Responsible AI in Africa Network Virtual Workshop

28 May 2020

Introduction and Ethics in AI

Prof. Dr. Christoph Lütge
Dr. Caitlin Corrigan
TUM Institute for Ethics in Artificial Intelligence
Technical University of Munich

Ing. Dr. JJ Kponyo
Faculty of Electrical and Computer Engineering
Kwame Nkrumah University of Science and Technology
Virtual Meeting Etiquette

✓ Please **turn off** your **microphone** and **camera** during the session!

✓ Please submit your **questions** in the discussion period **via chat or by “raising a hand”**.

✓ Please let us know in the chat if you have **technical difficulties**!

Responsible AI in Africa Network Virtual Workshop:

Introduction to RAIN-Africa and AI for Sustainable Development

Dr. Caitlin Corrigan

TUM Institute for Ethics in Artificial Intelligence
Technical University of Munich
For years TUM has been a driving force in researching the mutual interactions of science, technology and society and has anchored “Human-Centered Engineering” as a central point in its new strategic guidelines.

The IEAI was announced in 2019, as an integral part of the Munich Center for Technology in Society (MCTS), one of the leading centers of Science and Technology Studies in Germany.
IEAI’s Mission and Approach

• What should be possible in AI?
• How do we ensure that as many people as possible benefit from the rewards of AI?

Generation of global, fair and interdisciplinary guidelines for the ethical development and implementation of AI throughout society.
The Responsible Artificial Intelligence Network–Africa (RAIN-Africa) was launched in February 2020, with the IEAI and KNUST as coordinating partners.

- Bring together Africa’s AI stakeholders from:
  - Policy
  - Academia
  - Industry
  - Civil Society
  - International Community
- Provide a platform for cooperation between stakeholders:
  - Research projects – within and external to RAIN-Africa
  - Workshops and conferences
  - Growing network
- Partners:
  - TUM
  - KNUST
  - More to follow soon
THE RESPONSIBLE AI FORUM 2020

www.responsibleaiforum.com
Focus on AI and Sustainable Development
Focus on the use of AI in Africa

Figure 5.3

Firms' preparedness for the Fourth Industrial Revolution

The majority of African firms report moderate to very low levels of business preparedness for five key 4IR technologies. Notably, firms are least prepared for artificial intelligence/robotics and blockchain technologies. Experts say that the low levels of preparedness stem from the inability of firm leadership to develop effective digital strategies, as well as low levels of education and skills of employees.

Preparedness for (% of respondents)

- AI/robotics
- Internet of things (IoT)
- Big data/data mining
- 3D printing
- Blockchain

• SDG 2: No Hunger
  • Advancements in Agriculture

• SDG 3: Good Health and Well Being
  • Do more with less – triage, diagnostics
  • Individualized treatments
  • Reproductive Health- education (related to SDG 5)

• SDG 4: Quality Education
  • Example: Muse App
• SDGs 7 and 11 – Clean Energy and Sustainable Cities
  • Smart and Interconnected Cities and Homes
  • Smart Grids
  • Interconnect Traffic and Autonomous Driving
SDG: AI for Environment

- SDG 14/15: Life Below Water and Life on Land
- SDG 13: Climate Action

**Figure 4.1**

The impact of climate change on sub-Saharan Africa’s GDP

Climate change is predicted to significantly decrease Africa’s GDP through mechanisms such as lowered crop yields, reduced agricultural and labor productivity, and damage to human health. Assuming no major changes in the world’s social, economic, and technological trends, climate change resulting in a 3°C temperature increase will decrease Africa’s GDP by as much as 8.6 percent per year after 2100. If climate change is limited to the 1.5°C agreed to in the Paris Agreement, the decrease in GDP will be significantly less—only 3.8 percent per year after 2100.

![Graph showing the impact of climate change on GDP](image-url)

Working Together

• SDG 17: Partnership For the Goals
  • RAIN-Africa
What does this mean in the African Context?

Ing. Dr. JJ Kponyo
Faculty of Electrical and Computer Engineering
Kwame Nkrumah University of Science and Technology
The Opportunities and Challenges of AI in the African Context

Ing. Dr. Jerry John Kponyo
Dean, FECE, KNUST
Presentation Overview

1. Introduction to KNUST
2. Overview of AI
3. Opportunities and Strategies for Africa
4. Areas of Collaboration for RAIN Africa
5. Conclusion
<table>
<thead>
<tr>
<th>1951 Establishment</th>
<th>1961 Fully-fledged University</th>
<th>Location: Kumasi, the Capital City of the Ashanti Region of Ghana</th>
<th>55,580 Students</th>
<th>840 Full Time Teaching Staff</th>
<th>06 Colleges</th>
<th>06 Students’ Halls</th>
</tr>
</thead>
</table>

**KNUST at a glance!**
The Fourth Industrial Revolution

“We stand on the brink of a technological revolution that will fundamentally alter the way we live, work and relate to one another. In its scale, scope and complexity, the transformation will be unlike anything humankind has experienced before”

Klaus Schwab
Founder & Executive Chairman,
World Economic Forum
Overview of Industrial Revolutions

The Dawn of the Fourth Industrial Revolution

1. Digitization / Integration of value chains
2. Digitization of product and service offerings
3. Digital business models and customer access

1.0 1760-1840
Steam Engineering

2.0 1830s-1915
Assembly Line

3.0 1969-2010s
Computing / Internet Nuclear Energy

1. Mobile Devices
2. IoT platforms
3. Location detection technologies
4. Advanced human-machine interfaces
5. Authentication & fraud detection

Augmented Reality
Cloud Computing
Multilevel customer interaction and customer profiling
Big data analytics
Smart Sensors
3D printing
Facts About AI

1. Artificial intelligence has come into the mainstream because it allows us to make sense of an increasingly large amount of data in real time. Indeed, worldwide data will grow 61% to 175 zettabytes by 2025. (Source: IDC)

2. Thirty-seven percent of organizations have implemented AI in some form. That's a 270% increase over the last four years. (Source: Gartner)
Facts About AI

3. By 2021, 80% of emerging technologies will have AI foundations. *(Source: Gartner)*

4. Worldwide spending on artificial intelligence systems was forecast to reach $35.8 billion in 2019, an increase of 44% over the amount spent in 2018. *(Source: IDC)*

5. Marketing and sales prioritize AI and machine learning higher than any other department in enterprises today (40%). *(Source: Forbes)*
6. Demand for AI talent has doubled in the last two years. Technology and financial service companies are currently absorbing 60% of AI talent. (Source: MMC Ventures)

7. Sixty-three percent of people prefer to message a chat bot vs. talk with a human when communicating with a business. (Source: G2 Crowd)
Artificial intelligence (AI) can improve various aspects of healthcare.

- Reduction in annual expenditure
- Useful in early detection of diseases
- Provide round-the-clock monitoring of chronic disorders
- Help limit the exposure of healthcare professionals in contagious environments.
AI in Healthcare

Developments in AI will significantly improve:

- Health services
- Diagnostics
- Personalized Medicine
AI in Agriculture An Opportunity

Artificial intelligence allows farmers to increase yields by spotting pests and diseases early and has led to a new wave of indoor growing startups.

Farmers are using artificial intelligence to help grow food by combatting disease and pests.
AI in Agriculture

According to Markets and Markets, an Indian research company, in 2018 the worldwide AI in agriculture market was valued at €545 million and, by 2025, is expected to reach €2.4 billion as more and more smallholder farmers adopt new, data-driven technologies.
AI in Agriculture

AI provides opportunities in Africa to;

1. Analyze farm data
2. Tackle Labor challenges
3. Protection against pests
AI in Education An Opportunity

AI in education generally focuses on identifying what a student does and doesn’t know through diagnostic testing and then developing personalized curricula based on each student’s specific needs.
AI in Education

AI has the potential to revolutionize education in Africa in the following areas:

1. Simplifying administrative tasks
2. Smart content
3. Personalized Learning
AI in Business, Another Opportunity

Artificial intelligence is already widely used in business applications, including automation, data analytics, and natural language processing. Across industries, these three fields of AI are streamlining operations and improving efficiencies.
AI in Business

Other common uses for AI in business include:

1. Transferring and cross-referencing data; updating files
2. Consumer behavior forecasting and product recommendations
3. Fraud detection
4. Personalized advertising and marketing
5. Customer service via telephone or chatbots
Strategies for AI in Africa

1. Increased digital access
2. Building a robust labor force who are tech savvy
3. Developing digital platforms for rolling out digital services
4. Digital financial services to enhance trade
5. Creating an ecosystem to promote digital entrepreneurship and innovation

The idea is for Africa to become the “startup continent”
RAIN Africa Can Collaborate in the following:

1. Reinforcement learning
2. Ethics in AI
3. Natural Language processing
4. Convergence of AI and other emerging technologies
5. Neural networks
6. Digital platforms for Agriculture
7. Digital platforms for health
8. Digital platforms for education
9. Digital platforms for business
Conclusion

With AI leading the 4th industrial revolution, discussions AI and the responsible use of it is crucial and RAIN Africa couldn’t have come at a better time.
Thank You
RAIN-Africa Workshop: Ethics and Artificial Intelligence

Challenges and Guidelines

Prof. Dr. Christoph Lütge
TUM Institute for Ethics in Artificial Intelligence
Technical University of Munich
Challenges of AI

Technical challenges

- Dependence on the accuracy of technical systems, danger of technical errors
  - Telemedicine
  - Autonomous driving
  - Loss of once autonomous decisions
- Increased vulnerability against cyber attacks, danger of cyber wars
- Privacy and danger of data misuse
  - Development of GDPR
  - Differing perceptions of the problem
- Digital literacy
  - Education in digital literacy often lacking
Challenges of AI

Fakes and Manipulations

- Deepfakes
  - Software to generate or edit video material, e.g. creating a fake video of a politician’s speech
- DeepMasterPrints
  - Generate fake finger prints to e.g. unlock devices
- Fake social media bots
  - Bots acting like humans in social media to influence public opinions
- Preventing crime
  - Automated blackmailing, intelligent hacking attacks
Challenges of AI

Ethical problems

- Hiring software used by Amazon 2014 – 2017
  - Discrimination against women
- Airbnb smart pricing algorithm may widen ethnic gap
- Facebook’s Software supports separation based on race

**FACEBOOK IS LETTING ADVERTISERS EXCLUDE USERS BY RACE**

*PACIFIC STANDARD STAFF* · UPDATED: JUN 14, 2017 · ORIGINAL: NOV 2, 2016

- Cheat software for pollution tests at VW
Inner Ethical Conflicts

Care robots

• Care robots – do they endanger or preserve human dignity?

• Some studies indicate that older adults prefer robot assistance over human assistance for many instrumental (e.g., housekeeping, laundry, medication reminders) activities of daily living

• However, older adults seem less open to robot assistance for other activities of daily living (e.g., shaving, hair care)

• Elderly people might find it more dignified to be treated by robots rather than by humans – in some contexts
Inner Ethical Conflicts

Acceptance of robots

![Bar charts comparing acceptance of robots in Japan, the U.S., and Germany across different contexts such as offices, hotels, stores, nursing care, medical service, and education. Each context is represented with bars indicating the percentage of respondents who very much want to use, somewhat want to use, somewhat don't want to use, don't want to use at all, and those who don't know.](image-url)
Inner Ethical Conflicts

**Internet bots**

- $4.5$ Billion is expected to be invested in enterprise intelligent assistants by 2021. (Opus Research, 2017)
- Should bots be required to identify themselves as bots?
- Does the bot serve me or the service provider?
- Who owns the information/data shared with a chatbot?
- Does the benefit of a chatbot get lost when the user knows it is a chatbot and hence does not trust it?

Inner Ethical Conflicts

**Autopilot**

- Should it be indicated when a vehicle is on autopilot?
- For a car?
- In an airplane?
- Or would it make people feel more uncomfortable?
Inner Ethical Conflicts

Trust bias

- Often people have a bias towards trusting human control vs autonomous systems.

- Autonomous vehicles
  - Statistically, AVs would drive more safely than humans, yet people have a hard time giving up control.
  - Of 38 accidents where an AV was involved, in 37 cases the accident was made by the human driver (Martin, 2018).
Relevance of Ethical Guidelines

Was the programmer at Bosch aware of the entire consequences his piece of code would have?

Dieselgate

• Consequentialist view: What consequences could my piece of code have (worst case)?

Affected thousands of people

➢ Financial damage
➢ Pollution / health
➢ Reputation lost
➢ Time effort to replace cars

Often consequences are underrated and the amount of people affected is underestimated – hence there is a need for explicit codes of ethics
**Ethical Considerations**

**Principles for an ethical AI:**

**Human centered approach**

1. **Beneficence**
   - Promoting well-being, preserving dignity, and sustaining the planet

2. **Non-maleficence**
   - Ensuring privacy, security and “capability caution” (upper limit of future AI capabilities)

3. **Autonomy**
   - Striking a balance between the decision-making power we retain for ourselves and that which we delegate to AI

4. **Justice**
   - Creating benefits that are (or could be) shared, preserving solidarity

5. **Explicability**
   - Enabling the other principles through intelligibility and accountability
Principles for an ethical AI

Example: OECD Principles on AI (adopted May 2019)

OECD Expert Group

- 36 OECD members as well as Argentina, Brazil, Colombia, Costa Rica, Peru and Romania
- AI should be fair, transparent and accountable
- Companies to disclose enough about how their systems work
- Guidelines are not legally binding
- Policy observatory on AI planned to be launched in 2019
IEEE (Institute of Electrical and Electronics Engineers)

IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems

• “Ethically Aligned Design“

• General principles
  • Human rights, well-being, data agency, effectiveness, transparency, accountability, awareness of misuse, competence

• Extensive ethical and philosophical consideration of morality of amoral systems

• Ethics Certification Program for Autonomous and Intelligent Systems (ECPAIS)
Principles for an ethical AI

Core Principles for AI: Example Autonomous Driving

German National Ethics Commission on Automated and Connected Driving

- 14 experts from a wide range of disciplines, civil society and companies
- In June 2017, the Commission presented its final report in Berlin
- The core of the report consists in 20 “ethical rules for automated and connected vehicular traffic”
Core Principles for AI: Example Autonomous Driving

Rule 9

„In the event of unavoidable accident situations, any distinction based on personal features (age, gender, physical or mental constitution) is strictly prohibited. It is also prohibited to offset victims against one another. General programming to reduce the number of personal injuries may be justifiable. Those parties involved in the generation of mobility risks must not sacrifice non-involved parties.“
Economic Rationale

Corporate Social Responsibility

Consider systematically the interests of others who may be relevant for the company's long term future
Ethical AI in Companies – Critical Considerations

Source: Capgemini Research Institute.
Technology creates new ethical problems

The moral machine: findings

- Cross-cultural ethical variations: 3 major clusters of countries (see figures)
- „Differences correlate with modern institutions and deep cultural traits“
- E.g. „Collectivist cultures, which emphasize the respect that is due to older members of the community, show(ing) a weaker preference for sparing younger characters.“
- Authors state that their findings can „contribute to developing global, socially acceptable principles for machine ethics.“
Critical considerations for professionals

**Awareness to misuse**

- How could one try to hack into the system?
- Can the system be used to serve a purpose different from the programmer’s intention?
  - Use personas to access how the system could be misused
- Do the algorithms prevent unethical use of the software?
- What happens if the system fails? What potential damage can it cause? Worst case scenario?

---

Reaction to Facebook’s race excluding advertising software

While coding, access the risk of unethical misuse of the software
Critical considerations for professionals

Awareness to accessibility

- The ability to make use of computer software gets more and more critical in nowadays
- Software developers should develop programs, that don’t exclude disabled people
- Make use of a software accessibility checklist
  - Colorblindness
  - Size of text
  - Deafness
  - Instructions how to use software
Action steps towards ethical AI

• Data security
  • AI’s need data, often quite sensitive data. Companies must protect themselves against data breaches, and use state of the art data security measures. Further data should not be sold without the permission of the consumer, e.g. Cambridge Analytica

• Data bias
  • Often, the training sets introduce human biases. More data analysts should explore the data sets get rid of the bias. E.g. IBM’s Trusted AI toolkits

• XAI
  • Companies should make use of for example counterfactuals to avoid algorithmic biases, e.g. Google’s What-If tool
Example: Explainable AI (XAI)

https://www.datanami.com/2018/05/30/opening-up-black-boxes-with-explainable-ai/
Conclusion: What is needed for an ethical AI?

Trust in AI (internal and external) is key!

- More data is needed to assess algorithmic impact
- Rethink: Human-machine collaboration instead of AI replacing humans
- Focus on meeting societal acceptance.
- Guidelines and frameworks for ethical AI
- Tools for companies to follow ethical frameworks and establish trust (internal and external)
AI will not fly without ethics!
Thanks for your attention!
Please submit your questions in the discussion period via chat or by “raising a hand”.

Discussion Points:

• GIZ and the FAIR Forward Project
• African Digital Right Hub and UNESCO's Human Centred AI
• What are some of the challenges to networking on this topic?
• What would be useful activities to promote exchange and collaboration?
• Are there specific topics related to responsible AI you are interested or would like to form sub-groups on?