**Reading comprehension**

|  |
| --- |
| *With the implementation of a new AI designed to assist HR with CVs, there arises the question of factors to be considered during both the development and implementation phases. This is to elucidate the optimal approach for implementation. An article from Harvard Business Review seems useful to initiate preparation for an upcoming meeting with the HR department on this topic.* |

Sum up the relevant aspects **in English**. For your compilation you can use keywords.

1. Four suggestions to support the development of ethical AI (4P)
2. Two explanations for the challenges of implementing AI ethical risk mitigation strategies (2P)
3. Three reasons for assembling a senior-level working group for implementation (3P)
4. Five demands that AI ethics place on organizations (5P)
5. Two ethical implications of using AI in HR (2P)
6. Two practical implications of using AI in HR (2P)
7. Two suggestions for AI training in the context of HR (2P)

*Business Ethics*

**Ethics and Al: 3 Conversations Companies Need to Have[[1]](#footnote-1)**

by Reid Blackman and Beena Ammanath

March 21, 2022

Over the past several years, concerns around Al ethics have gone mainstream. The concerns, and the outcomes everyone wants to avoid, are largely agreed upon and well documented. No one wants to push out discriminatory or biased AI. No one wants to be the object of a lawsuit or regulatory investigation for violations of privacy. But once we've all agreed that biased, black box, privacy-violating Al is bad, where do we go from here? The question most every senior leader asks is: How do we take action to mitigate those ethical risks?

Acting quickly to address concerns is admirable, but with the complexities of machine learn-ing, ethics, and of their points of intersection, there are no quick fixes. To implement, scale, and maintain effective AI ethical risk mitigation strategies, companies should begin with a deep understanding of the problems they're trying to solve. A challenge, however, is that conversations about AI ethics can feel nebulous. The first step, then, should consist of learning how to talk about it in concrete, actionable ways. Here's how you can set the table to have AI ethics conversations in a way that can make next steps clear.

**Who Needs to Be Involved?**

We recommend assembling a senior-level working group that is responsible for driving AI ethics in your organization. They should have the right skills, experience, and knowledge such that the conversations are well-informed about the business needs, technical capacities, and operational know-how. At a minimum, we recommend involving four kinds of people: technologists, legal/compliance experts, ethicists, and business leaders who under-stand the problems you're trying to solve for using AI. Their collective goal is to understand the sources of ethical risks generally, for the industry of which they are members, and for their particular company. After all, there are no good solutions without a deep understanding of the problem itself and the potential obstacles for proposed solutions.

You need the technologist to assess what is technologically feasible, not only at a per product level but also at an organizational level. That is because, in part, various ethical risk mitigation plans require different tech tools and skills. Knowing where your organization is from a technological perspective can be essential to mapping out how to identify and close the biggest gaps.

Legal and compliance experts are there to help ensure that any new risk mitigation plan is compatible and not redundant with existing risk mitigation practices. Legal issues loom par-ticularly large in light of the fact that it's neither clear how existing laws and regulations bear on new technologies, nor what new regulations or laws are coming down the pipeline.

Ethicists are there to help ensure a systematic and thorough investigation into the ethical and reputational risks you should attend to, not only by virtue of developing and procuring AI, but also those risks that are particular to your industry and/or your organization. Their importance is particularly relevant because compliance with outdated regulations does not ensure the ethical and reputational safety of your organization.

Finally, business leaders should help ensure that all risk is mitigated in a way that is compatible with business necessities and goals. Zero risk is an impossibility so long as anyone does anything. But unnecessary risk is a threat to the bottom line, and risk mitigation strat-egies also should be chosen with an eye towards what is economically feasible.

**Three Conversations to Push Things Forward**

Once the team is in place, here are three crucial conversations to have.

One conversation concerns coming to a shared understanding of what goals an Al ethical risk program should be striving for. The second conversation concerns identifying gaps between where the organization is now and where it wants to be. The third conversation is aimed at understanding the sources of those gaps so that they are comprehensively and effectively addressed.

**1) Define your organization's ethical standard for Al.**

Any conversation should recognize that legal compliance (e.g. anti-discrimination law) and regulatory compliance (with, say, GDPR and/or CPA) are table stakes. The question to address is: Given that the set of ethical risks is not identical to the set of legal/regulatory risks, what do we identify as the ethical risks for our industry/organization and where do we stand on them?

There are a lot of tough questions that need answers here. For instance, what, by your organization's lights, counts as a discriminatory model?

Suppose, for instance, your Al hiring software discriminates against women but it discrimi-nates less than they've been historically discriminated against. Is your benchmark for suffi-ciently unbiased "better than humans have done in the last 10 years"? Or is there some other benchmark you think is appropriate? Those in the self-driving car sector know this question well: "Do we deploy self-driving cars at scale when they are better than the average human driver or when they are at least as good as (or better than) our best human drivers?"

Similar questions arise in the context of black box models. Where does your organization stand on explainability? Are there cases in which you find using a black box acceptable (e.g. so long as it tests well against your chosen benchmark)? What are the criteria for determining whether an AI with explainable outputs is otiose, a nice-to-have, or a need-to-have?

Going deep on these questions allows you to develop frameworks and tools for your product teams and the executives who green light deployment of the product. For instance, you may decide that every product must go through an ethical risk due diligence process before being deployed or even at the earliest stages of product design. You may also settle on guidelines regarding when, if at any time, black box models may be used. Getting to a point where you can articulate what the minimum ethical standards are that all AI must meet is a good sign that progress has been made. They are also important for gaining the trust of customers and clients, and they demonstrate your due diligence has been performed should regulators investigate whether your organization has deployed a discriminatory model.

**2) Identify the gaps between where you are now and what your standards call for.**

There are various technical "solutions" or "fixes" to Al ethics problems.

A number of software products from big tech to startups to non-profits help data scientists apply quantitative metrics of fairness to their model outputs. Tools like LIME and SHAP aid data scientists in explaining how outputs are arrived at in the first place. But virtually no one thinks these technical solutions, or any technological solution for that matter, will sufficiently mitigate the ethical risk and transform your organization into one that meets its Al ethics standards.

**3) Understand the complex sources of the problems and operationalize solutions.**

Many conversations around bias in Al start with giving examples and immediately talking about "biased data sets." Sometimes this will slide into talk about "implicit bias" or "unconscious bias," which are terms borrowed from psychology that lack a clear and direct application to "biased data sets." But it's not enough to say, "the models are trained on biased data sets" or "the AI reflects our historical societal discriminatory actions and policies."

The issue isn't that these things aren't (sometimes, often) true; it's that it cannot be the whole picture. Understanding bias in AI requires, for instance, talking about the various sources of discriminatory outputs.

That can be the result of the training data; but how, exactly, those data sets can be biased is important, if for no other reason than that how they are biased informs how you determine the optimal bias-mitigation strategy. Other issues abound: how inputs are weighted, where thresholds are set, and what objective function is chosen. In short, the conversation around discriminatory algorithms has to go deep around the sources of the problem and how those sources connect to various risk-mitigation strategies.

Productive conversations on ethics should go deeper than broad stroke examples descried by specialists and non-specialists alike. Your organization needs the right people at the table so that its standards can be defined and deepened. Your organization should fruitfully marry its quantitative and qualitative approaches to ethical risk mitigation so it can close the gaps between where it is now and where it wants it to be.

And it should appreciate the complexity of the sources of its Al ethical risks. At the end of the day, AI ethical risk isn't nebulous or theoretical.

It's concrete. And it deserves and requires a level attention that goes well beyond the repetition of scary headlines.

1. <https://hbr.org/2022/03/ethics-and-ai-3-conversations-companies-need-to-be-having> [↑](#footnote-ref-1)