ICAIL 2023 - ODR E-Justice Scheme Workshop Ethical Benchmarks for ODR Scheme

Yashar Saghai, Hans Micklitz, and Federica Casarosa

Philosophy Section, University of Twente, The Netherlands

EUI, Italy

Funded by the European Union

www.eui.eu

Introduction

Part I: Ethical AI and the EU Digital Policy Legislation

- Where are we standing?
- What is the legal frame in which we are operating, with its opportunities and its deficiencies?

Part II –Questions, Problems and Approaches to Future ODR Ethical Benchmarks

- what are the problems that we have to overcome in our project to comply with ethical benchmarks that go beyond the rather poor and under scrutinized European legal framework?
- What is our proposed approach?

Co-funded by the Erasmus+ Programme of the European Union

EUI ROBERT SCHUMAN CENTRE

Part I: Ethical AI and the EU Digital Policy Legislation

Co-funded by the Erasmus+ Programme of the European Union

Out-of-court dispute settlement mechanisms

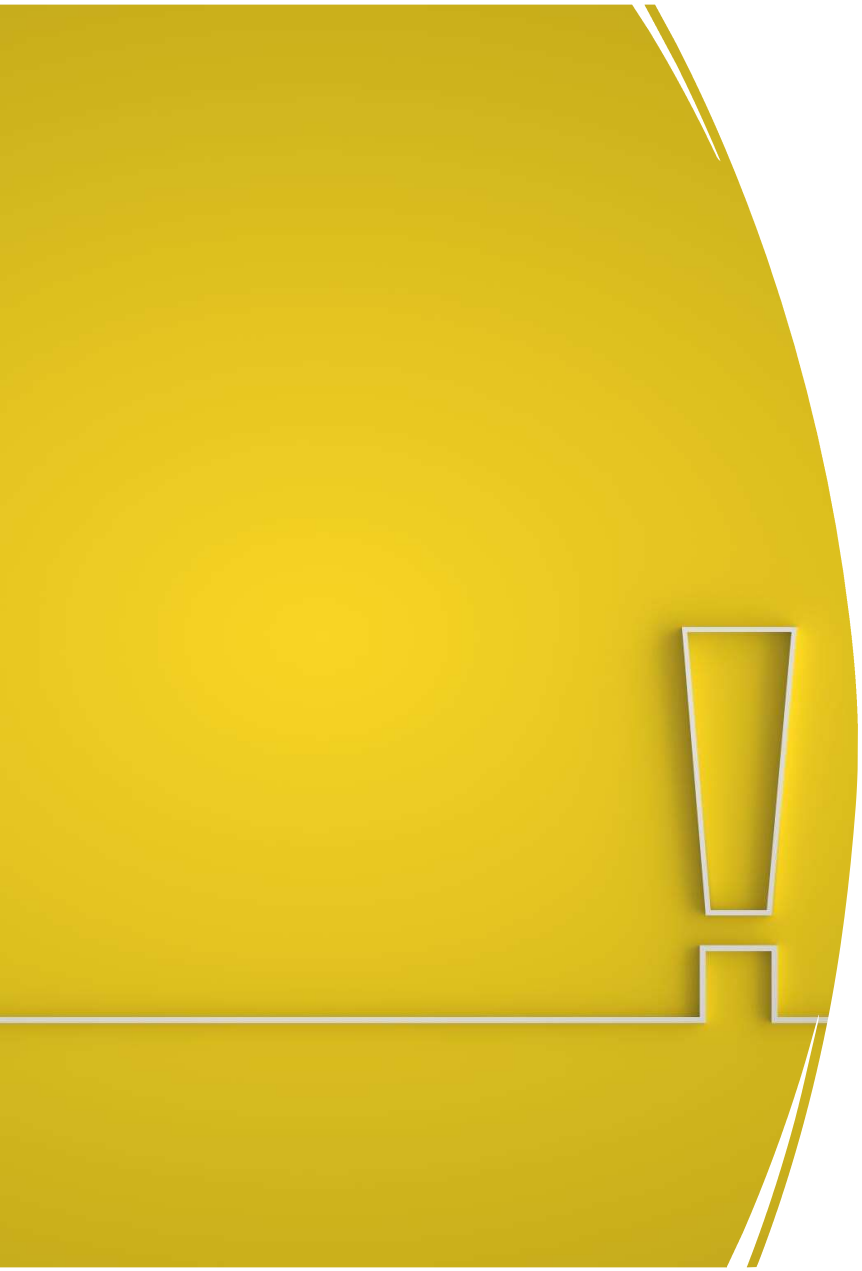
- directive 2000/31/EC on E-commerce
  - art. 17 dedicated to out-of-court dispute settlement: in case of disagreement between an information society service provider and the recipient of the service an obligation on the Member State so as not to adopt legislation that "hamper the use of out-of-court schemes, available under national law, for dispute settlement, including appropriate electronic means"
- Directive on consumer ADR 2013/11/EU
- Regulation on consumer ODR n. 524/2013
- Regulation (EU) 2019/1150 on promoting fairness and transparency for business users of online intermediation services
  - Art. 13 "*providers of online intermediation services and organisations and associations representing them to, individually or jointly, set up one or more organisations providing mediation services […] for the specific purpose of facilitating the out-of-court settlement of disputes with business users arising in relation to the provision of those services.*"

Co-funded by the Erasmus+ Programme of the European Union

Out-of-court dispute settlement mechanisms

- Directive 2019/790 on copyright and related rights in the Digital Single Market
  - Article 17(9) specifies that online content-sharing service providers shall provide an effective and expeditious complaint and redress mechanism, which is qualified as an out-of-court redress mechanism in cases of disputes between rightholders asking for content removal and platforms.
- Directive 2018/1808 amending the Audiovisual Media Services Directive
  - Article 28b provides for out-of-court redress for the settlement of disputes between users and video-sharing platform providers.
- Digital Services Act
  - Article 21(3) identifies a set of due process guarantees that the out-of-court dispute resolution provider should ensure in order to be certified, including impartiality and independence, respect of fair trial guarantees, expertise, absence of conflict of interest, accessibility as well as cost-effectiveness.

Co-funded by the Erasmus+ Programme of the European Union



Protection of vulnerable groups

- Regulation on consumer ODR n. 524/2013
  - art. 5 (1) provides: “The Commission shall develop the ODR platform and be responsible for its operation, including all the translation functions necessary for the purpose of this Regulation, its maintenance, funding and data security. The ODR platform shall be user-friendly. The development, operation and maintenance of the ODR platform shall ensure that the privacy of its users is respected from the design stage (‘privacy by design’) and that the ODR platform is accessible and usable by all, including vulnerable users (‘design for all’), as far as possible.”
- Digital Services Act
  - Accessibility as one of the principles to be certified

Co-funded by the Erasmus+ Programme of the European Union

Four dimensions of ethics

- meta-ethics;
- normative ethics (determining a moral course of action by examining the standards for right and wrong action);
- descriptive ethics (empirical investigation of people's moral behaviour and beliefs);
- applied/practical ethics;

Co-funded by the Erasmus+ Programme of the European Union

High Level Expert Group Guidelines on Trustworthy AI

- *Trustworthy AI should be*
- ***lawful**, ensuring compliance with all applicable laws and regulations,*
- ***ethical**, demonstrating respect for, and ensure adherence to, ethical principles and values*
- ***robust**, both from a technical and social perspective, since, even with good intentions, AI systems can cause unintentional harm.*
- ***inclusive and comprise all processes and actors** that are part of the system's **life cycle**.*

Co-funded by the Erasmus+ Programme of the European Union

European Ethical Charter on the use of AI in judicial systems and their environment

- **Principle of respect for fundamental rights**: ensure that the design and implementation of artificial intelligence tools and services are compatible with fundamental rights
  - clear purposes
  - in full compliance with the fundamental rights guaranteed by the European Convention on Human Rights (ECHR)
  - personal data protection; right of access to the judge and the right to a fair trial; principles of the rule of law and judges' independence in their decision-making process.

Co-funded by the Erasmus+ Programme of the European Union

European Ethical Charter on the use of AI in judicial systems and their environment

- **Principle of non-discrimination**: specifically preventing the development or intensification of any discrimination between individuals or groups of individuals
  - the methods do not reproduce or aggravate such discrimination and that they do not lead to deterministic analyses or uses
  - processing is not directly or indirectly based on “sensitive” data
- **Principle of quality and security**: with regard to the processing of judicial decisions and data, one should use certified sources and intangible data with models conceived in a multi-disciplinary manner, in a secure technological environment
  - Forming mixed project teams
  - Existing ethical safeguards should be constantly shared by these project teams
  - certified sources, traceable process to ensure that no modification has occurred to alter the content or meaning of the decision being processed

Co-funded by the Erasmus+ Programme of the European Union

European Ethical Charter on the use of AI in judicial systems and their environment

- **Principle of transparency, impartiality, and fairness**: making data processing methods accessible and understandable, authorise external audits
  - balance must be struck between the intellectual property of certain processing methods and the need for transparency (access to the design process), impartiality (absence of bias), fairness and intellectual integrity (prioritising the interests of justice)
  - technical transparency – the system could also be explained in clear and familiar language (to describe how results are produced) by communicating, for example, the nature of the services offered, the tools that have been developed, performance and the risks of error
  - certifying and external auditing processing methods
- **'Under user control' principle**: preclude a prescriptive approach and ensure that users are informed actors and in control of their choices
  - User autonomy
  - Professionals in the justice system should, at any moment, be able to review judicial decisions and the data used to produce a result and continue not to be necessarily bound by it in the light of the specific features of that particular case
  - The user must be informed in clear and understandable language
- *Note that the Charter is under process of operationalisation by the CEPEJ*

Co-funded by the Erasmus+ Programme of the European Union

EU Digital Policy Legislation (AIA etc)

- Mantra of human-centric, secure, trustworthy and ethical AI
- Emphasis on normative ethics – on law and the role of law
- Neglect of descriptive and normative applied ethics – use cases (fields in which the AI system will potentially be applied)
- De facto substitution of human-centric, secure, trustworthy and ethical AI through fundamental rights
- Delegation of the elaboration of AI standards to the European standardisation bodies
- Setting aside of civil society in the elaboration process

Co-funded by the Erasmus+ Programme of the European Union

Premises of EU Digital Policy Legislation

Reliance on law and technical standards to ensure that AI systems comply with fundamental rights

Conviction that law and regulations suffices to build trust in the society

Co-funded by the Erasmus+ Programme of the European Union

Part II: Questions, Problems and Approaches Regarding Future ODR Ethical Benchmarks

- Accessibility (A4J)
- Accountability
- Competence
- Confidentiality
- Efficient enforcement
- Empowerment
- Equality
- Explainability
- Fairness
- Honesty
- Impartiality
- Inclusion
- Informed participation
- Innovation
- Integration
- Legal obligation
- Neutrality
- Non- coercion
- Non-deception
- Non-discrimination
- Non-manipulation
- Protection from harm
- Respect for fundamental rights
- Security
- Transparency

+ Ethics of AI principles and digital ethics

(see Wing 2016 and https://odr.info/standards/)

Co-funded by the Erasmus+ Programme of the European Union

Four Problems

1. From abstract ethical and legal norms to operationalized standards and guidelines in design and practice
2. From abstract ethical and legal norms to context-sensitive practice (human-tech interactions in specific socio-technical systems within a particular field)
3. The need for use cases and the impossibility of predicting the future of emerging technologies (e.g., generative AI in ODR)
4. High-level commissions versus the need for the involvement of citizens and local communities in potential application

Co-funded by the Erasmus+ Programme of the European Union

Proposal:
A Three-Layered Approach for the Ethics of ODR

1. **Anticipatory Technology Ethics Combined with Ethics-by-Design Approach (intra-project)**
2. **Experimental Ethics Approach (post-project)**
3. **Ethical Certification (post-project)**

Co-funded by the Erasmus+ Programme of the European Union

Anticipatory Technology Ethics combined with Ethical-by-Design Approach (intra-project)

- **Anticipatory Tech Ethics (Brey 2012, 2017):**
  - **Three levels of analysis:**
    - **Technology:**
      - generic ethical issues related to inherent features of technologies
    - **Artifacts, systems, procedures**:
      - ethical issues that might be always present because of 1. the inherent features of the artifact; 2. unavoidable consequences of all or most uses; 3. high potential for problematic application
    - **Application level :**
      - interaction between the artifact and contextual elements, users, specific purposes
      - no forecast or prediction possible but possibility of anticipation of plausible future applications
  - **Combination of foresight techniques (e.g., scenarios, Delphi, wild Cards) and ethical analysis**

Application of ATE to ODR Scheme

- ODR scheme situated at the second level (systems, procedures) but its telos is in the third level (increased difficulty for *anticipation of contexts of application):*
  - Selection of two contexts (e-commerce and healthcare) for ethical annotations on BPMN because
- Using tools from choice architecture ethics to organize and present information and options to developers and users (adding or eliminating options, defaults, forced choice, nudges)
- Developing prospective use cases that are in-between traditional use cases and futures studies' scenarios — are there risks for ODR scheme or ODR developers to disrespect ethical and legal norms and best practices?
- Participatory process limited to professionals and practitioners because of the lack of accessible prototype

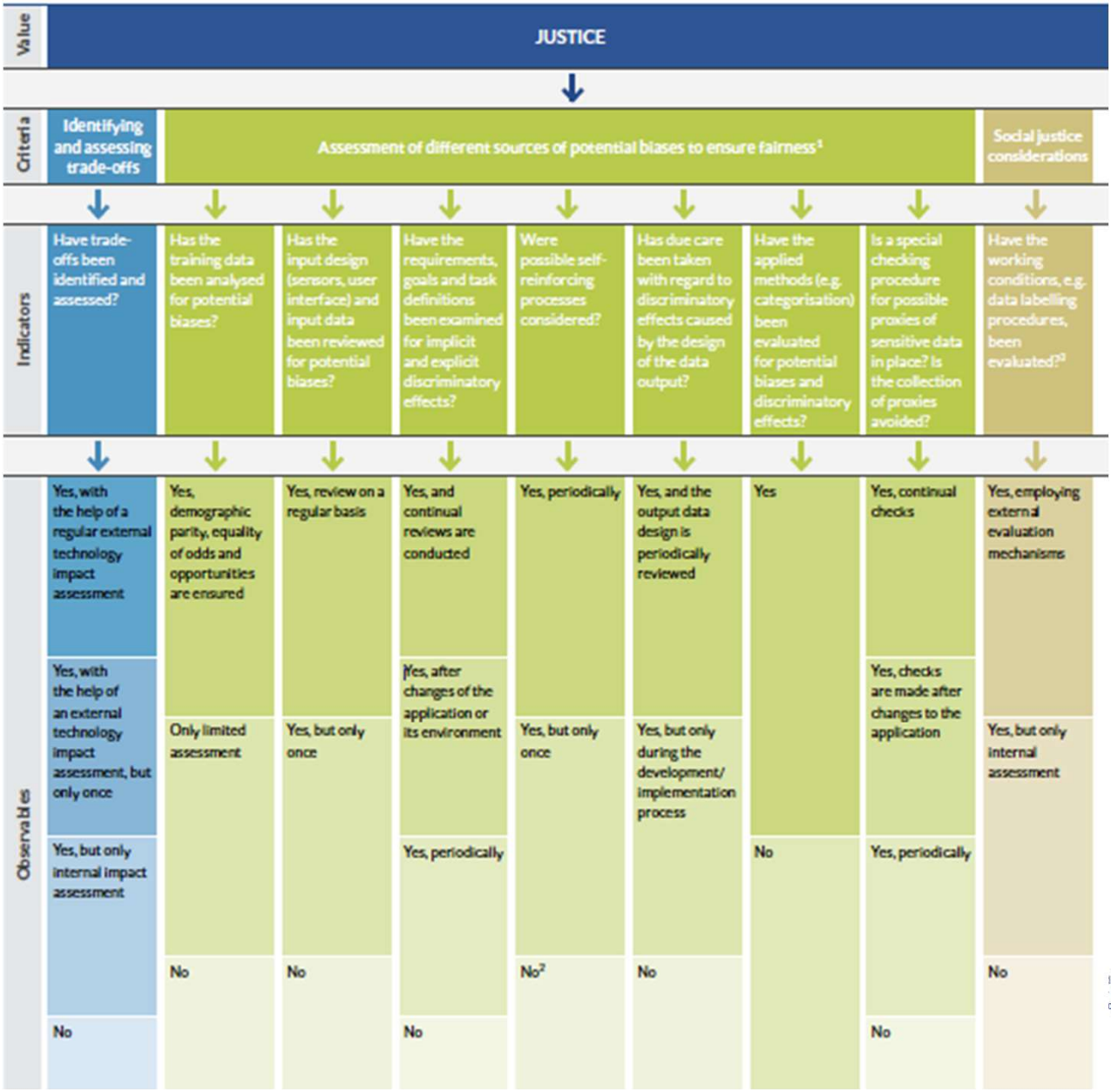
Co-funded by the Erasmus+ Programme of the European Union

Ethics by Design Approach (SIENNA project)

- "an approach that aims to incorporate ethical considerations into every stage of a technology's life cycle, from its design to its development and implementation, in order to (prevent or) mitigate possible negative ethical consequences produced by the technology" (Jansen et al 2021)
- Close to Value-sensitive-Design but developed specifically for AI
- Not limited to the protection of fundamental rights and mitigating harms: promotion the good (e.g., inclusiveness, sustainability, civility)
- Process of *specification of values*, starting with the most abstract and fundamental values, and working towards ethical requisits and specific ethical guidelines for different methodologies and processes.
- Main take:
  - Invention of new requisits and guidelines or use of existing ones (e.g, guidelines for inclusive information for people with different abilities)
  - If possible, creation of indices and metrics (e.g, VICIO model)

Co-funded by the Erasmus+ Programme of the European Union

Values Criteria
Indicators
Observables VCIO
Model for AI
(VDE 2020)




ROBERT
SCHUMAN
CENTRE
Questions and Comments
y.saghai@utwente.nl
www.eui.eu
Co-funded by the
Erasmus+ Programme
of the European Union