch zwischen Mensch und Maschine ausmacht. Unser Ziel ist, dass KI in ihren Erklärungen Reaktionen und Nachfragen von Nutzer\*innen berücksichtigt. Für diese Aufgabe haben wir im TRR 318 ein breites interdisziplinäres Team mit Forschenden aus Linguistik, Psychologie, Medienwissenschaft, Wirtschaftswissenschaft und Informatik aufgestellt.“



*Prof. Dr. Katharina Rohlfing,  
Sprecherin des TRR 318 und Linguistin  
Teilprojekte [A01](#), [A05](#) und [Z](#)*

### **Video: Workshop zu KI**

Der erste Co-Construction Workshop (kurz: CCWS) blieb nicht undokumentiert. Der Film gibt einen Einblick in den CCWS zu künstlicher Intelligenz, der sich an Schulklassen richtet.

[Anschauen](#)

### **Paper: "Explanation as a Social Practice: Toward a Conceptual Framework for the Social Design of AI Systems"**

Der derzeitige Anstieg des Interesses an der Erklärbarkeit von künstlicher Intelligenz (XAI) wird nicht nur durch technologische Fortschritte im Bereich des maschinellen Lernens vorangetrieben, sondern auch durch regulative Initiativen zur Förderung von Transparenz bei der algorithmischen Entscheidungsfindung. [Weiterlesen](#)

### **Paper: "Modeling Feedback in Interaction With Conversational Agents - A Review"**

Intelligente Systeme, die mit Menschen durch Konversation interagieren (z. B. Roboter, verkörperte Gesprächssysteme oder Chatbots), müssen vom Menschen Feedback erhalten, um sicherzustellen, dass die von ihnen kommunizierten Handlungen die beabsichtigten Folgen haben. Gleichzeitig wird der Mensch, der mit dem System interagiert, ebenfalls um Feedback bitten, um sicherzustellen, dass die von ihm kommunizierten Handlungen ebenso die beabsichtigten Folgen haben. [Weiterlesen](#)

### **Paper: "(De)Coding Social Practice in the Field of XAI: Towards a Co-constructive Framework of Explanations and Understanding Between Lay Users and Algorithmic Systems"**

Die Fortschritte in der Entwicklung Künstlicher Intelligenz und dessen Anwendung in immer mehr gesellschaftlichen Bereichen sorgen dafür, dass sich die Leute immer mehr mit den Technologien auseinandersetzen und sie zumindest zu einem gewissen Grade verstehen müssen. [Weiterlesen](#)

### **Teilnehmende für mehrere Studien**

Eine Auswahl der Studien:

- Gesellschaftsspiele erklären und erklärt bekommen,
- sich von virtuellen Agenten anleiten lassen,
- Alltagsbegegnungen mit künstlicher Intelligenz festhalten,
- Robotern Bewegungen beibringen.

[Zur Webseite](#) für mehr Informationen zu den aktuellen Studien.

Oder: Mailingliste mit Einladungen zu neuen Studien [abonnieren](#).

## Schulklassen für KI-Workshops an der Universität Paderborn

 ab Juni 2022  Schulklassen

Ein TRR-Forschungsteam aus dem Fachbereich Didaktik der Informatik bietet weitere Workshops zu KI an. Vorwissen ist nicht erforderlich, die Termine werden individuell vergeben. Interessierte Lehrkräfte und Schulklassen sind eingeladen, sich per E-Mail anzumelden: [communication@trr318.uni-paderborn.de](mailto:communication@trr318.uni-paderborn.de)

## Veranstaltungen

### 23.-24. Juni 2022: 1. TRR 318 Konferenz „Explaining Machines“

Universität Bielefeld: [zur Veranstaltung](#)

### 06. Juli 2022: Workshop zu lernenden Maschinen bei „50 Mosaik“

Universität Paderborn: [zur Veranstaltung](#)

## Was habe ich gelernt?

„In einem Gespräch mit Kolleg\*innen ist mir bewusst geworden, wie unterschiedlich wir akademisch sozialisiert und trainiert sind durch unsere Ausbildungen in unseren jeweiligen Disziplinen. Das birgt große Chancen, aber auch Herausforderungen; nicht nur für die inhaltliche Arbeit über Disziplinen hinweg, sondern auch im gemeinsamen Arbeiten und Lernen. Für einige von uns waren das Lesen und Diskutieren von Aufsätzen in unserem bisherigen Studium täglich Brot, bei anderen kam das fast gar nicht vor. Einige haben gelernt, dass mathematische Logik der Inbegriff von Wissenschaftlichkeit ist, andere fassen die Philosophie als Grund- oder Leitdisziplin auf. Diese Unterschiede führen zu unterschiedlichen Wahrnehmungen und Bewertungen, was manchmal frustrierend sein kann. Ich denke, dass wir als junge und angehende Wissenschaftler\*innen wachsen können, jetzt und erst recht zukünftig reflektierter interdisziplinär zu arbeiten.“

*Annedore Wilmes, Doktorandin im [Teilprojekt C04](#)*

---

TRR digital



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

Newsletter [abonnieren](#).

## Developing explanations together | 01.2022

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



### Kickoff for studies on the explainability of artificial intelligence

At TRR 318, an interdisciplinary team is working on understandable explanations. Their findings are intended to contribute to making AI more comprehensible. The research group includes actual AI users in its studies. Interested parties can attend workshops in which they explore AI or participate in studies as test subjects, giving them a glimpse into the research taking place at Transregio 318. [Continue Reading \(in German\)](#)



### First workshop on Artificial Intelligence

Exploring AI hands-on: Around 15 11th grade students from the ESG Gütersloh attended the first workshop on artificial intelligence (AI) held by TRR 318 at the University of Paderborn. First, students deciphered the functions of an AI-equipped system. This was followed by a three-hour hands-on program on AI topics.



## Special edition of AI Journal

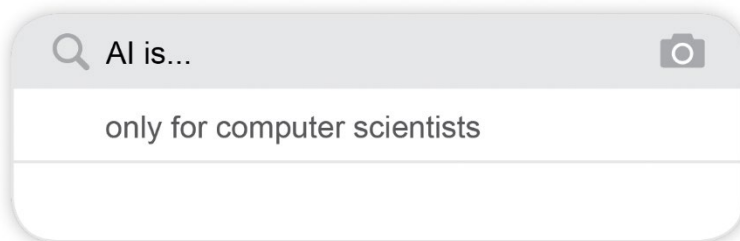
Springer Verlag will publish a special edition of its AI Journal dedicated to current research developments in explainable artificial intelligence (XAI). This special edition features some articles TRR 318 researchers. [Continue reading](#)



## Event on teachable machines

As part of the "50 Mosaics" event series celebrating the 50th anniversary of the University of Paderborn, a TRR team is offering an open workshop for families. In the hands-on event, participants will work with teachable machines. [Continue reading](#)

More News



"AI systems are programmed by computer scientists, that is true. But AI is used in many areas, for example in medicine. For many users, AI systems are often complex and opaque. The findings from our research contribute to better understanding explanation processes and thus to making AI more interactive and comprehensible. The first step here is to investigate what constitutes comprehensible explanations between humans, but also between humans and machines. Our goal is to ensure that AI takes into account user reactions and questions in its explanations. For this task, we have established a broad interdisciplinary team at TRR 318, including researchers from linguistics, psychology, media studies, economics, and computer science."



*Prof. Dr. Katharina Rohlfing,  
Spokesperson of TRR 318 and linguist  
Subprojects [A01](#), [A05](#) and [Z](#)*



### **Video: First Co-Construction Workshop**

The first Co-Construction Workshop (CCWS) was filmed to provide a glimpse into this participatory and exploratory workshop on artificial intelligence, which is aimed at school classes. [Watch here \(in German\)](#)

### **Paper: "Explanation as a Social Practice: Toward a Conceptual Framework for the Social Design of AI Systems"**

The recent surge of interest in explainability in artificial intelligence (XAI) is propelled by not only technological advancements in machine learning but also by regulatory initiatives to foster transparency in algorithmic decision making. [Continue reading](#)

### **Paper: "Modeling Feedback in Interaction With Conversational Agents - A Review"**

Intelligent agents interacting with humans through conversation (such as a robot, embodied conversational agent, or chatbot) need to receive feedback from the human to make sure that its communicative acts have the intended consequences. At the same time, the human interacting with the agent will also seek feedback, in order to ensure that her communicative acts have the intended consequences. [Continue reading](#)

### **Paper: "(De)Coding Social Practice in the Field of XAI: Towards a Co-constructive Framework of Explanations and Understanding Between Lay Users and Algorithmic Systems"**

Advances in the development of AI and its application in many areas of society have given rise to an ever-increasing need for society's members to understand at least to a certain degree how these technologies work. [Continue reading](#)

## Events & Activities

### **23.-24. Juni 2022: 1st TRR 318 Conference on "Explaining Machines"**

Bielefeld University: [Further information](#)

### **06. Juli 2022: "Teach me if you can" - a Workshop on Designing and Exploring AI**

Paderborn University: [Further information](#)

## What have I learned?

"In a conversation with colleagues, I realized how differently we are socialized and academically trained through our education in our respective disciplines. This presents great opportunities, but also challenges – not only in terms of the content when collaborating across disciplines, but also in terms of working and learning together. For some of us, reading and discussing essays was commonplace in our previous studies; for others, it almost never happened. Some have learned that mathematical logic is the epitome of scholarship, while others see philosophy as a foundational or guiding discipline. These differences lead to different perceptions and evaluations, which can be frustrating at times. I think we can grow as young and aspiring scholars to be more reflective in our interdisciplinary work now and even more so in the future."

*Annedore Wilmes, PhD student in [subproject C04](#)*

### TRR digital



Or [message us directly](#) for questions or feedback.

[Subscribe](#) to the newsletter for free.



### TRR 318 „Constructing Explainability“

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

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