

Ética da IA na escola: os seis princípios e a mentalidade centrada no ser humano

O marco da UNESCO (Organização das Nações Unidas para a Educação, a Ciência e a Cultura) condensa a ética da IA em seis princípios que o professor deve entender, aplicar e ajudar a adaptar. Aqui, agrupados em cinco páginas e lidos pela fricção cognitiva ótima: cada princípio já tem um lugar no desenho do sistema.

1

Não causar dano e proporcionalidade

Toda ferramenta deve conhecer seus impactos, e a proteção deve crescer com o risco do uso.

2

Não discriminação

A exclusão costuma estar embutida no design; cabe garantir uso inclusivo por e para todos os alunos.

3

Sustentabilidade

O ciclo de vida da IA tem custo ambiental; a validação deve priorizar ferramentas responsáveis.

4

Determinação humana

Seleção, uso e monitoramento da IA sob controle e responsabilidade de pessoas nomeadas.

5

Transparência e explicabilidade

O usuário tem o direito de questionar como a IA chega às conclusões; a regra precisa ser legível.

Ética não é anexo do contrato: é mentalidade

Quatro pontos de partida do marco para a mentalidade centrada no ser humano.

DESFAZER O ALARDE

O design e o uso da IA são **conduzidos por humanos**: quem cria os sistemas decide se eles protegem e ampliam capacidades ou se carregam vieses que violam direitos. Só o design intencionalmente ético e a implantação bem regulada avançam capacidades, inclusão e sustentabilidade.

ENTENDER AS AMEAÇAS DE DESIGN

Os caminhos algorítmicos atuais desafiam direitos humanos e privacidade, e o conteúdo gerado por IA vem minando conhecimentos, culturas e línguas indígenas. Entender como os modelos funcionam é condição para proteger a **diversidade linguística e cultural**.

FAZER PREVALECER VALORES HUMANOS

Algoritmos movidos a lucro enfraquecem a coesão social ao isolar as pessoas do mundo real e umas das outras. Empatia, altruísmo, justiça e solidariedade são essenciais; a IA **não deve desestimular o contato** com os outros e com o mundo.

Conduzir a IA para o desenvolvimento humano. O uso de IA sem orientação pedagógica pode enfraquecer o desenvolvimento intelectual dos alunos; a meta deve ir além do acesso à informação e das respostas padronizadas, rumo ao enriquecimento da investigação e da capacidade. É a mesma tese da fricção cognitiva ótima.

1 Não causar dano e proporcionalidade

A dose de proteção cresce com o que está em jogo.

O PRINCÍPIO

Toda ferramenta deve ser desenhada com **compreensão clara de seus impactos** sobre direitos, dignidade, segurança e bem-estar; e a resposta regulatória deve ser proporcional ao risco, como nos quatro níveis do AI Act europeu.

NA ESCOLA

Risco moderado (lição de casa comentada, sem nota): o professor lê cada comentário antes de liberar. Alto risco (nota final, vestibular, decisão sobre um aluno): decide o professor; a IA não dá nota a produção autoral, e cada nota de item objetivo é validada, com **registro auditável**.

NA FRICÇÃO ÓTIMA

É o hábito 4: quanto maior o custo do erro, maior a fricção. Na ontologia operacional, vira princípio das camadas de regras e ações: **ação mais governada onde o risco é maior**.

No IA para escolas. Os oito *Tipos de correção com IA* são uma escala de proporcionalidade: dos comentários sem nota (risco moderado) à correção validada (itens objetivos, com cada nota validada antes de o aluno ver) e à leitura por critérios (produções autorais, em que a IA não atribui nota e o professor decide).

2 Não discriminação

A exclusão costuma estar embutida no design, não na intenção de quem usa.

O PRINCÍPIO

Exclusão e discriminação estruturais são frequentemente **embutidas no design e no uso** da IA. Cabe ao professor garantir uso inclusivo por e para todos os alunos, independentemente de gênero, etnia, habilidades ou condição socioeconômica e migratória.

NA ESCOLA

Observar vieses das ferramentas que possam marginalizar alunos com deficiência ou de grupos vulneráveis e **reportá-los** aos gestores; priorizar ações que reduzam o impacto negativo sobre equidade e inclusão.

NA FRICÇÃO ÓTIMA

Equidade é calibragem: a dificuldade desejável para um aluno pode ser barreira incapacitante para outro. Para quem tem dificuldades de aprendizagem, **mais apoio é equidade**, não facilitação.

No IA para escolas. Vieses não se eliminam, reconhecem-se e administram-se: a correção assistida torna visíveis cinco vieses humanos (simpatia, rotulação, cansaço, letra, reclamação) e os calibra contra o viés algorítmico, sob auditoria (*Viés humano e viés algorítmico*).

3 Sustentabilidade

O custo ambiental da IA é parte da decisão pedagógica, não nota de rodapé.

O PRINCÍPIO

O "não causar dano" inclui o ambiente: o ciclo de vida e a cadeia de valor da IA podem agravar a crise climática. A validação deve priorizar ferramentas **ambientalmente responsáveis**, e o professor deve conhecer as emissões de carbono da IA.

NA ESCOLA

Perguntar que finalidade justifica o custo ambiental de cada uso e preferir usos proporcionais à tarefa. Levar o tema à aula: o marco sugere que alunos criem ferramentas para **rastrear o consumo de energia** da escola ou as emissões dos serviços de IA.

NA FRICÇÃO ÓTIMA

Sustentabilidade e fricção convergem: se a IA não deve entregar o produto pronto, também não precisa rodar em todo passo da tarefa. **Menos chamadas à máquina, mais esforço do aluno**: a pedagogia que preserva a aprendizagem também poupa recursos.

No IA para escolas. A regra "tente primeiro, IA depois" e a retenção do produto cognitivo reduzem, por desenho, o número de vezes que a máquina é chamada; e o critério de *IA que conecta* manda reinvestir o que a automação economiza em capacidade humana, não em mais automação.

4 Determinação humana

Quem decide continua sendo uma pessoa, com nome e responsabilidade.

O PRINCÍPIO

Determinação humana na interação humano-IA: a seleção, o uso e o monitoramento do impacto da IA devem ser **controlados e responsabilizados por humanos**. O professor permanece responsável pelas decisões pedagógicas no uso da IA.

NA ESCOLA

Nos circuitos de decisão com IA (gestão, avaliação, estratégias de ensino, interação dos alunos), a responsabilização humana é **obrigação legal**; resistir a que saídas e previsões da IA usurpem decisões do professor e o pensamento dos alunos.

NA FRICÇÃO ÓTIMA

É o analista no circuito: quem define a verdade operacional **governa, não apenas referenda**. E é o hábito 1: para terceirizar, é preciso saber avaliar; sem capacidade avaliativa, a terceirização vira dependência.

No IA para escolas. "Quem revisa é quem governa": em qualquer tipo de correção, o professor permanece a autoridade que define o que conta como boa devolutiva nesta escola; a ontologia registra o que exige aprovação humana, o que fica auditável e o que a IA só pode sugerir.

5 Transparência e explicabilidade

Uma nota que não se explica não se contesta; uma regra que não se lê não se governa.

O PRINCÍPIO

Modelos que produzem respostas convincentes podem não ser explicáveis e esconder riscos. O usuário tem o **direito de questionar a explicabilidade**; instituições e fornecedores respondem pela transparência do desempenho, dos resultados e dos impactos.

NA ESCOLA

Exigir que a ferramenta mostre como chegou à conclusão; **rotular o conteúdo gerado** por IA; observar se as regulações locais acompanham as iterações da tecnologia e cotejá-las com os princípios éticos.

NA FRICÇÃO ÓTIMA

A proteção não pode viver só no prompt: regras em linguagem natural são burladas e não se auditam. Codificar objetos, regras e ações em **formato executável e auditável** é o que torna a ética verificável.

No IA para escolas. Explicitar os critérios de correção é, em si, formativo: obriga a tornar pública a verdade operacional da correção, antes tácita no bom senso de cada docente (*Fricção cognitiva ótima*, Seção onze). A devolutiva traz o critério aplicado, não só a nota.

O que a escola assume

- ✓ Toda ferramenta passa por validação antes de chegar ao professor; a responsabilidade ética começa na instituição.
- ✓ Vieses são procurados ativamente, registrados e reportados; a inclusão de alunos com deficiência é critério de adoção.
- ✓ Toda decisão com IA tem um responsável humano nomeado, e o professor pode recusar o uso.
- ✓ A supervisão humana cresce com o risco: sem nota final, decisão sobre aluno ou exame de alto risco por IA autônoma.
- ✓ O custo ambiental entra na decisão: a IA é chamada onde agrega, não em todo passo da tarefa.
- ✓ Regras e critérios são escritos, auditáveis e revisados; o conteúdo gerado por IA é rotulado.

Créditos e referências

Terceiro material da série *Posicionamentos*, que lê marcos normativos pela lente do documento *Fricção cognitiva ótima: referencial teórico para o design de ferramentas de IA na Educação Básica*. Os seis princípios vêm do bloco 2.1 do marco (Princípios éticos), e os quatro pontos de partida da mentalidade centrada no ser humano, do capítulo 2, que os ancora na Recomendação sobre a Ética da Inteligência Artificial (UNESCO, 2022). A leitura de cada princípio pela fricção cognitiva ótima é do **IA para escolas**.

MIAO, F.; CUKUROVA, M. AI competency framework for teachers. Paris: UNESCO, 2024. DOI 10.54675/ZJTE2084.

UNESCO. Recommendation on the ethics of artificial intelligence. Paris: UNESCO, 2022.

UNESCO. Guidance for generative AI in education and research. Paris: UNESCO, 2023.

Oldimar Cardoso • iaparaescolas.com.br • 2026