Accel

The Four AI Ratios

There is no optimal level for AI spend — measure consumption and hold it against business performance

October 2026

Philippe Botteri, Sam Fold, Bilal Mobarik, Bodi Tent

Every technology company today is wondering if it is optimising the use of AI for their product and internal productivity, but benchmarks have been scarce and difficult to tie to business metrics. In our opinion, it is very hard to define a target for AI cost or token consumption given that:

- technology is evolving faster than anything we have seen in tech before
- token costs are decreasing quickly
- use cases are multiplying with model performance gains
- adoption is growing rapidly

How can you set targets when the foundation keeps changing at such a fast pace?

This does not mean that AI consumption cannot be measured. The focus should be on tracking consumption and understanding how increases relate to business performance. Instead of trying to define a metric to optimise towards a number, we are defining AI usage ratios that we can then tie to business metrics. Ultimately, a business is a business, and leveraging AI does not change the business metrics that drive value creation: revenue growth, churn, customer acquisition costs, EBITDA, and cash flows. What matters is how AI consumption impacts value creation.

In summary, piloting your business' AI adoption comes down to two fundamental questions:

1. Is the company maximising its AI adoption?
2. Is the use of AI driving business benefits?

The first question concerns measurement and the second concerns performance. They have to be answered in that order; you cannot tie AI to business outcomes until you can measure the AI investment.

We have developed four ratios to help answer the first question, which are set out in the next section. We will then discuss how they can be tied to business performance. All four ratios are built primarily from two inputs: the cost of token[1] and the cost of labour, which means they can be computed easily from the ledger rather than assembled through a survey.

[1] Cost of token includes the cost of model providers, the cost of AI subscriptions, as well as the cost of AI infrastructure a company uses to train or deliver open source or proprietary models.

RATIO	DEFINITION	SUB-QUESTION IT ANSWERS	HELD AGAINST
AI intensity	Cost of tokens used in the delivery of the product ÷ revenue	Is the company leveraging AI to deliver intelligent products?	Growth rate and churn rate
AI burn	Total AI costs ÷ revenue	How many points of EBITDA is the company investing in AI?	Business performance (EBITDA margin or growth)
AI labour	Cost of efficiency tokens ÷ (cost of efficiency tokens + labour)	What share of production is done by AI?	Each function's own output metric
AI maxing	Cost of efficiency tokens ÷ headcount	How much is each individual or team leveraging AI?	Peer teams/individuals in the same function

To better understand AI costs and calculate the ratios, we first need to distinguish between three AI cost buckets: (1) the tokens used to deliver the product, (2) the tokens used internally and (3) the investment made to develop proprietary models. For the purposes of this framework, these three buckets make up total AI costs

TOTAL AI COST		
DELIVERY TOKENS	EFFICIENCY TOKENS	MODEL TRAINING
The cost of tokens and inference used to deliver the product or service to customers, including the cost of any self-hosted GPUs used for this purpose	The cost of tokens used internally by employees to perform their work or automate internal processes. This represents the company's use of AI to improve internal productivity	Training and fine-tuning: The R&D cost of training or fine-tuning proprietary models during the reporting period

PART 1

Is the company maximising its AI adoption?

Let's look at each ratio and what insights they provide on the AI usage for a given company.

1) AI intensity ratio

The AI intensity ratio focuses on a business' top line and measures the extent to which a company is leveraging AI to deliver intelligent products. A technology company's product or service used to be lines of code. Now, these lines of code are complemented by intelligence, quantified in tokens. This ratio calculates how much token is used by customers to deliver the service and is using the following formula:

$$\text{AI intensity ratio} = \frac{\text{cost of tokens used in the delivery of the product}}{\text{revenue}}$$

WHAT ARE WE SEEING TODAY

If the company is a SaaS company, its ratio starts at zero and increases as its products start to include new intelligent functionality powered by AI models. If the company is AI-native, the ratio will typically be in the 50–80% range at the beginning.[2] In our view, as technology matures, the cost of token is likely to continue to decrease by roughly an order of magnitude every two years at a constant level of capability. It will therefore be interesting to see where the optimal ratio settles. SaaS companies should see their ratio increase as they develop more AI functionality, while AI-native companies should benefit from using cheaper models and more software functionality.

[2] Based on Accel's current unaudited estimates rather than a verified benchmark.

2) AI burn ratio

The second ratio focuses on the company's total AI investment. To make it easy to understand, we have defined it as the number of EBITDA points the company is investing in AI. It is calculated using the following formula.[3]

$$\text{AI burn ratio} = \frac{\text{total AI costs}}{\text{revenue}}$$

The total AI costs in the formula include the token delivered through the product, the efficiency tokens used internally, the tokens used to train or fine-tune proprietary models, and the tooling and inference infrastructure around them. As these costs are expensed as they are incurred, the ratio directly reflects the EBITDA impact of the AI investment. A company with an AI burn ratio of 6% is spending six percentage points of EBITDA margin on AI.[4] This makes the ratio an easy representation of the company's relative investment in AI.

The AI intensity ratio is included within the AI burn ratio because the tokens delivered through the product are part of total AI costs. The two ratios are therefore not independent, and the relationship between them is itself informative. Intensity divided by burn is the share of the AI investment that reaches the customer. A low figure indicates that the company is spending most of its AI budget on internal optimisation.

WHAT ARE WE SEEING TODAY

For a legacy SaaS company, the ratio is typically a few percent or less. For an early AI-native company it can exceed 100%. If the AI intensity ratio alone is 50–80% with internal token consumption and inference infrastructure on top, total AI costs can be larger than revenue. It is among the clearest arithmetic statements of what it costs to build an AI-native business today. We would expect it to compress meaningfully as these companies scale and as token prices fall.

3) AI labour ratio

The third ratio has been based on the assumption that tokens are a form of labour provided by a machine. If we consider the total work going into a company, the natural question

[3] The ratio's denominator is revenue, and it reads as EBITDA points as we consider for this analysis that all AI costs are expensed rather than capitalised

[4] This reading holds only where AI costs are expensed rather than capitalised – see the caveat in Part 2.

becomes: how much of the work is actually being done by machines rather than by people? The AI labour ratio answers this question and is calculated using the following formula:

$$\text{AI labour ratio} = \frac{\text{cost of efficiency tokens}}{\text{cost of efficiency tokens} + \text{labour cost}}$$

This AI labour ratio tells you what share of a company's total production comes from AI. Unlike the AI intensity ratio, which can vary significantly by company based on the service they deliver, this ratio is directly comparable across functions and companies.

To ensure the numbers are meaningful, token costs should include only tokens used for internal consumption. Typically, model providers combine all token costs on their invoices, so it is important to have tagged tokens separately at the source to differentiate between delivery tokens and efficiency tokens. The tagging taxonomy should be defined based on the granularity of the analysis you would like to conduct (function, team, etc.). Without proper tagging, it will be difficult to compute the ratios accurately, and the split cannot be reconstructed afterwards. The labour cost must be fully loaded.

WHAT ARE WE SEEING TODAY

Most companies today sit in the low single digit percentages. For example, a team spending $600 per month per engineer on AI tooling against a fully loaded cost of $15,000 a month runs an AI labour ratio of 4%. The most aggressive engineering organisations are pushing into double digits. We expect this ratio to increase significantly over the next few years, making it the ratio where the spread between leaders and laggards becomes most visible and comparable.

4) AI maxing ratio

The first three ratios take a broad view across the organisation. The fourth ratio zooms in and provides more granularity on how each team or individual is leveraging AI in their work. It shows how much AI is used by a person or a team, typically over a month or a year, and is calculated by the following formula:

$$\text{AI maxing ratio} = \frac{\text{cost of efficiency tokens}}{\text{headcount}}$$

This ratio produces a dollar figure per employee per month or per year, which is immediately communicable. It is not distorted by where employees sit within the organisation and shows

the relative intensity of AI spend by each person or team. Where the ratio is monitored on a monthly or quarterly basis (as shown in the Part 3 table); the efficiency token cost should be annualised so that it is consistently expressed on a per-employee-per-year basis.

The AI maxing ratio and the AI labour ratio are two views of the same underlying numbers, but they serve different purposes. The AI labour ratio is anchored around financial cost and can be compared against business output. The AI maxing ratio can be used to compare teams and individuals inside a single function and can be tied to specific team KPIs or individual Management by Objectives (MBOs).

WHAT ARE WE SEEING TODAY

In our experience, most companies are currently spending a few hundred dollars per employee per year on AI. Engineering-heavy organisations are, in our estimation, in the thousands, while the frontier is spending an order of magnitude above that. This number varies internally by team and by function, and the benchmarking should therefore be based on high- versus low-performing teams and examine how they use AI to enhance their performance. This can help raise performance across the organisation and spread best practices.[5]

If the ratio is applied to larger teams, the mean is interesting, but it should also be considered alongside the distribution.

[5] Based on Accel's current unaudited estimates rather than a verified benchmark.

PART 2

Is the use of AI driving business benefits?

The four ratios provide strong insights into how an organisation uses AI and how its usage compares to other organisations in the same industry. As we discussed in the introduction, no optimal level has emerged yet for any of them, but they can be extremely helpful as a company holds them against a business performance metric, over time, and compared to the company's historical performance. The correlation – or not – of the increase in AI usage and business performance is the finding.

> *For fast-growing start-ups, many things are changing over a short period of time, which can make it hard to isolate the benefits of AI adoption. Our benchmarking exercise addresses this by comparing companies within a sector and reading performance against AI adoption across the peer group. More at* ***ai-benchmark.accel.com***

RATIO	MEASURED AGAINST
AI intensity	Growth rate and churn rate
AI burn	Business performance (EBITDA margin or revenue growth)
AI labour	Each function's own output metric
AI maxing	Peer teams/individuals in the same function

Is AI improving the top line?

The AI intensity ratio should be measured over time against revenue growth rate: is increasing token delivery through your product or service accelerating adoption? It should also be measured against churn rate: is delivering more token through your product making it stickier and more differentiated? There is no optimal target today because the technology is changing quickly, but over time an optimal level should emerge, enabling technology companies to achieve a healthy gross margin while delivering an intelligent product or service.

Is the total AI investment driving business performance?

The AI burn ratio calculates the total investment a company is making in AI and should be measured against the main business objective the company is driving towards.

For a SaaS company trying to reaccelerate its top line, the AI burn ratio provides the EBITDA investment to assess against the additional growth rate achieved. For example, EBITDA deteriorated by 10 points, while growth accelerated by 20 points.

For a company aiming to maintain its growth rate but improve profitability, the AI burn ratio should be compared to the improvement in EBITDA margin. If AI costs rise by three percentage points of revenue and the EBITDA margin remains flat, the rest of the business has improved by three percentage points, and the AI investment has effectively paid for itself exactly. If a rise in AI costs correlates with an increase in EBITDA margin, this suggests that the AI investment is likely making the company more productive.

This is a simple way to measure if AI is driving any positive impact in the business overall. The failure case is equally legible: an AI burn ratio rising while the EBITDA margin falls or growth stays flat, is likely indicating that the investment is not being recovered anywhere in the P&L.

There are two cautions. The identity assumes AI costs are expensed rather than capitalised, so any capitalised development should be added back before the comparison works. It also attributes everything else that changed in the period to “underlying improvement”, which is why the comparison should be read over several quarters rather than one.

Is AI impacting the productivity of each function?

For this comparison the AI labour ratio should be calculated by function rather than blended, across the company, as this provides more granular insights. The use and benefits of AI vary widely by function, and each function has its own output metric to assess it against:

FUNCTION	MEASURED AGAINST	THE QUESTION IT ANSWERS
Total	(cost of efficiency tokens + labour) ÷ gross margin	Is my combined machine and human labour making the company more productive?
S&M	CAC ratio — customer acquisition cost per $ of gross margin acquired	Are my customer acquisition costs for each dollar of gross margin going down?
R&D	Number of pull requests	Is each engineer shipping more (assuming equal code quality)?
G&A	Total G&A cost as % of gross margin	Is the cost of running the company falling as a share of what the company produces?
Support	Deflection rate (tickets never touched by a human ÷ total tickets), and support cost as % of gross margin	Is AI actually resolving tickets, and is that showing up in the cost line?

Support carries two comparisons deliberately. Deflection alone can be improved by making the service worse — the ticket closes, the customer comes back through a channel you measure less carefully, and the dashboard looks better than the business. The cost line helps confirm whether the deflection represents a genuine improvement.

R&D comparison requires particular care. The number of pull requests is the most natural output measure, but it is also the one AI can inflate most directly. It is, therefore, worth pairing with the share of merged code that is reverted or rewritten within thirty days. Otherwise, the ratio and its comparison could improve while the codebase quietly degrades.

Another point to note: the four headline ratios - AI intensity, AI burn, AI labour and AI maxing - use revenue or headcount as their denominator, as defined above. However, when measuring these ratios against functional business metrics — such as CAC, G&A cost or support cost — those business comparisons should be computed with gross margin, rather than revenue. The use of AI in the product is likely to have an important impact on the gross margin of the business and gross margin — not revenue — is paying the bills.

Are the individuals and teams using AI doing better than their peers?

The AI maxing ratio is the only one of the four that is assessed across the company at a point in time rather than against its own historical performance. To use it, take a single function, divide the teams or individuals by their maxing ratio, and measure them against the same output metric. For example, compare engineering teams in the top decile of AI spend per head against those in the bottom decile based on pull requests per engineer per month, or sales pods in the top decile against those in the bottom on pipeline creation or quota attainment.

As the function, the output metric and the period are all held constant, this is the closest thing to a controlled comparison a company can run. Every other comparison in this framework is a time series, which means it is exposed to everything else that changed over the period — a pricing change, a new product, a reorganisation. The cross-sectional version is not.

It is worth keeping an eye on selection bias. From our experience, the teams that adopt AI most aggressively are frequently the strongest teams already, so a level comparison may overstate the effect of AI and understate the effect of talent. The correction is to compare trajectories rather than levels: not whether the high-maxing teams are performing better, but whether their output improved faster from the point at which their maxing ratio diverged. Where the ratio has diverged sharply and recently, that is a natural experiment worth reading carefully.

This comparison is quickly actionable, because it helps determine what to do next. If the high-maxing teams are pulling away, the constraint is likely to be diffusion, and the answer is enablement. If they are not, the constraint is the skillset and buying more tokens will not fix it.

One caveat: cost is falling

As noted above, we estimate that the price of token falls by roughly an order of magnitude every two years at constant capability. In this context, we expect that all four ratios will drift downwards over time even when usage is rising sharply. A flat AI labour ratio through a year of aggressive adoption is not stagnation — it is what progress looks like once the price decline is netted out.

For companies that want to correct for it, the token price can be fixed over the period. However, the price decrease will not affect comparison between companies at a point in time, nor the true magnitude of the business impact, as the ratios are compared against business metrics.

PART 3

Where to start?

If you would like to start using this framework for your company or start-up, our suggestion would be to focus on the following 11 metrics. You can then progress by analysing with much more granularity by function, but this is a good starting point:

LEVEL	RATIO	FORMULA	PERIOD	COMPARE TO
Company wide	AI intensity ratio	delivery tokens ÷ revenue	Quarterly	Quarterly growth and churn
	AI burn ratio	total AI costs ÷ revenue	Quarterly	Key business performance metrics: growth, churn, EBITDA or free cash flow
	AI labour ratio	efficiency tokens ÷ (efficiency tokens + labour)	Quarterly	(efficiency token + labour) ÷ gross margin
Sales & Marketing	S&M AI labour ratio	S&M efficiency tokens ÷ (S&M efficiency tokens + S&M labour)	Quarterly	CAC ratio
	S&M AI maxing ratio	S&M efficiency tokens ÷ S&M headcount	Monthly	Quota attainment and/or pipeline generation for quota-carrying teams, BDRs/SDRs and lead-gen teams
R&D	R&D AI labour ratio	R&D efficiency tokens ÷ (R&D efficiency tokens + R&D labour)	Quarterly or monthly	Total pull requests per month or quarter
	R&D AI maxing ratio	R&D efficiency tokens ÷ R&D headcount	Monthly	Pull requests per engineer per month, split by backend and front end
G&A	G&A AI labour ratio	G&A efficiency tokens ÷ (G&A efficiency tokens + G&A labour)	Quarterly	Total G&A cost as % of gross margin
	G&A AI maxing ratio	G&A efficiency tokens ÷ G&A headcount	Quarterly	Individual performance within legal, finance and HR against MBOs

LEVEL	RATIO	FORMULA	PERIOD	COMPARE TO
Support	Support AI labour ratio	support efficiency tokens ÷ (support efficiency tokens + support labour)	Monthly or quarterly	Deflection rate (tickets never touched by a human ÷ total tickets), support cost as % of gross margin and cost per ticket solved
	Support AI maxing ratio	support efficiency tokens ÷ support headcount	Monthly	Tickets solved per agent

Accel

Legal Disclosures

This document is prepared by Accel Partners Management LLP ("Accel"), a limited liability partnership incorporated in England and Wales with registered number OC303424, with its registered office at 1 New Burlington Place, London, W1S 2HR. Accel is authorized and regulated by the Financial Conduct Authority with registration number 222503.

This document is not directed at and should not be distributed, published, or used in any jurisdiction where such distribution, publication, or use would be unlawful. It may not be reproduced (in whole or in part), transmitted, modified without Accel's prior written permission. It does not constitute an offer, or a solicitation or recommendation, to purchase or sell any security, investment product or service and should not be treated as such.

This document is a thought-leadership and opinion piece, provided for information and discussion purposes only. It sets out Accel's opinions, estimates and views on certain markets, sectors, companies, trends and methodologies, including the AI ratios framework described herein. It is general in nature, does not take account of the objectives, circumstances or needs of any particular recipient or company, and must not be relied upon as the basis for any business, operational, strategic, financial, investment or other decision. The framework, ratios and any analysis described are conceptual and illustrative, may not be appropriate for any particular recipient or situation, and should not be applied or relied upon without independent verification. This document does not contain personalized recommendations or advice and should not be construed or relied on as investment, legal, business, operational, strategic, accounting, tax or other professional advice. Before taking or refraining from any action, or entering into any transaction, recipients should carry out their own independent analysis, consider carefully the suitability of any course of action to their own circumstances, and, where necessary, obtain independent professional advice in respect of the risks and any legal, regulatory, credit, tax, accounting and commercial consequences.

Nothing in this document, and no receipt or use of it, creates any advisory, fiduciary, client or other relationship between Accel and any recipient, or gives rise to any duty of care on the part of Accel to any recipient. Any recipient acting on the contents of this document does so entirely at its own risk.

The recipient is solely responsible for ensuring that any implementation of the framework, ratios or methodologies described in this document, including the collection, tagging, sharing, pooling or benchmarking of any cost, usage or performance data, complies with all laws and regulations applicable to the recipient, including competition and antitrust laws and data protection laws.

Except as otherwise indicated herein, this document speaks as of the date of publication and neither Accel nor any of its affiliates or representatives assumes any obligation to provide any recipient with subsequent revisions or updates to any historical or forward-looking information contained in these materials to reflect the occurrence of events and/or changes in circumstances after the date hereof.

This document contains forward-looking statements, projections, estimates and expectations (including as to the future cost of tokens, the evolution of the AI ratios and market and technology trends). Such statements are inherently uncertain, reflect Accel's views only as at the date of publication, are subject to change without notice and may prove to be incorrect; actual outcomes may differ materially. Any examples, figures, benchmarks and case scenarios are provided for illustrative purposes only, do not represent any actual company, result or outcome, and are not a guarantee or reliable indicator of future performance. No third party or benchmark data referred to has been independently verified by Accel.

Any benchmark, comparative or aggregated data referred to in, or made available alongside, this document is indicative only, may be based on a limited or self-reported sample, is not independently verified, may not be representative of any particular company or of the recipient's own position, and is provided without any representation or warranty as to its accuracy, completeness or fitness for any purpose. It is not a substitute for the recipient's own measurement, analysis and judgement.

Any information used in this document is believed to be reliable at the date of publication, but accuracy cannot be guaranteed. No representations or warranties are given or implied with respect to information contained in this document and no reliance should be placed on the contents of this document. To the fullest extent permitted by law, each recipient agrees that neither Accel nor its members, partners, directors, officers, employees and agents shall have any liability for any misstatement or omission of fact or any opinion expressed herein, or for any loss or damage of any kind arising from any use of, or reliance on, this document or its contents, including the framework, ratios, methodologies and examples described in or accompanying it. Nothing in this document excludes or limits any liability that cannot be excluded or limited under applicable law.

This document, and any dispute or claim (including any non-contractual dispute or claim) arising out of or in connection with it or its subject matter, including the disclaimers, exclusions and limitations set out above - are governed by, and shall be construed in accordance with, the law of England and Wales.