

THINKING PIECE · OPERATING MODEL

Why your operating model will determine whether your AI investment works

AI strategy is only as good as the organisation running it. The businesses getting this right start with structure, before they start with technology.

Most AI investment fails for a structural reason. The organisation delivering it was built for a different kind of work, and the technology moved before the organisation did.

Professional services businesses were built for a world with a clear end state. You scope a project, build to a waterfall plan, go live, and hand off to a separate support function that runs a fixed-scope, break-fix service from that point on. AI has no clean handoff. A model that goes live still needs monitoring for drift, retraining as data changes, and evaluation against compliance requirements, for as long as it's running.

AI adoption has consistently failed to deliver the return businesses expected of it. The model, the infrastructure, and the governance policy get the attention and the budget. The organisational structure underneath them is where this piece argues the gap actually sits.

The mismatch

The ongoing work on a live model needs the same skill as the build, often the same people. Most delivery structures were never built to provide continuous improvement at all. The people who understand the model move on to the next project, and the best specialists get pulled off new deals to firefight old ones, because there's no capability in the structure built to absorb the tuning work. That gap costs sales capacity every time it happens, and it grows as the AI pipeline grows.

Decision rights, not just ownership

Governance tells you that someone needs to own AI risk and AI performance. It doesn't tell you how your organisation is structured so that ownership functions day to day.

Who approves a non-standard AI commitment before it's made to a customer? Who has the authority to pause a system when performance degrades? Who's accountable when a model that performed well at go-live is producing worse outputs six months later, and who finds out before the customer does? Most businesses have a clear answer for who signs off a discount or who owns a client escalation. Almost none have defined the equivalent for AI, because the questions themselves didn't exist until recently. Governance poses the question. The operating model needs to answer it unambiguously.

The capability problem

Most delivery businesses split technical capability into professional services and managed services: two teams, two cost structures, two career paths. That split made sense when the work was genuinely different, build once, then maintain a stable system. AI removes the stable system. The build gets shorter. The ongoing work, monitoring, tuning, drift correction, becomes the harder and more valuable part, and it ends up with the team that was staffed and priced to handle break-fix tickets.

A single technical resource pool, spanning pre-sales, build, and ongoing tuning, fixes this by letting skilled people move to wherever the work actually is. Most businesses still account for the professional services and managed services split as a cost efficiency. In practice, it's what caps how much AI-enabled work they can deliver at any given time.

The commercial model has to follow the structure

Fixed-price, fixed-outcome pricing and licence-based support both assume delivery cost is static once you're live. AI breaks that assumption in three specific ways: there's no natural completion point, because monitoring and tuning never really stop; inference and hosting costs move on the vendor's schedule, not your contract term; and usage is driven by what the end customer does with the system, conversation length, complexity, volume, none of which you control. Price an AI-enabled engagement like a SaaS licence and your margin erodes in a way you never forecasted, because the thing driving your cost now sits outside your control entirely.

Sales has to sell differently. Finance has to forecast differently. The account has to be monitored differently after signature. This is the commercial consequence of the structural problem above. It's usually the first place a business feels the pain, and the easiest one to mistake for the whole problem.

CUSCOE®

CXBridge's answer to this is CUSCOE®, a peer-reviewed operating model framework, accepted for the proceedings of KMIS 2026 (Copeland and Mogford, 2026), built around four functions: Portfolio, Customer Success, Sales, and Centres of Excellence.

HOW CUSCOE® ADDRESSES IT

Resourcing. The professional services / managed services split is replaced by a single technical pool, so skilled people move to the work instead of the work waiting for whoever's free.

Accountability. Customer Success carries real P&L accountability for the account, tracking consumption against forecast so margin risk gets caught early rather than at renewal.

Commercial model. Three layers instead of one: a fixed-price MVP, an ongoing service agreement for monitoring and tuning, and consumption billed at cost plus margin, so the pricing structure actually reflects where the cost and the risk sit.

Get the structure right and the pricing holds, the specialists have somewhere to sit, and someone is actually watching the account before it becomes a problem. Get it wrong and none of that shows up in the pilot. It shows up on the account, live, about six months in.

REFERENCE

Copeland, D. and Mogford, K. (forthcoming) 'Designing a Customer-Centric Operating Model Framework for AI-Readiness', in Proceedings of the 18th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management: KMIS. SciTePress.

Bringing hard-won CX and operational experience to AI-enabled transformation, people first, always.

CXBridge | cxbridge.co.uk