



# Manifesto da Metodologia MVG de Governança de Agentes de IA

**Paulo Henrique E. S. Silva**

**P2 Consultoria Brasil**

Livre para acesso, estudo, difusão e uso interno não comercial, com atribuição obrigatória.  
Uso comercial somente mediante autorização prévia e expressa.



A inteligência artificial está deixando de ser apenas uma ferramenta de apoio e passando a atuar como agente nos processos, decisões e operações das organizações.

Acreditamos que essa transformação não deve ser bloqueada pelo medo, nem conduzida pela pressa sem controle.

A governança de agentes de IA deve nascer mínima, prática e evolutiva: suficiente para proteger, leve o bastante para não engessar, incremental para crescer com a organização.

**Por isso, defendemos que:**

1. **Todo agente deve ter um dono humano.**

Nenhum agente de IA deve operar sem responsabilidade clara, supervisão e accountability.

2. **Toda autonomia deve ter limite.**

A liberdade operacional do agente deve ser proporcional ao risco, ao impacto e à criticidade da ação.

3. **A governança deve começar mínima e evoluir por maturidade.**

Ninguém implanta governança perfeita no primeiro dia. O caminho é incremental.

4. **O MVG deve proteger sem engessar.**

A Governança Mínima Viável existe para permitir inovação segura, não para bloquear experimentação.

5. **Cada agente deve ter seu próprio nível de maturidade.**

A maturidade deve considerar risco, autonomia, dados acessados, impacto e criticidade de cada agente.

6. **Todo prompt deve ser tratado como artefato de governança.**

Prompts, modelos, parâmetros, ferramentas e fontes de contexto precisam de controle, versionamento e evidência.

7. **Todo risco deve ser visível e avaliado.**

Riscos podem ser mitigados, transferidos, aceitos ou evitados conforme o apetite a risco. Não deve existir risco invisível.

8. **Toda decisão relevante gera evidência.**

Ações críticas de agentes devem deixar trilhas decisórias auditáveis, com contexto, fontes, aprovações e responsáveis.

9. **Toda aprovação humana deve ser real.**

Human-in-the-loop não pode ser apenas um clique automático. Deve representar julgamento humano efetivo.

10. **Todo agente deve poder ser contido.**

Agentes precisam de mecanismos de bloqueio, redução de autonomia, rollback ou kill switch.

11. **Toda exceção deve ter dono, prazo e justificativa.**

Exceções podem existir, mas não devem ser informais, permanentes ou invisíveis. Exceções devem ser registradas.

12. **Todo agente deve ter governança cognitiva.**

É preciso testar e registrar como mudanças no modelo, versão, parâmetros, fallback ou provedor de LLM afetam o comportamento, a confiabilidade, a segurança e a autonomia do agente.



Não buscamos uma governança perfeita no primeiro dia.

Buscamos agentes visíveis, riscos conhecidos, autonomia proporcional, evidências rastreáveis, supervisão humana efetiva e evolução contínua da maturidade.

**Paulo Henrique E. S. Silva**

**P2 Consultoria Brasil**

**Metodologia MVG de Governança de Agentes de IA**

Livre para acesso, estudo, difusão e uso interno não comercial, com atribuição obrigatória.

Uso comercial somente mediante autorização prévia e expressa.