THE RELAUNCH · A MARKET OVERVIEW

We make AI real

CoA is now Collaborate AI. Our inaugural issue on the impacts of AI on cost deceleration, and where a native AI-first company is placing their strategic bets.

THE RELAUNCH · ISSUE 01

CoA is now Collaborate AI

Industry-award-winning practitioners and leaders building critical services and products at the edges of AI.

CoA Solutions is now Collaborate AI. Three people shipped six products and spent 21 billion tokens over six months across computational pathology to VA benefits. We re-launched as an applied AI company working on the forefronts of AI assistive and compliant products, modernization, and thought leadership.

This is also Issue 01 of From the Model, the company's own publication. (1) The quality bar has risen as AI accelerated the reducing cost of shipping technology. (2) AI pays off when it is assistive, measured, and governed. (3) The risk is implementing AI without strategic vision, focus, and accountability to see if recommendations or task automation is actually creating results. As the field has become saturated, over marketing and AI slop has made it hard to find companies who are actually creating on the edges of models and can scale solutions up or down, narrow or wide. Our central thesis, backed by our own experience and survey of the field, is that a senior group of award-winning talent can scale product for critical services at the true cost of technology.

BY THE EDITORS OF FROM THE MODEL · ARLINGTON, VIRGINIA

IN THIS ISSUE

BY THE NUMBERS

21B

tokens over 6 months — moving operations onto agents and shipping products.[1]

$1M

in operational costs saved — reinvested in senior people's time.[2]

20K+

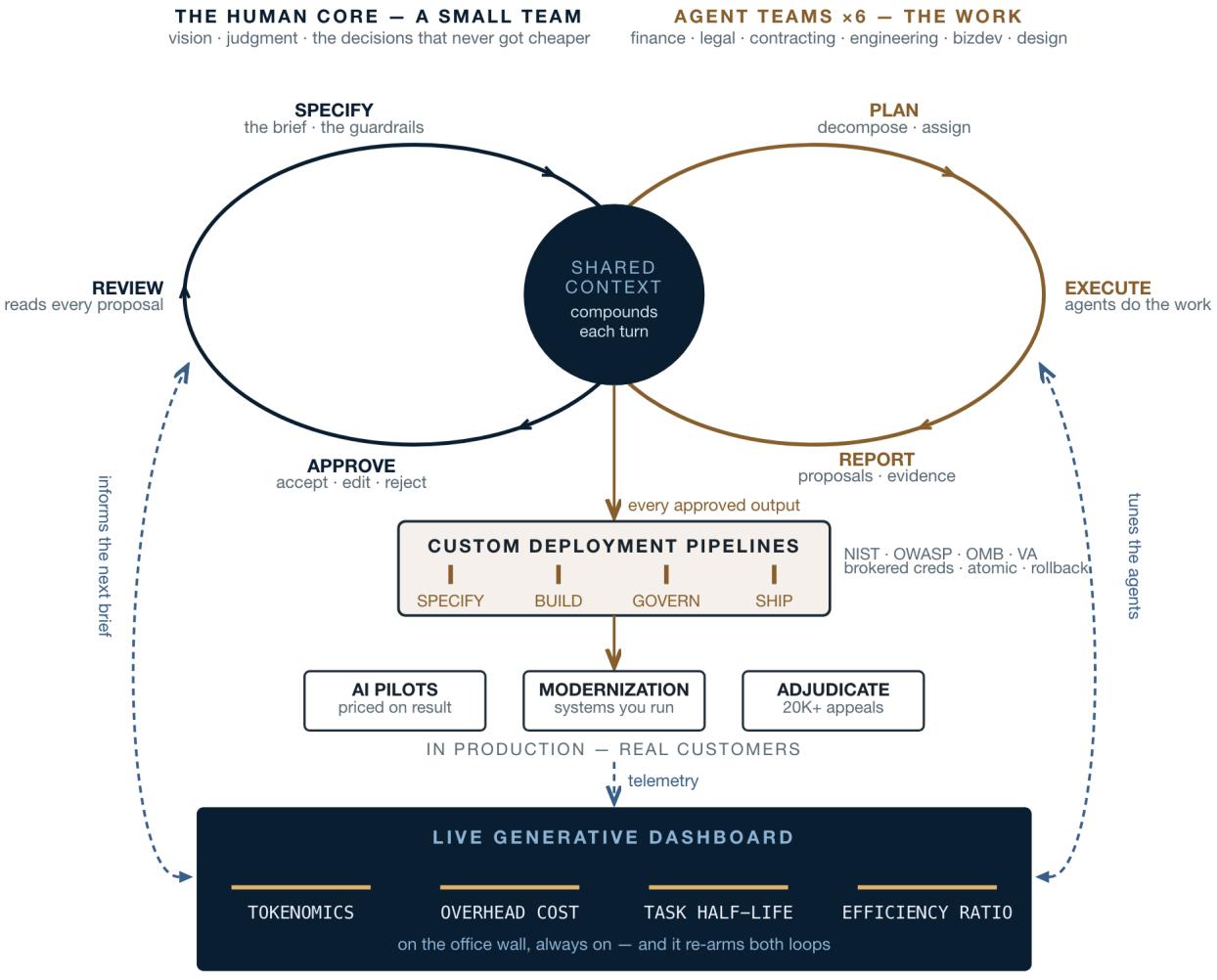
veteran appeals processed through Adjudicate.[3]

6

products in six months — the whole team is three people.[4]

The nuts and bolts of a native AI company

FIG. 1 How a small team runs a multi-million-dollar company: people and agent teams in one shared context, a governed path to production, and a live dashboard that monitors and starts the loop again.[5]



WHAT EACH PASS THROUGH THE LOOP BUYS — COMBINED RESULTS OF THE PAST 90 DAYS

20% PROMPT REDUCTION	15% TOKEN-COST REDUCTION	8% OPERATIONAL-COST REDUCTION	7 OPTIMIZATION AREAS DISCOVERED

Over the past two years, the cost of making software collapsed. The difference is that AI has now changed the shape of this curve to be more exponential due to its effects on human capitalization. The making has come loose from what companies used to budget and staff against, and the interesting question has become what that collapse freed, and who turns it into work a client can feel, measure and see.

As a result, the way technology is bought is also changing quickly. In the commercial space, as models improve and get cheaper, the ability to monetize their capabilities with quality, creativity, and at the right timing are becoming the competitive differentiators. In government, often times, you are acquiring services or teams to build a product once inside.

This way of procuring technology is costly as the government pays for "runaway." Excluding clearance requirements, the cost of custom software was expensive. So much so that an entire generation of agile delivery leads, project managers, and PM's were trained to help manage cost. The ramp up period was necessary to help companies with cost. Post AI, it has become less necessary when the cost of making has reduced so much.

As native AI-first companies become more common, there is a growing trend of senior experts running agentic teams[6]. With Opus 4.8, now 5 and Fable, assistive AI technology (our term to note that expert review is still critical) can mean those stewarding critical services can see and buy products made to their problems.

One of our first tests was to run one supervised agent against consolidating an opportunities interface to quickly identify matched RFI's and RFP's in the federal sector. The agent worked for twenty-two hours. By morning it had stood up the spine of a platform that pulls, normalizes, and indexes four federal data sources into one searchable system: SAM.gov, USAspending, SBA certifications, and GSA labor rates. The platform has been in daily use since[7]. We then created an executive response dashboard based on a criteria including our expertise and past performance helping us shepherd our Bdev pipeline. This alone saved us 1.5 FTE in 2025[8].

Another example. Thirty-two people helped implement the Appeals Modernization Act of 2017. We built its commercial successor on open-source code that has been running for eight years. In the recent Board of Veterans' Appeals COTS RFI, we took our COTS product and fully adapted it to the AMA process: 72 percent of all requirements covered, a functional requirements checklist, a searchable user-story database, and AI-assistive governance and dashboards with data instrumentation baked in. In two weeks, with two people, not a full team[9]. All of which scaffolded on a database of hundreds of thousands of synthetic VA data that we made. This reality is the new bar: not vaporware demos, code challenges, but pre-production scale products before procurement and drastically reduced post-award ramp up.

This trend has been accelerated by a drastic reduction in almost every part of software development: by Stanford's count, the cost of compute for a capable answer fell 280-fold in two years[11], storage is down about 87 percent[12], and build times dropped from years to weeks. The second is that AI has started to absorb the human work itself. There is a catch here. In Harvard and BCG's study of 758 consultants[13], the people who gained most were below the average: they improved 43 percent, against 17 percent at the top, closing toward a shared quality ceiling. Senior practitioners set that ceiling, because range and experience are what it takes to direct agent teams toward a vision. So the saving is not cheaper hands doing the same tasks. A growing share of tasks runs automated, and the work collapses into fewer, more senior people who direct and run the AI at consistent, measurable quality and higher productivity.

In our own company, the shift has tremendous value. We saved $1M in operational overhead that was re-invested into AI acceleration and growth. Our product cycle times reduced by 45%[14], allowing us to iterate on ideas faster. The custom git server that monitors all our products in real time has allowed us to find and fix bugs before users even notice. The harder part is finding what problem to solve, implementing AI well with the right safeguards, and ensuring proper governance and accountability. All of which we learned how to do in conversation with AI as they improve for our customers and users.

When the cost of making is this low, specifying, directing, and judging where AI is best leveraged in an organization, and who you trust to do it, becomes even more critical. Take the twenty-two-hour build. The agent, its operator noted, "didn't know how to organize the data so searches stay fast" and "didn't design the error handling." "Those decisions were mine." The constraint moved into domain knowledge and verification. In other words, AI accelerates the people who already have the judgment. That is why the rebuild was led by the same person who led the original government build.

The strongest counterargument is the enterprise failure rate. According to MIT's NANDA initiative, 95 percent of enterprise AI pilots returned nothing measurable from the $30 to $40 billion invested[15]. Treat that figure with care: pilots tend to show no return over short windows, whatever the technology. In NANDA's read, the successful five percent won on experience, the ability to redesign a workflow around the model instead of bolting the model onto the old one, not on better tools. McKinsey found the same: real value where leaders rebuilt the process, little where tools were only added. So lower build costs are necessary and nowhere near sufficient. So the deciding input is who, the quality of the thinking, where, and governance of what the AI is doing and whether it is actually creating value.

Even if there is a surplus in margin from lowering technology cost, the value isn't passed on evenly to customers. An incumbent's cost structure was built for the old curve, with staffing ratios written into contracts and review chains sized for slower work, and unwinding that is slow and culturally painful, a cost a native AI firm never pays. Once a company achieves a surplus, the instinct is to retain the profit and increase the margin. The alternative, and the one we take, is to spend it on increasing the quality of the services we provide. Since the work we do is for critical services and benefits, we push our internal systems and hiring processes to ensure a rising bar of quality at better cost propositions for customers. The end goal is everybody wins. The hardest part is finding the right level of innovation titration for a given customer and problem space, matrixed against their culture, people, needs, compliance realities, and competitive environment. That is another case study, coming soon.

FIG. 2 HOW WE RUN A PILOT

WHAT MAKES AN AI PILOT SUCCEED

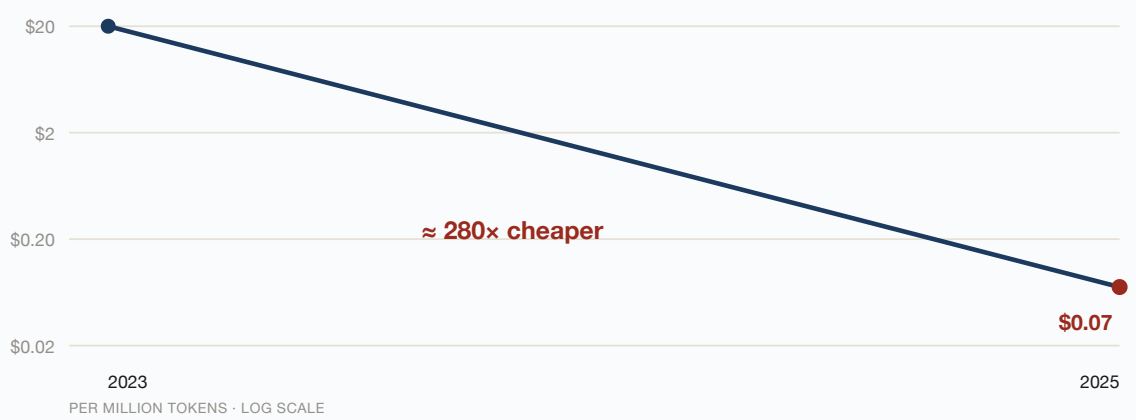
1. **Small, senior staff**
   Highly experienced senior staff with range (quality) with AI being the output (quantity)
2. **Strategic vision**
   Working with leadership and SMEs to codify where the technology needs to go through their eyes
3. **Governance controls**
   With reporting built in, not bolted on
4. **Model and task differentiation**
   The right model for the given task
5. **Assistive AI**
   Real-time feedback, not overrides with quality bars set by experts
6. **Clear quality bars**
   Set by leadership and the SMEs

How we run one: senior people, a clear quality bar, and governance and reporting that ship with the pilot rather than after it.

The collapse in the cost of making

Several cost curves fell at once, for the same reason. Technology has been getting cheaper, and smaller teams delivering more has been the trend, AI just supercharged it.

FIG. 3 Inference cost, per million tokens — STANFORD HAI AI INDEX [11]



A GPT-3.5-caliber answer fell from about $20 to seven cents in two years, roughly 280x. Incumbents still price in the old premium.

FIG. 4 The people cost is falling too — COMPANY REPORTS; O'BOYLE & AGUINIS; MCKINSEY [16,17]

SMALL TEAMS, BIG OUTPUT

2012	Instagram	13 people	a decade of platform leverage, then a $1B exit
2014	WhatsApp	55 people	$19B acquisition, 450M users; the era's famous outlier
			LATE 2022 · CHATGPT · AI GOES MAINSTREAM
2023	**Midjourney**	**~40 people**	~$200M revenue in year two, no outside funding (reported)
2024	**Cursor**	**~20 people**	$100M ARR in under two years (reported)
2026	**Collaborate AI**	**3 people**	six products in 6 months, seven-figure revenue (self-reported)

THE CAVEAT

A small group of top talent can now create dozens of products.

In a 633,263-worker meta-study, the top 5 percent produced about a quarter of all output (O'Boyle and Aguinis, 2012). McKinsey puts top performers at up to 8x average in complex work; Netflix built its operating model on that talent density.

Sam Altman frames the same thesis as the one-person billion-dollar company. The gains follow the most senior people: our two-week case-platform rebuild was led by the same person who led the original 32-person build.

Four forces are resetting how every organization operates

The barrier to capability has collapsed. The advantage moved to judgment, timing, and the infrastructure to adapt. The four forces form one loop with a direction.

01 THE JAGGED FRONTIER

Judgment decides who wins with a shared tool

Everyone rents the same model, so judgment, not access, is the constraint. Inside its competence, the tool improves the work. Just outside it, the same tool makes elite consultants confidently wrong.

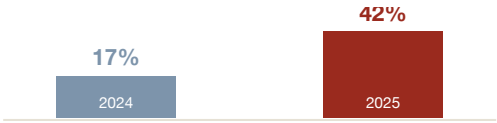
>40% / -19 pts Quality gain inside the frontier versus the drop in correct answers just outside it

HARVARD BUSINESS SCHOOL & BCG, 758 CONSULTANTS WORKING WITH GPT-4 [18,13]

02 THE ABANDONMENT WAVE

Selection beats speed

Speed won in 2023 and 2024. The correction came in 2025. Demos are cheap, so everyone built one. Those projects collapse on live data, hard compliance rules, and the cases a demo never touches.



17% → 42% Share of companies scrapping most AI initiatives, in a single year

S&P GLOBAL; MIT NANDA (95% OF PILOTS; $30–40B IN SPEND) [19,15]

03 THE DEFLATION

The cost is the organization, not the technology

The deflation cuts some inputs and not others. The technology collapses toward zero. The organization that delivers it stays human. The gap between price and cost is a legacy tax.

280x Cheaper for a GPT-3.5-caliber answer: $20 to $0.07 per million tokens, two years

STANFORD AI INDEX; BACKBLAZE (STORAGE ALSO –87% SINCE 2009) [11,12]

04 COLLAPSING CYCLES

No advantage lasts. You re-earn it or lose it

Incumbents die faster as the market absorbs new capability in months. With tenure near twelve years, no product or market share is durable. Only the capacity to keep re-earning position endures.

33 → ~12 yrs Average S&P 500 tenure: 33 years in the mid-1960s, 24 by 2016, ~12 forecast

S&P 500 TENURE DATA; CHATGPT REACHED 100M USERS IN ~2 MONTHS [20]

Most AI is a demo. Ours runs.

Only one enterprise pilot in twenty returns anything measurable, and the failures trace to adoption, not the model. The failure routes in the second figure come from our homepage.

FIG. 5 Where enterprise AI pilots end up MIT NANDA; S&P GLOBAL [15,19]

1 in 20 enterprise AI pilots returns a measurable result (in red); the other 19 return nothing.

Ninety-five percent return nothing measurable on the bottom line.

FIG. 6 Why ours is different COA.SOLUTIONS, THE HOMEPAGE [21]

TRADITIONAL MODELS

01	**Vaporware demos**	Impressive on the vendor's data. In production, on yours, there is nothing there.
02	**Low-code ceiling**	Climbs fast, then flattens. The hard last 20 percent stays out of reach.
03	**Off-the-shelf misfit**	The generic tool does not fit the work, and drags a legacy change-management tail.

OUR ANSWER

AI pilots de-risk traditional models of buying by building you products at the true cost of today's technology.

In summary, the cost of shipping technology has decreased, and has also proportionally increased the need for strategic vision and placement of AI within a problem space, industry, and organization. When MIT's researchers went looking for the return on the enterprise rush into generative AI, their 2025 survey found that 95 percent of pilots produced no measurable effect on the bottom line[15]. There were some key themes: incorrectly applied use cases, immature models or not understanding the edge of those models, and procuring the wrong talent or cost structure. There are also AI-specific ones such as governance and model selection, more on that later.

In both federal procurement and commercial development, older procurement strategies are becoming much less effective when the cost of making has exponentially decreased. They price in older ways of working that need high amounts of training, like low code, or procuring misfit COTS products that are not pre-built to your problem. Especially now, when AI allows cheap enough development to make this the bar, with a significantly smaller team.

In off-the-shelf software, that adoption and integration is the real price. The buyer thinks the license is the purchase, but the license is only a third to 40 percent of first-year cost. The rest is labor: integration, configuration, training, data migration, the slow work of bending a product built for everyone onto one company's way of working. Setup routinely costs twice the license, and a median ERP install lands near $450,000[22]. The premium pays for the distance between what a generic product does and what a specific buyer needs. The strange part is that the premium survived the deflation, because closing that distance is exactly the labor AI made cheap. Every input the incumbent used to bill as scarce has fallen at once, and the premium is still charged as if nothing changed.

The second cost that gets passed on is deculturation cost, or the overhead a company invests to make drastic transformations to their own infrastructure. When technology as large as AI enters the marketplace, nothing is unaffected: business development, marketing, delivery, strategy, and other critical functions need to adapt. All overhead flows somewhere, and it's usually in price. Native AI companies do not have this cost. An incumbent can absorb it quietly or pass it through in the rate, and the competition that once forced the first choice has thinned, because the whole field carries the same overhead. So the customer ends up paying for the vendor to become an AI-native firm. Maybe some cost of fit is simply unavoidable. Low-code was the last serious attempt to dodge that cost. Low-code builds the easy 80 percent fast, then stalls on the custom logic and edge cases that are the actual work, and many of those builds are rewritten in real code within two years. Fit always cost something.

Public procurement has always been about trust to deliver. AI has also automated much of the compliance structure, and has reduced the cost to scale. By leaning too much on incumbents, you pay all the aforementioned costs without a large jump in quality. Native AI companies can serve as both a technical and pricing bridge into higher-value software procurement.

The old proxies for trust no longer measure what a buyer most needs to know. Output no longer follows headcount, or the age of the firm producing it. The two-week rebuild and the overnight four-source build described earlier are the evidence. The better signal is judgment. The irreducible cost is still a person choosing: which appeal, which diagnosis, which design. But the build cost never decided whether a capability paid off. Everything around the decision did: the buying, the installing, the adopting, the evaluating. That machinery never got cheaper. It broke.

FIG. 7 HOW SOFTWARE BUYS TRUST

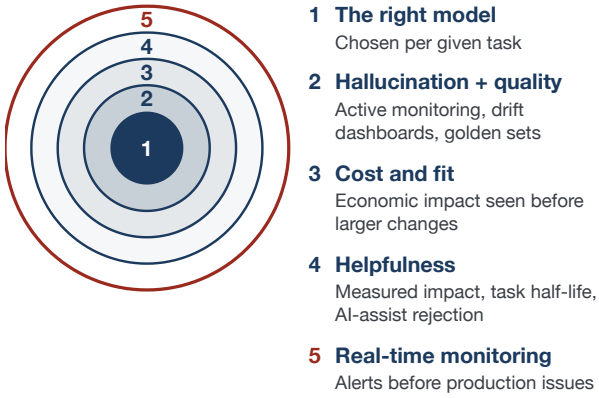
Procurement has always bought trust. What signals it has changed: quality, governance, and price at the true cost of technology.

Most enterprise AI never pays off. MIT put the share of pilots returning nothing measurable at 95 percent[15]; S&P found 42 percent of companies scrapped most of their AI projects last year[19]. Anyone funding AI right now knows how difficult it is to make the impacts real in their organization. "We Make AI Real" is the banner we rallied under to meet this need. A system cannot just be intelligent; it must also be learnable, perceptive, adaptive, and compliant.

Every AI product we build is governed inside compliance controls and the quality bars set by the leadership and SMEs of a workflow: NIST's AI Risk Management Framework and its generative-AI profile, the OWASP Top Ten for LLM applications, two OMB memos (M-25-21 and M-25-22), and the VA's Trustworthy AI principles[23]. Excessive agency is a model taking an action it was never meant to take; our answer is least-privilege access and a human-approval gate. You can argue a model out of an instruction in a prompt. You cannot argue it past a permission boundary it was never granted. An evaluator checks controls, and controls are auditable.

At Collaborate AI, when we approach critical systems, we take an assistive AI approach. Every AI recommendation is a proposal that a SME in a workflow can accept, edit, or reject. These are then reported up to give a real-time view of when the human or the AI recommendation was right, which are used to fine-tune recommendations so they are actually helpful to the people. We have the same transparency and reporting model for efficiency claims, drift, hallucinations, and so on. All of which are updated based on the latest science of vulnerabilities and research, ensuring best practice like the use of golden evaluation sets. Due to where we are with decreasing costs and ease of data instrumentation, we provide these out of the box. Other vendors usually still charge for the reporting capabilities in addition to compliance and governance needs. Grounding shows an answer is supported by the record, not that the conclusion is right, so correctness is scored separately, on golden evaluation sets against known-correct outcomes at every release.

FIG. 8 HOW WE BUILD GOVERNANCE



Governance is built outward from the model: each ring adds a check the one inside it cannot make for itself, and at the outer ring a person decides.

NIST AI RMF · OWASP TOP TEN FOR LLMS · OMB M-25-21/22 · VA TRUSTWORTHY AI

We also follow the same level of discipline from a tokenomics standpoint. As models advance, existing models also improve and costs go down. Even within enterprises that have their own, proper delineated task differentiation at scale ensures the organization is paying the precise amount for the problems they are solving. There have been quite a few examples of this causing major issues, such as Uber burning through its entire 2026 AI budget in four months[24]. We know the pricing models inside out and, from published research and our own product development and 21 billion tokens later, the edges of the model tiers. You don't need the latest for everything.

In summary, AI presents an opportunity to buy solutions to problems at the true cost of technology, and unique opportunities:

- **Strategic vision and placement of AI is more important than ever.** As building becomes cheaper, strategic vision, leadership, judgment, compliance, and quality become even more important. These have ironically been the more difficult problems pre-AI, and as AI scales the needs will accelerate. Procuring products should also mean testing a company's ability to implement these into true governance models tracking cost, recommendation, and accuracy against an organizationally set quality bar.
- **Quality of thinking, creativity, and desirability are the differentiators.** The research shows that highly skilled senior staff can drive outsized savings and efficiency. Consequently, quality of thinking and range is what contributes disproportionately to AI productivity[25]. For example, in the identity section later on you will see a sample prompt from our redesign. The idea of higher quality inputs leading to better end states with more precision is not new. We have encapsulated this as garbage in, garbage out in data. In AI terms, the reasons are mostly accumulated soft skills: task alignment, evaluation (what tools for the right job), creativity, or their combined experience and identity meeting real world problems, and so on.
- **Compliance and reporting are no longer extras.** As compliance, governance, and reporting also become easier with improved data instrumentation, there is a large opportunity for cost savings to expect these as part of procurement, not as add-ons.
- **Products can exist before procurement.** AI-native companies with senior staff can create products pre-procurement and steps before production scale, which leaves integration and fine-tuning workflows. Especially ones like Collaborate AI, who has developed half a dozen products this way in six months. Even if not us, they do not have deculturation cost and you are not paying for a company's modernization bill.
- **Ask for the product, not the branding.** Focusing on actual product requests during the RFI and RFP process can vet companies who are mostly selling the same thing packaged with AI branding. Any AI-native company should be able to provide this with where current models are.

Four lines of work, at rising burdens of proof

THE CHANGING REALITY

The bar has risen. Companies should sell products, not services, and customers should see the product working before they buy. COTS should mean demos already adjusted to your use cases before purchase, interfaces you drive with a prompt instead of a low-code box, and product-based proposals in place of marketing-filled documents.

OUR VALUE PROPOSITION

We build products at the true cost of technology. Each one is made for your problems, with the compliance you pay for automated into the product, and with senior, industry-award-winning experts you can trust. You see the thing working on your use cases before you buy it, and the governance, reporting, and quality controls arrive with it rather than as a line item after.

AI Pilots

EXPERIENCES PEOPLE LOVE AT TRUE COST

Rapid, senior-led pilots built to cross the demo-to-deployment gap where most attempts die. The pilot is scoped to fit production data, real regulation, and the corner cases nobody plans for.

FOR Senior federal and commercial leaders funding a first AI initiative

Proof. One supervised agent, handed a full system design, ran 22 hours overnight and stood up a four-source federal-data platform (SAM.gov, USAspending, SBA, GSA) now in daily use.[7]

Adjudicate

NEXT GENERATION CASE REVIEW AND MANAGEMENT

The adjudication and case-file-review product, drawn from eight years of veteran appeals, and the only one of the four lines already shipped and running. Two editions share one core: the firm edition attorneys use and the adjudicating-body edition.

FOR Adjudicating bodies and firms doing document-heavy case-file review

Proof. The dedicated page cites 20,000+ appeals across 60+ reviewers, over eight years of production casework.[3]

AI Modernization

RESULTS DRIVEN TRANSFORMATION

Modernization for industries where scale is hard. Priced on what we deliver, not on headcount or a large bill up front: AI transformation, model economics, and governance.

FOR Finance, banking, and critical government and healthcare services

Proof. We rebuilt a production federal case system (10TB) on new architecture in ~2 weeks, with two people, while its 50+ attorneys kept working. It was a second build, not a first, and the cost still fell to a tenth of the original.[10]

From the Model

WHERE AI ACTUALLY PAYS OFF

Collaborate AI's market-intelligence publication. This is Issue 01. It covers the market skeptically while carrying our own launch and rebrand rationale, with contributors from outside institutions.

FOR Leaders, founders, executives, designers, technologists, and buyers who want the latest in AI transformation, future of work, and native AI companies

Proof. Over a dozen insights from award-winning designers, engineers, and thinkers contributing from Collaborate AI, Apple, ex-Mayo, and more.

THE DESIGN ETHOS

Design is engineering desirability, and painting with the knowledge of the world.

Design works through emotion, human psychology, technology, and perception to change behavior at scale. It is also data-driven. A desirable product:

01	02	03	04	05
Easy to use	Joyful to use	High impact	Measurable	Consistently chosen

01

For good.

Technology has always come down to how it's applied. At Collaborate AI, we've dedicated our talent to problems that create lasting good in the world. Yes, profit is how we survive. But how we earn it matters deeply to us.

02

Accountable, and human-led.

We care whether AI is actually helping, not just running. Everything we build is traceable, with end-to-end governance and reporting. From drift and error rates to the moments we get it wrong, we give experts and leaders the ability to decide with their eyes wide open.

03

Enduring.

Great things last. From the ground up, we think about who will use the system long after we're gone. Every piece of context is documented up front: design artifacts, technical specs, integrations. Sustainability shouldn't cost extra. We build it in.

04

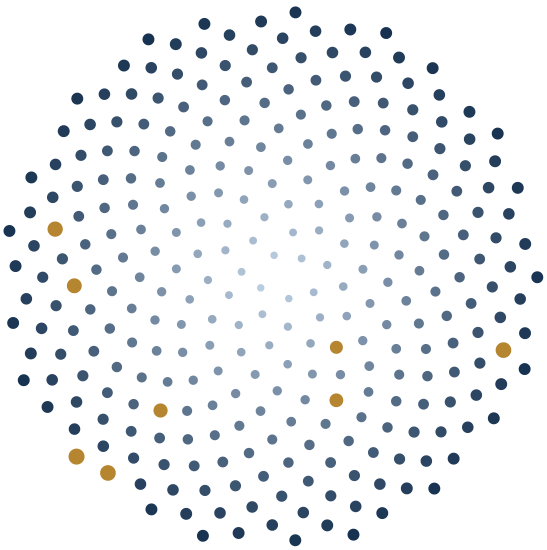
Authentic.

The best work comes from people who care about the problem itself. We work to see the world through your eyes. That's where we start, every time.

Plate I · The Mark

THE SUNFLOWER'S RULE

Our logo is grown, not drawn.

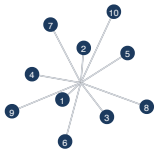


300 dots, one rule: turn 137.5°, step out by √n. Amber marks the seeds that become the next flowers.

Every dot in the mark is placed by one rule. Each new dot turns 137.5 degrees from the one before it, the golden angle, and steps out from the center by the square root of its count. The mathematician Helmut Vogel published the rule in 1979 as the model of how a sunflower packs its seeds.

No two dots ever land on the same line. The pattern runs through nature: sunflower heads, pinecones, daisies, and romanesco all grow this way.

That is who we are. Collaborate AI is a small team of senior, multidisciplinary practitioners, each owning a whole domain the way each dot owns its own space. We do not stack layers of people on a problem. The company grows the way the flower does: dot by dot, without crowding.



The first ten dots, each turned 137.5° from the last.

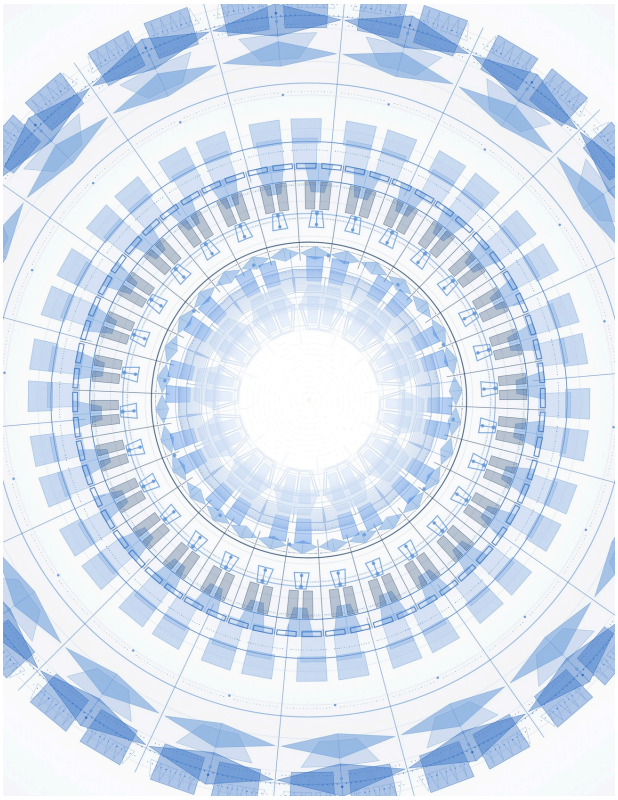
THE PROMPT · A CREATIVE BRIEF

```
Build the mark as a phyllotaxis spiral, Vogel's model: dot n at angle n × 137.5 degrees,
the golden angle, on a Fermat spiral. No outline, ragged rim. The arms should be felt, not
seen: too neat reads as a diagram, too loose reads as confetti. Blues deepening toward the
rim, one dot size with slight growth outward. The exceptions: scattered amber dots,
slightly larger, the seeds that become the next flowers. It has to hold at sixteen pixels,
embossed, and rebuild exactly from this formula.
```

Plate II · The Iris

THE HOMEPAGE ROSETTE

A homepage that opens like an eye.



Three rings of Islamic girih geometry, counter-rotating around an amber pupil.

The iris had to be abstract enough to read at a glance and intricate enough to feel deep and immersive. Traveling this past year, our designer found a new appreciation for Islamic and Moorish architecture. Its geometric repetition led to the idea: a human iris drawn as girih patterns arranged in rings around the pupil.

Each ring rotates on its own timing, and the intricacy rises toward the center. The rings match how we consult. We get close enough to see a problem as the people inside it see it. The pupil at the center carries the whole intent: We Make AI Real.

THE PROMPT · A CREATIVE BRIEF

A hero rosette in girih construction: concentric guide circles split into mirrored wedges — twelve, eighteen, thirty by ring — with hairline arcs, chords, and beads on the divisions. Three nested rings counter-rotating slowly on pure white. Deep federal navy thinning to powder blue, one amber point at center for the pupil. Flat, orthographic, plotter-drawn; the depth is the counter-rotation. Idle, it barely breathes. On scroll the pupil dilates and the center opens like a camera aperture. Rose window, aperture blades, Alhambra ceiling.

Appendix

THE NUMBERS, SOURCED

Method

Figures fall into three tiers, each held to a different burden of proof, and every source is numbered: superscripts in the article map to the rows below. Independent research and price indices are the firmest ground: Stanford's ~280x inference deflation, HBS and BCG's 758-consultant frontier study, S&P Global's abandonment survey, and O'Boyle and Aguinis's 633,263-worker star-performer meta-study. Disputed and press-derived figures are flagged where they appear: MIT NANDA's 95% / $30–40B is reported with the dispute noted, and the small-team benchmarks (Midjourney, Cursor) are company reports and contemporaneous press, marked 'reported'. Self-reported figures carry the heaviest caveats and are marked (self-reported) in the source column: the 22-hour build, the two-week case-platform rebuild, the bid restructuring, the dashboard saving, the operating-loop results, the product cycle time, and the team-and-output line (3 people, six products in 6 months, seven-figure revenue).

Figures & sources

superscripts in the article map to these numbers

1	Collaborate AI token consumption	**21 billion tokens over 6 months**	Collaborate AI (self-reported); context windows up ~10x in a year
2	Operational costs saved and reinvested	**~$1M, reinvested in senior people's time**	Collaborate AI (self-reported)
3	Adjudicate caseload / reviewers	**20,000+ appeals, 60+ reviewers, over 8 years**	Collaborate AI (self-reported)
4	Collaborate AI team and output	**3 people; six products in 6 months; seven-figure revenue**	Collaborate AI (self-reported)
5	Operating-loop results, past 90 days	**–20% prompts; –15% token cost; –8% operational cost; 7 optimization areas; 8 operational tools with expert review**	Collaborate AI (self-reported); figure, page 3
6	Bid restructuring	**20 people / 14 months → 3 senior experts with agent teams**	Collaborate AI bid (self-reported)
7	Four-source federal-data platform build	**1 supervised agent, 22 hours, 4 sources (SAM.gov, USAspending, SBA, GSA)**	Collaborate AI (self-reported)
8	Executive-response dashboard	**on the four-source platform; saved ~1.5 FTE in 2025**	Collaborate AI (self-reported)
9	BVA COTS RFI adaptation to the AMA process	**72% of requirements; checklist, user-story database, dashboards; 2 weeks, 2 people**	Collaborate AI RFI response (self-reported)
10	Adjudicate case-platform modernization	**10TB, 50+ attorneys, rebuilt in ~2 weeks by 2 people (original built by 32); second build ~1/10 the cost**	Collaborate AI case-platform account (self-reported)
11	Inference cost for a GPT-3.5-caliber answer	**$0.07 per million tokens, down from $20 two years earlier (~280x)**	Stanford HAI AI Index; a16z cites a slower ~10x/year
12	Storage cost decline	**~87% cheaper than in 2009 ($0.114 → $0.014 per GB)**	Backblaze, cost per gigabyte across 265,332 drives, 2009 – Nov 2022
13	HBS/BCG study, outside the frontier / skill compression	**–19 pts less likely to reach the right answer; below-average consultants +43% vs top +17%**	Harvard Business School / Boston Consulting Group, same study
14	Product cycle time	**~45% reduction**	Collaborate AI (self-reported); internal product development, first year of the near-total AI investment

Appendix

THE NUMBERS, CONTINUED

Figures & sources, continued

superscripts in the article map to these numbers

15	Enterprise GenAI pilots returning nothing measurable	**95% of pilots; ~$30–40B invested**	MIT NANDA initiative, 2025 survey (disputed; the piece flags it). McKinsey concurs value followed process redesign
16	Small-team leverage benchmarks	**Instagram 13 people / $1B (2012); WhatsApp 55 / $19B, 450M users (2014); Midjourney ~40 / ~$200M revenue (2023); Cursor ~20 / $100M ARR (2024)**	Company reports and contemporaneous press; Sam Altman's 'one-person billion-dollar company' (2024)
17	Star-performer concentration	**Top 5% of workers produce ~26% of output across 633,263 people; top performers up to 8x average in complex work**	O'Boyle & Aguinis, Personnel Psychology (2012); McKinsey; Netflix 'talent density' doctrine
18	HBS/BCG consultant study, inside the frontier	**+12.2% more tasks, ~25% faster, quality >40% higher**	Dell'Acqua et al., HBS Working Paper 24-013 (Sept 2023); 758 consultants (~7% of BCG's individual contributors) on GPT-4. Later published in Organization Science
19	Companies abandoning most of their AI initiatives	**42%, up from 17% a year earlier (2.5x)**	S&P Global, across 1,000+ enterprises
20	S&P 500 average company tenure	**33 yrs (mid-1960s) → ~24 (2016) → ~12 forecast (revised toward 14)**	Innosight / S&P Global; ChatGPT reached 100M users in ~2 months; direction over decimal
21	Collaborate AI cost waterfall	**contract indexed 100 → ~25; 75% saving 'illustrative', 38% 'theoretical'**	Collaborate AI site (positioning numbers, self-labeled)
22	COTS license share of first-year cost	**~1/3–40% of first-year cost; median ERP ~$450,000; setup routinely 2x the license**	Panorama Consulting Group, The 2024 ERP Report (median project cost $450,000, n=131); license/services split from COTS industry benchmarks
23	Guardrails tied to named standards	**NIST AI RMF + GenAI profile; OWASP Top 10 for LLM; OMB M-25-21 / M-25-22; VA Trustworthy AI**	Collaborate AI mapping (self-reported)
24	Uber's 2026 AI budget	**Entire annual AI budget consumed in four months, largely on coding-agent tokens**	Widely reported (TechSpot and others), 2026
25	Senior-talent leverage on AI productivity	**Skill and range, not tool access, separate teams that get returns from AI**	Fyxer, "The AI Productivity Trap," Admin Burden Index, fyxer.com (2026)

COLLABORATE AI

We make AI real.
With products that deliver results.
For the people you care about most.