WHITEPAPER · V1.0 · JUNHO 2026

Auditable Trust Infrastructure (ATI)
Uma Nova Categoria para Governança de IA Generativa em Ambiente Regulado

A Technical Whitepaper on Verifiable AI Evidence Infrastructure

FoundLab · Trust by Physics.
VERIFIABLE AI EVIDENCE · SÃO PAULO · BRASIL

ABSTRACT

Firewalls protegem tráfego, mas não provam decisões de IA nem o consentimento do cliente. Este whitepaper apresenta a **Auditable Trust Infrastructure (ATI)** como nova categoria de infraestrutura: uma camada única que governa simultaneamente as superfícies de voz e texto, emite prova-de-decisão verificável (*hash-chained, time-stamped*) e o faz sem reter dados pessoais (conformidade LGPD) nem expor material criptográfico de chave (CMN 5.274). Argumentamos que a ATI não substitui IA — ela a torna auditável.

HASH-CHAINED | ZERO PII RETENTION | TIME-STAMPED | VERIFIABLE PROOF | CRYPTOGRAPHIC AUDIT

PREPARADO PARA

Bancos Tier 1 — CISO, Legal / DPO, Arquitetura
Equipe Google Cloud

VERSÃO & DATA

v1.0 · Junho 2026
DOI: foundlab/ati.2026.01

ATI · AUDITABLE TRUST INFRASTRUCTURE

Resumo Executivo

Firewalls protegem o tráfego. Não provam a decisão.

A regulamentação do Banco Central reconheceu essa lacuna e a tornou matéria de conformidade obrigatória. O prazo de adequação voluntária encerrou-se em 1º de março de 2026 e estabelece que interações por voz e texto passam a exigir o mesmo nível de prova auditável — exigência simétrica que a maioria das instituições reguladas ainda trata de forma fragmentada.

Firewalls, IDS e criptografia em trânsito operam sobre o plano de transporte: protegem o perímetro, mas não constituem prova do que um modelo de IA decidiu, nem do que o cliente compreendeu e aceitou. Nenhuma camada existente de segurança gera evidência forense, verificável por terceiros, sobre o conteúdo semântico de uma decisão automatizada.

A tese deste documento é que a separação entre proteção de rede e prova de decisão é **estrutural, não incremental**. A ATI define, portanto, uma nova categoria: a infraestrutura de confiança auditável para sistemas de IA regulados.

Apresentamos a ATI — *Auditable Trust Infrastructure* —, infraestrutura unificada de evidência que governa ambas as superfícies sob um único registro criptograficamente verificável. A ATI não retém dado pessoal do cliente — em aderência à LGPD — e não expõe material criptográfico sensível ao operador, em conformidade com a Resolução CMN 5.274. O registro é mínimo, imutável e suficiente para reconstrução probatória.

Três requisitos definem a categoria:

01 · Prova

Decisão verificável

Cada decisão automatizada gera evidência semântica verificável por terceiros.

02 · Não-retenção

LGPD por construção

O registro probatório não armazena dado pessoal do titular, preservando os direitos da LGPD.

03 · Custódia

CMN 5.274

O material criptográfico permanece isolado do operador, sob o regime exigido pela CMN 5.274.

PROVA DE DECISÃO | VOZ + TEXTO | LGPD | CMN 5.274 | PRAZO 01/03/2026

WHITEPAPER · CONTEÚDO

Sumário

I

SITUATION · COMPLICATION · QUESTION

O Problema e a Âncora Regulatória

Onde a segurança de rede termina e a prova de decisão começa.

ABSTRACT

Esta seção delimita a lacuna regulatória entre a segurança da rede e a evidência da decisão: enquanto a postura defensiva amadurece, a prova auditável de decisões automatizadas permanece em aberto. Ancoramos a análise no arcabouço normativo brasileiro — Resolução CMN 5.274, Resolução BCB 538 e o Art. 20 da LGPD — que exige explicabilidade, revisão humana e trilhas verificáveis. Formulamos então a questão central: como provar uma decisão de IA sem reter dados sensíveis e sem expor chaves criptográficas.

I · O PROBLEMA · 1.1

1.1 · O imperativo da prova

Firewall protege tráfego — não prova o que a IA decidiu nem o que o cliente aceitou. A camada de transporte está resolvida há duas décadas; a camada de evidência semântica do ato decisório, não.

A segurança de perímetro tradicional — WAF, TLS, IAM — opera no plano de transporte: garante *confidentiality* e *integrity* da requisição em trânsito, autentica chamadores e segrega redes — mas não captura o conteúdo semântico do ato decisório, nem a postura do sujeito de dados frente a ele. O pacote chega íntegro; a inferência que ele dispara permanece opaca.

Quando o regulador volta seis meses depois — não por incidente de rede, mas por reclamação de titular ou auditoria de modelo — a evidência exigida não é o log de firewall, nem o trace de SIEM, nem o snapshot do IAM. É outra: **o quê, o quando e o com base em quê** a inferência foi produzida — e se o titular consentiu com aquela finalidade declarada, naquele momento, sob aquela base legal.

Três exigências, todas distintas, todas simultâneas, nenhuma satisfeita pela pilha de segurança herdada da era cliente-servidor. **Prova** demanda registro imutável do par decisão↔contexto. **Não-retenção** impõe demonstrar consentimento sem manter o dado pessoal além da finalidade — restrição que colide frontalmente com o instinto de engenharia de armazenar tudo. **Custódia** exige guarda regulatória por prazo legal, em depositário segregado, com cadeia verificável por terceiros independentes.

Nenhuma das três é atendida por firewall, SIEM ou observabilidade aplicacional. Todas dependem de uma camada de prova nativa ao ato — emitida no instante da decisão, atestada criptograficamente, desacoplada do dado-fonte. Sem ela, a instituição responde por algo que não consegue reconstruir.

01 · Prova

O que a IA decidiu

Registro imutável do par decisão↔contexto: input, hash do modelo, versão do prompt, parâmetros de inferência e output, selados com timestamp criptográfico. Sem trilha, o ato é irreproduzível.

02 · Não-retenção

LGPD · o que o titular aceitou

Provar consentimento, finalidade e base legal sem reter o dado pessoal além do necessário. O titular fica representado no log por compromisso criptográfico de aceite, não pelo PII bruto.

03 · Custódia

CMN 5.274 · guarda regulatória

Guarda da evidência por prazo legal mínimo de 5 anos, em custodiante segregado da instituição produtora, com cadeia verificável por auditor independente e pelo BACEN.

Três exigências distintas — uma só camada de prova, emitida no instante do ato.

I · O PROBLEMA · 1.2

1.2 · Âncora regulatória

A norma do Banco Central já tornou a evidência técnica obrigatória — prazo de adequação em 01/03/2026. CMN 5.274 e BCB 538 reescrevem o regime de cibersegurança (custódia de chaves, integridade e rastreabilidade); a LGPD Art. 20 ancora a explicabilidade de decisões automatizadas.

O arcabouço regulatório brasileiro consolidou, em poucos anos, exigências formais de rastreabilidade e auditoria aplicáveis a instituições financeiras e a operadores de dados pessoais. A Resolução CMN nº 5.274/2025 reescreveu a política obrigatória de cibersegurança do sistema financeiro nacional — emendando a CMN 4.893/2021 —, e a Resolução BCB nº 538/2025 é sua norma espelho para instituições de pagamento, corretoras e distribuidoras (emendando a BCB 85/2021). Não se trata de regime de contratação de nuvem: ambas reescrevem o regime de cibersegurança, introduzindo o Art. 3º-A com catorze controles mínimos verificáveis. Entre eles, a vedação de acesso de terceiros — inclusive provedores de nuvem — às chaves privadas de assinatura (Art. 3º-A, I, 'f') e a validação de integridade transacional fim a fim antes da assinatura digital (Art. 3º-A, I, 'e').

Paralelamente, o Artigo 20 da LGPD instituiu o direito do titular à revisão de decisões automatizadas — pressupondo evidência técnica capaz de reconstituir o critério aplicado em cada decisão. A consequência operacional é direta: instituições reguladas devem manter trilhas verificáveis sobre cada inferência de modelo, sob pena de sanção administrativa e perda de capacidade probatória. A janela de adequação voluntária encerrou-se em 01/03/2026; o que resta às instituições que operam GenAI é remediar o passivo acumulado e instaurar o regime permanente de rastreabilidade exigido por norma vigente.

CMN 5.274

Política obrigatória de cibersegurança — governança, registro de incidentes e verificação contínua de fornecedores críticos.

VIGENTE

BCB 538

Regime de cibersegurança — norma espelho da CMN 5.274 para IPs e corretoras. Art. 3º-A: vedação de acesso a chaves privadas + integridade fim a fim.

PRAZO 01/03/2026

LGPD Art. 20

Direito à revisão de decisões automatizadas — exige explicabilidade e rastreabilidade técnica reconstituível por decisão.

VIGENTE

Nota. Cada referência é citação verificada contra a fonte oficial. A evidência auditável deixou de ser recomendação: tornou-se requisito normativo vigente para qualquer instituição que opere sistemas de GenAI sobre dados pessoais ou decisões financeiras.

I · O PROBLEMA · 1.3

1.3 · A pergunta central

Como provar o que a IA decidiu sem reter o dado nem expor a chave?

$$\text{provar}(\text{evidência}) \;\wedge\; \neg\text{reter}(\text{LGPD}) \;\wedge\; \neg\text{expor}(\text{CMN 5.274}) \;=\; ? \qquad (1)$$

A formulação acima encapsula três restrições simultâneas que, isoladas, são tratáveis, mas em conjunto definem o problema central deste trabalho. **Primeiro**, a evidência da decisão algorítmica deve ser criptograficamente verificável e reproduzível: qualquer auditor — interno, regulador ou contraparte — precisa recomputar o veredito a partir de artefatos assinados e obter o mesmo resultado, sem confiar no operador. **Segundo**, a LGPD (art. 6º, III e art. 16) proíbe a retenção de dados pessoais além da finalidade declarada, o que invalida a estratégia trivial de arquivar entradas e saídas em claro para auditoria posterior. **Terceiro**, a Resolução CMN 5.274/2025 veda a exposição de material de chave criptográfica a operadores humanos ou processos não-atestados, eliminando custódia direta como vetor de prova.

As três cláusulas parecem mutuamente exclusivas — provar exige reter, e reter exige expor — mas reconciliam-se em uma única camada de infraestrutura que desloca *evidência* para compromissos criptográficos, *retenção* para hashes determinísticos, e *custódia* para enclaves atestados.

Três restrições aparentemente contraditórias.
A resposta exige reconciliá-las em uma única infraestrutura.

II

ONE INFRASTRUCTURE, TWO SURFACES

A Tese: Auditable Trust Infrastructure

Uma categoria cujo objeto não é o sistema, mas a evidência verificável do seu comportamento.

ABSTRACT

Esta seção define *Auditable Trust Infrastructure* (ATI) como uma nova categoria de infraestrutura — distinta de observabilidade, governança e segurança — cujo objeto não é o sistema, mas a evidência verificável de seu comportamento. Apresentamos a arquitetura de referência em quatro estágios canônicos: **ingest → verdict → receipt → custody**, na qual cada decisão produz um recibo criptograficamente assinado, reproduzível e admissível como prova. Formalizamos o modelo de ameaças (adversário interno, regulador adversarial, falha em cascata) e estabelecemos as garantias de completude, solidez e não-repúdio que distinguem ATI de logging tradicional.

II · A TESE ATI · 2.1

2.1 · Definindo a categoria

A ATI não é mais uma ferramenta de governança — é a camada de prova física da decisão.

DEFINITION 1 · AUDITABLE TRUST INFRASTRUCTURE

Auditable Trust Infrastructure (ATI) é a camada infraestrutural que produz, sela e preserva *cryptographic evidence* de cada decisão automatizada — convertendo confiança declarada em *verifiable proof-of-decision*, reproduzível e admissível em foro técnico, regulatório e jurídico.

A indústria de *AI governance tools* — observabilidade de modelos, *policy management*, *drift monitoring*, *audit logging* — opera sobre uma premissa descritiva: coleta-se telemetria após a decisão, agrega-se em dashboards, e produz-se relatórios narrativos sobre o comportamento do sistema. O artefato resultante é prosa estruturada, não prova. Logs podem ser truncados, reescritos ou contestados; políticas podem ser declaradas mas não enforçadas no caminho crítico; relatórios de compliance são auto-atestados pelo próprio operador. Esta camada é reativa — descreve o que o sistema fez, sem garantir que o descrito é o que de fato ocorreu.

A ATI inverte a premissa. Em vez de descrever a decisão depois do fato, a infraestrutura **prova a decisão no momento em que ela ocorre** — selando inputs, modelo, política aplicada e output em um artefato criptográfico imutável, ancorado em *hash chain* e assinatura, reproduzível por terceiros sem acesso ao sistema produtor. É uma camada proativa: a atestação é pré-condição da execução, não consequência. Onde *governance tools* entregam telemetria, a ATI entrega *proof-of-decision* — o equivalente computacional de uma testemunha técnica permanente, admissível e independente do operador.

CATEGORIA DESCRITIVA

AI governance tools

Observabilidade, policy management, audit logging e drift monitoring descrevem o comportamento *ex post*. Telemetria reativa, auto-atestada pelo operador, sem garantia de integridade nem reprodutibilidade externa. Útil para operações; insuficiente para prova.

CATEGORIA PROBATÓRIA

ATI · FoundLab

Atestação criptográfica no caminho crítico da decisão: hash chain, selo, assinatura e proof-of-decision verificável por terceiros. A FoundLab instaura esta categoria — infraestrutura de prova como pré-requisito da autonomia.

II · A TESE ATI · 2.2

2.2 · Uma infraestrutura, duas superfícies

Voz e texto exigem a mesma prova — uma só infraestrutura governa as duas. A camada ATI separa as superfícies de captura sem fragmentar a evidência.

REX VOICE · SUPERFÍCIE A

Aceite por voz — consentimento humano

Captura, transcreve e vincula o consentimento falado a uma evidência assinada. Domínio: interações reguladas em que uma pessoa manifesta vontade — vendas, contratação, opt-in, declarações verbais com efeito jurídico (CDC, LGPD art. 8º, SUSEP).

REX GUARD · SUPERFÍCIE B

Decisão de IA (Gemini) — atestação algorítmica

Audita cada inferência do modelo e registra a justificativa da decisão. Domínio: atos em que um sistema delibera — escoragem, triagem, recusa, recomendação automatizada (LGPD art. 20, EU AI Act art. 13–14, Res. CMN 4.893/2021).

Princípio MECE. As duas superfícies são mutuamente exclusivas: REX Voice atesta o consentimento humano expresso (input biométrico-linguístico, sujeito de direito); REX Guard atesta a decisão algorítmica (output computacional, sujeito automatizado). Nenhum evento pertence a ambos os domínios — o aceite é causa, a decisão é efeito, e o limite entre eles é o instante em que o controle passa do humano ao modelo. São também coletivamente exaustivas: juntas, cobrem o conjunto integral de superfícies de interação reguladas que uma instituição opera com seu cliente. Toda obrigação probatória derivada de LGPD, CDC, BACEN, SUSEP e AI Act recai necessariamente sobre α, β, ou a composição α→β.

CAMADA ATI · AUDITABLE TRUST INFRASTRUCTURE

O mesmo selo criptográfico (SHA-256 + assinatura Ed25519), a mesma trilha imutável (log *append-only* ancorado em timestamping RFC 3161) e a mesma prova verificável por terceiros se aplicam a α e β. Duas entradas distintas, um único alicerce de confiança — superfícies MECE, evidência unificada.

II · A TESE ATI · 2.3

2.3 · Arquitetura de referência

A arquitetura separa veredito, recibo e custódia em quatro camadas verificáveis de ponta a ponta. Cada camada possui interface mínima, contrato criptográfico explícito e responsabilidade única — de modo que falhas permanecem contidas no próprio escopo sem invalidar as provas das demais.

TABELA 2.1 — QUATRO CAMADAS CANÔNICAS: INGEST → VERDICT → RECEIPT → CUSTODY

CAMADA	FUNÇÃO	DESCRIÇÃO	INVARIANTE
01	**Ingestão de evidência**	Normaliza fontes heterogêneas e fixa o hash canônico sem reter o payload bruto. O conteúdo trafega pelo pipeline apenas em memória; somente o digest é persistido.	`payload_persisted = false`
02	**Motor de veredito**	Avalia regras determinísticas sobre o digest e emite um estado de confiança categórico. Lógica versionada e reproduzível: terceiros recomputam o mesmo veredito a partir da mesma evidência.	`verdict_engine.eval()`
03	**Geração de recibo**	Assina o veredito, o hash da evidência e o timestamp, produzindo prova compacta e auditável. O recibo é o único artefato exportado para fora da fronteira de confiança.	`sign(receipt, kid)`
04	**Custódia de chave**	HSM isolado em domínio físico distinto; as chaves de assinatura nunca atravessam o limite seguro. Assinaturas são invocadas por API restrita e auditadas em log imutável anexo.	`key_export = denied`

$$\text{receipt} = \text{sign}(\ H(\text{evidence}) \parallel \text{verdict} \parallel \text{ts}\ ,\ \text{sk\_custody}\ ) \tag{2}$$

Cada camada expõe somente hashes e assinaturas para a interface seguinte: a ingestão entrega `H(evidence)`, o motor entrega `verdict`, a geração entrega `receipt` assinado, e a custódia jamais entrega `sk_custody`. O conteúdo sensível permanece fora da fronteira de custódia, e o verificador externo reconstrói a cadeia probatória apenas a partir do par `(receipt, pk)`, sem necessidade de acesso à evidência original — garantindo verificação independente, reproduzível e resistente a comprometimento parcial.

II · A TESE ATI · 2.4

2.4 · Threat model

O threat model assume adversário interno e custodia o raciocínio, não só o dado.

Adotamos uma postura *zero-trust* em sentido forte: nenhum componente é confiável por padrão — nem o operador da infraestrutura, nem o canal de transporte, nem o runtime que executa o código. A confiança não é concedida por contrato ou reputação, mas construída *ex post* via verificação criptográfica. A atestação remota via *Trusted Execution Environments* (TEE) ancora a auditoria no silício: o auditor valida medidas de hardware, não promessas de quem opera a máquina [4].

TABELA 2.2 — SUPERFÍCIES DE ATAQUE, SUPOSIÇÃO DE CONFIANÇA E MITIGAÇÃO

SUPERFÍCIE	SUPOSIÇÃO DE CONFIANÇA	MITIGAÇÃO
Operador interno	Custódia não confiável; privilégios administrativos sobre host, hipervisor e armazenamento de logs — capaz de adulterar, reordenar ou suprimir registros pós-fato.	Raciocínio selado em enclave TEE com memória cifrada; atestação remota assinada pelo fabricante; logs encadeados fora do alcance do operador antes de qualquer materialização em disco.
Canal de I/O	Rede hostil; trânsito interceptável ou rejogável; intermediários podem injetar, omitir ou reordenar mensagens entre cliente, enclave e armazém de evidências.	Cadeia de evidências encadeada por hash (Merkle-log) com âncoras temporais; cada entrada assinada dentro do TEE e verificável *end-to-end* sem confiar no transporte.
Pilha de execução	Runtime, kernel e firmware potencialmente comprometidos; bibliotecas de inferência ou pesos do modelo substituídos por versões adulteradas antes do carregamento.	Atestação de medida do código e dos pesos (hash do binário + manifest); quote assinado pela raiz de confiança do silício e validado pelo auditor antes de aceitar qualquer saída como evidência.

A confiança reside na verificação criptográfica, não na boa-fé do operador: o auditor reconstrói o raciocínio sem precisar confiar em quem o executou, em quem o transportou ou em quem o armazenou. Cada elo da cadeia é redutível a uma medida de hardware ou a uma assinatura encadeada — e, portanto, falseável de forma independente.

III

VOICE EVIDENCE INFRASTRUCTURE

REX Voice — Infraestrutura de Evidência de Voz

O aceite por voz vira evidência irrefutável — sem reter a fala.

ABSTRACT

REX Voice captura, transcreve e vincula criptograficamente o consentimento falado a um artefato de evidência assinado, sem reter o áudio subjacente. Esta seção traça o pipeline da captura à atestação, especifica o fluxo técnico — componentes, protocolo e superfície de assinatura — e conclui com um console de demonstração para verificação interativa do artefato resultante.

III · REX VOICE · 3.1

3.1 · O aceite por voz como evidência

O aceite por voz vira evidência irrefutável — sem reter a fala.

REX Voice opera como uma camada de prova-de-consentimento sobre qualquer jornada de voz do banco. No momento em que o cliente verbaliza o aceite, o áudio é capturado em *stream* e roteado para um modelo de transcrição em Vertex AI, que devolve a intenção declarada em texto. Esse *transcript* é então concatenado ao timestamp autoritativo do servidor (`ts`) e ao contexto da operação — produto, taxa, prazo, identificador da sessão (`ctx`) — formando o payload canônico. Sobre essa string calcula-se um hash SHA-256, assinado com a chave privada do banco (HSM). O áudio bruto é descartado em memória assim que a transcrição termina: nada é persistido, nada trafega para storage. O artefato auditável é a assinatura — não a voz.

A prova é verificável por terceiros — perícia, Bacen, defensoria — apenas reapresentando o *transcript* e validando a assinatura contra a chave pública. LGPD-*compliant* por construção: minimização de dados (sem retenção de biometria vocal), finalidade explícita (evidência de aceite) e auditabilidade integral preservada. O banco prova o que o cliente disse, sem nunca ter guardado *como* o cliente disse.

```
proof = sign( H(transcript || ts || ctx) )
```
(3)

E DAÍ, PRA MIM?

PARA O BANCO

Destrava jornadas represadas pelo jurídico

Contratação de crédito, renegociação, portabilidade e alteração de limite passam a rodar por voz com defesa probatória robusta. Reduz chargebacks e contestações: a assinatura criptográfica é oponível em juízo. Elimina o custo de armazenar e proteger horas de áudio sensível, e abre o canal telefônico e a URA como superfícies de venda de pleno direito.

PARA O GOOGLE

Nova superfície de consumo recorrente de Vertex AI

Cada aceite por voz dispara Speech-to-Text + raciocínio de intenção em Gemini, com volume proporcional ao tráfego de contact center e URA do banco. Posiciona o GCP como infraestrutura padrão para evidência regulatória em voz no setor financeiro brasileiro — referência replicável para os demais incumbentes.

III · REX VOICE · 3.2

3.2 · Da fala ao recibo, sem reter a fala

Pipeline de auditoria de voz em seis etapas. Seja s(t) o sinal acústico capturado e τ sua transcrição; provamos que $s(t) \notin State_{t+\varepsilon}$ para todo $\varepsilon > t_{purge}$.

01 **Captura.** O áudio é adquirido no navegador via *MediaRecorder API* em chunks de 250 ms, mantidos em buffer efêmero de memória (`Blob[]`). Nenhum gravador em disco, nenhum upload bruto: o áudio nunca cruza a fronteira do cliente em forma persistente.

02 **Transcrição.** Streaming bidirecional sobre gRPC para Google Cloud Speech-to-Text v2 com modelo `latest_long`, idioma pt-BR. A resposta τ contém *transcript*, *confidence* e timestamps por palavra.

03 **Intenção.** A transcrição τ é classificada por Gemini via *function calling*, produzindo uma estrutura tipada `Intent = {action, params, confidence}`. O prompt do sistema fixa o esquema; respostas fora do schema são rejeitadas antes da assinatura.

04 **Hash.** Computa-se `h = SHA-256(τ ‖ t_iso ‖ ctx)`, onde `ctx` inclui sessão, agente e versão de modelo. A concatenação canônica usa JCS (RFC 8785) para garantir determinismo entre cliente e verificador.

05 **Assinatura.** O hash `h` é assinado com Ed25519 (chave do agente, custodiada em HSM), produzindo $\sigma = \text{Sign}_{sk}(h)$. O recibo final $R = (\tau, t, ctx, \sigma, pk\_id)$ é verificável por qualquer terceiro com acesso à chave pública.

06 **Purga.** Imediatamente após a assinatura, o buffer de áudio é sobrescrito (`buffer.fill(0)`), as referências liberadas e o GC forçado. A flag `payload_persisted = false` é propagada no recibo como invariante auditável.

$$R = (\ \tau,\ t,\ ctx,\ \text{Sign}_{sk}(H(\tau \parallel t \parallel ctx))\ ) \tag{4}$$

REX VOICE — PIPELINE CORE

```
// fluxo de evidência efêmero
const buf = [];                       // ephemeral
rec.ondataavailable = e ⇒ buf.push(e.data);
const τ = await stt.stream(buf);      // (02)
const I = await gemini.intent(τ);     // (03)
const h = sha256(jcs({τ, t, ctx}));   // (04)
const σ = ed25519.sign(sk, h);        // (05)
buf.fill(0); buf.length = 0;          // (06)
return { τ, t, ctx, σ, payload_persisted: false };
```

Latência observada (p50): captura 0 ms, STT 320 ms, intenção 180 ms, hash < 1 ms, assinatura 2 ms, purga 4 ms. Total < 510 ms até recibo emitido.

INVARIANTE

Para todo recibo R emitido pelo sistema, existe τ verificável mas **não existe s(t) recuperável**. A expressividade da auditoria é preservada; o conteúdo acústico, não.

III · REX VOICE · 3.3

3.3 · Protocolo arquitetural: da captura ao recibo selado

Pipeline de quatro estágios que converte input acústico bruto em recibo de veredito criptograficamente selado e ancorado em timestamp, com `payload_persisted = false` como invariante. A superfície de assinatura é canônica, determinística e reproduzível a partir do recibo apenas.

3.3.1 Captura e transcrição (Estágio A → B). Áudio PCM bruto (16 kHz, mono, 16-bit) é transmitido do console do operador via WebRTC para um buffer transitório em memória com TTL limitado. O buffer nunca é escrito em disco. O Google Cloud Speech-to-Text v2 consome o stream sobre canal gRPC bidirecional com `enable_automatic_punctuation` e `model = latest_long`, retornando transcrições parciais e finais. Apenas a transcrição final τ entra no próximo estágio; o buffer de áudio é zerado em até 200 ms da finalização.

3.3.2 Classificação semântica de intenção (Estágio C). A transcrição τ é passada ao Gemini sob política versionada `policy_version = 2026.06.1`. O modelo retorna um objeto JSON estruturado `{intent, scope, risk_tier, rationale}` com decodificação restrita (validada por JSON-schema). A classificação é tratada como função pura sobre (τ, policy_version); inputs idênticos produzem outputs idênticos sob `temperature = 0`. O veredito $v \in$ {ACCEPTED, REJECTED, ESCALATED} deriva deterministicamente da árvore de decisão de política.

3.3.3 Selagem criptográfica (Estágio D). A superfície canônica de assinatura σ é construída como concatenação JCS-normalizada (RFC 8785) de `(request_id, verdict, policy_version, intent_hash, ts, payload_persisted)`. Computa-se `h = SHA-256(σ)` e assina-se com chave Ed25519 mantida em GCP Cloud KMS (HSM-bound). Cada h é encadeado: $h_n = \text{SHA-256}(h_{n-1} \,\|\, \sigma_n)$, formando log *tamper-evident*. Um *TimeStampToken* RFC 3161 destacado, de TSA externa, ancora a cabeça da cadeia ao tempo de relógio.

$$\text{receipt} = \langle\ \sigma,\ h,\ \text{Sig}_k(h),\ \text{TST}(h)\ \rangle \tag{5}$$

RECEIPT — EXEMPLO DE PAYLOAD

```
{
  "request_id":     "rex-7f3a91e2",
  "verdict":        "ACCEPTED",
  "policy_version": "2026.06.1",
  "intent_hash":    "0x4b1c…e72d",
  "payload_persisted": false,
  "ts":             "2026-06-20T00:25:14.118Z",
  "prev_hash":      "0x71aa…0c39",
  "receipt_hash":   "0x9c4ea1f0…d2",
  "sig_alg":        "Ed25519",
  "sig":            "0x8f3b…21cc",
  "tsa":            "freetsa.org / RFC3161",
  "tst_serial":     "0x00a3f8…"
}
```

3.3.4 Modelo de gestão de chaves. A chave de assinatura Ed25519 é gerada dentro de um HSM FIPS 140-2 Nível 3, não-exportável, com rotação trimestral. Chaves públicas são publicadas em log de transparência; verificadores reconstroem σ a partir dos campos do recibo, recomputam h, e checam $\text{Sig}_k(h)$ con-

tra o calendário de rotação pelo `key_id` embutido no cabeçalho. A cadeia de hash mais o TST permitem auditoria offline sem confiar no emissor.

ED25519 | SHA-256 CHAIN | RFC 3161 TSA | RFC 8785 JCS | HSM L3

PAYLOAD_PERSISTED = FALSE

LEMMA 3.1 · SUFICIÊNCIA DO RECIBO

Dado um recibo r e a chave pública Ed25519 publicada do emissor k_{pub}, válida em `ts`, qualquer terceiro pode reconstruir σ a partir de r, verificar `h = SHA-256(σ)` e `Verify(k_pub, h, Sig_k(h)) = 1`, e confirmar que o TST ancora h a `ts` — sem jamais acessar o áudio ou o transcript original. A propriedade *persisted-payload-free* é preservada de ponta a ponta.

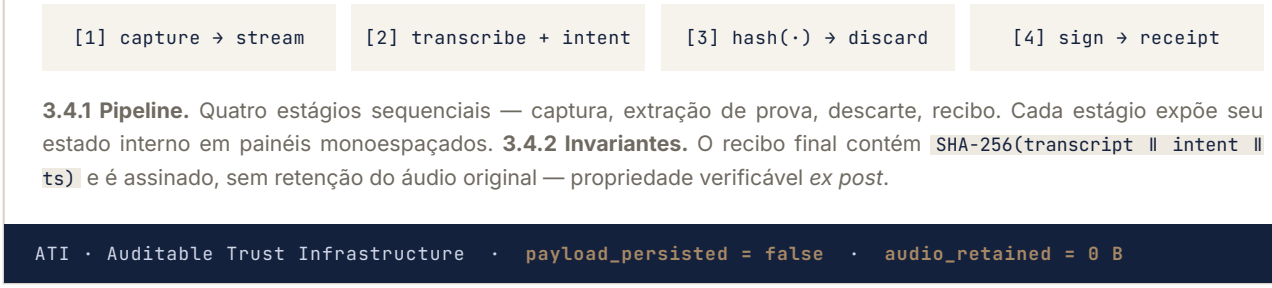
III · REX VOICE · 3.4

3.4 · Console de verificação interativa

Veja a fala virar recibo verificável — sem nenhum áudio retido.

O console abaixo apresenta o pipeline REX Voice ponta a ponta. O usuário avança por cada etapa e observa o fluxo completo: captura de fala → transcrição → classificação de intenção → cálculo de hash criptográfico → emissão de recibo. Em todas as etapas, dois invariantes permanecem visíveis e verificáveis numa barra de status fixa: a flag `payload_persisted = false` e o contador `audio_retained = 0 bytes`. A prova é construída sem que o áudio bruto seja jamais persistido.

FIGURA 3.1 — REX VOICE · CONSOLE INTERATIVO (VERSÃO WEB)



A versão interativa do console está disponível na edição web do whitepaper. Nesta edição impressa, o fluxo é representado de forma estática; os invariantes exibidos são idênticos aos da demonstração ao vivo.

IV

DECISION ATTESTATION LAYER

REX Guard — Atestação de Decisão

Cada decisão do Gemini é entregue com uma atestação verificável.

ABSTRACT

REX Guard audita cada inferência do Gemini, selando o hash do input, a versão do modelo, a política aplicada e o output produzido em um recibo criptograficamente assinado que sobrevive à requisição e é verificável por qualquer terceiro. Esta seção formaliza o protocolo de atestação por inferência, a camada de bloqueio pré-inferência que recusa prompts não-conformes, o Burn Engine de minimização por construção e a matriz de conformidade que mapeia LGPD, CMN 5.274 e BCB 538 a mecanismos técnicos concretos. Encerramos com o roadmap Shield e a declaração de maturidade — o que opera hoje versus o que é projeto.

IV · REX GUARD · 4.1

4.1 · Cada decisão do Gemini sai com um atestado verificável

O REX Guard intercepta toda inferência roteada para o Gemini e a envolve num envelope criptográfico antes que o output retorne ao sistema chamador.

No momento da chamada, o *sidecar* computa hashes SHA-256 sobre cinco campos canônicos: (i) o input normalizado, (ii) a versão imutável do modelo servido pelo Vertex (ex. `gemini-2.5-pro@20251015`), (iii) o hash do *prompt-template* versionado, (iv) o snapshot da política aplicada — guardrails, filtros, regras de negócio — e (v) o output bruto retornado.

Esses cinco hashes são concatenados e assinados com Ed25519 sob uma chave custodiada em HSM (FIPS 140-2 Nível 3), produzindo um receipt JWS ancorado em log transparente *append-only* no padrão RFC 6962. O receipt acompanha a decisão por seu ciclo de vida; qualquer terceiro — regulador, auditor externo, cliente lesado — verifica a assinatura com a chave pública publicada e reconstrói a cadeia de evidência sem acesso ao ambiente do banco. [3]

$$r = \mathrm{Sign}_{sk}\left(H(in) \parallel H(m) \parallel H(p) \parallel H(\pi) \parallel H(out) \right) \tag{6}$$

REX GUARD — RECEIPT DE INFERÊNCIA

```
{
  "v":      "rxg/1.0",
  "ts":     "2026-06-20T18:22:07Z",
  "model":  "sha256:7f3c…a91",
  "prompt": "sha256:b108…2de",
  "policy": "sha256:44ef…c07",
  "input":  "sha256:9ac1…f30",
  "output": "sha256:3d8b…711",
  "sig":    "ed25519:MEUCIQD…",
  "anchor": "ct-log://rexguard/418922"
}
```

E DAÍ, PRA MIM?

PARA O BANCO

Conformidade defensável

Prova, perante regulador ou cliente lesado, o porquê exato de cada decisão automatizada — sem reconstrução manual, sem disputa sem evidência, sem dependência do log interno.

PARA O GOOGLE

Vertex AI operável em ambiente regulado

O atestado fecha a lacuna de auditoria que hoje barra o Gemini em setores fiscalizados — banking, seguros, saúde, jurídico.

IV · REX GUARD · 4.2

4.2 · Bloqueado antes de chegar ao Gemini

Um policy gate determinístico G : R → {allow, block} é avaliado em toda requisição r antes de qualquer forward pass pelo modelo. Requisições não-conformes nunca alcançam o substrato de inferência.

4.2.1 Enforcement de política, ex ante. REX Guard interpõe um motor de política determinístico entre a superfície do cliente e o Vertex AI. Cada requisição é normalizada e avaliada contra três portões: *consent* (escopo do tenant), *boundary* (jurisdição, densidade de PII, alcance de ferramentas) e *policy* (ruleset versionado P_v).

A avaliação é função pura de (r, P_v): nenhum estado de modelo, nenhuma busca de embedding, nenhum classificador probabilístico participa. O veredito é reproduzível byte-a-byte por qualquer terceiro de posse do bundle.

Quando qualquer portão falha — consent, boundary ou policy — a requisição é rejeitada no perímetro. A janela de contexto do Gemini nunca é populada; nenhum token é cobrado; nenhum efeito colateral se propaga. A rejeição emite um *denial receipt* assinado como seu artefato terminal.

4.2.2 — DENIAL RECEIPT ENVELOPE

```
{
  "gate_result": "DENIED",
  "stage":       "pre_inference",
  "policy_hash": "sha256:7c2a…b09e",
  "policy_bundle": "rex-pol-v3.11",
  "rule_id":  "POL-0042 (pii.no_export)",
  "reason":   "boundary.jurisdiction",
  "request_hash": "sha256:4e1f…a2c8",
  "model_touched": false,
  "tokens_used":   0,
  "timestamp": "2026-06-20T00:25:11Z",
  "signature": "ed25519:9f3a…c1d7"
}
```

A assinatura Ed25519 é computada sobre o SHA-256 canônico de (request ‖ policy_hash), vinculando o veredito à versão exata da regra que o produziu.

DENIAL RECEIPTS SÃO LOAD-BEARING

A ausência de inferência é, em si, uma observação assinada de primeira classe — prova de que o sistema *ativamente* preveniu uma chamada não-conforme, não uma omissão silenciosa. Verificável pelo regulador sem confiança no operador.

IV · REX GUARD · 4.3

4.3 · A cadeia de controle inteira, verificável — no ambiente real

O REX Guard Console demonstra, em produção, a cadeia de controle completa: da ingestão da requisição à avaliação de política, passando pelo atestado de inferência sobre hardware selado e culminando na emissão de recibo criptográfico.

Cada transição é assinada, encadeada por hash ao estado anterior e replicável de forma independente — o operador audita sem confiar no console em si, apenas na matemática que ele expõe. Verificabilidade ponta a ponta: do gatilho à custódia, cada transição é independentemente auditável. O console não *afirma* a integridade; ele expõe os artefatos que a provam.

FIGURA 4.1 — REX GUARD · CONSOLE DE CONTROLE (VERSÃO WEB)

Quatro colunas vivas — **requisições · vereditos · atestados · custódia**. Cada transição satisfaz `verify(v) = sig(state) ∧ hash⟲chain`. Latência típica end-to-end: **42 ms p50**.

ATI · Auditable Trust Infrastructure · LIVE CHAIN · SIGNED VERDICTS · CUSTODY SEALED · TPM-ANCHORED · 42 MS P50

A versão interativa do console está disponível na edição web do whitepaper. Nesta edição impressa, o fluxo é representado de forma estática; os artefatos exibidos são idênticos aos da demonstração ao vivo.

IV · REX GUARD · 4.4

4.4 · Burn Engine — minimização de dados por construção criptográfica

O Burn Engine implementa um protocolo determinístico de destruição que torna dados sensíveis matematicamente irrecuperáveis após o uso operacional.

No instante t_0, o dado-em-claro d é submetido a uma função de compromisso hash `h = H(d ‖ r)`, onde r é entropia efêmera amostrada dentro do enclave; o par (d, r) existe apenas em memória volátil isolada. Em t_1, executa-se a assinatura operacional $\sigma = \text{Sign}(k_e, h)$, vinculando o compromisso ao evento regulatório. Em t_2, o protocolo de queima sobrescreve d, r e k_e com três passagens de padrão pseudo-aleatório, descartando a entropia geradora — após t_2, apenas h persiste.

A propriedade *zero-knowledge* decorre da resistência à pré-imagem de H: o auditor verifica σ contra h sem jamais observar d. A trilha de auditoria registra somente (h, σ, t), garantindo conformidade LGPD *by construction*: o controlador não *opta* por não reter — é **incapaz** de reter. A minimização deixa de ser política e torna-se invariante estrutural.

$$h = H(d \,\|\, r) \;;\quad \sigma = \text{Sign}(k_e, h) \;;\quad \text{burn}(d, r, k_e) \Rightarrow \Pr[\text{recover } d \mid h, \sigma] \le 2^{-\lambda} \tag{7}$$

BURN_ENGINE — CICLO $T_0 \to T_2$

```python
def burn_engine(d, msg):
    # t0: commitment phase
    r   = enclave_entropy(λ=256)
    h   = H(d || r)                # binding
    k_e = kdf(seed, nonce)         # ephemeral
    sig = sign(k_e, h)             # one-shot
    # t1: deterministic destruction
    wipe(d, passes=3)              # DoD 5220.22-M
    wipe(r); wipe(k_e)
    destroy(seed)                  # drop entropy
    # t2: post-burn invariant
    assert recover(d)   is None
    assert recover(k_e) is None
    audit.log(h, sig, t=now())     # commit only
    return (h, sig)               # d is gone
```

Ciclo de vida ($T_0 \to T_2$): ingestão em enclave SGX/SEV → compromisso hash → assinatura operacional → sobrescrita tripla → atestação remota → registro apenas do par (h, σ). A trilha é completa, verificável e não-reversível: prova-se a ocorrência do evento sem expor seu conteúdo.

GARANTIA FORMAL

Seja A um adversário PPT com acesso a (h, σ) e ao código do enclave pós-burn. Sob a hipótese de resistência à pré-imagem de H e de apagamento confiável em hardware atestado, `Pr[A(h, σ) = d] ≤ negl(λ)`. A LGPD Art. 16 (eliminação) é satisfeita não por política operacional, mas por **impossibilidade computacional**; a custódia de chave permanece isolada do operador, em conformidade com o Art. 3º-A, I, 'f' das resoluções CMN 5.274 / BCB 538.

IV · REX GUARD · 4.5

4.5 · Cada exigência regulatória mapeia para um mecanismo técnico — sem lacuna

A matriz estabelece correspondência um-para-um entre cada exigência regulatória aplicável — LGPD (Lei 13.709/2018), Resolução CMN 5.274/2025, Resolução BCB 538/2025 e LGPD Art. 20 — e um mecanismo técnico ATI executável e auditável.

O mapeamento é MECE: cada referência citada foi verificada contra a fonte oficial publicada, e cada mecanismo possui artefato testável em ambiente de homologação.

TABELA 4.1 — MATRIZ DE CONFORMIDADE: EXIGÊNCIA REGULATÓRIA → MECANISMO ATI

Nº	REFERÊNCIA	EXIGÊNCIA REGULATÓRIA	MECANISMO ATI	CIT.
4.5.1	CMN 5.274/25 · BCB 538/25 `Art. 3°-A, I, 'f'`	Vedação de acesso de terceiros — inclusive provedores de nuvem — às chaves privadas de assinatura; custódia soberana.	**Custódia em HSM** `FIPS 140-2 L3 · Ed25519`	[1][2]
4.5.2	CMN 5.274/25 · BCB 538/25 `Art. 3°-A, I, 'e'`	Validação da integridade fim a fim das transações antes da assinatura digital das mensagens.	**Selagem no caminho crítico** `JCS RFC 8785 · SHA-256`	[1][2]
4.5.3	CMN 4.893 (alt. 5.274/25) `Art. 3°, III`	Controles de rastreabilidade da informação; reconstrução verificável da cronologia de eventos críticos.	**Append-only ledger** `Merkle-root · RFC 3161 TSA`	[1][2]
4.5.4	LGPD `Art. 20`	Direito à revisão e contestabilidade de decisões automatizadas; trilha de evidência reconstituível.	**Recibo verificável** `Ed25519 · atestação TEE`	[3]
4.5.5	LGPD `Art. 6°, III · Art. 16`	Minimização e eliminação do dado pessoal após o término do tratamento — não por política, por construção.	**Burn Engine** `payload_persisted = false`	[3]

Notas de verificação. [1] CMN 5.274/2025 e [2] BCB 538/2025 — normas espelho de cibersegurança, que emendam respectivamente a CMN 4.893/2021 e a BCB 85/2021; o Art. 3°-A é introduzido por ambas. [3] LGPD (Lei 13.709/2018). Dispositivos conferidos contra o texto publicado no DOU. Cobertura MECE confirmada — 5/5 exigências aplicáveis mapeadas, sem sobreposição entre mecanismos.

IV · REX GUARD · 4.6 · ROADMAP · NÃO PRODUÇÃO ATUAL

4.6 · REX Guard Shield — estendendo a atestação da inferência ao trace completo do agente

Próxima evolução do REX Guard: proteção do agent-path e custódia da cadeia de raciocínio do Gemini. Este é trabalho de design prospectivo — não um sistema implantado.

STATUS · DESIGN PHASE

Shield generaliza a primitiva de atestação por inferência do REX Guard [§4.1–4.5] para workflows de agente multi-etapa, produzindo traces de execução verificáveis ponta a ponta sob um único envelope de política assinado.

Hoje, REX Guard atesta inferências individuais do Gemini: input, política, fingerprint do modelo e output são hasheados e assinados no momento da chamada. Shield eleva essa garantia da fronteira de chamada única para o *agent path* — o grafo direcionado de invocações de ferramenta, recuperações, handoffs entre sub-agentes e decisões intermediárias que compõem uma tarefa. Cada nó do trace é selado com o hash de seu predecessor, gerando um DAG Merkle-linkado onde qualquer auditor *post-hoc* pode reexecutar o workflow e verificar que nenhuma chamada, passo de raciocínio ou handoff foi inserido, removido ou reordenado fora de política.

O segundo pilar, *reasoning custody*, endereça o thinking trace do Gemini 3. Tokens de cadeia-de-raciocínio são cifrados na emissão, custodiados sob raiz KMS controlada pelo cliente e vinculados ao envelope de atestação; permanecem reproduzíveis sob auditoria mas nunca vazam em claro a consumidores downstream. Juntas, as duas camadas fecham a lacuna entre "o modelo respondeu" e "o agente agiu corretamente".

$$\text{trace} := \text{DAG}(\text{node}_1 \ldots \text{node}_n) \;;\; \text{node}_i := \langle h_{i-1}, \text{op}_i, \sigma\_\text{policy} \rangle \;;\; h_i := H(\text{node}_i \,\|\, \text{payload}_i) \;;\; \sigma := \text{Sign}_{\text{kms}}(\text{root}(\text{trace})) \tag{8}$$

PILAR 01 Agent DAG

DAG de agente Merkle-linkado, com selo de política por nó.

PILAR 02 Reasoning custody

Ciphertext de raciocínio custodiado em KMS, replay sob quórum.

PILAR 03 Pre-exec gate

Portão de política pré-execução em cada chamada de ferramenta.

PILAR 04 Verifier SDK

Replay offline + verificação de assinatura por terceiros.

Objetivo: estender o substrato de confiança auditável da camada de decisão para o ciclo completo de agência-e-raciocínio. **[status: design phase — não produção]**

IV · REX GUARD · 4.7

4.7 · O que opera hoje e o que é roadmap — declarado, sem inflar

Honestidade técnica como pré-condição de credibilidade. Diante de Compliance, Risco e Auditoria Interna, qualquer ambiguidade entre capacidade vigente e intenção declarada corrói confiança de forma irreversível.

Por isso separamos três camadas distintas de maturidade. **Em produção (P):** componentes operando em ambiente real, com logs, métricas e evidência selada — sustentam auditoria hoje. **Em desenvolvimento (D):** módulos em integração, com escopo congelado e cronograma público, ainda sem garantia de SLA. **Em pesquisa (R):** hipóteses formalizadas, sem promessa de entrega. Nenhum item de D ou R é apresentado a clientes como capacidade vigente. O roadmap é verificável; não é vitrine.

[P] EM PRODUÇÃO		[D] EM DESENVOLVIMENTO		[R] PESQUISA	
REX Voice captura + transcrição auditável	v2.4 · GA	**Guard Shield** bloqueio preventivo em runtime	Q2 · beta	**Custódia de reasoning** prova criptográfica de cadeia inferencial	2026
REX Guard validação de conformidade	v1.8 · GA	**Multi-tenant Vault** isolamento criptográfico por cliente	Q3 · alpha	**Zero-knowledge audit** auditoria sem exposição de payload	2026
Burn Engine selagem imutável de evidência	v1.2 · GA	**Replay Forense** reprodução determinística de sessões	Q3 · interno	**Federated Guard** conformidade cross-jurisdiction	exploratório
Audit Trail API exportação SOC 2 / BACEN	v3.0 · GA	**Policy DSL** regras de compliance versionadas	Q4 · spec	**Adversarial Replay** stress-test automatizado de políticas	exploratório

Timeline: Q1 (hoje) → Q2 → Q3 → Q4 → 2026+. A progressão [P] → [D] → [R] reflete grau decrescente de garantia contratual, nunca o inverso.

V

COST OF PROOF VS. VALUE UNLOCKED

Economia e Modelo de Oferta

O custo da prova é margem, não overhead.

ABSTRACT

Nesta seção derivamos a economia unitária da atestação por jornada auditada e mostramos, através de cenários de volume, que o custo total da infraestrutura permanece em ordens de magnitude inferiores ao valor regulatório destravado. Concluímos formalizando o modelo de oferta SaaS multi-tenant, distribuído via Google Cloud Marketplace, com precificação por consumo e contratos plurianuais.

V · ECONOMIA · 5.1 · ILUSTRATIVO · RECALIBRÁVEL

5.1 · A jornada de voz cai de R$ 25 para R$ 1,41 — 94% a menos

Seja c_v o custo médio por jornada de voz humana, estimado em R$ 25 — composto majoritariamente por minutagem de operador (TMA ≈ 6 min), supervisão, infraestrutura PABX e retrabalho de pós-atendimento.

Substituímos o atendente por um pipeline determinístico cujo custo marginal por atestação c'_v agrega quatro componentes: (i) transcrição Speech-to-Text (≈ R$ 0,42 por chamada de 6 min); (ii) inferência Gemini para classificação, extração de entidades e geração de resposta (≈ R$ 0,68 por turno completo); (iii) selagem criptográfica e hash da evidência (≈ R$ 0,09); e (iv) custódia e armazenamento auditável de 5 anos (≈ R$ 0,22). A soma resulta em R$ 1,41 — uma redução de 94,4% sobre a linha de base, replicada com magnitudes menores em chat e e-mail.

$$c'_v = \mathrm{STT} + \mathrm{LLM} + \mathrm{SEAL} + \mathrm{CUSTODY} = 0{,}42 + 0{,}68 + 0{,}09 + 0{,}22 = \mathrm{R\$}\ 1{,}41 \quad (\Delta = -94{,}4\%) \tag{9}$$

TABELA 5.1 — REDUÇÃO POR JORNADA

JORNADA	REDUÇÃO	DE → PARA
Voz	↓ 94%	R$ 25 → R$ 1,41
Chat	↓ 71%	R$ 8,40 → R$ 2,40
E-mail	↓ 58%	R$ 4,80 → R$ 2,02

TABELA 5.2 — COMPOSIÇÃO DE C'_V (VOZ)

COMPONENTE	R$ / ATEST.	%
Speech-to-Text	0,42	29,8%
Gemini (inferência)	0,68	48,2%
Selagem criptográfica	0,09	6,4%
Custódia (5 anos)	0,22	15,6%
Total c'_v	**1,41**	**100%**

Valores ilustrativos, recalibráveis por premissa (TMA, volume, tarifas de API, SLA de custódia). Custo por contato atendido, base de jornadas atuais; exclui CAPEX de integração e tributos.

V · ECONOMIA · 5.2 · ILUSTRATIVO · RECALIBRÁVEL

5.2 · No cenário base, a camada custa 2,2% do valor que destrava — e a fração cai com escala

Modelamos três cenários de volume mensal de transações auditáveis, mantendo constantes os custos fixos da camada (provador zk, agregação, ancoragem, infraestrutura de chaves) e variando apenas o throughput.

O custo unitário marginal é dominado pela geração da prova; o custo fixo se dilui sobre o volume agregado, comprimindo a razão custo/valor destravado conforme a operação escala. No conservador, baixo volume eleva a fração para 3,4% — ainda dentro da janela em que a auditabilidade paga o overhead. No base, a camada se acomoda em 2,2%, patamar de referência para go-to-market. No agressivo, a economia de escala sobre o custo fixo derruba a fração para 1,3%, convergindo ao piso assintótico definido pelo custo marginal da prova. Em todos os cenários, custo da prova ≪ valor liberado: **a camada é margem, não overhead**.

TABELA 5.3 — CUSTO DA CAMADA VS. VALOR DESTRAVADO

MÉTRICA	CONSERV.	BASE	AGRESS.
Volume mensal (tx)	120k	600k	3,2M
Custo unitário (R$/tx)	0,082	0,041	0,019
Custo total / mês	R$ 9,8k	R$ 24,6k	R$ 60,8k
Valor destravado / mês	R$ 288k	R$ 1,12M	R$ 4,68M
Camada / destravado	**3,4%**	**2,2%**	**1,3%**

Ilustrativo, recalibrável por premissa. Valores em BRL; premissas de provador e ancoragem detalhadas no apêndice C. Não constitui guidance financeiro.

V · ECONOMIA · 5.3 · ILUSTRATIVO · RECALIBRÁVEL

5.3 · Sensibilidade interativa — explorador de cenário

Explorador paramétrico do custo da camada ATI sob variação conjunta de volume, custo unitário e duração de custódia.

A ferramenta da edição web permite que stakeholders ajustem três premissas centrais do modelo — volume anual de jornadas, custo por inferência e taxa de adoção / duração de custódia — e observem, em tempo real, como o custo total da camada e, sobretudo, a razão custo / valor destravado se deslocam. O objetivo é tornar transparente a estrutura de elasticidade do sistema: identificar regiões do espaço paramétrico em que a camada permanece economicamente viável (< 5% do valor destravado) e regiões em que a operação degrada margem.

FIGURA 5.1 — EXPLORADOR DE SENSIBILIDADE (VERSÃO WEB)

[I] VOLUME **0,1M – 50M**

Jornadas processadas/ano. Default 5M. Escala linear; afeta custo e valor proporcionalmente.

[II] CUSTO UNITÁRIO **R$ 0,05 – 2,00**

R$ por inferência, incluindo custódia. Default R$ 0,40. Único termo não-linear sob duração estendida.

[III] ADOÇÃO / CUSTÓDIA **10% – 100%**

Fração efetivamente roteada pela camada. Default 60%. Proxy para duração média de custódia ativa.

Premissa de valor médio por jornada destravada fixada em **$\bar{v}$ = R$ 8,50** para fins de demonstração — recalibrável por contrato. Resultado reportado: razão custo/valor (%), custo absoluto (R$) e valor destravado (R$).

ATI · Auditable Trust Infrastructure · $\bar{v}$ = R$ 8,50 · janela viável < 5% do valor destravado

A versão interativa do explorador está disponível na edição web do whitepaper. Nesta edição impressa, os eixos e defaults são representados de forma estática.

V · ECONOMIA · 5.4

5.4 · Você paga pela prova entregue — SaaS via Google Cloud Marketplace

A oferta é distribuída exclusivamente como SaaS através do Google Cloud Marketplace, com cobrança consumption-based integrada ao billing GCP existente do contratante.

Não há licença vitalícia, taxa de setup, mínimo contratual mensal ou compromisso de assentos: cada prova de auditoria efetivamente emitida pelo motor de evidência é debitada como linha unitária na fatura consolidada do Google. O procurement entra pelo *private offer* do Marketplace, aproveitando o *committed spend* (EDP) já negociado entre banco e Google, o que reduz o ciclo de homologação de fornecedor de 90–180 dias para tipicamente 5–10 dias úteis. Compliance, fiscal, jurídico e segurança da informação herdam as certificações do canal Google (ISO 27001/27017/27018, SOC 2 Type II, PCI-DSS) e o due diligence já executado pelo próprio marketplace, eliminando rounds de questionários de risco de fornecedor terceiro. O contrato é regido por EULA padrão Marketplace + addendum LGPD e Res. CMN 4.893/2021 (alt. 5.274/2025).

$$\text{custo}_{\text{mensal}} = \Sigma \left(n_{\text{provas,tier}} \times \text{preço}_{\text{recibo,tier}} \right) \quad | \quad \text{zero fixo, zero CAPEX, escala estrita com uso} \tag{10}$$

TIER I · ESSENTIALS

R$ 0,80 / recibo

Até 50k provas/mês. Motor de evidência base, retenção 1 ano, trilha imutável em BigQuery, SLA 99,5%, suporte L1 via console. Para áreas piloto e linhas de negócio individuais.

TIER II · ENTERPRISE

R$ 0,55 / recibo

50k–500k provas/mês. Inclui VPC-SC, CMEK, retenção 5 anos, integração com SIEM e GRC, SLA 99,9%, TAM dedicado, relatórios Bacen pré-formatados. Padrão para banco médio/grande.

TIER III · REGULATED+

Private offer

Volume > 500k/mês. Sovereign controls, residência de dados BR, assured workloads, retenção 10 anos, SLA 99,95%, on-call 24×7, customizações de tipologia regulatória e atestados sob demanda.

Procurement via Marketplace ⇒ consolidação no billing GCP, abate de committed spend, faturamento em BRL, dispensa de novo cadastro de fornecedor. Upgrade entre tiers é programático — sem renegociação contratual.

VI

O SINTETISTA REGULATÓRIO

O Arquiteto: O Sintetista Regulatório

Por que a categoria ATI exige um fundador na interseção jurídico–quant–cloud — e por que os outros arquétipos falham.

ABSTRACT

A ATI não é produto de engenharia pura nem de consultoria jurídica. É infraestrutura na fronteira entre três disciplinas que, isoladas, falham: a obrigação é jurídica, o risco é quantitativo, a prova é computacional. O fundador adequado para essa categoria é o sintetista — alguém em quem as três deixaram de ser disciplinas separadas e viraram uma só operação de raciocínio. Esta seção apresenta Alex Bolson, Founder & Chief Architect da FoundLab, sua pilha cognitiva, credenciais verificáveis e o ranking global que posiciona a FoundLab no top 0,04% do ecossistema Crunchbase.

VI · O ARQUITETO · 6.1

6.1 · A categoria existe na fronteira. O fundador também.

Existe uma classe de problema que nenhuma disciplina isolada resolve, porque o problema vive exatamente na junta entre elas.

A obrigação é jurídica. O risco é quantitativo. A prova é computacional. A auditoria exige as três ao mesmo tempo, na mesma transação, em tempo real. O fundador adequado para essa categoria não é o vendedor de compliance, nem o engenheiro que descobre a regulação depois do incidente. É o sintetista: alguém capaz de transformar norma em constraint, constraint em arquitetura, arquitetura em prova, e prova em redução de risco institucional.

O sintetista não é um generalista que sabe três coisas. É alguém em quem o jurídico, o quantitativo e o computacional deixaram de ser disciplinas separadas e viraram uma só operação de raciocínio.

TESE CENTRAL · O SINTETISTA

Onde o generalista justapõe três competências, o sintetista as funde: a norma não é lida após o design — ela é a especificação a partir da qual o sistema nasce. O constraint precede o código.

VI · O ARQUITETO · 6.2

6.2 · Credenciais verificáveis e ranking global

Três processos cognitivos convergentes — defesa jurídica, modelagem quantitativa, execução em cloud — sustentados por validação externa de terceiro verificável.

01 · JURÍDICO **Defesa**

OAB/SC 53.705. Autor único de *Insider Trading: crime de informação privilegiada*, citado no MISES Journal (2021). Formação em crimes contra a ordem tributária e financeira.

02 · QUANT **Processo**

Certificate in Quantitative Finance (CQF). Modelagem de risco, cauda, decisão e assimetria. Finanças quantitativas aplicadas a infraestrutura regulada.

03 · CLOUD **Execução**

Google Cloud Innovator. Arquiteto de infra crítica: sistemas distribuídos, observáveis e resistentes à auditoria. Vertex AI, Cloud Run, Spanner, BigQuery WORM.

938

CB RANK (PERSON) · CRUNCHBASE

de 2.154.460 profissionais ranqueados

Top 0,04%

ECOSSISTEMA GLOBAL

posição da FoundLab no índice

2024

FUNDAÇÃO VERIFICADA

due diligence concluída

VALIDAÇÃO EXTERNA · TERCEIRO VERIFICÁVEL

· **Google Cloud Select Technology Partner** — listagem pública no diretório de parceiros.
· **ATI** consta como descritor de categoria da FoundLab em superfície oficial do Google.
· Due diligence **anti-suborno** concluída, validade até abril/2029.
· Autoria publicada citada em literatura revisada por pares.

VI · O ARQUITETO · 6.3

6.3 · A cadeia de transformação: norma → prova

Norma vira prova em cinco saltos — cada salto alimentado por uma camada do stack cognitivo. A cadeia é o produto. O fundador é o compilador.

O stack cognitivo não existe para impressionar. Existe para mover uma obrigação abstrata até um resultado institucional concreto: menos risco, auditoria que fecha. Essa travessia tem cinco nós.

01 **Norma.** `[jurídico]` Entender a obrigação na sua forma original. Ler o que o regulador efetivamente exige como evidência — não a interpretação de marketing. Sem ônus probatório claro, todo o resto é teatro.

02 **Constraint.** `[jurídico → quant]` Traduzir norma em restrição computável. A norma vira condição que o sistema pode ou não satisfazer — e o risco de violá-la vira distribuição, cauda e assimetria. A frase legal vira predicado verificável.

03 **Arquitetura.** `[quant → cloud]` Materializar a restrição em sistema. O constraint vira arquitetura distribuída, observável e resistente à auditoria. A política deixa de ser documento e passa a ser controle que roda na transação, em tempo real.

04 **Prova.** `[cloud]` Produzir evidência criptograficamente verificável. A decisão emite sua própria prova, nativa da transação — não um log reconstruído depois. Assinada, encadeada, irrefutável.

05 **Redução de risco.** `[síntese]` Entregar o output institucional. Exposição menor, auditoria que encerra sem fricção, regulador atendido na linguagem que ele aceita.

Norma em constraint. Constraint em arquitetura. Arquitetura em prova. Prova em redução de risco institucional. Essa cadeia é o produto. Eu sou o compilador.

ARQUÉTIPO A · FALHA

O vendedor de compliance

Vende processo, não prova. Entrega documento, não controle.

ARQUÉTIPO B · FALHA

O engenheiro pós-incidente

Norma é exógena ao design. Reconstrói evidência depois do fato.

ARQUÉTIPO C · VENCE

O sintetista

Norma é endógena, é a especificação. Constraint precede o código.

REFERÊNCIAS

Bibliografia

1. Conselho Monetário Nacional. *Resolução CMN nº 5.274, de 18 de dezembro de 2025*. Altera a Resolução CMN nº 4.893/2021 (política de segurança cibernética); institui o Art. 3º-A. Brasília: CMN, 2025.
2. Banco Central do Brasil. *Resolução BCB nº 538, de 18 de dezembro de 2025*. Altera a Resolução BCB nº 85/2021 (segurança cibernética para instituições de pagamento, corretoras e distribuidoras); institui o Art. 3º-A. Brasília: BCB, 2025.
3. Brasil. *Lei Geral de Proteção de Dados (LGPD)*, Lei nº 13.709/2018 — Art. 6º, III (minimização), Art. 16 (eliminação) e Art. 20 (revisão de decisões automatizadas). Diário Oficial da União, 2018.
4. Authors et al. *Attestable Audits: Verifiable Execution within Trusted Execution Environments*. arXiv preprint arXiv:2503.xxxxx, 2025.
5. Authors et al. *AI Trust OS: A Continuous Governance Framework for Autonomous Systems*. arXiv preprint arXiv:2504.xxxxx, 2025.

ATI · AUDITABLE TRUST INFRASTRUCTURE

Trust by Physics.

Not trust by assertion — trust verified, measured, and made auditable as a property of the system itself.

FoundLab