

ODIHR – NHRI Academy

2022: AI and Human Rights

20-24 June 2022, Tirana, Albania

Day 2 – Session 5

Human Rights Impact Assessment on AI:

Introduction and Existing Models

Trainers: Francesca Fanucci and Line Gamrath Rasmussen



European Center for
Not-for-Profit Law



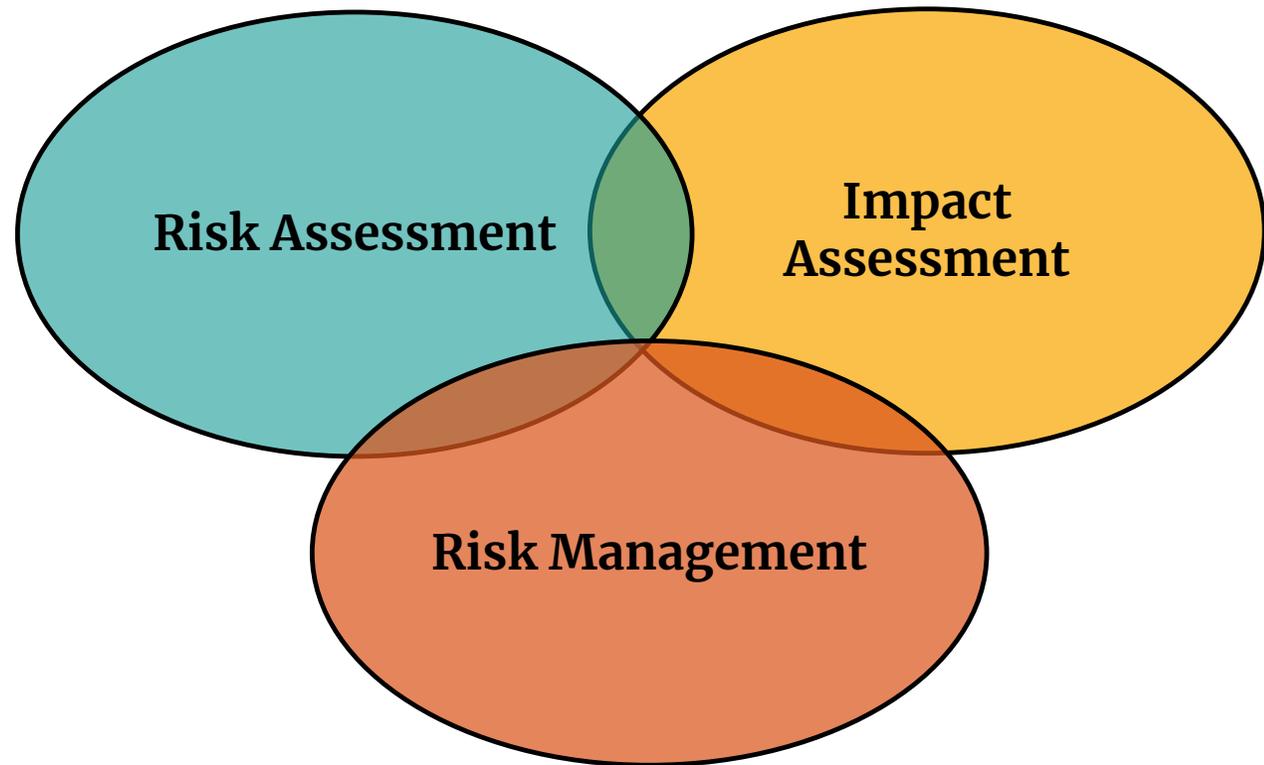
www.ecnl.org



@enablingNGOLaw

Three interconnected types of assessment

- **Risk Assessment:**
 - identification, analysis and evaluation of threats and vulnerabilities;
- **Impact Assessment:**
 - considers implications, both positive and negative, for people and their environment;
- **Risk Management:**
 - Identification of resources to minimise the likelihood of harms and negative implications;



Risk-Based Approach vs Rights-Based Approach

Risk-Based Approach:

- ✓ First, identifies a situation and/or sector that is likely to pose a potential threat/risk to a specific interest (not necessarily to human rights)
- ✓ Then, assesses and determines the scope and scale of the threat/risk
- ✓ Risk level determines applicable obligations

Rights-Based Approach:

- ✓ First, identifies series of rights potentially impacted, regardless of level of impact/risk
- ✓ Then, assesses level of impact/risk to such rights (scope, scale...)
- ✓ Potential additional obligations applicable depending on level



Example of Risk-Based Approach



European Commission – proposal for an EU AI Act:

Two cumulative criteria to preliminary identify situations with potential threat (“high risk”):

1. Sector in which AI systems is developed/deployed
2. Intended use:
 - ✓ “used in such a manner that significant risks are likely to arise” (EU White Paper on AI);
 - ✓ “significant risks to the health and safety or fundamental rights of persons” (AI Act)

If AI systems is high-risk = higher regulatory threshold for protection



Example of Rights-Based Approach

General Data Protection Regulation (GDPR):

- ✓ **Chapter 3:** Identifies rights granted to data subjects (that need to be respected regardless of level of risks incurred through data processing), e.g.:

Transparency

Right of access

Right of erasure

Right of objection

Right to data portability

Right to be forgotten

Right of rectification

- ✓ **Article 35:** where type of processing is likely to pose high risks to those rights = impact assessment



Example of Rights-Based Approach:

✓ **Article 35:** “Where a type of processing **in particular using new technologies**, and **taking into account the nature, scope, context and purposes of the processing**, is likely to result in a **high risk to the rights and freedoms** of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data”

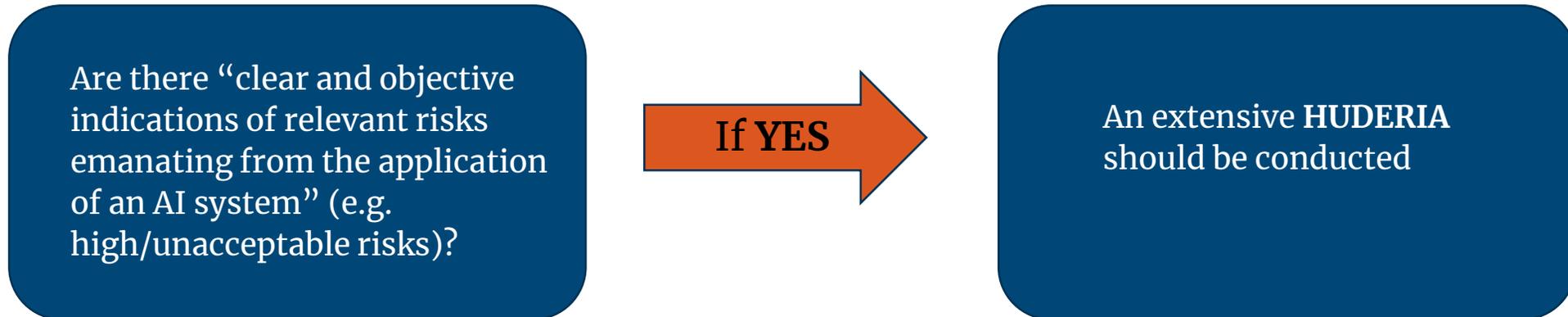
⇒ **DPIA minimum criteria:**

- Description of the processing operations, their purpose and legitimate interest of data controller;
- Necessity and proportionality;
- **Risks to rights and freedoms of data subjects affected;**
- Risk mitigation measures (risk management)



A “hybrid” model? CoE Ad hoc Committee on Artificial Intelligence (CAHAI) HUDERIA (Human Rights, Democracy and Rule of Law Impact Assessment)

- ✓ All AI systems should undergo an initial review (“risk assessment”) to assess “potential risk on the enjoyment of human rights, the functioning of democracy and the observance of the rule of law” & determine **risk classification** (e.g. low, high, unacceptable).



The extensive CoE HUDERIA should follow at least for main steps:

1. **Risk Identification** (relevant risks for human rights, rule of law and democracy)
2. **Impact Assessment** (likelihood and severity of the effects on those rights and principles)
3. **Governance Assessment** (Roles and responsibilities of duty-bearers, right holders and stakeholders in implementing and governing the mechanisms to mitigate the impact)
4. **Mitigation and Evaluation** (Identification of suitable mitigation measures and ensuring a continuous evaluation)



- ✓ context and purpose;
- ✓ level of autonomy of the AI system;
- ✓ underlying technology of the AI system, usage of the AI system (both intended and potentially unintended use);
- ✓ complexity of the AI system (e.g., part of multiple deep neural networks/building on other AI systems);
- ✓ transparency and explainability of the system and the way it is used;
- ✓ human oversight and control mechanisms for the AI provider and AI user;
- ✓ data quality;
- ✓ system robustness/security;
- ✓ involvement of vulnerable persons or groups;
- ✓ scale on which the system is used;
- ✓ geographical and temporal scope;
- ✓ reversibility of harm



Three experimental HRIA (pilot) models:

HRIA Model and Methodology

Polytechnic University of
Turin (Italy)

Impact Assessment Human Rights and Algorithms (IAMA)

Utrecht University (The
Netherlands)

Human Rights Impact Assessment Guidance and Toolbox

Danish Institute of Human
Rights



HRIA Model and Methodology - Polytechnic University of Turin (Italy)

- ✓ Based on empirical analysis of cases and guidelines by DPAs in Europe
- ✓ Suitable for AI systems embedded in globally distributed products/services (e.g., AI virtual assistants, autonomous cars, recruiting AI-based software, Internet Of Things, etc.)

1. Planning and Scoping



2. Data Collection and Analysis of Impact on Human Rights

Description/analysis of type of product (including data flows and processing purposes)

Contextualisation within local human rights law/jurisprudence

Identification of relevant stakeholders

Risk identification

Risk likelihood

Risk severity



Questions for break-out groups:

- **Question 1:** Which stakeholders' rights/freedoms do you think are going to be/are likely to be affected and how?
- **Question 2:** Which criteria and what methodology would you use to determine your assessment?
- **Question 3:** Can the risks you identified (if any) be managed/mitigated – and if yes, how?



Sources/reading materials

- European Commission, *White Paper on Artificial Intelligence: a European approach to excellence and trust* (2020), https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf
- Proposal for a Regulation of the European Parliament and the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act), (2021), Titles II and III, <https://artificialintelligenceact.eu/the-act/>
- Council of Europe – Ad Hoc Committee on Artificial Intelligence (CAHAI), *Possible elements of a legal framework on artificial intelligence, based on the Council of Europe’s standards on human rights, democracy and the rule of law* (2021), Chapters V and XII, <https://rm.coe.int/possible-elements-of-a-legal-framework-on-artificial-intelligence/1680a5ae6b>
- A. Mantelero, M.S. Esposito, *An evidence-based methodology for human rights impact assessment (HRIA) in the development of AI data-intensive systems* (2021) <https://www.sciencedirect.com/science/article/pii/S0267364921000340?via%3Dihub#tbl0002>
- Impact Assessment for Human Rights in the Use of Algorithms (IAMA),(2021) <https://www.rijksoverheid.nl/documenten/rapporten/2021/02/25/impact-assessment-mensenrechten-en-algoritmes>
- Danish Institute of Human Rights, *Human rights impact assessment guidance and toolbox* (2020), <https://www.humanrights.dk/tools/human-rights-impact-assessment-guidance-toolbox>
- Data & Society, ECNL, *Recommendations for Assessing AI Impacts to Human Rights, Democracy, and the Rule of Law* (2021), <https://ecnl.org/sites/default/files/2021-11/HUDERIA%20paper%20ECNL%20and%20DataSociety.pdf>
- Data & Society, ECNL, *Mandating Human Rights Impacts Assessments in the AI Act* (2021), <https://datasociety.net/wp-content/uploads/2021/11/HUDIERA-5-Pager-FinalR1.pdf>
- Mantelero, A., *Beyond Data: Human Rights, Ethical and Social Impact Assessment in AI* (2022), <https://link.springer.com/book/10.1007/978-94-6265-531-7>

