# Leverage Leadership's Role In Shaping Responsible AI Regulations

fractal

# Overview

Top-level executives in the C-suite and CXO positions are just beginning to explore the extensive capabilities of Generative AI (GenAI) to enhance and optimize their operational processes. Amid the enthusiasm for the limitless possibilities offered by GenAI, there is a heightened awareness of the risks associated with its deployment for external purposes. Companies predominantly harness AI to boost internal productivity rather than generate content for public release.

This caution indicates a broader industry trend where companies are increasingly mindful of producing content that is accurate, ethical, and compliant with legal standards. This meticulous approach to utilizing GenAI, especially for external-facing content, underscores the commitment of many organizations to a considerate and Responsible AI (RAI) implementation.

The great paradox of responsible technology use lies in its universal accessibility versus the potential for misuse. GenAI is no different.

fractal

# The current challenges

With the proliferation of open-source GenAI, people are becoming adept at using these tools. However, users are often not informed about the potential inaccuracies, the reliance on probabilities, or issues like content originality and potential copyright infringement. There is also the potential for GenAI to perpetuate discrimination and bias. For example, if you ask ChatGPT for jokes about a specific gender, it may respond differently based on the gender mentioned. The same applies to toxicity; the nature of the input influences the nature of the AI's response. If the prompt is toxic, the response mirrors that.

There is no easy solution because tools like ChatGPT, for example, mirror information they've been trained on — usually data already on the internet. To tackle this issue, we need to address the root of the problem, i.e., remove such content from the internet itself — an impractical, if not impossible, task. Having observed potential misuse, AI creators must become more cautious. Future iterations of tools like ChatGPT should be crafted to prevent misuse, ensuring they refrain from responding to inappropriate queries.

This is an ongoing process, and until that happens, the onus is on users to understand how to interact appropriately with GenAI. This makes educating users about what they shouldn't do with AI vital.

fractal

# Education and guardrails:
# The critical keys to RAI

Evidently, every technology, from electricity to nuclear power, has a propensity for misuse. For GenAI, this issue is further complicated by our limited exposure to the potential problems it can create.

The key lies in educating people about what not to do. We must devise strategies and processes that effectively communicate the potential long-term problems of misuse of AI. Take electricity, for example; we've been taught basic safety measures like not dropping water on live wires or touching them with bare hands. Similarly, guidelines for AI usage are crucial. For example, the nuances of framing prompts, what to expect from AI responses, and understanding the ethical and safety considerations are critical. Over time, it can be fairly expected that we will see progress in this area, but it's a journey that requires conscientious efforts from all involved in AI development and use.

In the meantime, AI development should include strict guardrails. These should not only be seen as temporary measures but should be designed with the future in mind. Technology evolves rapidly, making it impractical to update safety measures constantly. Our focus should, therefore, be on creating guardrails that are not necessarily perfect but sufficiently robust to last for the next decade or so. The objective is to build a framework that's adaptable and resilient enough to embrace technological progress while ensuring safety and ethical considerations are not compromised. This is critical while we wait for regulators to catch up.

fractal

# Global regulation and its issues

Governments are already working towards this integrative approach, but the groups informing policy seem limited, confined to a few specific skill sets. To fully tap into the possibilities of AI and navigate its complexities, we need to widen this circle of expertise. It's not just about having more people involved; it's about having the right mix of people, each contributing their distinct experiences and perspectives to the table. This will be instrumental in developing robust, effective, and forward-thinking AI policies.

But this is only the tip of the iceberg, and as AI technology develops further, developers, governments, and regulators will become more proactive. Globally, from the G20 to individual countries like the US, UK, India, and China, there's increasing dialogue about privacy, ethical use, and RAI. Policies are emerging, and although they might lack enforcement power, they will eventually evolve, imposing penalties for misusing technology.

When discussing the role of regulators in AI, it is necessary to note that they're often in the back seat rather than at the steering wheel. Technology is advancing rapidly, but regulators seem to lag, struggling to keep pace. They must shift from a reactive stance to a proactive, futuristic approach.

Looking forward, the future versions of GenAI tools will address many of the current RAI issues. In the interim, AI developers and the organizations using their tools shouldn't wait for regulatory bodies to catch up. Companies already know the potential harms of GenAI technology and should proactively self-regulate.

fractal

# It all starts with self-regulation

While it's true that self-regulation may not always be stringent or comprehensive, its fundamental value lies in the intention behind it. The mere act of a company or an individual choosing to self-regulate is a significant step, as it demonstrates a conscious commitment to avoiding actions that could be harmful or unethical.

Another vital aspect is public disclosure. Companies often share financial reports and other corporate information. Similarly, they should be transparent about their AI policies and practices. Publishing AI metrics would foster public trust and invite scrutiny and confidence in these systems. To make AI genuinely responsible and ethical, transparency — not accuracy — is the most critical aspect currently.

# 80% explainability is good enough

Even with just a sliver of understanding about what a model is doing, you can work towards making it transparent. Let's take a familiar example. The concept of AI as a black box, where we don't really see what's happening under the hood, is increasingly outdated. With technological advancements, numerous methods and techniques are now at our disposal that can transform this black box into a white box.

Once we've made this initial move towards transparency, we can continually refine and improve these white-box models. The goal is to make them clearer, more explainable, more transparent, and ultimately, more user-friendly.

Moving towards this goal starts with high-level oversight and decision-making, and industry leaders must take up and lead the way as we tackle this challenge.

fractal

# Conclusion

## It starts with organizational leadership

Addressing the complexities of AI begins with a fundamental step: acknowledgment. It's crucial to recognize the potential harm AI can cause, including reputational, revenue, and regulatory risks.

Understanding AI is akin to knowing the basics of car maintenance. We know when to refuel, service the car, or what to do if there's a puncture. Similarly, we should know whether an AI system is biased, how it handles our data, and the safeguards to prevent misuse or data breaches. Once this acknowledgment is made, it paves the way to address issues related to self-regulation effectively.

Leadership is pivotal role in this process. It's not sufficient to wait for external directives on how to proceed. The impetus must come from the top — CEOs, CTOs, and other C-suite executives must step forward. They need to acknowledge their organizations' challenges with AI openly and set clear strategies to tackle them. It's about leading by example and setting a precedent for the entire organization.

This will likely create a snowball effect: when significant players in the industry begin to regulate themselves effectively, it sets a standard. It encourages other companies to follow suit, fostering a broader culture of responsibility and ethical AI use across the industry. This becomes a chain reaction, where the commitment to self-regulation at the top cascades throughout the sector, promoting widespread adoption of ethical AI practices.

## Author

### Sray Agrawal
Principal Consultant,
Strategic Center

fractal

# About Fractal

Fractal is one of the most prominent providers of Artificial Intelligence to Fortune 500® companies. Fractal's vision is to power every human decision in the enterprise, and bring AI, engineering, and design to help the world's most admired companies.

Fractal's businesses include Asper.ai (AI for revenue growth management) & Flyfish (generative AI for Sales). Fractal incubated Qure.ai, a leading player in healthcare AI for detecting Tuberculosis, Lung cancer, heart failure and stroke.

Fractal currently has 4500+ employees across 17 global locations, including the United States, UK, Ukraine, India, Singapore, Middle East and Australia. Fractal has been recognized as 'Great Workplace' and 'India's Best Workplaces for Women' in the top 100 (large) category by The Great Place to Work® Institute; featured as a leader in Data Engineering services 2024 & Data Science Services 2024 by Information Services Group, Leader in AI and Analytics Services Specialists Peak Matrix Assessment 2021 by Everest Group, Leader in Customer Analytics Service Providers Wave™ 2023 by Forrester Research, Inc

For more information, visit fractal.ai

Corporate Headquarters
Suite 76J,
One World Trade Center,
New York, NY 10007

## Get in touch

fractal