SHAKUDO

How to Use Agentic AI in Utilities

A Practical Guide to Deploying Autonomous AI Agents for Grid Management, Asset Operations, and Customer Service

January 29, 2026

White Paper

Table of Contents

Executive Summary

Agentic AI represents a fundamental shift in how utilities operate—moving from reactive manual processes to autonomous, goal-directed systems that perceive, reason, and act with minimal human intervention. For water, gas, and electric utilities facing workforce shortages, aging infrastructure, and climate volatility, agentic AI offers a path to do more with less.

Early adopters are already seeing measurable results: 30% reductions in energy consumption through autonomous process optimization, significant improvements in mean time to restore (MTTR) during outages, and operational cost reductions of 40-60%. Unlike traditional automation that follows rigid rules, agentic AI adapts to changing conditions—coordinating distributed energy resources during peak demand, isolating faults autonomously using SCADA data, or orchestrating leak detection workflows from sensor alert to work order dispatch.

The business case is compelling. CIOs anticipate up to 179% ROI on AI investments as agentic capabilities scale. However, 73% of organizations lack comprehensive governance frameworks, and only 21% currently deploy agentic AI in production. Success requires addressing three critical challenges: integrating with decades-old legacy infrastructure, maintaining data sovereignty within heavily regulated environments, and building trust through transparent governance. Utilities that act now—with the right deployment architecture and pre-integrated tooling—can compress 6-18 month infrastructure builds into days, positioning themselves to lead in reliability, sustainability, and operational efficiency.

Overview

Agentic AI is not just another analytics tool or chatbot. It represents autonomous systems capable of perceiving their environment through sensors and data streams, reasoning through complex scenarios, planning multi-step actions, and executing those plans by interacting with other software or physical systems. In utility operations, this means an AI agent can monitor SCADA telemetry, correlate weather patterns with asset health data, predict an equipment failure, automatically generate maintenance work orders, and even adjust grid configurations—all without waiting for human approval at each step.

This capability matters now because utilities face converging pressures that manual processes cannot address at scale. The "silver tsunami" of workforce retirements is removing decades of operational expertise just as climate change intensifies extreme weather events. Meanwhile, the grid itself is transforming—distributed energy resources like rooftop solar and battery storage, electric vehicle charging demand, and aging infrastructure create complexity that exceeds human capacity to manage in real-time. Utilities are being asked to deliver higher reliability, faster restoration, and lower carbon intensity with fewer people and tighter budgets.

What makes agentic AI different from previous automation waves is its ability to reason across domains rather than optimizing isolated processes. Traditional SCADA systems monitor and control specific equipment. Advanced process control optimizes a single treatment process. Agentic AI can simultaneously consider operational parameters, energy market prices, maintenance schedules, weather forecasts, and customer impact—then orchestrate actions across multiple systems to achieve defined objectives within safety constraints.

The technology is maturing rapidly, but adoption remains early. According to recent enterprise research, 89% of CIOs consider agent-based AI a strategic priority for automation and decision-making, yet only 21% currently use agentic AI in production environments. The gap between strategic intent and operational deployment stems from real barriers: legacy infrastructure that lacks modern APIs, unstructured data trapped in proprietary formats, regulatory requirements that prohibit data from leaving controlled environments, and organizational concerns about AI decision-making in safety-critical systems.

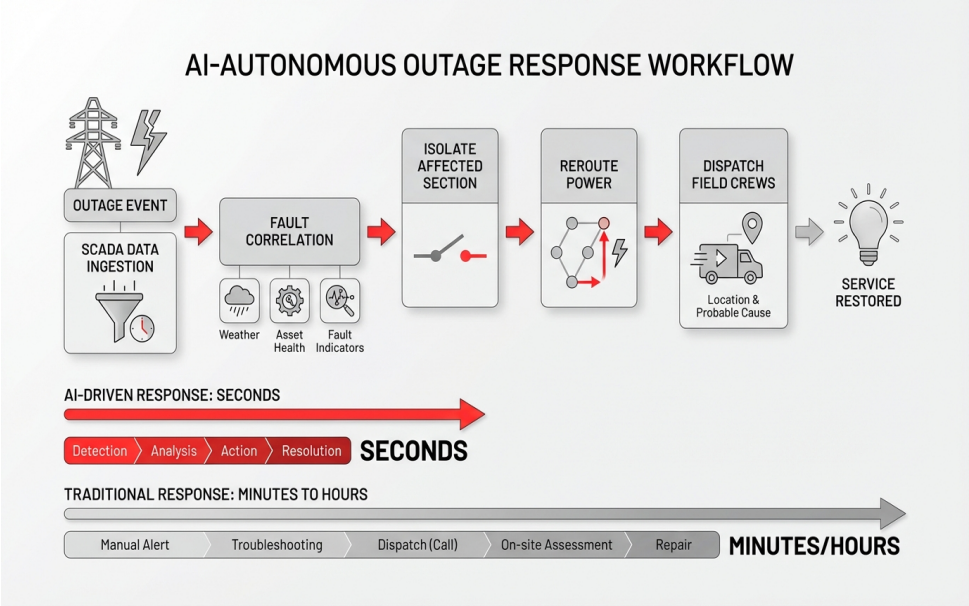
Platforms like Shakudo are addressing these deployment barriers by providing pre-integrated ecosystems of 200+ data and AI tools that can be deployed within customer environments—whether private cloud, VPC, or on-premises—in days rather than months. This sovereign deployment model allows utilities to experiment with and scale agentic AI capabilities without compromising data privacy, regulatory compliance, or operational security. The key insight is that agentic AI adoption is less about the algorithms themselves and more about the infrastructure, governance, and integration architecture that enables safe, scalable deployment within utility operating constraints.

Critical Use Cases Transforming Utility Operations

Agentic AI is already demonstrating value across the utility value chain, from generation and transmission to distribution and customer engagement. Understanding where autonomous agents deliver the highest impact helps organizations prioritize deployment and build internal capabilities systematically.

Autonomous Outage Response and Grid Resilience

Electric utilities are deploying agentic AI for Fault Location, Isolation, and Service Restoration (FLISR)—one of the most compelling use cases with direct customer impact. When an outage occurs, an AI agent ingests real-time SCADA data, correlates fault indicators with weather conditions and asset health records, isolates the affected section, reroutes power through alternate feeders, and dispatches field crews with specific location and probable cause information. All of this happens in seconds rather than minutes or hours.



Autonomous outage response workflow compresses restoration time from hours to seconds through coordinated AI-driven actions.

The business impact is substantial. Traditional outage response involves manual analysis of alarms, phone calls between dispatchers and field supervisors, and sequential troubleshooting. Agentic AI compresses this timeline dramatically, improving MTTR and reducing customer minutes of interruption—key regulatory performance metrics. Some utilities report restoration times improving by 40-50% for certain fault types.

Beyond reactive response, agentic AI enables predictive grid management. Agents continuously monitor sensor data from substations, transformers, and distribution equipment, identifying anomaly patterns that precede failures. Rather than simply generating alerts, these agents can automatically schedule preventive maintenance, order replacement parts, and even adjust load distribution to reduce stress on at-risk assets—all within predefined operational constraints that ensure safety and reliability.

Distributed Energy Resource Orchestration

The proliferation of rooftop solar, battery storage, electric vehicles, and demand response programs is transforming the grid from a one-way power delivery system into a dynamic, bidirectional network. Managing these distributed energy resources (DERs) manually is impossible at scale. Agentic AI acts as an orchestrator, balancing generation, storage, and demand in real-time to optimize grid performance, minimize costs, and maintain voltage and frequency stability.

During peak demand events or grid emergencies, an AI agent can dispatch battery storage, curtail non-essential loads through demand response agreements, adjust EV charging schedules, and coordinate with utility-scale generation—all while considering energy market prices, weather forecasts, and customer comfort parameters. This multi-objective optimization across thousands or millions of endpoints represents exactly the kind of complex, real-time decision-making where agentic AI excels beyond human or traditional automation capabilities.

For water utilities, similar orchestration opportunities exist in pump scheduling and energy optimization. Treatment plants consume significant electricity for aeration, pumping, and chemical dosing. Early implementations of AI-driven optimization have demonstrated 30% energy reductions by continuously adjusting process parameters based on flow rates, water quality sensors, and energy prices. Organizations deploying these capabilities through platforms like Shakudo benefit from pre-integrated connections between time-series databases, optimization engines, and operational technology interfaces—avoiding the 6-18 month integration projects that typically delay pilot-to-production transitions.

Predictive Asset Management and Maintenance

Utility asset bases are aging, with much critical infrastructure decades past its original design life. Reactive maintenance is expensive and risky; preventive maintenance on fixed schedules wastes resources. Predictive maintenance has long been a goal, but agentic AI elevates this from predictive alerts to autonomous action.

An agentic system monitoring gas pipeline sensors can detect pressure anomalies suggesting a developing leak, cross-reference the location with GIS data and maintenance history, assess risk based on proximity to population centers, automatically generate a work order with priority classification, route it to the appropriate field crew, and provision them with relevant asset documentation and safety protocols. The human role shifts from data analysis and administrative coordination to field execution and exception handling.

For water utilities, leak detection provides a parallel use case. Current acoustic monitoring systems place dots on a map indicating potential leaks. Human analysts must verify these alerts, eliminate false positives, prioritize based on severity and location, and initiate work orders. Agentic AI can automate this entire workflow, from sensor data interpretation through work order dispatch, scheduling trucks and crews, and even integrating with smart meter data to confirm leak repair through consumption pattern changes. The cycle time from detection to resolution can compress from days to hours.

Customer Service and Operational Support

Conversational AI agents are transforming customer interactions in utilities, handling everything from billing inquiries to service requests to outage notifications. Unlike traditional chatbots that follow decision trees, agentic AI can understand context, access multiple backend systems, reason through complex scenarios, and take actions like scheduling appointments, adjusting payment plans, or ordering meter inspections.

For water utilities, high bill inquiries provide a practical example. When a customer receives an unexpectedly high bill, an AI agent can analyze consumption history, identify usage spikes, correlate with smart meter data to detect continuous flow patterns suggesting leaks, explain findings to the customer in natural language, offer to schedule a meter inspection, and provide conservation guidance—all in a single interaction. Multilingual capabilities ensure these services reach diverse customer populations, particularly important for essential communications like water quality advisories or service interruptions.

Internally, operational teams use conversational agents to query asset databases, retrieve compliance documentation, log field observations, and access technical manuals—accelerating decision-making and reducing the learning curve for new employees facing workforce turnover. These internal agents become institutional knowledge repositories that preserve expertise even as experienced workers retire.

Navigating Legacy Infrastructure and Integration Challenges

The promise of agentic AI collides with a harsh reality in utilities: much of the critical infrastructure that agents must interact with was designed decades ago, prioritizes stability over interoperability, and lacks the modern APIs and data standards that AI systems expect. Successfully deploying agentic AI requires pragmatic strategies for bridging this gap without disrupting operations or compromising safety.

The Operational Technology Challenge

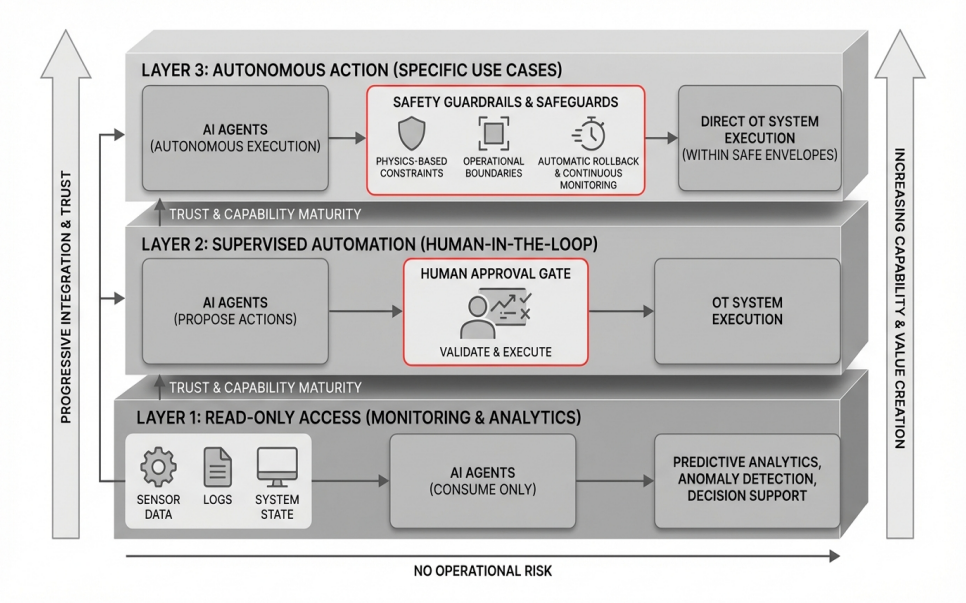
Utility operational technology (OT)—SCADA systems, energy management systems (EMS), outage management systems (OMS), programmable logic controllers (PLCs), and remote terminal units (RTUs)—often dates back 20-40 years. These systems were built for reliability and deterministic behavior, not for integration with AI agents that make probabilistic decisions. Many lack modern APIs entirely, communicating through proprietary protocols or flat file exports. Even when connectivity exists, latency, data quality, and format inconsistencies create barriers.

An AI agent designed to optimize pump operations in real-time may struggle to ingest telemetry from PLCs at remote water treatment facilities that output data in legacy formats at irregular intervals. Similarly, an agent orchestrating grid response during outages needs millisecond-level SCADA data, but older systems may batch updates every few seconds or minutes—introducing lag that undermines autonomous decision-making.

The risk of disruption compounds these technical challenges. Utility operations are safety-critical. A software integration that crashes a SCADA server or inadvertently sends control commands could cause service outages, equipment damage, or safety incidents. This creates understandable organizational caution around connecting AI systems to operational technology, even when the potential benefits are clear.

Practical Integration Strategies

Successful deployments typically adopt a layered approach that isolates AI systems from direct OT interaction initially, gradually increasing integration as trust and capability mature. The first layer involves read-only data access—agents consume sensor data, logs, and system state information but cannot issue control commands. This allows value creation through predictive analytics, anomaly detection, and decision support without operational risk.



The three-layer integration approach balances AI capability advancement with operational safety and organizational trust-building.

The second layer introduces supervised automation, where agents generate recommended actions that require human approval before execution. A grid management agent might propose a switching sequence to reroute power around a fault, displaying the plan to a human operator who validates and executes it. Over time, as accuracy and reliability are demonstrated, certain low-risk actions can be delegated to autonomous execution within predefined safety envelopes.

The third layer enables autonomous action for specific use cases with comprehensive safeguards—physics-based constraints, operational boundaries, automatic rollback mechanisms, and continuous monitoring. This is where agents truly operate autonomously, but only after extensive testing, validation, and governance framework development.

Data integration middleware plays a critical role in all three layers. Rather than requiring custom connectors for every OT system, utilities benefit from platforms that provide pre-built integrations to common SCADA vendors, time-series databases, asset management systems, and GIS platforms. Shakudo's ecosystem of 200+ pre-integrated tools includes many of these connectors, allowing organizations to establish data pipelines between legacy OT and modern AI platforms in days rather than the months typically required for custom integration projects. This acceleration is particularly valuable for pilot programs where rapid iteration and learning are essential.

Data Quality and Governance

Even when technical connectivity is established, data quality issues undermine AI agent performance. Sensor drift, missing values, inconsistent timestamps, duplicate records, and conflicting information across systems create challenges that affect autonomous decision-making. An agent optimizing energy usage based on flawed meter data might make recommendations that actually increase costs or compromise service quality.

Addressing this requires data governance frameworks focused on specific use cases rather than attempting enterprise-wide standardization before any AI deployment. For a leak detection pilot, define quality thresholds for acoustic sensor data, implement validation rules, establish metadata standards for GIS integration, and create monitoring dashboards that track data pipeline health. Expand governance scope as additional use cases are deployed.

Utilities should prioritize making data machine-readable in knowledge repositories—technical manuals, maintenance procedures, compliance documentation, and operational logs. When this information is structured and accessible, AI agents can reference it during decision-making, and human operators can be AI-assisted to be more productive. The alternative—unstructured PDFs scattered across file shares—limits agentic AI to sensor data alone, missing the rich contextual information that enables sophisticated reasoning.

Organizations deploying agentic AI within sovereign infrastructure environments maintained by Shakudo address data governance more effectively because all data remains within controlled boundaries. There's no need to sanitize datasets for external SaaS vendors, implement complex data residency controls, or navigate cross-border data transfer regulations. The AI tools, data storage, and operational systems all reside within the same governed environment, simplifying compliance while enabling sophisticated AI capabilities.

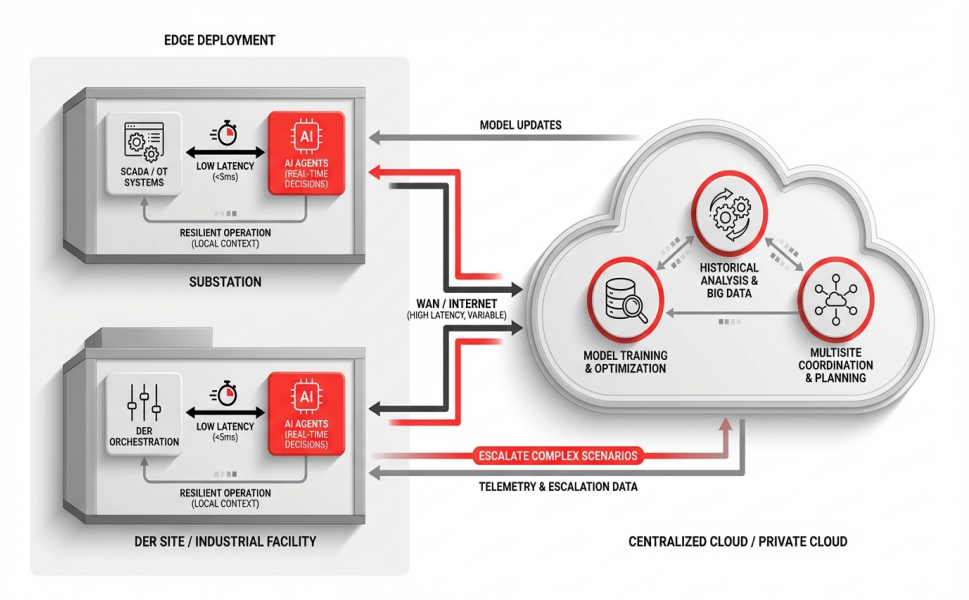
Architecture Principles for Production Deployment

Moving agentic AI from pilot to production requires architectural decisions that balance flexibility, security, scalability, and operational resilience. Utilities cannot afford experimental infrastructure for safety-critical systems, yet they also cannot wait years for perfect architectures before capturing value. The following principles guide successful production deployments.

Hybrid Infrastructure for Latency and Sovereignty

Utility agentic AI workloads span a spectrum from real-time operational control requiring millisecond latency to batch analytics and model training that can tolerate hours of processing time. A one-size-fits-all cloud or on-premises approach fails to serve both extremes effectively. Hybrid architectures that place compute close to where decisions and actions occur, while leveraging centralized resources for heavy processing, provide the optimal balance.

Edge deployment is essential for latency-sensitive use cases like autonomous grid switching or real-time DER orchestration. Agents making these decisions must run on infrastructure co-located with SCADA systems and operational technology, eliminating network latency and ensuring operation even during connectivity disruptions. These edge agents operate within narrow, well-defined parameters and escalate to centralized systems for complex scenarios requiring broader reasoning.



Hybrid edge-cloud architecture enables millisecond-latency decisions at operational sites while leveraging centralized compute for complex analytics and model training.

Centralized cloud or private cloud infrastructure handles model training, large-scale optimization, historical analysis, and coordination across multiple sites. This is where utilities aggregate data from hundreds of substations or thousands of water meters, train machine learning models that detect equipment failure patterns, and run simulations that test agent decision-making under various scenarios before deployment.

Data sovereignty considerations drive many utilities toward private cloud or on-premises deployments rather than public cloud SaaS solutions. Regulatory requirements in energy and water sectors often mandate that operational data, customer information, and critical infrastructure details remain within controlled environments. Additionally, cybersecurity concerns about exposing OT systems to internet-connected services push organizations toward air-gapped or highly restricted network architectures.

Platforms like Shakudo address this by deploying the entire AI operating system—including the 200+ pre-integrated tools for data engineering, ML operations, and agent orchestration—within the customer's VPC, private cloud, or on-premises environment. This eliminates data residency concerns, simplifies compliance, and allows utilities to leverage sophisticated AI capabilities without compromising sovereignty or security. The alternative—building and maintaining this toolchain in-house—typically requires 12-18 months and dedicated platform engineering teams, delaying time-to-value and diverting resources from use case development.

Agent Orchestration and Workflow Management

Production agentic AI systems rarely consist of a single agent operating in isolation. Complex utility operations require multiple specialized agents coordinating through orchestration frameworks. An outage response system might involve separate agents for fault detection, crew dispatch, customer notification, and load rebalancing—each with distinct responsibilities but requiring coordination to achieve the overall objective of minimizing customer impact.

Orchestration platforms manage agent lifecycles, handle inter-agent communication, maintain state across multi-step workflows, and provide observability into agent decision-making. When a grid event occurs, the orchestration layer ensures that the fault detection agent completes its analysis before the crew dispatch agent attempts to assign work orders, prevents duplicate notifications to customers, and coordinates load rebalancing only after isolation is confirmed.

Workflow management becomes particularly important when human-in-the-loop (HITL) approval is required for certain decisions. The orchestration layer must pause agent workflows at decision points, present recommendations to human operators with sufficient context for evaluation, capture approval or modification, and resume autonomous execution. As confidence grows, specific workflow steps can transition from HITL to fully autonomous within governance frameworks.

Observability throughout these orchestrated workflows is non-negotiable for production systems. Utilities need real-time visibility into what agents are perceiving, how they're reasoning, what actions they're planning, and why. When an agent makes an unexpected recommendation or takes an action that operators question, detailed logging and audit trails enable investigation and continuous improvement. This transparency builds organizational trust and satisfies regulatory documentation requirements.

Scalability and Resource Optimization

Agentic AI workloads exhibit highly variable compute demands. During normal operations, agents may monitor data streams with modest resource requirements. During grid events, outages, or peak demand periods, those same agents may need to scale instantly to process massive data volumes, run complex

simulations, and coordinate hundreds of actions simultaneously.

Containerized and serverless architectures that dynamically scale with workload demand provide the operational efficiency and cost control that utility CFOs require. Rather than provisioning infrastructure for peak load that sits idle most of the time, agents run in container orchestration platforms that automatically add compute resources during demand spikes and scale down during quiet periods.

Workload orchestration strategies further optimize resource utilization by shifting non-urgent compute tasks—model retraining, historical analysis, simulation runs—to off-peak hours when electricity costs are lower and infrastructure utilization is higher. This approach reduces operational costs while improving overall system efficiency.

Data pipeline optimization matters as well. Unnecessary data transfers between regions, redundant processing of unchanged datasets, and inefficient query patterns waste compute resources and increase costs. Organizations benefit from platforms that provide built-in pipeline optimization, compression, and caching—features that become increasingly important as agentic AI scales from pilot projects processing gigabytes to production systems handling terabytes daily.

Security, Access Control, and Cyber Resilience

Agentic AI introduces new cybersecurity considerations that extend beyond traditional IT security models. Agents are software entities that access corporate systems, query databases, initiate workflows, and in some cases control physical infrastructure. They require network access, system credentials, and permissions that create potential attack vectors if compromised.

Zero-trust architectures that treat agents as distinct identities with specific, limited permissions provide a security foundation. An agent optimizing pump operations should have read access to sensor data and write access to control systems for those specific pumps—but no access to customer billing systems, corporate email, or unrelated operational technology. Least-privilege principles and micro-segmentation limit the blast radius if an agent is compromised or malfunctions.

Continuous monitoring for anomalous agent behavior provides an additional security layer. If an agent suddenly begins accessing systems outside its normal pattern, making unusual API calls, or exhibiting decision-making that deviates from established baselines, automated systems should flag this for investigation and potentially isolate the agent pending review.

Cyber resilience planning must address scenarios where agentic AI systems themselves are compromised or unavailable. Utilities need fallback procedures, manual override capabilities, and operational continuity plans that don't assume agent availability. This might mean maintaining traditional SCADA interfaces alongside agentic control, or ensuring that human operators can step in with complete system visibility if AI systems fail.

Building Organizational Readiness and Governance

Technology and architecture are necessary but insufficient for successful agentic AI deployment in utilities. The organizational dimension—culture, skills, governance, and change management—determines whether pilot projects scale to enterprise-wide transformation or stall in the innovation theater phase.

Workforce Implications and Skills Development

Agentic AI fundamentally changes what utility workers do, shifting emphasis from routine monitoring and manual coordination to exception handling, strategic decision-making, and continuous system improvement. For some employees, this represents an exciting opportunity to move from repetitive tasks to higher-value work. For others, it creates anxiety about job security and concerns about whether they can adapt to new ways of working.

Successful deployments address these concerns proactively through transparent communication about how roles will evolve, investments in reskilling programs, and demonstrations of how AI augments rather than replaces human expertise. A SCADA operator whose job currently involves monitoring dozens of screens for anomalies might become a grid optimization specialist who works with AI agents to continuously improve response strategies, develops new decision-making rules based on operational experience, and handles complex edge cases that agents escalate.

Skills development programs should focus on three areas. First, foundational AI literacy—helping operational staff understand what agentic AI can and cannot do, how agents make decisions, and how to interpret agent recommendations critically. Second, tool-specific training on the platforms and interfaces that employees will use to interact with agents, review agent decisions, and provide feedback. Third, new analytical and problem-solving skills that become more valuable as routine tasks are automated—systems thinking, root cause analysis, and strategic planning.

Organizations should expect workforce resistance, particularly in safety-critical environments where operators have been trained to trust their own judgment and maintain direct control. Building trust requires transparency about agent decision-making, starting with low-stakes use cases where errors have minimal consequences, and celebrating early wins where agents and humans collaborate effectively. Employee involvement in pilot design, testing, and feedback loops increases buy-in and surfaces practical concerns early.

Establishing AI Governance Frameworks

Only 27% of organizations deploying AI have comprehensive governance frameworks in place—a significant gap given the risks of autonomous systems making decisions that affect public safety, environmental compliance, and customer service. Utilities need governance structures that balance innovation speed with appropriate oversight, particularly as agents gain autonomy.

Governance frameworks should define clear accountability structures—who is responsible when an agent makes a decision that leads to negative outcomes? In most cases, the answer is human leadership overseeing the AI system, not the AI itself, but this needs explicit documentation. Similarly, decision rights must be

established: which types of decisions can agents make autonomously, which require HITL approval, and under what conditions can an agent's recommendation be overridden?

Risk assessment processes should evaluate each agentic AI use case for potential failure modes, worst-case scenarios, and appropriate safeguards. A leak detection agent that incorrectly identifies a false positive wastes field crew time—annoying but not dangerous. A grid switching agent that incorrectly opens a breaker during peak load could cause cascading outages affecting thousands of customers—a scenario requiring extensive safeguards, testing, and initially, human approval for all actions.

Audit trails and explainability requirements ensure that agent decisions can be reconstructed, understood, and evaluated. When regulators investigate a service disruption or compliance incident, utilities must be able to demonstrate exactly what data the agent perceived, how it reasoned, what alternatives it considered, and why it took the action it did. This requires logging at a level of detail that balances observability with storage costs and performance.

Governance should also address ethical considerations and equity. Do agent decisions inadvertently disadvantage certain customer populations? Are response times during outages influenced by neighborhood demographics? Do optimization algorithms prioritize cost savings over environmental impact? These questions require ongoing monitoring, diverse perspectives in governance committees, and willingness to adjust agent objectives and constraints as issues emerge.

For organizations deploying agentic AI through platforms like Shakudo, governance is simplified by built-in audit capabilities, role-based access controls, and the ability to enforce policies consistently across all AI tools within the environment. Rather than implementing governance separately for each tool in a fragmented stack, unified platforms allow policy definition once and enforcement everywhere.

Pilot Design and Scale Strategies

Most utilities begin agentic AI adoption with pilots—limited scope deployments that demonstrate value, surface integration challenges, and build organizational confidence before enterprise-wide rollout. Effective pilot design balances ambition with achievability, targeting use cases where success is measurable, impact is meaningful, and failure is tolerable.

Selecting the right pilot use case matters enormously. Ideal candidates have clear business metrics that can demonstrate ROI (energy costs reduced, MTTR improved, leak repair cycle time decreased), access to quality data that doesn't require extensive cleanup, limited integration complexity with a manageable number of systems, and stakeholder support from both technical teams and business owners. Conversely, avoid pilots that require integration with highly complex legacy systems, involve politically sensitive decisions, or have vague success criteria.

Time-bound pilots with defined success criteria prevent endless experimentation. A 90-day pilot to deploy an agent for automated leak detection work order generation, with success defined as 80% accurate leak verification and 50% reduction in time from detection to work order, provides clarity. Teams know what they're trying to achieve, by when, and how success will be measured.

Once a pilot demonstrates value, scaling requires deliberate planning. Will the same use case be deployed to additional regions, or will new use cases be developed? What infrastructure investments are needed to support broader deployment? How will training and change management scale beyond the pilot team? Organizations that compress pilot-to-production timelines from 12-18 months to 3-6 months typically leverage platforms that eliminate infrastructure setup delays. When the tooling, orchestration, and integration frameworks are already in place, teams can focus on use case refinement and organizational adoption rather than platform engineering.

Regulatory Compliance and Stakeholder Communication

Utilities operate in heavily regulated environments where transparency, reliability, and public trust are paramount. Deploying autonomous AI systems that make decisions affecting service quality, public safety, and environmental impact requires proactive engagement with regulators, clear communication with customers and community stakeholders, and documentation that demonstrates responsible AI practices.

Regulators increasingly ask questions about AI decision-making in utility operations: How do you ensure that AI recommendations are accurate? What safeguards prevent AI systems from causing service disruptions? How do you maintain human accountability? Can you explain AI decisions to customers and during investigations? Addressing these questions requires preparation—developing documentation that describes agent capabilities and limitations, maintaining detailed audit trails, and establishing clear escalation paths when agents encounter scenarios requiring human judgment.

Public communication about agentic AI should emphasize benefits—faster outage restoration, improved service quality, lower costs—while acknowledging that humans remain accountable for outcomes. Customers care less about the technical details of how agents work and more about whether service improves. Framing AI as a tool that enables utility staff to serve customers better, rather than a replacement for human workers, tends to resonate more effectively.

For regulated industries facing strict compliance requirements around data privacy, operational security, and service quality, the ability to deploy AI capabilities within sovereign environments becomes a competitive advantage. Utilities using platforms like Shakudo can demonstrate to regulators that data never leaves controlled environments, that all processing occurs within compliant infrastructure, and that audit trails capture every decision—addressing regulatory concerns that might otherwise delay or prevent AI adoption.

Measuring Success and Continuous Improvement

Deploying agentic AI is not a one-time project but an ongoing operational discipline that requires measurement, learning, and refinement. Organizations that treat AI as a deployed-and-done technology fail to capture compounding value. Those that build continuous improvement cultures around their AI systems see performance gains accelerate over time.

Defining Meaningful Metrics

Success metrics for agentic AI should tie directly to business outcomes rather than technical performance alone. Model accuracy, prediction latency, and system uptime matter, but executives and board members care about operational efficiency, cost reduction, reliability improvements, and customer satisfaction. Effective measurement frameworks connect technical metrics to business KPIs.

For outage response agents, relevant metrics include mean time to detect (MTTD) faults, mean time to isolate (MTTI) affected areas, mean time to restore (MTTR) service, customer minutes of interruption (CMI), and crew utilization rates. Comparing these metrics before and after agent deployment, and tracking improvement trends over time, demonstrates value quantitatively.

For energy optimization agents in water treatment, measure kilowatt-hour consumption per million gallons treated, cost per gallon, energy cost as a percentage of operating budget, and compliance with effluent quality standards. The agent should reduce energy costs without compromising treatment quality—capturing this requires metrics across both dimensions.

For customer service agents, track first-contact resolution rates, average handle time, customer satisfaction scores, escalation rates to human agents, and operational cost per interaction. If agentic AI increases first-contact resolution from 60% to 85% while maintaining or improving satisfaction scores, the business case is clear.

Leading indicators provide early signals of agent performance before business outcomes fully manifest. For predictive maintenance agents, track the percentage of failures predicted before occurrence, the false positive rate of alerts, and the percentage of predicted failures that result in actual work orders. These indicators help teams tune agent decision-making and adjust alert thresholds before equipment failures occur.

Feedback Loops and Agent Refinement

Agentic AI systems improve through continuous learning from operational experience, but this requires deliberate feedback mechanisms. When agents make recommendations that humans override, capturing the reasons for override—and feeding this information back into training datasets—helps agents learn edge cases and improve future reasoning. When agents take actions that lead to suboptimal outcomes, root cause analysis should identify whether the issue was data quality, reasoning logic, or insufficient constraints.

Active learning approaches where agents flag uncertain decisions for human review create structured feedback. Rather than escalating every decision, agents identify scenarios where their confidence is low or where the stakes are high, asking for human input on those specific cases. This selective escalation provides

high-quality training data in the areas where agents most need improvement, without overwhelming operators with constant approval requests.

Model retraining schedules should balance the need for continuous improvement with operational stability. For some use cases, weekly or monthly retraining with recent operational data keeps agents current with changing conditions. For others, quarterly retraining with careful validation provides sufficient currency without introducing instability. The key is having infrastructure and processes that make retraining routine rather than a major project each time.

Organizations deploying agentic AI within integrated platforms like Shakudo benefit from MLOps tools that automate model versioning, A/B testing, gradual rollout, and rollback capabilities. When a newly retrained agent model is ready for deployment, teams can test it against the current production model on a subset of traffic, compare performance metrics, and promote the new model only if it demonstrates measurable improvement. This systematic approach to continuous improvement reduces risk while accelerating learning cycles.

Cost-Benefit Analysis and ROI Tracking

Executive support for scaling agentic AI depends on demonstrating return on investment that justifies continued platform investment and expansion to new use cases. Comprehensive ROI calculations should capture both direct cost savings and indirect value creation, while also accounting for implementation costs and ongoing operational expenses.

Direct cost savings include reduced labor costs for automated tasks, lower energy consumption from optimization, decreased equipment repair costs from predictive maintenance, and reduced regulatory penalties from improved compliance. These are typically the easiest benefits to quantify with financial precision.

Indirect value creation includes improved customer satisfaction leading to better regulatory performance scores, enhanced workforce productivity allowing teams to focus on higher-value activities, faster time-to-market for new services enabled by flexible AI infrastructure, and risk reduction from better decision-making. These benefits are real but require more sophisticated measurement approaches, such as comparing customer satisfaction trends before and after agent deployment or surveying staff about productivity improvements.

Implementation costs include platform licensing or infrastructure costs, data engineering and integration expenses, model development and testing resources, training and change management investments, and ongoing operational costs for monitoring, maintenance, and continuous improvement. Organizations often underestimate the "tax" of building and maintaining AI infrastructure in-house—the hidden costs of integration work, tool compatibility issues, and platform engineering team time.

Utilities that deploy agentic AI through pre-integrated platforms typically reduce total cost of ownership by 40-60% compared to building equivalent capabilities in-house. The platform handles tool integration, version compatibility, security patching, and infrastructure scaling, allowing utility data science and engineering teams to focus on use case development. This dramatically improves the cost side of the ROI

equation while accelerating time-to-value.

Tracking ROI over time reveals important patterns. Early pilots may show modest returns as teams learn and refine approaches. As the organization develops AI capabilities, subsequent use cases should show accelerating ROI because the foundational infrastructure, skills, and governance frameworks are already in place. If ROI isn't improving with successive deployments, it suggests that platform or organizational issues are creating friction that needs to be addressed.

Long-Term Strategic Positioning

Beyond immediate operational improvements, agentic AI positions utilities strategically for the future. As grids become more complex with distributed generation and storage, as climate volatility increases operational uncertainty, and as workforce demographics shift, organizations with mature AI capabilities will outperform competitors still operating manually.

Building this strategic advantage requires thinking beyond individual use cases to enterprise AI architecture. What common data platforms, orchestration frameworks, and governance structures enable rapid deployment of new agents as needs emerge? How can insights from one agent be shared with others to create compounding learning effects? What partnerships with technology providers or industry consortiums accelerate capability development?

Sustainability considerations increasingly factor into strategic planning. Agentic AI that optimizes energy usage, reduces water waste, and minimizes carbon emissions aligns with corporate sustainability commitments and regulatory expectations around environmental performance. Organizations should track and report the environmental impact of their AI deployments—energy consumed by AI infrastructure itself, as well as energy and resource savings that AI enables in operations.

Transparency about AI carbon footprints and water usage by data centers is becoming an expectation from investors, regulators, and customers. Utilities deploying AI in their own controlled environments can select efficient infrastructure, commit to renewable power purchase agreements, and optimize workload placement to minimize environmental impact—a level of control that's difficult with public cloud SaaS providers.

Ultimately, success in agentic AI is measured not by the technology itself but by the outcomes it enables—more reliable service, lower costs, improved sustainability, and better customer experiences. Organizations that maintain focus on these outcomes, while building the infrastructure, skills, and governance to deploy AI safely and effectively, position themselves to lead in an increasingly autonomous future for utility operations.

Ready to Get Started?

Shakudo enables enterprise teams to deploy AI infrastructure with complete data sovereignty and privacy.

shakudo.io

info@shakudo.io

Book a demo: shakudo.io/sign-up