HRPA HR POLICY ASSOCIATION®

HR Policy Association's AI Principles

Introduction

Data privacy, and more broadly the ethics of AI and digitization in the workplace, are topics garnering increasing interest from Congress, regulators, watchdog agencies and the media and public at large. As the development of technology has outpaced our ability to regulate, moderate or even fully understand it, policymakers struggle to stem the tide.

In addition to the European Union's General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA), both of which have been at least partially praised by tech giants Microsoft, Facebook and Apple, global initiatives focusing on the ethics of AI use proliferate. In the United States, state and federal legislative and regulatory proposals aiming to regulate corporate use of machine-learning and employee data are well underway, including the following:

- Illinois' Artificial Intelligence Video Act, which went into effect in January of 2020, places requirements on companies regarding notice, consent, confidentiality and transparency when using AI to analyze video interviews.
    - This is particularly timely given the recent FTC complaint filed by tech watchdog EPIC (Electronic Privacy Information Center) against HireVue, which uses face-scanning technology in combination with AI to assess job applicants.
- A similar new initiative in California, the California Privacy Rights Act, would require companies to increase disclosures around "profiling algorithms" pertaining to employment.
    - San Francisco has already passed an outright ban on the use of facial recognition software by police and other government agencies, but not private companies (yet).
- The Algorithmic Accountability Act of 2019, the first federal effort to regulate bias in corporate use of algorithms, would require companies to assess how they use "automated decision systems" to mitigate bias, discrimination and privacy violations.
- In January of 2020, the White House proposed regulatory principles for AI targeted at federal agencies, instructing them to consider factors such as public trust in AI, public participation, scientific integrity, risk assessment, disclosure and transparency among others.

At the time of writing, 42 countries including the U.S. had adopted the new OECD Principles on Artificial Intelligence, the first ever set of intergovernmental policy guidelines on AI. These five value-based principles, summarized, state that:

- AI should benefit people and the planet through inclusive growth, sustainable development and well-being.

- AI should be human-centered and respect the rule of law, human rights, and democratic values including diversity.
- Those designing AI systems should commit to transparency and responsible disclosure so that stakeholders are aware of and can challenge AI-based outcomes.
- AI systems should be robust, secure and safe, with systematic risk management employed.
- AI creators and users must be accountable for the function and compliance of AI with the above principles.

Company Approaches

Separately from the avalanche of potential legislative and regulatory initiatives concerning AI, companies are showing great innovation and leadership in developing principles to guide and govern their use of data in all forms. For example, Google has published their seven principles for the use of AI:

1. Be socially beneficial
2. Avoid creating or reinforcing unfair bias
3. Be built and tested for safety
4. Be accountable to people
5. Incorporate privacy design principles
6. Uphold high standards of scientific excellence
7. Be made available for uses that accord with these principle

Building upon the principles currently in use by many HR Policy Association member companies as well as the OECD standard, we have developed some sample principles to help companies who are just beginning the process of putting their practice regarding the use of data and machine learning into words. We hope these principles will help guide thinking about this complex topic.

Sample Principles

For most companies, principles for the use of employee data and AI will coalesce around the following five categories. We therefore offer these as a framework and starting point for companies to leverage in their own environments.

Principle 1: Privacy and Security

Although most companies currently have an existing data privacy policy, these are often universal in scope or geared toward customers and consumers. Principles for the use of data and AI should include a statement specific to employee privacy and security, and may explicitly state that data may not be used for the purpose incompatible with the specific purpose for which it was collected without employee consent.

Principle 2: Transparency

A recent Accenture survey found that 92% of employees are "open to the collection of data on them and their work" if it is done in a way that benefits them, such as in exchange for an "improvement in their productivity, wellbeing or other benefits."[1] The intended uses of data should be able to be clearly understood, explained and shared, including the impact on decision-making and the processes for raising and resolving any issues. In some cases, this may include an explanation of the algorithms involved in machine learning assisted analysis and how those algorithms are developed and "trained" to analyze employee data.

Principle 3: Integrity

The principle of integrity is interpreted in a variety of different ways by companies according to their culture but is rooted in the concept of "positive intent." In addition to committing to the use of data in a highly responsible way, companies may also specify that the purpose of all AI is to augment and elevate humans rather than replace or diminish them, and that data usage should be sensitive to cultural norms and customs and aligned with company values.

Principle 4: Bias

Although AI has been touted as the solution to unintended bias in many people-related processes, such as hiring, performance management and promotion, the risk of unintentional bias occurring within AI algorithms or the datasets used to train them is concerning. Principles around data and ethics should commit to continuous monitoring and correction for unintended bias in machine learning.

Principle 5: Accountability

Individuals should be accountable for the proper functioning of AI systems and for unintended consequences arising out of its use. Companies should ensure that everyone involved in the lifecycle of an AI system is trained in AI ethics and that ethics is part of the product development and operation of an AI system. This may include the coders and developers responsible for creating the software, the data scientists responsible for training it, or the management of the company.

---

[1] Decoding Organizational DNA, Accenture: https://www.accenture.com/us-en/insights/future-workforce/workforce-data-organizational-dna.

Further Questions

Beyond the key principles outlined above, additional questions abound:

1. Who should manage the ethics principles and monitor compliance for the company?
2. Can the full potential of AI systems ever be utilized without risking worsening or creating new inequalities and biases?
3. What will the potential "chilling" effect be on the development and application of AI in business if individuals involved in its creation are held personally accountable for unintended adverse outcomes?

The answers to these questions are far from fully-formed and will continue to evolve rapidly (though perhaps not at the pace of the technology itself).

However, we hope these guidelines are a useful starting point for the complex decisions and thinking needed to tackle the Gordian knot of artificial intelligence, data analytics and ethics.