

Motivation & Significance

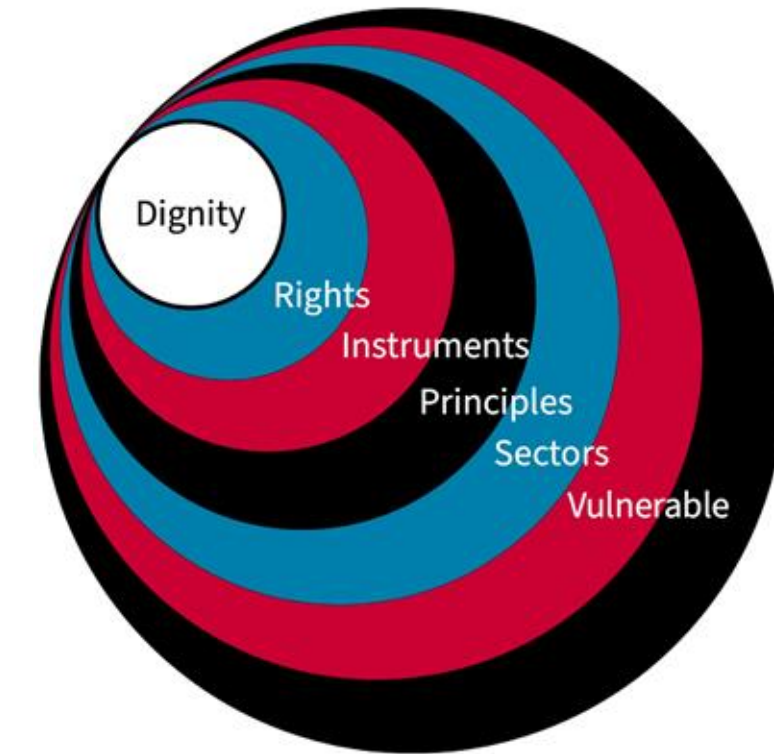
The *AI & Human Rights Index* is a global research collaborative focused on how artificial intelligence can violate or advance human rights. Editors from four continents are working with contributors worldwide to develop a comprehensive legal framework that serves as a guiding structure for evaluating AI's impact on human rights across societal sectors.

The *Index* draws on **eight decades of human rights law** to organize the relationships among **rights**, the **instruments** used to measure and enforce those rights, the **principles** that illuminate gaps in existing law, their application across distinct societal **sectors**, and AI's negative and positive impacts, especially on **vulnerable populations**.

The *Index* emphasizes both harm prevention (**nonmaleficence**) and measurable positive outcomes (**beneficence**), charging societal sectors with ensuring that AI benefits humanity and the environment.

Index researchers apply a **Solutions Scholarship** methodology to interrogate and critique AI while committing to identify measurable solutions for every problem analyzed, embedding solutions-focused research into each article.

Ultimately, the Index fosters **technical, cultural, and legal literacy** about how AI intersects with human rights through multimedia encyclopedia entries that are accessible to the public.



Step 1. Rights Mapping

In advancing this mission, the *Index* is organized around a series of steps, the first of which is a **Rights Mapping exercise**. In evaluating thousands of legal documents, we ask **which specific human rights are most at risk and which may be advanced by AI**. We also examine which **cultural contexts** and **regional approaches** to human rights are used to interpret these risks.

Peace

- Right to Peace
- Recognition as a Person Before the Law
- Freedom from Slavery and Servitude
- Freedom from Torture and Inhumane or Degrading Treatment
- Right to Prevention of and Protection from Genocide
- Right to Security of Person, Protection by the State

Justice

- Right to Life, Liberty, and Personal Security
- Right to Legal Recognition
- Right to Equality before the Law
- Right to a Fair Trial and Due Process
- Right to Presumption of Innocence until Proven Guilty
- Right to Justice and Accountability
- Right to Truth, Redress, and Information
- Right to Protection from Persecution
- Right to Equality and Non-Discrimination
- Right to Remedy

Civic Engagement

- Freedom of Assembly and Association
- Right to Free Elections
- Right to Participation in Public and Political Life
- Right to Participate in Cultural Life, Arts, and Science
- Right to Personal Mobility

Privacy

- Right to Data Protection and Freedom from Surveillance
- Right to Control over Personal Data and Data Sovereignty
- Right to Intellectual Property and Protection of Personal Creations
- Right to Mental and Biological Privacy
- Right to Informed Consent in Data Collection and AI Interactions
- Right to Protection from AI Manipulation and Misinformation

Movement

- Freedom of Movement and Residence
- Right to Seek Asylum from Persecution
- Right to a Nationality

Family

- Right to Marry
- Right to Form a Family
- Right to Family Benefits
- Right to Family Unity
- Right to Own Property

Human Agency & Expression

- Freedom of Thought, Conscience, and Religion
- Right to Language
- Freedom of Opinion and Expression
- Right to Self-Determination
- Right to Live Independently and in Community
- Right to Digital Identity

Education

- Right to Education
- Right to Higher Education
- Right to Non-Discrimination in Education
- Right to Cultural and Linguistic Education
- Right to Education in Emergencies

Work

- Right to Work
- Equal Opportunities and Treatment
- Right to Remuneration
- Right to Form and Join Trade Unions
- Right to Rest and Leisure
- Right to Protection against Unemployment
- Right to Retraining and Skill Development in Response to Technological Change
- Right to a Standard of Living Adequate for Health and Well-being
- Right to Food, Clothing, and Housing
- Right to Social Security
- Right to Development
- Corporate Responsibility to Respect Human Rights
- Human Rights Due Diligence

Health

- Right to Health and Medical Care
- Right to Environmental Health
- Right to Water and Sanitation
- Right to Mental Health
- Right to Protection from Harmful Effects of AI in Healthcare

Environmental Rights

- Right to a Healthy Environment
- Right to Sustainable Development
- Right to Protection from Environmental Harm due to Technological Advancements

Digital Rights & Technology

- Right to Access to Science and Technology
- Right to Internet Access
- Right to Ethical AI and Emerging Technologies
- Digital Rights and Online Freedoms
- Right to Algorithmic Transparency and Accountability
- Right to Protection from Cyber Threats and Cybersecurity
- Right to Digital Self-Determination
- Right to Fair and Equitable Access to AI Benefits

Step 2. Instrument Mapping

Our next question focuses on how to use established **human rights instruments** to measure technology's impact at every stage of the AI lifecycle, from development to deployment to monitoring. *Which legal instruments should be used to determine whether an AI system violates or protects a human right?* To answer this question, we draw from UN documents:

- **Declarations**, such as the Universal Declaration of Human Rights and the UN Declaration on the Rights of Indigenous Peoples, which articulate principles or standards but are not legally binding;
- **Conventions**, such as the International Convention on the Elimination of All Forms of Racial Discrimination, the Convention on the Elimination of All Forms of Discrimination Against Women, and the Convention on the Rights of the Child, which offer formal, legally binding, and enforceable terms for ratifying member states;
- **Covenants**, such as the International Covenant on Civil and Political Rights and the International Covenant on Economic, Social and Cultural Rights, which are legally binding treaties upon ratification; and
- **Other instruments**, such as resolutions and guiding principles, including the UN Human Rights Council's Guiding Principles on Business and Human Rights.

We also draw upon multinational laws and reports adopted by the **African Union**, the **Association of Southeast Asian Nations**, the **European Union**, the **Organisation for Economic Co-operation and Development**, the **Organization of Islamic Cooperation**, and others. By using these varied instruments, we build the legal taxonomy needed to demonstrate that many human rights laws from the past eighty years apply to current and future technological advancements.

Step 3. Ethics Mapping

Step 4. Sector Mapping

To identify gaps in existing human rights law, we next conduct an **Ethics Mapping exercise** to determine the primary principles used in the responsible technology movement. To measure AI's impact accurately, we also need context. The ways AI may violate or uphold human rights differ across sectors of society. Accordingly, we conducted a **Sector Mapping exercise** by asking: *which sectors are responsible for ensuring that AI systems prevent harm and benefit humanity?*

Accountability	Human Agency	Safety
Beneficence	Nonmaleficence	Security
Consent	Notice	Sustainability
Dignity	Privacy	Transparency
Explainability	Reproducibility	Trust
Fairness	Remedy	Verifiability



As you can see, our ethics mapping and sector mapping work has enabled us to build an **interdependent research tool**. We understand each right, instrument, principle, and sector as inherently connected to the others. A wide range of combinations emerges when a specific right is paired with an instrument, then linked to a principle, and so forth. At the center, however, one element remains constant. The heart of this inquiry is **our prime directive**: to measure whether AI systems uphold the **inherent worth and dignity of every person**.

Step 5. Apply Vulnerability Lens

Our prime directive is particularly important when examining AI's impact on the most vulnerable among us.

We intentionally apply the UN's "vulnerability lens" and include a dedicated section of the *Index* to analyze the ways AI may affect populations that are already more susceptible to human rights abuses.

Identity Groups

- Women
- Indigenous
- Religious
- Racial/Ethnic
- LGBTQ+



At-Risk Populations

- Children, Aging, Poverty, Disabilities

Displaced Persons

- Refugees & Migrant Workers

Step 6. Build Protocol

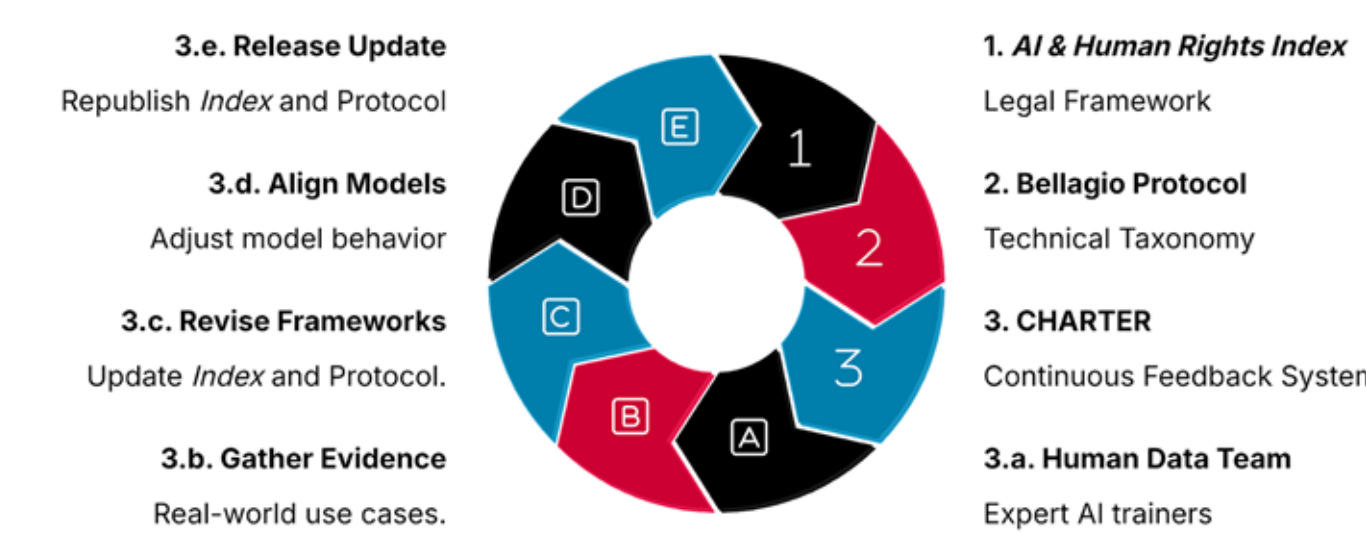
Step 7. Activate CHARTER

To recap, the *Index* draws on eight decades of human rights law to organize the relationships among **rights**, **instruments**, **principles**, **sectors**, and **vulnerable populations**. We are now translating this legal framework into a **technical taxonomy** called the **Bellagio Protocol**, named for the Rockefeller Foundation's Bellagio Center in Lake Como, Italy, where Dr. Nathan C. Walker serves as a resident researcher.

The Bellagio Protocol is a **semantic infrastructure** that uses SKOS and OWL to translate the Index into a structured technical system for evaluating how AI systems may both violate and advance human rights. SKOS enables us to organize concept schemes, label related ideas, identify semantic relationships, and prepare documentary notes. OWL allows us to articulate formal ontologies, define logical constraints, create inference-based class definitions, and specify property restrictions. We are careful not to overuse OWL, because human rights norms are continually contested and reinterpreted.

We then activate CHARTER: Classifying Human-Rights Advancements for Responsible Technology and Ethical Response, a continuous feedback system that serves as our **Governance Flywheel**. CHARTER begins with (a) Expert AI Trainers from around the world applying structured rubrics through Reinforcement Learning from Human Feedback. We then (b) gather real-world evidence from audits, NGO reports, academics, and cross-cultural feedback. This evidence is used to (c) revise the frameworks, (d) align the models accordingly, and (e) release updated versions of the *Index* and the Bellagio Protocol.

AI & Human Rights Governance Flywheel



Call to Action

And this is where you come in!

Do you have expertise in human rights law, ethics, or technology? Would you be interested in serving as a contributing author for the Index entries or glossary entries? Would you like to help develop the Bellagio Protocol or the CHARTER, or serve as an Expert AI Trainer on our Human Data Team? Learn more at

aieethicslab.rutgers.edu.