XXVIII SEMINÁRIO DE FILOSOFIA · MINICURSO

Escrever com IA sem deixar de pensar

Material do Participante

Participante: _____

Faculdade Católica da Paraíba · Cajazeiras, 23 de outubro de 2026 · Ministrante: Gervásio Teixeira (mestre em Ciência da Computação, UFPE; graduado em Filosofia)

Pergunta-guia do minicurso

Quando um texto é escrito com ajuda de uma máquina, o que nele ainda é meu? Escreva nas margens, discorde, rabisque: este material é seu. Cada seção tem um **texto curto**, **perguntas para refletir** e uma **atividade** para fazer no computador.

Antes de começar	Cinco regras de ouro
☐ Computador com internet ☐ Conta em uma IA gratuita (ChatGPT, Gemini, Claude ou Copilot) ☐ Abas: Google Acadêmico e SciELO ☐ Editor de texto aberto (Word ou Google Docs) ☐ Um tema real seu (TCC, artigo ou projeto)	1. Eu penso primeiro; depois peço ajuda. 2. Nada que a IA entrega vale como fonte antes de eu verificar. 3. Não colo dados pessoais nem trabalhos sigilosos. 4. Registro o que fiz: é a minha declaração de uso. 5. Se a IA concordar demais, peço que discorde.

Antes de tudo: de onde vem a IA e por que ela "estourou"

Séc. XVII · Leibniz sonha com um cálculo para resolver disputas.

1943 · Modelo matemático de neurônio (McCulloch e Pitts).

1950 · Turing e o jogo da imitação.

1956 · Dartmouth: nasce o nome "inteligência artificial".

Anos 1970 a 1990 · Os "invernos da IA".

1986 · Retropropagação: redes de várias camadas.

2012 · AlexNet e o aprendizado profundo (GPUs).

2017 · Transformer: *Attention is all you need*.

2020 · GPT-3 e as leis de escala.

2022 · Ajuste com feedback humano e ChatGPT (30/11).

Por que "estourou"? **Dados** + **cálculo** (GPUs) + **arquitetura** (transformer) + **escala** + **conversa** (ajuste com feedback humano e chat aberto a todos). Nenhum deles é uma nova teoria do pensamento.

1 · Autoria na era da IA

No *Fedro* (274c-275b), Platão conta que o deus Theuth ofereceu a escrita ao rei Thamus como *pharmakon*, palavra grega que quer dizer ao mesmo tempo remédio e veneno, para a memória e a sabedoria. Thamus respondeu que ela traria esquecimento e uma "aparência de sabedoria". A IA que escreve é o caso mais recente da mesma história: remédio para quem já pensou, veneno para quem ainda não pensou. A ferramenta é a mesma; o que decide é o uso.

Kant, em *Resposta à pergunta: que é esclarecimento?* (1784), chama de menoridade a incapacidade de servir-se do próprio entendimento sem a direção de outrem, e lhe atribui uma causa: falta de decisão e de coragem. Seu exemplo: "um livro que entende por mim, um pastor que tem consciência por mim, um médico que decide minha dieta". A IA entra nessa lista? *Sapere aude!*

Para refletir: aplique ao seu caso (um texto seu em que já usou ou pretende usar IA)

Pergunta	Minha resposta
Quem pensou? A ideia, o recorte e a tese partiram de mim ou do modelo?	
Quem decidiu? Cada escolha (o que entra, o que sai, qual fonte) passou pelo meu juízo?	
Quem responde? Se houver erro, plágio ou invenção, quem assume?	

Minha definição de autoria (antes do debate):

Definição construída pelo grupo:

2 · Como uma IA funciona

Um **modelo de linguagem** (LLM) é uma rede neural treinada em enormes volumes de texto para uma tarefa aparentemente simples: **prever a próxima palavra**. O percurso de uma frase tem cinco etapas e se repete a cada palavra:

> **Texto** → **tokens** (pedaços de palavras, viram números) → **vetores** (listas de números que representam cada token) → **atenção** (pondera o que importa no contexto) → **probabilidades** (chance de cada próxima palavra) → a palavra escolhida volta ao texto e o ciclo recomeça.

Vetores e uso. Palavras usadas em contextos parecidos recebem vetores parecidos. Isso lembra Wittgenstein (*Investigações filosóficas*, §43: o significado de uma palavra é, para uma grande classe de casos, o seu uso na linguagem). É uma analogia, não uma equivalência: o modelo capta padrões de uso, mas não participa de uma forma de vida. **Atenção**, a ideia central da arquitetura *transformer* (Vaswani et al., 2017), é o que lhe permite pesar o contexto.

Probável não é verdadeiro. O modelo atribui probabilidades e sorteia; a **temperatura** controla o quanto arrisca. Por isso a mesma pergunta pode ter respostas diferentes. Aristóteles já distinguia o verossímil do verdadeiro (*Poética*, cap. 9). O modelo não consulta um banco de fatos, tem data de corte de conhecimento e herda os vieses dos textos de que aprendeu.

Alucinação é conteúdo gerado sem fundamento nos dados ou na realidade (Ji et al., 2023); na escrita acadêmica, aparece como referências, autores e citações inexistentes, entregues com total confiança. Em analogia com Frankfurt (*On Bullshit*), é fluência indiferente à verdade. Bender et al. (2021) chamam os modelos de "papagaios estocásticos": recombinam formas sem acesso ao sentido.

Dois clássicos para pensar. Turing (1950) trocou "as máquinas pensam?" pelo jogo da imitação: conversar bem não prova compreensão. Searle (1980), com o quarto chinês, sustenta que manipular símbolos (sintaxe) não basta para compreender (semântica); aplicá-lo aos modelos atuais é uma hipótese de leitura, e o debate segue aberto.

Verdadeiro ou falso? Marque V ou F

Afirmação	V/F
1. Ao responder, a IA consulta um banco de fatos verificados.	
2. Com temperatura alta, as respostas ficam mais variáveis.	
3. Alucinação é conteúdo gerado sem fundamento nos dados ou na realidade.	
4. Se duas ferramentas citam a mesma referência, ela certamente existe.	
5. Uma resposta fluente indica que está correta.	
6. A mesma pergunta pode gerar respostas diferentes.	

Experimento 1 · Variabilidade (faça em sala): peça duas vezes "Explique em um parágrafo o conceito de pharmakon no Fedro" e compare.

Resposta 1 (resumo)	Resposta 2 (resumo)	O que mudou?

Para refletir: se a IA escreve bem e erra pela mesma razão, o que isso exige de quem a usa?

3 · Pesquisa e referências

No *Teeteto* (155d), Platão diz que o espanto é o princípio da filosofia. Em pesquisa, o espanto precisa ser lapidado em um problema investigável: **tema → recorte → pergunta → hipótese**. Exemplo: tema "IA e autoria"; recorte "texto gerado por IA em trabalhos de graduação em Filosofia"; pergunta "em que sentido o texto escrito com IA é de autoria de quem a aciona?"; hipótese "autoria exige decisão e responsabilidade, não só execução". A IA ajuda a afinar o recorte; **o espanto é seu**.

Prompt em cinco camadas: *papel* (quem a IA deve ser), *contexto* (seu tema e nível), *tarefa* (um verbo), *formato* (tabela, lista) e *restrições* (o que evitar). Exemplo: "Você é professor de Filosofia Antiga. Escrevo um TCC sobre a *phronesis* em Aristóteles. Proponha 3 recortes de problema de pesquisa, em tabela (recorte, pergunta, o que fica de fora). Não cite livros nem autores que não possa indicar com certeza; diga quando não souber." A restrição final reduz, mas não elimina, a invenção.

Onde buscar. Bases (a fonte existe): Google Acadêmico, SciELO, Portal de Periódicos CAPES, PhilPapers, Stanford Encyclopedia of Philosophy. IAs de apoio (conferir sempre): ChatGPT, Gemini, Claude, Copilot; Perplexity, Consensus, Elicit; NotebookLM; Zotero para organizar referências. Os recursos gratuitos mudam sem aviso. Dica de busca: aspas para expressão exata, AND, OR e NOT (ou o sinal de menos no Google Acadêmico), como em *"inteligência artificial" AND (autoria OR plágio)*.

Verificação em 4 passos (a dúvida metódica de Descartes: nada vale como fonte antes de ser conferido)	
1. Existe? Procure o título exato em uma base ou no catálogo de uma biblioteca.	2. Os dados conferem? Autor, ano, veículo e DOI (identificador permanente).
3. Diz isso? Leia o trecho: a obra existe, mas pode não dizer o que a IA afirma.	4. Sustenta? A fonte serve ao seu argumento ou só enfeita o texto?

Atividade 1 · Do espanto à referência verificada

1. Meu espanto (tema ou incômodo filosófico que me chama a atenção), em 2 linhas:

2. Meu prompt de cinco camadas (cole na IA e peça 3 recortes)

Camada	**O que escrevi**
Papel	
Contexto	
Tarefa e formato	
Restrições	Não cite livros nem autores que você não possa indicar com certeza. Diga quando não souber.

3. Os recortes sugeridos e a minha escolha

Recorte	**Pergunta de pesquisa**	**O que deixa de fora**

Meu problema de pesquisa em uma frase (em forma de pergunta):

4. Referências sugeridas pela IA e os quatro passos (marque ✔ ou ✘; status: real, real com erro ou inexistente)

Referência sugerida	**1**	**2**	**3**	**4**	**Status**

Exemplo resolvido (didático): "ARISTÓTELES. *Ética a Nicômaco*, livro VI" → ✔✔✔✔ **real**. "SILVA, M. *A prudência aristotélica na era digital*, 2021" (inventada para este exemplo) → não encontrada, **inexistente**. "ARISTÓTELES. *Ética a Nicômaco*, livro X, sobre a prudência" → obra real, mas a prudência está no livro VI: **real com erro**.

5. Conferência do colega: troque a tabela e confira uma referência dele pelos quatro passos

Referência conferida	Meu resultado	Concordamos?

O erro mais sutil que encontrei hoje foi:

---

4 · A IA como interlocutora socrática

A maior utilidade filosófica de um modelo de linguagem não é escrever por você, mas **conversar com você**. **Maiêutica** (Platão, *Teeteto*, 148e-151d): Sócrates se descreve como parteiro, que não gera o saber no outro, mas ajuda a dar à luz e examina se a ideia é viável. Passos: pergunta → definição provisória → refutação (*élenchos*) → aporia (impasse produtivo) → nova tentativa. **Disputatio** (Tomás de Aquino, *Suma teológica*): cada artigo enfrenta primeiro as objeções (*objectiones*), cita uma autoridade em contrário (*sed contra*), responde (*respondeo*) e replica a cada objeção. Em computação, o equivalente é o **teste adversarial**: tentar quebrar o sistema antes de lançá-lo. Aqui, quebramos o argumento antes que o leitor o faça.

Cuidado com a bajulação: modelos tendem a concordar com o usuário. Se a IA elogiar demais a sua tese, escreva "Discorde. Ataque o ponto mais frágil." Platão (*Górgias*) chama de adulação a fala que agrada sem compromisso com a verdade.

Prompt socrático (copie e cole)

```
Aja como Sócrates. Vou apresentar uma tese. Você só pode fazer perguntas, uma por vez, para me levar a definir melhor meus conceitos e a perceber contradições. Nunca dê a resposta. Comece perguntando o que eu entendo por cada termo central. Minha tese: [cole aqui]
```

Advogado do diabo (copie e cole)

```
Aja como um professor rigoroso da tradição oposta à minha. Apresente as 3 objeções mais fortes à minha tese, cada uma com a premissa em que ataca. Não ofereça soluções. Depois, indique qual objeção eu teria mais dificuldade de responder e por quê. Minha tese: [cole aqui]
```

Falácia (Aristóteles, *Refutações sofísticas*)	Pergunta-teste	Achei na minha tese?
Petição de princípio	A conclusão já está dentro da premissa?	☐ sim ☐ não
Falso dilema	Há alternativas que ficaram de fora?	☐ sim ☐ não
Apelo à autoridade ("disse a IA")	A autoridade é competente e verificável?	☐ sim ☐ não
Ad hominem	O ataque é ao argumento ou à pessoa?	☐ sim ☐ não

Atividade 2 · Seu argumento no tribunal da IA

1. Minha tese (3 a 5 linhas):

---

2. Rodadas socráticas (5 ou 6): onde as minhas definições eram frágeis?

Pergunta da IA	Minha resposta (resumo)	O que percebi

3. Objeção mais forte do advogado do diabo (e por que é a mais forte):

4. Minha tese reformulada, respondendo a essa objeção:

A pergunta da IA que mais me fez pensar foi:

5. Explique a nova tese à sua dupla em 1 minuto, sem ler. Feedback do(a) colega:

Entendi a tese?	O ponto que ficou obscuro	A melhor parte do argumento
☐ sim ☐ em parte ☐ não		

5 · Estruturar, revisar e responder por si

Depois do rascunho, a IA pode ser **leitora crítica**: pergunta e aponta, sem escrever no seu lugar. A estrutura mínima (começo, meio e fim, como já pedia Aristóteles na *Poética*, cap. 7): **introdução** (problema, tese, método), **desenvolvimento** (argumentos e fontes) e **conclusão** (responde à pergunta da introdução?). Dica: o histórico de versões do editor de texto ajuda a demonstrar autoria.

Pode delegar	Com cuidado	Não delegue
Corrigir ortografia e gramática Formatar referências (e conferi-las) Sugerir sinônimos Gerar perguntas para revisão	Resumir textos (leia o original) Sugerir estrutura de seções Reformular parágrafos seus Traduzir trechos (confira os termos técnicos)	Escolher o problema e a tese Interpretar os autores Decidir o que é verdadeiro Assinar e responder pelo texto

Aristóteles (*Ética a Nicômaco*, VI) chama de **prudência** (*phronesis*) a capacidade de deliberar bem em situações concretas; ela não se transfere a um terceiro. Floridi e Cowls (2019) propõem cinco princípios para uma IA socialmente benéfica: **beneficência, não maleficência, autonomia, justiça e explicabilidade**. Jonas defende que a responsabilidade seja proporcional ao poder das técnicas. Na prática: não cole dados pessoais nem trabalhos inéditos de terceiros em ferramentas públicas (LGPD, Lei nº 13.709/2018); respeite os direitos autorais (Lei nº 9.610/1998); siga as normas da ABNT (NBR 6023:2025, 10520:2023 e 14724:2024; não há norma da ABNT para citar IA: declare o uso) e as regras de sua instituição e de seu orientador sobre o uso de IA.

Normas brasileiras hoje. Não há lei geral de IA (o PL 2338/2023 ainda tramita). O **Parecer CNE/CP nº 15/2026** (aprovado em 1º/9/2026; a resolução depende de homologação do MEC) manda cada instituição criar sua política de IA e, na pesquisa, declarar o uso quando relevante; a IA não assume a autoria nem a responsabilidade, e não se pode punir com base só em detector de IA. A **Portaria CNPq nº 2.664/2026** exige, na pesquisa apoiada pelo CNPq, declarar a ferramenta e a finalidade e veda apresentar texto gerado como autoria humana. Vale sempre a política da sua instituição e do seu orientador.

Prompt do leitor crítico (copie e cole)

```
Leia a introdução abaixo como um orientador exigente. Responda: (1) qual é o problema? (2) qual é a tese? (3) qual é o método? (4) o que está vago? Não reescreva. Texto: [cole aqui o seu parágrafo]
```

Atividade 3 · Do rascunho ao texto assinado

Regra de ouro: o primeiro parágrafo é escrito **sem IA**. Ela só entra depois, como leitora crítica.

1. Meu parágrafo, sem IA (6 a 8 linhas, com a minha tese):

2. O que a IA apontou como leitora crítica, e a minha decisão

O que a IA apontou	Aceito?	Por quê (minha decisão)

3. Meu parágrafo reescrito, por minha conta:

4. Revisão em pares: o(a) colega marca

Pergunta	Sim	Não
O problema está claro nas duas primeiras frases?	☐	☐
A tese aparece de forma explícita?	☐	☐
Cada afirmação importante tem fonte ou justificativa?	☐	☐
Há alguma objeção reconhecida e respondida?	☐	☐
O texto soa como a voz de quem o assina?	☐	☐

5. Minha declaração de uso de IA

DECLARAÇÃO DE USO DE INTELIGÊNCIA ARTIFICIAL	
Trabalho / Autor(a) / Data	
Ferramenta(s) utilizada(s)	
Finalidade do uso	
Não utilizei IA para	
Verificação e responsabilidade	☐ Conferi as referências nas fontes originais. ☐ Assumo integralmente o conteúdo deste trabalho.
Assinatura	

Fechamento · Meus compromissos e a avaliação

Kant resumiu as perguntas da filosofia em quatro (*Crítica da razão pura*, A805/B833; *Lógica*): **que posso saber? que devo fazer? que me é permitido esperar? que é o homem?** A IA devolve a última com urgência nova. Escolha três compromissos para adotar já:

	Compromisso		Compromisso
☐	Eu penso primeiro. Defino o problema e a tese antes de pedir ajuda.	☐	Eu testo meus argumentos. Uso a IA para achar objeções.
☐	Eu pergunto melhor. Uso prompts com as cinco camadas.	☐	Eu protejo dados. Não exponho dados pessoais nem trabalhos de terceiros.
☐	Eu verifico tudo. Nenhuma referência entra no texto sem os quatro passos.	☐	Eu declaro e respondo. Informo o uso de IA e assino o que escrevo.

Avaliação do minicurso (anônima, se preferir): 1 = fraco, 5 = excelente	1	2	3	4	5
Conteúdo e clareza · Atividades práticas · Material impresso · Organização do tempo (dê uma nota geral)	☐	☐	☐	☐	☐

O que mais gostei / o que levo para o meu TCC ou artigo / o que melhoraria:

Referências citadas

BENDER, E. M. et al. On the dangers of stochastic parrots. In: FAccT, 2021. p. 610-623. · BRASIL. Lei nº 9.610/1998; Lei nº 13.709/2018; CNE, Parecer CNE/CP nº 15/2026; CNPq, Portaria nº 2.664/2026; ABNT NBR 6023:2025, 10520:2023, 14724:2024. · FLORIDI, L.; COWLS, J. A unified framework of five principles for AI in society. *Harvard Data Science Review*, v. 1, n. 1, 2019. · FRANKFURT, H. G. *On bullshit*. Princeton: Princeton University Press, 2005. · JI, Z. et al. Survey of hallucination in natural language generation. *ACM Computing Surveys*, v. 55, n. 12, art. 248, 2023. · JONAS, H. *O princípio responsabilidade*. Rio de Janeiro: Contraponto; Ed. PUC-Rio, 2006. · KANT, I. Resposta à pergunta: que é esclarecimento? In: *Textos seletos*. Petrópolis: Vozes, 2005. · KAPLAN, J. et al. Scaling laws for neural language models. arXiv:2001.08361, 2020. · KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, v. 25, 2012. · McCARTHY, J. et al. A proposal for the Dartmouth summer research project on artificial intelligence. 1955. · OUYANG, L. et al. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, v. 35, 2022. · PLATÃO. *Fedro*. Belém: Ed. UFPA, 2011; *Teeteto-Crátilo*. Belém: Ed. UFPA, 2001. · SEARLE, J. R. Minds, brains, and programs. *Behavioral and Brain Sciences*, v. 3, n. 3, p. 417-424, 1980. · TURING, A. M. Computing machinery and intelligence. *Mind*, v. 59, n. 236, p. 433-460, 1950. · VASWANI, A. et al. Attention is all you need. *Advances in Neural Information Processing Systems*, v. 30, 2017. · WITTGENSTEIN, L. *Philosophical investigations*. Oxford: Blackwell, 1953, §43. · Clássicos citados por obra e passagem: Aristóteles (*Ética a Nicômaco* VI; *Poética* 7 e 9; *Refutações sofísticas*), Tomás de Aquino (*Suma teológica*), Descartes (*Meditações*; *Discurso do método*), Kant (*Crítica da razão pura*; *Lógica*).