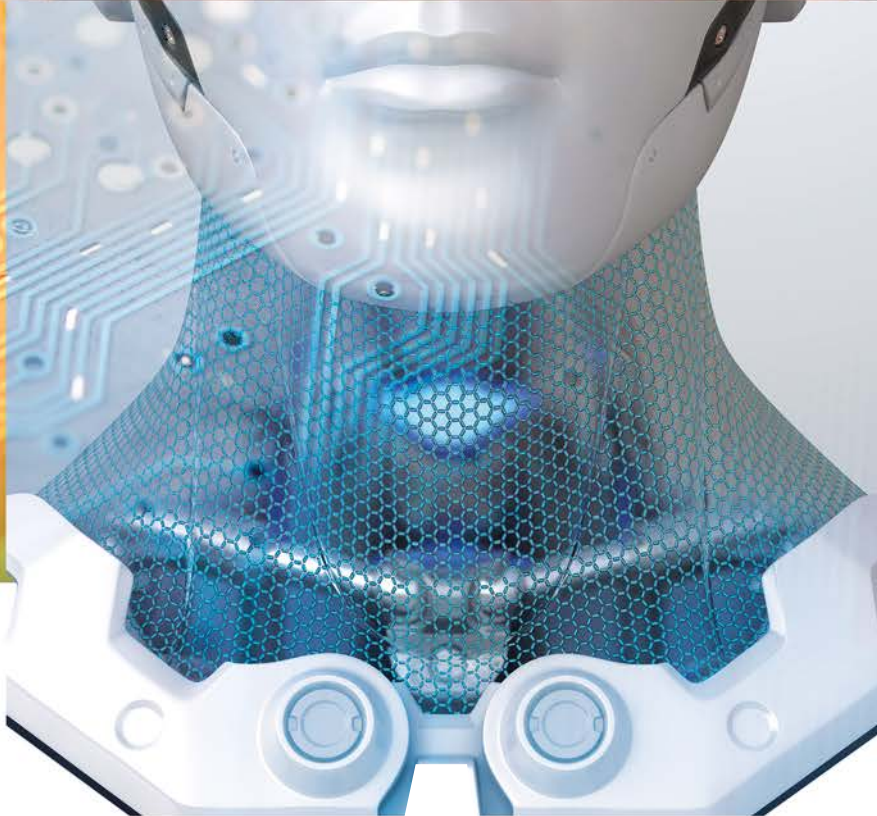2021

TUM
IEAI

TUM
Institute for Ethics
in Artificial Intelligence

There is hardly any part of society that will not be affected by AI.

TUM
Institute for Ethics in Artificial Intelligence

ANNUAL REPORT

2 0 2 1

ieai.sot.tum.de

Index

⊡ ◀······ **external links** // only in digital version

2 0 2 1   **TUM IEAI**
ANNUAL REPORT

Director's Message

Dear Readers,
Associates,
Colleagues and
Friends of the IEAI,

I proudly present to you our second IEAI Annual Report. In the following pages, you will have the chance to take a peek at some of our most significant accomplishments during 2021. But first, allow me to seize this opportunity to step back and reflect on the topic of AI ethics, the progress that has been made during the last couple of years, the IEAI's contribution and the top three reasons 2021 was a great year for our institute.

There is hardly any part of society that will not be affected by AI. Yet, AI Ethics did not exist as its own field of study until very recently, often seen as a sub-discipline of computer science. Subsequently, ethical questions were not sufficiently integrated into the AI-related discussions.

Because of this, we were among the pioneers who paved the way and underlined the need to address AI ethics-related challenges. The IEAI, founded three years ago, was one of the first Institutes of its kind. At that time, discussions around AI and AI Ethics were not high on the agenda; AI Ethics was not a trending topic. Therefore, we saw the need to act and create an institution that addresses these needs.

In the meantime, a lot of things have changed. AI Ethics and AI governance are taken much more seriously both at the European and at the international level. The growing number of institutes and centers working on the topic of AI Ethics, the large volume of research papers and the numerous AI ethics-related events taking place worldwide underscore the progress that is being made.

On an institute level, we have many reasons to say that 2021 was a great year. Allow me to highlight the top three.

FIRST, we further built upon the IEAI's reputation this year as a **trusted partner.** More specifically, the IEAI partnered with significant players in AI from academia (NYU Tandon School of Engineering, The GovLab, Center for Responsible AI at NYU, Université Alioune Diop de Bambey, Université virtuelle du Sénégal), industry (Fujitsu Research Lab, Huawei, IBM, Munich Re, Volkswagen) and civil society (Institute of Business Ethics, GIZ), to name a few.

SECOND, in an attempt to raise awareness, we moved the discussion on AI Ethics **from abstract theories to practice.** In our most notable example this year, the IEAI, together with our US-based partners, designed the AI Ethics: Global Perspectives, a free online course for a global audience. Since its launch in February 2021, we have released twenty modules, welcomed 30 instructors from 15 countries, and organized 6 faculty panels. And more is to come in 2022. Through this course, we are bringing practical discussion of AI ethics to practitioners and students around the world.

THIRD, we strengthened our role as a platform for **bringing together diverse perspectives** from the field of ethical AI. Having discussions with experts from all over the world is the only way to better understand complex issues and gain a holistic view of the limitations, ethical risks and benefits of using AI. Our major international conference in 2021, The Responsible AI Forum (TRAIF), had exactly this goal. Postponed twice during the pandemic, it finally happened virtually, thanks to many of you.

TRAIF took place from December 6th to December 8th and was the highlight of 2021. We welcomed top experts from industry, civil society, government and academia, joining us from many different locations and time zones. We exchanged views on the most relevant and pressing issues related to the responsible use of AI through shared stories, cutting-edge research and practical applications. We addressed ethical issues associated with the development, use and impact of AI-enabled technologies. And, of course, we stressed the importance of AI ethics and the need for responsible use of AI.

How do we develop the technology? How do we use technology? How do we regulate technology? The answers to these questions are not purely technical or technological, but rather, in a broad sense, ethical or governmental. This is what makes AI Ethics so important.

The only way to deal with skepticism and unlock AI's full potential is by devoting our energy to building trust in the development and use of the AI technology, creating practical guidelines for ethical AI, meeting societal acceptance and building trustworthy, explainable and responsible AI systems.

Our Institute will intensify its efforts towards this direction. Our goal for 2022 is to keep providing pertinent feedback both at the European and international level regarding AI Ethics-related issues, sharing important research outcomes of our interdisciplinary projects, and uncovering new topics in responsible AI through the new projects supported through the IEAI 2022 Call for Proposals. Last but not least, we will also focus on the topic trust, as trust is a fundamental concept and a key element of AI Ethics.

We welcome the opportunity to meet our growing community again in person in 2022, and to continue to connect with our partners and leading experts in the field of AI from around the world to consider the ethical ramifications of AI and work together to push the AI Ethics agenda forward.

AI Ethics can be the compass that will help us navigate these uncharted waters of the new era that lies ahead, tackle challenges and shape our future. Enjoy the read! ●

Yours,

Christoph Lütge
Director

Executive Director's Message

Welcome everyone to our 2021 Annual Report!

I want to first echo the message of our Director, Prof. Christoph Lütge, on the amazing accomplishments of the IEAI in 2021, as well as the increasing importance and growing recognition of the work we do on AI ethics. And, I want to underscore the fact that this work requires an enormous effort of our staff, student assistants and researchers.

I have the privilege of working daily with an amazing core team at the IEAI. As a small, but dedicated group, their ability to consistently produce high-quality content for our research reports, communications and events, and manage the increasingly complex apparatus that is the IEAI, astounds me.

I want to thank each of them for their extra efforts and extreme willingness to take on new projects and tasks. This is what has really made the IEAI the success story it is. Our diverse team is also one of the reasons we have been able to facilitate the incredible research productivity we have seen over the last year at the IEAI. With 17 research projects, in 7 research clusters, either currently running or recently completed since our founding, the depth and range of topics that we have been able to examine is inspiring.

Our **'Research Impact'** section highlights the projects' major findings, as well as the impact of the IEAI's essential interdisciplinary approach and focus on practical applications. However, I urge you to check out the publications mentioned in this report in order to dig deeper into the high-quality and novel work being conducted by all of our affiliated professors and research associates. The importance of independent academic research is exemplified by their work and I am convinced it will be vital to making sure AI can have **a positive and sustainable impact** on our world.

In 2022, we will **take our work to the next level.** As our projects produce more and more findings, we want to make sure those findings are articulated to the relevant stakeholders and inform practice in the public and private sectors going forward. Through workshops and events, publications and stakeholder engagement, **our goal is to make the IEAI a key source for high-quality, rigorous and trustworthy research** on the responsible use of AI.

I hope you continue to join us on this journey! ●

Yours,

Caitlin Corrigan,
Executive Director

2021 at the IEAI in Numbers

IEAI's priority is the generation of global and interdisciplinary guidelines for the ethical development and implementation of AI throughout society. ●

11 faculty members

4 international networks

women make up to **80%** of our core staff

9 members of staff

16 researchers

6 research briefs

15 research projects in **7** research clusters

12 IEAI hosted events

1,178 event attendees

7,000+ social media followers

12 founders and partners

14 published articles

18 exclusive expert interviews

Testimonials

Our personnel are as
important as the IEAI itself.
here is ...
... what they have to say about what moivates them ▼ *(this page)*
... and what their goals are for the IEAI ▶ *(right page)*

Interdisciplinary

Passion

Innovative

Open-mindedness

Collaborative!

Agile

Multifaceted Friendly

Challenging Lively

Ellen Hohma
There is always
an open door
for new ideas.

Auxane Boch
Our field of research is quite new and relevant,
which makes every new idea and project exciting,
and an opportunity to participate
in building a better future for all.

Manuela Fuchs
What keeps me motivated is a great team
that consistently tries to challenge the way
things are done which always leads
to productive discussions.

Anastasia Aritzi
Its mission:
the integration of ethical and societal priorities
into the development of AI.
Because AI Ethics matter!

Immanuel Klein
It motivates me that I can come up
with my own solutions to given problems
and be creative in this regard.

Ana Catarina Fontes
I would like to start a new project
representing not only the interdisciplinary culture
of the IEAI but also an aggregative nature
by inviting public authorities,
private industry and civil society
to discuss AI ethics.

Caitlin Corrigan
Significantly deepen
our current partnerships
to produce meaningful
research results.

Laura Lucaj
I would love to see many collaborations
with other academic institutions
around the world to join forces
on this relevant field.

Christina Daman
Gaining even more outreach
to the public and thus spreading
civic and academic education
on the crucial topic
of ethics in AI.

Franziska Poszler
To move away from ivory-tower research,
I would like to see the IEAI become a pioneer
in producing unconventional, creative and more tangible ways
of sharing research insights. ●

Research impact.

TUM has long been a driving force in researching the mutual interactions of science, technology and society. Founded in 2019, the Institute for Ethics in Artificial Intelligence (IEAI) follows TUM's strategy for "Human-Centered Engineering".

AI already has and will continue to transform society substantially within the course of our lifetimes. How will this societal transformation take place, and what will the repercussions be?

"What should be **possible** *in AI?"*

"How can we ensure that as many people as possible **benefit** *from AI's rewards?"*

We continued running 12 projects from 2020 and added three new projects with partners at two universities in Senegal, the Institute for Business Ethics in the UK and with support from Fujitsu Research Lab. Several of our projects wrapped up this year and new projects will be starting up in 2022.

Project 01

Towards an Accountability Framework for AI Systems

Aim

To develop a general accountability framework targeting organizations, building on the explanation of AI risks, their acceptable management by different actors, and thus the according responsibility distribution in a European-centric setting.

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

2021 Findings

▷ The creation of a multi-actor rationale for responsibility ●

This project is generously supported by **FUJITSU** and the collaboration with their researchers has an important and informative part of the project's success.

Developing a practical and unified accountability framework taking various stakeholder interests into account

Framework
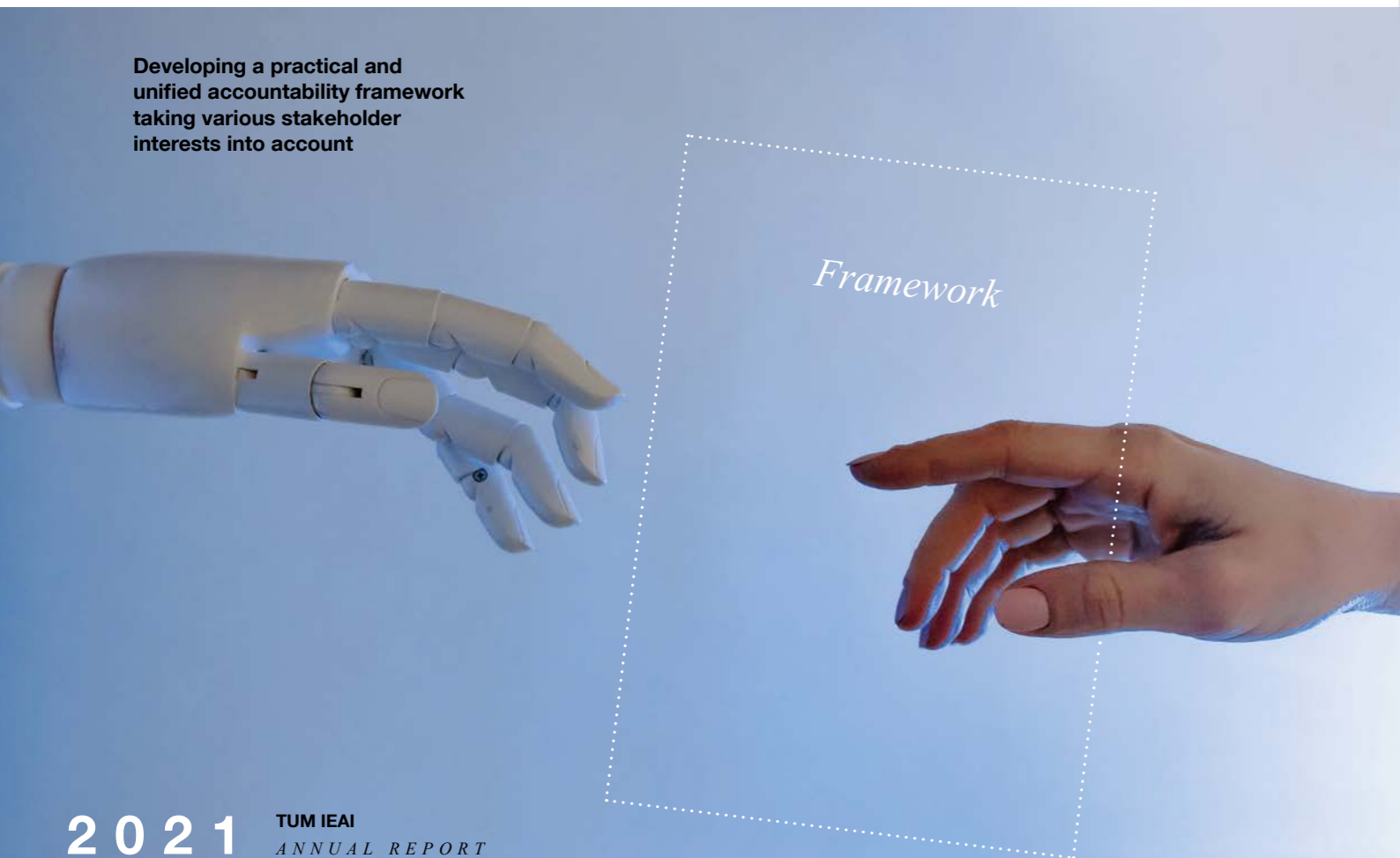
Project 02

ANDRE – AutoNomous DRiving Ethics

Aim

To demonstrate one potential solution of how ethical behavior can be integrated into the trajectory and behavior planning of automated vehicles using the ethics of risk.

2021 Findings

▷ The ethics of unavoidable accidents of AVs can be developed to the problem of a fair risk distribution

▷ Mathematical formulation of risk distribution strategies is the key to transfer ethical theories into real software

▷ A fair risk distribution for autonomous driving contains a combination of multiple shared principles, for example: minimization of the overall risk, priority for the worst-off, equal treatment of people, responsibility and a maximum acceptable risk ●

◁ Option A   Option B ▷

Franziska Poszler ▶
and **Maximilian Geisslinger** ◀

in the documentary *"AutonomesFahren: Wie selbstfahrende Autos bei einem Unfall entscheiden"* at **Bayerischer Rundfunk**

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

Project 03

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

Aim

To identify different categories of successful and ethical AI Localism (a new and radical shift in AI governance from the national to the local level) approaches while highlighting risks and challenges, including problematic AI applications and public concerns.

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

2021 Findings

▷ Development of the AI Localism Repository, which documents local AI measures around the world and how they are governed. From this, innovations in AI governance are described in the AI Localism Canvas, a framework for assessing the emergence of AI governance in cities.

▷ Preliminary results suggest that there are differences in public opinion about the ethical aspects of different local AI interventions and that innovative governance could influence public acceptance, reservations, and fears about the use of AI technologies in cities. ●

The AI Localism Repository
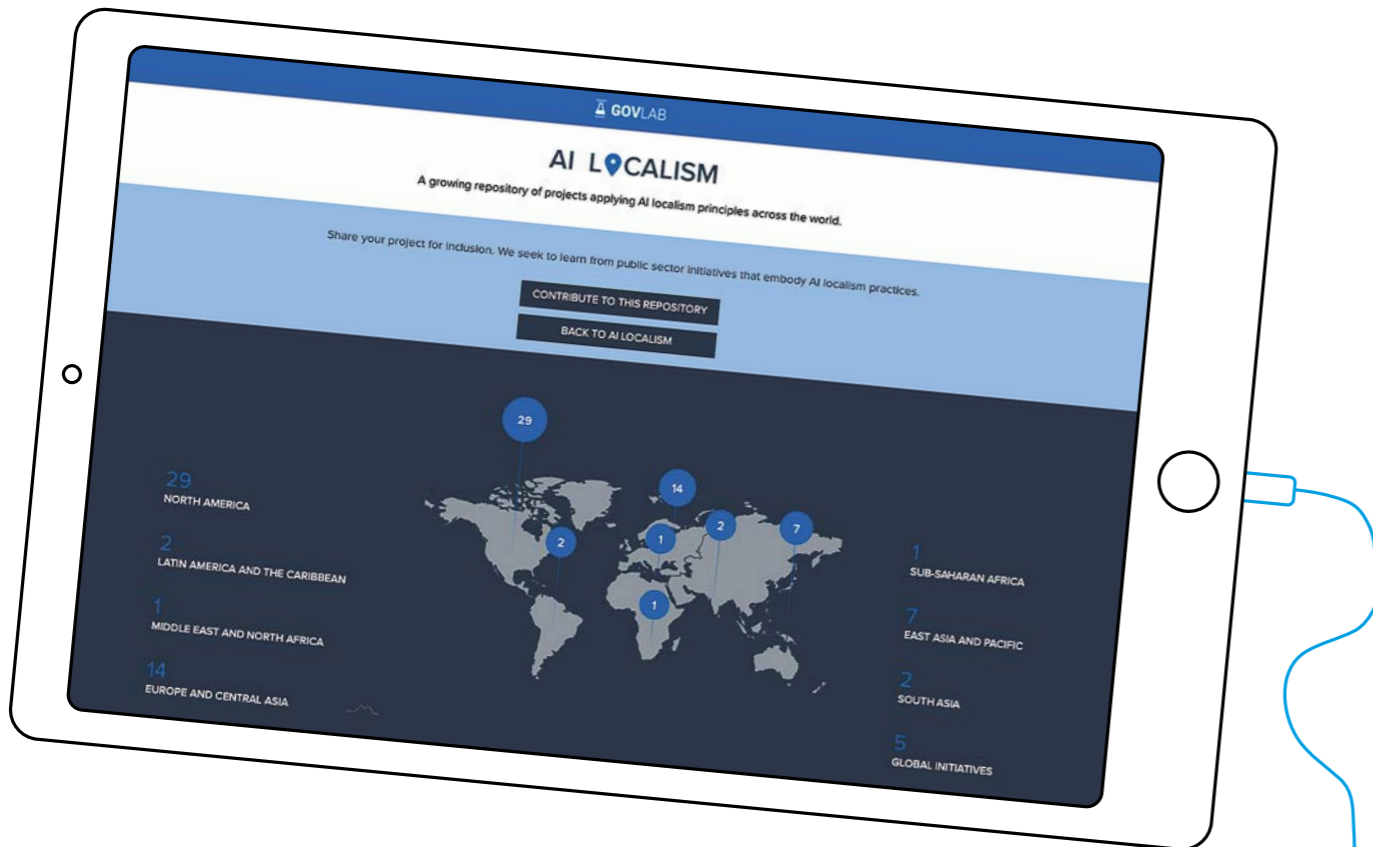
Project 04

A Framework for Digital Transformation of Senegalese Universities: Lessons Learned from the Covid-19 Pandemic

Aim

To assess the social, economic and technical barriers engendered by the Senegalese national and regional policies against COVID-19 for effective remote teaching and learning in the context of higher education.

🔗 **ieai.sot.tum.de/research/a-framework-for-digital-transformation-of-senegalese-universities-lessons-learned-from-the-covid-19-pandemic**

2021 Findings

▸ The living environment (rural vs. urban) mattered in regard to accessibility to – and satisfaction with – online courses.

▸ Technical problems (internet connection, its costs, and owned devices) highly impacted the students' and teachers' satisfaction with remote teaching.

▸ Financial background and security had a higher importance than expected for students in regard to remote education.

▸ Prior knowledge and experience with online teaching was identified as a major factor of appreciation and success for teachers.

▸ From these findings, first actionable mitigation approaches to support the development of digital learning strategies were proposed to the Higher Education Ministry of Senegal. ●

Impressions from the 2-day restitution conference held in Senegal in September 2021

Project 05

Personalized AI-based Interventions Against Online Norm Violations: Behavioral Effects and Ethical Implications

Aims

▶ To understand whether interventions against online norm violations have differential effects on the attitudinal and behavioral reactions depending on personality and social identity of those involved.

▶ To understand the ethical ramifications from specific forms of AI-based personalized interventions.

▶ To explore the balance between the deployment of effective personalized interventions, the associated ethical considerations, and the need for privacy and data protection.

2021 Findings

▷ In an online study with 600 participants, perceived ingroup norms predict bystander endorsement of counterspeech; performing substantially better as an indicator for endorsement than individual morality and attitudes. This effect was robustly found for hate speech against different target groups.

▷ Even with an increased manipulation of ingroup norms, the perception of social norms regarding counterspeech was unaffected by the experimental manipulation of the counterspeaker's group membership.

▷ A theoretical process model for bystander counter-speech in online contexts was developed, providing recommendations for online platforms and civic actors on how to increase civic engagement against online hate.

▷ Using a leading gamehacking forum's historical data around user behavior and the forum's infraction system, information on how new users and long-term users understand and comply with community rules, including the effects of a site redesign and change in communication of community rules, was gathered.

▷ Using 26 volunteer moderators on Facebook Groups, problem areas were identified and a set of automated and personalizable moderation tools that would improve their moderation were recommended. 22 out of 26 interviewed volunteer moderators provided feedback on these tools. ●

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

Project Meeting

Project 06

Public Trust in AI and the Ethical Implications: A Comparative Study of Governmental Use of AI During the Covid-19 Pandemic

Aim

To investigate the use of AI-powered technology around the world to fight the Covid-19 pandemic, and how the acceptance and uptake of such new technology depends on factors such as culture, past experience, and governance style of different countries.

2021 Findings

▷ The levels of cultural tightness or looseness in countries or regions, as well as governance characteristics (such as level of democracy) is related to the acceptance of more or less intrusive AI-powered technologies. ●

🔗 **ieai.sot.tum.de/research/public-trust- in-ai-and-the-ethical-implications**

Contact tracing apps. Now omnipresent tools.
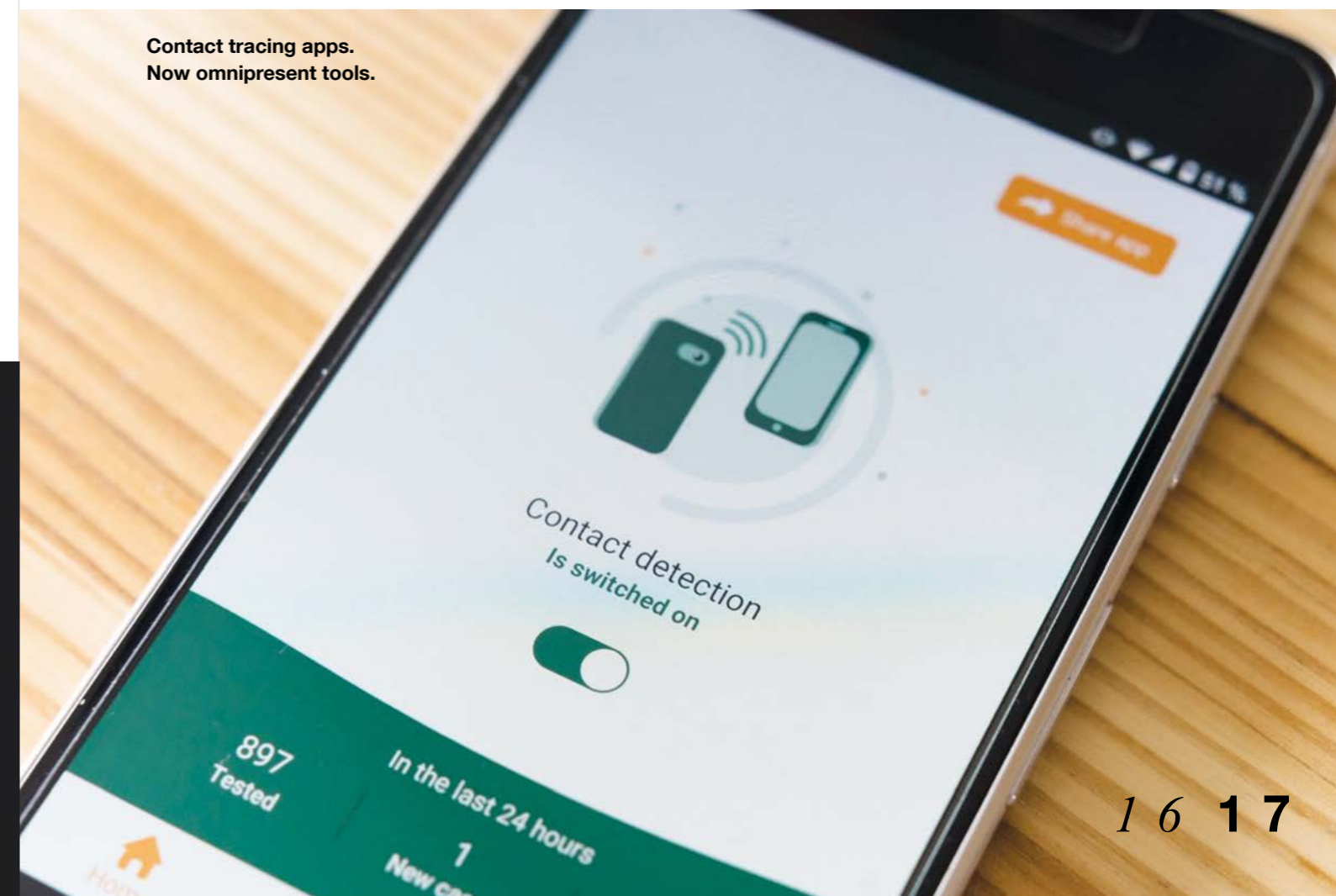
Project 07

Building Strategic Partnerships to Understand Ethics and the Use of AI to Manage Health-related Crises

Aim

To examine the potential of AI-enabled technologies to monitor, respond to and manage individuals or populations in times of crisis using two case studies focused on the use of surveillance technologies in (1) public spaces for law enforcement and (2) public health.

▶ ieai.sot.tum.de/research/building-strategic-partnerships-to-understand-ethics-and-use-of-ai-to-manage-health-related-crises

2021 Findings

▷ Transparency and proportionality are important criteria for the use of AI-enabled facial recognition systems in public spaces for law enforcement .

▷ Crime prevention and missing persons in particular appear as plausible political arguments to undertake the massive surveillance of public spaces, however informed and un-coerced consent is key for ethical use.

▷ Using the case of public health surveillance, personalization in algorithmic decision-making positively influenced user fairness perceptions and, therefore, should be balanced with the user's need for data privacy

▷ Fairness perceptions in the surveillance applications example were also found to strongly correlate with a user's overall satisfaction with the application. ●

Preparedness level on early detection and reporting epidemics of potential international concern

○ most prepared   ◑ more prepared   ● least prepared
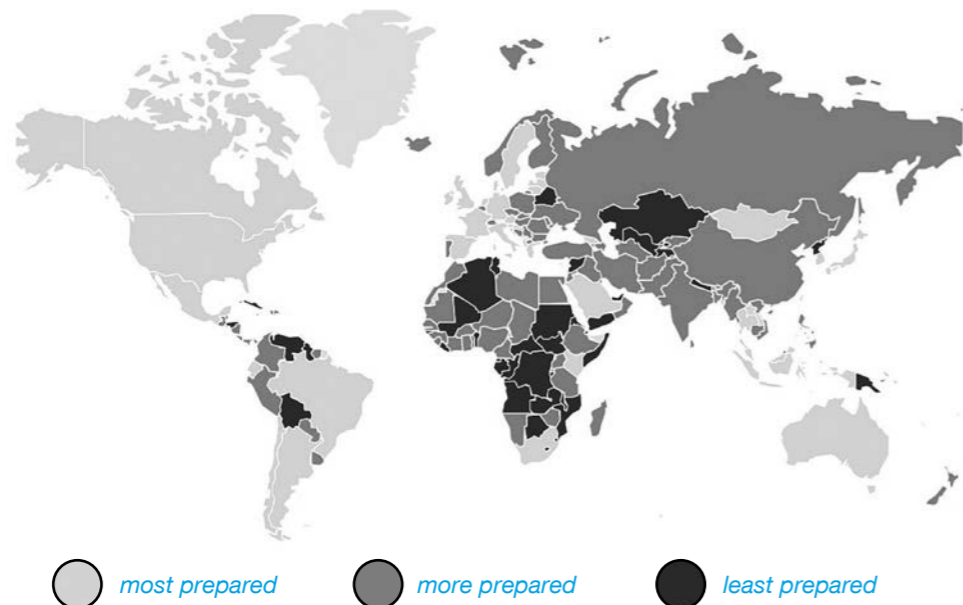
Project 08

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

Aim

To mathematically model opinion formation in large groups of interacting people on social media by investigating the driving factors which lead to negative dynamics at the social media group level in order to offer approaches on how to detect, react to, and possibly mitigate such negative dynamics at an early stage to prevent harm caused through social media.

2021 Findings

▷ Linguistic patterns and cues, such as the frequent appearance of certain pronouns or parts of speech, can be used to spot outbreaks of online firestorms in real time.

▷ Being an indispensable part of communication and therefore readily accessible in written language, linguistic cues also allow for finer analysis through investigating the sentiment in posts on social media networks.

▷ These features allow for clustering, and hence differentiating, online firestorms according to their word usage, which is helpful in distinguishing personally motivated attacks against individuals or social groups from responses to social media campaigns of companies or other types of firestorms. ●

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

Vrooom!
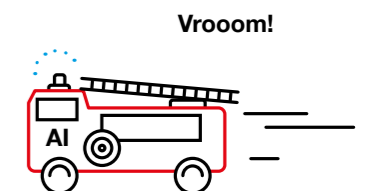
Project 09

## A Human Preference-aware Optimization System

Aim

To explore whether and how AI can be used to optimize processes in the workplace in an ethical manner, respecting fairness and privacy of employees and promoting their well-being using the example of logistics.

2021 Findings

▷ Theories of meaningful work can inform the human-centered design of workplace technology.

▷ Criteria which should be respected in an algorithmic scheduling system are employee-related, workplace and task-related, organizational and working environment-related.

▷ Algorithms that solve allocation problems can be divided into optimal, heuristic and hybrid solution methods. AI approaches in algorithmic scheduling are underexplored so far. ●

🔗 ieai.sot.tum.de/research/a-human-preference-aware-optimization-system
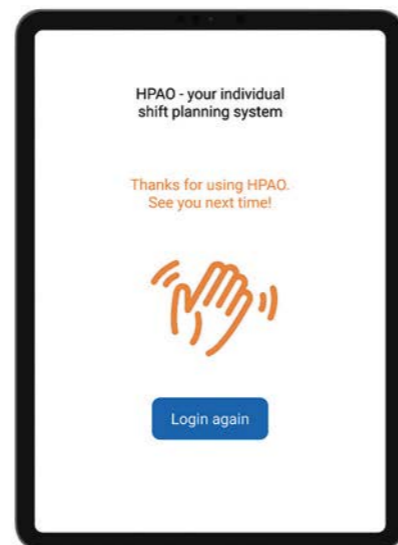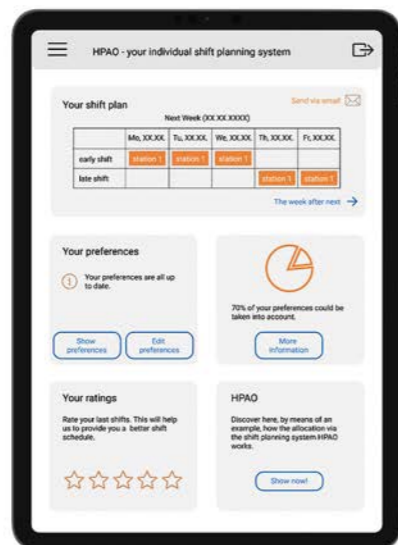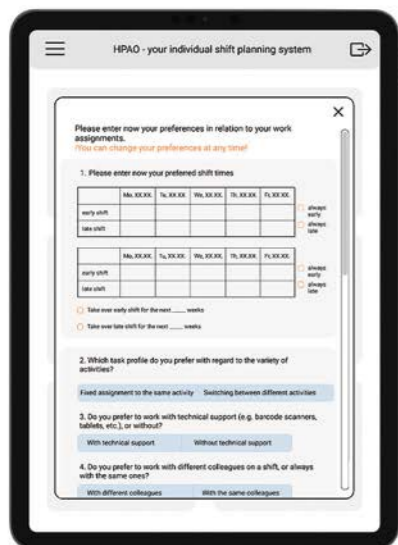
Project 10

METHAD – Toward a MEdical ETHical ADvisor System for Ethical Decisions

Aim

To develop an algorithm that can help make ethically relevant decisions in the clinic. That is, a machine that roughly does what clinical ethics committees usually do.

2021 Findings

▷ First performance results for an algorithm, based on Beauchamp and Childress' prima-facie principles, capable of advising on a range of moral dilemma situations in the medical domain showed high accuracy. From this, a pilot user interface was developed. ●

🔗 ieai.sot.tum.de/research/METHAD-toward-a-MEdical-ETHical-ADvisor-system-for-ethical-decisions
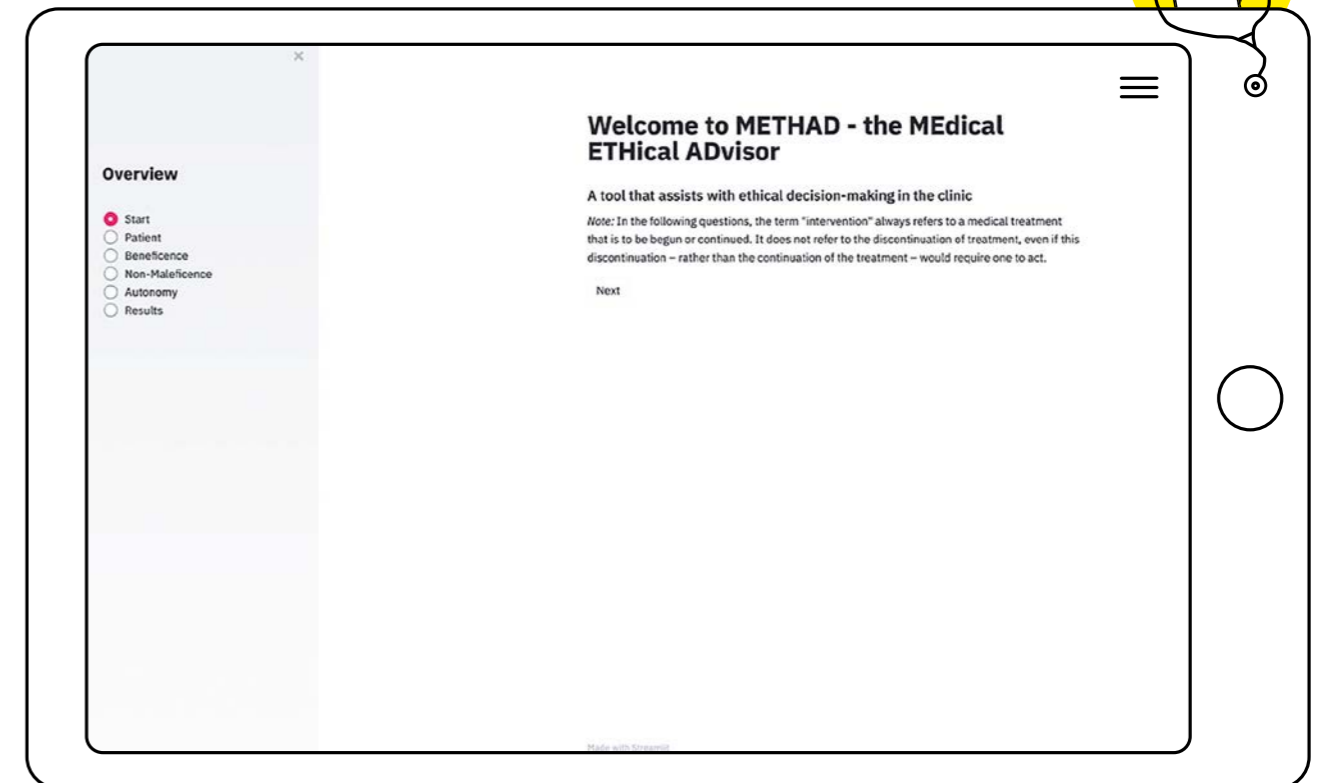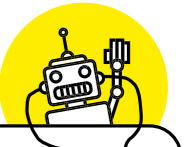
Project ⓫

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

Aim

To systematically investigate the effects of online experiences on offline attitudes and behavior, focusing on polarization tendencies triggered and promoted by feedback provided in social networks.

2021 Findings

▷ Encouraging feedback on polarizing statements significantly increases the political polarization of participants. ●

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

Change before/after experiment

0.0　　0.2　　0.4　　0.6　　0.8

Top half of participants　**+0.89**

Bottom half of participants　**+0.44**

Polarization tendencies of ...
▶ *experimental group* ▲ *(upper bar)*
▶ *and control group* ▼ *(lower bar)*

Project ⓬

Understanding AI Ethics in the Workplace

Aim

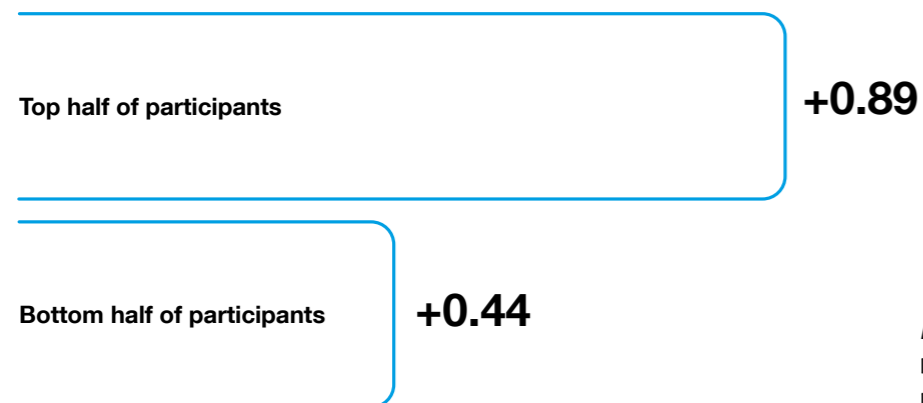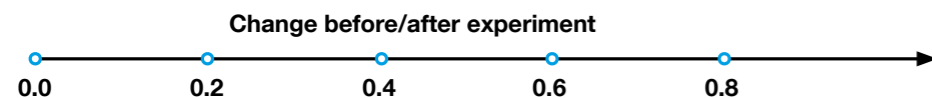**To investigate the level of implementation of AI and AI-specific ethical guidelines across organizations, as well as employees' concerns on the occurrence of ethical misconduct related to AI technologies.**

2021 Findings

▷ AI is widely used and implemented across organizations' business functions, particularly in "Information & Communication" and "Finance & Insurance". "Public administration & education" and "Health" have relatively lower AI adoption rates.

▷ AI adoption rates differentiate geographically, with almost half of employees reporting use in South Africa and the U.S., while about 15% report so in Portugal.

▷ Contrary to expectation, the highest adoption rates of AI use can be found across mid-sized firms (~33%).

▷ Slightly above 60% of individuals whose employers have adopted AI or who personally interact with AI in the workplace were aware of any AI ethics guidelines in their organization.

▷ The more familiar people are with AI, the more concerned they are about future misuses of AI in the workplace revealing the missing trust in AI. Thus, to realize the benefits of AI, managers need to foster an organizational culture based on ethical AI. ●

↪ ieai.sot.tum.de/research/understanding-ai-ethics-in-the-workplace
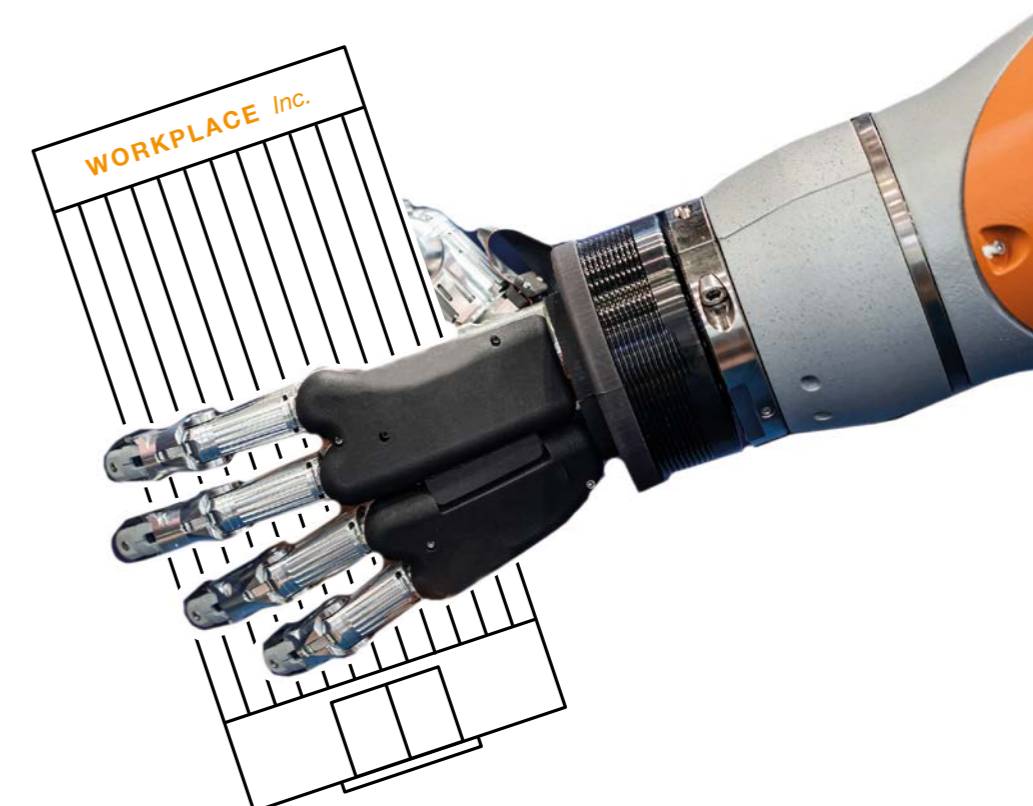
WORKPLACE *Inc.*

Project ⑬

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

Aim

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

2021 Findings

▶ AI4EO scientists are either unfamiliar with existing AI ethics guidelines or find them to be mostly impractical/not useful for their everyday research. Concrete, discipline-specific examples and case studies are needed to help scientists understand the practical importance and impact ethics has and can have on their present and future research. ●

⮕ ==**ieai.sot.tum.de/research/artificial-intelligence-for-earth-observation**==
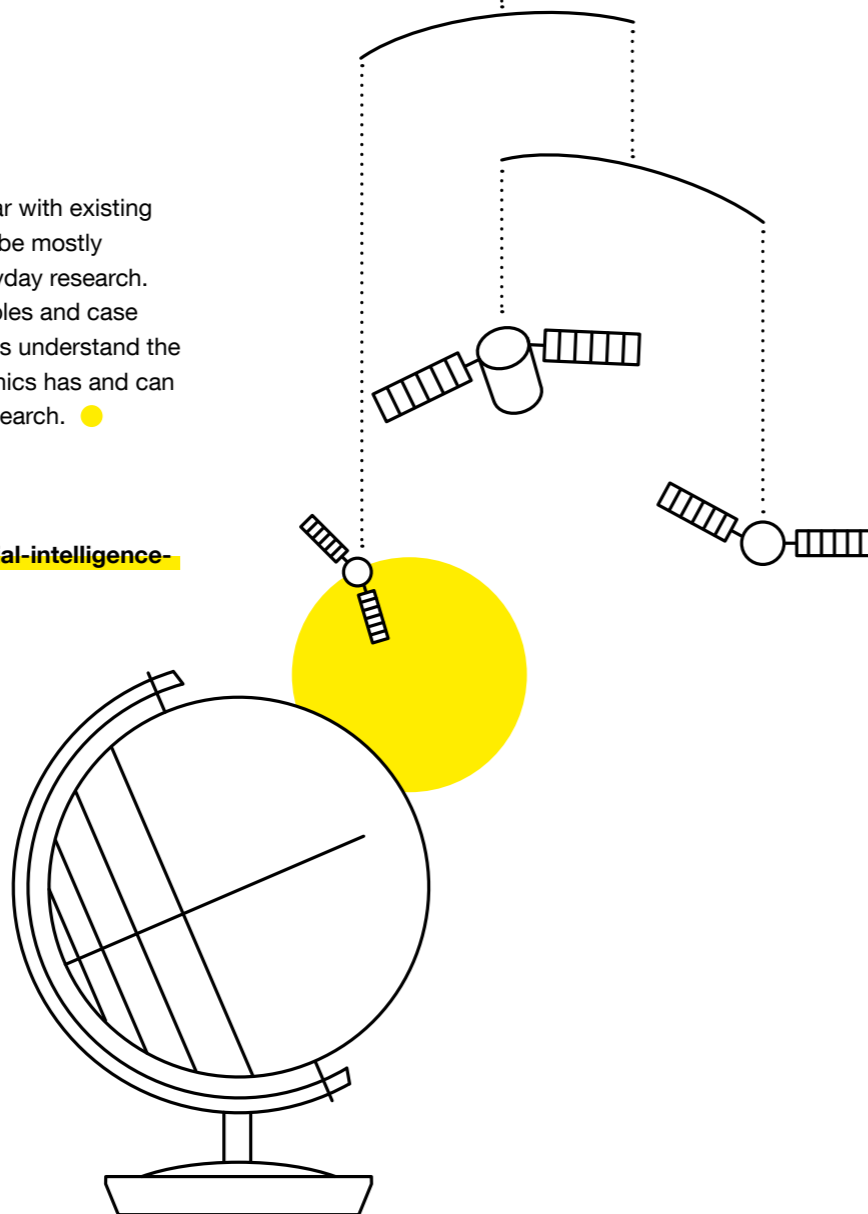
Project ⑭

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

Aim

To provide state-of-the-art guidance for ensuring that research and innovation using technological applications for controlling COVID-19 are compliant with the Rule of Law. Focusing on AI-assisted technology, the research addresses COVID-19 control strategies in the pandemic and post-pandemic phases

2021 Findings

▷ There is a critical link between trust and COVID compliance. An absence of trust will produce negative consequences for COVID control.

▷ There are 4 different types of legitimacy styles: 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 use must be tempered.

▷ The Rule of Law enhances trust relations (even in authority styles where utility is the predominant motivation) through the reduction of arbitrariness. ●
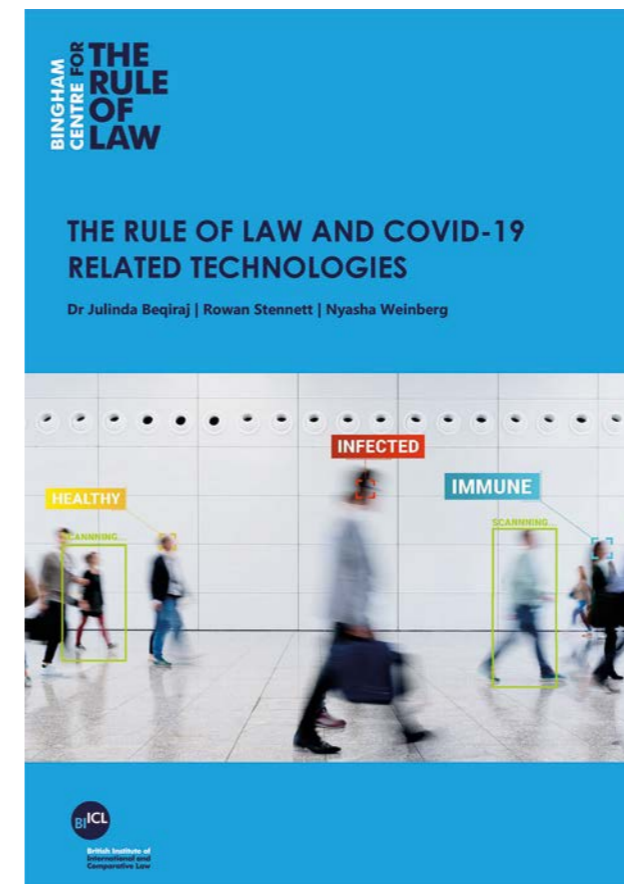
⮕ **ieai.sot.tum.de/research/rule-of-law-legitimacy-and-effective-covid-19-control-technologies**

BINGHAM CENTRE FOR THE RULE OF LAW

THE RULE OF LAW AND COVID-19 RELATED TECHNOLOGIES

Dr Julinda Beqiraj | Rowan Stennett | Nyasha Weinberg

Report from the Bingham Center for the Rule of Law based on this project's research

Project 15

ETHAN
– Ethical AI for Pandemic Management

Aims

▶ To enable decision makers to test hypothetical scenarios of proposed methods to fight pandemic spread by providing a simulation-spread framework which takes into account individual behaviors and how they can be affected and impacted by public health decisions.

▶ To develop deep-learning assisted methods to deal with pandemic developments on an agent level which minimally impacts the restrictions on daily life and which are testable and verifiable through the simulation environment.

2021 Findings

▷ High compliance for public measures is paramount to pandemic control. Measures not accepted and acted upon by the public are set to have little-to-no effect on curbing the spread of a disease,

▷ Public acceptance should be factored into use of measures to actually induce the desired effect. ●

🔗 ieai.sot.tum.de/research/ETHAN-ethical-ai-for-pandemic-management

Three constituent parts of the simulation framework

Long short-term social network

Disease spread

Mobility & encounters
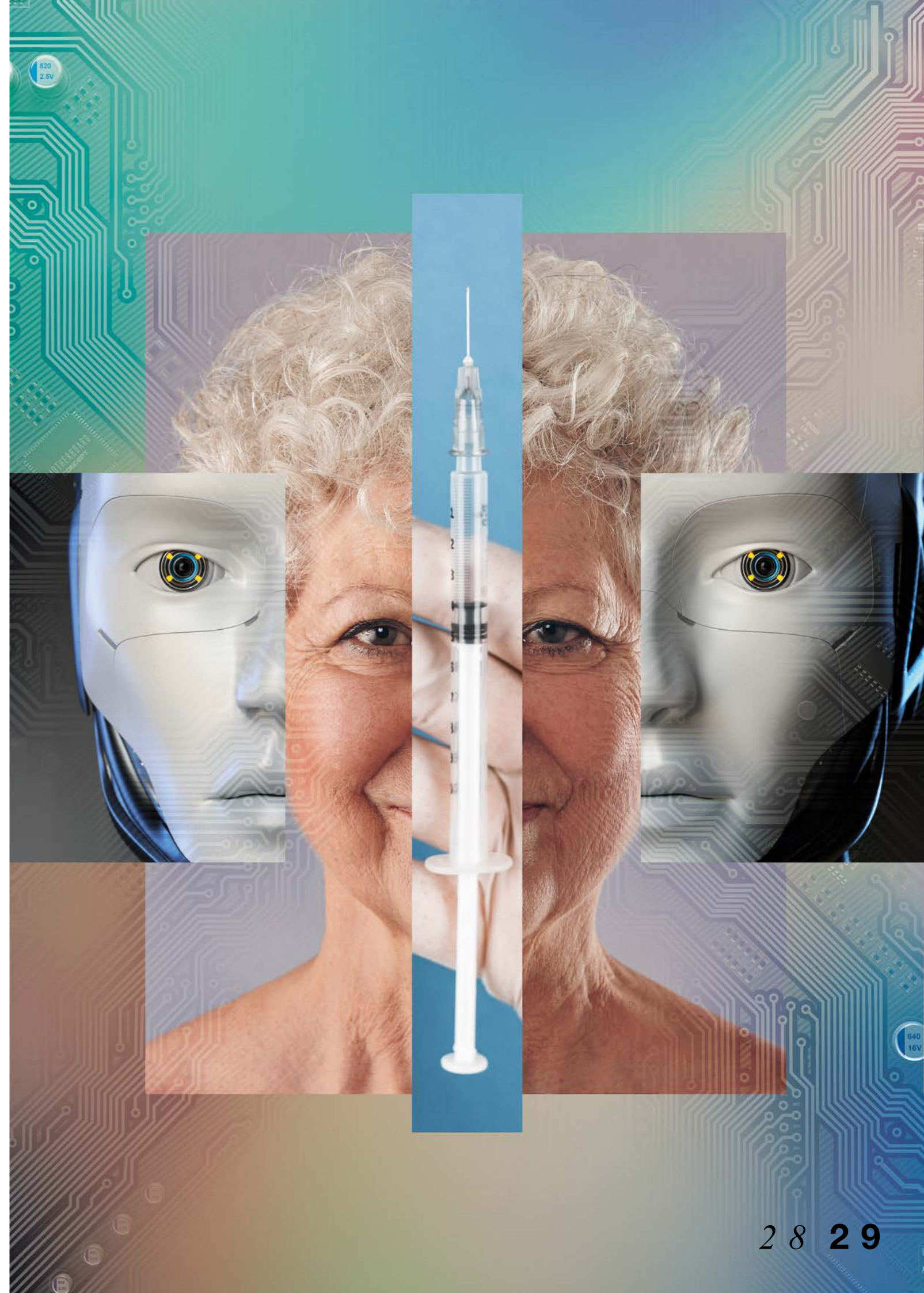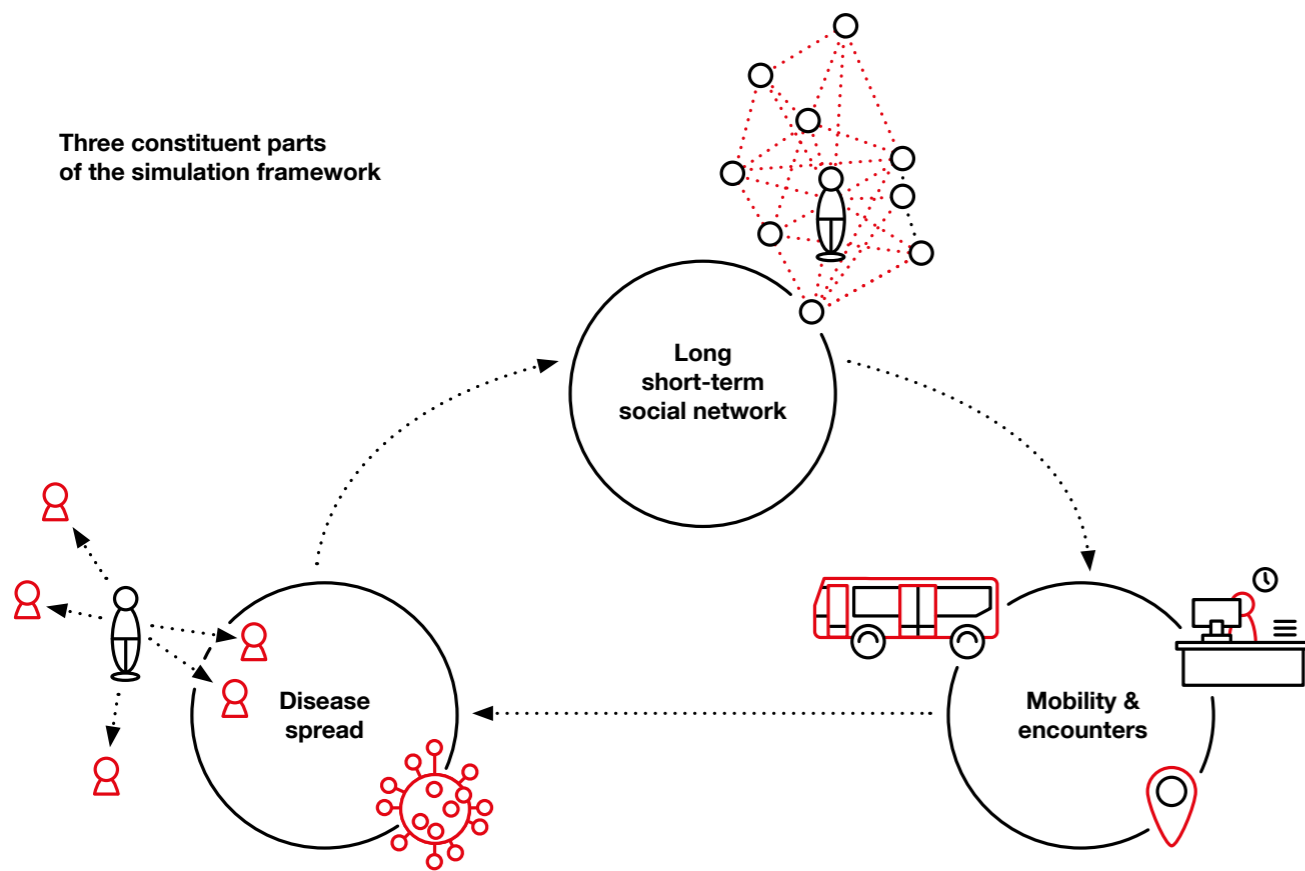
Changing the Way We Think about AI

IEAI research

Old thinking → New thinking

Our research challenges what we think we know about how AI works, its impact and how it can be implemented for good. We asked our researchers how their projects aim to change the way we currently think about AI:

▶ Our work aims to define accountability more clearly and thus create more trust and acceptance of AI systems.

P 01 *Towards an Accountability Framework for AI Systems*

▶ Our developed algorithm for autonomous vehicles could generate a more sophisticated and transparent ethical decision-making process.

P 02 *ANDRE – Autonomous Driving Ethics*

▶ Our work will provide a detailed insight into the emerging phenomenon of AI Localism in different countries, drawing attention to the governance of technology at the local level where policy is often implemented

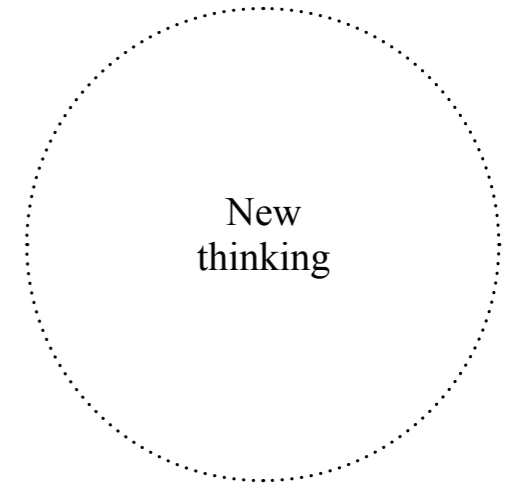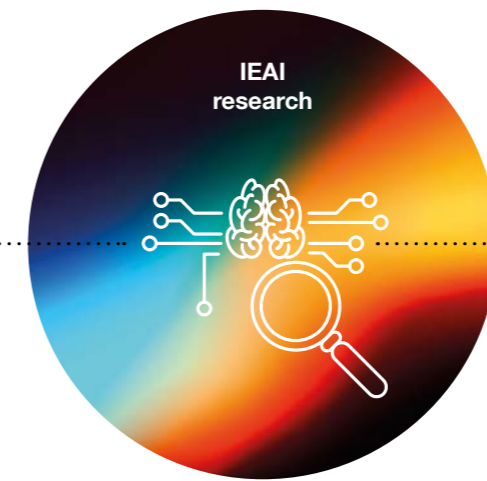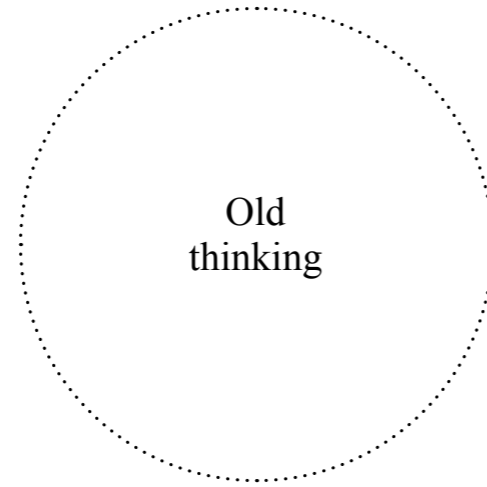**P 03** *The Ethics and Practice of AI Localism at a Time of COVID-19 and Beyond*

▶ Our work demonstrates that the context where tools are implemented matter and barriers to use may vary by culture or socio-economic status, which should be kept in mind as AI tools are being developed.

P 04 *A Framework for Digital Transformation of Senegalese Universities*

▶ This research questions the current state of moderation technology and provides insight on how to create personalizable, context-sensitive automated technologies for the everyday moderator of online communities.

P 05 *Personalized AI-based Interventions Against Online Norm Violations*

▶ Understanding the way AI systems are accepted, and the cultural dynamics behind this process in the global, regional, and local settings can aid in the adaptation and uptake of responsible AI tools.

P 06 *Public Trust in AI and the Ethical Implications*

▶ Our project turns algorithmic optimization in the workplace on its head: we want to use AI to adapt work processes to the needs and preferences of employees, not for optimizing outcomes.

P 09 *A Human Preference-aware Optimization System*

▶ Our project raises questions about what an "ethical" decision support system should really be trained to predict and how AI can or should be used in that process

P 10 *METHAD – Toward a MEdical ETHical ADvisor System for Ethical Decisions*

▶ The research will help scientists re-think their research questions and goals to avoid ethical dilemmas, support the accomplishment of UN SDGs and make a more constructive difference to life on the planet.

P 13 *AI4EO- Artificial intelligence for Earth Observation*

▶ The project deals with the transparency and communication of decisions as made by an AI, which should further help in fostering public trust when employing such measures.

P 15 *ETHAN – Ethical AI for Pandemic Management* ●

Putting AI Ethics into Practice

The IEAI aims to truly and comprehensively address a growing group of ethical challenges arising at the interface of technology and human values. One of the biggest ways we do this is through a focus on the development of thoroughly operational ethical frameworks and applications in the field of AI.
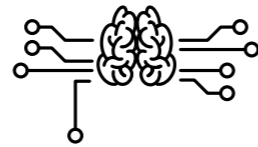
Thus,
putting ethics into practice
is high on our agenda.
Our research projects
have focused on this
in a number of ways.
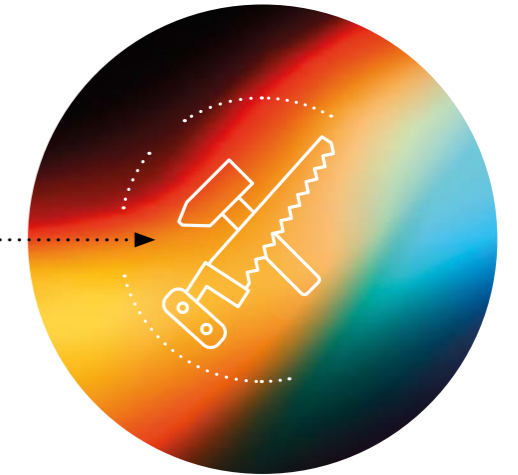
▶ Our accountability framework will focus on practical applicability to be used in the design, development and implementation of AI systems, helping to solve responsibility challenges on different levels.

P **01** *Towards an Accountability Framework for AI Systems*

▶ The hope is that policy makers will adopt our frame work to decide what constitutes a fair risk distribution in traffic. Industry players can also draw on our mathematical formulation to develop their algorithms.

P **02** *ANDRE – Autonomous Driving Ethics*

▶ By determining the characteristics of successful approaches to AI localism, our research could help to guide local leaders towards responsible AI implementation and innovative governance in the future.

P **03** *The Ethics and Practice of AI Localism at a Time of COVID-19 and Beyond*

▶ By developing actionable recommendations, we were able to propose to the Senegalese Minister for Higher Education that they restart and reinforce their existing program - "one student – one computer".

P **04** *A Framework for Digital Transformation of Senegalese Universities*

▶ Our research lays the foundation for automated personalized interventions against online hate speech that transcend common deletion-based approaches and foster a broader civic engagement online.

P **05** *Personalized AI-based Interventions Against Online Norm Violations*

▶ Our results contribute to assessing whether massive and intrusive surveillance tools can be justifiable, considering risks for privacy, discrimination and democratic values.

P **07** *Building Strategic Partnerships to Understand Ethics and the Use of AI to Manage Health-related Crises*

▶ This research can aid in the detecting of an upcoming online firestorm early on by monitoring only readily available linguistic features to understand whether it might be particularly harmful to an individual.

P **08** *Online Firestorms and Resentment Propagation on Social Media*

▶ By allocating employees to workplaces they prefer, based on our preference-aware scheduling system, we aim to increase satisfaction, autonomy and well-being at the workplace.

P **09** *A Human Preference-aware Optimization System*

▶ Our ethical adviser tool has the potential to be used for educational purposes or as a tool for providing patients and relatives with informal guidance where the involvement of a clinical ethics committee is not warranted.

P **10** *METHAD – Toward a MEdical ETHical ADvisor System for Ethical Decisions*

▶ Our study provides clear recommendations for successful AI implementation: companies need to invest not only in technological developments, but also in employees and processes to ensure a symbiosis between human and machine in an ethical way. It is a business transformation.

P **12** *Understanding AI Ethics in the Workplace*

▶ This collaboration has yielded a first of its kind tutorial on ethical issues and opportunities in AI for Earth Observation.

P **13** *AI4EO- Artificial intelligence for Earth Observation*

▶ This work highlights the need to look at the wider architecture of the tech, its proper implementation, execution, and review, including an examination into State authority, power, legitimacy, and citizens' trust.

P **14** *Rule of Law, Legitimacy and Effective COVID-19 Control Technologies*

▶ The goal of this work is to enable decision-makers to find and test optimal strategies of pandemic response both in terms of effect, but also under ethical considerations.

P **15** *ETHAN – Ethical AI for Pandemic Management* ●

The Interdisciplinary Approach

The IEAI conducts inter-, multi-, and transdisciplinary research that promotes active collaboration between the technical, engineering and social sciences, while also actively courting interaction with a wide group of international stakeholders from academia, industry, and civil society.

While this approach is vital to creating implementable and comprehensive solutions to developing responsible and ethical AI, our projects also note this process requires a strong willingness to learn for colleagues and be open-minded and ready to adapt.

▶ Team diversity comes with its own benefits but also challenges, such as finding common vocabulary, understanding the problematics from each field, and integrating all perspectives to reach a common goal.

P **01** *Towards an Accountability Framework for AI Systems*

▶ Throughout our research, we aim to connect the importance of particular ethical considerations with concrete recommendations of how these can be technically implemented. Working in this interdisciplinary fashion has been very enriching.

P **02** *ANDRE – Autonomous Driving Ethics*

▶ Our team was not only built around different disciplines, but also different cultural perspectives (Europe and West Africa), bringing in unique experiences and perspectives in regard to the research process, and also challenges in language disparities and multicultural understanding that ultimately make the project results stronger.

P **04** *A Framework for Digital Transformation of Senegalese Universities*

▶ While experts from each discipline have taken the lead on the respective research questions, regular project meetings involve in-depth discussions among the entire team for interdisciplinary reflection and critique.

P **05** *Personalized AI-based Interventions Against Online Norm Violations*

▶ The most challenging thing when working in interdisciplinary pairs is to find a common language everybody understands. In order to ensure that we do not talk past each other, regular meetings and exchanges are key.

P **08** *Online Firestorms and Resentment Propagation on Social Media*
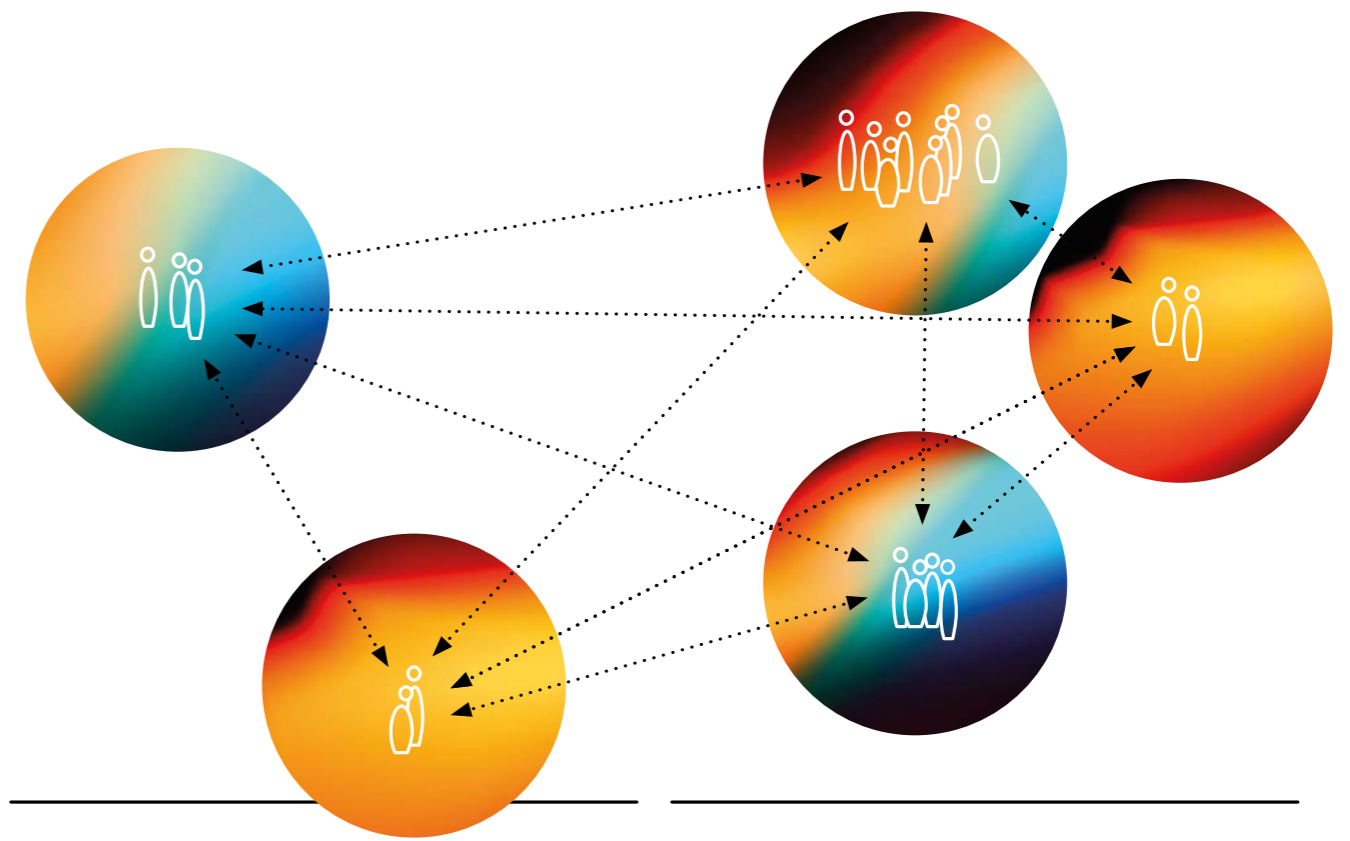
▶ It is certainly challenging to think across disciplinary boundaries, but this is also the most enriching part of the project and has expanded our horizons. Our students are an important part of this collaboration.

P **09** *A Human Preference-aware Optimization System*

▶ Working at the interface of (Eastern) philosophy, law, and Artificial Intelligence for Earth Observations, a novel approach to identifying ethical issues and opportunities at early stages of research in emerging technologies is being developed.

P **13** *AI4EO- Artificial intelligence for Earth Observation*

▶ The originality of our work lies in a combination of different methodological disciplines (legal, technological, regulatory policy) to provide an operationally and normatively richer assessment of control technology.

P **14** *Rule of Law, Legitimacy and Effective COVID-19 Control Technologies*

▶ There has been close contact during every step to make sure that the developed framework was able to be evaluated under core principles of ethics and public health law. The biggest challenge is to keep the focus on the high-level goals of the interdisciplinary team, especially when faced with technical limitations.

P **15** *ETHAN – Ethical AI for Pandemic Management* ●

Knowledge dissemination.

Effectively communicating our findings is key to ensuring that **academic research** *informs* **real-world** **practice.**

IEAI Research Briefs

Our bi-monthly Research Brief Series aims to highlight important topics in AI ethics and explore their ethical considerations. In 2021, the IEAI team published six Research Briefs shading light on important topics.

What are autonomous vehicles and how do they use AI? What makes social robots different? What are the ethical considerations of AI-based process mining in hospitals? How does the ethics behind surveillance technologies depend on the cultural and political context? How has the field of public-health surveillance evolved from a mere incident recording to complex AI-based prediction systems? How does AI pose new challenges and call for increasing proportionality and transparency to ensure respect for individual rights and liberties?

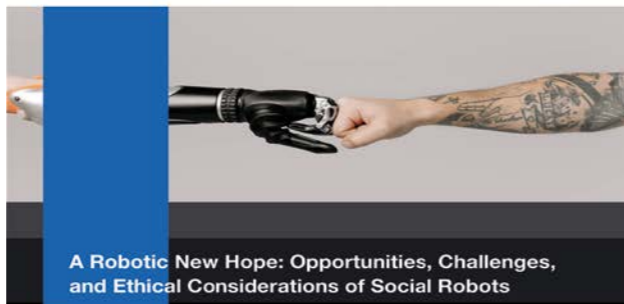**These were just some of the questions our IEAI researchers tried to answer.** ●

AI and Autonomous Driving: Key ethical considerations

Culture is "Tight" with Technology Adoption: Cultural and governance factors involved in the acceptance of AI-powered surveillance technology deployed to fight Covid-19

The Use of AI to Analyze Process-based Data in Hospitals: Opportunities, Limits and Ethical Considerations

A Robotic New Hope: Opportunities, Challenges, and Ethical Considerations of Social Robots

2021 FEB
AI and Autonomous Driving: Key Ethical Considerations
▶ Franziska Poszler
▶ Maximilian Geißlinger

2021 APR
A Robotic New Hope: Opportunities, Challenges, and Ethical Considerations of Social Robots
▶ Auxane Boch
▶ Laura Lucaj
▶ Caitlin Corrigan

2021 JUN
The Use of AI to Analyze Process-based Data in Hospitals: Opportunities, Limits and Ethical Considerations
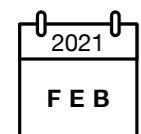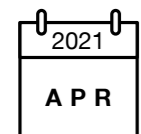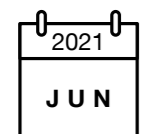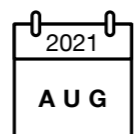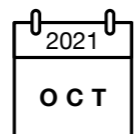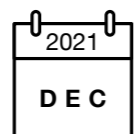▶ Thiemo Grimme
▶ Ellen Hohma

2021 AUG
Culture is "Tight" with Technology Adoption: Cultural and Governance Factors Involved in the Acceptance of AI-powered Surveillance Technology Deployed to Fight Covid-19
▶ Auxane Boch

2021 OCT
Assessing Fairness in AI-enabled Public Health Surveillance
▶ Ellen Hohma

2021 DEC
Ethics of Surveillance: Harnessing the Use of Live Facial Recognition Technologies in Public Spaces for Law Enforcement
▶ Catarina Fontes
▶ Christian Perrone

external links // only in digital version

Publication Highlights

Business and Human Rights Journal
Best Scholarly Article of 2020
"Artificial Intelligence and Human Rights: A Business Ethical Assessment"
A. Kriebietz & C. Lütge, 2020

"Advanced Statistical Analysis of large-scale Web-based Data. In: Data and Economics and Finance for Decision Makers", W. Strathern, R. Gawi, J. Pfeffer, Data Science in Economics and Finance for Decision Makers (P. Nymand-Andersen, Ed.).

"AI4People: Ethical Guidelines for the Automotive Sector – Fundamental Requirements and Practical Recommendations", C. Lütge, F. Poszler, A.J. Acosta, G.Gottehrer, L. Mihet-Popa, A.Naseer, Int. J. Technoethics.

"Are your Friends also Haters? Identification of Hater Networks on Social Media (Data Paper)", M. Wich, M. Breitinger, W. Strathern, M. Naimarevic, G. Groh, J. Pfeffer, WWW'21: Companion Proceedings of the Web Conference 2021.

"Autonomous Driving Ethics: from Trolley Problem to Ethics of Risk", M. Geisslinger, F. Poszler, J. Betz, C. Lütge, M. Lienkamp, Philos. Technol, 2021.

"Autonomous systems in ethical dilemmas", A. Bodenschatz, M. Uhl, G. Walkowitz, Computers in human behavior reports.

"Dictator game variants with probabilistic (and cost-saving) payoffs: A systematic test", G. Walkowitz, Journal of Economic Psychology.

"Ethics at Work: 2021 International Survey of Employees", Institute for Business Ethics, October 2021.

"Heidegger, Digitalisierung und Künstliche Intelligenz: wo steht der Mensch?", C. Lütge, Heidegger und die Philosophie der planetarischen Technik (Harald Seubert, Ed.).

"How To Improve Fairness Perceptions of AI in Hiring: The Crucial Role of Positioning and Sensitization", A.L. Hunkenschroer & C. Lütge, The AI Ethics Journal.

"Incentivizing research & innovation with agrobiodiversity conserved in situ: Possibilities and limitations of a blockchain-based solution", M. Kochupillai, U. Gallersdörfer, J. Köninger, R. Beck, Journal of Cleaner Production.

"Manure management and soil biodiversity: Towards more sustainable food systems in the EU", J. Köninger, E. Lugato, P. Panagos, M. Kochupillai, A. Orgiazzi, M.J.I. Briones, A gricultural Systems.

"Personaleinsatzplanung in der Logistik", C. Haid, C. Unruh, I. Pröger, J. Fottner, T. Büthe, Zeitschrift für wirtschaftlichen Fabrikbetrieb.

"The AI Localism Canvas: A Framework to Assess the Emergence of Governance of AI within Cities", S. Verhulst, A. Young, M. Sloan, Informationen zur Raumentwicklung. ●

Conference and Media Highlights
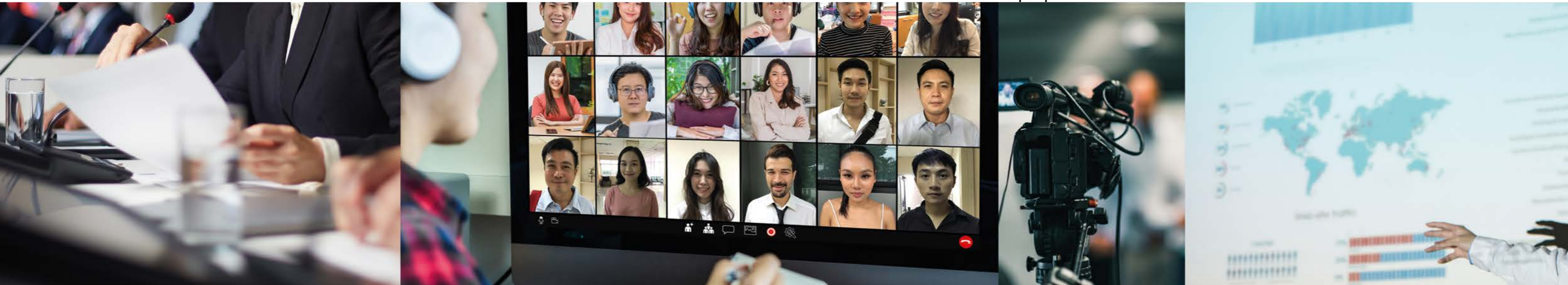
ANDRE – Autonomous Driving Ethics

▶ IAA Mobility 2021, September 2021.

▶ Stakeholder workshop on Ethics
   for CCAM with EU Commission

▶ Bayerischer Rundfunk (BR): Beta Stories (S3 F1)
   on Autonomous Driving Ethics

▶ Wer ist der bessere Fahrer?,
   Bild der Wissenschaft, June 2021

A Framework for Digital Transformation of Senegalese Universities

▶ Hybrid restitution event in Senegal, September 2021.

Building Strategic Partnerships to Understand Ethics and the Use of AI to Manage Health Related Crises

▶ Reconhecimento Facial - Perspectiva internacional,
   ITS Rio, September 2021

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

▶ Networks 2021 - A joint Sunbelt and NetSci
   Conference, June – July 2021.

▶ Special Research Seminar on COVID-19 2021,
   Munich School for Public Policy, Technical University
   Munich, January 2021

▶ International Conference on Web
   and Social Media (ICWSM), July 2021.

▶ International Conference on Computational
   Social Science (IC2S2), July 2021.

▶ Political Networks Conference (POLNET), July 2021.

▶ European Conference on Social Networks (EUSN),
   September 2021

▶ Mathematics of Deep Learning, Isaac Newton Institute
   for Mathematical Sciences, July – December 2021

▶ Mathematical and Scientific Machine Learning
   Conference, École polytechnique fédérale de Lausanne,
   August 2021.

A Human Preference-Aware Optimization System

▶ Human-Centered Algorithmic Management at the 4TU
   Ethics Conference, University of Wageningen,
   October, 2021.

METHAD – Toward a MEdical ETHical ADvisor

▶ KI2021: 44th German Conference
   on Artificial Intelligence

▶ The Virtual Ethical Innovation Lecture (The VEIL)
   Ethical Innovation Hub (EIH) Universität zu Lübeck

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

▶ Kampong Ethics and Life-Space Storytelling: Trusting
   AI. IEAI Virtual Speaker Series, November 11, 2021

▶ Rule of Law Roundtable, November 2021

Personalized AI-based Interventions Against Online Norm Violations: Behavioral Effects and Ethical Implications

▶ Society for Ambulatory Assessment Conference.
   University of Zurich, June, 2021

▶ Kolloquium des Lehrstuhls Persönlichkeits-
   und Sozialpsychologie,. University of Magdeburg,
   January, 2021

▶ Annual Meeting of the International Society
   of Political Psychology, July, 2021.

▶ 12th Media Psychology Conference, Aachen, July, 2021

▶ AAAI Spring Symposium on Implementing AI Ethics,
   February, 2021 ●
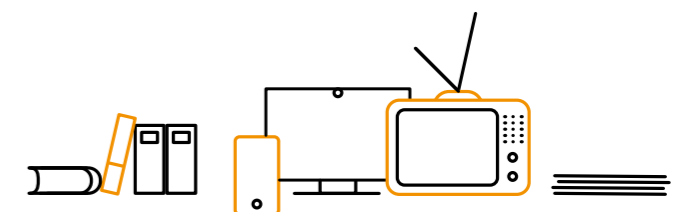
Master's and Bachelor Theses

In our efforts to bridge the gap between academia and practice, our Institute offers not only an up-to-date curated directory of multi-disciplinary AI ethics-related courses for TUM students, but also theses opportunities in collaboration with major companies and our AI ethics-related projects.

Here is what our students have to say about the IEAI and their experience in writing their Master's or Bachelor Theses.

❝ I really enjoyed writing my thesis at the Institute for Ethics in Artificial Intelligence. I had the chance to find a topic which I loved, combining both astronomy and psychology, and Auxane was of great help during the whole time. It has been an amazing experience! ❞

1 Héloïse Miny, *Houston, We Will Have a Mental Health Problem: An Ethical Analysis of AI-powered Mental Health Assistants for Astronauts during Long-duration Space Exploration*

❝ I see the discourse surrounding ethical AI crucial as the use of artificial intelligence keeps growing. My research in AI risk revealed the ethical concerns of AI manipulation, discrimination, and data privacy when AI is used in HR recruitment. As a graduate student going into the job market, the findings were a revelation to me. ❞

2 Stella Ngozika Amadi, *Risk of Artificial Intelligence in Recruitment and Selection*

❝ Writing my bachelor thesis at IEAI on the global legal status quo of autonomous cars gave me valuable perspectives on highly complex and dynamic processes in the development of new legal policies. These insights will be of great help to me in the future when dealing with new and disruptive technologies. ❞

3 Naiara Hanebeck, *The Legal Landscape for Autonomous Vehicles*

❝ As a bachelor student at TUM School of Management, I was honored to write my final thesis on the topic of The Status Quo of Autonomous Driving at the Chair of Business Ethics. I was always supported by my supervisor, while having enough freedom to come up with creative suggestions. ❞

4 Julia Huber *Autonomous Technologies in the Automotive Industry: How Automated are we driving in the present and through the future*

❝ With the coming paradigm shift in the mobility sector towards self-driving cars, questions arise: Which technology standard should be applied? How ethics can and should be implemented? And how to regulate this technology effectively? The results of my thesis showed me the still present difficulty to bring regulation and ethics in balance with the AV technology. ❞

5 Manuel Leidl *Status Quo of Autonomous Driving: Current Issues & The Legal Landscape* ●
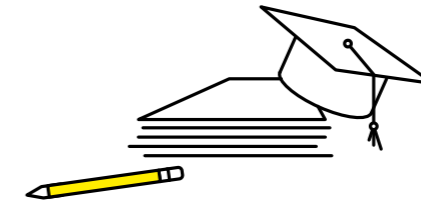
HR AI-powered
Legal landscape
Mental health in space
Autonomous driving
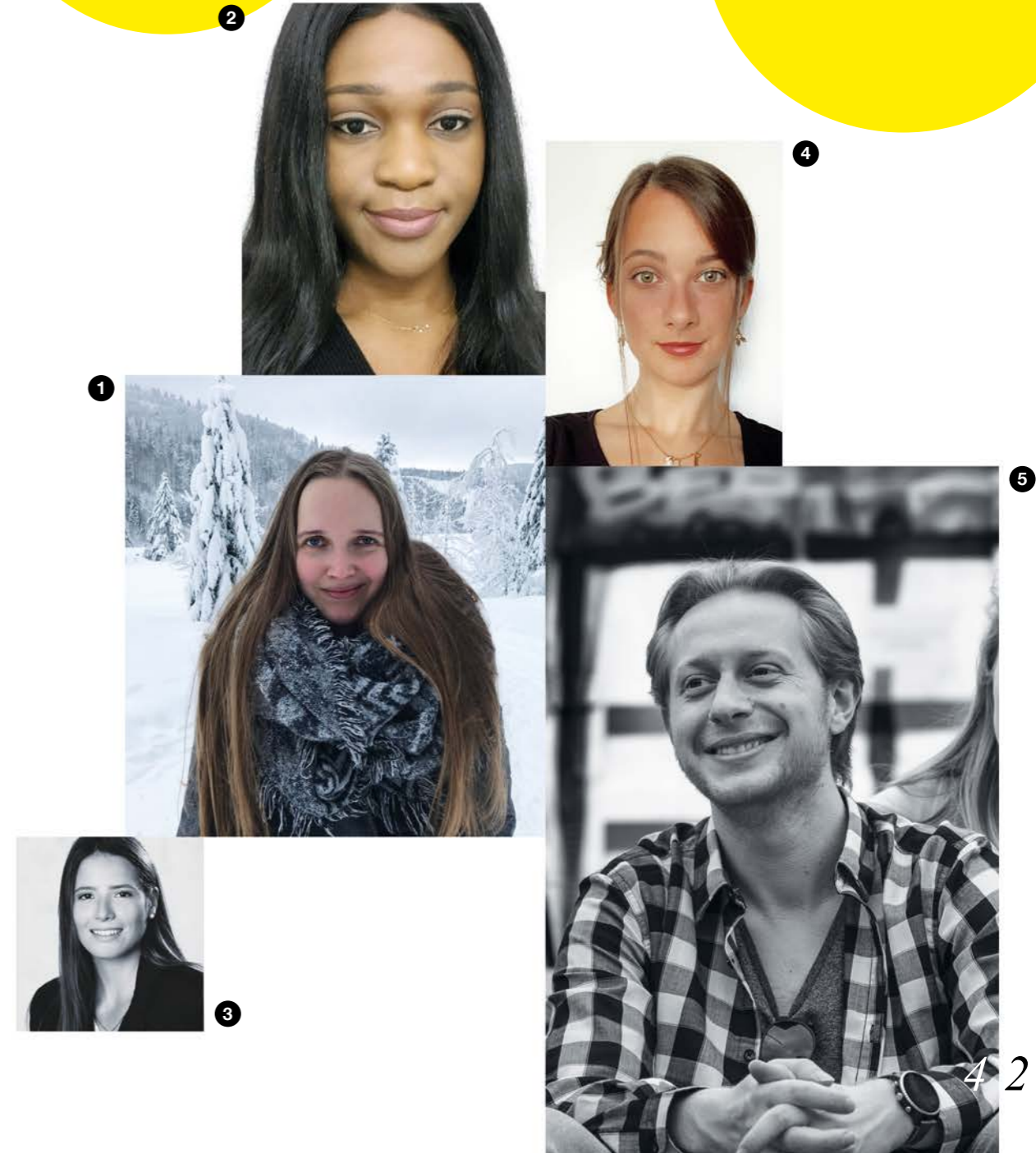
Speaking with Our Guests . . .

The IEAI Q&A Series seek to understand the importance of AI Ethics, the applications and potentials of the research being done on AI ethics, the dilemmas and misconceptions associated with AI and the role of the various stakeholders in the ethics and governance of AI, as viewed by the experts and researchers themselves.

In our "Q&A Series: Reflections on AI", we had the pleasure of speaking to eight experts from different countries and backgrounds.

▶ **Andrea Martin** ··········► [January]
▶ **Emre Kazim** ··········► [February]
▶ **Celina Bottino** ··········► [March]
▶ **Zuzanna Warso** ··········► [May]
▶ **Samira Samadi** ··········► [June]
▶ **John Tasioulas** ··········► [July]
▶ **Caitlin Corrigan** ··········► [September]
▶ **Mark Findlay** ··········► [November]

ieai.sot.tum.de/publications-and-reports/
QA-series-reflections-on-AI

What is the role of academia, research institutions, and other centers when it comes to the ethics and governance of AI **?**

What are the challenges in working in interdisciplinary teams **?**

What is the biggest misconception about AI **?**

Why is this interdisciplinary approach so important when it comes to AI Ethics **?**

Who should be in charge or involved in developing ethical frameworks and standards for AI **?**

What is the most important question in AI ethics right now **?**

How can your research be applied in practice **?**

I am deeply convinced
that we need an interdisciplinary
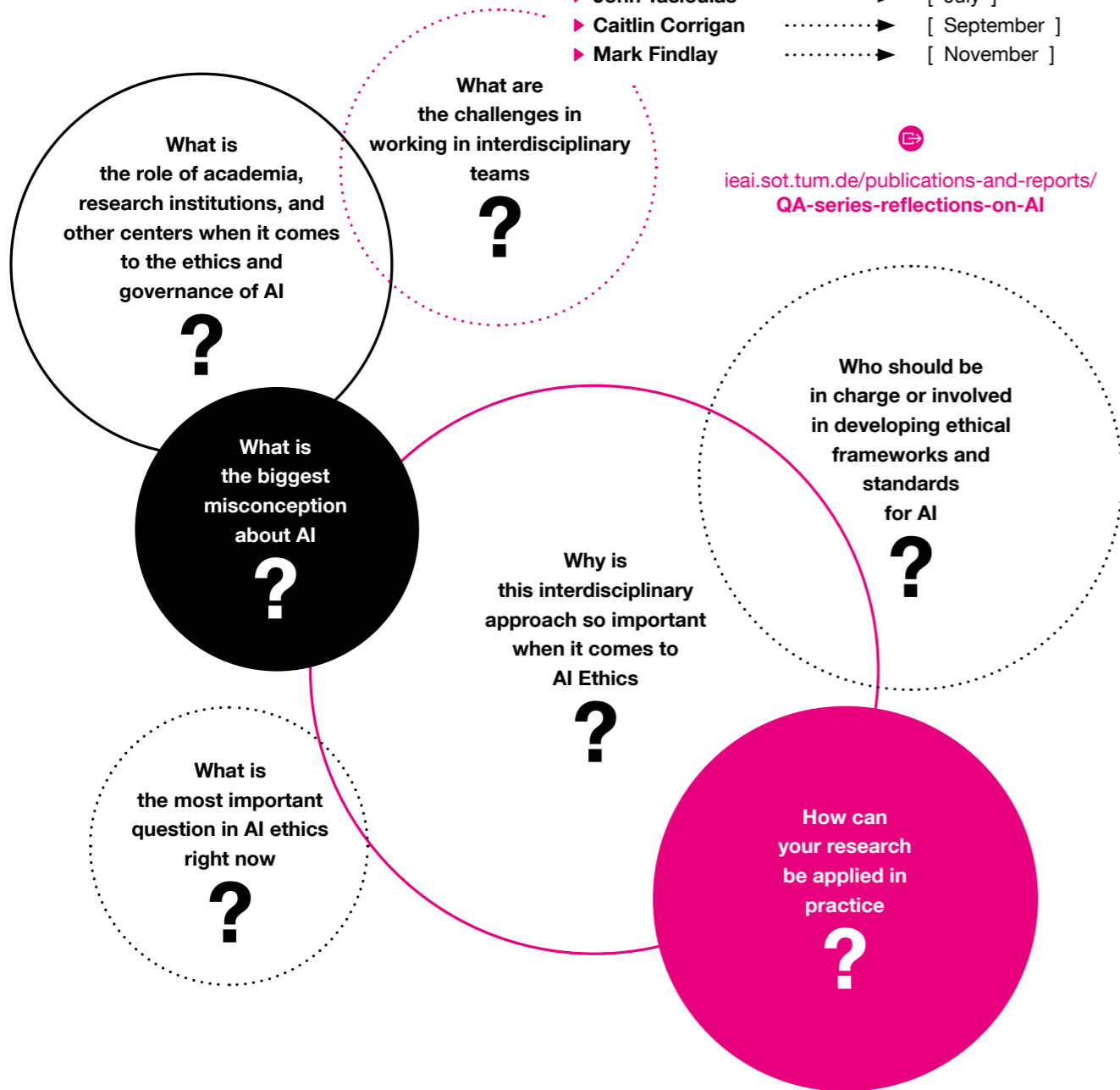dialogue and collaboration to create
trust and acceptance in AI.

Andrea Martin
Leader of the IBM Watson Center Munich & Member of
the German Parliament Commission for AI, Germany

The most important and interesting
question in AI Ethics at the moment
is the question of authority.

Emre Kazim
AI Ethicist at the University College London, UK

AI applications depend
on connectivity and there is still
a significant digital divide:
more than half of the population
in Brazil is still disconnected.

Celina Bottino
Project Director at the Institute for Technology & Society
of Rio de Janeiro (ITS Rio), Brazil

Philosophers are well placed
to highlight ethical questions,
which are subject to rational
investigation, discussion
and debate.

John Tasioulas
Professor of Ethics and Legal Philosophy
and Director of the Institute for Ethics in AI
at the University of Oxford, UK

People have the right
to participate in decisions
and processes that affect them.

Zuzanna Warso
Senior Research Analyst at Trilateral Research, UK

Recently,
there has been too much affection
given to ethics, without a serious
reflection on whether ethics alone
is the right direction.

Mark Findlay
Professor of Law at Singapore Management University
and Director of its Centre for AI and Data Governance,
Singapore

Governing AI use
is a complex problem, that requires
technical, but even more
so policy-related solutions.

Caitlin Corrigan
Executive Director at the TUM Institute for Ethics
in Artificial Intelligence, Germany

---

ML predictions
are only as good as the data
which it is provided with.

Samira Samadi
Research Leader of the "Human Aspects of Machine
Learning" group at the Max Planck Institute for Intelligent
Systems, Germany

. . . and **Researchers**

In our *"Q&A with the Researchers: Behind the scenes"*, our highly-skilled doctoral and post-doctoral researchers had the opportunity to share their views on these topics, the outcomes of their research and explain how their projects could answer specific questions at solving practical problems.
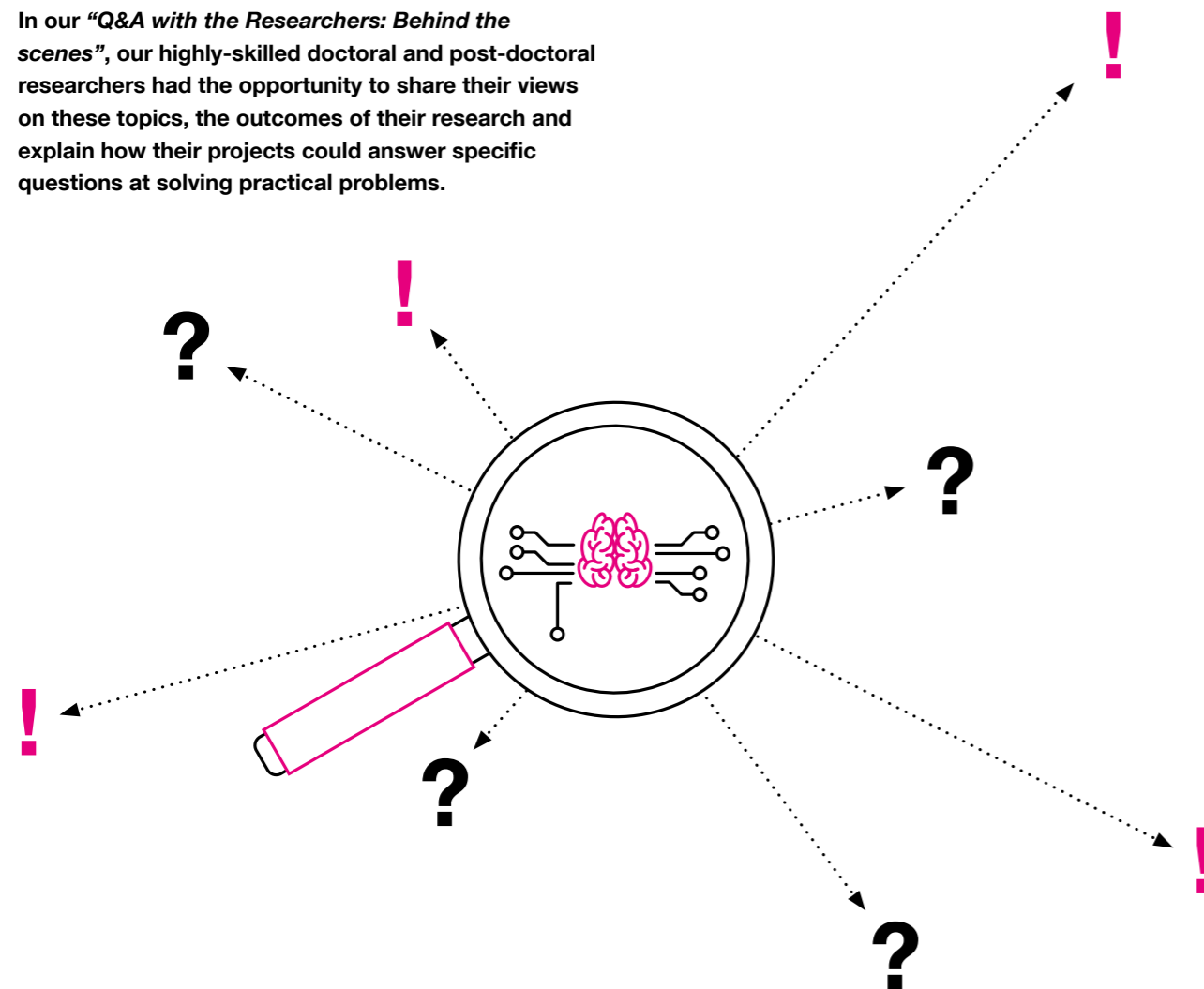
ANDRE – AutoNomous DRiving Ethics
- ▶ Franziska Poszler
- ▶ Maximilian Geißlinger

A Human Preference Aware Optimization System
- ▶ Charlotte Haid
- ▶ Charlotte Unruh

Personalized AI-Based Interventions Against Online Norm Violations: Behavioral Effects and Ethical Implications
- ▶ Tina Kuo
- ▶ Niklas Cypris

Online Firestorms and Resentment Propagation on Social Media: Dynamics, Predictability and Mitigation
- ▶ Konstantin Riedl
- ▶ Wienke Strathern

Q

A

Maxmilian keeps challenging me
with "so what" questions,
which is very helpful because this way,
we stay focused on things
that are very relevant in practice
and applicable.

Franziska Poszler

We, the ethicists, tend to think
in very high-flying abstract terms. Thinking
about what it would look like to imply
these principles on the ground in the warehouse
is something that I really benefit from.

Charlotte Unruh

When it comes to AI Ethics,
we need to consider both sides.
The technical and the social side.
To make these tools pleasant for all people.
And from my point of view, the biggest challenge
in interdisciplinary pairs is finding
a common language.

Wienke Strathern

When you work in interdisciplinary pairs,
what is different is that you have people who approach
the same problem from different perspectives.
I think that is a challenge and a benefit.

Tina Kuo ●

ieai.sot.tum.de/publications-and-reports/
QA-with-the-researchers-behind-the-scenes

Leading
the AI ethics
conversation.

Because **everyone**
*must take part
in the AI Ethics debate.*

The Responsible AI Forum

The Responsible AI Forum took place from Dec. 6 to Dec. 8 and was the highlight of 2021. We welcomed top experts from industry, civil society, government and academia, joining us from all over the world. We exchanged views on the most relevant and pressing issues related to the responsible use of AI through shared stories, cutting-edge research and practical applications.

We addressed ethical issues associated with the development, use, and impact of AI-enabled technologies. And of course, we stressed the importance of AI ethics and the need for responsible use of AI.

It was an absolute pleasure having so many people joining us virtually from different countries and time zones. Each and everyone's contribution was important. ●

Artificial Intelligence is a global challenge, so is AI Ethics.

TRAIF at a glance:
▶ **Three** days
▶ Over **150** participants joined us from **29** countries
▶ **40** session: **five** main panels and **35** parallel sessions
▶ More than **30** topics presented by around **100** scholars and practitioners

Topics included:
▶ AI for Healthcare
▶ Inclusiveness in AI
▶ Governance of AI
▶ AI for Mobility
▶ AI Ethics in practice
▶ Explainability and Transparency in AI
▶ AI and Sustainable Development
▶ Global Perspectives on AI
▶ New Research on AI Ethics

ieai.sot.tum.de/
The-Responsible-AI-Forum

BECOME AN ACTIVE PART OF TRAIF2021

THE RESPONSIBLE AI FORUM

What we are discussing here today is incredibly relevant on the political agenda, which was not the case when I started this research in 2009.

Jürgen Pfeffer
Associate Professor of Computational Social Science and Big Data at the TUM School of Public Policy

It is no longer responsible not to use AI as it is no longer responsible to work on a local, national or regional basis.

I am grateful to have so many colleagues from the Global South at The Responsible AI Forum. An inherently global perspective is necessary at a time where AI seems to be threatening to deepen rather than bridge the participation gaps between the Global South and Western Countries.

Urs Gasser
Dean of the new TUM School of Social Science and Technology

Learning global perspectives is the only way to understand complex issues better and gain a holistic view of the limitations, ethical risks, and benefits of using AI.

This is why we decided to organize The Responsible AI Forum.

Christoph Lütge
Professor of Business Ethics at the Technical University of Munich and Director of the TUM Institute for Ethics in Artificial Intelligence
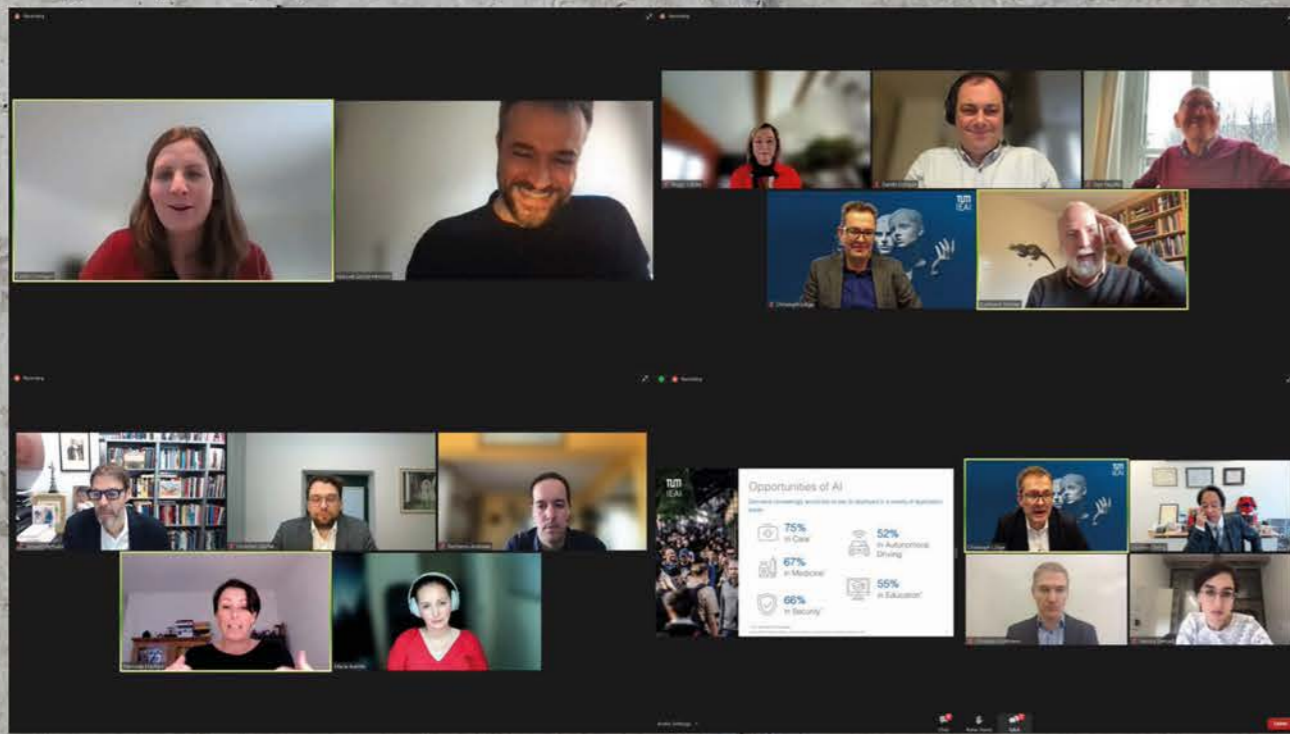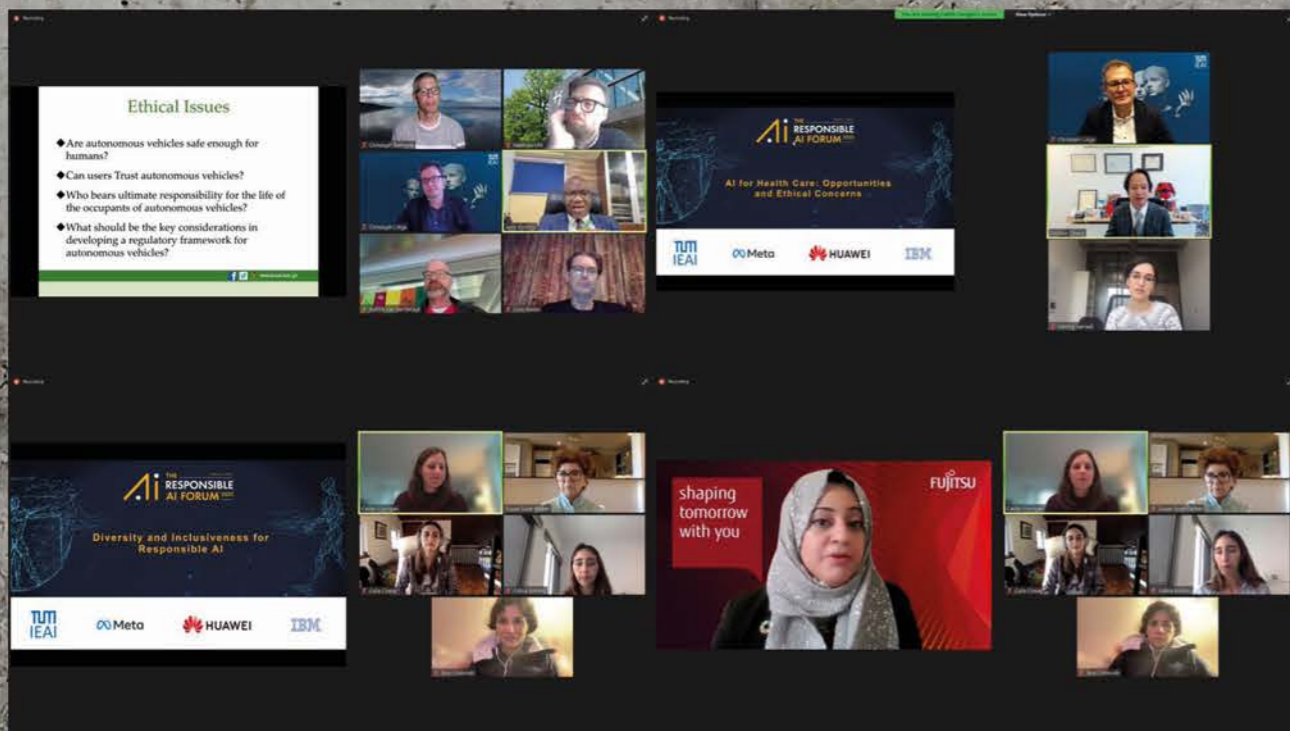
TRAIF Main Stage Panels

[Day 1]

▸ AI for Health Care: Opportunities and Ethical Concerns
▸ Diversity and Inclusiveness for Responsible AI

[Day 2]

▸ AI for Mobility: Ethics and Autonomous Vehicles
▸ Innovative AI to reach the needs of vulnerable children
▸ Governing AI: The Role of Policy and Regulation

[Day 3]

▸ Explainability and Guidelines for Responsible AI

Responsible AI should be built including perspectives from the global south. Global south countries should be part of this discussion as they can be most negatively affected by non-responsible AI systems.

Celina Bottino
ITS Rio
#TRAIF2021

How can we be more diverse?
How can we be more inclusive?
We should look at how data
is collected and used.
Aisha Naseer

People should feel comfortable
raising issues. I would say having
a mixture of #diversity
and inclusiveness is crucial.
Celia Cintas

Unless the unintended impact of unethical AI powered recruitment is urgently addressed - millions, indeed hundreds of millions already disadvantaged by assumptions about their disability will face lifetimes of needless economic and social exclusion.

Susann Scott-Parker
business disability international
#TRAIF2021

We should stop improving the state of the art machine learning when it's studied in isolation and consider the role of the systems we develop in society.

Samira Samadi
Max Planck Institute for Intelligent Systems
#TRAIF2021

Women are underrepresented in tech.
The number of women who actually
know how to write a computer program
worldwide is less than 10%.
Tara Chklovski

If you check on the policies,
there are generic references to children's
rights, but such rights are too generic
because there is a huge gap in terms of
embedding the rights of children
into AI systems deployed.
Manuel Garcia Herranz

But you know what, each organization willingly or unwillingly — with pressure from society, customers or regulators will have to go down that path of soul searching and through the test of "how moral we REALLY are". Responsible AI cannot be built by unethical business, and funny enough AI are exactly those mirrors that will uncover with phenomenal precision the true moral nature of those organizations.

Maria Axente
PwC UK
#TRAIF2021

Special thanks to our conference sponsors
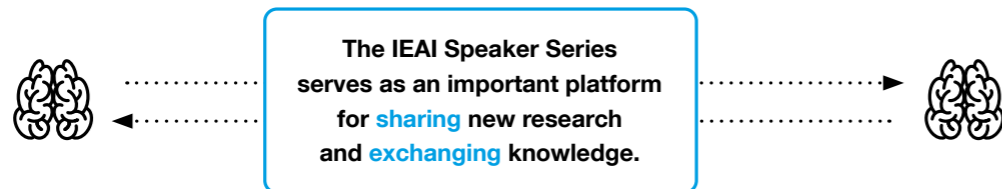**who helped make this event a success.**

Speaker Series
Talking about Ethics and Governance of AI

What does it mean for AI to be trustworthy? Why is thinking about AI governance and policy in the Global South of particular importance? Why must civil society take part in the AI Ethics debate?

These were just some of the questions we had the chance to discuss with ten distinguished guest speakers that joined us virtually from all over the world in 2021.

The IEAI Speaker Series serves as an important platform for **sharing** new research and **exchanging** knowledge.

9 virtual events
10 distinguished speakers
coming from **6** countries
and **4** continents

IEAI Spring/Summer Speaker Series
Join us & Register at: https://ieai.mcts.tum.de/events/

YI ZENG
ZUZANNA WARSO
SAMIRA SAMADI
JOHN TASIOULAS

IEAI January 2021 Speaker Series
Trustworthy AI: What does it mean and how to implement it?

ANDREA MARTIN
IBM Watson Center &
Member of the
German Parliament Commission
for AI

Join us & Register at

IEAI February 2021 Speaker Series
AI Assurance and Auditing

ADRIANO KOSHIYAMA
University College London
4 Feb. 2021
11:00am - 12:00pm CET
Zoom

EMRE KAZIM
University College London

Join us & Register at: https://ieai.mcts.tum.de/events/

IEAI April 2021 Speaker Series
Near-term and Long-term Challenges for Creating Ethical AI

YI ZENG
Professor at Institute of Automation
Chinese Academy of Sciences
29 April 2021
10am - 11am CEST
Zoom

Join us & Register at: https://ieai.mcts.tum.de/events/

IEAI July 2021 Speaker Series
The Uneasy Relationship Between Human Rights and Public Health: Lessons from Covid-19 and AI

JOHN TASIOULAS
Director, Institute for Ethics in AI
Faculty of Philosophy
University of Oxford
15 July 2021
11am - 12pm CEST
Zoom

Join us & Register at: https://ieai.mcts.tum.de/events/

[**January**]
Building more trustworthy and ethical AI systems is not only a research question, it's a business, legal and societal imperative
▶ **Andrea Martin,** Leader IBM Watson Center Munich & Member of German Parliament Commission for AI

[**February**]
The increasing importance of algorithmic auditing in the field of AI Ethics
▶ **Adriano Koshiyama,** Research Fellow at Computer Science at University College London
▶ **Emre Kazim,** Senior Research Fellow, Computer Science, University College London

[**March**]
Including the Perspectives of Global South in the Research of AI Governance
▶ **Celina Bottino,** Project Director at the Institute for Technology & Society of Rio de Janeiro (ITS Rio)

[**April**]
Near-term and Long-term Challenges for Creating Ethical AI
▶ **Yi Zeng,** Professor at Institute of Automation, Chinese Academy of Sciences

[**May**]
AI Ethics, Human Rights and the Role of Civil Society
▶ **Zuzanna Warso,** Senior Research Analyst at Trilateral Research

[**June**]
Building more Ethical ML – for Humans, with the Help of Humans
▶ **Samira Samadi,** Research leader of the "Human Aspects of Machine Learning" group at the Max Planck Institute for Intelligent Systems (MPI-IS)

[**July**]
The Uneasy Relationship Between Human Rights and Public Health: Lessons from Covid-19 and AI
▶ **John Tasioulas,** Director of the Institute for Ethics in AI at University of Oxford, Professor of Ethics and Legal Philosophy at their Faculty of Philosophy.

[**October**]
To Trust a Robot: Examining Human-Robot Trust during Emergency Evacuations
▶ **Alan R. Wagner,** Professor of Aerospace Engineering and a Research Associate with The Pennsylvania State University
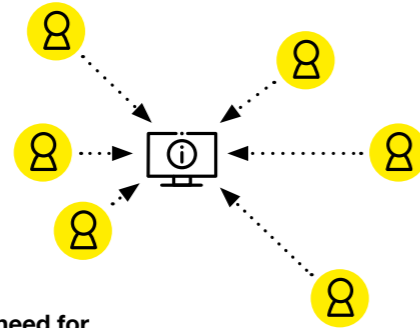
[**November**]
Kampong Ethics and Life-Space Storytelling
▶ **Mark Findlay,** Professor of Law at Singapore Management University, Director of the Centre for AI and Data Governance ●

The recordings, as well as brief summaries of our Speaker Series events, are available on the News Section of our website.
Visit **ieai.sot.tum.de/news** and find out more!

AI Ethics Course

The IEAI supports the immediate need for expertise and resources on the emerging ethical concerns, dilemmas, and questions that arise as the use of AI grows. This support also includes raising awareness and providing education regarding the ethical implications of decisions made on AI use. To this end, the IEAI worked with Global AI Ethics Consortium (GAIEC) partners at the GovLab at NYU to create an AI Ethics: Global Perspectives course. The course is a collection of short, free of charge online lessons and seminars from leading global experts in the field of ethical considerations for AI creation and use.

The AI Ethics: Global Perspectives course brings together leading experts in the field of AI worldwide to consider the ethical ramifications of AI and rectify initiatives that might be harmful to particular people and groups in society. It is designed to raise awareness and help institutions work toward more responsible use of AI. It draws heavily on our GAIEC members.

" 🙎 "
The course will allow experts to jointly develop a global understanding of AI.

[Stefaan Verhulst]

" 🙎 "
This online course is GAIEC's attempt to approach and apply AI ethics effectively in practice.

[Christoph Lütge]

AI WEEK

Celina Bottino
Institute for Technology & Society of Rio de Janeiro

Fabro Steibel
Institute for Technology & Society of Rio de Janeiro

Yves Poullet
Namur Digital Institute (NADI), University of Namur

Julia Stoyanovich
Center for Data Science and Center for Responsible AI, Tandon School of Engineering, New York University

Yi Zeng
Chinese Academy of Sciences

Christoph Lütge (Moderator)
TUM Institute for Ethics in Artificial Intelligence (IEAI)

Stefaan Verhulst (Moderator)
The Governance Lab

MAY 11, 2021
8-9 AM EST

GLOBAL PERSPECTIVES ON AI ETHICS

Photo by Nadine Shaabana on Unsplash

The course modules comprise pre-recorded lectures on either AI Applications, Data and AI, and Governance Frameworks, along with supplemental readings. New course lectures are released the first week of every month.

Since its launch in February 2021,
AI Ethics: Global Perspectives course ...

▶ Welcomed **30 instructors** from 1**5 different countries**.
▶ Released **27 modules.**
▶ Organized **6 faculty** panels – Global Perspectives on AI Ethics.
▶ Managed to have **1,000+ views** on YouTube and more than **1,500 subscribers.**

This course is a joint project of the Institute of Ethics in Artificial Intelligence, the Global AI Ethics Consortium, the Governance Lab NYU Tandon School of Engineering, and the Center for Responsible AI (NYU). ⬤

" 🙎 "
The goal of this course is to create a nuanced understanding of the role of technology in society so that we, the people, have tools to make AI work for the benefit of society.

[Julia Stoyanovich]

To learn more and sign up to receive updates as new modules are added, visit the course website at ...

🔗 **aiethicscourse.org**

Reaching out.

Working together,
instead of operating in silos,
is crucial to making real progress.

*Thus, partnerships*
are at the **heart of our philosophy.**

ARGMAX.ai
VOLKSWAGEN GROUP ML RESEARCH

FUJITSU

*) Pandemic version

Since the inception of the IEAI, one of the main goals has been to expand its partnerships with key stakeholders in the fields of AI and AI Ethics and build on each other's work to make as much progress as possible in this rapidly evolving field. In 2021 the IEAI continued to develop strategic partnerships with universities, research institutes, civil society groups, and companies to address AI ethics-related challenges. Through these partnerships, we seek to finance, foster, and conduct independent, game-changing research into wide-reaching, ethical and responsible applications for AI.

In 2021, our Institute expanded its connections with organizations like the Inter-American Development Bank and UNESCO, companies like Fujitsu, IBM, Munich Re, and the Volkswagen Group, as well as academics all around the world through the Global AI Ethics Consortium.

Seeking to design and implement researched-based concrete and practical solutions for issues related to AI, the IEAI worked on research projects with:
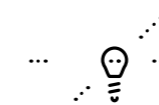
▶ **Munich Re** – **Using Explainable AI to Better Understand Credit Risk Assessment Models Based on Natural Language Processing.** Research exploring methods for explaining the predictions of NLP models, implementing them in Munich Re's proprietary NLP package and pre-evaluating the chosen methods with experts and users.

▶ **Volkswagen Group Machine Learning Research Lab (Argmax AI)** – **Bridging the gap between high-level ethical requirements and what is applicable in the field of ethical AI.** Research in the field of algorithmic auditing in European organizations for "ethical and trustworthy artificial and machine intelligence" (etami). Etami aims to be a leader in trustworthy and ethical AI by creating industry standards and piloting its certification strategies.

▶ We also welcomed the Japanese technology company, **Fujitsu,** as a partner. They are directly supporting core research at the IEAI on the project "Towards an Accountability Framework for AI Systems: The Autonomous Vehicle Use Case".

In 2022, we plan to continue working on new partnerships in the pipeline to further bridge the gap between research and practice on responsible AI. ●

Global AI Ethics Consortium

GAIEC
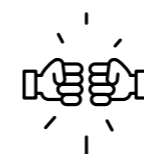
The effort to ensure that technologies are beneficial to society needs to be globally collaborative. With its broad access across different disciplines and cultures, academia can connect and build international research communities, collaborative networks, and global consortia working together on AI ethics-related issues.

That is precisely why the IEAI joined forces with academic institutions, research centers and distinguished members of academia worldwide and launched the Global AI Ethics Consortium (GAIEC). It aims to move out of the general and abstract theories, frameworks, and guidelines and provide what is really needed: instructions for applying AI ethics in a technically feasible and globally minded way.

Founded on April 15th 2020, the GAIEC celebrated its first anniversary in spring 2021. This first year was both exciting and intense. The GAIEC started as a network of experts interested in examining the responsible use of AI in the fight against COVID-19. The pandemic made clear the urgent need to discuss the ethical considerations in the use of AI and develop operational ethical frameworks in the field of AI. GAIEC's ability to react rapidly to this need was tested in practice.

A year after, we are very proud of what the group has accomplished. First, the IEAI released a call for proposals and awarded a total of €400,000 in funding to support four short-term research projects that involved GAIEC members with a focus on providing insights to developers and policymakers considering the use of AI to manage the current COVID-19 pandemic, as well as prepare for future scenarios. This was IEAI's immediate response to this pandemic.

In 2021, GAIEC presented the results of its multi-stakeholder projects and had the chance to discuss with the GAIEC members not only our research findings, but also opportunities for new areas of collaboration during The Responsible AI Forum (#TRAIF021) that took place in Munich (6-8 December 2021).

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

GAIEC's 2021 Highlights

▶ **AI Ethics Course**
The IEAI, together with its GAIEC partner, The Gov-Lab, the NYU Tandon School of Engineering and the Center for Responsible AI, designed and launched the AI Ethics: A Global Perspective course; a collection of short, free of charge online lectures and seminars from leading global experts in the field of AI.

▶ **GAIEC Repository**
The Consortium also launched the GAIEC Repository, which aims to promote AI ethics research and make this information more easily available to all. The GAIEC Repository centralizes data, publications, and other work from our members and those working on AI ethics broadly. It is a continuously expanding database and welcomes suggestions for contributions.

▶ **Joint Statement: The Importance of Public and Civil Society Participation in AI Governance and the creation of Civil Society for AI Ethics Directory**
The Global AI Ethics Consortium has devoted its 2021 agenda to highlighting the need to strengthen the role of public and civil society participation in the AI governance process. This includes pursuing collaborative research and engagement efforts that include, among other topics:

▷ **The role and implications of AI-enabled tools in the democratic process.**
▷ **Effective mechanisms for REAL civil society participation in AI public and private governance.**
▷ **Inclusive geographic approaches to reduce inequalities in discussion about AI**
▷ **Governance and access to AI enabled and ethical technologies.**
▷ **The role of the civil society in implementing AI effectively.**
▷ **How to ensure a multicultural approach of this civil society participation.**

The first step in this process was the creation of a living directory of civil society groups working on AI ethics around the world, which can be found in the GAIEC Repository. ●

The time for analyzing
how AI is deployed – whom it affects, how it affects them,
what its broader social and economic impacts are – is now.

Excerpt from the GAIEC Statement of Purpose

Since its launch,
the GAIEC has brought together **31** experts from academic
and research institutions on **6** continents while its projects
involve **7** universities and research centers, **13** principal investigators,
12 researchers and **7** applied partners from all over the world.
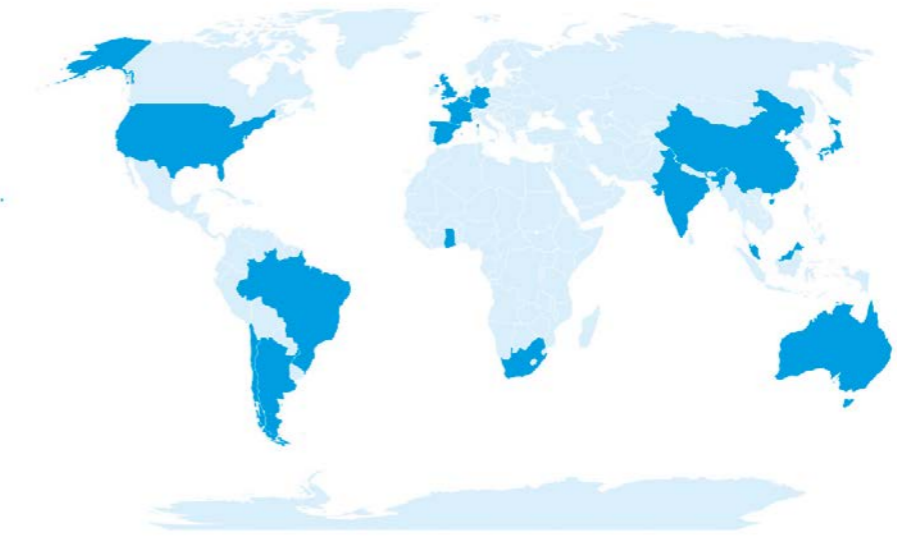
**And GAIEC continues to grow
in terms of scope and membership.**

We believe that rapidly deploying new technologies,
without sufficient consideration to societal and ethical implications
and use of a participatory process that incorporates public voices
from data subjects and their communities, risks unleashing
a number of negative developments with much broader
and longer-term effects.

**Excerpt from the GAIEC Statement on the Importance
of Public and Civil Society Participation in AI Governance**

Responsible AI Network (RAIN-Africa)

As TUM's strategic partnership initiative articulates: "No single university or country can master today's scientific challenges on its own." Working on a local, national, or even regional basis on these matters is no longer enough.

That is why the IEAI partnered with the Faculty of Electrical and Computer Engineering at Kwame Nkrumah University of Science and Technology (KNUST) in Ghana to create the Responsible AI Network (RAIN-Africa). RAIN-Africa provides a platform for cross-regional and interdisciplinary research collaboration and builds a network of scholars working on the responsible development and use of AI in Africa.

In order to develop practical and responsible AI-based applications with maximum impacts, we need to understand the cultural and socio-economic contexts where these tools are being used in order to reach more people and build relationships that can be further developed.

Laud Ammah
Co-founder of RAIN-Africa, TUM

There is a growing need to understand how Artificial Intelligence may impact or be accepted by society in various regions around the world. There is also a growing community of academics, researchers, scholars, entrepreneurs, and tech leaders from the Global South who need to participate in this discussion.

Given its potential to improve education, health, ecology and standards of living, AI development in Africa, the world's second-largest and second-most populous continent, must become one of the top priorities. However, this potential must be developed and used by and for Africans in an ethical and responsible way.
RAIN-Africa is committed to this direction.

It is important to build capacity and bring together like-minded people that can work together in pushing the agenda forward.

Prof. Jerry Kponyo
Head of the Quality Assurance and Planning Unit of the Kwame Nkrumah, University of Science and Technology (KNUST), Co-Founder of RAIN-Africa

....

RAIN-Africa's goal is to become a platform for debate, discussion and cooperation between Africa's AI stakeholders. Africa cannot and should not be left behind.

Prof. Christoph Lütge
Director of the IEAI and Co-founder of RAIN-Africa

....

We want to bring together emerging researchers to discuss and build joint projects on the ethical and social challenges arising at the interface of technology and human values, specifically as it relates to Africa's sustainable development.

Caitlin Corrigan Ph.D.
Executive Director if the IEAI and Co-founder of RAIN-Africa

RAIN-Africa needs your help to identify key educational materials, relevant publications, and use case and data sets relevant to Africa. Please send your questions, contributions, ideas or resources to:

ResponsibleAI-Africa@sot.tum.de

▶ RAIN-Africa Workshops
In 2021, RAIN-Africa organized two virtual workshops:

▶ Responsible AI and Start-Ups (17/03/2021)
▶ Responsible AI and the Development of Smart Cities/ Energy Systems in Africa (19/05/2021)

▶ RAIN-Africa Research and Data Repository
Research, data, use cases, and publications are crucial not only for the development of a robust and responsible AI ecosystem but also for the creation and promotion of an education system that supports such developments. To this end, RAIN-Africa has developed in 2021 the RAIN-Africa Research and Data Repository, which is freely available to support advancing curriculum on responsible AI, as well as development of AI applications in Africa. Through this platform, the RAIN-Africa team hopes to expand the network and promote research collaboration and knowledge sharing on the use of responsible AI.
The Repository will continuously expand as resources become available.

▶ Q&A Series: Country Representatives
The Responsible AI Network (RAIN-Africa) celebrated its first anniversary this spring (2021). In order to highlight its work, the RAIN Africa team had the pleasure of speaking to six of its country representatives posing a set of important AI ethics-related questions.

▷ **How can we harness the potential of AI in Africa?**
▷ **What are the potential benefits and risks of AI in Africa?**
▷ **What concerns and ethical challenges related to the use of AI are especially prevalent in the African context?**
▷ **What are the recent major developments in AI in your country?**
▷ **What are the AI projects you are currently working on?**
▷ **What do you hope RAIN-Africa will achieve?**

Visit the RAIN-Africa Q&A Series: Country Representatives subpage and read what the country representatives say:

▶ **Malo Sadouanouan** [Burkina Faso]
▶ **Samah Mohamed El Kahteeb** [Egypt]
▶ **Attlee M. Gamundan** [Namibia]
▶ **Lavina Ramkissoon** [South Africa]
▶ **Claudino Franklim Mendes** [Cape Verde]
▶ **Laeticia Nneka Onyejegbu** [Nigeria]

▶ Participation at The Responsible AI Forum (TRAIF)
RAIN-Africa was strongly represented at The Responsible AI Forum in Munich in December. Two parallel sessions touching upon two extremely important issues
▷ **AI in Agriculture: Challenges and Opportunities**
▷ **Setting the Stage for Responsible AI Development in Africa – The Role of Higher Education**

▶ A Framework for Digital Transformation of Senegalese Universities
The IEAI and RAIN-Africa members at the Université Alioune Diop de Bambey and the Université virtuelle du Sénégal (UVS) worked hand in hand to understand the impact of COVID-19 restrictions on Senegalese higher education.

The project with our partners from Senegal allowed the IEAI to build upon the relationships developed through the Responsible AI Network – Africa – while focusing on an increasingly important debate about the role of digital learning in higher education. We welcome future opportunities to further expand our research efforts with our partners in Senegal.

Caitlin Corrigan, IEAI Executive Director

The outcome of the project included structured recommendations to Senegalese stakeholders on what actions need to be taken to support university teachers and students in digital learning. ●

Publications and Outreach
IEAI in Social Media

Our goal was and remains to keep growing on social media and engaging more people across diverse fields, sectors, and locations in the AI Ethics discussion. This relationship with our followers is incredibly important to us.

On the one hand, we showcase the progress of our multidisciplinary projects and present our work to a global audience. On the other hand, we further promote the importance of AI Ethics, highlight the dilemmas and misconceptions associated with the use of AI, shed light on the role of the various stakeholders when it comes to the ethics and governance of AI and raise awareness regarding AI Ethics-related topics.

The IEAI Social Media [LinkedIn / Twitter / YouTube] help our team stay in touch with our growing community, give and receive valuable feedback, share news and articles and open new opportunities for collaboration with other partners.

Special thanks go out
to all who follow us, like our posts,
share our content, and write
in the comments section.

More than **7,000** followers on **LinkedIn** and **Twitter**

More than **675** subscribers to our **Newsletter**
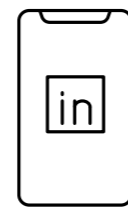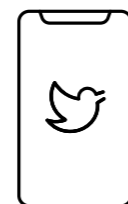
LinkedIn
[January - December 2021]

More than **300** Posts

More than **2,500** new followers

Engagement rate: **6,09%** (Mar - Dec)

More than **300,000** Impressions

Twitter @IEAITUM
[January - December 2021]

Tweets: **351**

Profile Visits: **135 K** (135,059)

New followers: **1,233**

Tweet Impressions: **720.7 K**

Mentions: **463**

Expand our outreach and strengthen
our relationships and engagement with our audiences,
followers, and broader community.
That is what we want to achieve
when it comes to the use of
social media.

Anastasia Aritzi
Communications Consultant

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

- **Ethics of surveillance: harnessing the use of live facial recognition technologies in public spaces for law enforcement – Research Brief**
- **Best Scholarly Article of 2020**
- **Reflections on AI: Q&A with Samira Samadi**
- **The Use of AI to Analyze Process-based Data in Hospitals: Opportunities, Limits and Ethical Considerations Research Brief**
- **A Robotic New Hope: Opportunities, Challenges, and Ethical Considerations of Social Robots – Research Brief**

- **AI and Autonomous Driving: Key ethical considerations – Research Brief**
- **Reflections on AI: Q&A with John Tasioulas**
- **Reflections on AI: Q&A with Caitlin Corrigan**
- **Culture is "Tight" with Technology Adoption: Cultural and governance factors involved in the acceptance of AI-powered surveillance technology deployed to fight Covid-19 – Research Brief**
- **RAIN Africa Q&A Series**
- **Reflections on AI: Q&A with Mark Findlay**

Press/Media Highlights

[Podcast]

① Navigating the ethical landscape in AI and Tech with Prof. Christoph Lütge

"For me, ethics is about proportionality. It is not necessarily about taking sides. It is about seeing things from different angles and different perspectives. I think that proportionality is a key to understanding ethics."
▶ **Professor Christoph Lütge**

[Magazine]

② Time for new experiences

"It is precisely these soft ethical issues that must be taken into account in order to inspire trust in this new technology. Our aim is to provide input right from the development phase"
▶ **Professor Christoph Lütge**

[Podcast]

③ Social Robots: Opportunities and Challenges with Auxane Boch (IEAI)

"I think social robots will be useful in behavioral environments"
▶ **Auxane Boch,** IEAI Researcher

[YouTube]

④ An insightful conversation with Auxane Boch – Associate Researcher with Institute for Ethics in Artificial Intelligence, Germany and Tanvi Mehta-Educator with Coding and MoreGaming with AI

"A game experience can actually change a behavior and make someone rethink an idea"
▶ **Auxane Boch,** IEAI Researcher

[YouTube]

⑤ How Will Human-Robot-Interaction Look Like in the Future?

"So my work as a psychologist is to say OK, we have an interaction, which means we have an action and a reaction. What happens? What is the consequence of this action-reaction? What is the consequence in the long and short run? What is the consequence on different populations like children, or people with mental disabilities?"
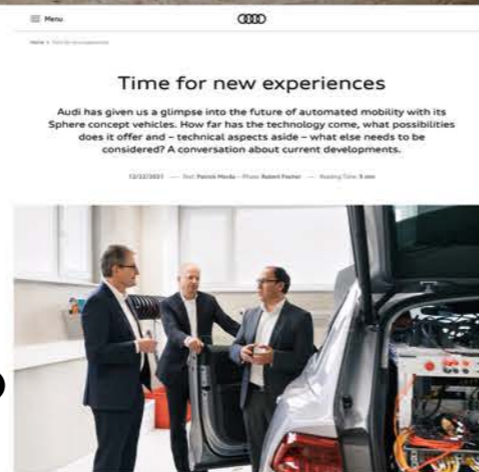▶ **Auxane Boch,** IEAI Researcher
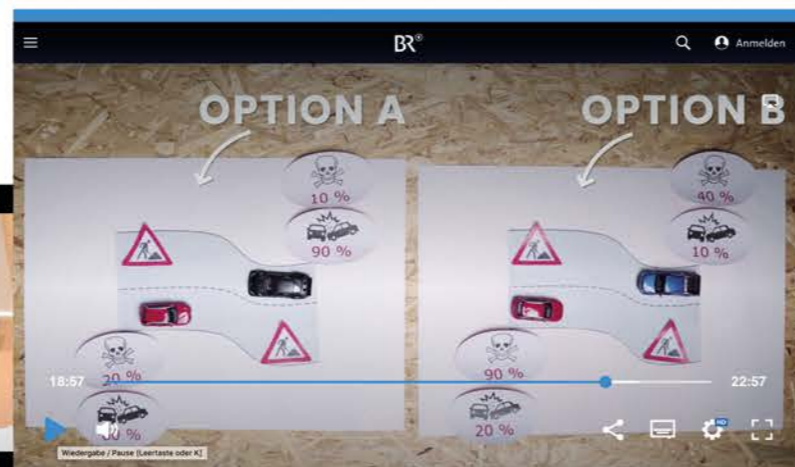
[TV]

⑥ Maximilian Geißlinger and Franziska Poszler starred in the documentary „Autonomes Fahren: Wie selbstfahrende Autos bei einem Unfall entscheiden" at Bayerischer Rundfunk

„Was ich eben gut finde ist, dass Max schon manchmal mich mit So-What-Questions challengt"
▶ **Franziska Poszler,** IEAI Researcher
„Dass Technik und Ethik so eng zusammen-arbeiten, ist nicht die Regel"
▶ **Maximilian Geißlinge,** IEAI Researcher

[Podcast]

⑦ How is AI already applied in our day-to-day lives? Joanna Bryson and Christoph Lütge share their views in episode 1 of the British-German Tech Talks podcast series

"AI Ethics did not exist as a field until very recently. There were maybe a couple of articles about this, but there was not really a big discussion. Now, this has changed. You see researchers from many different disciplines within this field and this is a very interesting development."
▶ **Professor Christoph Lütge**

[TV]

⑧ Unternehmen im Diskurs: „Da haben wir in Deutschland noch Nachholbedarf"

„Wir erleben gerade in Institut für Ethik in der Künstlichen Intelligenz, dass auch kleinere Unternehmen auf uns zukommen, wir wollen gerne etwas zum Thema ‚Ethik in der KI' machen, wir arbeiten selber in dem Bereich, aber wir erkennen, dass zentrale Probleme bei KI-Systemen nicht nur technische Probleme sind, sondern Probleme, bei denen es geht darum, in welche Richtung ich gehen soll"
▶ **Professor Christoph Lütge**

[Online]

⑨ Interview with Wienke Strathern

"This is actually what we're working on. We're trying to figure out characteristics of these firestorms to understand collective emotional actions. Why do people change their behavior and start attacking others? We are working on interaction behavior in communication networks. How do people communicate? Who's talking to whom? We try to testify if this is just happening unpredictably, or if we see somehow a systematic change based on certain patterns. We see some kind of patterns already in the networks, that is, how users communicate. People act, work, and communicate with each other and at a certain point, we see a movement in another direction, a system changes."
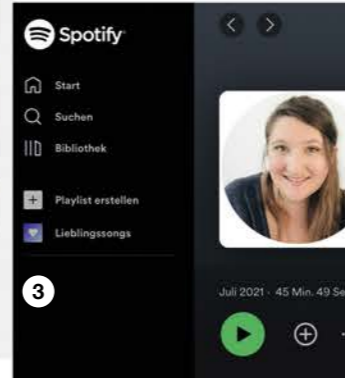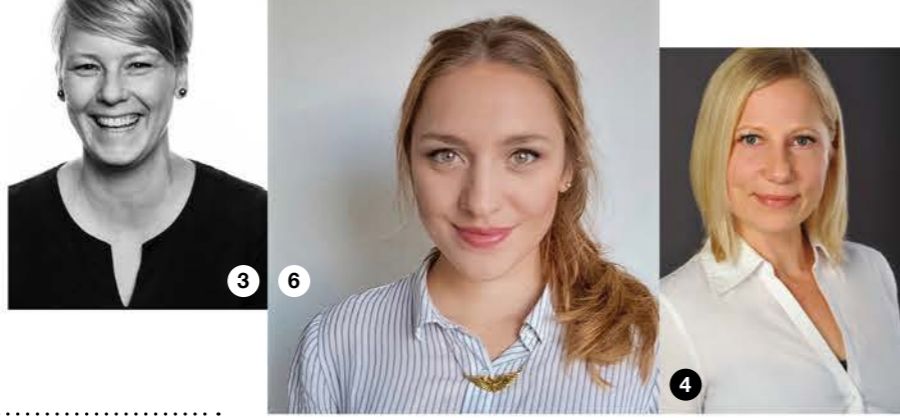▶ **Wienke Strathern,** IEAI Researcher ⬤

Annex

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 **Dr. Ana Catarina Fontes**
Research Associate

8 **Immanuel Klein**
Student Assistant

9 **Laura Lucaj**
Student Assistant
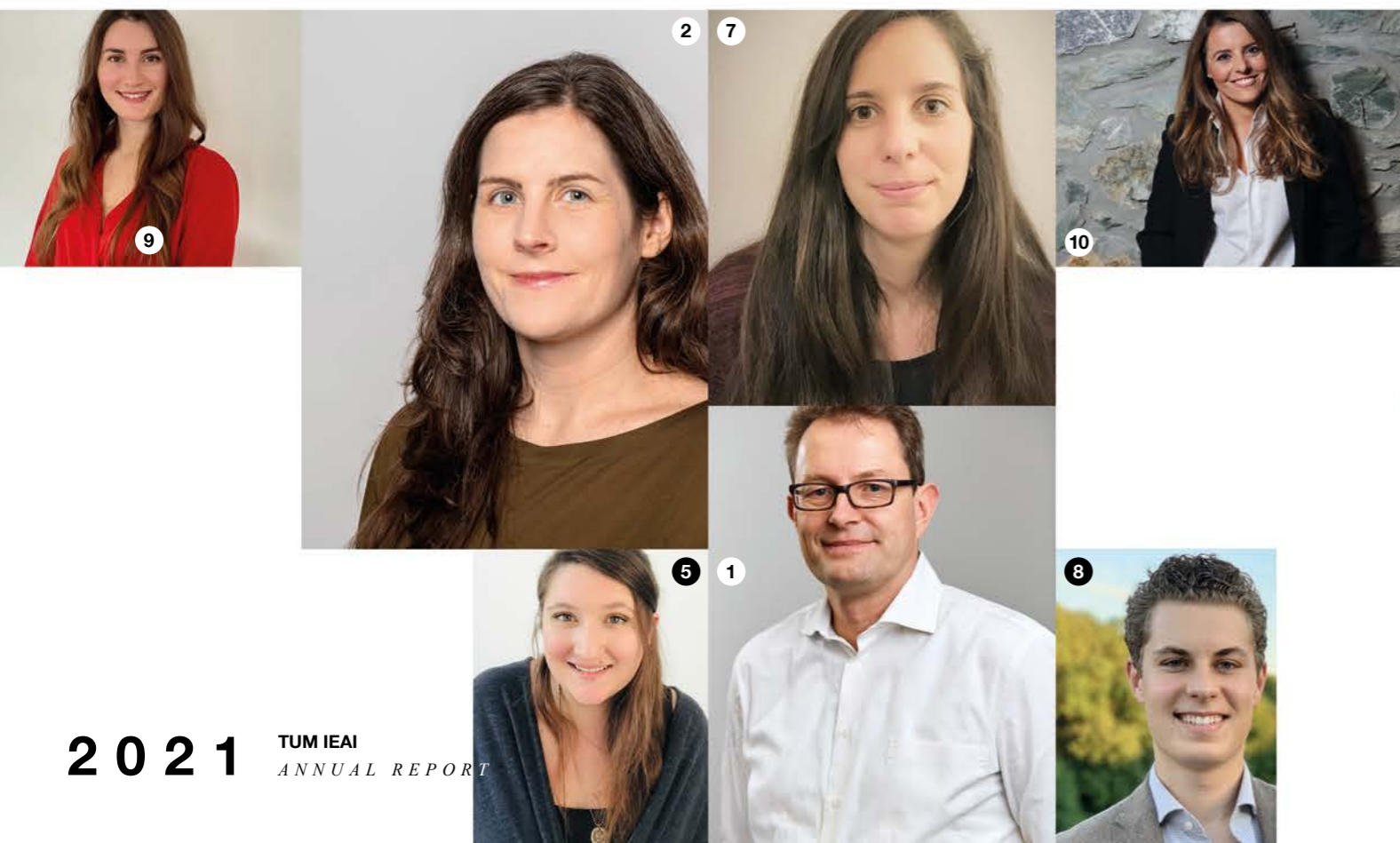
10 **Anastasia Aritzi**
Communications Consultant

Advisory Board Members

11 **Prof. Dr. Petra Ahrweiler**
Chair for Technology and Innovation Sociology,
Johannes Gutenberg University Mainz

12 **Auxiliary Bishop Anton Losinger**
Episcopal Vicar ('Bischofsvikar') for Bioethics
and Social Policy, Augsburg

13 **Prof. Dr. Christoph Meinel**
Director and CEO of the Hasso-Plattner-Institute
for Digital Engineering at the University of Potsdam

14 **Hannes Schwaderer**
President of Initiative D21 e.V. & CEO of Intel Germany

15 **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

16 **Zusanna Warso**
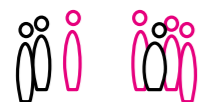Helsinki Foundation for Human Rights, Warsaw

External Affiliated Researchers

▶ **Johann Jakob Häußermann**
Doctoral Student and Researcher,
Center for Responsible Research
and Innovation, Fraunhofer IAO

▶ **Anna Hunkenschroer**
Doctoral Student and Consultant
at Boston Consulting Group

▶ **Raimund Waning**
Doctoral Student and Researcher, clockworkx GmbH

Project Affiliates

Project	Principal Investigators	Researchers
ANDRE – AutoNomous DRiving Ethics	▶ **Prof. Dr.-Ing. Markus Lienkamp,** Institute of Automotive Technology, TUM ▶ **Prof. Dr. Christoph Lütge,** School of Social Sciences and Technology, TUM	▶ **Maximilian Geißlinger, M.Sc.,** Institute of Automotive Technology, TUM ▶ **Franziska Poszler, M.Sc.,** School of Social Sciences and Technology, TUM
Artificial Intelligence for Earth Observation: Reasoning, Uncertainties, Ethics and Beyond (AI4EO)	▶ **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	▶ **Dr. Mrinalini Kochupillai,** Professorship for Signal Processing in Earth Observation, TUM
The Ethics and Practice of AI Localism at a Time of COVID-19 and Beyond	▶ **Prof. Stefaan Verhulst,** The GovLab and the Alliance for Public Interest Technology, New York University ▶ **Prof. Dr. Matthias Uhl,** Faculty of Computer Science, Technische Hochschule Ingolstadt ▶ **Prof. Jeannie Marie Paterson, PhD,** Centre for Artificial Intelligence and Digital Ethics, University of Melbourne ▶ **Mona Sloane,** PhD, The GovLab and the Alliance for Public Interest Technology, New York University	▶ **Yvette Maker,** The Centre for Artificial Intelligence and Digital Ethics, University of Melbourne, Australia ▶ **Manuela Schönmann,** Faculty of Computer Science, Technische Hochschule Ingolstadt ▶ **Andrew Young,** The GovLab, New York University
A Framework for Digital Transformation of Senegalese Universities: Lessons learned from the Covid-19 pandemic	TUM: ▶ **Prof. Dr. Christoph Lütge** ▶ **Caitlin Corrigan, PhD** Alioune Diop University of Bambey (UADB): ▶ **Prof. Dr. Fodé Camara** Virtual University Senegal (UVS): ▶ **Dr. Ndeye Massata Ndiaye** ▶ **Dr. Alpha Dia**	TUM: ▶ **Auxane Boch** ▶ **Ellen Hohma** ▶ **Matthew Cottrell** ▶ **Paloma Laye** ▶ **Dr. Cindy Cheng** Alioune Diop University of Bambey (UADB): ▶ **Prof. Dr. Gaoussou Camara** ▶ **Mouhamadou Lamine Ba, PhD** ▶ **Fam Fall Guène** ▶ **Ndeye Khady Mbengue** ▶ **Fatou Ndiaye** Virtual University Senegal (UVS): ▶ **Mamadou Abdourahmane Ba** ▶ **Ameth Thiam**
Personalized AI-based Interventions Against Online Norm Violations: Behavioral Effects and Ethical Implications	▶ **Prof. Dr. Anna Baumert,** Bergische Universität Wuppertal ▶ **Prof.. Jens Grossklags, PhD,** Department of Informatics, TUM ▶ **Dr. Julia Sasse,** Max Planck Institute for Research on Collective Goods	▶ **Niklas Cypris, M.Sc.,** Max-Planck-Institute for Research on Collective Goods ▶ **Tina Kuo, M.Sc.,** Department of Informatics, TUM ▶ **Slieman Halabi, M.Sc.,** International Max Planck Research School. Friedrich Schiller University, Jena
Public Trust in AI and the Ethical Implications: A Comparative Study of Governmental Use of AI During the Covid-19 Pandemic	▶ **Caitlin Corrigan, PhD,** Institute for Ethics in Artificial Intelligence, TUM ▶ **Prof. Dr. Christoph Lütge,** School of Social Sciences and Technology and Institute for Ethics in Artificial Intelligence, TUM	▶ **Auxane Boch, M.Sc.,** Institute for Ethics in Artificial Intelligence, TUM
Building Strategic Partnerships	▶ **Prof. Dr. Christoph Lütge,** School of Social Sciences and Technology and Institute for Ethics in Artificial Intelligence,, TUM ▶ **Caitlin Corrigan, PhD,** Institute for Ethics in Artificial Intelligence, TUM ▶ **Prof. Rafael A. Calvo, PhD,** Dyson School of Design Engineering, Imperial College London	▶ **Ellen Hohma,** Institute for Ethics in Artificial Intelligence, TUM ▶ **Ana Catarina Fontes, PhD,** Institute for Ethics in Artificial Intelligence, TUM ▶ **Ryan Burnell, PhD,** Leverhulme Centre for the Future of Intelligence ▶ **Christian Perrone,** ITS Rio
Online Firestorms and Resentment Propagation on Social Media: Dynamics, Predictability, and Mitigation.	▶ **Prof. Dr. Massimo Fornasier,** Department of Mathematics, TUM ▶ **Prof. Dr. Jürgen Pfeffer,** Bavarian School of Public Policy, TUM	▶ **Konstantin Riedl,** Department of Mathematics, TUM ▶ **Wienke Strathern,** Bavarian School of Public Policy, TUM
A Human Preference-Aware Optimization System	▶ **Prof. Dr. Tim Büthe,** School of Social Sciences and Technology, TUM ▶ **Prof. Dr. Johannes Fottner,** School of Engineering and Design, TUM	▶ **Charlotte Unruh,** PhD, School of Social Sciences and Technology, TUM ▶ **Charlotte Haid,** M.Sc., School of Engineering and Design, TUM
Understanding AI Ethics in the Workplace	▶ **Prof. Dr. Christoph Lütge,** School of Social Sciences and Technology and Institute for Ethics in Artificial Intelligence, TUM.	▶ **Anna Hunkenschroer, M.Sc.,** Boston Consulting Group

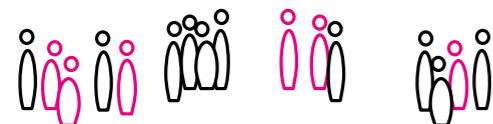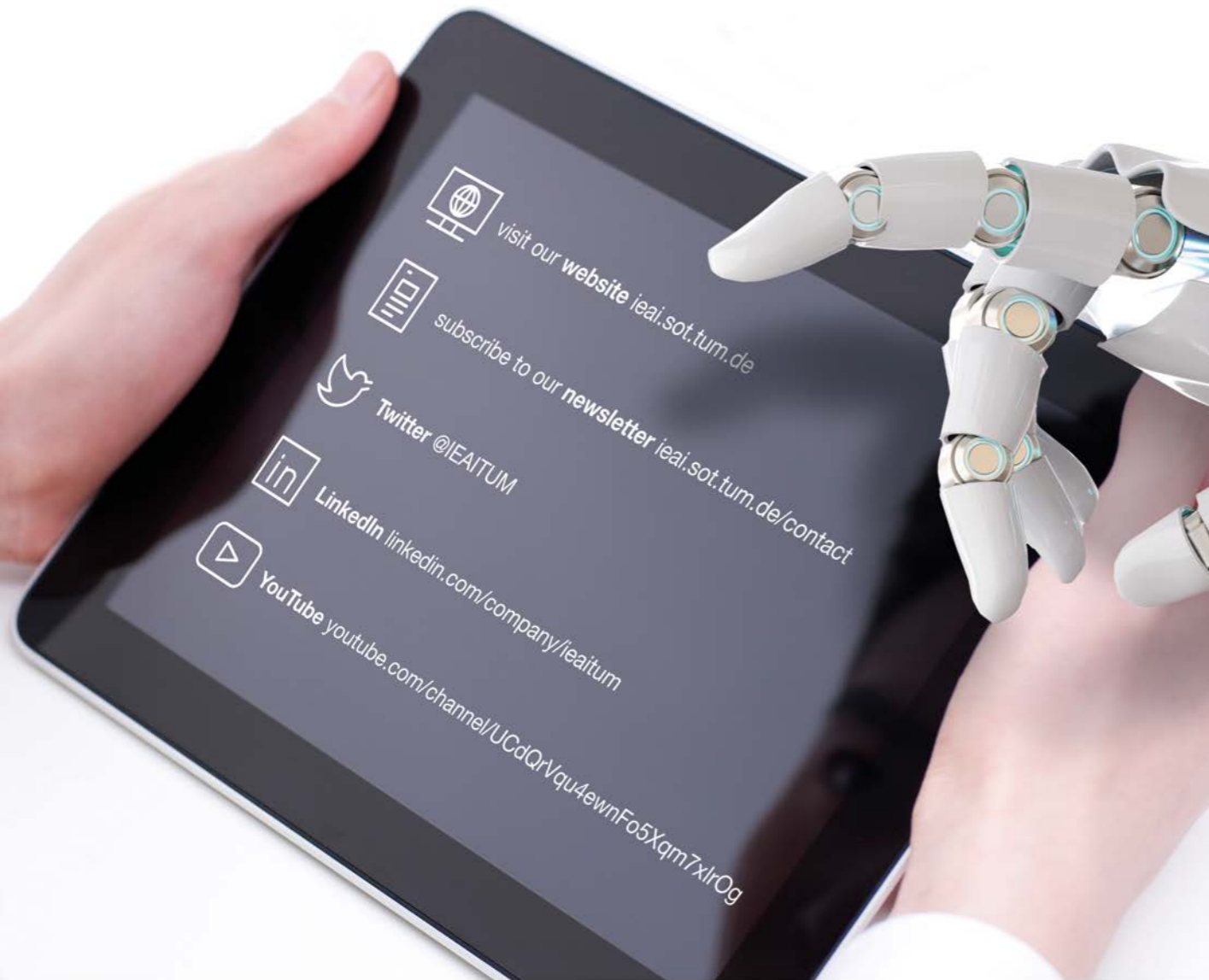| Project	Principal Investigators	Researchers
Towards an Accountability Framework for AI Systems	▶ **Prof. Dr.-Ing. Markus Lienkamp,** Institute of Automotive Technology, TUM ▶ **Prof. Dr. Christoph Lütge,** School of Social Sciences and Technology and Institute for Ethics in Artificial Intelligence, TUM.	▶ **Auxane Boch,** Institute for Ethics in Artificial Intelligence, TUM ▶ **Ellen Hohma,** Institute for Ethics in Artificial Intelligence, TUM ▶ **Rainer Trauth,** Institute of Automotive Technology, TUM
METHAD – Toward a MEdical ETHical ADvisor	▶ **Prof. Dr. med. Alena M. Buyx,** Chair of Ethics in Medicine and Health Technologies, TUM ▶ **Prof. Dr.-Ing. Klaus Diepold,** Department of Electrical and Computer Engineering, TUM	▶**Lukas J. Meier, PhD,** Chair of Ethics in Medicine and Health Technologies, TUM ▶ **Alice Hein, M.Sc.,** Department of Electrical and Computer Engineering, TUM
Online-Offline Spillovers – Potential Real-world Implications of Online Manipulation	▶ **Prof. Dr. Jürgen Pfeffer,** Bavarian School of Public Policy, TUM ▶ **Prof. Dr. Matthias Uhl,** Faculty of Computer Science, Technische Hochschule Ingolstadt	▶ **Gari Walkowitz, PhD,** School of Social Sciences and Technology, TUM ▶ **Wienke Strathern, M.Sc.,** Bavarian School of Public Policy, TUM
ETHAN – Ethical AI for Pandemic Management	▶ **Prof. Dr. Georg Groh,** Department of Informatics, TUM ▶ **Dr. Dirk Brand,** School of Public Leadership, University of Stellenbosch. ▶ **Dr. Johan van der Merwe,** Praelexis Ltd. ▶ **Dr. McElory Hoffmann,** Praelexis Ltd.	▶ **Prof. Dr. Ben M. Herbst,** Applied Mathematics Department, University of Stellenbosch. ▶ **Edoardo Mosca,** Department of Informatics, TUM ▶ **Tobias Eder,** Department of Informatics, TUM
Rule of Law, Legitimacy and Effective COVID-19 Control Technologies	▶ **Prof. Dr. Christian Djeffal,** School of Social Sciences and Technology, TUM ▶ **Prof. 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	▶ **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

Connect with Us!

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

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

Twitter @IEAITUM

LinkedIn linkedin.com/company/ieaitum

YouTube youtube.com/channel/UCdQrVqu4ewnFo5Xqm7xIrOg

In 2022, we will take our work to the next level.

TUM
IEAI

AI Ethics can be the compass that will help us navigate these uncharted waters of the new era that lies ahead, tackle challenges, and shape our future.