QuantumRise

GOVERNANCE

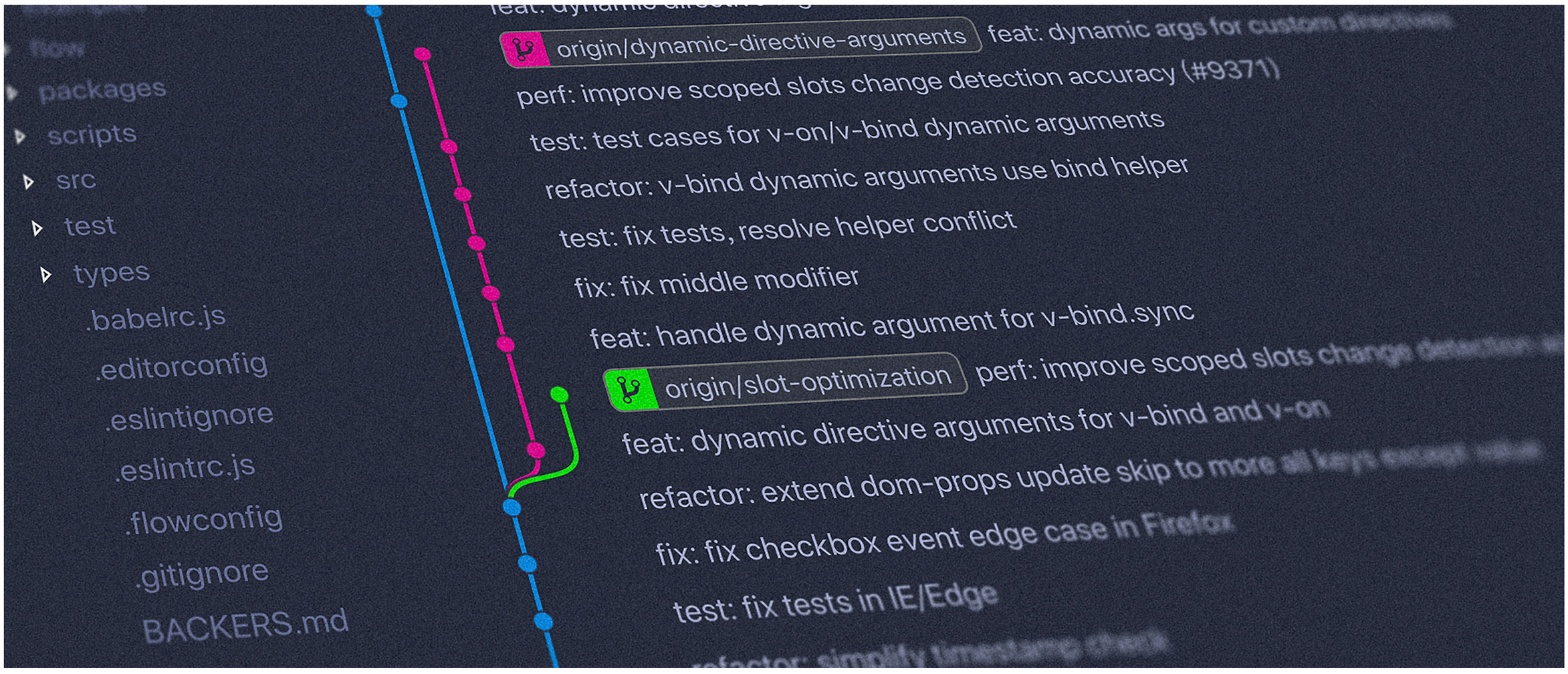
Most AI ethics work is theater

Here’s the real situation today at mid-market companies.

Julian Berman
Vice President, Client Innovation

John Swift
Director, Data Engineering

Visit www.quantumrise.com to learn more.



Here’s the real situation today at mid-market companies.

Most mid-market companies we come across have little, or nothing, to prove they have done Artificial Intelligence ethics work. All of them are having conversations with us because they want or need to deploy AI. But they frequently admit that their coverage of AI ethics and the governance is not even a serious conversation yet.

It would be invigorating if someone shared a screen listing the dead internal AI projects. Ultimately, the goal is to stop unethical AI before it occurs, but starting with AI monitoring would help. Governance is required to review and stop AI projects and usage before they are allowed to go forward, and that means humans actively engaging with what it means to be ethical with AI.

Most mid-market companies have not done this yet, which means they are essentially engaged with ethics theater. Delivering anything short of a governance process that says ‘no’ at least some of the time is like rehearsing without ever staging the play.

Why ungoverned AI is risky

There is a real risk in mid-market companies of AI decisions happening out of sight, and on entirely different criteria than what is on the AI ethics principles statement in a board presentation deck. A performative pretense of AI governance, staged for the benefit of the board, the auditor, the cyber insurance carrier, and the conscience of the company’s leadership creates the potential for massive legal jeopardy.

Laws are still trying to catch up to AI, but data residency, PII, PHI, GDPR, and other existing regulatory risk can easily be violated by an undisciplined, ungoverned AI choice.

After observing how governance and ethics play out in businesses heavy with meetings designed to discuss, and light on empowered decision-making, we don’t think this is anyone’s fault, exactly.

The structures available to mid-market companies for managing risk were built for different kinds of problems. Compliance frameworks handle well-defined and measured questions: did we break a specific law, did we submit a false filing, did this product injure this customer. AI ethics is mostly about questions that don’t have these bright lines and obvious metrics.

- Should we let the sales team paste customer call notes into a consumer AI tool because it summarizes them beautifully?
- Should we deploy a hiring filter from a vendor who can’t explain how it scores or filters candidates?
- Should we ship an AI feature that works for most customers and behaves erratically for a few?

These questions describe risks. Of course, policies need to exist but that's not enough. Monitoring and enforcement are needed to make sure the answer means something. What does not help is having a chat tool produce a statement of principles. This only appears to answer a question without actually doing so.

The reality of operating in the US mid-market makes this policy-process-enforcement cycle harder in a particular way. A Fortune 500 can afford to staff an AI ethics function, even if that function is constrained. A $300M company usually cannot spare the human capital. AI governance lands on whoever has available time, which could be the CIO, the General Counsel, or a director who already has back-to-back meetings all day. The result is a policy that exists on paper, lacks accountability, process and the capacity to reject anything.

Self assessment

We put together a list of signals that your AI ethics is theater. These common patterns we see in AI governance appear as natural outcomes of diligent work that is disconnected from the AI work it is supposed to govern.

Consider this list a guide to see whether your company is being theatrical or real with AI ethics.

- The AI policy exists but no one references it in a meeting or uses it to alter a project;
- Vendor due diligence asks the right questions, but nobody verifies models, bias, or training data sets;
- The model- or vendor-evaluation happens before purchase, never again;
- Prohibitions of AI use are ignored by employees, introducing risk;
- The "human in the loop" is whoever happens to be available, measured on speed, so flagging a problem works against their own numbers;
- No one can produce a current inventory of which AI tools and models are in use; the official list and the truth parted ways long ago;
- An AI output issue was solved with a patch, but the policy that allowed the problem to occur was never updated.

Real AI governance and ethics

What would meaningful AI governance look like in a mid-market company? We will give you our honest answer. We have not seen it in the wild yet. The answer is still somewhat theoretical.

What we do have a strong opinion on is that anyone who claims to be certain what mature AI ethics looks like is selling something. That said, there are a few features that distinguish effective governance, and they translate equally to companies with $200M and $20B revenue.

Governed AI ethics:

Has a kill switch and uses it

Saying 'no' is not theoretical. If nothing has ever been killed or refused, the function is advisory and should at least be honest about that.

Has purchasing authority

The person responsible for managing AI risk needs to be able to block a contract, require a security addendum, or hold a deployment until conditions are met.

Involves people who can lose something

The review group includes the executive whose bonus depends on the AI rollout, the operating leader of the affected business unit, and legal in a serious review position.

Is specific

Governance and ethics demand to know what data the vendor trained on, whether the vendor will indemnify outputs, whether the system can be audited, and what the rollback plan is if it produces something the company has to defend. Specific questions force specific answers.

Is documented at decision time, not at audit time

The decisions about acceptable risks, approvals, and what triggers a roll-back are all written down before the AI functionality ships or a contract renews. Documentation created when requested is performative.

The obvious objection

Every week someone tells us, “This sounds expensive and slow.” That is a fair concern for a company that does not have these skills or people to spare, but it does not have to be expensive or slow. The cost to establish a governance program with authority is likely higher up front. This is typical of many infrastructure investments, and AI ethics is exactly that. The benefits can show up quickly, however. Governed AI initiatives can actively prevent wasted effort, time, and expense (tokens, defense, and even lost customers).

If you do this work now, while the stakes and volume are still manageable, you will be nearly alone in your ability to speak substantially about your AI ethics and governance program. Everyone else will be doing a theatrical song and dance routine.

Governed AI initiatives can actively prevent wasted effort, time, and expense

Here’s our advice: Make AI policies shorter and enforce specific review and monitoring processes. State the principles and support them with enforced thresholds. Then you can talk about the programs you’ve shut down because they didn’t meet your standards.

We have not seen many mid-market companies getting far into that transition yet, but we have seen some large enterprises who are. We are helping a handful of mid--market companies build things like agent registries, and agent monitoring tools. The companies leading the way on AI ethics and governance now are going to be leading their industries sooner than they realize.

Julian Berman

Vice President, Client Innovation

View LinkedIn profile

Julian is the VP of Client Innovation at Quantum Rise. Julian brings deep expertise in combining technology with business innovation, with a solid focus on outcomes. He has over a decade of experience in designing, building and operating high-volume, revenue critical AI and machine learning systems, and in leading others in doing so across multiple functional areas. He previously held roles as Vice President, Decisioning Automation at Deloitte Digital, where he led all aspects of the machine learning offering for large enterprise customer experience clients, and at Magnetic where he lead innovation of real-time systems for advertising. He also served in an Open Source Software fellowship role at Postman, and is an adjunct professor at Columbia University where he teaches in the MBAxMS program. He holds a B.A. in mathematics (and loves trying to keep it up as a hobby).

John Swift

Director, Data Engineering

View LinkedIn profile

John is the Director of Data Engineering at Quantum Rise, where he leads a team of talented engineers to architect, build, and deliver productized data services. His passion is to deliver governed business insights at scale by combining industry best data practices, statistical process controls, AI/ML, data quality, and data governance. Previously, John was a Principal Consultant and Data Strategy & Data Management leader at Analytics8. Through a nearly three decade career, John’s insatiable curiosity and love of automation drove him to deep understanding of all phases of the data lifecycle and to deliver data-driven business insights through many different roles and industries using a wide array of technologies.