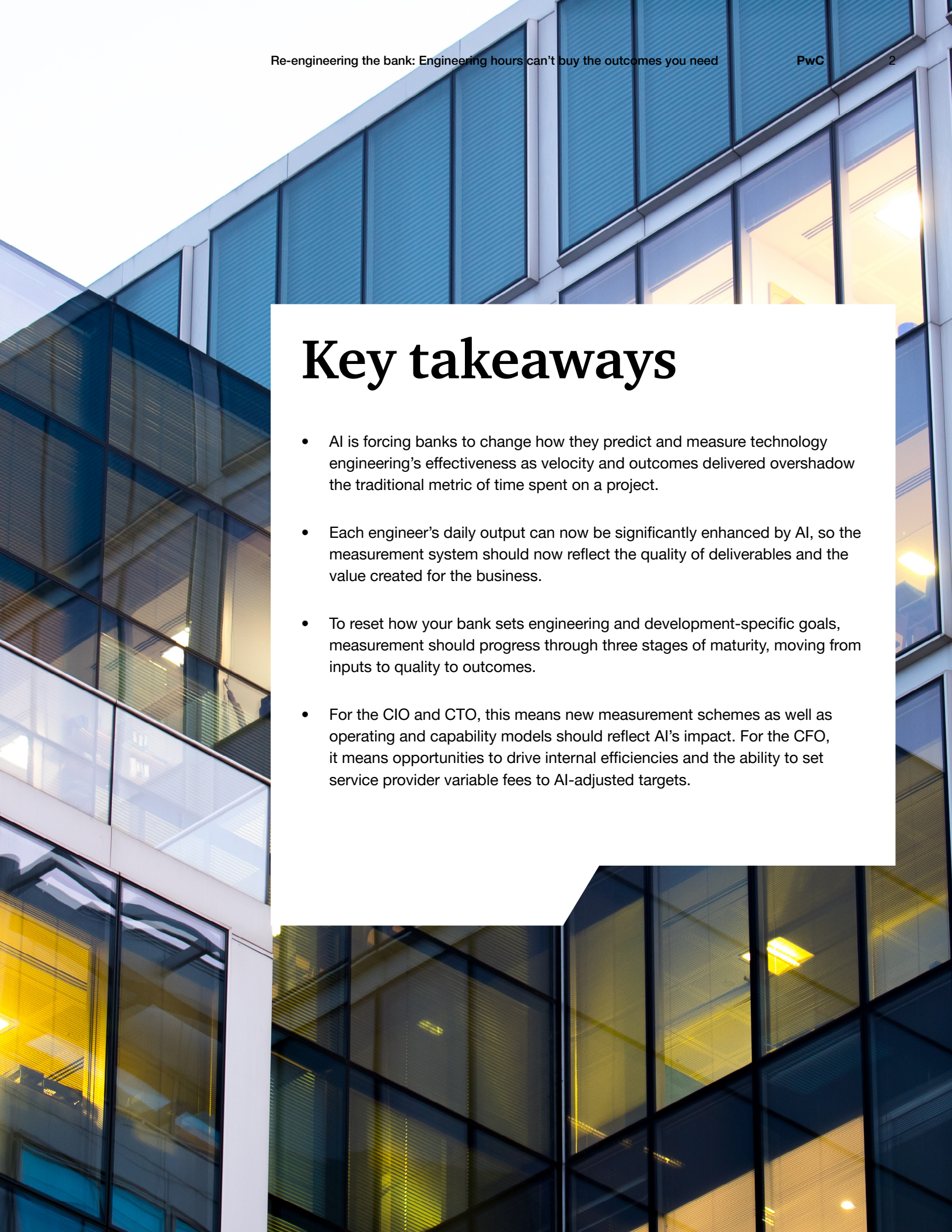
pwc

Re-engineering the bank: Engineering hours can't buy the outcomes you need





Key takeaways

- AI is forcing banks to change how they predict and measure technology engineering's effectiveness as velocity and outcomes delivered overshadow the traditional metric of time spent on a project.

- Each engineer's daily output can now be significantly enhanced by AI, so the measurement system should now reflect the quality of deliverables and the value created for the business.

- To reset how your bank sets engineering and development-specific goals, measurement should progress through three stages of maturity, moving from inputs to quality to outcomes.

- For the CIO and CTO, this means new measurement schemes as well as operating and capability models should reflect AI's impact. For the CFO, it means opportunities to drive internal efficiencies and the ability to set service provider variable fees to AI-adjusted targets.

Banks are using technology engineering more than ever as they attempt to automate more processes and modernize legacy environments. AI in engineering is the de facto investment, but these new tools can't produce the ROI you're looking for on their own. The issue, in our experience, is that AI's revolutionary power is not being paired with a revolutionary change in ways of working.

For example, the variance of output of each individual can be greatly reduced by standardizing output at the beginning of a sprint or the creation of a story point where agent-first development is guided by human oversight. The benefit shows up during the integration phase, when predictable levels of quality create a smooth process with far less rework.

The "secret sauce" is pairing AI enablement tools with modern ways of working, automation best practices, and common standards and controls. Together, they can turbocharge outcomes.

Yet the language that's often used to buy, manage, and evaluate the value of contracted engineering still boils down to more traditional measures including rate per hour.

A service provider contract's default vocabulary of full-time equivalents (FTEs), rate cards, hours, and locations is structured around time and materials—and frequently includes a minimum commitment to certain vendors in return for rate reductions.

Those legacy metrics increasingly are misaligned and cause confusion in the value exchange with service providers. They may miss the mark when they don't fully align to the operating model or the business outcomes engineers can provide in an AI-first culture. Faster workflows, increased automation, and reusable use cases are just a few of the AI benefits banks should be measuring rather than hours.

Institutions that measure hours get hours. Institutions that measure outcomes get outcomes. The unit of measurement defines the behavior of the technology engineering operating model.

AI as the forcing function

As AI continues to rapidly evolve, it's forcing significant change in how your bank uses engineering and the productivity of that resource. AI is the forcing function that's reshaping operating models as well as the measurement tools that gauge the performance of those models.

It used to be that an engineer's time was the handcuff limiting their output. That's no longer the case. AI coding assistants can increase individual developer productivity on routine tasks such as code generation, test writing, documentation, and so on. What took hours now takes minutes.

Taking the place of time is quality. AI is making a material difference when it comes to what "good" work practices look like. Engineering teams using AI-assisted development show higher deployment frequency, increased speed to value, and improved test coverage. Pre-AI measurement practices that focus on hours and other similar measures aren't designed for a world where time spent is becoming less important, while spending too much on AI is a growing risk.

AI goes far beyond accelerating engineering output. It redefines what and how rapidly engineering can deliver for the business. Institutions clinging to legacy measurement systems won't be able to gather the information that reveals the engineering value derived from AI.

If time is no longer that important, what's critical? Minimizing project ambiguity, collaborating to define the desired target state with high precision, executing on quality outcomes, and embedding control approvals in the process rather than as a toll gate at the end are today's pillars supporting ROI from AI. To do those things well, AI almost requires that you co-locate engineers with the lines of business so that they're enmeshed with one another.

That close relationship is an essential but unappreciated aspect of AI-leveraged engineering. As the volume of engineering output expands at much lower costs, it's human intelligence that's essential. It's able to define intent, exercise judgment at the right control points, and verify that AI-generated work meets the business objective. What banks need now are yardsticks that capture the effectiveness of human-led, AI-enabled engineering and updated contracts that incentivize the delivery of quality outcomes.

Changing how you measure AI in engineering

The path out of legacy engineering measures is a progression through three stages of maturity in how your bank sets goals, moving from inputs to quality to outcomes. Getting there requires decisions from different leaders at various levels of the institution.

For the CEO, the definition of engineering's contribution to the business should expand to include categories that were not possible or economical before AI, such as progress toward delivering personalized financial advice at scale or real-time fraud detection with explainability. For the CIO and CTO, this means new measurement schemes should help increase baseline productivity to reflect today's ways of working. For CFOs, it means variable fees in provider contracts should be indexed to AI-adjusted productivity or quality targets.

The contract maturity framework we use describes how your engineering measurement system progresses from hours to quality and then to business outcomes.

Stage 1

Inputs: Hours and FTEs. The focus is on headcount, FTEs, blended rates, and onshore/offshore mix. "How many people do I get?" is usually the primary question.
Goal: This is an opportunity to normalize the process and ways of working to set a measurement baseline.

Stage 2

Quality: Throughput per unit. The focus in this stage is more diverse, ranging from cost per story point, lead time from code commit to production, deployment frequency, change failure rate, and percentage reuse of components and platforms. This stage is transitional as it builds the factual baseline your bank needs to be ready for outcome-based contracts.
Goal: Focus on the incremental gains through automation and AI and understand the benefits, such as reduced rework expense, more stable cost per feature, and nonlinear capacity scaling.

Stage 3

Outcomes: Business value. In this stage, engineering leadership should be in business reviews and strategy sessions.
Goal: To calculate the revenue and margin impact tied to a platform, cost-to-serve, cost-per-transaction. Also to examine risk outcomes, such as losses avoided and to measure client outcomes, including digital adoption and time-to-decision.

This progression aims to build your bank's measurement muscle and then connect it to business value as quickly as the organization can absorb the shift from stage 2's quality to stage 3's outcomes.

A note of caution. It's tempting to stop at stage 2 because engineering is beginning to demonstrate business value, but it's still possible to "game" the system. For one thing, there's no standard definition of a story point, and that ambiguity creates room for drift. Cost per story point can be bid down by a vendor while the quality and substance of the underlying code declines. Or providers could offer cheaper rates per hour and add more bodies, producing the appearance of efficiency yet repeating the same input-driven economics of stage 1.

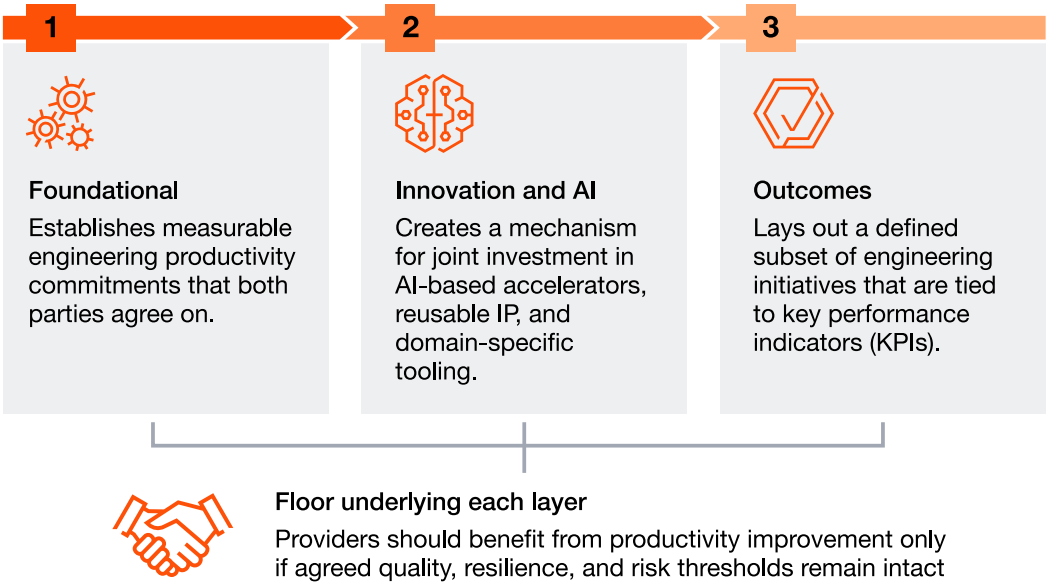
To avoid this, stage 2 metrics should focus on a combined view of code volume, code quality, and delivery velocity. It should be measured at the individual engineer level, such as code commits per engineer, quality of code delivered, and AI adoption rates (percentage of code that's AI-assisted; percentage of tests that are AI-generated). These metrics should appear in real-time dashboards that include benchmarks against peers so leaders can distinguish what's genuinely productive from what's not. eaders typically care about.

Without the shared language of engineering productivity, outcome-based contracts lack the operational foundation to succeed. First, build the measurement muscle and then connect it to business outcomes as quickly as the organization can absorb that shift.

The engineering contract of the future

Future engineering contracts should have three layers that your bank can use to align provider incentives with your desired outcomes.

Three successive layers that modernize computer engineering contracts



Foundational layer—productivity and flow targets

The foundational layer establishes measurable engineering productivity commitments. This layer corresponds to stage 2 of the maturity framework. It creates the shared language and factual baseline that's required before outcome-based contracting can work.

Both parties agree on baseline metrics and improvement trajectories for:

- **Deployment frequency**
- **Lead time from commit to production**
- **Change failure rate**
- **Cost per story point or feature delivered**
- **Component and platform reuse rates**

Innovation and AI layer—shared investment in acceleration

Building on the foundational layer, you can now create a mechanism for joint investment in AI-based accelerators, reusable IP, and domain-specific tooling. This layer acknowledges that AI is changing the economics of engineering faster than traditional contracting cycles can accommodate. Shared investment creates the incentive to adopt and scale. Examples include:

- **Jointly built code-generation patterns tailored to the bank's stack that establishes the agentic harness for AI-assisted engineering**
- **AI-generated requirements at a sufficient level of detail and specificity that AI tools build quality code that serves the business**
- **A modern data platform with domain-specific data products and feature stores**
- **Testing automation frameworks that leverage generative AI**

Outcomes layer—business KPIs

You can now move to the most mature phase, setting up a defined subset of engineering initiatives tied explicitly to key performance indicators (KPIs). This layer requires that business stakeholders co-define the criteria for success with engineering and external providers. Examples include:

- **Reduction in time-to-onboard for new customers**
- **Increase in digital adoption for a product or channel**
- **Reduction in operating expense for a specific process**
- **Improvement in settlement speed or exception rates**
- **Measurable gains in customer satisfaction or net promoter score**

Not every initiative can be tied to a business KPI from Day One, which is why the foundational layer exists as the floor.

Engineering contract operating model: Product-aligned, platform-powered.

Metrics alone don't produce outcomes. The organizational structure should support the measurement system. This means moving from function-organized engineering (a platform team, a data team, a channels team connected by handoffs) to domain-aligned engineering where cross-functional teams own business outcomes end-to-end.

Four elements define this operating model.



Cross-functional product teams. Engineers, product managers, data engineers, and control partners form stable teams organized around business domains, such as deposits, lending, payments, and onboarding. Each team should be in control of its outcomes.



Internal developer platform (IDP). A shared engineering foundation provides paths for infrastructure provisioning, continuous integration/continuous deployment (CI/CD) pipelines, security controls, and observability. It's here that the benefits of standardized testing, security, policy, observability, and evidence can be applied consistently. The benefit is that the IDP reduces the cognitive load on domain teams, letting them focus on business logic rather than infrastructure.



Establish the agentic harness. The IDP and CI/CD, respectively, define the control plane and pipelines, but it's the agentic harness where engineering work occurs in concert with AI, creating an autonomous, multi-step system to manage execution and feedback loops, promote compound engineering, control over memory and context, and guardrails and safety.



Clear ownership of shared capabilities. Shared services—such as identity, orchestration, data pipelines, API management—are explicitly owned, with clear service level agreements (SLAs) and roadmaps. Domain teams consume these services while platform teams invest in making them better.

This operating model is the delivery-side expression of the continuous delivery model described in a companion article. Domain teams deliver "value drops." The platform enables speed. Metrics confirm the system is working.

Reassess where engineering work is done

AI-driven productivity gains create an opportunity to rethink the geography of engineering work. Innovation and change initiatives—especially those that touch customer experience, product design, or front-office workflows—benefit from co-locating engineering teams with the business they serve. Working in close proximity accelerates feedback, sharpens requirements, and improves quality where that kind of competitive differentiation matters.

You can still leverage offshore delivery models for maintenance, platform operations, and foundational infrastructure where cost efficiency matters more. The contract of the future should align the sourcing strategy with the performance strategy. If you source hours offshore but measure performance on code quality, speed, and business outcomes, the disconnect will surface in every metric. Location decisions should follow the work, not the rate card.



CEO: Redefine what engineering delivers

Changing the measurement of what engineering delivers starts at the top. If the CEO sets the institutional definition of value as hours and headcount, every downstream decision—budgeting, contracting, governance, team design—follows suit.

There are at least four broad business outcomes that can be used to measure the value of engineering investments.

- **Financial outcomes.** Revenue and margin impact attributable to a platform or capability.
- **Efficiency outcomes.** Cost-to-serve, cost-per-transaction, operating expense reduction tied to automation and simplification.
- **Risk outcomes.** Losses avoided, capital efficiency gains, reduction in operational risk events.
- **Client outcomes.** Net promoter score, digital adoption rates, time-to-decision for customers and relationship managers.

These outcomes are part of the third layer of the engineering contracting maturity framework and only the CEO can make this a reality. When the executive committee requires that every engineering investment articulate its expected business outcome, the rest of the company reorganizes to deliver against that standard.

Practically, this means engineering leaders join business reviews, quarterly investment decisions are informed by data on outcomes, and "done" is defined by adoption and measured impact. This shift mirrors the broader reengineering agenda. As articulated in our Re-engineering the bank for growth article, capacity for growth is built by engineering outcomes, not projects

If engineering is a cost to be managed, banks will struggle to measure its value. A bank that sees engineering as a capability to be invested in, and measured by its impact on the business and its customers, changes the trajectory.

CFO and procurement: Restructure how engineering is bought

The commercial model is where measurement becomes contractual. How a bank buys engineering—the structure of vendor agreements, the metrics embedded in statements of work, the governance mechanisms attached to spending—powerfully reshapes behavior.

The three-layer contract structure requires a corresponding evolution in commercial terms.

- **Base fees**. Cover the foundational layer—the cost of landing and onboarding core team members, training, tooling, and infrastructure required to deliver within the bank's environment.
- **Variable fees.** Paid based on achieving or exceeding thresholds defined in the foundational layer. AI adoption can be a starting point as a leading indicator of the true reward which is the value created by AI such as flow, quality, unit economics and business outcomes. Performance penalties kick in when thresholds are missed—tied to the same dashboards that track code quality, delivery velocity, and AI adoption.
- **Gain share on outcomes.** Engineers embedded with business and ops share in the realized business value from the outcome layer, creating direct alignment between provider economics and institutional impact. This works only when business outcomes can be reasonably isolated and measured.

Governance shifts accordingly with jointly shared objective and key results (OKRs) replacing statement of work (SOW) tracking. Shared dashboards provide real-time visibility into productivity and outcome metrics, and regular portfolio reviews replace monthly status meetings.

The CFO's role evolves from negotiating rates to governing a value exchange. This directly extends the funding model described in the companion piece discussing continuous delivery. Domain-level funding tied to quarterly outcomes is the internal analog. The contract of the future is the external one. Together, they create a coherent financial architecture for creating engineering value.

CTO, CIO: Measure engineering excellence, not effort

The CTO owns how engineering is organized, how productivity is defined internally, and how technical capabilities compound over time. If the CEO sets the destination (outcomes) and the CFO restructures the commercial model (contracts), the CTO builds the operational engine that makes both possible.

Moving from headcount to engineering metrics

The internal measurement system should morph from tracking inputs, such as headcount, utilization rates, and hours logged, to tracking the efficiency and quality of what engineering is delivering. This is the second layer of the engineering contract maturity framework and it's the CTO's responsibility to establish.

Core metrics include:

- **Lead time.** How quickly a code change moves from commit to production. This measures the efficiency of the entire delivery pipeline—build, test, review, deploy.
- **Deployment frequency.** How often the team ships to production. Higher frequency correlates with smaller batch sizes, lower risk per release, and faster feedback.
- **Change failure rate.** What percentage of deployments cause degradation requiring remediation. This measures builds quality and testing rigor.
- **Cost per story point or feature.** The fully loaded cost to deliver a unit of business functionality. This connects engineering effort to economic efficiency.
- **Component and platform reuse.** What percentage of new capabilities leverage existing components, services, or platform capabilities rather than building from scratch.
- **AI in engineering index.** The proportion of code that is AI-generated versus human-written, the percentage of tests and documentation maintained by AI, and the degree to which AI is embedded in code review, incident triage, and developer tooling. This index isn't a secondary dashboard. It's a core engineering metric that directly affects every other metric on this list.

While several of these metrics aren't new, what's novel is making them the primary language of engineering performance. When the CTO reports to the executive committee on lead time and deployment frequency rather than team size and project milestones, the conversation changes. coherent financial architecture for creating engineering value.

The AI in engineering index

AI is reshaping both engineering productivity and outcomes, but most institutions lack a systematic way to measure them.

The AI in engineering index that we propose serves multiple purposes. First, it provides visibility into AI adoption—which teams are using AI effectively, which are not, and where the greatest productivity gains are occurring. Second, it creates the baseline needed for outcome measures such as declining rework, change failure, review effort, and cost per accepted change. Third, it's the foundation for renegotiating provider contracts as AI changes the economics and outcomes of engineering.

The index should be constructed from a simple, consistent set of metrics that track how deeply AI is being embedded in the engineering workflow and how that's improving quality over time.

- Percentage of code that's AI-generated versus written code
- Percentage of tests generated or maintained by AI
- Percentage reduction in testing backlog drag
- Percentage of incidents triaged or resolved with AI assistance
- Percentage of code reviews augmented by AI analysis
- Percentage of documentation generated or maintained by AI
- Time savings from AI-powered developer tooling per sprint
- Percentage increase of predictability of delivery quality
- Reduction in time-to-value per feature.

If your engineering organization can't quantify how AI is changing its productivity, you're flying blind into negotiations with your providers.

Achieving 100% in each metric is unrealistic. People are still essential in all areas of engineering. Instead, your bank should be setting a target that aligns productivity with strategic business outcomes. Some outcomes, say the approval of routine payments, can be nearly fully automated while others, such as loan approvals, likely will require more human involvement (See PwC's Consumer Lending Radar survey for more detail on what lenders need to know about using AI in lending operations and user experience).

The entry point for change

The shifts we describe here are structural and they touch executive mandates, commercial contracts, operating models, and measurement systems. Such a change won't happen all at once, but you can start in one place.

The entry point for change is to choose one domain, one provider, one quarter, and make good on those three commitments. Establish the baseline, change the measurement and commercial model, prove the economics and controls, and then scale based on evidence.

Define an outcome the CEO owns

Select a single business domain—deposits, payments, onboarding, lending—and define the engineering and control investment in terms of a business outcome. Not "deliver a new platform" but "reduce time-to-onboard by 40 percent" while keeping fraud, operational-risk, and control outcomes within your bank's tolerance. Your executive committee agrees on the metric and reviews progress.

A caution on metrics translation: Engineering productivity gains do not map linearly to budget reductions. If AI enables engineers to code 30 percent faster, a CEO might reasonably expect engineering budgets to fall 30 percent. In reality, the gains do not apply uniformly across the portfolio. New costs emerge (AI tooling, token consumption, infrastructure) and the most valuable use of freed capacity in these early days of AI adoption is often reinvestment—not savings. The outcome metric must reflect what the institution actually gains, not a simplified arithmetic that overpromises on cost and underdelivers on value.

Apply the three-layer model to one provider engagement

Take one provider relationship in that domain and restructure it around the three-layer contract. Establish the foundational productivity metrics. Define one or two outcome KPIs. Create a small innovation/AI investment pool. The commercial terms don't need to be perfect but they do need to be directionally correct and measurable.

Establish the engineering productivity baseline

The CTO establishes the stage 2 metrics—lead time, deployment frequency, change failure rate, reuse rate—for the selected domain. Simultaneously, begin tracking the metrics that go into your AI in engineering index. This baseline is the factual foundation for every conversation with your executive committee, with providers, and with your board.

Prove it works in one domain. Prove the metrics are real, the contract structure is viable, and the outcomes are measurable. Then expand.

The opportunity is not simply to buy engineering more cheaply, it's to redesign how engineering value is defined, verified, funded, and rewarded.



The re-engineering the bank for growth series

This is the fourth in a series of articles examining how banks can build capacity for growth through structural re-engineering—a larger process than investing in technology, applications or data management.

- Re-engineering the bank for growth. The parent piece establishes the case for horizontal coordination and capacity for growth across the enterprise.
- Re-engineering data architecture for real-time intelligence. The data piece addresses how to reengineer the data foundation from regulatory compliance to real-time intelligence—shifting from transaction-oriented architectures to domain-owned data products.
- Re-engineering the bank: From programs to continuous delivery. The delivery piece addresses how to shift from multiyear programs to continuous value delivery—restructuring budgets, governance, and team design around quarterly outcomes.
- **Re-engineering the bank:** From hours to outcomes (this piece). This piece addresses how to re-engineer the measurement, commercial, and operating model for engineering itself—moving from hours and headcount to outcomes and impact.

Together, they form a connected agenda: Redefine the data foundation, reshape the delivery model, and re-engineer how engineering value is defined, bought, and measured.



Contact us

Dan Goerlich
Partner, Banking and Capital Markets Advisory Leader
PwC US

Ryan Battles
Partner,
PwC US

Chris O'Hara
Partner, Data & Analytics
PwC US

Christian Iantoni,
Cloud Acceleration Leader
PwC US

© 2026 PwC. All rights reserved. PwC refers to the PwC network and/or one or more of its member firms, each of which is a separate legal entity. Please see www.pwc.com/structure for further details. This content is for general purposes only, and should not be used as a substitute for consultation with professional advisors.