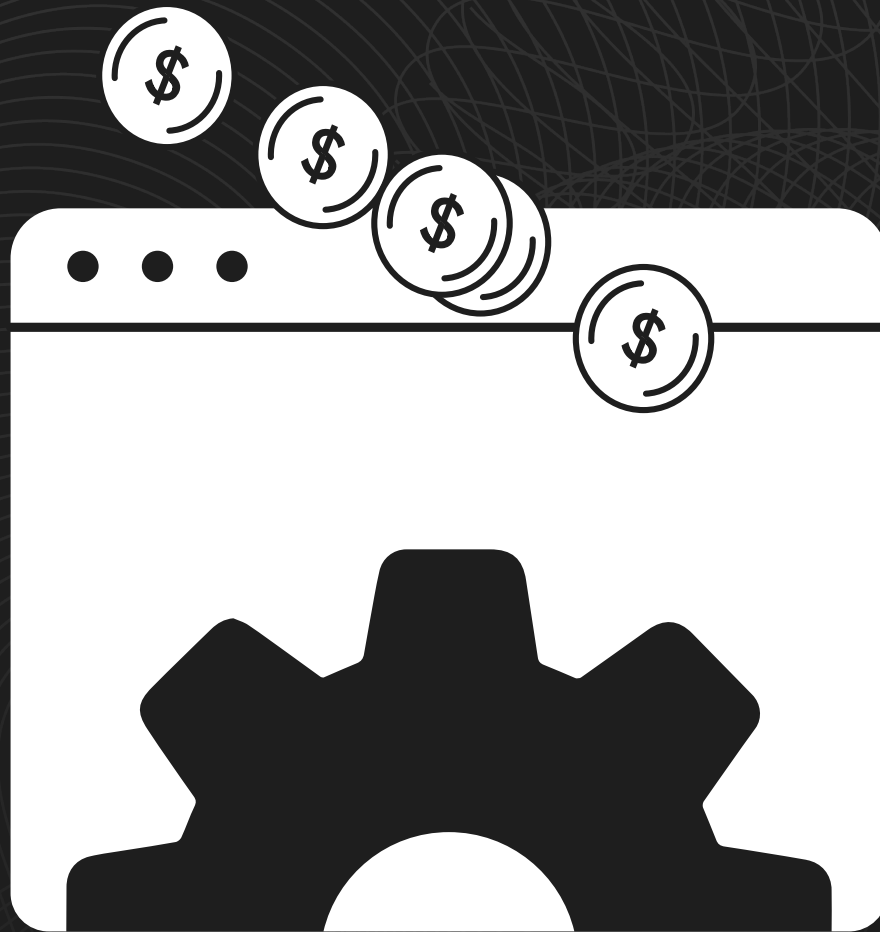




PART 4

The Real Total Cost of Ownership

K-1 BUILD TRAP



The Token Cost Reality

One of the most significant underestimations in K-1 build evaluations is the ongoing infrastructure cost of processing documents through LLM APIs at production volume. The “it’s just API calls” framing obscures what the actual cost structure looks like when applied to real K-1 documents at real volumes.

K-1 DOCUMENT CHARACTERISTICS

A K-1 package is not a short document. A typical Schedule K-1 with accompanying footnotes, supplemental schedules, and state allocations runs 8 to 15 pages. In the context of token pricing:

- **Dense financial and legal PDFs** tokenize in the range of 800 to 1,200 tokens per page.
- **A typical 10-page K-1 package generates** 8,000 to 12,000 input tokens before any system prompt or extraction schema is included.
- **Each API call also requires** a system prompt (500 to 1,000 tokens), an extraction schema, and a validation pass, adding 1,000 to 3,000 tokens per call.
- **A production-grade extraction pipeline** requires 3 to 5 separate API calls per K-1: an OCR/extraction call, a validation pass, an exception classification call, and at minimum one downstream structuring call.
- **Output tokens** — the structured JSON the system generates — are priced 4 to 5 times higher than input tokens on frontier models. A verbose structured extraction output for a complex K-1 can run 1,000 to 2,500 output tokens.

PER-K-1 TOKEN COST ESTIMATES

Using current API pricing (GPT-4o at **\$2.50/M input** and **\$10.00/M output**; GPT-4.1 at **\$2.00/M input** and **\$8.00/M output**; Claude Sonnet 4 at **\$3.00/M input** and **\$15.00/M output**):

DOCUMENT TYPE	GPT-4O	GPT-4.1	CLAUDE SONNET 4
Simple K-1 (5 pages, 3 calls)	~\$0.08–\$0.14	~\$0.07–\$0.12	~\$0.10–\$0.17
Typical K-1 (10 pages, 4 calls)	~\$0.22–\$0.38	~\$0.18–\$0.31	~\$0.27–\$0.46
Complex K-1 (15+ pages, K-3 footnotes, 5 calls)	~\$0.48–\$0.82	~\$0.39–\$0.67	~\$0.59–\$0.98
Weighted average (mixed portfolio)	~\$0.28–\$0.46	~\$0.23–\$0.38	~\$0.34–\$0.56

These estimates assume a well-optimized pipeline. Teams earlier in build maturity — without prompt caching, batching optimization, or right-sized model routing — can expect costs **1.5x to 2.5x higher** during development and early production seasons.

Pricing as published at time of analysis. Frontier model pricing changes frequently. The cost ratios in this table are more durable than the absolute figures.



VOLUME SCALE: WHAT TOKEN COSTS LOOK LIKE IN PRACTICE

ANNUAL VOLUME	YEAR 1 TOKEN COST	YEAR 2 TOKEN COST (VOLUME GROWTH)	PLANNING NOTE
1,000 K-1s	\$280–\$460	\$230–\$380	Low visibility cost
5,000 K-1s	\$1,400–\$2,300	\$1,150–\$1,900	Noticeable but not alarming
10,000 K-1s	\$2,800–\$4,600	\$2,300–\$3,800	Meaningful budget line
25,000 K-1s	\$7,000–\$11,500	\$5,750–\$9,500	Comparable to platform licensing
50,000 K-1s	\$14,000–\$23,000	\$11,500–\$19,000	Material annual infrastructure cost
100,000 K-1s	\$28,000–\$46,000	\$23,000–\$38,000	Exceeds typical platform subscriptions

Token costs are only a portion of the Total Cost of Ownership.

The Growth Problem

Token costs grow with volume, and K-1 volume in private markets is growing. The SEC reported 54,392 private funds as of Q3 2025, up 7.1% year over year, and a 2025 Executive Order directs regulators to open alternative investments to over 90 million retirement savers. An organization that builds today at 10,000 K-1s per season will process materially more within 3 to 5 years. Token costs scale linearly with that volume. Purpose-built platform pricing does not.

The pricing trajectory is the larger risk. LLM API pricing has been declining, but frontier pricing for complex structured extraction has held far steadier than pricing for simple tasks, and projections that assume continued decreases for production-grade work may be optimistic. Underneath those projections is a harder problem: today's token prices do not reflect the true cost of building and serving these models. Frontier providers are pricing aggressively to win share, and independent analyses describe current pricing as an artificially low floor that does not yet cover the enormous training and infrastructure investment behind these systems. That floor holds only while the land grab continues. Once the market settles into defined winners and losers, prices are widely expected to normalize upward, and no one can size that increase in advance.

The industry has run this play before. When enterprises moved to the cloud, early pricing was low enough to drive near-universal adoption. Once workloads and dependencies were entrenched, costs climbed, usage and egress fees compounded, and many firms are now repatriating workloads to regain control of the bill. A firm that builds its K-1 pipeline on a single provider's API inherits the same pattern: a core cost line it does not control, cannot forecast, and cannot easily exit. A purpose-built platform prices the outcome, not the token, which insulates the buyer from a cost curve set entirely by someone else.



The Full TCO

WHAT THE BUILD COSTS OVER THREE SEASONS

The following analysis compares an internal build against K1x, which brings roughly a decade of machine learning built specifically for K-1 and K-3 data, across three tax seasons for an organization processing 10,000 K-1s annually.



COST CATEGORY	BUILD (3-YEAR EST.)	K1X PLATFORM (3-YEAR EST.) ¹	NOTES
Initial build / development	\$150,000–\$350,000	\$0 (included)	Build scoping almost always underestimates footnote, K-3, and validation complexity
Engineering maintenance & enhancements	\$200,000–\$400,000	\$0 (included)	Annual IRS updates, state law changes, integration maintenance, bug cycles
Annual regulatory update cycle	\$60,000–\$120,000	\$0 (included)	Form revisions, new guidance, GILTI/UBTI/163(j) rule changes – every year
Token / API infrastructure (Yr 1 @ 10K K-1s)	\$2,800–\$4,600	\$0 (included)	Scales linearly with volume; grows 7-15% annually with K-1 volume growth
Token / API infrastructure (Yr 2 + 3)	\$6,800–\$12,600	\$0 (included)	Assumes 15% annual volume growth; cost grows faster than volume with complexity
7216 compliance architecture	\$40,000–\$80,000	\$0 (included)	Consent framework, data protection agreements, audit logging, SOC 2 alignment
Integration maintenance (CCH/GoSystem/UltraTax)	\$30,000–\$75,000	\$0 (included)	Updates across one or more tax platforms break integrations; requires dedicated maintenance cycles
K1x platform licensing (3 years)	\$0	\$120,000–\$180,000	Predictable, bundled, inclusive of all above categories
K1xCelerator adoption program	\$0	Included	Workflow design, deployment support, value realization tracking
TOTAL (3-Year Range)	\$489,600–\$1,042,200	\$120,000–\$180,000	Build TCO runs 4x – 6x platform cost when all costs are captured

¹ Illustrative platform licensing range modeled by Genius Drive for a composite organization processing 10,000 K-1s annually over three seasons. Used for TCO comparison and not a pricing proposal. Actual pricing varies by volume, modules, and term.



Several Observations on This Comparison:

The build costs shown above assume a competent engineering team executing efficiently. Teams without prior K-1 domain experience – which describes most internal teams – can expect initial build costs 40 to 60 percent higher, with additional time and cost for domain learning.

The TCO gap widens in years 2 and 3. Year 1 build costs may not be dramatically higher than platform licensing. The structural cost divergence occurs as maintenance, regulatory updates, and volume-driven token cost growth compound.

Activation and training costs are comparable between models and are not shown in the table above. They do not favor the build.

KEEP READING. [CLICK HERE FOR PART 5](#)
TIMING, DIFFERENTIATION, AND THE DECISION

About K1x, Inc.



K1x is the leading data distribution platform for alternative investment tax compliance as well as the industry standard, IRS-certified solution for seamless tax-exempt filings.

The fintech company's patented, AI-powered SaaS solution digitizes and distributes tax data seamlessly—connecting investors, accounting firms, tax software, IRS and state taxing authorities—simplifying complex processes, accelerating filings, reducing costs, and delivering greater control and

accessibility. K1x is battle-tested by the best, and trusted by more than 8000 organizations including 44 of 100 largest institutional investors in the US, 16 of the top 25 accounting firms, 11 of the top 100 private foundations, 36 of the top 100 university endowments, and 8 of the top 40 health systems.

About Genius Drive



Genius Drive develops the financial valuation models and tools that technology buyers and providers use to quantify value and total cost of ownership. The team was founded by former Gartner TCO experts and brings more than 30 years of experience building technology financial valuation models and tools. The TCO ranges in this paper follow that discipline: sourced inputs, conservative-to-aggressive ranges rather than single-point estimates, and a full accounting of structural costs, not just the visible ones.

SOURCES

TIGTA, Ref. No. 2026-308-011, March 18, 2026.

<https://www.tigta.gov>

IRS Office of Professional Responsibility, Alert 2026-19, June 24, 2026. <https://content.govdelivery.com/accounts/USIRS/bulletins/41d6e70>

SEC Division of Investment Management, Private Fund Statistics, Q3 2025. <https://www.sec.gov/files/investment/private-funds-statistics-2025-q3.pdf>

Rev. Proc. 2025-32, 2026 inflation-adjusted items. Chelli et al., J Med Internet Res, 2024;26:e53164. <https://www.jmir.org/2024/1/e53164>

Executive Order 14330, August 7, 2025.