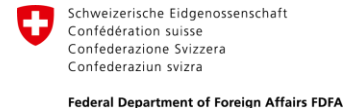


# ARTIFICIAL INTELLIGENCE, ETHICS AND HUMAN RIGHTS: IDENTIFYING GAPS AND MITIGATING GLOBAL RISKS.

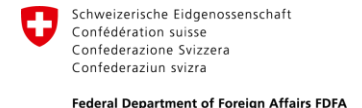
Dr. Alexander Kriebitz

Ludwig Maximilian University of Munich  
Technical University of Munich



# Short Introduction

- Post-Doctoral Researcher at Ludwig Maximilian University
- Co-founder of **iuvenal research GmbH**
- IEEE SA Co-Lead of the Ad-hoc Committee AI and Human Rights
- Scientific coordinator of the **Munich Convention on AI, Data and Human Rights**
- Publication on AI and Human Rights with the UN Global Compact Network Germany



# ETHICS AND HUMAN RIGHTS



# Ethics - What is it?

## Ethics and Human Rights – A Special Relationship

1. Ethics defines principles, norms and means of individual and collective conduct along the lines of good and bad.
2. It discusses principles and norms to prevent harm (e.g. physical harm, financial losses) and to create desirable outcomes.
3. In moral dilemmas, we need to balance between these two ends. **Example:** In order to contain the spread of the virus, national governments restricted individual autonomy.
4. Human rights specify the meaning of harm and desirable outcomes and are codified in international conventions (e.g. Universal Declaration of Human Rights, UDHR).

### Examples of Human Rights:

- Right to Human Dignity (UDHR, Art. 1)
- Right to Individual Freedom (UDHR, Art. 3)
- Right to Equal Treatment (UDHR, Art. 7)
- Right to Freedom of Expression (UDHR, Art. 19)
- Right to Occupational Safety (UDHR, Art. 23)
- Right to Health (UDHR, Art. 25)

# THE ETHICAL IMPACT OF AI



# Artificial Intelligence – A Working Definition

“Artificial intelligence (AI) refers to systems that **display intelligent behaviour** by analysing their **environment** and **taking actions** – with some **degree of autonomy** – to achieve **specific goals**”

*(High-Level Expert Group on AI)*

## Use cases of AI:

- Large Language Models (Chat GPT)
- Diagnosis of patients
- Human Resources
- Autonomous driving
- Satellite imagery processing and analysis
- Military use of AI

# The Ethical Impact of AI

Underlying issues associated with AI:

- **Data use**
- Automated decision making
- Algorithmic bias
- Algorithmic opacity

## Example 1:

An AI system evaluates communication patterns of individuals in order to determine their mental state.

**Adverse Human Rights Impact:**

UDHR: Article 12

*Right to Privacy*



# The Ethical Impact of AI

Underlying issues associated with AI:

- Data use
- **Automated decision making**
- Algorithmic bias
- Algorithmic opacity

## Example 2:

An individual is unaware that AI generates a personal profile, which is then used to determine their compensation.

### Adverse Human Rights Impact:

UDHR: Article 23

*Right to Work and Fair  
Compensation*

# The Ethical Impact of AI

Underlying issues associated with AI:

- Data use
- Automated decision making
- **Algorithmic bias**
- Algorithmic opacity

## Example 3:

A person is denied medical treatment because of a bias in the system, due to the ethnicity of the person.

**Adverse Human Rights Impact:**

UDHR: Article 25

*Right to Health and Well-being*

# The Ethical Impact of AI

Underlying issues associated with AI:

- Data use
- Automated decision making
- Algorithmic bias
- **Algorithmic opacity**

## Example 4:

A sophisticated AI system decides not to hire a specific candidate, but the reasoning behind the decision is opaque and cannot be explained.

### Adverse Human Rights Impact:

UDHR: Article 23

*Right to Work*

# AI Ethics and AI Governance – The State of Play

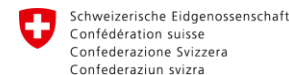
## Emerging landscape of AI governance

- Montreal Declaration for Responsible AI (2018)
- IEEE Ethically Aligned Design (2019)
- NIST AI Risk Management Framework (2023)
- US AI Bill of Rights (2023)
- EU AI Act (2024)
- Framework Convention of the Council of Europe (2024)
- UN General Assembly Resolution A/78/L.49 (2024)



## Relevant principles and concepts

- Dignity
- Privacy
- Human oversight
- Transparency and explainability
- Non-discrimination and equal treatment
- Safety, Accuracy and Reliability
- Definition of High Risk AI (e.g. employment, health)



Federal Department of Foreign Affairs FDFA

# AI IN THE MILITARY CONTEXT



# Artificial Intelligence – A Working Definition

Current approaches to AI regulation **exclude military operations from their scope**. However, the convergence of algorithmic bias, opacity, and automated decision-making can result in violations of international humanitarian law:

- **Algorithmic bias** may prevent AI from distinguishing between combatants and civilians, leading to disproportionate harm.
- **Intentional misuse**: AI systems could be exploited to escalate collateral damage or to develop weapons targeting specific racial or ethnic groups.
- **Incitement of ethnic hatred**: Additionally, AI can be weaponized for propaganda, inciting war or ethnic hatred, especially through social media platform.

## Applicable norms in international law:

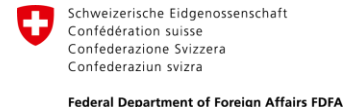
- Additional Protocol I to the Geneva Conventions (1977): Articles 48, 51(2), and 52(2)
- Convention on the Prohibition of Certain Conventional Weapons
- International Covenant on Civil and Political Rights: Article 20(2)

# Conclusion

## The bottomline.

In sharp contrast to the growing body of AI regulation, there is currently **no binding global framework governing the use of AI**, particularly in military operations. Although international conventions and customary international law offer general norms on the use of technology in warfare, more specific guidelines are necessary to **strengthen accountability and safeguard human rights**.

One aim of our Munich Convention is to drive this conversation and to anchor international human rights law and international humanitarian law as pillars of global AI governance!



# THANK YOU

