

AIMM · Modelo de Maturidade em IA — 2026

Documento de referência · Versão 1.0

Escopo

Este documento define o AIMM, um modelo para avaliar a maturidade de uma organização no uso de inteligência artificial. A avaliação cobre oito dimensões, cada uma classificada em cinco níveis segundo critérios observáveis por um terceiro. O modelo distingue o nível *atual* do nível-*alvo* de cada dimensão e estabelece um teto que limita a maturidade global à solidez de sua base.

Destina-se a organizações que precisam determinar, de forma auditável, onde estão e qual é o próximo passo defensável. Não é objeto do modelo prescrever tecnologias, fornecedores ou arquiteturas: o AIMM mede estado e expõe risco; a decisão do que construir é subsequente à avaliação.

Termos

- **Dimensão** — um eixo de maturidade avaliado de forma independente dos demais.
 - **Nível atual** — o estágio observável em que a organização se encontra em uma dimensão, no momento da avaliação.
 - **Nível-alvo** — o estágio adequado àquela dimensão, definido pela estratégia e pelo apetite a risco da organização. Não é necessariamente o nível 5.
 - **Piso** — as três dimensões (Dados, Testes e validação, Governança) que limitam a maturidade global. Ver §3.4.
 - **Exposição** — operação de IA em produção acima do nível sustentado pelo piso. Distingue-se de maturidade.
-

1. Contexto

A adoção de inteligência artificial deixou de encontrar seu principal limite na tecnologia. A capacidade de construir sistemas está disponível e é acessível. O limite predominante passou a ser organizacional: a maioria das organizações não dispõe de uma forma verificável de declarar onde está, o que a separa do estágio seguinte, e em que ponto a velocidade de adoção passou a constituir exposição a risco. Sem essa medida, as decisões de expandir ou conter o investimento carecem de base observável.

O AIMM existe para produzir essa medida — e para produzi-la de modo que resista a auditoria externa, não à leitura complacente de quem a encomendou.

2. Limitações dos modelos atuais

Os modelos de maturidade em uso corrente foram formulados entre 2019 e 2023, período em que "inteligência artificial" designava, na prática, os sistemas analíticos que as organizações já operavam havia anos — crédito, fraude, recomendação. Medem esse domínio com competência. A premissa que os sustenta, porém, deixou de valer: uma organização pode ser avançada nesses sistemas e iniciante em IA generativa, que constitui um domínio distinto, com riscos distintos. Um índice único de maturidade oculta essa distinção e atribui a condição de "maduro" a quem domina o antigo e apenas iniciou o novo.

A esse problema de premissa somam-se duas falhas de método:

Medição por adjetivo. Classificações como "avançado" ou "transformacional" não possuem critério observável. Dois avaliadores atribuem notas diferentes à mesma organização, e nenhuma das duas é verificável por terceiro. Uma classificação que não pode ser auditada não é uma medida.

Agregação em índice único. Uma nota global elevada pode ocultar uma base próxima de zero. Uma organização que opera dezenas de casos de uso sobre dados desorganizados e sem governança não é madura: está exposta. O índice único registra essa exposição como progresso.

O reconhecimento dessas falhas tem um custo, que o AIMM assume deliberadamente: um modelo que as corrige produz um retrato mais desconfortável, porque revela a exposição que a nota agregada dissimulava. Um instrumento cuja finalidade seja tranquilizar uma instância de governança não deve adotar este modelo.

3. Estrutura do modelo

3.1 Escala de níveis

Cada dimensão é avaliada na mesma escala de cinco níveis. Os descritores abaixo são genéricos; §3.2 especifica o objeto de cada dimensão, e §3.5 detalha os critérios observáveis do nível 5 para as dimensões do piso.

Nível	Designação	Condição observável
1	Inicial	A prática ocorre por iniciativa individual, sem padrão nem registro.
2	Repetível	Existe prática documentada, aplicada de forma inconsistente entre áreas.
3	Definido	A prática é padronizada e adotada; os critérios são explícitos, mas a verificação é manual ou pontual.
4	Gerenciado	A prática é medida por indicadores objetivos e acompanhada ao longo do tempo.
5	Auto-corretivo	O sistema detecta e corrige desvios dentro de limites definidos; os critérios são verificáveis por terceiro.

A escala segue a tradição dos modelos por níveis, da qual o AIMM herda a estrutura e recusa o destino — ver §6.

3.2 As oito dimensões

Dimensão	Objeto da avaliação
Estratégia e portfólio	Como a IA se conecta à decisão de negócio, e o que a organização escolheu não fazer.
Dados e conhecimento	Se o dado que alimenta a IA existe, está organizado e é confiável.
Testes e validação	Se há um teste que impede uma versão pior de entrar em produção.
Agentes e autonomia	O quanto o sistema decide sozinho, e dentro de que limite.
Operação e custo	Se o sistema é monitorado em uso, e se o custo por decisão é conhecido.
Governança e risco	Quem responde quando a IA erra, e o que a impede de errar caro.
Adoção e redesenho	Se o processo de trabalho foi de fato redesenhado, ou apenas recebeu uma ferramenta.
Valor	O que efetivamente mudou no resultado — o realizado, não a promessa.

3.3 Nível atual e nível-alvo

Cada dimensão recebe dois valores: o nível atual, observado, e o nível-alvo, definido pela estratégia e pelo apetite a risco da organização. O alvo não é uniformemente o nível 5. Elevar uma dimensão a um nível superior ao necessário consome recurso sem gerar valor correspondente; em dimensões de baixo risco, o nível-alvo adequado pode ser 3.

Esta separação distingue o AIMM da maior parte dos modelos publicados, que tratam "mais maduro" como invariavelmente melhor. A condição em que ela falha: se a organização define o alvo por conforto — abaixo do risco que de fato corre —, o modelo registra uma lacuna aparentemente pequena sobre uma exposição real. O alvo é uma decisão de risco, e responde como tal.

3.4 O teto

A maturidade global de uma organização é limitada pela solidez de sua base. Três dimensões constituem o piso — Dados, Testes e validação, Governança — e a nota global obedece à regra:

$$\text{Maturidade global} \leq \text{menor}(\text{Dados, Testes e validação, Governança}) + 1$$

A justificativa é operacional. É admissível conduzir casos de uso um nível à frente do piso enquanto a base é construída; dois níveis à frente configura operação sem instrumentação — decisão sem verificação, ação sem responsável, dado sem confiança. Casos de uso em produção acima do teto não constituem maturidade: constituem exposição, e o modelo os classifica dessa forma.

A regra é uma heurística projetada, não uma lei medida. O termo "+1" expressa uma tolerância a risco, e uma organização pode decidir excedê-lo — desde que como decisão explícita, registrada e com responsável,

não como resultado acidental. O que o modelo impede é que a exposição transite por maturidade sem que alguém a tenha escolhido.

3.5 Critério verificável e o nível 5

O nível 5 de cada dimensão é definido por fatos verificáveis por um terceiro, não por adjetivos. Para as três dimensões do piso, os critérios observáveis são:

- **Dados** — a origem de cada dado é rastreável, do ponto de entrada ao uso; a qualidade é medida de forma contínua; o acesso é governado e registrado.
- **Testes e validação** — existe um conjunto de testes automatizados que impede a promoção de uma versão que regrida; resultados e cobertura são registrados; nenhuma versão entra em produção sem passar por ele.
- **Governança** — cada sistema de IA em produção tem responsável nomeado e limite de decisão definido; um mecanismo o impede de agir fora desse limite; incidentes são registrados e revisados.

Um critério que não possa ser observado por fora não é admitido no modelo. Essa é a razão pela qual nenhum nível é descrito por adjetivo — e o custo dessa escolha é que o modelo não captura o que não se deixa observar, como a qualidade de uma decisão estratégica isolada.

4. Anti-padrões de maturidade

O modelo nomeia quatro configurações que aparentam maturidade sem a constituir. Cada uma corresponde a um perfil reconhecível de avaliação.

Shadow AI em escala. Parte significativa da organização já emprega ferramentas de IA por iniciativa própria, sobre dados internos, sem registro de onde e com qual dado. Manifesta-se como adoção alta com governança baixa — uso disseminado que a organização não consegue nomear nem conter.

Museu de PoCs. Sucessão de provas de conceito bem-sucedidas em demonstração e nunca promovidas a produção. Manifesta-se como atividade intensa em experimentação com valor realizado próximo de zero.

Fortaleza vazia. Estrutura de governança, comitês e políticas bem constituída, incidente sobre um volume mínimo de sistemas em produção. Manifesta-se como controle robusto exercido sobre quase nada — o inverso do Shadow AI.

Piloto automático cego. Sistema em produção que decide de forma autônoma sem monitoramento correspondente. Manifesta-se como operação estável até o primeiro desvio não detectado, momento em que a ausência de instrumentação se converte em incidente.

5. Aplicação

Unidade de análise. A avaliação incide sobre uma capacidade de negócio — uma linha, um processo —, não sobre a organização como um todo. A maturidade varia entre áreas a ponto de tornar um índice organizacional único destituído de sentido operacional.

Participantes. A avaliação requer quem define a estratégia, quem responde pelos dados e quem responde pelo risco. Não é um exercício conduzido isoladamente pela área de tecnologia; conduzido assim, mede a percepção de uma função, não o estado da organização.

Produto do diagnóstico. A avaliação produz três resultados: o mapa *atual vs. alvo* nas oito dimensões; a posição do teto, que localiza a exposição; e a ação seguinte — qual dimensão do piso elevar primeiro. O produto não é um relatório destinado a arquivo, mas a base observável para uma decisão de construção.

6. Limitações do modelo

O AIMM é um instrumento, e um instrumento tem condições de falha que devem ser declaradas.

O teto é uma heurística. O intervalo de um nível acima do piso é uma escolha de tolerância a risco, não uma constante derivada de medição. Organizações com apetite a risco distinto podem calibrá-la — assumindo a decisão.

O modelo mede estado, não trajetória. Registra a posição no momento da avaliação; não prevê velocidade de progresso nem retorno de investimento, e não deve ser lido como tal.

Todo modelo por níveis tende à conformidade vazia. O CMMI, em parte relevante do mercado de software, degradou-se em teatro de conformidade: a obtenção do nível tornou-se o objetivo, e a prática permaneceu inalterada. O critério verificável reduz esse risco — não se audita um adjetivo, mas se audita um teste que barra um deploy —, porém não o elimina. Um instrumento honesto só produz resultado honesto quando quem o aplica busca a resposta, e não o carimbo. Essa condição é organizacional, e está fora do alcance do modelo.

7. Uso

A avaliação não se encerra em si mesma. Encerra-se em uma escolha: qual dimensão do piso elevar primeiro, e o que construir para isso. É nesse ponto que o modelo deixa de ser um retrato e passa a orientar trabalho.

O critério de nível 5 adotado por este modelo — sistemas que se autocorrigem dentro de limites definidos — deriva de sistemas em produção, não de formulação teórica. Quem estabeleceu a régua a estabeleceu operando IA sob carga real.

Avaliação da sua operação → [contato]