FAST
FORWARD
XPERIENCE

web summit
LISBON

DESTAQUES DE NATÁLIA SCHIFINO, CURADORA DA FFX

PALCO	DIA	LISBOA
AI SUMMIT	**10 NOV 2026**	**13:40**

A CONTA DA IA CAI QUANDO A EMPRESA ESCOLHE, NÃO QUANDO O TOKEN BARATEIA

IA · IA AGÊNTICA · EFICIENCIA · GOVERNANCA DE DADOS

Arvind Jain

Fundador e CEO da Glean, plataforma de IA corporativa que conecta busca interna, assistente e agentes sobre os dados da empresa.

Muhammad Lila

Founder, Goodable

EM 30 SEGUNDOS

Arvind Jain, CEO da Glean, plataforma de IA corporativa, vê a IA disputando orçamento com gente pela primeira vez, e a conta subindo mês a mês mesmo com o token mais barato.

Quase toda empresa com que a Glean conversa sente que gasta demais com IA sem retorno; para Jain, o retorno vem quando a empresa para de dar ferramentas a todos e escolhe onde usar.

O OLHAR DA CURADORIA

Natália Schifino

Curadora da FFX

DESTACOU a economia de tokens que vem de investir na escolha do modelo certo para cada tarefa.

APONTOU que montar os agentes de cima para baixo aumenta a eficiência: pensar nos cinco que de fato mudam o negócio e trabalhar neles.

NÚMEROS DO PALCO

95%

Tarefas corporativas que modelos 10 a 20 vezes mais baratos resolvem

mais de 90%

Corte no custo de IA com dados prontos ao agente e modelo certo por tarefa

90% a 95%

Trabalho feito com IA que não agrega valor, para Jain

PARA LEVAR

1. **Troque a ferramenta de IA para todos por uma lista curta de agentes, decidida pela diretoria.**

2 **Mande cada tarefa ao modelo mais barato que dá conta dela e deixe o mais caro para o que exige.**

3 **Entregue ao agente os documentos e dados da empresa já organizados, para ele não gastar tokens procurando.**

O ARTIGO

3 MIN DE LEITURA

Esperar que a IA fique mais barata não reduz a fatura de uma empresa: o uso cresce mais rápido que o preço cai.

O diagnóstico é de Arvind Jain, fundador e CEO da Glean, plataforma de IA corporativa, em entrevista no AI Summit do Web Summit Lisboa. **Natália Schifino, curadora da FFX, destacou duas saídas dele: o modelo certo para cada tarefa e os agentes escolhidos de cima para baixo.**

O preço por unidade de inteligência vai continuar caindo, admite Jain, inclusive com modelos abertos feitos para tarefas específicas. Mesmo assim, a IA chegou a um lugar inédito no orçamento: "É a primeira vez que tratamos software e tecnologia no mesmo patamar das pessoas em termos de quanto custam." Fora do palco, uma simulação da Accenture de julho de 2026 mostra o efeito: se o token (a unidade em que os modelos cobram) ficasse 25% mais barato, só 15% das organizações aproveitariam a economia.

A primeira saída é gastar menos tokens escolhendo o modelo. Para Jain, quase todas as tarefas de uma empresa cabem em modelos que custam um décimo ou um vigésimo dos mais caros. A segunda é o contexto. Um assistente ligado aos sistemas da empresa gasta tokens só para achar documentos e dados antes de trabalhar; se recebe esse material pronto, organizado numa camada de contexto, consome metade. **Somadas, as duas frentes cortam mais de 90% da conta, segundo Jain.** Vale lembrar que a camada de contexto é o que a Glean vende.

A economia com o modelo certo, porém, pode não chegar à fatura. A própria Accenture conta que sua plataforma interna manda cada tarefa ao modelo adequado por cerca de um sexto do custo dos modelos mais potentes. **Mesmo assim, quando o padrão era um modelo intermediário, grande parte dos funcionários ainda escolhia o mais potente.**

Cinco passos para a fatura de IA encolher

1 **Mapear**
Achar os processos internos que mais custam

2 **Priorizar**
Fechar uma lista de cinco agentes

3 **Modelo**
Usar o mais barato que dá conta da tarefa

4 **Contexto**
Entregar dados prontos e gastar metade dos tokens

5 **Resultado**
Conta mais de 90% menor, segundo Jain

Quando cada funcionário decide como usar a IA, o resultado é desperdício. Jain ouve com frequência que 80% do trabalho feito com IA não agrega valor e avalia que a fatia chega a 90% ou 95%: documentos de 50 a 100 páginas, agentes e habilidades de baixa qualidade. A saída que Natália apontou é definir os agentes de cima para baixo, concentrados nos cinco que de fato pesam no negócio.

Arvind Jain, da Glean, conversa com Muhammad Lila no palco do AI Summit. Foto: curadoria FFX no Web Summit 2026

> **QUANDO VOCÊ MUDA SUA ESTRATÉGIA PARA SER MAIS INTENCIONAL, MAIS DE CIMA PARA BAIXO, É AÍ QUE COMEÇA A TER RETORNO REAL EM VALOR DE NEGÓCIO.**
>
> **Arvind Jain**, Founder & CEO da Glean

O método de Jain começa pelos processos internos mais caros e termina numa lista curta: "estes são os cinco agentes que precisamos construir para automatizar estes cinco processos de negócio diferentes". **Jain reconhece o custo da escolha: o uso espalhado, mesmo cheio de lixo, ensina as equipes e muitas vezes as faz trabalhar mais rápido.**

Quem decide o uso decide a conta

IA de cima para baixo	**IA aberta a todos**
Decisão Quem paga escolhe processos, modelo e contexto	**Decisão** Cada funcionário escolhe sem ver a fatura

Foco Cinco agentes ligados ao que pesa no negócio	**Foco** Até 95% do trabalho sem valor
Retorno Valor de negócio mensurável	**Ganho** Equipes aprendem e às vezes aceleram

Para a FFX, os dois pontos de Natália levam à mesma pergunta: quem decide como a IA é usada. Na IA de cima para baixo, quem responde pelo custo escolhe os processos, o modelo e o contexto. Na ferramenta aberta a todos, escolhe quem não vê a conta. Enquanto for assim, cada queda no preço do token vira mais uso, e a fatura volta a subir.

LEITURA FFX

O preço do token é do fornecedor; quantos tokens se gastam, e em que modelo, é decisão da empresa.

No Brasil, poucos olham essa conta: só 45% das empresas acompanham o custo por token, embora 90% esperem consumir mais nos próximos 12 meses (pesquisa IDC para a Dell, Mobile Time, 2026).

O C6 Bank montou uma estrutura própria que troca de modelo sem depender de um único fornecedor e reserva os modelos mais caros às tarefas complexas, porque usá-los em tudo seria desperdício (Mobile Time, 2026).

Antes de liberar mais verba, faça como o C6: defina qual modelo cada caso de uso pode chamar e cobre de cada área o custo por token do que ela consome.

PROVOCAÇÃO FFX

SE A SUA EMPRESA SÓ PUDESSE TER CINCO AGENTES DE IA, VOCÊ SABERIA DIZER QUAIS?

RETRATO DA EMPRESA **GLEAN**

Plataforma de IA corporativa que conecta documentos, e-mail, CRM e tickets da empresa e oferece busca, assistente e agentes sobre vários modelos de linguagem.

US$ 300 milhões de receita recorrente anual em 2026, o triplo de 15 meses antes (Latent Space, 2026)

US$ 200 milhões de receita recorrente anual em dezembro de 2025 (Fortune, 2025)

Avaliação de **US$ 7,2** bilhões na Série F de US$ 150 milhões, em junho de 2025 (Glean, 2025)

Leandro Lima, ex-Snowflake Brasil, contratado como VP para a América Latina em julho de 2026 (Baguete, 2026)

FONTES

1. Accenture: IA Tokenomics para geração de valor empresarial (29 jul. 2026)
2. Mobile Time: Quase 60% das empresas brasileiras usam IA, aponta pesquisa da Dell com IDC (5 nov. 2026)
3. Mobile Time: Gasto com tokens preocupa empresas no uso da IA (24 jul. 2026)
4. Latent Space: Glean model routing (ago. 2026)
5. Fortune: Glean hits $200 million ARR (8 dez. 2025)
6. Glean Press: Série F de US$ 150 milhões a US$ 7,2 bilhões (jun. 2025)
7. Baguete: Glean agora tem VP Latam (17 jul. 2026)
8. IBM, Cost of a Data Breach Report 2026, recorte Brasil (Canaltech, 15 out. 2026)