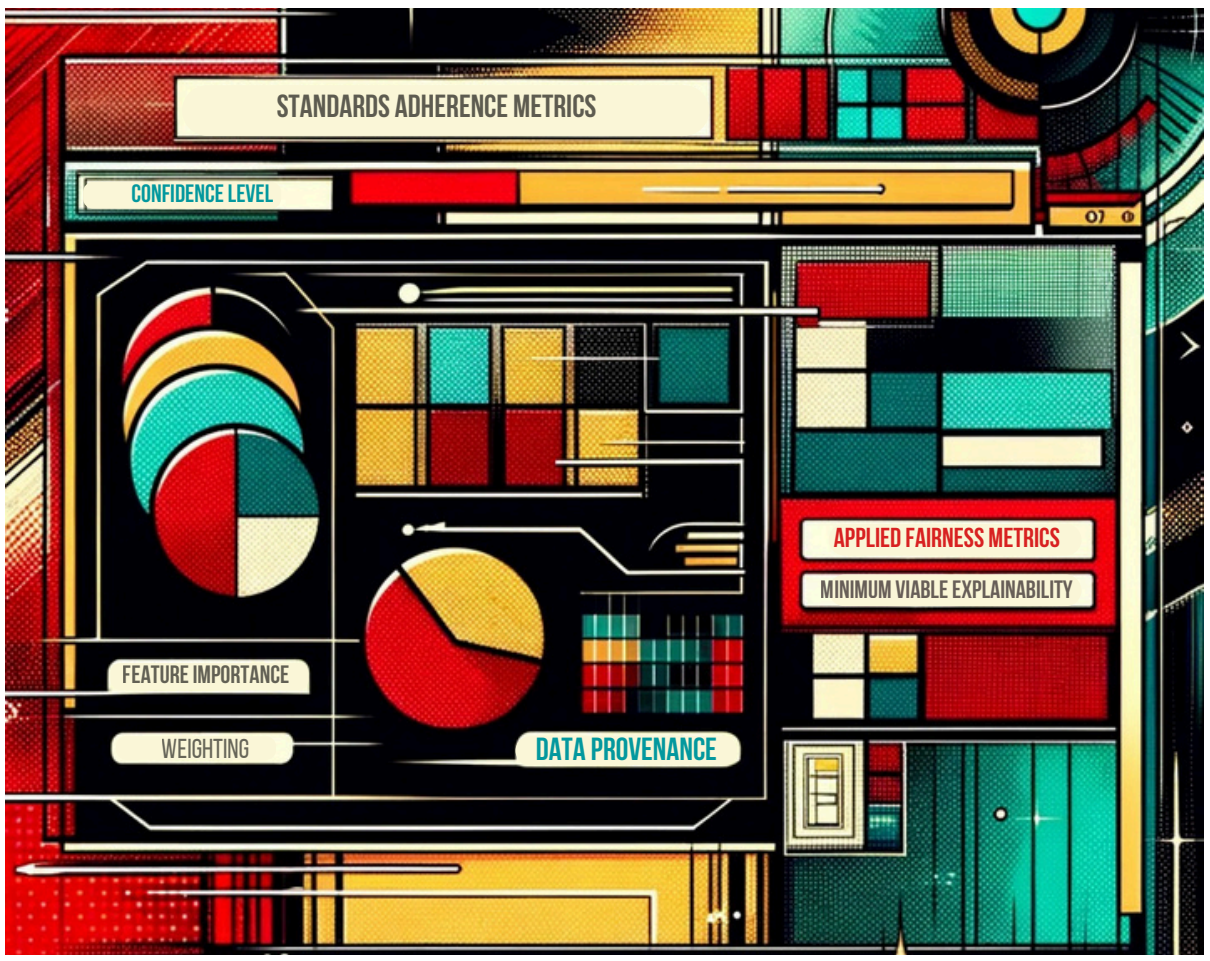
EXPLAINABILITY by DESIGN



EXPLAINABILITY ADDENDA AND NOTICES

AND WHY YOU NEED TWO IMPACT ASSESSMENTS

...AND A MICROPHONE.

Shoshana Rosenberg

A.I. DE STATEMENTS PRIVACY

IMPACT ASSESSMENTS

LET'S START HERE

The requisite AI Impact Assessment is not a subset of the Privacy Impact Assessment.
We know this, I just needed to set it out here so that we can get to the good stuff. And what if your organization is resistant to a distinct AI Impact Assessment? We will get into that too. In the end, your organization is always right- until the regulators say otherwise, but it would help if you had a microphone that wasn't just about potential fines .

PRIVACY IMPACT ASSESSMENT

The Privacy Impact Assessment is the proper tool for the proper job - of evaluating and mitigating risks to privacy. It is an in-depth undertaking that identifies and documents the processing activities and types of personal data collected; the necessity, justification, and proportionality of collecting and processing that data; and the analysis of, and safeguards and mitigations for, the privacy related risks and potential harms to individuals.

(It does more than that, in truth, and can make you coffee, but it is not the droid you are looking for to document the risk of harms that can arise from AI.)

Why does the PIA do all of these things?
It is focused (intentionally and fundamentally so) on the future handling, processing, and protection of personal data within, or tied to ,an organization's operations- which absolutely will include AI systems.

A PIA is a living document which will change alongside various factors (just as an AI Impact Assessment will need to), but it is sufficiently endowed and burdened with exactly the information and remit needed to do the worthy and important work that it does.

(Also, the focused legal relevance of a PIA allows organization to provide targeted evidence relevant to a privacy issue for a straightforward examination of the case at hand without complicating legal proceedings with unrelated information. Foreshadowing as to my thinking for those of you who think the AI Impact Assessment and the AI Risk Review should be one and the same.)

AI IMPACT ASSESSMENT

The AI Impact Assessment will need to specifically examine the implications of using AI technologies and will have a very broad scope, covering a range of impacts that include social, ethical, legal and operational aspects. ***It is also really important to remember that generative AI is not the whole of the moon here, not by a long shot.***

This AIIA ("A2A"?) will need address and aim to document and mitigate the broad universe of challenges and potential risks associated with the relevant AI type and model, such as algorithmic biases, reliability, decision-making processes, data handling specific to AI operations and certain models, impacts on individuals, the organization using it, stakeholders, and broader societal impacts.

Further, this document will need to evaluate and rule out the potential for and likelihood of a huge number of risks to fundamental human rights. Privacy is one of them - always *(cue PIA!- and the PIA says- bias! discrimination! security! which is RIGHT AND NEEDED but only for the aspects derived from the personal data to be provided by or through or at the request of the organization and it would lead back over to the A2A (yes I will stop trying to make "fetch" happen))* but in truth there is no fundamental human right that is immune to the potential to be impacted by AI.

I will leave that there instead of giving you a laundry list. I stand by it, 100%

Importantly, though, I think that AI can also be used to have a positive impact on almost all of them (some are a stretch to be sure AI will do more good than harm- privacy and several types of freedoms among them.)

Shoshana Rosenberg

AI RISK REVIEWS

UNDERSTANDING THE AI RISK REVIEW IN ORGANIZATIONAL CONTEXT

Here I speak to the AI Risk Reviews that must be undertaken by organizations integrating or otherwise working with and incorporating external AI technologies or components. This would not change dramatically for organizations who start to build their own machine learning tools, but the goal is to gear this to the 99% of Privacy and AI Governance teams, not the 1. By the time that distribution changes and most companies have their hands in building AI models of one sort or another, the EU will have this all sorted for us anyway.

Prudent organizations must navigate this new era with foresight and preparedness. Organizations around the world are realizing that they will need to engage in an ongoing process of determining and discerning their thresholds and protocols for different levels of AI Risk Reviews. This has to go beyond the mere assessment of impacts and delve into the multifaceted risks and changes associated with the deployment and integration of AI systems anywhere within the business infrastructure.

AI IS NOT STATIC OR FIXED. OUR PROGRAMS AND RISK REVIEWS CANNOT BE STATIC.

Once (and if) you have determined your initial organizational stance and controls with regard to generative AI tools, which also should be regularly and continually reviewed, AI deployment and use opportunities and new tools should lead to Risk Reviews that evaluate potential risks to an organization's operational integrity, strategic objectives, and overall risk profile.

W**hile the AI Impact Assessment focuses on the broader implications of a given use or tool, the AI Risk Review is distinctly business-centric.** It should include the Impact Assessment prior to any final decision, but must focus on the factors that matter most to that organization and its strategic objectives, and any risks (including those raised in an AI Impact Assessment) to its operations, reputation, and financial health.

In practical terms, conducting an AI Risk Review - whether the risks are low level and the review streamlined , or the risks are high level risks and it is a more involved, technical and robust review- requires a multidisciplinary approach. Vendor assessment and integration risks are paramount and should include anticipating the risks of operational dependency on external AI systems and any contingency planning.

The reviews and processes will need to be iterative, adapting as both the organization, the regulatory climate, and AI technologies evolve. As markets and shifts in strategies evolve, so too should our understanding and management of AI risks.

The separation of AI Risk Reviews from AI Impact Assessments, however, is critical to ensure a focused and tailored approach to managing the unique risks brings to the table. (See also: strategic disclosure- though the cases are rarer where a risk review would not be pulled to a matter in where an AI Impact Assessment is required- it is good to know what kind of weapon to bring to any particular kind of fight.

AI RISK REVIEWS, AI IMPACT ASSESSMENTS,
AND THE REGULATIONS TO COME, WILL (should)
REQUIRE TWO NEW TOOLS.

Shoshana Rosenberg

EXPLAINABILITY ADDENDA

HOW CAN YOU JUST ADD 'EXPLAINABILITY " TO EVERYTHING?

I hear you. Stay with me.

To effectively assess the risks associated with a given AI model, whether directly integrated into a new or existing business operation or simply added as a component to a (read as: every) set of software tools or professional services the business relies on, organizations need a deep understanding of how the AI system functions.

In this moment before Explainability by Design, and long after, organizations should have an explainability addendum to their vendor contract that provides detailed and specific information about the AI or AI components, its decision-making processes, data usage and sourcing, feature importance, reliability, and environmental impact (its Minimum Viable Explainability, if you will.)

This detailed understanding is not just part of technical due diligence, it is the means by which purchasing organizations can ensure responsible AI usage, aligning with ethical standards and regulatory standards.

In this antediluvian period of time before AI and AI-integrated vendors have the answers, we should still be giving them these questions, these addenda, to evidence both their need to provide them, and the extent to which we are currently forced to make big and impactful decisions without critical information.

EXPLAINABILITY NOTICES

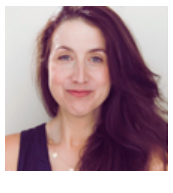
See? You are with me now. I don't even have to say it.

Without forcing organizations to set explainability notices in place on their websites, ideally in a tiered way that is accessible to individuals with one level of granularity and organizations doing risk and impact assessments at another, vendors and service providers can provide different answers to different organizations (very cynical, I know) but also- we have already established that transparency, accountability and trust are the only way forward. It also makes the job of those conducting the risk and impact assessments- or the vendor staff facilitating them or the addendum- immeasurably harder.

The availability of a public notice ensure compliance with the flood of emerging regulations that is headed this way and assures individuals and clients that they are working with transparent and ethical AI providers. Don't worry, they will get easier to create, and to read, with the advent of Explainability by Design.

WAIT. WHY DO YOU NEED A MICROPHONE?

A microphone is a metaphor. For AI Governance you need organizational support and empowerment. You need a platform to provide policies and resources, and you need to be able to speak over the din of opportunities and risks- to advocate, to educate, and to foster a culture that can support an agile and robust AI governance program. Don't wait for the fine structure to be determined under a given law to ask for what you need- because it isn't you that needs it- it is the organization you serve. (Also, they are handy and impactful for dropping alongside good advice, some say.)



Shoshana Rosenberg

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.

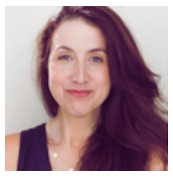
DATA IS THE KEY TO AI

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

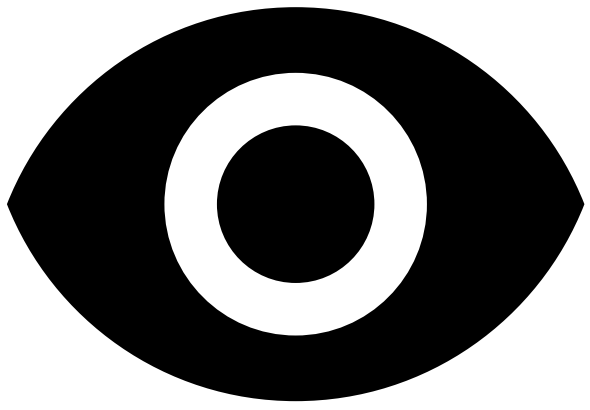
AI that is properly and thoughtfully designed with DEI principles can identify and help flag and rectify systemic disparities across any number of sectors and disciplines and processes, as well as helping to identifying gaps in the policies or tools that support them.

STATEMENTS



Shoshana Rosenberg

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



A.I. DETECTS STATE PRIVACY

Shoshana Rosenberg