The Role of the Software Development Lifecycle (SDLC) in Releasing Responsible Technologies

A journalistic paper examining the role of the software development lifecycle born out of the early Facebook-era product development ethos, featuring interviews with Thor Mitchell (CPO-in-residence at Balderton Capital, Former PM at Google) and Marc Leone (Head of Trust and Safety at GIPHY).

By William Bishop

The Role of the Software Development Lifecycle (SDLC) in Releasing Responsible Technologies

In the summer of 2024, a dozen employees of OpenAI, Anthropic, and Google signed a letter to the industry, warning of the dangers AI poses without proper transparency and protections for individual employees to speak up about concerns.[1] In this open letter to industry leaders, employees highlighted how Silicon Valley's major AI players over-rely on the "move fast and break things" culture that shaped the Facebook-era product development ethos, and assume serious risks by releasing early, imperfect versions of AI models.

As a former engineer who actively participated in the software development lifecycle (SDLC) at the likes of Zoom, Meetup, and other tech companies, I understand the risks of releasing an imperfect, minimum viable product (MVP) to leverage a living laboratory of users for rapid feedback. However, there's a big difference between testing the waters with a new Zoom virtual background and releasing early AI models that might share racist advice or encourage discourse on a conspiracy theory like "white genocide."[2] Such concerning behaviors of generative AI, while presented to us in a user-friendly app like ChatGPT, are introduced during the development of the models and data pipelines that feed the app. Much like doctors and clinicians play a role in advancing medical ethics, software engineers and product managers should play a similar role when it comes to the software development lifecycle (SDLC), a process that brings emerging technologies from ideation to fruition. But where does this role begin and end, and how can we best empower the builders at the forefront of software development to help protect us from AI's potential harms?

The Atlantic's Matteo Wong suggests that "reorienting the internet and society around imperfect and relatively untested products is not the inevitable result of scientific and technological progress—it is an active choice Silicon Valley makes every day." This sentiment gets at the heart of decision-making in big tech and highlights the question of responsibility more broadly. To explore this definition of "responsible tech" further, I sat down with a former product manager at Google and current Venture Capitalist, Thor Mitchell. Mitchell has a wealth of experience in the product domain, from the launch of the Google Maps API during his time as an individual product manager (PM), to his current role as Chief Product Officer at Balderton

---

[1] Pranshu Verma and Nitasha Tiku. "AI Employees Warn of Technology's Dangers, Call for Sweeping Company Changes." *The Washington Post*, 4 June 2024, https://www.washingtonpost.com/technology/2024/06/04/openai-employees-ai-whistleblowers.

[2] Matteo Wong. "The Entire Internet Is Reverting to Beta." *The Atlantic*, 18 June 2025, https://www.theatlantic.com/technology/archive/2025/06/ai-janky-web/683228/.

Capital (a UK-based venture capital firm), in which he advises tech clients and startups across a variety of products. When asked what he thought about the role of engineers and PMs in shepherding responsible tech, Mitchell starts by asking “responsible to whom?”

“Responsible to whom” does, in fact, seem like a logical place to start in attempting to understand the motivations of tech companies and the employees that shape their products. According to Mitchell, there are tiers of responsibility in any product-minded organization. First, there is a sense of responsibility to the shareholders and the business at large. Next, there is a responsibility to the users of the product. And usually, there are regulators who shape the confines of the business landscape, including industry-specific legal obligations, such as the General Data Protection Regulation (GDPR), domain-specific security practices, and the like. Mitchell shares that there is also a layer that exists inconspicuously in the middle, involving best practices, or things one *ought* to do to protect the business’s reputation and maintain user trust (but that companies aren’t necessarily mandated to do). In other words, a bit of ethics.

As Mitchell puts it, these best practices are “not necessarily going to get you sent to jail [if you don’t do them], but they will certainly destroy the reputation of your business or undermine the trust you have with your customers, suppliers, partners, or whoever [if you don’t do them].” He adds that “Unfortunately, a lot of that stuff is often missed in many cases, not for ill intent, but […] because people are not aware they should be doing it.” For anyone familiar with the impact of overwhelming, top-down priorities in large tech organizations, this likely resonates deeply. There are only so many hours in the day, and oftentimes, the things that don’t get done are written off as “nice to have” rather than “must have.” When Mitchell put this practice into words, I immediately recalled a time at a previous company where my team rushed to release a new set of APIs for external developers—a release that was achieved only because leaders determined that certain “nice to have” security layers, such as time sensitive authorization for application programming interfaces (APIs), were ultimately deemed non-essential.

Given tight deadlines and competing priorities, how can engineers and PMs do justice to end-users? It’s a difficult task, but there are some concrete things that Mitchell has gleaned from his tenure as a thought leader in tech and product development. One practice worth highlighting is the “black mirror” test. Mitchell describes the black mirror test as “essentially trying to anticipate all of the dark patterns and all of the ways in which your product might be used that you are not comfortable with.” In his experience, PMs are generally an “optimistic bunch of people, because they sort of have to be in order to keep the spirits of the team [up].” But being an optimist often means that PMs “are blind to what can happen when someone who is explicitly a bad actor gets hold of the product” and attempts to use it in ways that weren’t anticipated. In

this sense, one of Mitchell's key points in our discussion is that there is an optimism bias in product teams and "irresponsible" decisions are often ill-informed decisions—not necessarily a result of malicious intent, and not always due to a lack of capacity, either.

As a PM at Google, Mitchell has many stories top-of-mind, but one that stands out is the Google Street View WiFi case, which he describes as a product development situation in which the black mirror test could have been applied more aptly. In 2010, Google announced that it mistakenly collected large swaths of information over WiFi while its streetview cars were navigating public streets. Google's intent was challenged by the FCC (was it really a mistake?), but ultimately the tech giant was let off the hook.[3] According to a report by *Wired*, the engineer who was working on the streetview WiFi technology seemed to be aware of the potential capabilities of the product, and likely knew it could absorb large amounts of tangential data. In this scenario, even a member of that same team could have seemingly applied the black mirror logic to understand, with some ease, that this technology could be used for purposes beyond its original intent. Even though Google was ultimately unfazed by this situation, it is in its best interest to more thoroughly scrutinize its own capabilities to ensure the next scandal doesn't have more serious effects on its users or business reputation.

Failing to foresee risks in product development is a common refrain from experts focusing on the role of the SDLC. But not enough is being done in the culture of product development to normalize this process as one that can benefit from a bit of ethics. In the context of software like AI, the stakes are high. In a 2020 case study of software engineers working in an AI-oriented startup environment ("This is Just a Prototype"), researchers came to similar conclusions. As part of the study's exploration of prototypes and MVPs, it conducted interviews with several engineers working on health-tech related AI projects. According to the published report, developers did not often consider future implications of the technologies they were working on. One developer said: "The calculations are made in the algorithms, so it doesn't really make mistakes." This highlights the same optimism bias raised by Mitchell, and in this case, the report concluded that "Product misuse and error scenarios are only considered during development. They are not considered in terms of the future operational life of the system out [in] the field."[4] Not only is this a damning near-sightedness of the engineers working on this

---

[3] David Kravets. "An Intentional Mistake: The Anatomy of Google's Wi-Fi Sniffing Debacle." *Wired*, 2 May 2012, https://www.wired.com/2012/05/google-wifi-fcc-investigation/.

[4] Vakkuri, Ville. "'This Is Just a Prototype': How Ethics Are Ignored in Software Startup-Like Environments." *Agile Processes in Software Engineering and Extreme Programming: 21st International Conference*, vol. 383, 2020, pp. 195–210, https://doi.org/10.1007/978-3-030-49392-9_13.

technology, but it also shows a detachment that many employees feel from the consequences of their work—something that's seemingly baked into the culture.

This same report contained echoes of Mitchell in terms of accountability, too. Vakkuri et al emphasized in "This is Just a Prototype" that, when team members don't know enough about the capability they're working on, they tend not to feel as accountable to the potential repercussions of the technology, either. The report found that "Without understanding how the system works, it is impossible to establish why it malfunctioned" in the first place. This aligns with Mitchell's point of view, which essentially boils down to "you don't know what you don't know"—a dangerous approach to the development of emerging technologies like AI.

If the individuals working on developing the algorithms and models powering generative AI apps like ChatGPT conduct product development in the same way as the apathetic engineers in Vakkuri's study, shouldn't we be alarmed? How do we ensure individual contributors have more say during product development, both technically and ethically? And how can organizations foster an environment where engineers feel empowered to speak up if they're unsure of a technology's intent, potential for harm, or perhaps aren't even clear on how it's supposed to work?

My conversation with Marc Leone, Head of Trust and Safety at GIPHY, helps to shine some light on how creating an environment where technology, and particularly its privacy and safety concerns, can benefit from a proactive approach to governance that may lessen the impact of individual apathy. Prior to GIPHY, Leone led trust and safety initiatives related to content moderation, misinformation prevention, and dispute resolutions at Meetup, a community-oriented social platform. Leone's leadership style stems from an approachable, friendly demeanor that's concerned, first and foremost, with the lives of actual people; both his teammates *and* the end-users of the products he works to keep safe. It's no surprise, then, that he believes the key to responsible tech often requires embedding safety mechanisms into the product's infrastructure early on, rather than waiting for the results of an unpredictable release from which to draw conclusions.

"We're not just reacting—we're trying to prevent harm before it happens." This is Leone's main point when I ask him about how he's trying to leverage safety mechanisms at GIPHY. He explains that in the beginning of his time at GIPHY, a key concern was the infrastructure the company lacked for content moderation. He describes his early work as being "all self-policed, basically, where people would upload their content, rate it themselves, and that would determine whether people would report something if it wasn't okay [to them]." But in his view, this approach left a lot to be desired and highlighted the limitations of reactive content moderation. In the many

years that he's been working on moving GIPHY's content moderation capabilities forward, he's since started relying on more proactive moderation techniques.

One concrete suggestion from Leone, to be proactive and deal with the "don't-know-what-you-don't-know" mentality in emerging tech, is to seek out experts and partners. For him, this was a critical part of how the company moved from reactive content moderation to proactive moderation—by partnering with a third-party technology company that relied on best-in-industry machine learning models to proactively identify abusive content on GIPHY. In his words, "…putting machine learning and hash matching into the pipeline to be able to scale and increase the accuracy and quality (and the scope) of the [content] moderation program was critical. We [needed to] focus a lot on child safety and terrorist or violent extremist content. So we partnered up [with experts]." Here, he highlights what I believe is an invaluable suggestion for avoiding some of the indifference that can come from internal PMs and engineers who may be out of their league when making important decisions: if you don't have the capability in-house, look externally.

It's exactly this sort of reliance on experts and trust in the partners he surrounds himself with that brings me to one of the biggest takeaways from our conversation. When you have a leader who makes informed decisions by relying on partners, it's a testament to the culture of the organization. There's a sense of psychological safety that PMs and engineers in isolated companies don't benefit from when they're unable to consult external expertise. In "It's Just a Prototype," it was clear that the engineers who were building the health tech apps were not even experts in their *own* algorithms. Could a third party, or another team internally, have brought a better perspective? Was this self-reliance and apathy towards the product's success a part of the culture that was already ingrained in the organization? In both cases, it seems likely that the answer is yes.

In another study published in 2023, "It's about power: What ethical concerns do software engineers have, and what do they (feel they can) do about them," hundreds of software engineers were surveyed or interviewed in an attempt to understand how power dynamics in organizations affect individuals' abilities to enact change. A major takeaway from this study complements Marc Leone's thoughts, suggesting that the culture of an organization does indeed play a big role in the extent to which individual employees feel empowered to speak up. In the study, respondents "described how their organization's culture—including norms, expected practices, and communication styles—affected their willingness to raise concerns."

Importantly, respondents shared "trust and respect and an open-door policy [...] with execs" as organizational characteristics that empowered them to raise concerns. On the other

hand, an even larger share of respondents stated that “hostile, authoritarian/passive aggressive management style [and] hierarchical culture” made them feel less empowered to raise concerns.[5] This shows how, even across a variety of different tech environments, whether content moderation at GIPHY or the behavior of health tech algorithms, engineers’ and individual contributors’ sense of empowerment to make responsible decisions often depends on the culture of the organization.

There’s not necessarily a silver bullet that will yield responsible technology releases that are concerned, foremost, with the wellbeing of end-users and society at large. However, there *are* indicators from Mitchell, Leone, and the highlighted studies that the MVP culture that has dominated the SDLC in the tech industry for decades needs to be reevaluated. They make it clear that acknowledging to whom, exactly, organizations concern themselves with first, is key to appropriately injecting some ethical awareness into the process. This approach appreciates the hierarchy in many tech companies, and the powers they’re beholden to—whether shareholders, users, regulatory entities, or the organization’s own employees. At the same time, I believe engineers and PMs are most effective in an environment where they can challenge the status quo, seek outside expertise, and do so in a place where they feel psychologically safe. Perhaps Mitchell said it best: “The harm often happens because no one stopped to ask what could go wrong.”

[5] David Gray Widder, et al. “It’s about Power: What Ethical Concerns Do Software Engineers Have, and What Do They (Feel They Can) Do about Them?” *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT ’23)* [Chicago], 2023, https://doi.org/10.1145/3593013.3594012.

Interviewees

1. Thor Mitchell, CPO-in-residence at Balderton Capital, Former PM at Google
2. Marc Leone, Head of Trust and Safety at GIPHY

Footnotes

Kravets, David. “An Intentional Mistake: The Anatomy of Google’s Wi-Fi Sniffing Debacle.” *Wired*, 2 May 2012, https://www.wired.com/2012/05/google-wifi-fcc-investigation/.

Vakkuri, Ville. “‘This Is Just a Prototype’: How Ethics Are Ignored in Software Startup-Like Environments.” *Agile Processes in Software Engineering and Extreme Programming: 21st International Conference*, vol. 383, 2020, pp. 195–210, https://doi.org/10.1007/978-3-030-49392-9_13.

Verma, Pranshu and Tiku, Nitasha. “AI Employees Warn of Technology’s Dangers, Call for Sweeping Company Changes.” *The Washington Post*, 4 June 2024, https://www.washingtonpost.com/technology/2024/06/04/openai-employees-ai-whistleblowers.

Widder, David Gray, et al. “It’s about Power: What Ethical Concerns Do Software Engineers Have, and What Do They (Feel They Can) Do about Them?” *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT ’23)* [Chicago], 2023, https://doi.org/10.1145/3593013.3594012.

Wong, Matteo. “The Entire Internet Is Reverting to Beta.” *The Atlantic*, 18 June 2025, https://www.theatlantic.com/technology/archive/2025/06/ai-janky-web/683228/.