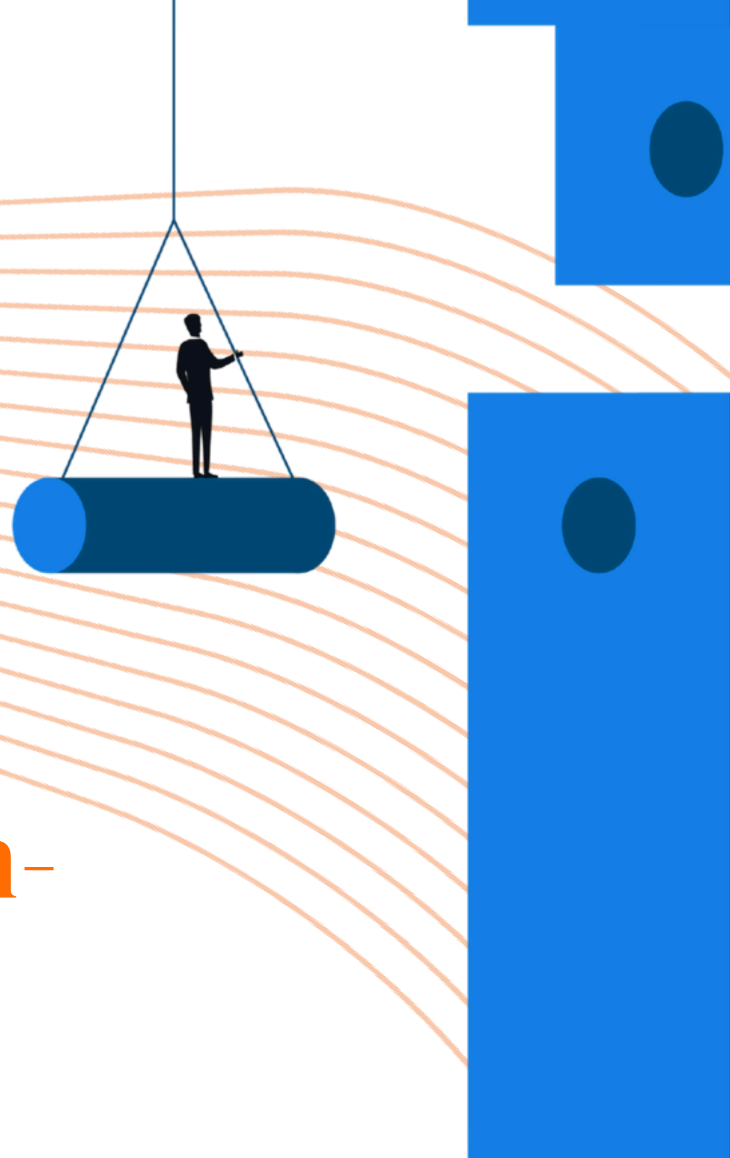
From Experimentation to Spend Discipline:

Managing Token-Based Spend



Introduction

AI token consumption is usage-based, decentralized, and difficult to forecast, but this creates an immediate opportunity for procurement to establish the visibility, governance, and commercial discipline required to scale AI adoption without suppressing innovation and experimentation.

This paper, written in collaboration with the team at Tropic, will examine both the challenges and opportunities of this unique moment and offer practical advice for developing a category strategy.

AI Spend Has Entered a New Phase

Many organizations encouraged their employees to “use AI” before establishing the policies, tools, training, or ownership models needed to understand the cost implications. Now they have to address affordability and control in an environment where usage is already widespread, and any change in access or oversight may be viewed as heavy-handed.

Michael Shields
Vice President of Procurement, Tropic

“AI tail is wagging the dog. What I mean by that is that when the invoice comes in, it needs to be paid, and then the rest of the budget is made to fit around that.”

This challenge is compounded by the fact that AI token spend tends to grow quickly and quietly, often without clear ownership or accountability. This category touches all corners of the business, including technology, finance, legal, and engineering. Nobody owns it, but everybody has a stake in its management. Budgets are forced to absorb the cost before anyone has a chance to govern the behavior that drove it.

> *The best procurement functions are modifying their processes to account for the world of AI and improving their data analysis skills to help measure total cost of ownership.*
>
> **Michael Shields**
> Vice President of Procurement, Tropic

That ambiguity and complexity may be uncomfortable for procurement, but they also create an opening. Procurement is being asked to make sense of a category that does not (yet) fit neatly into existing frameworks, and that is precisely the kind of problem the function is built to solve. Helping the business move from enthusiasm to discipline by building the visibility, governance, and commercial structures that will allow AI adoption to scale without the cost spiraling is a leadership opportunity that procurement is well placed to take on.

Why AI Token Spend Is Different from Traditional Software Spend

Consumption pricing is not new; cloud infrastructure and data platforms have used it for years. What is new is that token-based consumption is being triggered by employees across many functions, not just by specialized development or engineering teams.

When consumption models were tied to cloud infrastructure or data platforms, the people managing the spend were typically the same people generating it. AI consumption is harder to contain because it happens through chat interfaces, embedded workflows, API calls, and AI-enabled features inside SaaS tools that employees were already using. The spend is distributed, often invisible at the point of consumption, and largely disconnected from the budgets it eventually hits.

Most employees have no sense of what their usage costs or which behaviors burn through tokens the fastest. A long prompt, a complex query, a habit of regenerating outputs… these don't feel like spending decisions, but they add up in ways that compound quickly.

Utilization data has always been a challenge for procurement to work with, even when it was seat based. Variable consumption, driven by this many users across many entry points, creates a much bigger problem to overcome. Getting ahead of it will require procurement to develop genuine fluency in consumption-based pricing models, not just as a negotiation capability, but as a category management discipline in its own right.

Innovation Is Moving Faster Than the Guardrails

Organizations are telling employees to adopt AI, but education and controls have not kept up. This lag creates tension between executive mandates to adopt AI and finance/procurement mandates to control cost. When policy changes eventually arrive to address this gap, they can feel punitive to the people who were simply doing what they were encouraged to do in the first place.

Another part of the problem is an unanswered ownership question. IT is the default home for technology spend governance, but the heaviest AI token consumption often happens in engineering, product, sales, and marketing. Assigning ownership to a function that does not control the consumption is not a good governance model.

> *We're seeing the conflict in executive level remits, which is to use AI, make yourself more productive and efficient without the same kind of deployment and training on how to use it effectively and efficiently.*

Justin Etkin
Co-Founder and COO,
Tropic

Poorly managed AI spend does not stay in a technology budget line for long… it starts affecting hiring assumptions, headcount plans, and operating expense forecasts. Procurement's opportunity is to get ahead of that by helping business units translate their AI activity into proper business cases and by building budget governance structures that make spend visible before it becomes a problem. This will ensure that investment flows toward the use cases most aligned with company objectives.

Procurement's Emerging Role: From Sourcing Function to Spend Orchestrator

Procurement is well positioned to play a central role in token spend management because they already sit at the intersection of business need, supplier strategy, budget discipline, and risk management.

> *I'm all in on procurement taking ownership of this capability because to me, it really sits at the nexus of the technology strategy and general cost management across the organization.*

Justin Etkin
Co-Founder and COO,
Tropic

The impact of AI on costs and cost models also comes into play with SaaS software providers that are incorporating AI into their solutions. When there is an allowance for AI, especially for functionality that was not available when the original contract was signed, it is easy for users to run over their allotment without realizing it.

Some suppliers are using this new functionality to trigger early renewals. Companies can easily make short-sighted decisions in these cases. Procurement works to stem the overages, when what they really need to do is rethink the business need and the solutions available to meet it, and initiate a new sourcing event. That is the best way to both leverage new technology and secure good contract terms and commercial rates.

Visibility Comes First

AI token consumption can be opaque, especially when spend is spread across LLM providers, embedded AI features in SaaS platforms, and individual or departmental usage patterns.

Without a clear picture of who is using AI, for what purpose, at what cost, and against which budget, everything is guesswork. Knowing that the organization spent a certain amount on AI last quarter is not the same as knowing whether that spend generated value or simply accumulated passively. The more important question is whether consumption can be distinguished from unintentional waste.

Jacob Leichtman
Senior Director of Procurement Services, Tropic

> *"There's just more enterprise risk than there ever has been. And there's a big risk to not getting this right, to not having spend controls in place."*

Monitoring AI spend against commitments, forecasts, user groups, and expected business outcomes turns visibility from a reporting exercise into a management tool. Without that frame of reference, procurement has numbers but not insight, and numbers alone will not be enough to make the governance conversations that follow any easier.

Governance Without Gridlock

Thorough oversight does not have to shut down AI experimentation. It should create space for proof-of-concept workflows, business-case requirements, and differentiated controls based on role, use case, and business criticality. Procurement should ensure that they do not return to the 'spend police' role of the past, but instead help the organization identify and prevent wasteful spending.

In practice, that starts with asking the right questions before usage scales. For example, teams should be able to articulate their use case, the productivity gain they expect, the budget impact, and the timeframe over which they expect to see a return. That is not an unreasonable bar, and it creates the kind of accountability that makes AI investment easier to defend and optimize over time.

Proof-of-concept workflows are a useful mechanism for validating ROI before a broader deployment. Procurement should resist the temptation to respond with one-size-fits-all limits. A blanket token limit will frustrate teams like engineering or product development while doing little to address the casual usage that is actually driving waste. Controls applied at the user or department level, calibrated to the nature of the work, are considerably more effective. The goal is not to stop usage. It is to make that usage intentional.

Negotiation Matters, But It Is Not Enough

Negotiating with LLM providers is important, but procurement should not assume that negotiation alone will contain their spend growth problem. The greater opportunity is likely to come from usage optimization, education, model selection guidelines, and better contract structures.

> *Procurement should be laser focused and aware of what that spend trajectory and consumption against token looks like relative to commits that they have signed up for.*
>
> **Justin Etkin**
> Co-Founder and COO, Tropic

Although leverage with LLM providers is limited, it is not nonexistent. Multi-provider strategies can create meaningful commercial pressure because switching between models is relatively straightforward, and providers are aware of that. Procurement should use that dynamic rather than defaulting to a single-provider relationship out of convenience or inertia.

Fluency in how AI pricing works is a prerequisite for any of this to be effective. Seat-based access, API consumption, AI credits, and AI-enabled SaaS pricing are not interchangeable constructs; confusing them creates commercial exposure that shows up in invoices long after a deal is signed. Getting the commercial structure right from the beginning matters considerably more in this category than it does with traditional software spend.

Even with strong negotiation and sound contract structures, the bigger lever is likely internal. Cost savings will come from a combination of negotiation, consumption optimization, financial controls, and user education. If procurement focuses only on the supplier side of the equation, they will find that the demand side quickly outpaces whatever gains were made at the negotiating table.

Tactics for Reducing AI Token Consumption

- Match the model to the task rather than defaulting to the most expensive option available
- Use batch processing when real-time output is not required
- Cache prompts where appropriate to avoid redundant token consumption
- Train users on what drives token usage and which behaviors are most costly
- Create internal guidance around AI-efficient usage so teams can self-govern more effectively

Contracting Priorities for AI Token Spend

The contracts being signed today will define the organization's commercial flexibility for years. Getting the fundamentals right from the start is considerably easier than renegotiating from a position of dependency later.

Partner with Legal: Procurement needs to have a strong partnership with legal, in part because of the rate of change. No one is going to accept that the process must slow down for the sake of legal review, but terms are being established now that will have an impact in the future.

Plan for Continuity and Accountability: Procurement must be on the lookout for anything that might skew continuity, including data and training rights or model degradation. If procurement is going to create a sense of accountability on the part of the provider, it has to happen from the start. That may include visibility into consumption, updates, overages, and also planned model changes.

Prioritize Shorter Contract Terms: Given how quickly the market is moving, shorter contract terms are generally preferable. Procurement should avoid committing to one hundred percent of forecasted token usage, understand whether discounts continue after committed usage is exceeded, and ensure that pricing tables are codified in the contract rather than referenced through links to web pages that can be updated without notice.

Leverage Credit-Based Agreements: Rollover provisions, overage forgiveness, and the pricing of additional credits are all negotiable, and the difference between favorable and unfavorable terms in these areas can be significant. Data protection requirements, zero data retention provisions, and any industry-specific compliance obligations should also be addressed explicitly rather than assumed.

Revisit Legacy SaaS Contracts: As AI capabilities are added to products that were previously priced on simpler seat-based models, the commercial terms that made sense at signing may no longer reflect what the organization is actually getting, or what it is now being asked to pay.

Building a Culture of AI Efficiency

Enterprises don’t need to choose between AI adoption and cost control. The leaders will be those who create enough structure to make AI scalable.

Procurement has the opportunity to help the business move from enthusiasm to discipline. This means creating visibility, guiding demand, improving contracts, enabling experimentation, and helping the company understand where AI consumption creates value and where it simply creates cost.