

AI Ethics and Regulation 2023-2024

Coordinator: Prof. dr. Jan De Bruyne
KU Leuven CiTiP



AI and human enhancement

Erik Kamenjašević, LL.M.
Researcher & doctoral candidate – KU Leuven
AIER 28.03.2024

Outline

- Setting the scene
- PART I:
 - Definitional issues
- PART II:
 - Societal issues
 - Ethics
- PART III:
 - Current regulatory framework
- Recap
- Q&A

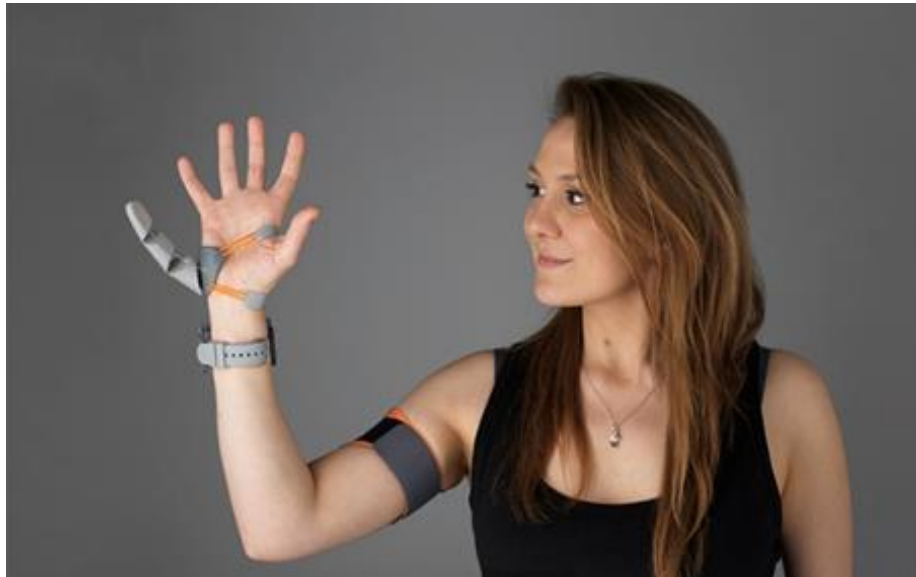
PART I

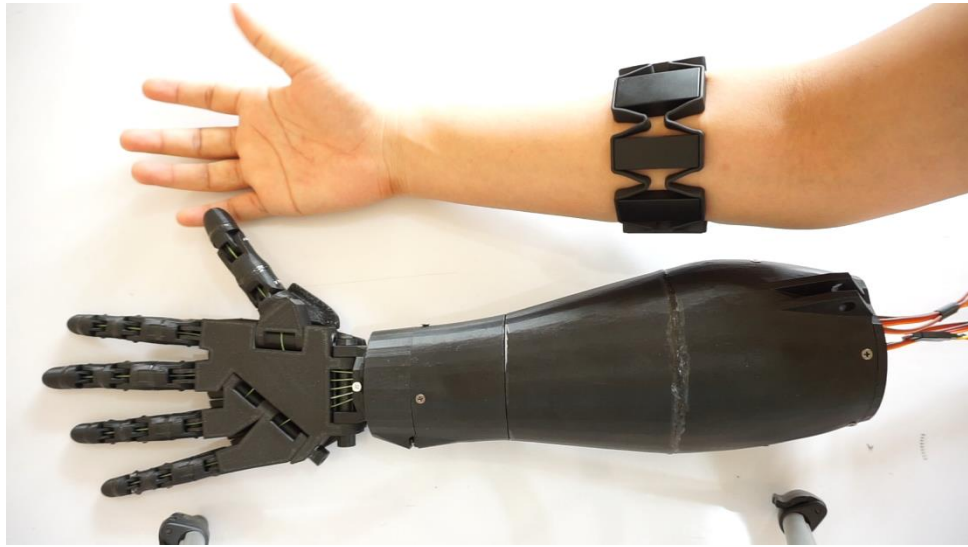




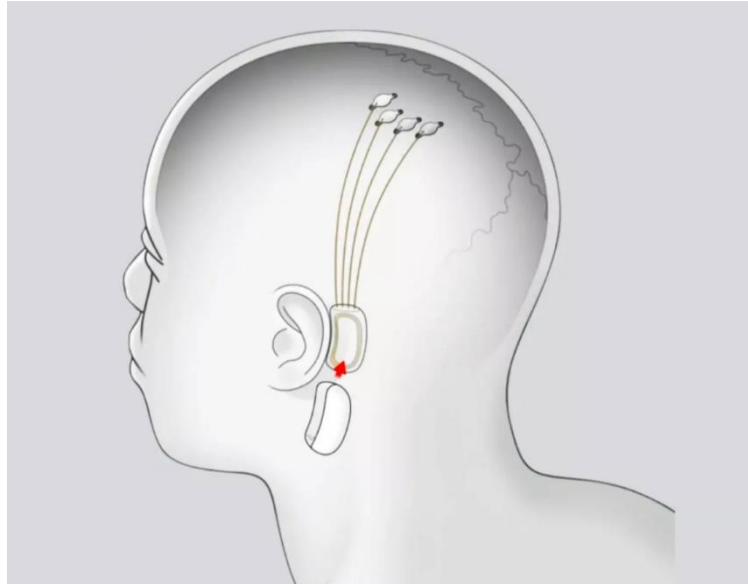




















Can you think of other examples?

Is it desirable to develop HETs?
Why (not)?

Defining human enhancement

- Literally: improvement of human beings (i.e., diet, books, medicine, technology)
- Many proposals to define HE
- Often omitted: the reference to the term normality (and distinction therapy – enhancement)
- For defining HE, it is the aspect of augmentation above levels of normality that is a distinctive characteristic



HE means **changing** physical, cognitive or emotional **human capacities** or the performance of these capacities **above levels of normality** by technology or supported by technology.

Normality

<i>Approach 1</i>	<i>Approach 2</i>
Species-typical functioning	Context-dependent
Statistical normality	Individual
Natural	Environment-dependent
Intuitive	Economic-dependent

Therapy or enhancement

- **Against** the distinction:
 - Pathologies or measurements that are not objectively or non-arbitrarily defined
 - Fluidity of notions normality and health causes vagueness
 - Nonetheless, sometimes useful for law/policy-making
- **In favor** of the distinction:
 - Practical reasons
 - Define moral zones without sharp boundaries
 - Define boundaries of medical practice

The notion of health



- **Health:** “a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity”.
- **The right to health:** an inclusive right that needs to be considered in the context of the social determinants of health. They are defined as “the conditions in which people are born, grow, live, work and age. These circumstances are shaped by the distribution of money, power and resources at global, national and local levels”.
- Controversial; but *why*?
- In the context of HETs
 - HETs impact on (mental) health
 - Suggestion to consider the two notions as one

Key points

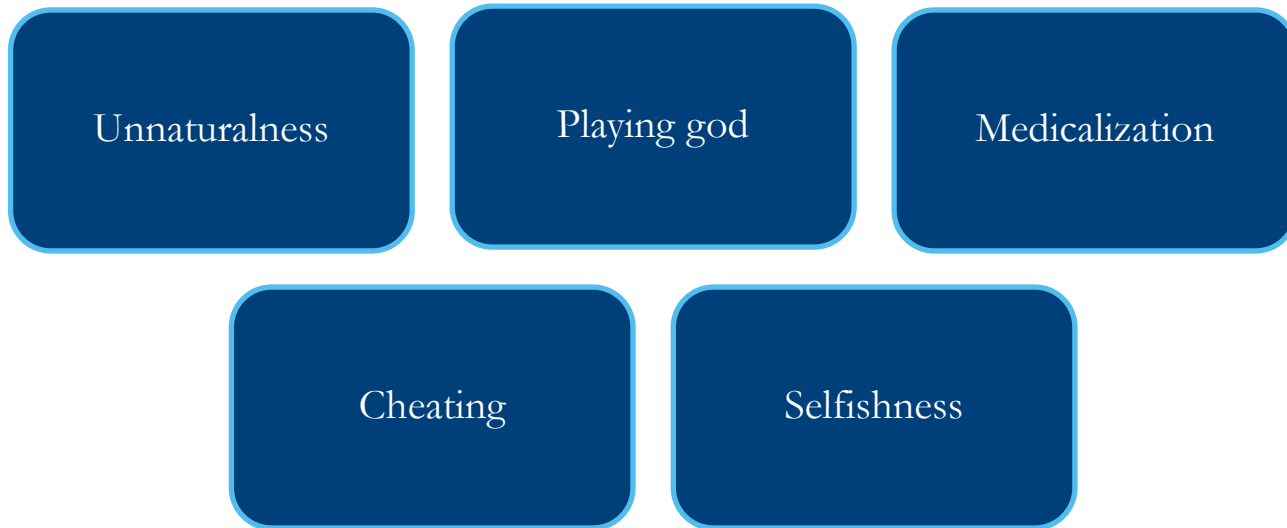
- 'Normality'
 - A distinctive notion in the HE definition
 - Multiple levels and context-dependent notion
 - For the permanent reflection on the well-being of people
- Therapy *vs* enhancement distinction
 - Used to raise a 'moral warning flag'
 - Limited usefulness
 - A reminder of the changing notion of HE

PART II

Ethics *vs* law

- Law as a formal regulatory system of governance within which different stakeholders (individuals, society, institutions) act
- Ethics as a basis for the law and a normative source for the interpretation of the law and guidance in addition to/in parallel with what the law demands (if the law exists)
- When the law does not exist/has not been interpreted in light of (the risks posed by) the new technology, ethics can support the management of known and potential risks, usage, development and deployment.

Religious, ideological and rhetorical arguments



Unnaturalness

Objection/argument:

- With HETs, people do not only change them but provoke changes in nature
- Discussed in the context of radical forms of enhancements (e.g., mind uploading) and together with other arguments (e.g., authenticity, identity, dignity, cheating, playing god) in order to show the severity of possible consequences for people who are trying to change nature
- Anything artificial must be rejected

Recommendations for policymakers and lawmakers:

- Understand what kind of HET the argument refers to
- Consider only a specific HET/a group of similar HETs at a time

Playing god

Objection/argument:

- By using HETs, people assume a god's role in creating or changing human nature
- Religious authors call for a precautionary stance toward HETs because people should not presuppose what is morally good for them and should fear what changes such technologies could bring

Recommendations for policymakers and lawmakers:

- Minor/no role in the legislative process due to a neutral approach the EU takes toward religion and does not support any specific confession

Medicalization

Objection/argument:

- Medicine or medically-related expertise should not be used for treating conditions and life experiences
- A negative connotation of the term causes seeing HETs as undesired

Recommendations for policymakers and lawmakers:

- Medicalization is not necessarily a bad thing since mood HETs can be beneficial in many ways, even if they could be considered as deepening the medicalization phenomena (e.g., self-empowerment, autonomy, independence, dignity)

Cheating

Objection/argument:

- Skills and achievements should be admired only when realized without a help of technology, medicine, enhancers

Recommendations for policymakers and lawmakers:

- Depending on the context, the cheating accusations change
- No rule prohibiting HETs, no cheating (e.g., doping in sports)

Selfishness

Objection/argument:

- People using HETs always put their personal interests ahead of others and cause some form of harm
- Implicitly mentioned as part of other objections and arguments (e.g., cheating, playing god)

Recommendations for policymakers and lawmakers:

- Mood HET users do not put their interests ahead of others nor cause any harm by using the HET

Ethical arguments

Justice

Identity

Autonomy

Dignity

Privacy

Safety and
prevention of harm

Justice

Objection/argument:

- Distributive justice: fear of new societal inequalities and the increase of the existing ones
- Resource allocation: fear that investments in HETs could result in disinvestments in prevention/care/disease treatments and technologies

Recommendations for policymakers and lawmakers:

- Analyze the possible pre-existing barriers to HETs: who is affected, why, and define mitigation measures
- Ensure incentives for R&D of the enhancement technology to be also used for therapeutic purposes

Identity

Objection/argument:

- Using HETs might change how people see themselves (e.g., a third-thumb study) and impact autonomy and authenticity by undermining self-expression abilities
- Fear of the so-called identity commodification

Recommendations for policymakers and lawmakers:

- Avoid arguments based on unrealistic scenarios about HETs
- Mood HETs are likely to impact identity in no other way than any other technology (identity is not static)
- Establish clinical trials to study specific HET's impact on identity

Autonomy

Objection/argument:

- 3 approaches:
 - Negative impact: HETs might undermine the human capacity to act freely and take responsibility; the question of the authenticity of the action; the coercion problem
 - Positive impact: to respect the autonomy (and dignity) of a person requires allowing enhancements (if they are safe and cause no harm to others)
 - Context-sensitive approach considers cognitive diversity of individuals

Recommendations for policymakers and lawmakers:

- Case-by-case analysis and qualification of the dimensions of autonomy that the HET could impact in order to avoid speculations
- Establish clinical trials to study specific HET's impact on autonomy

Dignity

Objection/argument:

- 2 approaches:
 - HETs positively impact dignity (and autonomy) because they improve the health and well-being of the user
 - HETs negatively impact dignity because they lead to coercion, medicalization, discrimination, stigmatization, and justice-related issues

Recommendations for policymakers and lawmakers:

- The concept of dignity can be used to define HETs' boundaries and help judge them as morally right or wrong (to decide about stigmatization, shame, distributive justice, etc.)

Privacy

Objection/argument:

- New types of data (e.g., mood, feelings, emotions) available for collection and processing are leading to new threats to privacy due to the technological advancements
- Data breaches might negatively impact the identity, autonomy, and dignity of a person as well as lead to new forms of discrimination and stigmatization

Recommendations for policymakers and lawmakers:

- Examine whether the currently applicable frameworks provide a sufficient level of privacy protection
- Reconsider the notion of mental integrity and explore the need for strengthening the notion of mental privacy

Safety and prevention of harm

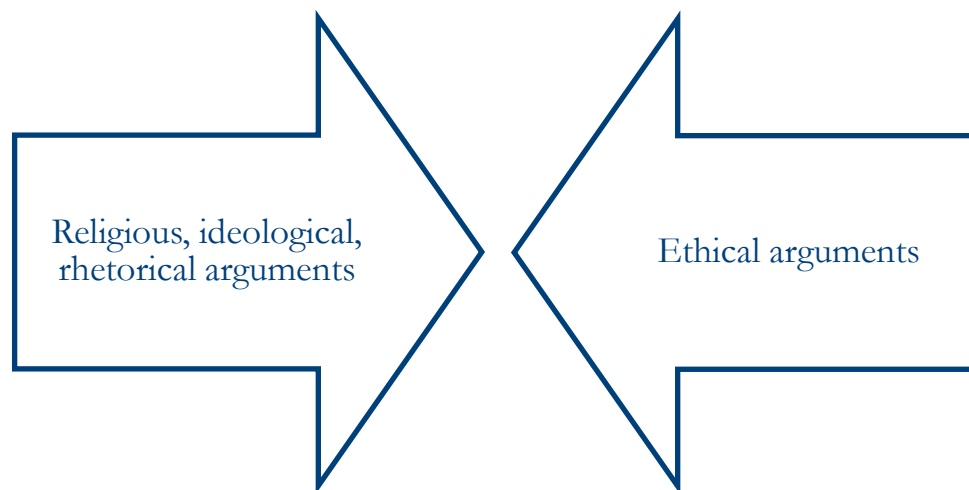
Objection/argument:

- Safety risks concern the individual using mood HETs, but they might affect society and the environment
- Risks depend on different HET's abilities
- Extra caution is needed when children and young adults use HETs

Recommendations for policymakers and lawmakers:

- Establish clinical trials in patients and healthy individuals
- Promote transparency about the safety risks and inform users who should make an informed decision based on the proportionality test
- Establish public-private partnerships to study HETs' benefits and safety risks
- Establish post-market surveillance mechanisms to monitor possible risks after HETs start circulating on the market (e.g., MDR)

Let's discuss!



What is the difference between these arguments?
Which arguments can/should be used for regulating HETs?
What should policymakers and legislators do with these different arguments?
Why (not)?

Overlaps and connections

Authors using religious, ideological and rhetorical arguments refer to ethical arguments and *vice versa*. They are often based on HETs seen in sci-fi films/books, which causes confusion.

Recommendations for policymakers and lawmakers:

- Ensure a clear description of the technology and, based on it, define potential risks and propose mitigating measures
- Clearly distinguish different types of arguments and focus on the ethical arguments first
- Religious, ideological and rhetorical arguments should be tackled only to the extent they can be translated into general principles
- The remaining risks should be tackled through policy and legislation

Key points & open ethical questions

- Refrain from religious, ideological, and rhetorical argumentation
- Clearly define the technology and its risks
- Address the pre-existing barriers (e.g., such as those leading to the distributive justice issue)
- Develop a mechanism to incentivize R&D of therapy and enhancement technology
- Establish clinical trials involving patients and healthy individuals for studying mood HETs' impact on the user's identity, autonomy, dignity, and safety
- Adapt existing or develop new rules dealing with privacy, discrimination, stigmatization, and equality fit for new technological advancements
- Establish a post-market surveillance mechanism

PART III

Current regulatory framework

United Nations

Council of Europe

European Union

United Nations

- International Bill of Human Rights:
 - Universal Declaration of Human Rights
 - International Covenant on Civil and Political Rights
 - Prohibition of experimentation without a valid consent for severe risks, freedom of thought, freedom of expression
 - International Covenant on Economic, Social and Cultural Rights
 - Right to self-determination → right to improve oneself via HETs
 - Right to the highest attainable standard of health; scientific benefits
- Universal Declaration on Bioethics and Human Rights
 - Requires minimizing harm and maximizing the benefits of research and technologies
 - Autonomy and the highest attainable standard of health

Council of Europe

- Oviedo Convention
 - Dignity, integrity, identity, autonomy, distributive justice, privacy
 - Reconceptualisation of valid consent and professional standards
- ECHR
 - Integrity, privacy, dignity, integrity, freedoms of expression and thought
 - Expansion of the notion of mental privacy, mental integrity
- Convention 108+
 - Data protection
 - Qualification of mood data as health data, valid consent for data processing
- European Social Charter
 - Safe and healthy working conditions, workers associations
 - Ban on imposing usage of mood HETs in the employment context

European Union

- Charter of Fundamental Rights of the European Union
 - Dignity (non-discrimination, non-stigmatization), informed consent, privacy, freedoms of thought and expression, right to health, distributive justice
- Secondary EU legislation:
 - GDPR
 - MDR
 - Directive 2001/83/EC
 - Regulation 726/2004
 - EU cybersecurity framework
 - MDR, NIS 2 Directive, CSA, GDPR

European Union

Forthcoming EU legislation relevant to mood HETs

- AI Act proposal*
- AI Liability Directive proposal
- EHDS proposal
- New Directive 2001/83/EC proposal
- New Regulation 726/2004 proposal
- SoHO Regulation proposal

AI Act proposal

- Prohibition of the emotion recognition systems in workplace and education
 - Recital 18: The notion refers to emotions or intentions such as happiness, sadness, anger, surprise, disgust, embarrassment, excitement, shame, contempt, satisfaction and amusement.
- Difference between emotion and mood?
 - Mood lasts longer
 - Mood is less intense than emotion
 - Mood is a reaction to a cumulative sequence of events
 - Emotion is more visible to others, mood is internal
 - Mood reflects the underlying feelings of people

Principles and rights addressed by the existing laws

- Dignity
- Integrity
- Identity
- Autonomy
- Prohibition of discrimination, stigmatization and coercion
- Distributive justice
- Privacy and data protection
- Self-determination
- Just conditions at work, safe and healthy working conditions
- Workers associations
- Right to health
- Right to enjoy scientific progress
- Prohibition of inhuman and degrading treatment
- Right to hold an opinion
- Right to peaceful assembly
- Freedom of thought
- Freedom of expression
- Safety and prevention of harm
- Public health

Key points & open legal questions

- **Obtaining valid informed consent:** lack of clinical trials implies a lack of information about risks and other characteristics of the technology
- **Professional standards and obligations:** who should prescribe or administer mood HETs will determine professional standards and obligations. Relevant also for obtaining valid informed consent.
- **Qualification of mood data as data concerning health:** mood HETs enable the collection and processing of mood data, which are not explicitly recognised as data concerning health
- **Reinterpretation of notions of mental privacy, mental integrity, and freedom of expression:** due to the novel risks to privacy posed by mood HETs
- **Addressing the distinction between therapy and enhancement:** implications for defining the application of MDR and manufacturer's obligations, safety and prevention of harm requirements, obtaining centralised marketing authorisation, but also for defining HETs.

- Definition of HE
- Normality
- Therapy vs enhancement

- Religious, ideological, rhetorical arguments
- Ethics
- Law vs ethics

- Regulation of HETs
- UN
- CoE
- EU





Thank you for your attention!

Erik Kamenjašević, LL.M.
Researcher & doctoral candidate – KU Leuven

erik.kamenjasevic@kuleuven.be

 [e_kamenjasevic](#)