**Generative AI and Professional Responsibility**

**StudentID:**21072686

## **Section 1: Research and Analysis – Generative AI and Professional Responsibility**

Generative Artificial Intelligence (GenAI) is no longer a distant innovation—it is now a driving force within the computing profession. From tools that write code and generate digital artwork to systems that simulate virtual worlds and assist in legal drafting, GenAI is redefining the boundaries of what software can create. These technologies raise critical questions about the responsibilities of computing professionals who build, implement, and oversee them. As GenAI systems become more integrated into industry workflows and public-facing applications, professionals are increasingly expected to navigate the legal, ethical, and sustainability challenges associated with their use.

### **1.1. Legal Responsibilities: Accountability and Ownership in a Generative Era**

A central issue for professionals working with GenAI is the unresolved nature of intellectual property rights and content ownership. These systems are trained on vast datasets—often harvested from the internet—without always obtaining appropriate permissions. This raises questions about the legality of both the training process and the outputs themselves. As Taeihagh (2025) notes, existing copyright laws are not designed to manage content that is algorithmically generated yet statistically derived from human-made works.

The opacity of model decision-making further complicates legal responsibility. Outputs generated by large language models (LLMs) are difficult to trace, making it challenging to assign liability when errors, misinformation, or harmful content are produced (Taeihagh, 2025). Professionals in computing now have a duty not just to create tools but to ensure traceability, explainability, and legal compliance across the AI development pipeline (Floridi and Cowls, 2022).

In high-risk domains such as law, healthcare, and finance, this professional responsibility becomes even more significant. The infamous 2023 court case where lawyers unknowingly submitted AI-fabricated citations demonstrates the potential consequences of failing to verify GenAI outputs (Bender et al., 2021). Such examples reinforce the need for engineers and developers to work in alignment with regulatory safeguards and verification protocols.

### **1.2. Ethical Implications: From Model Bias to Human Oversight**

GenAI systems reflect the data they are trained on—and that data often contains embedded biases. As Bender et al. (2021) argue, large language models act as "stochastic parrots", echoing the structures and stereotypes found in their training data without contextual understanding. This creates ethical challenges around bias amplification, fairness, and inclusivity in outputs.

In creative and content-based industries, this challenge is compounded by the blurring of authorship and originality. McCall (2024) highlights how generative systems are widely used to produce music, literature, and design content, raising concerns about the dilution of creative diversity and the commodification of human expression. For computing professionals, the task is no longer just about producing functional systems but ensuring that these systems do not perpetuate inequality or cultural homogenisation.

Laptev et al. (2025) offer a technical perspective on this issue through their work on interpretable model behaviour. By mapping feature flow within LLMs, their approach allows developers to understand how semantic features evolve across model layers, enabling more controlled and ethically guided model steering. This contributes to a growing toolkit for professionals seeking to balance technical power with ethical oversight.

Moreover, Jobin, Ienca and Vayena (2019) analysed over 80 global AI ethics guidelines and found a common emphasis on values such as transparency, accountability, and justice. These principles are now becoming embedded in the skill set expected of computing professionals, who are increasingly being asked to engage with interdisciplinary ethics as part of their technical practice.

### **1.3. Environmental Sustainability: A Professional Duty Often Overlooked**

The environmental cost of GenAI is significant, yet frequently under-discussed in professional settings. Training and deploying large-scale models requires vast computational power, often consuming millions of GPU hours and resulting in substantial carbon emissions. Strubell, Ganesh and McCallum (2019) revealed that the carbon footprint of training a single large NLP model could equal that of multiple transatlantic flights.

Lu et al. (2024), in their work on the Generative World Explorer, demonstrate the capabilities of vision-based GenAI systems to simulate exploratory behaviours in virtual worlds. While impressive, these systems also highlight the intense resource demands of high-fidelity generative tools. As professionals, developers must factor in energy efficiency, model size, and environmental impact when designing and deploying such systems.

Vinuesa et al. (2020) further argue that the role of AI in achieving—or undermining—global sustainability goals must be addressed at the design stage. For responsible professionals, this means actively considering how their models contribute to or detract from broader ecological well-being.

### **1.4. Professional Responsibility in Practice:**

Taken together, these issues challenge computing professionals to evolve their roles. They are no longer simply builders of technology, but stewards of ethical, lawful, and sustainable innovation. This shift is underscored by the call for explainability, fairness, and accountability not as optional add-ons but as core deliverables (Floridi and Cowls, 2022).

Whittlestone et al. (2019) argue that relying solely on high-level ethical principles is insufficient; true professional responsibility requires identifying and managing the tensions that arise in real-world scenarios. Rather than treating ethics as a checklist, professionals must adopt an approach that is sensitive to context, open to interdisciplinary dialogue, and responsive to conflicting priorities. This mindset encourages practitioners to reflect on their assumptions, question design decisions, and engage more deeply with the social consequences of their work.

As GenAI continues to evolve, professionals will increasingly be held accountable for the outcomes of systems they contribute to. This includes addressing not only what the technology can do, but what it should do, and understanding their role in navigating that distinction.

## **Section 2: Reflective Commentary – Generative AI and My Professional Development**

Exploring the intersection of Generative AI and professional responsibility has significantly influenced how I view my future in the computing profession. As a final-year Computer Science student specialising in Artificial Intelligence, I have been deeply immersed in technical development, research, and machine learning. However, the reflective and analytical process of studying the wider implications of Generative AI, as explored in Section 1, has expanded my understanding of what it truly means to be a responsible computing professional in the age of intelligent systems.

### 

### **2.1. From Technologist to Thoughtful Practitioner**

Much of my academic journey has focused on technical innovation—developing models, writing efficient algorithms, and improving performance metrics. My current final-year project, a machine learning-based fraud detection system, is no exception. Initially, I approached this work from a technical standpoint, focused solely on data analysis, predictive accuracy, and system optimisation.

However, researching issues of bias, transparency, and governance in Generative AI (Bender et al., 2021; Taeihagh, 2025) prompted me to reassess my priorities. Fraud detection systems, like generative tools, can have real-world consequences if deployed irresponsibly. A biased dataset or opaque prediction could lead to innocent users being wrongly flagged, while actual fraud slips through. This realisation shifted my mindset—from that of a developer focused on functionality to that of a practitioner concerned with fairness, explainability, and trust.

### **2.2. Applying Interpretability and Ethical Design**

A key insight that influenced this shift came from Laptev et al. (2025), whose work on interpretability in language models introduced techniques for tracing how outputs are formed. While my project is not generative in nature, the principles of transparency and feature traceability are directly applicable. I have begun integrating model interpretability tools such as SHAP into my workflow, not only to understand model behaviour but also to make results accessible to non-technical stakeholders—a practice I now view as a professional obligation.

This ties back to the broader ethical landscape discussed by Floridi and Cowls (2022), who argue for human agency and transparency in AI systems. In my view, these values are essential, particularly in systems that operate with significant autonomy or influence decisions about people’s finances, identity, or access to services.

### **2.3. Recognising Legal Responsibility in Practice**

The legal implications of Generative AI also impacted my approach to development. In Section 1, Taeihagh (2025) highlighted the unresolved nature of accountability in AI-generated content, especially regarding ownership and compliance. While my project does not produce original content in the same way generative models do, it still processes and learns from data—some of which originates from publicly available financial datasets.

This raised questions about data rights, licensing, and consent, which I had previously overlooked. I’ve since reviewed the legal terms associated with my data sources and incorporated data handling procedures that align with ethical and legal standards. This awareness will guide me in future roles where I’ll be expected not only to build systems, but to ensure compliance with evolving AI regulations, such as the AI Act or GDPR.

### **2.4. Sustainability as a Professional Priority**

Another aspect that reshaped my thinking was the environmental cost of AI systems. Reading the work of Strubell et al. (2019) and Lu et al. (2024) exposed the significant energy demands of training large-scale models. While my current project does not involve generative or high-resource models, this broader understanding of compute cost and carbon impact has encouraged me to explore lighter, more efficient architectures and prioritise performance without excessive energy use.

Sustainability, once a secondary consideration in my project planning, has now become a core value. I recognise that professionals must balance innovation with ecological responsibility, a concept that will stay with me as I move forward in both academic and industry settings.

### **2.5. Defining My Professional Identity**

This reflective process has helped me articulate a clearer vision of the kind of professional I want to become. I am no longer satisfied with being someone who simply builds things because they are possible. I want to be a computing professional who builds things that are ethical, legal, sustainable, and aligned with human values.

Over the next five years, I hope to develop further in areas like:

* **Model interpretability and fairness in machine learning**
* **Ethics-informed system design**
* **Legal frameworks for AI governance**
* **Sustainable AI practices**

These goals are not only relevant to my current fraud detection work but also to any future role I may take in AI development, research, or consultancy. I am also interested in roles like AI governance analyst, AI ethics engineer, or responsible AI advisor, where I can combine technical skills with an understanding of wider societal impacts.

### **3. Conclusion**

Reflecting on the role of Generative AI and the responsibilities it places on computing professionals has been transformative. The insights gained in Section 1—from legal uncertainties to ethical risks and environmental trade-offs—have directly shaped how I approach my studies, my project work, and my career goals. As I prepare to graduate and enter the field, I carry with me not only a technical toolkit but a set of guiding principles. These include a commitment to ethical awareness, regulatory compliance, environmental consciousness, and above all, professional accountability in every line of code I write and every system I help build.

## **4. References – Sections 1 & 2:**

* Bender, E.M., Gebru, T., McMillan-Major, A. and Shmitchell, S., 2021. On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pp.610–623.
* Floridi, L. and Cowls, J., 2022. A unified framework of five principles for AI in society. *Ethics and Information Technology*, 24(1), pp.1–17.
* Jobin, A., Ienca, M. and Vayena, E., 2019. The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), pp.389–399.
* Laptev, D., Balagansky, N., Aksenov, Y. and Gavrilov, D., 2025. *Analyze Feature Flow to Enhance Interpretation and Steering in Language Models*. arXiv preprint.
* Lu, T., Shu, T., Yuille, A., Khashabi, D. and Chen, J., 2024. *Generative World Explorer*. arXiv preprint.
* McCall, A., 2024. *Applications of Generative AI in the Creative Sector*. ResearchGate publication.
* Strubell, E., Ganesh, A. and McCallum, A., 2019. *Energy and Policy Considerations for Deep Learning in NLP*. arXiv preprint.
* Taeihagh, A., 2025. Governance of generative AI. *Policy and Society*, 00(00), pp.1–22.
* Whittlestone, J., Nyrup, R., Alexandrova, A. and Cave, S., 2019. The role and limits of principles in AI ethics: Towards a focus on tensions. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pp.76–82.