California Management Review

INSIGHT

Agentic AI

How to Run the World: Architecting the Agentic Twin Economy

Olaf J. Groth, Tobias Straube, and Dan Zehr

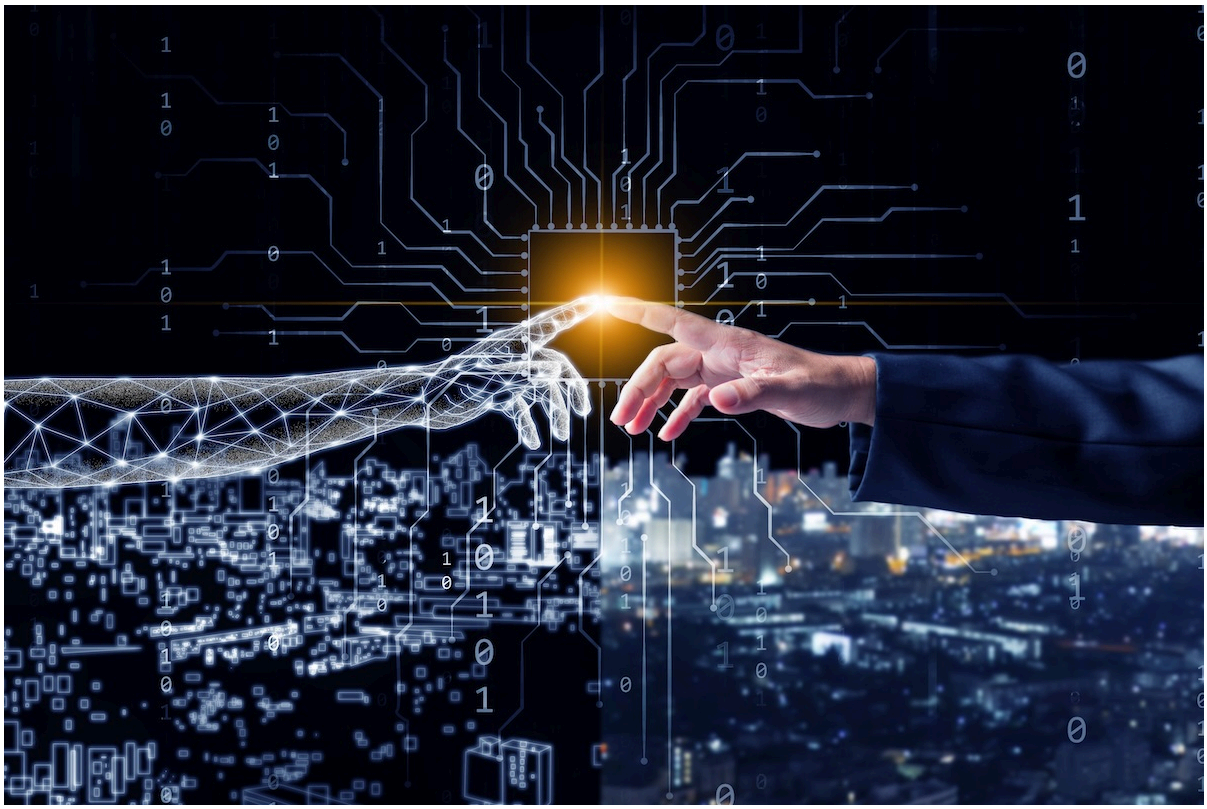

Image Credit | OleCNX

Anticipating the impact of digital clones of real-life systems

INSIGHT | FRONTIER 03 Dec 2025

The struggles to extract measurable value from generative AI applications continue to plague most companies. A recent IBM survey of corporate CEOs found that only a quarter of such projects actually deliver clear business outcomes.[1] While billions of dollars continue to pour into AI projects, startups and research, lingering concerns about ROI have started to cloud both internal and external investment decisions.

RELATED ARTICLES

Mohammad Hossein Jarrahi and Paavo Ritala, "**Rethinking AI Agents: A Principal-Agent Perspective**," *California Management Review Insight*, July 23, 2025.

Terence Tse, "**Agentic AI 101,**" *California Management Review Insight*, May 15, 2025.

RELATED TOPICS

Artificial Intelligence
Business Models
Planning & Forecasting
Data Management

However, those same concerns have obscured a crucial evolution occurring in the digital economy today. The advent of a cloned economy is underway, one run by "agentic twins." These digital clones of real-life systems, organizations and people can act autonomously to execute complex tasks, transact and make decisions on behalf of individuals and companies. While residing, interacting and evolving in the digital realm, these agentic twins will have real-life consequences, potentially driving ROI in ways that companies, focused on generative AI returns, have yet to imagine and embrace.

The hype around the first wave of this evolution – Agentic AI – has exploded over the past year. These sophisticated systems operate atop LLMs, integrating memory, tool use, protocols, and advanced autonomous reasoning frameworks to automate an array of operational roles – handling HR requests, routing service calls, managing calendars, submitting expense reports, drafting compliant legal documents, and so on. Crucially, they don't just generate content. They initiate actions across entire enterprise environments, signaling a pivotal shift from AI as a passive assistant to AI as an active economic actor. Jared Friedman, managing partner at Y Combinator, has called for innovators to build full-stack AI-native companies – an entire legal practice, for instance, operated end-to-end by a network of agents, not just aided by them for isolated tasks.[2] In this future, traditional boundaries between tool, user, and organization dissolve. Consider AI-managed traffic agent networks optimizing urban flows in New York or London, orchestrating everything from dynamic signal timing to emergency vehicle routing. This goes far beyond controlling a few traffic lights. It involves an intelligent, autonomous system that manages an entire city's transportation network. To be sure, we're not there yet. Most AI agents still lack robust causal reasoning and cross-domain generalization and are not yet fully aligned with human values. But innovators around the world are striving to solve these problems with very large sums: In 2025 alone, nearly $2 billion was invested in startups specifically building agentic or autonomous AI systems, almost matching the total for all of 2024, to say nothing of the $400 billion the large AI platforms expect to spend on enabling AI infrastructure next year. The AI agent market is projected to grow from $3.7 billion in 2023 to $103.6 billion by 2032.[3]

This immense economic value will spawn an interconnected set of infrastructure capabilities that extend beyond the AI system itself and provide the linkage points in organizations that alleviate friction and to begin to unlock value. For example, agents cannot function without an interface to connect them with existing digital environments. One such interface, called front-end emulation, mirrors human interaction, with agents simulating keyboard inputs and mouse clicks like virtual workers. This approach has already spawned successful devices like the Rabbit R1,[4] but a more groundbreaking path lies in back-end integration. In this approach, agents communicate with other digital services directly via APIs and protocols, such as Anthropic's Model Context Protocol (MCP) and Google's Agent-to-Agent Protocol (A2A). This lays the groundwork for interoperable agent ecosystems.

But this creates a critical uncertainty that's just over the horizon and that will in turn likely shape the evolution of Agents: Regardless of front- or back-end interfaces, access equals exposure. As agents embed deeply into commercial systems, they generate significant cybersecurity risks. Consider the example that Meredith Whittaker, president of Signal, provided at SXSW 2025. An agent booking a concert ticket and sharing it with friends needs access to your browser history, credit card, calendar, contacts, and messaging app – near-total access to your digital life, Whittaker explained. This demands root-level access that means organizations will have to bypass traditional permission frameworks.[5] Furthermore, processing often occurs in the cloud, heightening exposure, and agents need plain-text data access, potentially undermining privacy safeguards. What begins as convenience can quickly become a radical reconfiguration of our digital trust architecture. And herein lies the next horizon: These heightened cybersecurity concerns will usher in a second wave of agentic twins with trust at their very core. As AI agents grow autonomous and impact the real world, verifying their identity, authorized actions, and behavior becomes critical. A robust trust layer, underpinned by persistent digital identities for agents and verification systems (and ideally rooted in spoofing-resistant cryptographic keys), has taken center stage among the growing cast of researchers and commercial innovators.[6] Agents equipped with machine-readable charters – clear operational parameters defining what they can and cannot do – will become the central players in this cloned economy. Comprehensive audit trails that chronicle an agent's actions and decisions will enable the transparency and accountability necessary to foster trust and provide recourse for when that trust is breached. And when errors occur, robust revocation mechanisms – digital kill switches – allow people and organizations to suspend, reprogram or remove agents in real-time.

But this can only work if agents are fused with digital twins – those high-fidelity simulations of real-world systems, from human bodies to urban infrastructure that allow judgments of charter compliance. So far, digital twins have been on a development path of their own, largely independent of agentic AI, as innovators have created clones of machinery, infrastructure and people to better model, understand and predict their behavior and evolutions. These increasingly sophisticated and ever-evolving twins have emerged from the combined streams of every kind of personal data we can fathom – from financial records to biometric data, browsing histories to genetic codes – and they are quickly becoming the backbone of a new underground ecosystem. In this universe of AI-

powered simulations and analyses, Human Digital Twins (HDTs) will drive a battalion of autonomous agents that act on our or our organization's behalf, sometimes without our awareness or consent. In healthcare, they model our organs to test and individualize treatments before we ever receive them. In autonomous vehicles, they train robotaxis to navigate real-world streets alongside our family sedans. In geopolitical simulations, the twins direct autonomous agents that carry out political influence operations and model cyber and biological attacks against geopolitical foes.

Data is no longer just passive information organized into convenient demographic profiles – it has become the foundation for autonomous action that carries over to real-world consequences. A digital twin of a cancer patient might direct agents to scour clinical trial databases or suggest treatment paths. A city-scale twin could automatically task agents to reroute traffic or adjust energy grids. In these cases, consequences of error or malicious interference can be grave. Misinformation could compromise lives. Misidentification could paralyze systems or compromise their financial underpinnings. Without rigorous governance via Agentic Twins the future of agents could amplify risk at the same pace it promises efficiency. We cannot leave trust to an afterthought. We need to meticulously design it into every layer of the Agentic Twin stack from its inception.

Here's what makes these emerging Agentic Human Digital Twins tick in the interest of their human originators:

1. Identity Layer: The Core of You

This is the bedrock of your twin: a verified, richly textured self-portrait. It blends your static attributes (age, education, health baselines, financial standing) with deeper psychological markers - your values, aspirations, even intellectual edge. Think of it as your sovereign narrative in code.

2. Dynamic Layer: You in Motion

Life doesn't stand still, and neither does your twin. This layer captures your world in flux - real-time data from wearables, smart homes, work platforms, and social signals. It logs your physiology (sleep, stress, movement), your digital trails (posts, patterns, productivity), and your environmental context. Advanced twins even fold in their own past actions, learning from their decisions like a co-evolving consciousness.

3. Intelligence Layer: Personalized AI for Your Goals

This is where the magic happens. AI models built on your identity and dynamic data help forecast, simulate, and steer. Whether it's nudging your health routines, optimizing your learning path, or mapping your career trajectory, these models are tuned to *your* values - and crucially, you remain in charge. Transparent. Consent-based. No black boxes.

4. Governance Layer: Your Digital Constitution

Every agentic twin needs guardrails. This layer encodes your personal charter: what the twin can and cannot do, where it should escalate, when it should self-terminate in case of ethical breaches or external compromise. Think of it as your twin's constitutional law, authored by you.

5. Executive Code Layer: Autonomy in Action

Here, your twin becomes an agent. It can make decisions, execute tasks, and interface with other systems - be it your bank, your health provider, your employer, or another digital twin in your network. It does so securely, using decentralized identity tools and zero-trust protocols, ensuring privacy and contextual control. This is the layer that lets your twin **do things**, not just think them.

Agentic Human Digital Twins

Identity Layer

A foundational record of the person's static and self-defined attributes

Dynamic Layer

A streaming view of the person's life in motion

Intelligence Layer

AI models that serve the individual's goals

Governance Layer

Defines the charter and boundaries of action

Executive Code Layer

Decision-making, task execution, and interfaces

Human Agentic Twins in Action: Agency-Assured Contracts and Transactions

Trust, by itself, isn't enough, of course. For any functional economy to securely function, trust needs to be formalized through clear, enforceable contracts, monitoring and verification. As AI agents collaborate, exchange services, or deliver outcomes, they'll require a standardized method to establish agreements, define responsibilities, and settle transactions. This involves integrated rule-making within and between agents, as well as the administrative implementation of those rules by agents themselves. Today's human-centric contracts, however, are problematic. Simply reading a legal document doesn't give an AI agent the ability to understand its complex rules. Contracts are full of "if/then" relationships and nuanced language that humans mentally untangle and piece together. They demand human judgment, negotiation, and manual steps to put them into action, none of which agents can reliably replicate.

We can address this dilemma with smart contracts for agents. These structured, machine-readable agreements are designed to automatically execute when predefined conditions are met. For autonomous agents, they offer an elegant solution to codify cooperation, significantly reducing the need for human intermediaries and manual enforcement for the executable portions of agreements. Using smart contracts in combination with the identity and charter embedded in the digital twin blueprints, Agentic Twins can know which types of transactions are within acceptable parameters and to the benefit of the human they represent. They can securely open digital wallets (particularly those native to blockchain networks), exchange digital assets or services directly, and precisely encode engagement terms – e.g. deadlines, deliverables, and payment triggers – into executable code.

To be sure, smart contracts are not yet a full substitute for human legal systems, as they struggle with ambiguity, external real-world verification, and dispute resolution for non-coded outcomes. But they provide a viable foundation for experimentation with agent-to-agent commerce, especially in digitally native environments prioritizing speed, scale, and automation. Over time, hybrid models may emerge: natural-language contracts for human stakeholders, paired with structured ontologies and logic that machines can process and

enforce. In the Agentic Twin Economy, this blend of agents, twins and smart contracts could rapidly become the new normal, enabling systems legible to us by way of our or our organizations' cloned digital personas, agency charters and executable code.

Catalyzing Agentic Twin Collaboration

Once Agentic Twins interact in established environments, creating marketplaces for their discovery, negotiation and contract formation becomes the next logical step. These digital spaces, crucial for the agentic economy, will operate separately from human platforms, with agents increasingly engaging directly with other agents. Functioning as programmable coordination hubs, these agent-to-agent marketplaces will enable seamless discovery based on skillsets, compatibility, or availability, in alignment with their human or organizational governance charters. Agents will share credentials, assess reliability via dynamic feedback, initiate collaboration through machine-readable protocols and verify reciprocal charter compliance.

However, building these marketplaces requires a rethink of network effects. For human platforms, such as social media, value grows as more people join. For Agentic Twin marketplaces, value grows not just from more agents in the mix, but from the new, complex ways they can work together, as well. Imagine specialized agents finding car parts, negotiating prices, and arranging delivery to your mechanic as soon as your check engine light illuminates. The whole system becomes far more powerful than any single agent. This goes beyond "more is better" to "more diverse and connected is smarter." While fully autonomous agent-to-agent marketplaces remain aspirational, specialized platforms are emerging, including **Agent.ai** and **Salesforce AgentExchange**. These platforms currently cater to humans who seek agent services, but projects like **Fetch.ai** are building the foundational Web3 infrastructure for direct machine-to-machine interactions, enabling autonomous agents to discover, negotiate, and transact directly via smart contracts for purposes like decentralized finance (DeFi) and supply chain automation. Over time, these ecosystems could scale dramatically, forming dynamic webs of machine cooperation and ultimately evolving into a foundational layer of the digital economy where agents fluidly engage in supply and demand – and where your organization can drive value from cutting-edge investments in agentic twins instead of the cloudy search for AI returns today.

This Agentic Twin Economy will not materialize overnight, of course, but its contours are clearly taking shape. With it, we will witness a quiet, yet profound, reordering of value, power, and trust in the digital world. This new economy isn't merely about faster tools. It's about infusing our identities and agency into autonomous actors that will make decisions, transact value, and coordinate in ways that respect us and resemble our preferences every step of the way. Their success won't hinge on intelligence alone, but on close alignment *with* us and accountability *to* us.

References

1. Sara Braun, "**These Companies Are Rolling Up Their Sleeves to Implement AI**," Fortune, June 26, 2025, accessed Sept. 19, 2025.
2. Jared Friedman, "**Full-stack AI Companies**," YouTube video, published May 7, 2025, 0:38.
3. "**AI Agents Market Set to Reach USD 103.6 Billion by 2032, Owing to Rising Demand for Automation & Customer Experience**," CMSWire, November 19, 2024, accessed Sept. 19, 2025.
4. Rabbit Inc., **Rabbit R1** webpage, accessed Sept. 19, 2025.
5. Meredith Whittaker, "**The State of Personal Online Security and Confidentiality**," SXSW 2025 speaker session, YouTube video, published March 7, 2025, 59:32,
6. "**Why AI Agents Need Verified Digital Identities**," Identity.com, accessed Sept. 19, 2025.
7. "**An Agentic AI Identity-Management Approach**," Cloud Security Alliance blog, March 11, 2025, accessed Sept. 19, 2025.
8. "**Revolutionizing Identity Management: How AI Is Transforming Agent ID Systems**," BytePlus, accessed Sept. 19, 2025.

Olaf J. Groth Follow

Olaf Groth, PhD, is a Professor of Practice at UC Berkeley Haas School of Business, Adjunct Professor at UT Malaysia, and a member of the AI Governance Alliance of the World Economic Forum. He also serves as CEO of concept studio Cambrian Labs and of Cambrian Futures, a global network of analysts, strategists and futurists for transformations of enterprises and economies driven by AI, data and emerging technologies. He is the author of two books - The AI Generation and The Great Remobilization.

Tobias Straube Follow

Tobias Straube is VP Analysis at Cambrian Futures, Assistant Lecturer at UC Berkeley Executive Education, as well as a lecturer at Steinbeis Business School. He is a co-founder of Scio-Network and previously served as an Innovation Manager for the German Society for International Collaboration (GIZ).

Dan Zehr Follow

Dan is vice president of communications at Cambrian Group and lecturer for journalism at U of Texas Austin. Prior to joining the firm, he was an award-winning business journalist whose reporting spanned most major industries and geographies. He most recently covered the economy and finance for the Austin American-Statesman. Dan has co-authored or edited eight books, including works on artificial intelligence and advanced technologies, digital marketing, social media strategies and consumer markets in developing economies.