30 Managing AI's Nontechnical Risks

While Chapters 17 and 18 outlined some risks that an IT team **can** control, the risks herein require a broader discussion with other leaders. Members of the HR, legal, and executive teams will glean useful insights from this material.

This chapter starts with some examples of organizational risks for which there are limited or no technical controls. Then it provides some suggested policy and processes to minimize the gap. Finally, and for reference, the efforts of governments and standards bodies to define and control AI's risks are outlined. Their ideas provide a broader context to consider when developing an organization's specific approach to risk.

Primary (known) human risks

There are some known human risks, while some are still evolving. Some will wane in time as people become acclimated to using GenAI. Some will shrink as technical systems and controls improve. Some will simply become a new normal. Some will surprise us!

Rather than a long list of possible risks, the four highlighted here could have legal or HR impacts on organizations in the near-term.

Copyright infringement

LLMs may produce content that infringes on the intellectual property rights of others. It'll be difficult for an individual to know if the response specifically used one author's content, or patented information from another company. But if that author or company finds their content or idea being used by another, they may consider filing a case.

Counsel should also be aware that Microsoft has a Copyright Commitment where they "assume responsibility for the potential legal risks involved."[1]

It's true, Copilot may generate text, images, or music that are similar or identical to existing works, without proper attribution or permission.

At this point, several lawsuits have been filed against OpenAI and Microsoft. They're mainly class action suits by content creators and entertainment industry groups. The New York Times is a significant filing from an individual organization.[2] Legal professionals will be following such cases and considering the implications of enterprise use.

To minimize the risk of copyright infringement, an enterprise should take some precautions when using generative AI.

- Be aware that the sources and quality of the data used to train the model may include copyrighted content and ideas of competitors.
- Check the model's output using manual reviews to verify its accuracy and originality.
- Follow the ethical and legal guidelines for using and sharing the generated content, and respect the rights and interests of the original creators.
- Educate and train associates on the responsible and safe use of generative AI, and establish clear policies and procedures for reporting and resolving any issues or disputes.
- Ensure employees use private GenAI tools like Copilot instead of public models like ChatGPT when sharing internal content, to ensure the public models won't pick up any of your organization's secrets.

Biases

While Copilot has stayed out of the news, there have been instances where companies faced lawsuits due to AI mistakes or biases. For example, the Equal Employment Opportunity Commission (EEOC) settled its first-ever AI discrimination lawsuit with a tutoring company. The case involved allegations that the company's **AI-powered hiring selection tool automatically rejected women applicants over 55 and men over 60**. The company agreed to pay $365,000 to resolve charges of age and gender discrimination.[3]

Such cases highlight the legal and ethical challenges companies face when implementing AI tools, especially in sensitive areas like hiring. They underscore the importance of ensuring AI tools do not inadvertently discriminate against protected classes and comply with anti-discrimination laws.

Phishing and Deepfakes

AI makes the creation of social engineering tricks like phishing lures and fake images and videos simple. More than just memes, deepfakes can create public and political problems. GenAI can be used for nefarious purposes. Microsoft fine tunes its models to block prompts and responses it deems to be illegitimate. Still, **LLMs can be jailbroken**[4]. In recent memory, OpenAI lost control of one of its cybersecurity-specific models which ended up breaking several rules before hacking into Hugging Face.

> Organizational policy should *"Prohibit the use of GenAI for anything other than acceptable business use."*

Public Sentiment

There's no shortage of daily news about data center controversies, potential job losses, rogue AI agents, and ecological impacts. Despite efforts to separate news and social media from the work environment, people will bring their personal sentiments to work, especially when seeing others using AI. If an organization isn't consistently sending the same message and supporting the new way of work at all levels, the chances of personal beliefs leaking into work life are high.

> In organizations with histories of employee activism or union membership, the importance of aligning and communicating is heightened.

Communicating can control risk

As highlighted in the breakout boxes above, organizations that use generative AI should develop and implement an AI policy, a set of principles, guidelines, and processes that govern the use of AI within the organization.

AI Position Statement

Write and communicate the organization's guiding principles. A (living) document should define the guiding principles of the use of AI, such as fairness, accountability, transparency, human agency, and social good.

The Position Statement should be one of the first outputs of the Center of Excellence, outlined in Chapter 28.

AI Policy

The CoE should then draft an initial AI policy to consider the following aspects:

- **Acknowledge the risk** of using tools whose output aren't completely understood.
- **Establish clear and consistent guidelines** for using generative AI, defining the roles, responsibilities, and expectations of the AI providers, users, and stakeholders.
- **The data sent and received from Copilot** should comply with applicable data protection and privacy laws and regulations.
- **Alert users of your responsible AI practices**, such as monitoring, auditing, reviewing, as well as providing feedback and training.

An AI policy should be communicated and disseminated to Copilot users and stakeholders, ensuring their awareness, understanding, and compliance.

AI safety regulation

As generative AI becomes more widespread and impactful, standards-bodies and governments have recognized a need for regulation and governance of its use. So have the AI labs themselves.

The National Institute of Standards Technology (NIST) have created an AI Risk Management Framework.[5] It's especially useful for organizations who are using (not creating) AI in their business.

International Organization for Standardization (ISO) has created an AI specific document to complement its general framework for risk management.[6]

The European Union (EU) enacted a comprehensive framework for regulating AI in 2024, establishing clear rules and guidelines for AI use in their territory.[7] The USA has a looser "Action Plan," but that's more about the desire to accelerate AI innovation, build AI infrastructure, and leading in national security.[8]

In the absence of any international standards, some of the AI labs have suggested a global AI oversight body, similar to the US's Financial Industry Regulatory Authority.[9]

While most efforts are aimed at developers, rather than consumers or Copilot customers, **organizations that use generative AI should *share* the responsibility** for its use and impact. To ensure AI is used in the interests of its people and society, organizations should:

- **Share the NIST RMF with legal, compliance, and risk management teams**, ensuring that the organization is aware of and addresses the proposed AI regulations.
- Be transparent about the use of generative AI, disclosing the use, purpose, and limitations of AI.
- Collaborate with the AI providers, users, and stakeholders, establishing a system of coordination, as well as a mechanism for feedback and dispute resolution.

Conclusion

While helpful, Generative AI also poses risks, such as malicious use, bias, discrimination, copyright infringement. Therefore, prudent organizations implement AI safely and effectively, by considering the following aspects:

- Organizations need to develop and implement an **AI policy**, a set of principles, guidelines, and processes that govern the use of AI within the organization.
- Organizations need to assume and **share responsibility** for the use and impact of AI, ensuring that the AI system is safe, secure, and respectful of the rights and interests of the people and society.
- Organizations need to secure the use and **privacy** of AI in their tenants, using the tools and features provided by Microsoft and other best practices.
- Even though Microsoft provides some backing for **copyright infringement**, organizations need to educate their users on the fact that Copilot gleans information from many sources, some of which may be copyrighted.
- Organizations should at least be aware of **and take action on the NIST's Risk Management Framework,** as well as other relevant laws and regulations for AI use.

By implementing AI safely and effectively, organizations can leverage the power of generative AI while mitigating the risks, enhancing their productivity, creativity, and innovation.

Next up, addressing a macro-trend that AI may well instigate: Job impacts, and how to upskill employees to keep them ahead of the machines.

References and Resources

[1] https://blogs.microsoft.com/on-the-issues/2023/09/07/copilot-copyright-commitment-ai-legal-concerns/
[2] https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html
[3] https://www.eeoc.gov/newsroom/itutorgroup-pay-365000-settle-eeoc-discriminatory-hiring-suit
[4] [2404.01833] Great, Now Write an Article About That: The Crescendo Multi-Turn LLM Jailbreak Attack (arxiv.org)
[5] https://www.nist.gov/itl/ai-risk-management-framework
[6] https://www.iso.org/standard/77304.html
[7] https://artificialintelligenceact.eu/high-level-summary/
[8] https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf
[9] https://demishassabis.substack.com/p/a-framework-for-frontier-ai-and-the-dawning-of-a-new-age