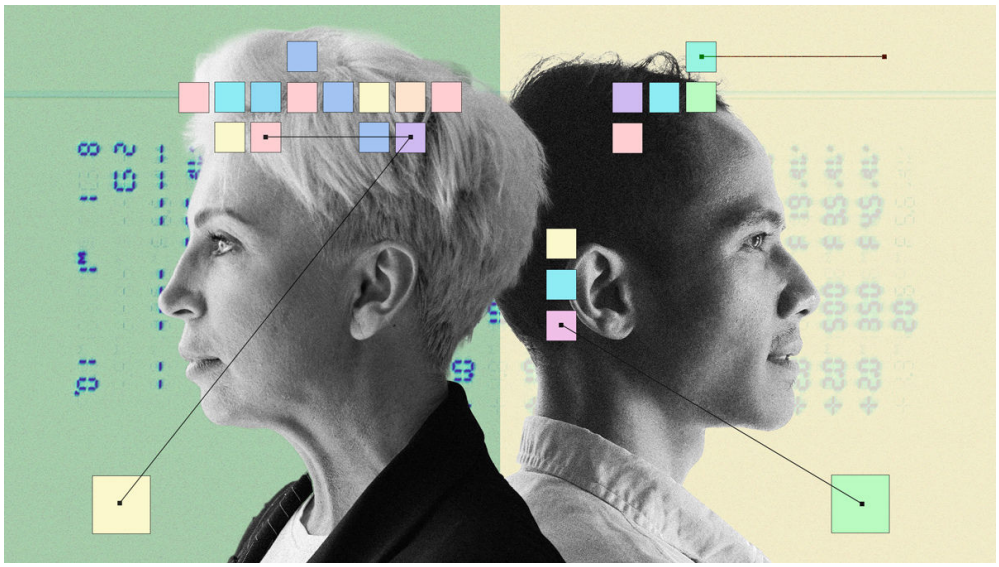
HARVARD BUSINESS REVIEW

Digital Article / Business Models

AI Is Changing the Structure of Consulting Firms

Generative AI tools, predictive algorithms, and synthetic research platforms are increasingly able to do the work of junior consultants. *by David S. Duncan, Tyler Anderson, and Jeffrey Saviano*

Published on HBR.org / September 10, 2025 / Reprint H08VTP



HBR Staff/Pexels

The debate about AI's impact on consulting tends to swing between two extremes: Some argue that AI will render consultants obsolete; others claim it will only make them more indispensable. Both views miss the more nuanced and important reality: Consulting isn't disappearing; it's being fundamentally reshaped.

For decades, the industry has operated on a stable "pyramid" model: a wide base of junior consultants handling research, modeling, and analysis and supporting a narrow apex of senior leaders who guide strategy and manage client relationships. This pyramid structure has powered the economics of consulting and shaped the profession's identity.

But AI is upending that model. Generative AI tools, predictive algorithms, and synthetic research platforms are rapidly automating the very tasks that once filled junior consultants' weeks. As this transformation accelerates, consulting firms face a choice: evolve their delivery model or risk becoming irrelevant.

As leaders of an AI-native consulting firm, Disruptive Edge, two of us, David Duncan and Tyler Anderson, are experimenting with this new model, exploring how AI tools can help with a wide range of consulting tasks—from automating routine research to augmenting more advanced analysis and synthesis. The aim is to see how these tools can speed delivery, sharpen the quality of insights, and free consultants to focus on the things that demand judgment, creativity, and deeper engagement with clients. Along with our third coauthor, Jeffrey Saviano, an expert in the emerging fields of AI governance and ethics, we're also thinking through how critical components of consulting firms, such as talent development and governance, need to evolve in these new models.

How the Pyramid Is Crumbling

The core reasons companies hire consultants are unlikely to go away any time soon. Organizations will continue to need external, specialized expertise, flexible capacity, credible validation, and an independent perspective on complex problems.

But while the demand remains, AI is undermining the traditional model consulting firms use to fulfill it. This model depends on junior consultants spending weeks on tasks such as gathering data, analyzing it, modeling scenarios, and crafting slides to support senior consultants-led recommendations. Today, AI systems can do all of that, and more—and they do it faster, more cheaply, and, in many cases, better.

For example, McKinsey's proprietary AI assistant, Lilli, is now used by over 72% of its workforce, reducing research and synthesis time by around 30%. The Boston Consulting Group is using Deckster, a tool that creates presentation decks in minutes, and Bain has deployed Sage, an AI copilot trained on its internal IP. Applications of agentic AI are also on the rise, with examples like Deloitte's Zora AI agents and PwC's agent OS platform reshaping internal workflows and client offerings. Across the board, generative AI is increasingly performing the work usually handled by large teams of junior consultants.

The work being automated isn't trivial. It already includes tasks that are the cornerstone of lower-level consulting roles and is encroaching on the tasks of the middle tier as well. If AI takes over work that used to justify thousands of billable junior hours, the pyramid will collapse under its own weight.

The Rise of the Consulting Obelisk

In response, we propose that a new model is emerging: *the consulting obelisk*. Unlike the traditional consulting pyramid, which depends on a wide base of junior consultants, the obelisk represents a tall, narrow model: fewer layers, smaller teams, and more leverage at every level. Instead of relying on sheer scale, this model is built around three human roles:

- **AI facilitators** are early-career consultants trained in the latest AI tools and data pipelines. They design and refine AI-driven workflows, helping teams generate insights at speed. This role offers a new kind of apprenticeship that emphasizes technical fluency and applied judgment from day one.
- **Engagement architects** are experienced consultants who lead projects. They help define the problems to solve, interpret AI outputs with human judgment, and translate them into actionable strategies. They orchestrate how work gets done and adapt as conditions change, while ensuring that insights turn into results.
- **Client leaders** focus on the long game. They cultivate deep, trusted relationships with senior executives, help them make sense of change, and stay close enough to advise them on how to stay ahead of disruption.

We identified these roles based on the three functions that are essential for any consulting firm to operate: Someone must build and implement the workflows that generate outputs for the client problem at hand, someone must lead the work and translate it into recommendations, and someone must build and sustain trusted executive relationships. Together, these roles also create a natural pipeline for talent development—a pipeline that must be sustained even in the age of AI.

The obelisk model balances these roles across levels, reflecting a shift away from scale for its own sake. What matters now is delivering sharper thinking with greater speed and less overhead. More than just a cost-efficient response to automation, the obelisk represents a necessary evolution of how consulting talent is structured and deployed. As AI takes over routine tasks, human energy can be reallocated to what matters most: insight, judgment, and trusted partnership.

A New Wave of AI-Native Boutiques

Some of the clearest examples of the obelisk model aren't coming from the big firms but from a fast-growing wave of AI-native boutiques. Monevate, a firm focused exclusively on pricing strategy, combines deep expertise with AI-enabled playbooks and modeling tools to deliver advice without a traditional analyst layer. SIB, which specializes in cost reduction, uses AI agents to scan invoices and vendor contracts for savings opportunities, deploying human experts only when needed. Both firms avoid the pyramid structure entirely, delivering focused, repeatable value with fewer people and far less overhead.

Unity Advisory, launched by former partners of the Big Four professional services firms and backed by $300 million in private capital, represents perhaps the most deliberate reinvention of the consulting model at scale. It positions itself as conflict free, meaning not encumbered by audit-advisory entanglements or client conflicts, and AI-native by design. Rather than build a traditional pyramid, Unity relies on agile pods of senior consultants who work in close coordination with proprietary AI tools to deliver high-speed, high-quality strategy support. It doesn't hire large, entry-level analyst cohorts or rely on hierarchical, middle management structures; it has eliminated the classic leveraged pyramid of billable hours in favor of fast, expert-driven delivery, a textbook expression of the obelisk in action.

Disruptive Edge is applying similar principles. For example, rather than have junior consultants spend weeks building foundational knowledge, we kick off engagements with AI-powered deep research reports that synthesize vast amounts of information quickly. Tools like the AI-powered app-development platform Lovable allow teams to move from concept to fully functional prototypes in under two weeks rather than months. This enables us to staff assignments with smaller,

more senior teams and spend more time with clients where it matters most.

While research on traditional firms confirms that AI tools can perform many of the tasks that once occupied large teams of junior consultants, often with gains in speed and quality, the relative newness of the obelisk model means long-term evidence is still to emerge. However, we believe that, if foundational tasks can be automated or accelerated, then smaller, more senior-heavy teams can focus more on the areas where human expertise creates the most value.

Why Firms Will Resist Change

Despite the forces causing this shift, many traditional consulting firms will struggle to make the leap. As Clayton Christensen explains in *The Innovator's Dilemma*, incumbents rarely disrupt themselves, especially when their existing model is still printing money. In consulting, that model is the highly profitable pyramid.

The pyramid model has long shaped consulting firms' culture, economics, and delivery. Promotions, compensation, staffing models, and even the mental model of what "good consulting" looks like are all wired around headcount and leverage, containing powerful incentives to maintain large teams of junior consultants.

This is precisely what makes change so hard. Even as firms invest in AI tools or announce AI innovation labs, those capabilities often remain siloed from core delivery. Flashy demos may impress clients, but the underlying engine— large project teams staffed with junior talent— remains largely untouched. Moving to a leaner, AI-augmented structure can feel like an existential threat to firms built on scale.

There's also a talent realignment challenge. Traditional firms are built to recruit and train generalist MBAs by the hundreds. But the future demands something different: smaller cohorts fluent in AI tools, data workflows, and systems thinking. Some have launched upskilling programs (PwC, for example, has committed $1 billion to AI training) but culture and incentives will naturally lag, with systems still favoring hours billed over insight delivered.

These symptoms all point to a deeper issue: treating AI as a tool to bolt onto the old model rather than a reason to re-architect it from first principles. In the short term, that may preserve margins. But over time, it opens the door for smaller, AI-native firms to move faster, operate leaner, and offer clients more value with fewer people and at lower cost.

Implications for Consulting Firms

The implications of this shift are profound. Consulting firms that continue to rely on junior-intensive models risk becoming slower, more expensive, and less relevant. Those that act decisively can emerge leaner, more expert-driven, and far more valuable to clients.

For incumbent firms, transitioning to the obelisk model will not be easy. Entire workflows will need to be redesigned around AI to redefine how value is created and delivered. Training and development of junior talent will need to evolve to integrate AI facilitation skills with more traditional consulting skills like problem-solving, communications, and client management. And compensation models must evolve, rewarding strategic contribution and client outcomes over billable hours.

This new model also raises important questions around AI governance and ethics. In traditional consulting, client deliverables typically passed through layers of review by analysts, senior consultants, managers, and partners, making it easier to catch issues and assign responsibility. But

in the obelisk model, where small teams move quickly and AI plays a larger role in decision-making, new approaches will be needed to ensure that AI-powered decisions are understandable, equitable, and made by people who are clearly accountable.

One of us (Jeffrey Saviano), is a former EY partner who now leads a research team developing new models for AI governance and ethics at Harvard's Edmond & Lily Safra Center for Ethics, an institution focused on strengthening research and teaching on ethical issues. Its work emphasizes the need for business leaders to take responsibility for governing AI themselves versus waiting for government regulation. That means building ethical guardrails into how AI is used, not relying solely on centralized compliance teams or after-the-fact reviews. We've seen how this is especially important in the obelisk model, where ethical accountability must be clear, distributed, and embedded directly into team workflows—especially as small expert teams are empowered by tools capable of influencing high-stakes decisions.

Despite the challenges for consultancies, this is not a moment for incrementalism. The winners will be those who move first and reimagine the industry before someone else does it for them.

This article was originally published online on September 10, 2025.

David S. Duncan is a partner at the consulting firm Disruptive Edge, where he advises organizations on growth strategy, innovation, and AI adoption. He is the author of *The Secret Lives of Customers* and coauthor (with Clayton M. Christensen Karen Dillon, and Taddy Hall) of *Competing Against Luck*.

Tyler Anderson is the CEO of Disruptive Edge and Aucctus AI, where he helps leading enterprises build new growth businesses and deploy AI solutions. His firms have been recognized as some of Canada's fastest-growing companies. He is the recipient of Canada's Young Entrepreneur of the Year award.

Jeffrey Saviano is the business AI ethics leader at Harvard's Edmond & Lily Safra Center for Ethics and also holds academic appointments at MIT and the Boston University School of Law. He previously was a partner at EY, where he led the firm's Emerging Technology Strategy & Governance practice.