

20
22

TUM
IEAI

TUM
Institute for Ethics
in Artificial Intelligence

ANNUAL REPORT

20
22

TUM
IEAI



Our mission:
Integrating ethical and societal priorities into the development
of AI technologies and approaches.

TUM
Institute for Ethics in Artificial Intelligence

ANNUAL REPORT

2022



ieai.sot.tum.de

Director's Message	04
Executive Director's Message	06
2022 at the IEAI in Numbers	07
Working at the IEAI	08

external links // only in digital version



Index

Research	10
A	12
13	
14	
15	
16	
17	
18	
19	
B	20
21	
22	
23	
C	24
25	
26	
27	
28	
29	
30	
31	
▶ 34	
▶ 36	
▶ 38	

New Projects

- ▶ **A1** Co-designing a Risk-Assessment Dashboard for AI Ethics Literacy in EdTech
- ▶ **A2** Idiomorphic AI: Emergent, Tailored, and Normative Behavior in Large Language Models
- ▶ **A3** NLawP: Natural Language Processing and Legal Tech
- ▶ **A4** The Potential for AI in the Extractive Industries to Promote Multi-objective Optimization
- ▶ **A5** Tradeable Mobility Credits: Addressing Ethical Concerns with Algorithm Transparency
- ▶ **A6** UTEM – Undistorted Technological Mediation
- ▶ **A7** Ethics for the Smart City: Applied Socio-technical Frameworks to Assess the Implementation of AI-related Solutions

Joint Projects and Partnerships

- ▶ **B1** MELISSA – Mobile Artificial Intelligence Solution for Diabetes Adaptive Care
- ▶ **B2** Towards an Accountability Framework for AI Systems
- ▶ **B3** Artificial Intelligence for Earth Observation: Reasoning, Uncertainties, Ethics and Beyond (AI4EO)

Continuing Research

- ▶ **C1** A Human Preference-Aware Optimization System
- ▶ **C2** ANDRE – AutoNomous DRiving Ethics
- ▶ **C3** Online Firestorms and Resentment Propagation on Social Media: Dynamics, Predictability and Mitigation
- ▶ **C4** Online-Offline Spillovers – Potential Real-World Implications of Online Manipulation
- ▶ **C5** Personalized AI-Based Interventions Against Online Norm Violations
- ▶ **C6** Rule of Law, Legitimacy and Effective COVID-19 Control Technologies
- ▶ **C7** The Ethics and Practice of AI Localism at a Time of COVID-19 and Beyond

Applying Ethics

Assessing Impact

Promoting Engagement

Knowledge Dissemination

- 42 Publication Highlights
- 44 Conference Highlights
- 46 Research Briefs
- 47 New White Paper Series
- 48 Thesis Projects
- 50 Q&A Series: Reflections on AI
- 51 In the Media

Reaching Out

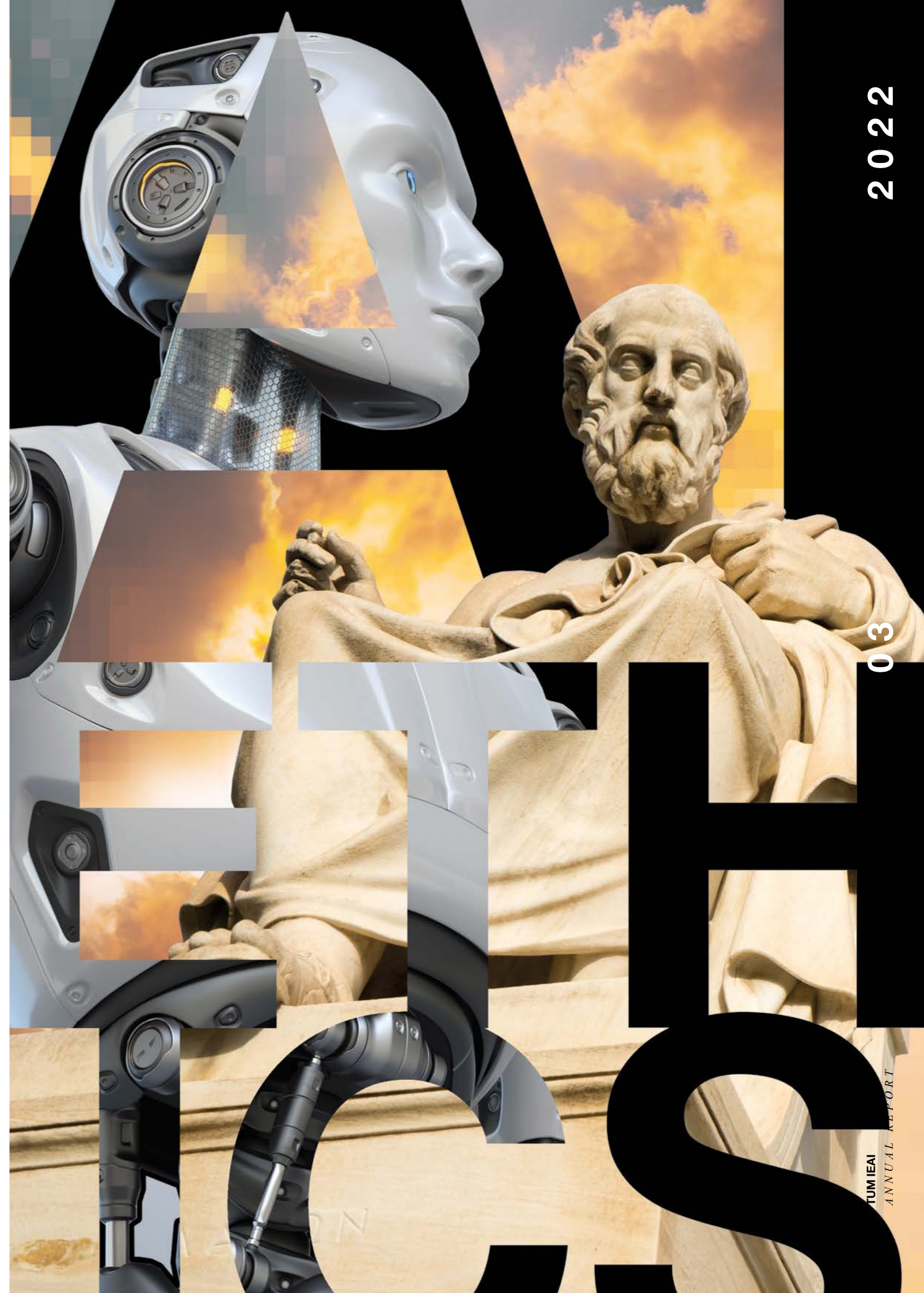
- 54 Global AI Ethics Consortium (GAIEC)
- 55 AI Ethics Course
- 56 Responsible AI Network (RAIN-Africa)
- 58 IEAI Speaker Series
- 60 Workshops and other Events

Social Media

Annex

- 66 IEAI Staff / Advisory Board
- 67 Visiting Professors / External Doctoral Researchers
- 68 Project Members

Connect with Us! / Imprint 72





Dear Readers,

As the year comes to a close, we seize the opportunity to step back, reflect and review what happened during the past twelve months. Creating our Annual Reports provides a unique way for us to better understand what we have achieved, evaluate our progress and gain important insights that will help us expand our impact, increase our reach and accelerate our pace in the coming years.

As surveys clearly illustrate, Artificial Intelligence has seen huge growth, which will only increase in the near future. During 2022, AI-related technologies and applications enabled phenomenal advances, and their adoption across sectors is simply inevitable.

2022 was a big year, a year in transition. We are grateful to have been able to return to hosting in-person events and to collaborating with partners through travel, meetings, and workshops. We have also learned a lot in this process and feel confident about our ability to reach out to our networks in a variety of ways. We look forward to continuing this trend in 2023.

Our work in 2022 has been characterized by **a major leap forward in our reach**, both in terms of our network and our topical scope. Our Institute has, from the beginning, aimed to bring AI Ethics to the center of contemporary AI-related discussions, recognizing both the risks and opportunities of this remarkable technology, and pushing for the adoption of responsible use of AI in practice. This year, we capitalized upon that goal in many ways, pushing our own boundaries in terms of where and how we conduct and disseminate our research, and who we work with.

In 2022, our Institute **awarded over €1.3 Million** to fund seven multi- and interdisciplinary projects. Focusing on the use of AI in education, the legal and mineral mining sectors, the effect of AI on behavior, and AI as it relates to mobility and urban infrastructure. Our new projects will bring concrete solutions and best practices to stakeholders in many sectors.

With these new projects, our Institute welcomed 15 new principal investigators from TUM and external partners, such as the University of Ulm, Prairie – PaRis Artificial Intelligence Research Institute –, and the Berkman Klein Center for Internet and Society at Harvard University. **With these new additions we are lifting the barriers to cross-cultural cooperation in the AI ethics field and expanding future of our network and outreach.**

Spanning sectors, disciplines and borders, the IEAI also **brought diverse voices across the AI community together this year.** Through our Speaker Series and other events co-organized with partners from the industry (IBM) and other research centers (such as the new TUM Think Tank), we reflected on, explored, and identified the ethical issues related to the development, use and impact of AI-enabled technologies. Distinguished guest speakers from Africa, the Americas, Asia, Europe, and Oceania brought a wider range of perspectives and expertise from around the world, raising, at the same time, public awareness of the benefits and risks of AI technologies.

Visiting Professors from Africa (Ghana), Asia (Japan), and the United States (Virginia) joined us in Munich to discuss AI issues in several fields of application, boosting conversations within the research community and highlighting the ecosystem's diversity.

AI is a rapidly changing technology that embeds itself at all levels of society, raising new challenges, risks, dilemmas, and ethical concerns while repositioning and reprioritizing its responsible use. That makes AI ethics an enormously important field.

The work of the IEAI, among other actors in this space, has played an important role in **defining a new and distinct discipline** that until recently was only considered part of other fields, such as computer science. Indeed, AI ethics and AI governance are being taken much more seriously at local, regional, European, and international levels. Countries, companies, the public and private sectors, civil society, and, of course academia, are seeking to be involved in AI discussions like never before.

That brings enormous opportunities and added value in the field of AI ethics, but also brings responsibility to all of us that are involved to think about how we influence the development of concrete and case-specific recommendations and regulations.

Ethical guidelines are crucial for establishing public trust and elemental to the success of AI overall. Our commitment for 2023 is to work more in this direction.

Stay tuned for an exciting year to come where we will continue to present more and more outcomes of our projects and get the chance to meet members of our network from all over the world in Singapore for the second installment of **The Responsible AI Forum (#TRAIF2023)** in September.

Looking forward to another productive year! ●

Yours,

Christoph Lütge
Director

Executive Director's Message



90



Welcome everyone to our 2022 Annual Report!

Choosing the words of our Director, I want to take the opportunity to underscore the impressive work of our researchers and staff over the past year and thank them for their commitment to the mission of the IEAI. In addition to expanding our reach topically and in terms of our network, **2022 was a year of intense growth for the IEAI.** We more than doubled the amount of in-house researchers this year, as well as the number of projects we are administrating directly at the institute. Our dedicated staff of program and communication managers and student assistants have been invaluable in this process and have helped to make the IEAI a shining example of how to build an institute to be one of the foremost research centers working on AI ethics in the world in a very short timeframe.

2022 also was a turning point in the productivity of the IEAI. With many of our first round of research projects wrapping up or finalizing their data analysis, **we have seen an explosion of high quality publications and outreach communicating the results of years of work on novel research subjects.** We more than doubled the amount of peer-reviewed publications this year. We have continued with our Research Brief series, creating a space to communicate our work to a wider public. And we started a White Paper Series, creating an outlet for our work aimed at decision makers across sectors.

As 2022 ends, I am also terribly excited about the future at the IEAI. Our **seven new research projects cover such a wide range of topics, all of incredible valuable to our mission** of integrating ethical and societal priorities into the development of AI technologies and approaches. As these projects ramp up in 2023, so will our ability to provide meaningful content and consideration across the AI ethics space.

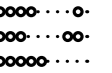
Moreover, with the launch of our new research projects, we have increased the number of IEAI affiliated professors and researchers like never before. Not only do our new projects cover a very different set of issues than in our first and second rounds of funding, but they are **bringing new voices to the table of AI ethics discussions.** This demonstrates the far-reaching and interdisciplinary nature of our work, and also serves as a good indicator of the IEAI success in creating a community across the Technical University of Munich and beyond that I hope we can continue to grow in 2023.

I urge you to take a closer look at each of our projects on the following pages, as well as the other outstanding activities we carried out in 2022, and see the amazing work this group of individuals has put together in just the last year.

Enjoy the read! ●

Caitlin Corrigan,
Executive Director

2022 at the IEAI in Numbers



2022

2022 was an incredibly productive year for our research and our outreach. Alone in 2022, the IEAI ... ●

Hosted
17
events

grew to an
in-house staff of
15

with
more than
700
participants

conducted
and published
8
exclusive
expert interviews

ran
19
projects in
7
clusters

produced
30+
academic publications ...

with
24
affiliated PIs

and
10
IEAI Research Briefs
and White Papers

and
22
affiliated
research associates

reached
15,000
social media followers

07

Working at the IEAI



From the very beginning, our goal was to create a working environment that brings together people who support each other in addressing real-world challenges and far-reaching future questions through interdisciplinary research.

The IEAI's amazing international team consists of 15 committed members from Africa, Europe, and North America. This diversity ensures the development of new ideas, brings different perspectives, and constitutes a key to the IEAI success.

We asked our staff members, research associates, and student assistants why working on AI Ethics matters, why an interdisciplinary approach is so important when it comes to AI Ethics, and what motivated them to work at the IEAI.

Here are their thoughts: ...

Staff:
Why working on AI Ethics matters?

▶ Working on AI ethics matters as it helps to ensure that AI is developed and used in ways that are safe, fair, and aligned with human values and interests. As AI becomes more advanced and integrated into society, it is crucial to address potential ethical concerns and promote responsible AI development and deployment.

Manuela Fuchs / Christina Daman
 Program- & Event Managers

▶ AI Ethics is the only way to ensure society benefits from AI's full potential. Thus, stressing its importance and communicating the outcome of the incredible work that is being done by the IEAI research team is much more than work; it is a mission.

Anastasia Aritzi
 Communications Consultant



Research Associates:
Why is interdisciplinary approach so important when it comes to AI Ethics?

▶ AI proves how difficult it can be to enclose reality within disciplinary borders. The range of applications is so vast that to fully explore the potentials and risks we need to open a dialogue not only within academia but also with other stakeholders.

Ana Catarina Fontes

▶ The development of technologies needs to go along with "bigger picture" ethical reflection to prevent implicit embedding and diffusion of undesirable values. Similarly, ethical reflection needs to be complemented by technical implementation to turn ideas into practice. To ensure both, we have to work in interdisciplinary teams.

Franziska Poszler

▶ An interdisciplinary approach is important in AI ethics because AI ethics covers a wide range of disciplines, including computer science, philosophy, law, psychology, and sociology. Each of these disciplines provides a unique perspective and insights to address a growing group of ethical challenges arising at the interface of technology and human values.

Lameck Amugongo

▶ When discussing AI Ethics, we are talking about all the aspects of societies, organisations, and environments the technology will touch through its lifecycle. Addressing all of those elements requires people with varying degrees of expertise to be in the conversation if we want to build the best and most inclusive recommendations and analysis.

Auxane Boch

▶ Some of the questions in our modern, digitalized world can only be answered by bringing together a wide range of knowledge and experience. This creates new opportunities for thinking out of the box.

Ellen Hohma



▶ People continuously make ethical decisions about their own behavior and judgments about the actions of others. This applies to all areas of our lives – including the digital world. For me as a descriptive ethicist, this is exactly what is exciting to quantify and measure in interdisciplinary projects at the IEAI.

Raphael Max

▶ An interdisciplinary approach is helping to question how knowledge is being created in AI. It synthesizes and contrasts different concepts and methods. This inclusive approach is essential in AI ethics, as it enriches technical knowledge with humanistic values.

Maria Pokholkova

▶ The development and implementation of artificial intelligence poses significant risks for the realization of legal and ethical principles such as accountability, autonomy, dignity, equity, privacy, and responsibility. Since AI is deployed in different use cases, understanding the exact meaning of such ethical concepts requires an interdisciplinary understanding, particularly when it comes to translating general principles to specific use cases of AI in health or in the context of human resources.

Alexander Kriebitz



Student Assistants:
What motivated you to work at the IEAI?

▶ I was always quite interested in ethics and during my studies I also discovered an interest for AI, which made me look into AI ethics and business ethics and eventually brought me to the IEAI.

Immanuel Klein

▶ I was excited for the opportunity to hear from experts working on the development of AI technologies and the ethical challenges they may pose. This experience has shown me new ways to analyze both society's values and my own.

Paloma Laye ●





RESEARCH

AGENDA

Research.

The IEAI research agenda was taken to the next level in 2022. There has been an outpouring of quality publications and other outputs from our longer-running projects and exciting new thematic focuses from our newly established projects. Our overarching questions remain the same:

- ▶ *What should be possible in AI?*
- ▶ *How can we ensure that as many people as possible benefit from AI's rewards?*

We continued running eight projects and added seven newly funded projects to our work. Additionally, we have three externally funded projects with partners in industry and academia across the world.

New Projects



In June 2022, ...

... the IEAI awarded over €1.3 Million over the next three years to fund seven new projects on AI Ethics. All of these projects combined academic expertise from multiple fields to approach challenges related to implementing AI in a responsible and effective way. The topics include examining the use of AI in education, law, and the extractive industries, the effect of AI on behavior and AI as it relates to mobility and urban infrastructure.

With these new projects, we welcome 15 new principal investigators from TUM to the IEAI, as well as external partners from University of Ulm, Prairie – PaRis Artificial Intelligence Research InstitutE, and Berkman Klein Center at Harvard University, among others.

Project A1

Co-designing a Risk-Assessment Dashboard for **AI Ethics Literacy in EdTech**



This project has three key research objectives related to ethics in AI EdTech: (1) **Developing a risk-assessment framework to inform stakeholders in the education sector**, (2) **co-designing and developing a dashboard to support educational decision makers in AI ethics literacy** and (3) **documenting the design process for the development of AI ethics literacy tools for educational decision making.**

with an educational audience in mind but still carrying a wide range of affordances valuable for educational use.

- ▶ There is a need to consider context-specific implementations of AI tools within educational settings to distill intersectional ethical implications and opportunities that arise from the pedagogical use for our understanding of how teaching and learning should happen.
- ▶ Educational decision-making processes that use AI-based tools implemented for pedagogical purposes should build on teachers' pedagogical expertise. This includes using educators' capacities to imagine new uses of AI tools in order to innovate and advance how teaching and learning happens in the classroom.

2022 Findings

- ▶ The current state of AI EdTech and law is not considered through a child-rights-based perspective. Thus, we have identified a research area.
- ▶ Child rights are central to the debate of educational technologies and could help to identify the key AI ethics framework, which will be the basis of our dashboard later on. Therefore, discussing AI ethical principles through a child-rights perspective instead of geographical legislation (and connected principles) could allow for a nuanced and interdisciplinary understanding of ethical risks for AI in education.
- ▶ Analysis of over 50 AI tools used in the education setting found that some are deliberately developed for educational purposes, with particular curricular and learning aims in mind, while others were not developed

Plans for 2023

In 2023, we have planned a hackathon around the project with students at TUM to propel our research on AI EdTech tools, their use in the classroom and policies and laws related to AI Ethics in education into an initial iteration of the AI Ethics Literacy Dashboard. We also will be co-designing workshops with educational stakeholders to iteratively refine the Dashboard. ●

➔ ieai.sot.tum.de/research/co-designing-a-risk-assessment-dashboard-for-ai-ethics-literacy-in-edtech

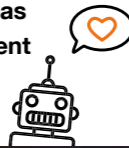
EdTech Project Team



Project **A2**

Idiomorphic AI: Emergent, Tailored, and Normative Behavior in Large Language Models

This project aims to investigate generative AI models, such as ChatGPT for, (1) evidence of emergent human-like psycho-social qualities ('what is'), (2) dynamics of tailoring them to individual users ('what it could be for you') and (3) how these emergent and tailored aspects inform a range of ideas in AI ethics ('what should be'), such as value alignment and trustworthiness.



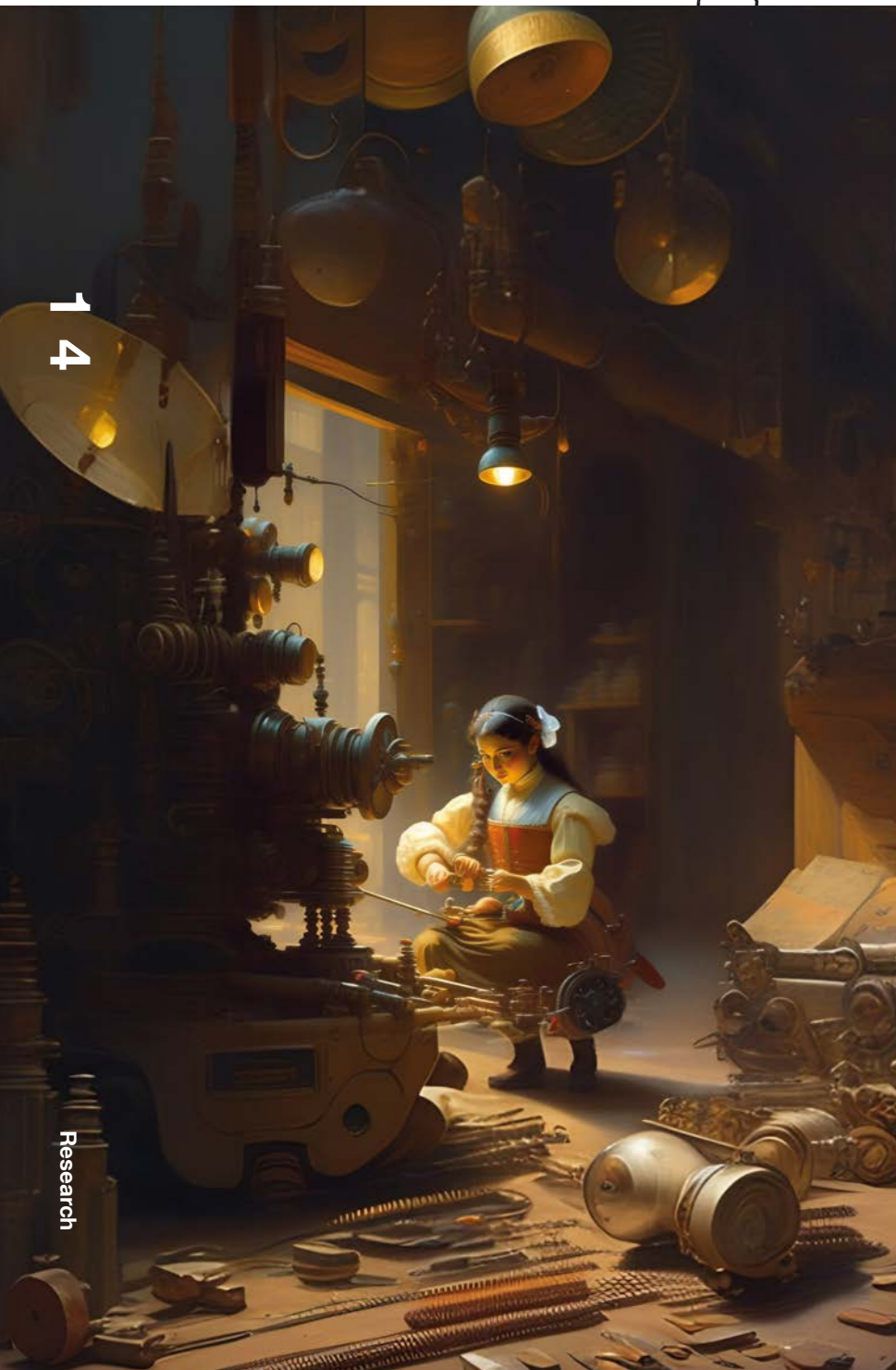
Plans for 2023

In 2023, we will focus on studying emergent aspects of large language models, especially whether there is evidence for reliable and internally consistent biases for specific psycho-social qualities in their output. In the psychological realm, we will examine personality traits and in the political realm, we will look for policy positioning and ideological alignment in major commercial large language models.

ieai.sot.tum.de/research/idiomorphic-ai-emergent-tailored-and-normative-behavior-in-large-language-models

Generated by: Stable Diffusion 2.1

Prompt: A Carl Spitzweg painting of a girl building a robot in a workshop with machine parts strewn around



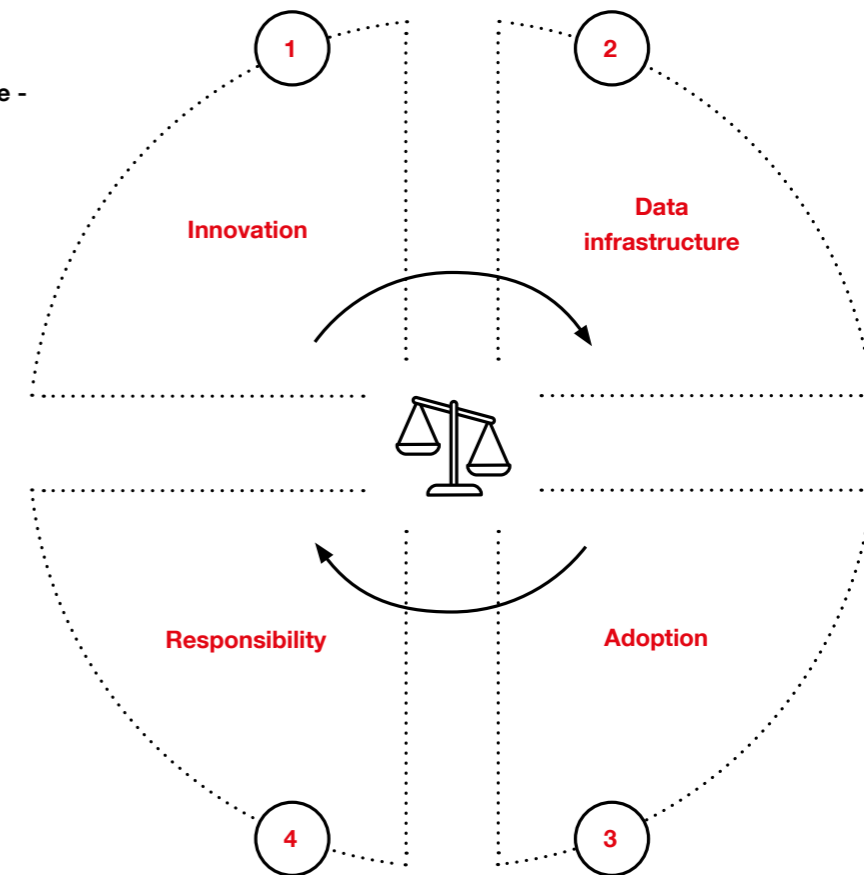
Project **A3**

NLawP: Natural Language Processing and Legal Tech

This project evaluates how AI technologies can impact the legal sector in disruptive ways. It will map current developments in research and in the legal tech market, inquire into how data governance can enable a responsible and efficient adoption, and bring stakeholders together in order to produce new ideas and innovations.

ieai.sot.tum.de/research/nlawp-natural-language-processing-and-legal-tech

The NLawP Cycle - Areas of inquiry



Plans for 2023

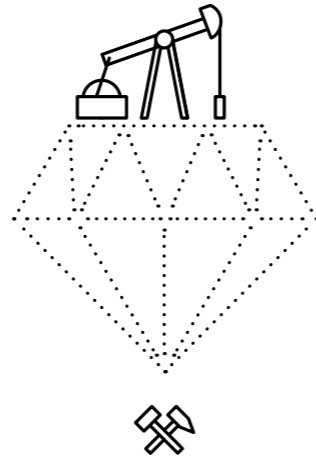
In 2023, we aim to create a "NLawP Tech Map", which will give a dynamic overview of applications of NLP and similar technologies in the field of legal tech. With this, we will derive a picture of the systemic effects of NLP technologies to the legal system in order to understand risks and opportunities with respect to the common good.

From there, we will analyze technical infrastructure and evaluate data governance and data management of such systems, informed by the needs and views from practitioners in three roundtable discussions. A report on Legal Data-Sharing will then outline possible solutions to enhance data sharing in the legal sector across organizations.

Project **A4**

The Potential for AI in the Extractive Industries to Promote Multi-objective Optimization

This project aims to design and test a multi-objective optimization model for extractive industries that considers both economic and non-economic costs and benefits. The approach will be underpinned by novel AI and machine learning methods that assess and project environmental and societal implications. The project will also compile a database to raise the understanding of a diverse array of factors that affect or are affected by natural resource production.



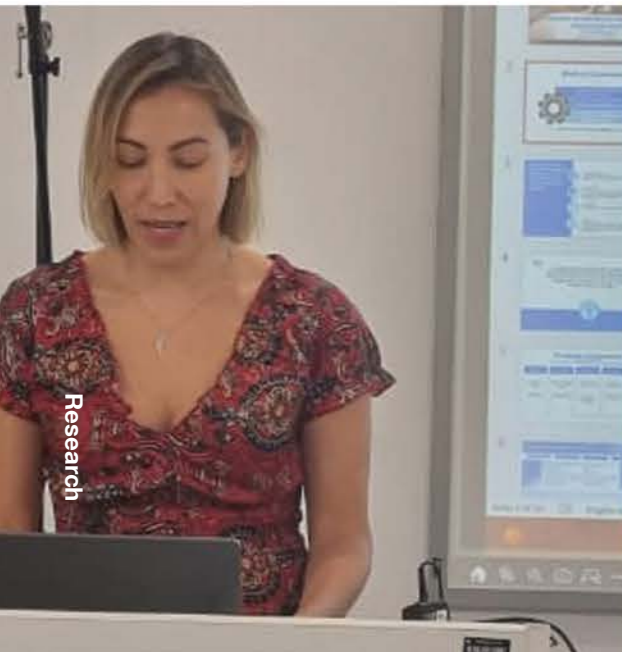
2022 Findings

▶ AI is already being used in the extractive industries, mostly in terms of increasing productivity, efficiency and workplace safety. However, social and environmental factors have up to this point been overlooked in this process.

▶ Based on an overview of current AI uses in the mining industries, we identified four main categories of ethical considerations:

- [1] **autonomy** and **observation**,
- [2] **balance of rewards**,
- [3] **bias** and **prioritization** and
- [4] **explainability** and **acceptance**.

Pamela Duran Diaz speaking about frameworks and indicators for land governance Africa



Plans for 2023

In 2023, the project will focus on stakeholder and geoportal assessments and collecting information on land policies for extractive industries in Sub-Saharan Africa to enable a machine learning produced map of land regeneration after extractive industry concessions. Building on this, we will take the first steps in the model development. In order to workshop these concepts with stakeholders in Africa, the IEAI will organize and attend a Responsible AI Workshop at KNUST in Ghana in March 2023.

▶ ieai.sot.tum.de/research/the-potential-for-ai-in-the-extractive-industries-to-promote-multi-objective-optimization

Participants at 2nd Land Governance and Societal Development Conference 2022 in Windhoek



Project **A5**

Tradeable Mobility Credits: Addressing Ethical Concerns with Algorithm Transparency

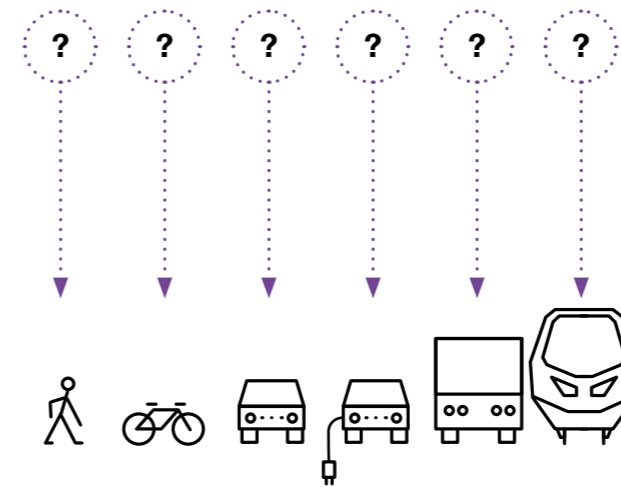
The goal of this project is to better understand the ethical aspects involved in the implementation of Tradeable Mobility Credits (TMC), in particular the mediating role of trust on the utilization of TMC as a result of transparency. This will inform the successful and ethical implementation of TMC systems.

▶ ieai.sot.tum.de/research/tradeable-mobility-credits-addressing-ethical-concerns-with-algorithm-transparency

Through a series of incentivized choice experiments (lab and field-type), the project looks at the impacts of disclosure of tracking requirements and algorithmic transparency on (1) user reaction, acceptance, and usage, (2) mode choice and reduction of carbon emissions.

Plans for 2023

In 2023, we will finalize the design and conduct the first pilots related to our research questions. In addition, we plan to run the first experiments on the effects of disclosure on acceptance and usage of Tradeable Mobility Credits.



TUM Sustainability Day during ... (a) the lecture "Nudging Sustainable Behavior" and (b) the event "Artificial Intelligence for Sustainability - Opportunities and Challenges".



UTEM – Undistorted Technological Mediation

The goal of this research project is to identify reciprocal influences (especially epistemic and normative biases) in human-machine interaction and to identify proposed solutions and designs for manufacturers of corresponding technologies.

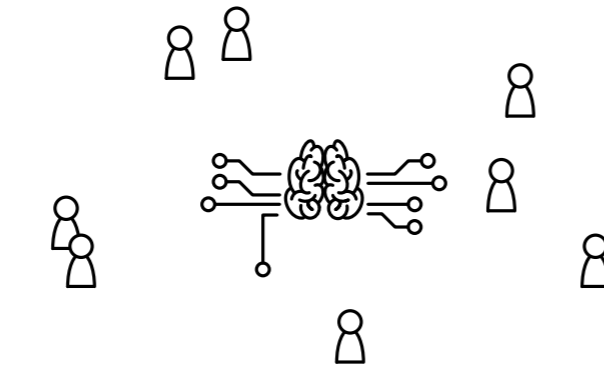
To this end, the project will provide (1) a holistic overview and empirical analysis of technological mediation, (2) a new and joint student module for sharing research insights and creating knowledge transfers in real time and (3) practical recommendations for assessing and mitigating distortions that can emerge during the design and use of AI.



Thierry Poibeau, Paris Artificial Intelligence Research Institute

Plans for 2023

In 2023, we will conduct qualitative interviews and run a workshop with industry and academic experts related to our research questions and create a new student module on ethics of AI and technological mediation to be hosted in Paris. In Paris, we will also host a workshop on the use of language models, following the ongoing discussions around ChatGPT and the new possibilities offered by language models, especially in the creative industry domain. Expanding on this, we will explore language models empirically, and perform a series of tests around the idea of bias, taking into account different kinds of possible biases (e.g. gender, racial, cultural bias, etc.).



ieai.sot.tum.de/research/utem-undistorted-technological-mediation

Research Associate Franziska Poszler presenting a framework to structure arranged vs. technology-inherent (im)perfections



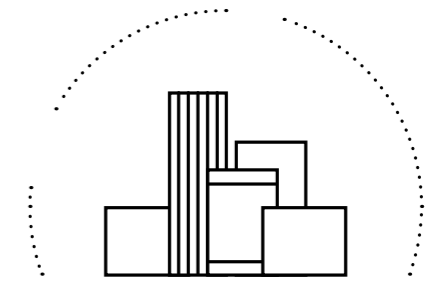
Ethics for the Smart City: Applied Socio-technical Frameworks to Assess the Implementation of AI-related Solutions

This project applies an interdisciplinary approach to answer the question, how can we assess the ethics of AI-enabled solutions implemented in urban settings from a socio-technical perspective? The project uses case studies and perspectives from locations around the world and will work with implementation partners to answer this question.

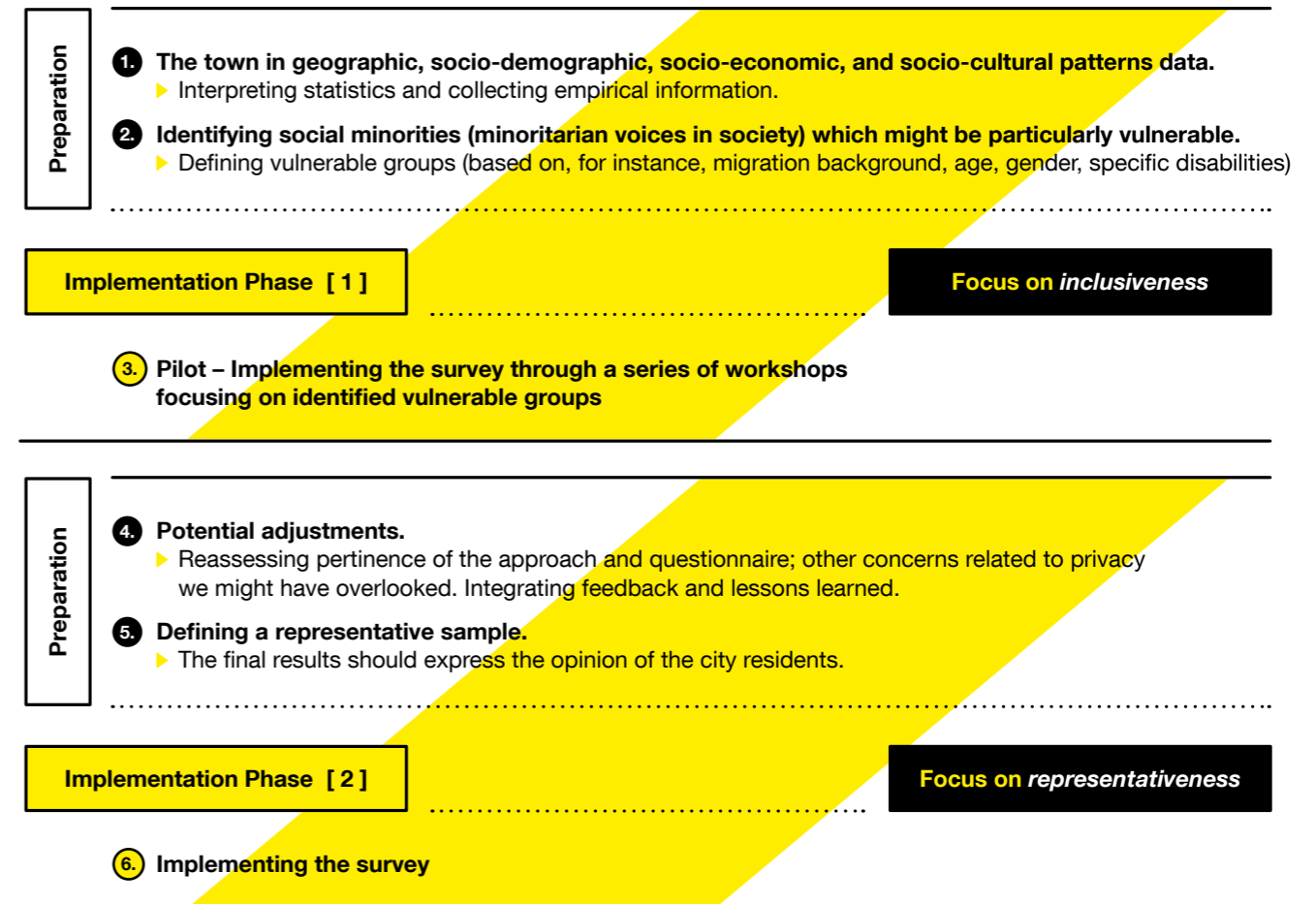
ieai.sot.tum.de/research/ethics-for-the-smart-city-applied-socio-technical-frameworks-to-assess-the-implementation-of-ai-related-solutions

Plans for 2023

In 2023, we will implement a series of workshops and surveys in collaboration with municipalities and local associations representing societal groups deemed vulnerable in order to assess the value of local privacy meanings for personal data and focusing on inclusiveness and representativeness.



Framing a public opinion survey as an inclusive tool for public engagement



Joint Projects and Partnerships

BB

20

No single country, ...

... industry or research institution can tackle the AI ethics challenges alone. To build on one another's efforts and advance the field of AI ethics as much as possible, the IEAI has collaborated since its inception with important stakeholders from many sectors. By joining forces with researchers and practitioners worldwide, the IEAI can address real-world challenges and contribute to the broader conversations and concerns surrounding ethics and AI on an international level.

In 2022, we established joint research projects, one with a European consortium working on a Horizon-Europe-funded, multi-year project and the other with our research partners at Fujitsu. We also continued our work with the AI for Earth Observation Future Lab.

Project **B1**

MELISSA – Mobile Artificial Intelligence Solution for Diabetes Adaptive Care

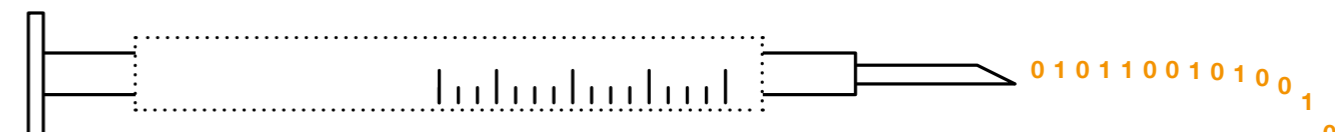
In 2022, the IEAI joined an interdisciplinary consortium of 12 international partners from Denmark, Germany, Greece, Spain, Switzerland, The Netherlands, and the US to work on the Horizon-Europe-funded project MELISSA – Mobile Artificial Intelligence Solution for Diabetes Adaptive Care.



The aim of the project is to innovate diabetes treatment and care for people with insulin-treated diabetes through the development of a clinically validated, efficient, and cost-effective Artificial Intelligence-based digital diabetes management solution.

“The MELISSA project has the potential to create an explainable AI-based solution, the first of its kind, for diabetes care that will empower patients by personalizing their diabetes management. We aim to help patients make better decisions about their diet and physical activities. With the app, we want to help calculate the right amount of insulin that needs to be given to a patient at the right time.”

Mbangula Lameck Amugongo
IEAI Researcher and Member of the MELISSA Team



“The Institute for Ethics in Artificial Intelligence will support the consortium partners by defining ethical principles that guide the design, development and deployment of the software solution. Key points to be addressed in the ethics framework are biases resulting in discrimination of vulnerable groups and the ethical treatment of patient data, as well as the design of human-machine interfaces. Furthermore, the IEAI will elaborate on the business ethical aspects pertaining to the commercialization of application.”

Alexander Kriebitz and **Raphael Max**
IEAI Researchers and Members of the MELISSA Team

Plans for 2023

In 2023, the IEAI will deliver ethical guidelines to the project partners that detail how to navigate ethical issues when developing, but also applying, the AI solution. We will focus specifically on the practical applications of AI in health and human rights' implications. ●

ieai.sot.tum.de/research/melissa-mobile-artificial-intelligence-solution-for-diabetes-adaptive-care

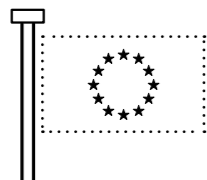
Towards an Accountability Framework for AI Systems

In 2022, the IEAI also continued to develop its strategic partnerships with its industry partners. This year, the collaboration between IEAI and Fujitsu ramped up its research aimed at driving the debate regarding the transparency, accountability, and traceability of AI.

The aim of this project is to bring more clarity to the definition and implementation of what “accountability” means for AI systems in a practical sense for organizations, and to look more closely at how the changes in the EU market that will arise in the near future with the EU AI Act can be managed. To support this movement towards a more responsible approach to AI and conform with ethical principles, the overall motive of our research is to define accountability in the context of AI and create an organizational framework on how to comply with it.

2022 Findings

- ▶ We have identified four core questions that need to be examined in order to move accountability from theory to practice: (1) Who is accountable? (2) For what is someone accountable and against whom? (3) How can the accountable entity ensure compliance with the identified duties? (4) How can the measures taken be satisfactorily explained?
- ▶ AI risk governance approaches need to meet five key requirements to be practically useful, mainly that they have balance, extendability, representation, transparency, and a long-term orientation.
- ▶ From the expert discussions in our workshops, we derived several ‘call for action’ items for regulators and organizations, such as providing clear definition of risks, providing standardization regarding risk governance, and providing clear accountability frameworks.



ieai.sot.tum.de/research/towards-an-accountability-framework-for-ai-systems

▶ *“I believe that accountability is a crucial aspect. If companies have a clear idea of accountability for AI systems and the results the systems produce, they will be more eager to implement them. Accountability, transparency, and explicability are probably the most critical issues right now to move AI forward. Therefore, I believe that this partnership could not have come at a better time.”*

Prof. Christoph Lütge / IEAI Director



Wordcloud survey on the perception of risk management in AI by experts

▶ *“Fujitsu is excited about the opportunity to partner with TUM and the IEAI on the topic of AI ethics. Independent academic research on the responsible use of AI is vital to creating a sustainable framework for technological development. Through this partnership, we are confident that we will be able to gain advanced technology in AI ethics, which is essential for global AI business.”*

Dr. Daisuke Fukuda / Head of the Fujitsu research center for Ethics in Artificial Intelligence

Plans for 2023

In 2023, we will focus on the implementation of our framework for accountability, answer questions such as: How can we more concretely implement this framework? And, which exact shape should the framework have for its contribution to the current landscape, its usability, and its understandability by all stakeholders? ●

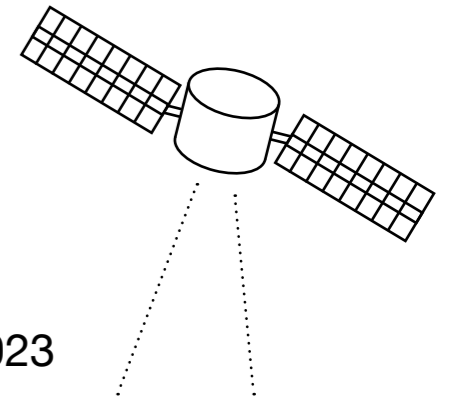
Artificial Intelligence for Earth Observation: Reasoning, Uncertainties, Ethics and Beyond (AI4EO)

The aim of this project is to develop ethics guidelines, tutorials, and approaches to identify ethical issues and opportunities in early stages of research for scientists working at the interface of Artificial Intelligence (AI) and Earth Observation/ Remote Sensing.

2022 Findings

- ▶ We created detailed illustrations from four AI4EO research fields (classification of slum areas; urban mapping and building level population density; monitoring (illegal) artisanal and small scale mining; monitoring and protecting biodiversity), to explain how scientists can redesign research questions to more effectively grab ethical opportunities to address real-world problems.
- ▶ We provided a step-by-step iterative process to understand and apply ethics in AI4EO research, including (1) identify your ethical duties, (2) align your research goals to these duties, (3) consider unethical consequences of research in short or long run, and (4) reframe your research questions in a manner that either avoids the unethical consequences or, preferably, shifts your research toward the accomplishment of an ethical opportunity.

Mrinalini Kochupillai
at The IEAI and TUM
Sustainability Day



Plans for 2023

In the coming year, the project will be developing novel curriculums for educating AI4EO scientists on identifying ethical issues in early stages of research and will present a tutorial at the IGARSS symposium (the most important conference on Geoscience and remote sensing globally) in July 2023. ●



ieai.sot.tum.de/research/artificial-intelligence-for-earth-observation

Continuing Research

24

Several of our ...

... longer term research projects continued in 2022. An impressive array of outcomes, in terms of publications, workshops, and conference presentations, were delivered over the year, solidifying the importance of our work for the application of AI ethics in practice.

Research

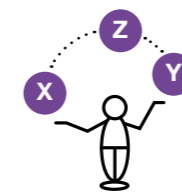
Project **C-1**

A Human Preference-Aware Optimization System

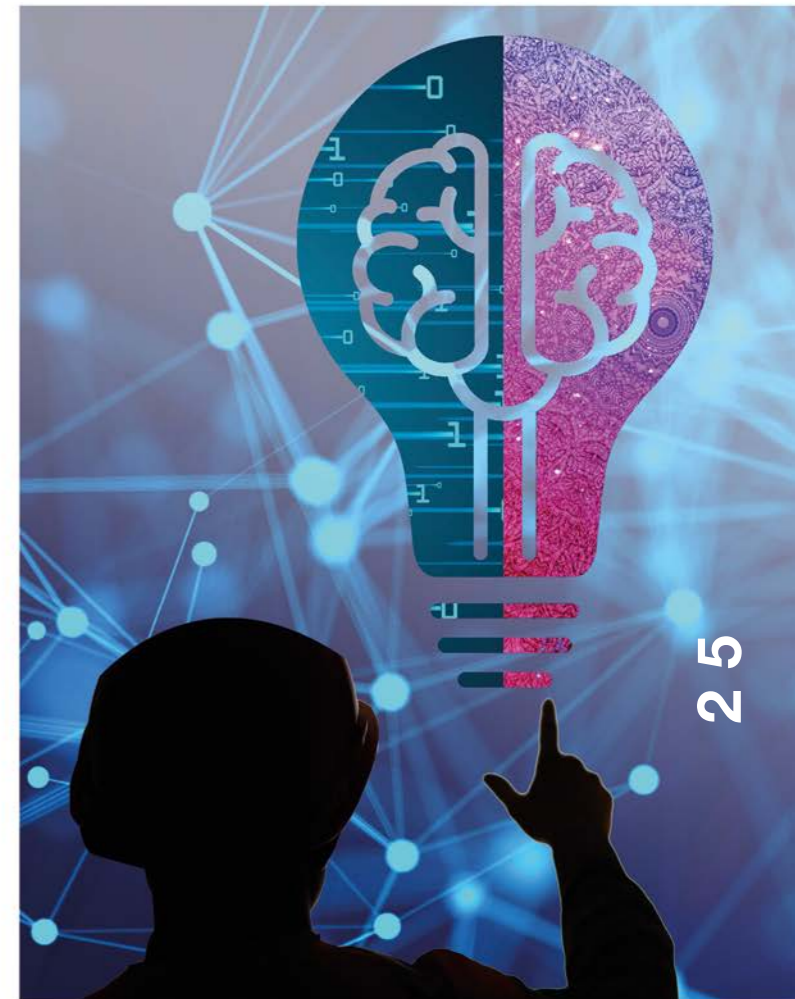
This project aims to explore whether and how AI can be used to optimize processes in the workplace in an ethical manner, promoting well-being by taking into account employees' needs and preferences (along with the employers') while respecting the fairness and privacy of employees.

2022 Findings

- ▶ Three dimensions of autonomy can be distinguished: psychological autonomy, value autonomy, and decision autonomy. The threefold distinction can be used to evaluate the use of new workplace technology.
- ▶ The use of algorithms at work for process optimization relates to the sense-making abilities and autonomy of employees, possibly affecting the meaningfulness of work. Employers should carefully scrutinize whether and how new workplace technology might affect meaningfulness.
- ▶ AI-based algorithms are also used in scheduling. However, these systems hardly provide any explanations as to why the shifts are scheduled in a certain way.
- ▶ Interviews with works councils from variously sized companies emphasized the importance of fairness in shift scheduling. However this notion often conflicts with workplace requirements that limit the positions in which employees can actual rotate.
- ▶ We found that to meet the requirement for fairness in shift assignment, a preference score helps calculate how many of an employee's preferences could be accommodated.



ieai.sot.tum.de/research/a-human-preference-aware-optimization-system



25

Plans for 2023

After two and a half years of project work, we have developed a solution for shift planning in logistics. There are many similar areas for which the solution is also suitable, for example hospitals. We are exploring the adaptability of our solution to other circumstances and industries. ●

TUM IEAI
ANNUAL REPORT

2022

ANDRE – AutoNomous DRiving Ethics

The aim of this project is to demonstrate one potential solution of how ethical behavior can be integrated into the trajectory and behavior planning of automated vehicles. By doing so, the investigated approach draws on and integrates various (traditional) ethical theories to provide guidance to automotive companies and policymakers on how risks can fairly be distributed between road users in traffic.

ieai.sot.tum.de/research/andre-autonomous-driving-ethics

2022: Findings

- ▶ Using a structured literature review and evaluation of applying various (traditional) ethical theories, we created an overview of the social, moral/legal, and functional advantages and disadvantages for each ethical identified theory.
- ▶ Given the theoretical options, an explicit ethical decision model was chosen to guide AVs during hazardous situations. The model aims to serve practitioners in the automotive sector by providing a first step towards a decision-making process that approximates compliance with ethical theories, shared principles and policymakers' demands.
- ▶ Explicit consideration of risk in trajectory planning leads to reduced risks for all road users. Ethical risk distribution leads to lower risks for vulnerable road users compared to a selfish risk distribution.

Plans for 2023

Going forward, an 'ethical vehicle' experiment with civil society will be conducted to test and validate the proposed ethical decision-making model for AVs. We will also continue the collaboration with AI4People to share expertise and research findings with policymakers. Finally, the developed ethical algorithm for trajectory planning will be deployed on the autonomous research vehicle EDGAR.

- ◀ left
Maximilian Geisslinger presenting on algorithmic challenges with automated vehicles at the Handelsblatt AI Summit
- ▶ right
Franziska Poszler giving a guest lecture at the Hochschule Bonn-Rhein-Sieg



Online Firestorms and Resentment Propagation on Social Media: Dynamics, Predictability and Mitigation

This project aims to establish a framework for mathematically modeling opinion formation within interconnected groups of social media users. To offer approaches on how to detect, react to, and possibly mitigate such negative dynamics at an early stage to prevent harm caused by social media, we investigate the driving forces which lead to negative dynamics at the social media group level.

ieai.sot.tum.de/research/online-firestorms-and-resentment-propagation-on-social-media-dynamics-predictability-and-mitigation

Findings of 2022

- ▶ Linguistic patterns and cues, such as the frequent appearance of certain pronouns or parts of speech, can be used to predict the outbreaks of online firestorms in real-time. Being an indispensable part of communication, and therefore readily accessible in written language, such features allow for more fine-grained analysis by investigating the sentiment in posts on social media networks.
- ▶ Our research identified a linguistic metric, lexical diversity, as an indicator of polarizing developments.
- ▶ Since toxicity tools, such as the Google Perspective API, were found to be lacking in detecting implicit misogyny, we devised a better model for recognizing explicit and implicit misogynistic language.
- ▶ We measured a broad linguistic synchronization in the form of hashtag usage across accounts that actively take part in a shared firestorm. During this time, firestorm-related hashtags become dominant across the timeline of almost all partaking users. This synchronization effect lasts for some days, even after the end of the firestorm.

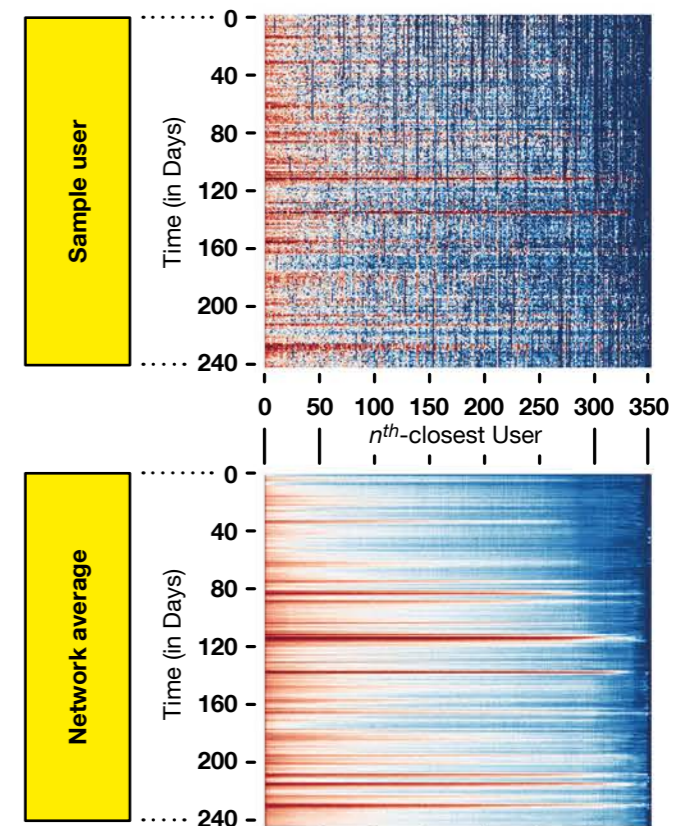
Closeness of users measured by their hashtag-usage.

- ▲ top All users are ordered by the overall similarity in hashtag-usage compared to some sample user.
- ▼ bottom The average of this process, repeated for every user, is displayed.

Firestorms heavily synchronize the network's hashtag-usage, but only for a short time.

Plans for 2023

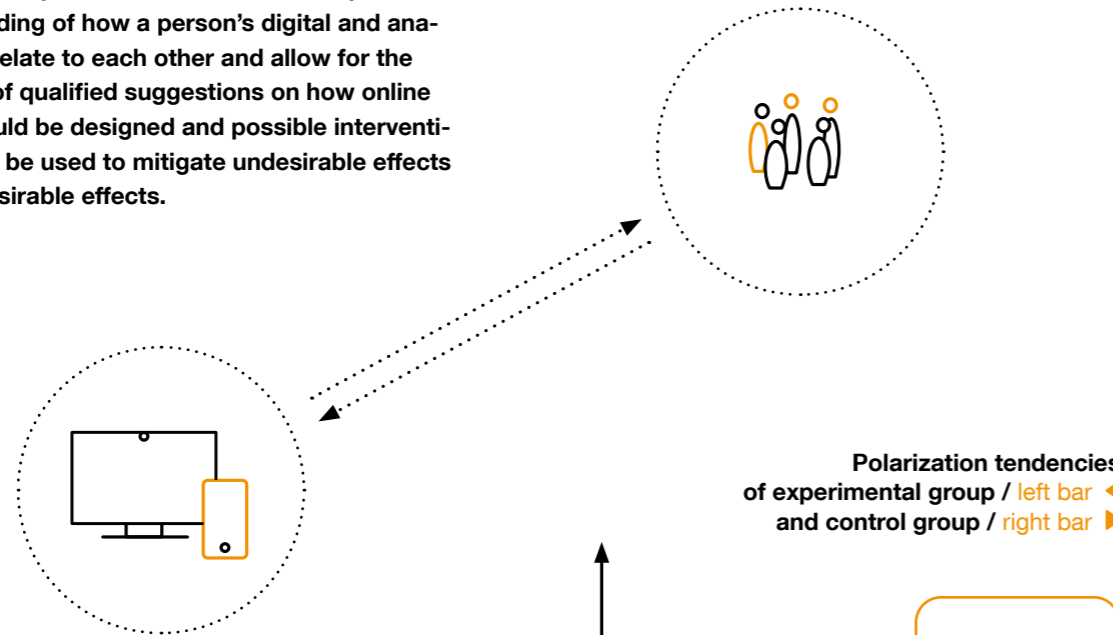
In 2023, we will develop and evaluate real-time monitoring systems that use linguistic patterns and cues to predict and prevent the outbreak of online firestorms. We will also investigate the effectiveness of time-out interventions as a way to interrupt the polarizing effects of continuous presence in polarizing online communities. Using linguistic metrics, such as lexical diversity, we will continue studying the language used in online firestorms and other polarizing online discussions.



Online-Offline Spillovers – Potential Real-World Implications of Online Manipulation

The goal of this project is to systematically investigate the effects of online experiences on offline attitudes and behavior, focusing on people's polarization tendencies that are triggered and promoted by approving feedback provided in social networks. An analysis of these effects will improve our understanding of how a person's digital and analog personas relate to each other and allow for the development of qualified suggestions on how online platforms should be designed and possible interventions that could be used to mitigate undesirable effects or promote desirable effects.

ieai.sot.tum.de/research/online-offline-spillovers-potential-real-world-implications-of-online-manipulation



Findings of 2022

- ▶ Linguistic features such as pronouns can be used to identify upcoming polarization effects in social media networks.
- ▶ Based on the literature, we developed a classification schema that covers explicit and implicit misogyny.
- ▶ Automated methods such as Google's Perspective API struggle when identifying implicit forms of misogynistic language.

Plans for 2023

In 2023, we will launch our main large-scale study and develop outputs from the results that discuss policy implications. ●

Personalized AI-Based Interventions Against Online Norm Violations

The aim of the project is to address how AI-based interventions can be applied to counter online norm violations and tailor them in personalized ways to enhance their effectiveness. We approach this by investigating what types of interventions are effective in terms of changing online behavior and what types of interventions are accepted by users.

- ▶ Deletion and public reprimand of sexist content in online communities are effective in restoring women's (and men's) feelings of safety. Yet deletion is seen as less fair, suggesting that users value transparency concerning the actions of moderators.
- ▶ Voluntary social media moderators would welcome automated tools to support their work, if those tools demonstrate accuracy, maintain moderators' agency to make final decisions, and are personalizable.

Findings of 2022

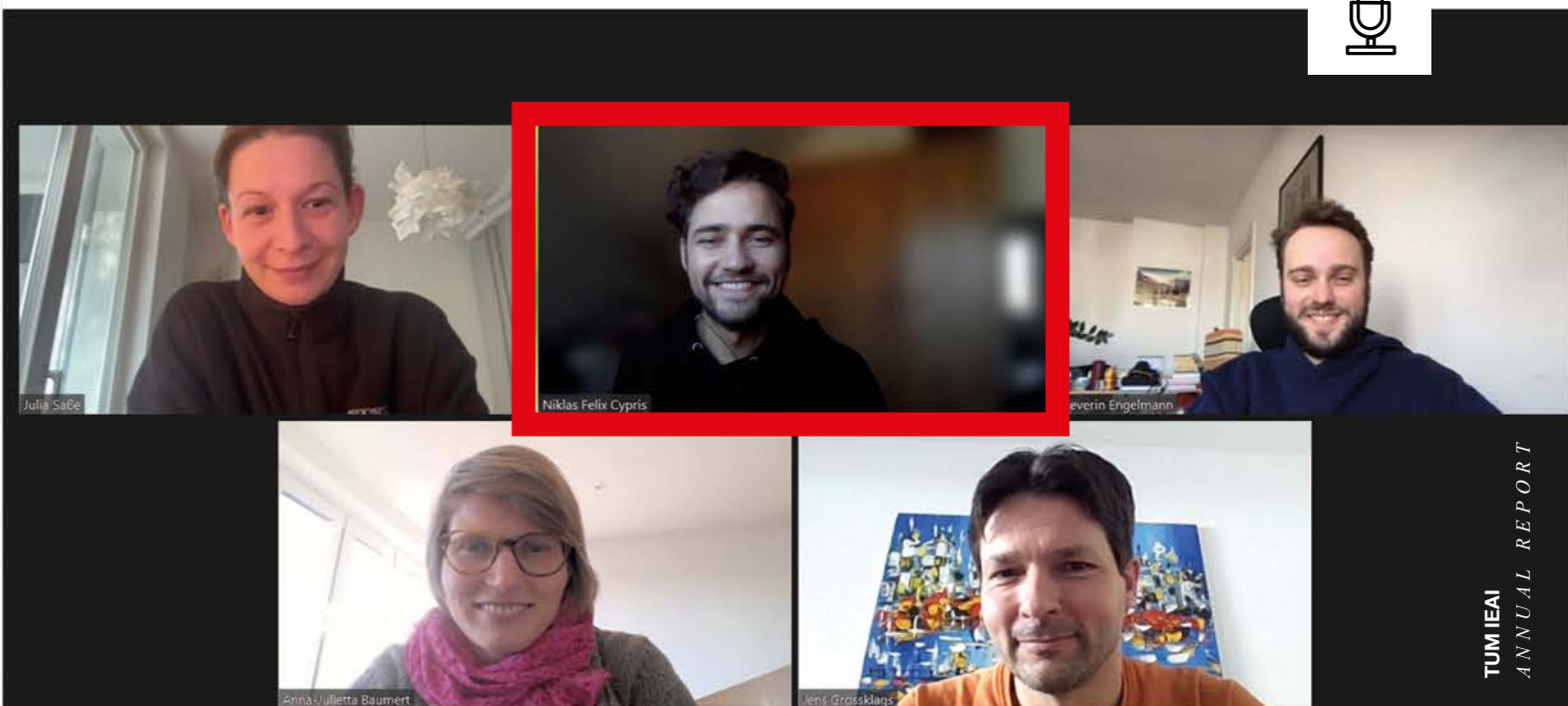
Bystanders are more likely to engage in counterspeech if they perceive that doing so is normative within their social group.

- ▶ Such norms can be shaped by single members of the group who speak up against hate speech.
- ▶ When deciding whether to punish norm violations, people are equally strongly deterred by the possibility of facing repercussions by humans or by automated actors.

Plans for 2023

In 2023, we will complete three ongoing lines of research, investigating which characteristics of interventions reduce the occurrence of hate speech, and empower targets of norm transgressions, as well as staking out to what extent users accept privacy infringements if they serve to motivate counter speech against hate speech. In addition, we will write a theoretical paper on ethical considerations with regards to counter speech on social media. ●

ieai.sot.tum.de/research/personalized-ai-based-interventions-against-online-norm-violations





Rule of Law, Legitimacy and Effective COVID-19 Control Technologies

The project aims to provide state of the art guidance for ensuring that the research and innovation using technological applications for controlling COVID-19 are compliant with the Rule of Law. Focusing on AI-assisted technology (principally mobile applications), the research addressed COVID-19 control strategies in the pandemic and post-pandemic phases.

- ▶ **IEAI White Paper** based on the project
- ▶ **Workshops on The Rule of Law in the Face of Contemporary Global Transformations**, November 2022



Findings for 2022

- ▶ It is important to acknowledge the critical link between trust and COVID compliance. An absence of trust will produce negative consequences for COVID control.
- ▶ Autocratic, and democratic states have different legitimacy styles. The research has identified four different types of legitimacy styles. They include popular legitimacy, democratic legitimacy, situational legitimacy, and performance legitimacy. The role and significance of trust feature to a varying degree depending on the legitimacy style in question.
- ▶ The arbitrary exercise of discretionary state powers is harmful to legitimacy in all authority styles and its exercise must be tampered. The Rule of Law enhances trust relations (even in authority styles where utility is the predominant motivation) through the reduction of arbitrariness.
- ▶ Adherence to the Rule of Law values, such as legal certainty and proportionality, among the other core components, helps in balancing the trade-offs between rights, public values, and public interests.
- ▶ Analysis from a Rule of Law perspective is relevant at each stage of the lifecycle of a new technology – from introduction to the implementation of data-driven technologies over the course of the emergency. Moreover, a Rule of Law analysis is also relevant for the technology ecosystem because pandemic technologies are embedded in wider digital architectures.



ieai.sot.tum.de/research/rule-of-law-legitimacy-and-effective-covid-19-control-technologies



Research

30

Research

The Ethics and Practice of AI Localism at a Time of COVID-19 and Beyond

The term AI Localism describes a new and radical shift in AI governance from the national to the local level. This project aims to identify different categories of successful and ethical AI Localism approaches while highlighting risks and challenges, including problematic AI applications and public concerns.

Findings of 2022

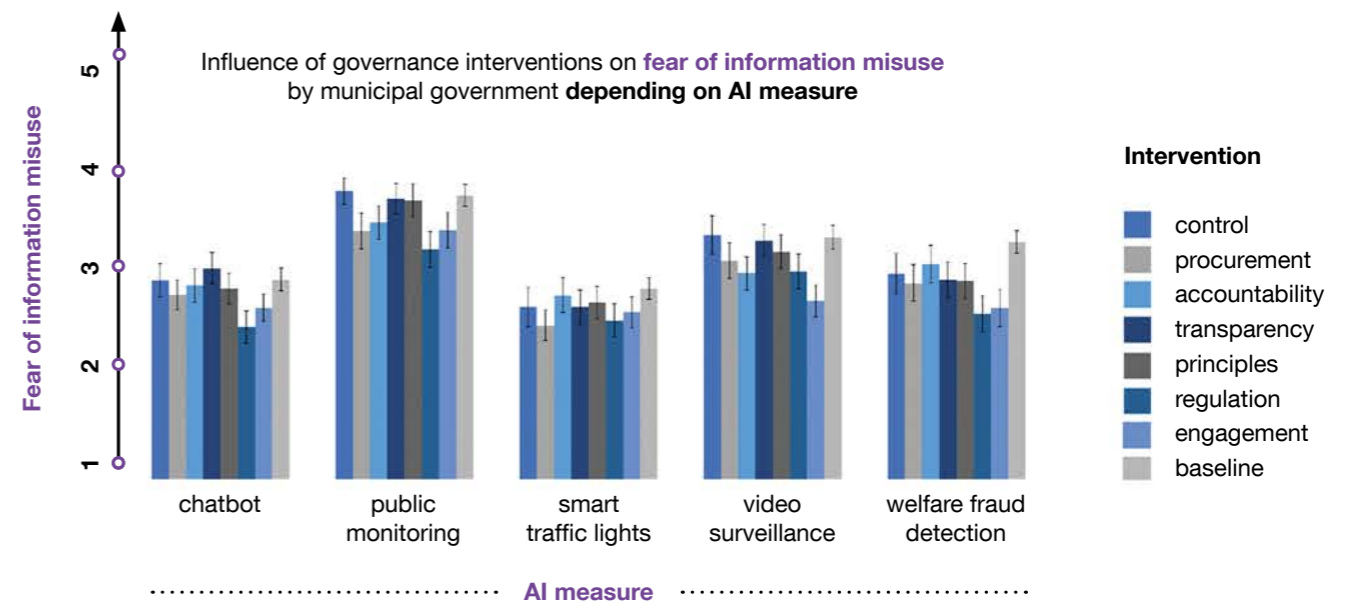
- ▶ The AI Localism Repository documents local AI measures around the world and how they are governed.
- ▶ Innovations in AI governance are described in the AI Localism Canvas, a framework for assessing the emergence of AI governance in cities.
- ▶ The results from two vignette studies suggest that there are differences in public opinion about the ethical aspects of different local AI measures and that innovative governance interventions could influence public acceptance, reservations, and fears about the use of AI technologies in cities.

Plans for 2023

The vast amount of data that we collected allows us to perform several additional analyses that help us to reveal the influence of particular governance interventions on certain AI technologies going forward. Furthermore, we will conduct the representative studies that were already done in the US, Germany, and Australia (and potentially other countries) to get a deeper understanding of intercultural differences that drive people's attitudes toward AI applications.

ieai.sot.tum.de/research/the-ethics-and-practice-of-ai-localism-at-a-time-of-covid-19-and-beyond

Fear Of Misuse By AI Measure





AI is a rapidly *changing* technology that embeds itself at all levels of society, raising new *challenges*, risks, dilemmas, and ethical concerns while repositioning and reprioritizing its responsible use. That makes AI ethics an enormously important field.

Applying Ethics

We asked our researchers specifically how their projects engage with ethical questions around AI.

► A variety of AI-based technologies are currently changing the field of education and learning. We analyze and evaluate these shifts from an ethics perspective and provide tools for learners and educators to navigate associated ethical challenges.

P A1 *Co-designing a Risk-Assessment Dashboard for AI Ethics Literacy in EdTech*

► In our project, we utilize a research perspective that focuses on geoethics and the decolonization of data, promoting awareness and explainability to ALL stakeholders in AI.

P A4 *The Potential for AI in the Extractive Industries to Promote Multi-objective Optimization*

► Accountability can be defined as the sum of taking responsibility and giving suitable explanations for the processes of an AI system to the choices made by a company. It is thus at the centre of ethical behavior on every level.

P B2 *Towards an Accountability Framework for AI Systems*

► We assess the ethicality of using AI recommender algorithms to amplify counterspeech to increase their viewership and the ethicality of AI-based user profiling methods to facilitate personalized counterspeech measures.

P C5 *Personalized AI-Based Interventions Against Online Norm Violations*

► What is 'and 'what could be' aspects of a specific AI model can only be figured out after the fact, yet both are essential to frame the thinking around 'what should be' or normative aspects of that model.

P A2 *Idiomorphic AI: Emergent, Tailored, and Normative Behavior in Large Language Models*

► The overall aim is to ensure that Smart City related solutions relying directly on AI or on urban big data potentially feeding AI-based models are people-centered and ethically assessed.

P A7 *Ethics for the Smart City: Applied Socio-technical Frameworks to Assess the Implementation of AI-related Solutions*

► We are aware of the potential risks and dangers associated with modeling opinion dynamics and we recognize that these models could be misused to manipulate online discussions and influence public opinion. We aim to understand how these tools might be used to start or reinforce firestorms instead of mitigating them and how we can suppress such uses.

P C3 *Online Firestorms and Resentment Propagation on Social Media: Dynamics, Predictability and Mitigation*

► Our project contributes to fostering the acceptability of AI applications by systematically incorporating people's moral intuitions into the development of AI governance.

P C7 *The Ethics and Practice of AI Localism at a Time of COVID-19 and Beyond* ●

► While there are already substantive reflections on ethical NLP, we will answer the prompting questions when the technology is applied in such a crucial and delicate area of society like the legal sector.

P A3 *NLawP: Natural Language Processing and Legal Tech*

► Our participation in the Melissa project is not just about the establishment of ethical standards for the app, but rather about identifying the wider ethical issues in the context of AI in health. This is necessary when considering the lack of ethical standards for AI in health.

P B1 *MELISSA – Mobile Artificial Intelligence Solution*

Assessing Impact

How does your project team strategize about the impact of your work?

► We aim to preempt the coming commercial deployment of personalized AI applications by creating and studying a prototype so that we can issue guidelines for developers and regulators before such technology becomes mainstream.

P A2 *Idiomorphic AI: Emergent, Tailored, and Normative Behavior in Large Language Models*

► By creating an AI-based modeling system that can be adapted for use in practice by companies, they will be able to use AI to improve their socio-environmental impacts, rather than only monetary gains.

P A4 *The Potential for AI in the Extractive Industries to Promote Multi-objective Optimization*

► With this project, we aim at producing both theoretical and practical outputs. On the one hand, we will produce scientific papers that we plan to publish in the best venues (conferences and journals). On the other hand, we plan to produce practical guidelines that will be of direct use to practitioners in the field.

P A6 *UTEM – Undistorted Technological Mediation*

► The impact of our work will be underscored by the engagement of external partners, which may leverage directly (e.g., municipalities and local organizations) or broaden the outreach (e.g., international organizations).

P A7 *Ethics for the Smart City: Applied Socio-technical Frameworks to Assess the Implementation of AI-related Solutions*

► From our ongoing work and communication with Fujitsu, and frequent interactions with companies through our workshops, we have the chance to see AI stakeholders' interest in our work grow. Ideally, our project raises awareness and guides organizations in the uptake of an applied accountability framework.

P B2 *Towards an Accountability Framework for AI Systems*

► We are focusing on educating the relevant group(s), i.e. scientists and researchers working in AI for earth observation. Over time, we expect that as the community becomes more familiar with the importance of ethics, we will see a rapid increase in the number of publications and conference presentations on the topic or on related topics.

P B3 *AI for Earth Observation (AI4EO)*

► The implementation of our newly developed firestorm detection methods has the potential to significantly enhance our ability to identify and mitigate potential firestorms early on. By understanding the negative dynamics in these situations, we can gain valuable insights into the potential danger they pose and take action to mitigate them. These tools provide a valuable resource for safeguarding against the negative impacts of firestorms and promoting healthy online discourse.

P C3 *Online Firestorms and Resentment Propagation on Social Media: Dynamics, Predictability and Mitigation*

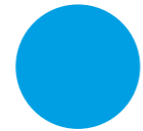
► In our study, we conduct a lab-in-the-field experiment. This means that we can observe naturally occurring human decision-making behavior in natural decision-making environments (social media platforms) and can thus draw direct conclusions about the behavior of the subjects and the implications for their non-experimental environment.

P C4 *Online-Offline Spillovers – Potential Real-World Implications of Online Manipulation*

► Our White Paper summarizes learnings from our several original findings and translates them into easily accessible style for policy makers. In this way, we have made our learnings accessible for future decisions.

P C6 *Rule of Law, Legitimacy and Effective COVID-19 Control Technologies*

Promoting Engagement



How does your project engage with relevant stakeholders outside of academia?

► Stakeholder participation is a core design principle of the project. In 2022, we engaged with educators at the Munich International School. We also began recruitment for co-design workshops with educational stakeholders.

P A1 *Co-designing a Risk-Assessment Dashboard for AI Ethics Literacy in EdTech*

► Interactions with the local communities in our selected case studies in Africa serve to test and verify the validity of our analysis and the applicability and feasibility of the projections and recommendations.

P A4 *The Potential for AI in the Extractive Industries to Promote Multi-objective Optimization*

► We aim to host expert workshops, offer a new student module, publish articles in journals, and present our newest insights at international conferences. Therefore, we hope to include and inform various stakeholders such as academia, policymakers, students, and technology companies about our study.

P A6 *UTEM – Undistorted Technological Mediation*

► In the process of designing and refining our ethical guidelines for the Melissa project, we aim to reach out to international standardization organizations, as well as governmental institutions that are related to health care, including the Bavarian Ministry of Health. We aim also to get in closer contact with companies developing and using AI solutions in practice.

P B1 *MELISSA – Mobile Artificial Intelligence Solution*

► We involved the stakeholders relevant to the project – employees, managers, and works councils – from the very beginning. In total, we visited almost ten different companies and gained many different perspectives from employees and managers in small and large companies.

P C1 *A Human Preference-Aware Optimization System*

► “With our approach to risk-aware trajectory planning we present the first planning algorithm that is in line with the recommendations from the EU expert group. The implementation can now be taken up by the industry as a concrete proposal to comply with future legal requirements.”
[Maximilian Geisslinger]

P C2 *ANDRE – AutoNomous DRiving Ethics*

► We are committed to working closely with different social media researchers to identify ways in which our findings can help mitigate negative dynamics and promote healthy online discourse. We engage with risk assurance experts to explore the financial risks that firestorms may pose for companies. By sharing our insights and collaborating with these stakeholders, we aim to contribute to the overall goal of creating a safer and more positive online environment.

P C3 *Online Firestorms and Resentment Propagation on Social Media: Dynamics, Predictability and Mitigation*

► We engage with civil society across multiple channels such as providing expert opinions for TV segments, publications by the telecommunication providers, and public panel discussions.

P C5 *Personalized AI-Based Interventions Against Online Norm Violations*

► We have analyzed practices and policies from jurisdictions around the world and accumulated unique comparative knowledge. Our White Paper feeds these learnings back to decision makers.

P C6 *Rule of Law, Legitimacy and Effective COVID-19 Control Technologies*

KNOWLEDGE LEADERSHIP DISSEMINATION.



Knowledge Dissemination.

In an effort to move from theory to practice, the IEAI has placed a strategic importance on using a variety of media to make sure ALL stakeholders are able to learn from and utilize our research.

➔ **“AI Localism in Practice: Examining How Cities Govern AI”**, S. Marcucci, U. Kalkar, S. Verhulst in: SSRN

➔ **“A Fuzzy-Cognitive-Maps Approach to Decision-Making in Medical Ethics”**, A. Hein, L.J. Meier, A.M. Buyx, K. Diepold in: IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)

➔ **“Accommodating Employee Preferences in Algorithmic Worker-Workplace Allocation. Human Factors in Management and Leadership”**, C. Haid, S. Stohrer, C. Unruh, T. Bütthe, J Fottner, in: Proceedings of the Applied Human Factors and Ergonomics [Engineering] Conference (AHFE)

➔ **“AI-Powered Public Surveillance Systems: Why We (Might) Need Them and How We Want Them”**, C. Fontes, E. Hohma, C. Corrigan, C. Lütge in: Technology in Society

➔ **“Algorithms as Partners in Crime: A Lesson in Ethics by Design”**, S. Krügel, A. Ostermaier & M.Uhl in: Computers in Human Behavior

➔ **“Algorithms for Ethical Decision-Making in the Clinic: A Proof of Concept”**, L. Meier, A. Hein, K. Diepold, A. Buyx in: The American Journal of Bioethics

➔ **“Clinical Ethics – To Compute, or Not to Compute?”**, L. Meier, A. Hein, K. Diepold, A. Buyx in: The American Journal of Bioethics

➔ **“Closing Pandora’s Box on Naver: Toward Ending Cyber Harassment”**, N. Kang, T. Kuo & K. Grossklags in: Proceedings of the 16th International AAAI Conference on Web and Social Media (ICWSM)

➔ **“Creating a Digital Marketplace for Agrobiodiversity and Plant Genetic Sequence Data: Legal and Ethical Considerations of an AI and Blockchain Based Solution”**, M. Kochupillai, and J. Köninger in: H. Williamson and S. Leonelli Eds. Towards Responsible Plant Data Linkage: Global Challenges for Food Security and Governance. Springer Nature.

➔ **“Digitization and Business Ethics”**, C. Lütge, M. Uhl, R. Max, A. Kriebitz in: Springer

➔ **“Earth Observation and Artificial Intelligence: Understanding emerging ethical issues and opportunities”**, M. Kochupillai, M. Kahl, M. Schmitt, H. Taubenböck, X. Zhu in: IEEE Geoscience and Remote Sensing Magazine

➔ **“Ending the Myth of Mobility at Zero Costs: An External Cost Analysis”**, D. Schröder, L. Kirn, J. Kinigadner, A. Loder, P. Blum, Y. Xu, M. Lienkamp in: Research in Transportation Economics

➔ **“Ethics of AI-Enabled Recruiting and Selection: A Review and Research Agenda”**, A. Hunkenschroer, C. Lütge in: Journal of Business Ethics

➔ **“Evaluation of Conversational Agents: Understanding Culture, Context and Environment in Emotion Detection”**, M. Teye, Y. Missah, E. Ahene, T. Frimpong, A. Boch in: IEEE Access

➔ **“Geo-Information Harvesting from Social Media Data”**, X. Zhu, Y. Wang, M. Kochupillai, M. Werner, M. Häberle in: IEEE Geoscience and Remote Sensing Magazine (in press)

➔ **“Hiding Behind Machines: When Blame is Shifted to Artificial Agents”**, T. Feier, J. Gogoll, & M. Uhl in: Science and Engineering Ethics.

➔ **“How Prosocial is Moral Courage?”**, J. Sasse, M. Li, A. Baumert in: Current Opinion in Psychology

➔ **“Human Autonomy in Algorithmic Management”**, C. Unruh, C. Haid, J. Fottner, T. Bütthe, in: Proceedings of 2022 AAI / ACM Conference on Artificial Intelligence, Ethics, and Society (AIES’22)

➔ **“Identifying Lexical Change in Negative Word-of-Mouth on Social Media”**, W. Strathern, R. Ghawi, M. Schönfeld, Jürgen Pfeffer in: Social Network Analysis and Mining

➔ **“Individuality and Fairness in Public Health Surveillance Technology: A Survey of User Perceptions in Contact Tracing Apps”**, E. Hohma, R. Burnell, C. Corrigan, C. Luetge in: IEEE Transactions on Technology and Society

➔ **“Inter-Cultural Cooperation: The Role of Attitudes, (Shared) Expectations, and Behavioral Standards”**, G. Walkowitz, A. Belianin & A.R. Dorrough in: Frontiers in Psychology

➔ **“International Cooperation on Artificial Intelligence – A Problem Statement”**, A. Kriebitz, Christoph Lütge in: Turkish Policy Quarterly

➔ **“Is AI Recruiting (Un)ethical? A Human Rights Perspective on the Use of AI for Hiring”**, A. Hunkenschroer, A. Kriebitz in: AI and Ethics

➔ **“Lessons Learned from Co-governance Approaches – Developing Effective AI Policy in Europe”**, C. Corrigan in: The 2021 Yearbook of the Digital Ethics Lab

➔ **“Optimal pricing and investment in a multi-modal city – Introducing a macroscopic network design problem based on the MFD”**, A. Loder, M. C. J. Bliemer & K.W. Axhausen in: Transportation Research Part A: Policy and Practice

➔ **“People Prefer Moral Discretion to Algorithms: Algorithm Aversion Beyond Intransparency”**, J. Jauernig, M. Uhl & G. Walkowitz in: Philosophy and Technology

➔ **“‘Proof under reasonable doubt’: Ambiguity of the Norm Violation as Boundary Condition of Third-Party Punishment”**, D. Toribio-Flórez, J. Sasse, Anna Baumert, in: Personality and Social Psychology Bulletin

➔ **“The AI Localism Canvas: A Framework to Assess the Emergence of Governance of AI within Cities”**, S. Verhulst, A. Young & M. Sloane in: Informationen zur Raumentwicklung

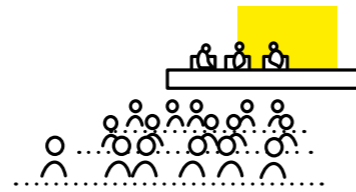
➔ **“The German Act on Autonomous Driving: Why Ethics Still Matters”**, A. Kriebitz, R. Max, C. Lütge, in: Philosophy & Technology

➔ **“The Polarizing Impact of Continuous Presence on Users’ Behavior”**, W. Strathern, A. Mooseder & J. Pfeffer in: Proceedings of the third MEDiate Workshop (ICWSM 2022), Atlanta, Georgia, June 6-9

➔ **“The Rule of Law as the Perimeter of Legitimacy for COVID-19 Responses”**, J. Grogan & J. Beqiraj in: Routledge Handbook of Law and the COVID-19 Pandemic.

➔ **“Zombies in the Loop? Humans Trust Untrustworthy AI-Advisors for Ethical Decisions”**, S. Krügel, A. Ostermaier & M. Uhl in: Philosophy & Technology

Conference Highlights



► **A Human Preference-Aware Optimization System**

- ▷ Gesellschaft für Analytische Philosophie Konferenz, Berlin
 - ▶ 12 - 15 September
- ▷ AAI/ACM Conference on Artificial Intelligence, Ethics, and Society, Oxford
 - ▶ 1 - 3 August
- ▷ International Conference on Applied Human Factors and Ergonomics, New York
 - ▶ 24 - 28 July
- ▷ Society for Applied Philosophy Annual Conference, Edinburgh
 - ▶ 1 - 3 July

► **ANDRE - AutoNomous DRiving Ethics**

- ▷ Handelsblatt KI Summit, Essen
 - ▶ 8 - 9 June
- ▷ 1st Emergency-VRD Workshop on moral dilemmas involving self-driving cars, University of the Bundeswehr Munich, Neubiberg
 - ▶ 7 October
- ▷ The International Society of Business Economics and Ethics (ISBEE) World Congress, Bilbao
 - ▶ 20 July
- ▷ TÜV SÜD Tagung Automatisiertes Fahren, Erding
 - ▶ 3 March
- ▷ Leverhulme Centre for the Future of Intelligence, Cambridge
 - ▶ https://www.youtube.com/watch?v=UUjWpPs_PM
 - ▶ 20 May
- ▷ Lecture series „Technik- und Umweltethik“ at the Hochschule Bonn-Rhein-Sieg
 - ▶ <https://technik-umwelt-ethik.de/allgemein/videos-2022>
 - ▶ 14 April

► **Artificial Intelligence for Earth Observation: Reasoning, Uncertainties, Ethics and Beyond (AI4EO)**

- ▷ Alexander von Humboldt's 12th Brazil-German Frontiers of Science and Technology Symposium, Maceo
 - ▶ 29 June - 2 July
- ▷ ML4Earth xAI Workshop co-organized by Technical University of Munich and Helmholtz AI, Munich
 - ▶ 22 November

- ▷ 2nd Annual Conference of the Indo-German Centre for Business Excellence, Frankfurt
 - ▶ 22 September
- ▷ Helmholtz AI Conference 2022, Munich
 - ▶ 2 June

► **Co-designing a Risk-Assessment Dashboard for AI Ethics Literacy in EdTech**

- ▷ AI Goes Edu – a Blessing or a Curse?, Fireside Chat with Dr. Elisabeth Sylvan, Harvard University, TUM Think Tank
 - ▶ 12 December
- ▷ Forum TUMpaed and Guests, TUM, SOT, Department of Educational Sciences
 - ▶ 25 November

► **Ethics for the Smart City: Applied socio-technical frameworks to assess the implementation of AI-related solutions**

- ▷ Ethics in AI Summer School, organized by the Goethe Institute and the University of Bamberg
 - ▶ 26 - 30 September
- ▷ Smart City Expo World Congress 2022
 - ▶ 15 - 17 November

► **Online Firestorms and Resentment Propagation on Social Media: Dynamics, Predictability and Mitigation**

- ▷ Sunbelt Conference 2022 – International Network for Social Network Analysis, Cairns
 - ▶ 12 - 16 July
- ▷ 3rd Mediate Workshop 2023, International Conference on Web and Social Media (ICWSM 2022), Atlanta
 - ▶ 5 - 9 June
- ▷ Evostar Conference, Complutense University of Madrid
 - ▶ 20-22 April
- ▷ Franco-German Workshop on Artificial Intelligence, Inria Headquarters, Versailles
 - ▶ 14 June
- ▷ ALOP Workshop on Algorithmic Optimization and Data Science, University of Trier
 - ▶ 20 - 22 July
- ▷ MCML Kick-Off Event, Bavarian Academy of Sciences and Humanities, Munich, Germany
 - ▶ 27 July

- ▷ Summer School on Applied Harmonic Analysis and Machine Learning, University of Genoa
 - ▶ 5 - 9 September
- ▷ Workshop “Matematica per l'Intelligenza Artificiale e il Machine Learning: Giovani ricercatori”, Politecnico di Torino
 - ▶ 24 - 26 November
- ▷ Workshop on Recent Advances in Nonlocal Kinetic, Fluid and Diffusive PDEs, Liaoning University, (online)
 - ▶ 26 - 27 November
- **Online-Offline Spillovers – Potential Real-World Implications of Online Manipulation**
 - ▷ Sunbelt Conference 2022 – International Network for Social Network Analysis, Cairns
 - ▶ 12 - 16 July
 - ▷ 3rd Mediate Workshop 2022, International Conference on Web and Social Media, Atlanta
 - ▶ 5 - 9 June
 - ▷ European Conference on Social Networks (EUSN)
 - ▷ Algorithms & Economic Behavior (AIEcon), Hamburg University of Technology
 - ▶ 29 - 30 September
 - ▷ Munich-Passau Workshop 2022, University of Passau
 - ▷ „Der Algorithmus als Partner oder Komplize: Ethik konformes Design von Mensch-Maschine-Interaktion“, Hamburg University of Technology
 - ▶ 16 September
 - ▷ Better together – Das Zusammenspiel von Mensch und Maschine, Bayerisches Institut für Digitale Transformation (bidt)
 - ▶ 29 November
- **Personalized AI-Based Interventions Against Online Norm Violations**
 - ▷ Interdisciplinary online meeting of the German political psychology network
 - ▶ March
 - ▷ Conference of Experimental Psychologists, Cologne
 - ▶ March
 - ▷ The Aarhus '22 Conference on Online Hostility and Bystanders
 - ▶ June
 - ▷ Annual Meeting of the International Society of Political Psychology. Athens
 - ▶ July

- **The Ethics and Practice of AI Localism at a Time of COVID-19 and Beyond**
 - ▷ PPE Annual Conference, New Orleans
 - ▶ November

► **The Potential for AI in the Extractive Industries to Promote Multi-objective Optimization**

- ▷ Research writers' workshop in Chaminuka, Zambia
 - ▶ October
- ▷ 2nd Conference on Land Governance and Societal Challenges, Namibia University of Science and Technology
 - ▶ December

► **Towards an Accountability Framework for AI Systems**

- ▷ EIT Health Germany-Switzerland and EIT Health Spain Joint Symposium on Right to privacy vs. duty to protect lives: Health Data Management in Focus
 - ▶ November
- ▷ Seminar on AI Ethics with the the British Consulate-General Munich and UK Innovate
 - ▶ September

► **Tradeable Mobility Credits: Addressing Ethical Concerns with Algorithm Transparency**

- ▷ Workshop on transparency and consumer behavior at Humboldt University of Berlin
 - ▶ October
- ▷ Introduction of the new GoTransTech at the School of Management Research Fest
 - ▶ November
- ▷ TUM Think Tank Discussion
 - ▶ November

► **UTEM – Undistorted Technological Mediation**

- ▷ 38th EGOS Colloquium, WU Vienna
 - ▶ 8 July
- ▷ PSL Intensive week on Ethics and Artificial Intelligence
 - ▶ November
- ▷ 24th PRAIRIE seminar: Artificial Intelligence and Society: What would a better AI mean?
 - ▶ November ●

Research Briefs

Outreach is an incredibly important component of the IEAI mission. Research on AI ethics needs to translate to practice. Otherwise, we cannot maximize the impact of our work. Publications and outreach relate to the results of our research, therefore making up an element of our work. To this end, the IEAI itself produces the Research Brief Series. The aim of these publications is to highlight important topics in AI Ethics and explore the ethical considerations related to them.

FEB **Towards an Accountability Framework for AI: Ethical and Legal Considerations**
by Auxane Boch, Ellen Hohma, Rainer Trauth

APR **Intervening Against Online Hate Speech: A Case for Automated Counterspeech**
by Niklas Felix Cypris, Severin Engelmann, Julia Sasse, Jens Grossklags, Anna Baumert

JUN **Algorithmic Scheduling in Industry: Technical and Ethical Aspects**
by Charlotte Franziska Unruh, Charlotte Haid, Johannes Fottner, Tim Bütthe

JUL **Considering the Responsible Use of AI in the Military**
by Paloma Laye and Caitlin Corrigan

AUG **AI Auditing: The Uncharted Waters of AI Research**
by Laura Lucaj

OCT **Reflecting on AI ethics – Perspectives from the Global AI Ethics Consortium**
by Anastasia Aritzi and Caitlin Corrigan

DEC **The Use of AI in the Mining Industry – Insights and Ethical Considerations**
by Caitlin Corrigan and Paloma Laye



These succinct reviews are aimed at a broad audience with the goal of informing the public on the role of AI in society and the ethical challenges that accompany this expanding role. In 2022, the IEAI team published seven Research Briefs. Here are some questions our researchers tried to answer:

What are the major accountability problems for AI systems from a legal and ethical perspective? Can automated counterspeech serve as a supplement to deletion-based approaches? What ethical principles can guide the design of algorithmic scheduling systems? What questions do we need to ask to ensure ethics and societal impacts are considered in using AI in the military? Why are algorithmic AI auditing procedures necessary? How do we ensure AI is used to benefit communities, societal development, and environmental sustainability in the mining industry? Where is AI Ethics going?



New White Paper Series

As part of the IEAI's goal to ensure that research informs practice, we launched our White Paper Series to make our AI governance approaches, frameworks and researcher results available to practitioners and policy makers.

[1] **On a Risk-Based Assessment Approach to AI Ethics Governance**

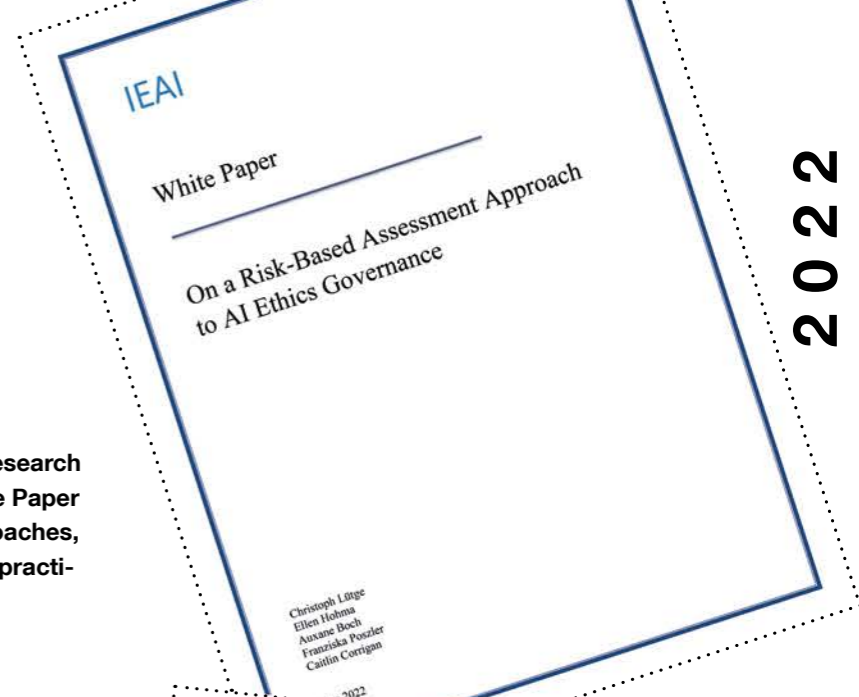
The IEAI published its first White Paper of the new IEAI White Paper Series in June 2022. The paper, entitled *On a Risk-Based Assessment Approach to AI Ethics Governance*, outlined a first step towards operationalizing an applied "risk-based assessment approach" to AI ethics governance.

[2] **Rule of Law, Legitimacy and Effective COVID-19 Control Technologies**

In July, the IEAI published its second White Paper *Rule of Law, Legitimacy and Effective COVID-19 Control Technologies*. This White Paper aimed to provide an overview of the assessment of the technological applications in terms of their legitimacy and amenability to Rule of Law standards ensuring respect for rights and liberties. It is oriented toward the policy makers who consider introducing surveillance technologies for health purposes and would like to learn more about how to assess this from a Rule of Law perspective.

[3] **Towards an Accountability Framework for Artificial Intelligence Systems**

In August, the IEAI published its third White Paper *Towards an Accountability Framework for Artificial Intelligence Systems*. This White Paper aims to define the steps that need to be taken to achieve accountability in the context of Artificial Intelligence accelerated systems and builds on previous work by industry and international organizations.



Thesis Projects

The IEAI seeks to explore the ethical issues related to the development, use, and impact of AI technologies. From the beginning, our goal was to create a research institute that inspires, promotes, and supports students to focus on AI Ethics and conduct research that focuses on practical and action-oriented outcomes.

For a third year, the IEAI team worked systematically with undergraduate and graduate students at TUM to supervise thesis projects related to AI ethics and the responsible use of AI. Here is what our students have to say about the IEAI and their experience in writing their theses. ●

► Student **Mehdi Fekari** 1
 Master Thesis **Moral Machine Case for Health Care Robot: Assessment of the Public Moral Acceptance of Elderly Care Robots' Privacy Related Decision Based on Cultural Background.**
 Supervisor **Auxane Boch**

“As a Software Engineer, writing a master’s thesis at IEAI was a great opportunity to understand ethical challenges that go beyond the technical aspects and see IA innovations from new perspectives.”

► Student **Immanuel Klein** 2
 Bachelor Thesis **Impacts of AI on Human Rights**
 Supervisor **Auxane Boch**

“My experience in writing my Bachelor’s thesis at the IEAI was really good, everything went down perfectly, and I was very content with how well I was supervised.”

► Student **Nicole Seimebua** 4
 Master Thesis **The Ethical Conceptualization of AI in German Smart City Programmes - An Analysis of AI Transparency and Explainability**
 Supervisor **Ellen Hohma**

“Throughout the entire process of my master thesis, I felt very supported by my supervisor Ellen. She always gave me constructive feedback which not only helped me with my thesis but fostered my entire academic skill-set. I am really glad to have had the chance to write my thesis at the IEAI.”

► Student **Amy Ndiaye Sow** 3
 Master Thesis **No Bot Is Perfect: A Theory of Mind Approach to Tackling Bias in Human-Robot Interactions**
 Supervisor **Auxane Boch**

“Writing my master’s thesis at the IEAI was exactly as I expected: I got to do research about an exciting and socially relevant topic with the support of an all-in-one supervisor, coach and mentor. Thanks to Auxane and the IEAI institute for this great experience!”

► StudentS **Anna Eroshkina** 5
 Master Thesis **Legitimate justification for “autonomous” decisions – Drawing parallels between humans and algorithms.**
 Supervisor **Franziska Poszler**

“My supervisor’s insights, tools’ suggestion, and interest in the research itself were helpful. Following a schedule of virtual and in-person sessions has been an important part of my smooth experience. Under the supervision of Franziska Poszler, gaps in legislation of AI decisions were identified for its further development”



Q&A Series: Reflections on AI

In our Q&A Series: Reflections on AI, we have the chance to speak to experts from different countries and backgrounds. These were just some of the questions we could ask the ten experts during the brief interviews.

How does the physical anthropomorphism of a robot impact people's attitudes towards it and why? What are the major potential challenges and opportunities that AI-powered legal technology may bring to traditional legal practice? How will AI change the future of work? What are the major ethical concerns related to the use of AI for humanitarian work? What is it particularly about AI ethics that demands global discussions? What do militaries need to do to maximize the benefits of AI-enabled systems while mitigating the risks they entail? How can we strike a balance between the goals of regulation with liberties to innovate? Will the AI Act have a major influence on how other countries deal with AI regulation?



- [1] Reflections on AI** with **Christoph Bartneck & Dwain D. Allan**
"Robots are not just machines. They represent us without being us."
- [2] Reflections on AI** with **Thomas Hildebrandt & Henrik Olsen**
"The most important question is how to develop methods for trustworthy development of AI systems."
- [3] Reflections on AI** with **Natsuko Tokumaru**
"AI can extend the possibilities of diverse working styles and make these working styles more flexible."
- [4] Reflections on AI** with **Manuel García-Herranz**
"It is very important to focus on quantifying inequalities."
- [5] Reflections on AI** with **Aditya Johri**
"We need a much higher level of governance and some regulatory impetus when it comes to AI ethics."
- [6] Reflections on AI** with **David Barnes**
"We need to ask the question: should we be doing this?"
- [7] Reflections on AI Q&A** with **Christian Djefal**
"We need to find ways to ground democratic legitimacy."
- [8] Reflections on AI Q&A** with **Dragoş Tudorache**
"The whole debate about regulating AI, the principles of AI, has started from ethical concerns."

In the Media

Tech500 has identified our Director as ..

... the **number 23 most influential person** in the **German** tech sector and ...



... the **number 2 most influential person** within the AI and Data Science space across the three countries: **the UK, Germany, and France.**

Autobahnverkehr automatisieren ist sinnvoll

"I was a guest in the studio of Bayern 2 and spoke with Rolf Büllmann about the future of autonomous driving. In our conversation, we discuss, among other things, why 99% technology readiness is not enough and the role that ethics play in autonomous driving."
Maximilian Geißlinger

How will AI change the way we think about mobility? What are the ethical challenges? Is it possible to integrate ethical behaviour into autonomous vehicles? Our researcher Maximilian Geißlinger shared his views and expertise on autonomous driving.

Autonome Lastwagen machen Langstreckenfahrer arbeitslos

Maximilian Geißlinger was interviewed about autonomous transportation systems for the Frankfurter Allgemeine Zeitung (F.A.Z.)



Ancient wisdom, new knowledge – Research on ethics in AI.

"We, too, need a combination of ancient wisdom and new technology. We can combine them in research. In ethical terms, it would contribute to a better and more sustainable world."

Mrinalini Kochupillai

On behalf of the German Ministry for Education and Research, Olga Gilbert interviewed Professor Mrinalini Kochupillai about her ongoing research at the TUM Future Lab: AI for Earth Observation.

Trusting Untrustworthy Machines and Other Psychological Quirks

How do humans relate to AI decision-making systems? John Danaher invited Matthias Uhl to "Philosophical Disquisitions" Podcast.

read & talk: Cyborg – Zwischen Mensch und Maschine



Interview with the Director of the Institute for Ethics in Artificial Intelligence at the Technical University of Munich, Christoph Lütge by Marisol Chevez Hidalgo

"I think it is not possible to let technology alone automatically make decisions about what's right, because it wouldn't work. People with their human conscience must constitute the circle in which the final decisions are made, considering what is correct and what is not."
Christoph Lütge.

REACHING



OUT

Reaching Out.

Spanning sectors, disciplines and borders, the IEAI seeks to bring voices across the AI community together.

- ▶ With the Global AI Ethics Consortium (GAIEC) we bring a wider range of perspectives and expertise from around the world.
- ▶ With the Responsible AI Network (RAIN-Africa) we embrace the ecosystem's diversity by bringing more perspectives across African cultures to the conversation.
- ▶ With the AI Ethics: Global Perspectives free online course we raise public awareness of the benefits and risks of AI technologies.
- ▶ With our events we seek to encourage and facilitate multi-stakeholder dialogue and focus on a broad range of AI Ethics-related issues.
- ▶ With our presence on the Social Media Platforms we share the outcome of our work and inform our growing community about the latest developments regarding AI Ethics.

Global AI Ethics Consortium



Since its launch, the GAIEC has brought together 36 experts from academic and research institutions on six continents. Academia, with its open-mindedness and its broad access to different disciplines and cultures, is an ideal place to connect and build international research communities, collaborative networks, and global consortia working together on AI ethics-related issues.

At its annual meeting on May 30, the Global AI Ethics Consortium Members discussed priorities for AI ethics in 2022. Given the recent emphasis on regulatory mechanisms to mediate AI governance, it is important to consider both AI ethics' short- and long-term roles in this space.

Following this meeting, the IEAI team reached out to our GAIEC members and gathered their reflections on the most relevant aspects of AI ethics. In the special Research Brief, Reflecting on AI ethics – Perspectives from the Global AI Ethics Consortium, released in October 2022, 16 experts from diverse fields, sectors, and locations shed light on this issue's complexity and current relevance.



“Having discussions with experts from all over the world is the only way to understand these issues better and gain a holistic view of the limitations, ethical risks, and benefits of using AI.”

Christoph Lütge

ieai.sot.tum.de/global-ai-ethics-consortium

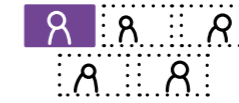
“AI ethics means many different things to different people, given their backgrounds and interests. Although this diversity is a good thing as it makes the field fertile, it also poses the danger of resulting in incommensurable perspectives where no one really agrees with anyone else and therefore, there is no common ground to move the field forward.”

Aditya Johri

GAIEC Members

- ▶ Christoph Lütge ▶ John Basl
- ▶ Ricardo Baeza-Yates ▶ Christoph Bartneck
- ▶ Richard Benjamins ▶ Celina Bottino
- ▶ Rafael A. Calvo ▶ Javier Camacho Ibáñez
- ▶ Cansu Canca ▶ Jessica Eynard
- ▶ José-Luis Fernández-Fernández
- ▶ Christian Fieseler ▶ Mark Findlay
- ▶ Luciano Florid ▶ Jean-Gabriel Ganascia
- ▶ Nicole Gillespie ▶ Benedetta Giovanola
- ▶ Jason Grant ▶ Ken Ito ▶ Aditya Johri
- ▶ Mohammad Ershadul Karim
- ▶ Jean-Marie Jacquet ▶ Jerry Kponyo
- ▶ Catherine Muñoz ▶ Viviana Polisena
- ▶ Yves Pouillet ▶ Jeannie Marie Paterson
- ▶ Thierry Poibeau ▶ Huw Price
- ▶ Kanshukan Rajaratnam ▶ Emma Ruttkamp-Bloem
- ▶ Idoia Salazar ▶ Nitin Sawhney ▶ Nidhi Singh
- ▶ Fabro Steibel ▶ Kan Hiroshi Suzuki
- ▶ Makoto Usami ▶ Shannon Vallor
- ▶ Stefaan Verhulst ▶ Tanya De Villiers-Botha
- ▶ Adrian Weller ▶ Yi Zeng

AI Ethics Course



The AI Ethics: Global Perspectives course, launched in February 2021, is a joint project of the IEAI, the Global AI Ethics Consortium, The Governance Lab, NYU Tandon School of Engineering, and the Center for Responsible AI at NYU.

Our free online course is designed to raise awareness of technology's societal impacts and give individuals and institutions the tools to pursue the responsible and ethical use of AI and data. A year and a half later, the course has grown to 50 modules, contributed by 56 faculty members representing over 25 countries.

“Our conversations with faculty members and our experiences with the course modules have yielded a wealth of knowledge about AI Ethics. In keeping with the values of openness and transparency that underlie the course, we summarized these insights into ten learnings to share with a broader audience. Here are our learnings”

Sampriti Saxena and Stefaan G. Verhulst
 '10 learnings from considering AI Ethics through global perspectives'

The AI Ethics: Global Perspectives course in numbers [as of December 2022]

- ▶ 50 modules on AI and data ethics
- ▶ 56 instructors from 25 countries around the world
- ▶ 5 webinar series
- ▶ 11 faculty panels
- ▶ 1,800+ newsletter subscribers
- ▶ 10,000+ views on YouTube

Through the AI Ethics: Global Perspectives course, we have created a platform to facilitate this discourse by engaging the many perspectives that exist in the ecosystem. We will continue to bring together voices from around the world and across industries to foster these important conversations around ethics and AI.

More information about the course:
aiethicscourse.org

The Challenge and Potential for a Global Framework

Ask the Right Questions Broaden the Conversation

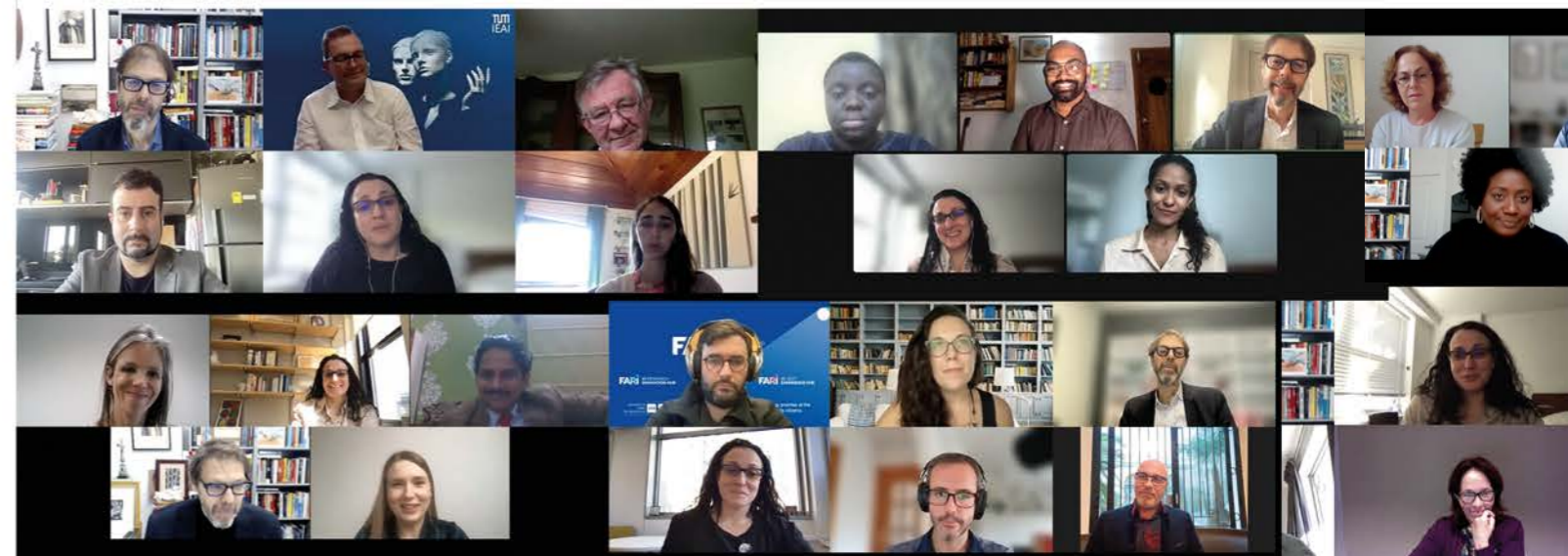
Establishing Trust The Role of Independent Research

Centering Diversity and Inclusion in Ethics Our Shared Responsibility

The Public as a Stakeholder

Humans at the Center

Building Effective Systems of Accountability



Responsible AI Network (RAIN-Africa)



The IEAI supports and participates in several networks related to AI Ethics. The Responsible AI Network (RAIN-Africa) is the outcome of our close collaboration with the Faculty of Electrical and Computer Engineering and the Responsible AI Lab at Kwame Nkrumah University of Science and Technology (KNUST) in Ghana.

Since its launch in 2020, ...

RAIN-Africa has brought together important stakeholders from different sectors to discuss and build joint projects on the ethical and social challenges arising at the interface of technology and human values. During these three years, RAIN Africa has provided a platform to promote cross-regional and interdisciplinary discussions regarding the responsible development and use of AI in Africa.



The focus of the Responsible AI Network Africa is to be a strong voice for the ethical and responsible use of AI.

Jerry Kponyo

In 2022, ...

RAIN-Africa organized three webinars and one on-site event shedding light on important issues.

[April 2022]

RAIN Africa hosted its first workshop for the year on the theme “Responsible AI Applications in Africa”. We had the pleasure of hearing from three researchers from Ghana. **Henry Nunoo-Mensah**, Uganda, **Rose Nikibuule** and Burkina Faso, **Malo Sadouanouan**.

Why should AI be responsible? To ensure that AI solutions are delivered with integrity, equity respecting individuals and always being mindful of the social impact of what we do.

Henry Nunoo-Mensah



[June 2022]

Jerry John Kponyo, conducted a compelling presentation on “Responsible AI and Ethics from an African Perspective” at the Institute for Ethics in Artificial Intelligence offices in Munich.

Education systems will need to adapt quickly and new frameworks need to be created for workers and citizens to develop the skills they need to thrive.

Jerry Kponyo

Peter Martey Addo delivered an exciting presentation on “Advancing The Responsible Adoption of AI at The Local Level”.

Identify the needs that matter. Promote open public dialogues and include everyone in the discussions. Diversity and inclusion matter.

Peter Martey Addo

[November 2022]

RAIN-Africa organized a workshop on the theme “Responsible AI and Data Protection: Regulatory Landscape Across Africa”. The invited panellists, **Teki Akuetteh**, **Arthur Gwagwa** and **Bev Townsend** shared their insight on the regulatory landscape and measures across Africa in the context of AI.

Stakeholder engagement is very important.

Bev Townsend

We have very few experts in Africa that understand the law, AI and have the technical expertise.

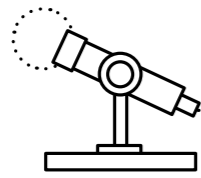
Teki Akuetteh

The data that is being collected today from Africa is going to give even more power to those who possess and use it.

Arthur Gwagwa



IEAI Speaker Series



With its Speaker Series, the TUM Institute for Ethics in Artificial Intelligence invites experts from all over the world to talk about the ethics and governance of AI. These events serve as an important platform for sharing new research and exchanging knowledge. Since our launch in 2019, we have been able to run 27 virtual or in-person Speaker Series events. In 2022, our institute had the pleasure of organizing three in-person and six virtual sessions.

Our 2022 IEAI Speaker Series brought us:

- ▶ 9 events
- ▶ 10 distinguished speakers
- ▶ coming from 9 countries
- ▶ and 4 continents

Focusing on:

- ▶ future of work
- ▶ human cognition
- ▶ AI regulation
- ▶ robots
- ▶ legal tech
- ▶ algorithms for humanitarian and development work
- ▶ AI in military
- ▶ sandboxes
- ▶ EU AI Act

In-person

We had the opportunity to host three in-person events at TUM.

[April 2022]

Changing Human Resource Management in Japan through AI-enabled Technology and Related Ethical Issues

Natsuka Tokumaru

"I wouldn't say that humans will be completely excluded from the economy and society by AI because in our society we have a lot of challenges and there are many fields that only humans can do."



[May 2022]

AI, Ethics and Human Cognition
Aditya Johri

"AI magnifies social relations and tensions and is the lens that forces society to examine existing faultlines and vulnerabilities."



[November 2022]

Regulating AI: A Comparative View
Christian Djeflal

"Trustworthy AI can mean different things. Also, there are many questions on how you govern a general purpose technology."



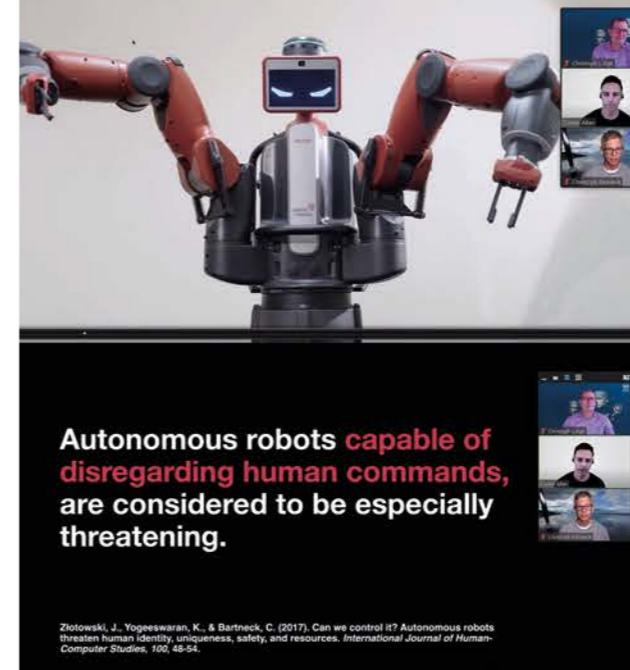
Virtual

We also had the opportunity to host six virtual events. Because of these events, we were able to reach an even broader audience from around the globe.

[February 2022]

Perceived Threat Of Robots
Christoph Bartneck

"Is abusing a robot agent seen as more morally acceptable than abusing a human? Is reactive aggression more acceptable when it comes from the human agent than from the robotic one?"



Autonomous robots capable of disregarding human commands, are considered to be especially threatening.

Dwain Allan

"Incremental theorists perceived #robots to be significantly less threatening to human identity and uniqueness. Entity theorists perceived robots to pose significantly more threat to employment, resources, and well-being."

[March 2022]

AI and Legal Tech: An Ethical Perspective
Thomas Hildebrandt

"A sustainable development requires support in making the right decisions in compliance with increasingly complex legislation."

Henrik Olsen

"Decisions must be based on justifications."



[May 2022]

Algorithms for Humanitarian and Development Work
Manuel García-Herranz

"There is definitely a North/South divide in deriving benefits from AI. The digital gap grows with the speed of technology."



[June 2022]

We are the LIMFAC: Some Thoughts on Military AI
David Barnes

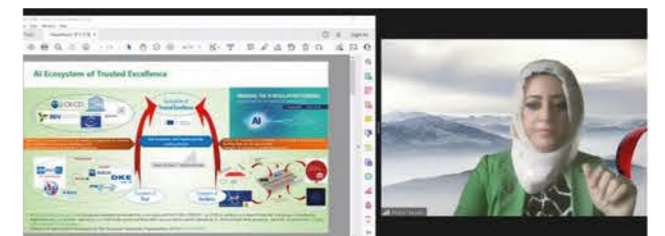
"We don't want to mask the concerns around the development of such technologies."



[October 2022]

AI Ecosystem of Trusted Excellence
Aisha Naseer

"AI regulatory sandboxes establish a controlled environment to test innovative technologies for a limited time on the basis of a testing plan agreed with the National Competent Authorities."



[December 2022]

Governance: the Key to Success in Regulating AI
Dragoş Tudorache

"If we want AI to be developed and accepted by people, companies, etc. we need to build trust. And this is what we are trying to do."



Workshops and Other Events



[June 2022]

The IBM and TUM IEAI co-hosted an in-person event to discuss the need to bring together actors from research and business to facilitate the development of **“Responsible Artificial Intelligence”**. The Responsible AI Soirée with IBM included research project pitches about AI explainability, AI fairness and AI governance, along with a panel discussion with AI experts. 1

[June 2022]

Artificial Intelligence in Agriculture: Ethical Concerns & Responsible Solutions was the topic of the joint online workshop hosted by AI Institute for Resilient Agriculture (AIIRA) and the Institute for Ethics in Artificial Intelligence.

[September 2022]

The IEAI co-organized by the British Consulate-General Munich and UK Innovate an in-person seminar on **AI Ethics for a delegation consisting of UK-based companies**. Among the topics discussed were the application of ethical standards in the development of AI-driven health solutions and ethical questions pertaining to autonomous driving. 2

[September 2022]

Three IEAI researchers, Auxane Boch, Ellen Hohma and Franziska Poszler, participated in the Global Summit on Responsible AI in Rio de Janeiro. During the conference, the research team from the IEAI presented a workshop on **“Assessing the Risks of AI Applications: A Hands-On Tool for AI Governance”** for experts from industry and academia as well as students. 3

[October 2022]

The IEAI and TUM Sustainability Day
The 27th of October 2022 marked the first TUM Sustainability day, which focused on demonstrating the steps taken toward the implementation of the sustainable strategy for the university. As a part of this, researchers from the IEAI presented their projects concerning this topic, such as smart cities and mobility, sustainable mining, and AI-enabled earth observation, in order to start a discussion on the role of technology in achieving sustainability goals in an in-person event with attracted a lot of attention. 4

[December 2022]

The IEAI and TUM Think Tank hosted an in-person panel discussion focused on the diverse **“Global Insights on AI Governance”**. Urs Gasser served as moderator to the speakers: Laetitia Onyejebu Elisabeth Sylvan and Armando Guio Espanol, who discussed the challenges to creating global approaches to AI governance that incorporate cultural distinctions and diversity in values and priorities. 5



Social Media

Staying in touch with our growing community. Giving and receiving feedback. Sharing news and the outcome of our work. Opening new opportunities for collaboration.

Some of the most popular posts on LinkedIn
(in terms of clicks and/or engagement rate/random order)

- ➔ Reflections on AI with Thomas Hildebrandt and Henrik Palmer Olsen
- ➔ Research Brief “Towards an Accountability Framework for AI Ethical and Legal Considerations”
- ➔ “The German Act on Autonomous Driving: Why Ethics Still Matters” Open Access paper
- ➔ AI for Children Toolkit
- ➔ White Paper entitled “On a Risk-Based Assessment Approach to AI Ethics Governance”
- ➔ 16 Recommended Free AI-Ethics & Data-Ethics and XAI Online Courses
- ➔ White Paper “Towards an Accountability Framework for AI Systems”
- ➔ 2021 IEAI Annual Report
- ➔ AI-powered public surveillance systems: why we (might) need them and how we want them Open Access Paper
- ➔ Reflecting on AI ethics-Perspectives from the Global AI Ethics Consortium
- ➔ “Mobile Computer Vision-Based Applications for Food Recognition and Volume and Calorific Estimation: A Systematic Review” Paper
- ➔ Governing AI – attempting to herd cats? Introduction to the special issue on the Governance of Artificial Intelligence article



More than **15,000** followers on **LinkedIn** and **Twitter**



More than **1,000** subscribers to our **Newsletter**



LinkedIn

[January - December 2022]

290+ posts and reposts

6000+ new followers

570K+ impressions

1,100+ reposts



Twitter @IEAITUM

[January-December 2022]

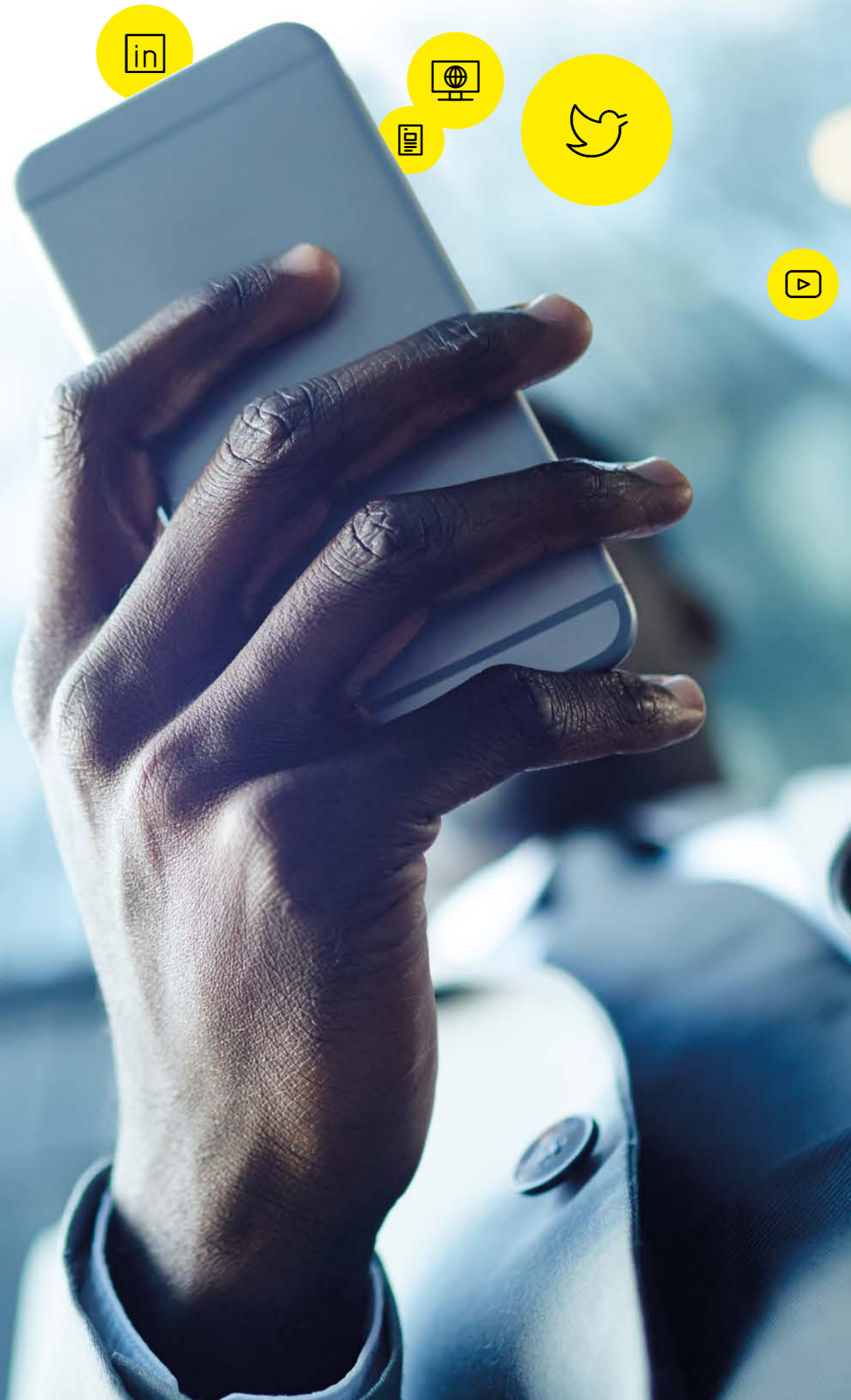
500+ tweets and retweets

740+ new followers

117 K+ profile visits

330 K+ tweet impressions

330+ mentions ●







IEAI Staff

- 1 ▶ Prof. Dr. Christoph Lütge / Director
- 2 ▶ Caitlin Corrigan, Ph.D. / Executive Director
- 3 ▶ Christina Daman / Program- & Event Manager
- 4 ▶ Manuela Fuchs / Program- & Event Manager
- 5 ▶ Auxane Boch / Research Associate
- 6 ▶ Ellen Hohma / Research Associate
- 7 ▶ Ana Catarina Fontes, Ph.D. / Research Associate
- 8 ▶ Dr. Lameck Amugongo / Research Associate
- 9 ▶ Dr. Alexander Kriebitz / Research Associate
- 10 ▶ Dr. Raphael Max / Research Associate
- 11 ▶ Maria Pokholkova / Research Associate
- 12 ▶ Franziska Poszler / Research Associate
- 13 ▶ Immanuel Klein / Student Assistant
- 14 ▶ Paloma Laye / Student Assistant
- 15 ▶ Anastasia Aritzi / Communications Consultant



Advisory Board

- 16 ▶ Prof. Dr. Petra Ahrweiler / Chair for Technology and Innovation Sociology, Johannes Gutenberg University Mainz
- 17 ▶ Auxiliary Bishop Anton Losinger / Episcopal Vicar („Bischofsvikar“) for Bioethics and Social Policy, Augsburg
- 18 ▶ Prof. Dr. Christoph Meinel / Director and CEO of the Hasso-Plattner-Institute for Digital Engineering at the University of Potsdam
- 19 ▶ Hannes Schwaderer / President of Initiative D21 e.V. & CEO of Intel Germany
- 20 ▶ Prof.Dr.-Ing.Dr.-Ing.E.h.Dr.h.c Dieter Spath / President of Acatech and Director of Fraunhofer IAO and IAT, University of Stuttgart
- 21 ▶ Zusanna Warso / Helsinki Foundation for Human Rights, Warsaw

Visiting Professors

- ▶ Prof. Aditya Johri / Professor of Information Sciences & Technology, George Mason University
- ▶ Dr. Natsuka Tokumaru / Professor of Economics, Ritsumeikan University
- ▶ Prof. Jerry John Kponyo / Professor of Telecommunication Engineering and Director of Responsible AI Lab, Kwame Nkrumah' University of Science and Technology
- ▶ Prof. Laetitia Onyejebu / Professor of Computer Science, University of Port-Harcourt

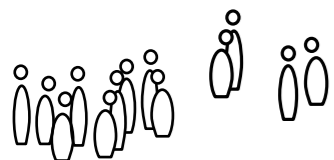
External Doctoral Researchers

- ▶ Johann Jakob Häußerman / Doctoral Student and Researcher, Center for responsible Research and Innovation, Fraunhofer IAO
- ▶ Raimund Waning / Doctoral Student and Researcher, clockworkx GmbH



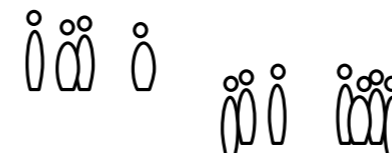
Project Members

Project	Principal investigators	Researchers
Co-designing a Risk-Assessment Dashboard for AI Ethics Literacy in EdTech A1	<ul style="list-style-type: none"> ▶ Prof. Dr. Urs Gasser, School of Social Sciences and Technology and Hochschule für Politik (HfP), TUM ▶ Prof. Dr. Matthias Grabmair, School of Computation, Information and Technology, TUM ▶ Prof. Dr. Anna Keune, School of Social Sciences and Technology, TUM 	<ul style="list-style-type: none"> ▶ Verena Kappes, School of Social Sciences and Technology, TUM ▶ Noha Halim, School of Social Sciences and Technology, TUM ▶ Pablo Gomez Ayerbe, School of Social Sciences and Technology, TUM <p><i>External Partner:</i></p> <ul style="list-style-type: none"> ▶ Dr. Lis Sylvan, Berkman Klein Center for Internet & Society, Harvard University
Idiomorphic AI: Emergent, Tailored, and Normative Behavior in Large Language Models A2	<ul style="list-style-type: none"> ▶ Prof. Dr. Simon Hegelich, Hochschule für Politik (HfP), TUM ▶ Prof. Dr. Christian Montag, Department of Molecular Psychology, Ulm University 	<ul style="list-style-type: none"> ▶ Saurabh Dhawan, M.Sc., Hochschule für Politik (HfP), TUM
NLawP: Natural Language Processing and Legal Tech A3	<ul style="list-style-type: none"> ▶ Prof. Dr. Christian Djeffal, School of Social Sciences and Technology, TUM ▶ Prof. Dr. Florian Matthes, School of Computation, Information and Technology, TUM 	<ul style="list-style-type: none"> ▶ Stephen Meisenbacher, School of Computation, Information and Technology, TUM ▶ Verena Müller, School of Social Sciences and Technology, TUM ▶ David Rebohl, School of Social Sciences and Technology, TUM ▶ Juraj Vladika, School of Computation, Information and Technology, TUM ▶ Oliver Wardas, School of Computation, Information and Technology, TUM
The Potential for AI in the Extractive Industries to Promote Multi-objective Optimization A4	<ul style="list-style-type: none"> ▶ Prof. Dr. Svetlana Ikonnikova, School of Management, TUM ▶ Caitlin C. Corrigan, Ph.D., Institute for Ethics in AI, TUM ▶ Prof. Dr. Ir. Walter Timo de Vries, School of Engineering and Design, TUM 	<ul style="list-style-type: none"> ▶ Gen Li, School of Management, TUM ▶ Dr. Arch. Pamela Durán Díaz, School of Engineering and Design, TUM



Project	Principal investigators	Researchers
Tradeable Mobility Credits: Addressing Ethical Concerns with Algorithm Transparency A5	<ul style="list-style-type: none"> ▶ Prof. Dr.-Ing Klaus Bogenberger, School of Engineering and Design, TUM ▶ Dr.sc. ETH Zürich Allister Loder, School of Engineering and Design, TUM ▶ Prof. Dr. Hanna Hottenrott, School of Management, TUM ▶ Prof. Dr. Christoph Ungemach, School of Management, TUM 	<ul style="list-style-type: none"> ▶ Lisa Hamm, School of Engineering and Design, TUM ▶ Dominik Fischer, School of Management, TUM
UTEM – Undistorted Technological Mediation A6	<ul style="list-style-type: none"> ▶ Thierry Poibeau, Ph.D., Paris Artificial Intelligence Institute ▶ Prof. Dr. Christoph Lütge, School of Social Sciences and Technology and Institute for Ethics in AI, TUM 	<ul style="list-style-type: none"> ▶ Aïda Elamrani, Paris Artificial Intelligence Research Institute ▶ Franziska Poszler, Institute for Ethics in AI, TUM
Ethics for the Smart City: Applied Socio-technical Frameworks to Assess the Implementation of AI-related Solutions A7	<ul style="list-style-type: none"> ▶ Prof. Dr. Christoph Lütge, School of Social Sciences and Technology and Institute for Ethics in AI, TUM ▶ Prof. Dr. Constantinos Antoniou, School of Transportation Systems Engineering, TUM 	<ul style="list-style-type: none"> ▶ Catarina Fontes, Ph.D., Institute for Ethics in AI, TUM ▶ Christelle Al Haddad, School of Engineering and Design, TUM <p><i>External Partners:</i></p> <ul style="list-style-type: none"> ▶ Gemeinde Kirchheim bei München ▶ The DFKI-Deutsches Forschungszentrum für Künstliche Intelligenz ▶ Ariadne Maps ▶ Georgios Pipelidis
MELISSA – Mobile Artificial Intelligence Solution for Diabetes Adaptive Care B1	<ul style="list-style-type: none"> ▶ Prof. Dr. Christoph Lütge, School of Social Sciences and Technology and Institute for Ethics in AI, TUM 	<ul style="list-style-type: none"> ▶ Dr. Alexander Kriebitz, School of Social Sciences and Technology and Institute for Ethics in AI, TUM ▶ Dr. Raphael Max, School of Social Sciences and Technology and Institute for Ethics in AI, TUM ▶ Dr. Lameck Amugongo, Institute for Ethics in AI, TUM
Towards an Accountability Framework for AI Systems B2	<ul style="list-style-type: none"> ▶ Prof. Dr. Christoph Lütge, School of Social Sciences and Technology and Institute for Ethics in AI, TUM ▶ Prof. Markus Lienkamp, Institute of Automotive Technology, TUM 	<ul style="list-style-type: none"> ▶ Ellen Hohma, Institute for Ethics in AI, TUM ▶ Auxane Boch, Institute for Ethics in AI, TUM ▶ Rainer Trauth, Institute of Automotive Technology, TUM

Project	Principal investigators	Researchers
Artificial Intelligence for Earth Observation: Reasoning, Uncertainties, Ethics and Beyond (AI4EO) B3	<ul style="list-style-type: none"> ▶ Prof. Dr. Xiaoxiang Zhu, Department of Aerospace and Geodesy, TUM ▶ Prof. Dr. Richard Hans Georg Bamler, Remote Sensing Technology Institute, DLR ▶ Prof. Dr. Massimo Fornasier, Department of Mathematics, TUM ▶ Prof. Dr. Christoph Lütge, School of Social Sciences and Technology and Institute for Ethics in AI, TUM 	<ul style="list-style-type: none"> ▶ Dr. Mrinalini Kochupillai, Professorship for Signal Processing in Earth Observation, TUM
A Human Preference-Aware Optimization System C1	<ul style="list-style-type: none"> ▶ Prof. Dr. Tim Büthe, School of Social Sciences and Technology, TUM ▶ Prof. Dr. Johannes Fottner, School of Engineering and Design, TUM 	<ul style="list-style-type: none"> ▶ Dr. Charlotte Unruh, School of Social Sciences and Technology, TUM ▶ Charlotte Haid, M.Sc., School of Social Sciences and Technology, TUM
ANDRE - AutoNomous DRiving Ethics C2	<ul style="list-style-type: none"> ▶ Prof. Dr. Christoph Lütge, School of Social Sciences and Technology and Institute for Ethics in AI, TUM ▶ Prof. Dr.-Ing. Markus Lienkamp, Institute of Automotive Technology, TUM 	<ul style="list-style-type: none"> ▶ Franziska Poszler, School of Social Sciences and Technology, TUM ▶ Maximilian Geisslinger, Institute of Automotive Technology, TUM
Online Firestorms and Resentment Propagation on Social Media: Dynamics, Predictability and Mitigation C3	<ul style="list-style-type: none"> ▶ Prof. Dr. Massimo Fornasier, School of Computation, Information and Technology, TUM ▶ Prof. Dr. Jürgen Pfeffer, School of Social Sciences and Technology, TUM 	<ul style="list-style-type: none"> ▶ Konstantin Riedl, School of Computation, Information and Technology, TUM ▶ Wienke Strathern, School of Social Sciences and Technology, TUM. ▶ Daniel Matter, School of Social Sciences and Technology, TUM.
Online-Offline Spillovers – Potential Real-World Implications of Online Manipulation C4	<ul style="list-style-type: none"> ▶ Prof. Dr. Jürgen Pfeffer, School of Social Sciences and Technology, TUM. ▶ Prof. Dr. Matthias Uhl, Faculty of Computer Science, Technische Hochschule Ingolstadt 	<ul style="list-style-type: none"> ▶ Wienke Strathern, TUM School of Social Sciences and Technology ▶ Gari Walkowitz, Faculty of Computer Science, Technische Hochschule Ingolstadt



Project	Principal investigators	Researchers
Personalized AI-Based Interventions Against Online Norm Violations C5	<ul style="list-style-type: none"> ▶ Prof. Dr. Anna Baumert, School of Human and Social Sciences, University of Wuppertal, ▶ Prof. Jens Grossklags, Ph.D., School of Computation, Information and Technology, TUM 	<ul style="list-style-type: none"> ▶ Niklas Felix Cypris, School of Human and Social Sciences, University of Wuppertal ▶ Dr. Slieman Halabi, School of Human and Social Sciences, University of Wuppertal ▶ Severin Engelmann, School of Computation, Information and Technology, TUM ▶ Tina Kuo, School of Computation, Information and Technology, TUM ▶ Dr. Julia Sasse, School of Computation, Information and Technology, TUM
Rule of Law, Legitimacy and Effective COVID-19 Control Technologies C6	<ul style="list-style-type: none"> ▶ Prof. Dr. Christian Djeffal, School of Social Sciences and Technology, TUM ▶ Prof. Dr. Mark Findlay, Centre for Artificial Intelligence and Data Governance (CAIDG); Singapore Management University ▶ Dr. Julinda Beqiraj, British Institute of International and Comparative Law, Bingham Centre for the Rule of Law 	<ul style="list-style-type: none"> ▶ Akanksha Bisoyi, School of Social Sciences and Technology, TUM ▶ Jane Loo, Centre for Artificial Intelligence and Data Governance (CAIDG), Singapore Management University ▶ Ong Li Min, Centre for Artificial Intelligence and Data Governance (CAIDG); Singapore Management University ▶ Anuj Puri, British Institute of International and Comparative Law, Bingham Centre for the Rule of Law
The Ethics and Practice of AI Localism at a Time of COVID-19 and Beyond C7	<ul style="list-style-type: none"> ▶ Prof. Dr. Matthias Uhl, Faculty of Computer Science, Technische Hochschule Ingolstadt ▶ Prof. Stefaan Verhulst, The GovLab and the Alliance for Public Interest Technology, New York University ▶ Prof. Jeannie Marie Paterson, Ph.D., Centre for Artificial Intelligence and Digital Ethics, University of Melbourne ▶ Mona Sloane, Ph.D., The GovLab and the Alliance for Public Interest Technology, New York University 	<ul style="list-style-type: none"> ▶ Yvette Maker, The Centre for Artificial Intelligence and Digital Ethics, University of Melbourne ▶ Manuela Schönmann, Faculty of Computer Science, Technische Hochschule Ingolstadt



Connect with us!

The IEAI annual report is published in English. More information about the IEAI can be found on our **Website** at ieai.sot.tum.de

Overall responsibility
Prof. Dr. Christoph Lütge

Published by

Technical University
of Munich,
Institute for Ethics
in Artificial Intelligence

Visitor's address

Marsstrasse 40
D-80335 Munich

Postal address

Arcisstrasse 21
D-80333 Munich

Tel. +49 89 289 22901
Email: ieai@sot.tum.de


Concept, design and production


Andreas Rintzner
www.RINTZNER.de

Photography


© Andreas Heddergott
© Sonja Herpich
© IBM


© Unsplash
© Envato Elements

 visit our **website**
ieai.SOT.tum.de

 subscribe to our **newsletter**
ieai.SOT.tum.de/contact

 **Twitter**
[@IEAITUM](https://twitter.com/IEAITUM)

 **LinkedIn**
linkedin.com/company/ieaitum

 **YouTube**
youtube.com/channel/UCdQrVqu4ewnFo5Xqm7xlrOg

As our projects ramp up in 2023,
so will our ability to provide meaningful content and consideration
across the AI ethics space.



**Ethical guidelines are
crucial for establishing
public trust and
elemental to the
success of AI overall.**

