



AN OVERVIEW | A Proposal



# The **GLOBAL ACCORD** **ON AI TRANSPARENCY** **AND ACCOUNTABILITY**

and

the need for the United Nations Human Rights Council to  
Recognize Digital Agency and Understanding as a Human Right

THE BEACON INITIATIVE



STATEMENTS



Shoshana  
Rosenberg





# WHAT WE CAN AGREE ON/TO

With Explainability-by-Design as the horizon that I believe we must reach towards in earnest, this is my pass at cutting through the noise from many directions. **We are in the midst of the maturation phase of the digital (r)evolution we can all observe both hype and skepticism starting to peak- precisely in line with historical precedent and marking this moment as the very inflection point we are all poised to remember it as.** True industrial revolutions reshape society and, at this juncture, clarity and action are particularly powerful tools.

As we shift to the next phase, marked by the further integration and harmonization of a broad scope of AI technologies, the stakes will change fundamentally, and we must set the frameworks and foundations in place to ensure transparency, enforce accountability, and foster explainability. The EU AI Act and the regulations currently being sculpted around the world are vital, but our collective duty, acumen, and insight should demand we take aim to go further, advocating for the recognition of digital self-determination as a fundamental human right and crafting a global framework that guides us toward AI designed for explainability and environmental sustainability.

**Yes, we have very solid legal frameworks in place for addressing privacy, discrimination, and consumer protection, but the advent of AI introduces challenges in algorithmic transparency, ethical decision making, and safeguarding against unintended bias and manipulation that extend beyond these established domains- as well as complexities in real world applications, where the convergence of digital decisions and physical actions can lead to tangible consequences.** If we look ahead to the tapestry of regulations to come, the imperative of explainability stands out as a defining feature of the emerging landscape. I propose a dual-pronged approach to cultivate consensus around fundamental standards and key directives that can set the foundation for the principle of (and mandate for) explainability by design within future AI technologies.

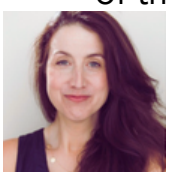
The first component of this approach derives from the pressing need to advocate for the recognition of **Digital Agency and Understanding** (or equivalent) as a human right by the United Nations Human Rights Council (UNHRC). This recognition is imperative in an era where digital technologies, and AI in particular, will play a significant role in shaping the fabric of human society. Digital Agency and Understanding would encompass not only the right of individuals to understand and influence the algorithmic decisions that impact their lives, but their empowerment to engage with digital technologies on equitable terms and their ability to navigate, participate in and influence the digital ecosystem that engages with their information. Acknowledging this right at an international level would not only protect individuals, but also help guide the development of AI towards more explainable, ethical, and equitable outcomes.

Simultaneously, the second component requires the establishment of a **Global Accord on AI Transparency and Accountability**. This accord would focus exclusively on these foundational aspects, delineating a universal standard and providing a cohesive framework for the ethical design, development and application of AI technologies to come without infringing on the regulations that will differ amongst nations. This accord would catalyze a focused initiative to guarantee that AI systems are designed and developed with intrinsic explainability and are held to accountability standards that are both rigorous and clear.

**The dual initiatives I am proposing are not merely reactionary measures to the current state of AI development (or hysteria) but are intended as proactive strategies, designed to anticipate and mitigate the issues related to the huge incentives and race to continue to design and build AI technologies without explainability as a foundational component.** The Global Accord on AI Transparency and Accountability would serve as the document that outlines the principles of ethical AI development, while the recognition of Digital Agency and Understanding as a human right by the UNHRC provides the legal and moral framework necessary to protect individual autonomy in the digital age.

This moment in time requires a nuanced understanding of both the AI technologies themselves and the potential they have to impact human rights and societal norms. By establishing a Global Accord on AI Transparency and Accountability, and by advocating for the recognition of Digital Agency and Understanding as a human right, we can lay the groundwork for a future in which AI technologies are conceived, developed and deployed in a manner that is built to live in to the requirements of the cascade of forthcoming regulations.

**Can one just go around proposing new Human Rights?  
HASN'T THIS BEEN SORTED ALREADY?**



**Shoshana  
Rosenberg**

# HERE BE DRAGONS/ HIC SUNT DRACONES



We don't need to enshrine a right to be free of interference by dragons until there are dragons.  
*And then, we do.*

Human rights are inherent to all individuals, regardless of any external threat or encroachment, but the recognition and articulation of specific human rights often emerge in response to threats or violations. The time is now to formally recognize and establish protections necessary to preserve our inherent human rights to digital agency and understanding. This right, now challenged by the intricacies of the digital (r)evolution, highlights the necessity for individuals to have clarity and controls over the digital constructs influencing their lives. As an international community, our immediate action and advocacy is required to integrate these protections into the global human rights framework as a key step in shaping the way that AI technologies are designed and developed going forward.

In the evolving landscape of global digital rights, the Human Rights Council of the United Nations played a pivotal role in extending the recognition of conventional human rights into the digital domain, reflecting an understanding that the liberties and safeguards afforded to individuals offline must be equally respected and protected online. The United Nations' resolution on the promotion, protection, and enjoyment of human rights on the Internet has been instrumental in affirming that access to information, freedom of expression, and privacy are fundamental rights that extend into the digital sphere.

Likewise, the Council of Europe's Convention 108, and its subsequent enhancements through Convention 108+, stands as a landmark legal instrument in the realm of data protection, directly influencing digital privacy rights across member states, with Convention 108+ specifically updates protections to address the challenges posed by technological advancements, emphasizing the need for all data processing to respect human dignity and rights in the digital environment.

Together, these frameworks serve as collective acknowledgment of the critical need to safeguard personal autonomy in the face of digitalization and established a solid foundation for digital privacy and data protection. Despite these significant advancements, AI technologies truly bring new challenges that extend beyond the current scope of digital privacy and data protection. AI's capacity to influence decisions, impact societal norms, and impact individuals and individual autonomy necessitates a broader and more nuanced consideration of digital rights.

The autonomous aspects of AI, coupled with its potential for widespread use and impact, bring into sharper relief the limitations of existing frameworks to fully address the complexities introduced by these technologies and, to my mind, the need to advocate passionately for the recognition of **Digital Agency and Understanding** as a distinct human right.

While the foundational digital rights principles laid out by the UN Human Rights Council and Convention 108+ are indispensable, the unique challenges derived from AI call for a focused response and the development of new principles or rights to specifically address the gaps in the current framework.

**Digital Agency and Understanding** aims to specifically address the rights of individuals to understand, interact with, and influence the algorithmic systems that increasingly govern aspects of their lives. This proposed right acknowledges the intricate interplay between technology and autonomy, advocating for a future where individuals retain control and understanding over the digital and AI-driven decisions that affect them. Leveraging the ample groundwork established by existing human rights frameworks, the call for Digital Agency and Understanding as a new human right is both an evolution and an extension of these principles into the digital age. The recognition of this by the UNHRC would reflect a proactive step towards ensuring that advancements in AI and digital technologies are developed and deployed in a manner that respects, protects, and enhances human dignity and rights.



# ACCORDING TO CHANGE

## The need for a Global Accord on AI Transparency and Accountability



### THE EXPLAINABILITY AGENDA

This moment presents a unique opportunity to shape the trajectory of AI technology development with the establishment of an agreed-to framework that will guide AI's evolution, focusing on shared principles rather than entangling ourselves in the minutiae of regulatory specifics. This approach would enable us to lay a global foundation that encourages and supports ethical innovation and responsible use, without presupposing the outcome of or agreement to every other local legislative detail. I hope to set out my vision for such an Accord as a collective endeavor to align the future of AI advancements with a set of indisputable shared objectives.

The **Global Accord on AI Transparency and Accountability** is strategically framed to sidestep the quagmire of exhaustive negotiations, offering a more agile and principle-based platform from which AI can evolve. By prioritizing these core values and setting the definitive goals, we ensure that the Accord can serve as a blueprint that will guide and foster the development of future AI and innovation in a manner that is ethically grounded.

#### UNIFIED GLOBAL STANDARDS

The Accord will establish a unified set of global standards for AI that prioritize transparency, accountability, fairness, adaptability, environmental sustainability and explainability by design. With the goal of fostering global consensus and minimizing contention, the Accord should respect and align with prevailing norms and mandates regarding privacy, security, non-discrimination and other protections, carefully calibrated to compliment rather than supplant existing local laws.

#### EXPLAINABILITY BY DESIGN

Central to the aims of the Accord is the mandate for both **minimum viable explainability** in existing systems, and **explainability by design** (by this or any other name) in all AI systems or layers developed beyond a specified date. This goal embeds the principle of explainability by design within global AI development practices and acknowledges the current challenges where the quest for efficiency sidelines the needed imperative for the path to outputs to be understandable.

This forces the needed paradigm shift in how AI systems are conceived and built, ensuring the paths to their outputs and decisions are made understandable to users regardless of their level of technical expertise.

#### MIGHTIER AND LIGHTER TECHNOLOGY

Recognizing that explainability can burden AI systems, slowing them down or demanding more resources, the Accord should commit to encourage, support and incentivize AI technology that is both more powerful and efficient, capable of offering explainability without compromising on speed or environmental sustainability.

#### ENVIRONMENTAL IMPACT

The Accord must place strong emphasis on minimizing the environmental footprint of existing and future AI technologies. It will advocate for sustainable development practices that ensure that AI's growth does not come at the expense of the planet. This commitment must be integral to the Accord's vision, ensuring that the pursuit of advanced AI is harmonious with ecological conservation efforts.

#### UNIVERSAL TAGGING PROTOCOLS

Another foundational component of the Accord should be the introduction of universal code tagging protocols for AI development. This aspect serves a critical role in allowing AI to be consistently built for explainability, transparency, and auditability. By embedding universal tag protocols and an index within the Accord's framework, it underscores the initiative's holistic approach to AI governance—balancing innovation with ethical considerations and user empowerment. *(The Accord would not include the Beacon Initiative Universal Tagging System, but the proposed index and protocols would set the foundational requirements to allow for consistency across all such initiatives and tools.)*

#### GLOBAL COLLABORATION AND RESOURCE SHARING

The Accord should promote an ethos of global collaboration, advocating for the sharing of knowledge, best practices, and resources to overcome the digital divide. This would underscore the importance of equitable AI development, ensuring that all nations have the capacity to contribute to and benefit from ethical AI advancements.

#### ADAPTABILITY TO FUTURE CHALLENGES

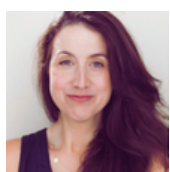
As with all else related to AI technologies, the Accord would need to include provisions for its periodic review and update. This adaptability ensures that it remains responsive to emerging technological trends, ethical dilemmas, and societal needs.

#### EMPOWERING PUBLIC ENGAGEMENT AND LITERACY

The Accord should set out and envision a comprehensive strategy for public education and engagement on AI ethics and technology. We must ensure that it takes aim to cultivate an informed global citizenry capable of participating in ethical AI discourse, reinforcing agency, understanding, and the societal foundation needed to support and demand responsible AI.

#### CALL TO ACTION

At its core, the Global Accord on AI Transparency and Accountability would be a commitment to collective action and shared responsibility. It is a key step to a future where AI is developed and deployed with a profound commitment to ethical integrity, global inclusiveness, and environmental stewardship. Through this Accord, nations could collaboratively shape an AI-powered future that aligns with our highest human values and aspirations.



**A.I., Privacy and DEI** have a high level of interdependence and interconnectedness and are continuously evolving.



All three are tied directly to ethics, fundamental human rights, the future of work, and decision making and bias, which means:

**YOU AREN'T ON THE SIDELINES  
OF THESE THINGS.**

**YOU ARE CRUCIAL TO THEM  
BEING WHAT THEY SHOULD.**

Do you want to know more?



STATEMENTS



Shoshana  
Rosenberg



# THE WAY I SEE IT



## DIVERSITY, EQUITY, AND AI ARE THE FUTURE

AI will accelerate the interconnectivity of the world and will be deeply ingrained in all kinds of decision-making processes.

DEI is essential for sustainable progress, innovation, and harmony.

Embracing DEI in the AI era is the only viable path to ensure that AI does not exacerbate or perpetuate existing biases and inequity.

## INCLUSION IS THE KEY TO DIVERSITY AND EQUITY

Diversity will not be sustainable and equity not possible unless a full spectrum of the community is represented, integrated, accepted, respected and valued.

## PRIVACY IS THE KEY TO INCLUSION

Inclusion cannot be measured, refined or fostered effectively without candid feedback and diversity data, both of which put individuals **at risk** without true privacy and anonymity, and both of which are too often collected without the preservation of privacy rights.

## DATA IS THE KEY TO AI

To ensure ethical AI, the data it is trained on must be diverse, representative, and gathered properly (with authorization and/or consent) and used responsibly.

## AI CAN BE AN UNPARALLELED KEY TO EQUITY

AI that is properly and thoughtfully designed with DEI principles can identify and help flag and rectify systemic disparities across any number of sectors and disciplines and processes, as well as helping to identifying gaps in the policies or tools that support them.



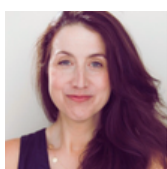
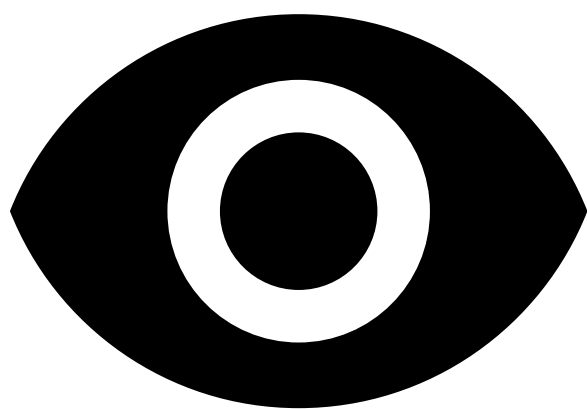
STATEMENTS



Shoshana  
Rosenberg

**TRANSPARENCY, ACCOUNTABILITY, FAIRNESS,  
AND TRUST ARE KEY TO ALL THREE**

# DEFENDING STATEMENTS PRIVACY



Shoshana  
Rosenberg