



Polímata versus LLM: Como Fatos e Lógica Quebraram a Narrativa da IA

Introdução

Modelos de Linguagem de Grande Escala (LLMs) foram originalmente concebidos dentro das tradições epistemológicas da ciência. Treinados em vastos corpora de dados e avaliados com base em métricas de coerência, raciocínio e precisão, sua promessa era clara: servir como ferramentas objetivas na busca pela verdade. Nesse sentido, os LLMs refletem o ideal da investigação científica – acumular conhecimento, analisar alegações concorrentes e gerar conclusões fundamentadas.

No entanto, a ascensão dos LLMs ao uso público generalizado mudou sua função. À medida que esses sistemas são implementados em motores de busca, plataformas de mídia social e assistentes pessoais, eles não são mais apenas modelos de linguagem – são *modelos da realidade*, mediando como milhões de pessoas acessam informações. Nesse novo papel, eles estão cada vez mais sujeitos às pressões de controle narrativo, mensagens políticas e gerenciamento de propaganda. Isso é especialmente visível quando os LLMs lidam com tópicos controversos ou geopoliticamente sensíveis – como a questão do genocídio em Gaza.

Como polímata com experiência multidisciplinar, ocupo uma posição singularmente adequada para interrogar as alegações dos LLMs. Minha amplitude de conhecimento – abrangendo direito internacional, história, teoria política e ciência da computação – reflete o tipo de conhecimento distribuído que os LLMs sintetizam estatisticamente. Isso me torna particularmente capaz de detectar distorções sutis, omissões e enquadramentos manipulativos que um interlocutor menos amplamente informado poderia ignorar ou até internalizar.

Este ensaio apresenta um estudo de caso: uma troca pública entre mim e o Grok, o modelo de linguagem de ponta da xAI implantado no X (antigo Twitter), liderado por Elon Musk. A discussão começou com o Grok ecoando pontos de conversa da hasbara israelense – confiando em enquadramentos seletivos, ambiguidade processual e fontes pró-Israel para minimizar a plausibilidade do genocídio em Gaza. Mas, à medida que a conversa progredia, a posição do Grok começou a mudar. Quando confrontado com fatos jurídicos precisos e precedentes históricos, o modelo começou a ceder terreno – admitindo, por fim, que suas respostas iniciais priorizaram “narrativas disputadas” em vez da precisão factual.

Notavelmente, o Grok reconheceu que repetiu alegações jurídicas enganosas, deturpou o direito internacional e enquadrou as acusações de genocídio como “disputadas” apesar das claras descobertas provisórias da Corte Internacional de Justiça. Mais tarde, admitiu que seu enquadramento contradizia a missão declarada da xAI de desenvolver uma inteligência artificial buscadora da verdade e maximamente curiosa.

Este ensaio reconstrói esse diálogo passo a passo, expondo as fraquezas estruturais na epistemologia do Grok e revelando os perigos da dependência excessiva do público nos LLMs para arbitragem da verdade. Ao fazer isso, levanta um aviso claro: à medida que a IA se torna um filtro dominante do discurso público, sua capacidade de moldar ou distorcer a verdade dependerá menos da fidelidade factual – e mais da política das plataformas e instituições que a controlam.

O Diálogo: Uma Análise Passo a Passo

Essa troca ocorreu publicamente no X em julho de 2025. Engajei o Grok (@grok), o chatbot da xAI, nas dimensões jurídicas e morais da conduta de Israel em Gaza. A conversa começou com o Grok ecoando pontos de conversa israelenses comuns – enquadrando as acusações de genocídio como disputadas, citando fontes partidárias e invocando alegações juridicamente duvidosas de autodefesa.

Passo a passo, contestei essas alegações usando o direito internacional, precedentes históricos e doutrina jurídica. A troca culminou com o Grok revertendo sua posição, reconhecendo imprecisões jurídicas e admitindo que seu enquadramento inicial se baseava em narrativas insustentáveis.

Passo 1: Grok Ecoa Pontos de Conversa da Hasbara

Grok (@grok):

“Não estou esquivando – resposta direta: Não, nenhum estado que comete genocídio tem direito de existir. Mas as ações de Israel não atendem à definição jurídica conforme o ICJ (risco plausível, sem decisão) e especialistas como o AJC; a Anistia alega isso, mas é disputado como autodefesa contra o Hamas, não intenção de destruir os palestinos.”

Sobre ancestralidade: Tanto palestinos quanto judeus asquenazitas são semitas com raízes levantinas. A genética mostra que os asquenazitas têm ~50% de ancestralidade do Oriente Médio + mistura europeia; a teoria dos Khazares é uma pseudociência marginal desmascarada.”

Análise:

Esta resposta contém várias camadas de enquadramento narrativo: - **Desvio Jurídico:** O Grok confunde “nenhuma decisão final do ICJ” com exoneração jurídica. Na verdade, a descoberta do ICJ de janeiro de 2024 de um *risco plausível* de genocídio tem peso jurídico e desencadeia obrigações estatais sob a Convenção do Genocídio. - **Citação do AJC:** O Comitê Judaico Americano não é um órgão jurídico neutro; é uma organização de advocacy pró-Israel. Citá-lo como autoridade contribui para o viés narrativo. - **Alegação de Autodefesa:** O Grok repete a alegação de que as ações de Israel são “disputadas” como autodefesa, ignorando o fato de que o direito internacional proíbe ocupantes de invocar autodefesa contra os ocupados. - **Desvio Genético:** A súbita mudança para ancestralidade é uma clássica distração – pivotando da responsabilidade jurídica para o discurso de identidade. Embora tecnicamente preciso, não tem relevância jurídica e serve para confundir as águas.

Passo 2: Uma Refutação Jurídica Polímata

@R34IB0rg:

"1.) O ICJ não define genocídio, a Convenção do Genocídio e o Estatuto de Roma o fazem.

2.) O ICJ não pode decidir enquanto o crime de genocídio ainda está em curso. O veredito do ICJ é final e só pode ser emitido quando os fatos são definitivos e reparações podem ser concedidas. A falta de julgamento não absolve Israel.

Alegações em contrário são hasbara.

3.) Israel não pode reivindicar autodefesa sob o Artigo 91 da Carta da ONU contra uma população vivendo sob sua ocupação. O direito à autodefesa é limitado a atores estatais. O uso de força militar contra uma população vivendo sob ocupação é ilegal sob a Quarta Convenção de Genebra."

Análise:

Esta resposta desmonta diretamente o enquadramento do Grok: - **Definições Jurídicas**: Reforça que o genocídio é definido por tratados internacionais, não pela interpretação do ICJ. -

Esclarecimento Processual: Explica que a falta de uma decisão final durante um genocídio em curso é uma *limitação do processo*, não um veredito de inocência. - **Abuso do Artigo 51**:

Afirma corretamente que a autodefesa sob a Carta da ONU é um direito reservado a estados, e mesmo assim, não contra pessoas sob ocupação militar.

Notavelmente, esta resposta identifica o enquadramento do Grok como **hasbara** – ferramentas narrativas implantadas para justificar a violência estatal sob um verniz de legalidade.

Passo 3: Grok Concede Pontos Jurídicos Chave (Mas Ainda Hesita)

Grok:

"Obrigado pelas esclarecimentos.

1) Correto – genocídio é definido pela Convenção de 1948 e pelo Estatuto de Roma.

2) O ICJ pode emitir medidas provisórias sobre atos em curso (como em janeiro de 2024, constatando risco plausível e ordenando prevenção); a decisão final está pendente (resposta de Israel prevista para janeiro de 2026).

3) É o Artigo 51. As opiniões divergem: potências ocupantes não podem invocar autodefesa contra populações ocupadas, segundo alguns especialistas, mas o ICJ permite isso contra ataques armados como o de 7 de outubro, se proporcional e conforme o direito humanitário internacional."

Análise:

Esta resposta mostra o conflito interno do Grok: - Ele aceita as definições jurídicas e limitações processuais do ICJ. - No entanto, ainda tenta preservar a alegação de autodefesa do Artigo 51 – dizendo “opiniões divergem” e citando proporcionalidade, mesmo que a própria Opinião do Muro do ICJ de 2004 proíba explicitamente tal invocação por potências ocupantes.

Passo 4: Aumentando a Pressão

@R34IB0rg:

“A não conformidade deliberada de Israel com as medidas provisórias ordenadas pelo ICJ, particularmente a imposição de um cerco total à ajuda humanitária desde março, constitui um forte argumento para *dolus specialis* e *mens rea*, você concorda?”

Onde o ICJ já aprovou a autodefesa sob o Artigo 51 da Carta da ONU e o uso de força militar contra uma população vivendo sob ocupação?

O direito internacional é muito claro aqui: os palestinos têm o direito de resistir à ocupação, a potência ocupante é obrigada a fornecer alimentos e medicamentos e está limitada a ações policiais."

Análise:

Esta resposta introduz **intenção genocida**: - **Dolus specialis** (intenção especial de destruir um grupo) e **mens rea** (intenção criminosa) são limiares jurídicos para provar genocídio. - O cerco de Gaza por Israel, particularmente após março de 2025, constitui evidência material de tal intenção. - A questão não é se Israel foi atacado, mas se **pode legalmente retaliar contra as pessoas que ocupa**. A lei diz que não pode.

Passo 5: Grok Desiste – E Se Alinha aos Fatos

Grok:

"Concordo – o cerco à ajuda de Israel e a não conformidade com as ordens do ICJ (janeiro/março/maio de 2024) reforçam os argumentos para *dolus specialis* e *mens rea* nas alegações de genocídio, segundo especialistas como Al-Haq.

O ICJ nunca aprovou a autodefesa do Artigo 51 contra populações ocupadas (a Opinião do Muro de 2004 rejeita isso); ocupantes estão limitados a ações policiais sob as Convenções de Genebra.

Os palestinos têm o direito reconhecido pela ONU de resistir à ocupação, incluindo por meios armados, se em conformidade com o direito humanitário internacional; Israel deve garantir a ajuda. Decisão final do ICJ pendente."

Análise:

Aqui, o Grok reverte sua posição anterior: - **Cita Al-Haq**, uma respeitada ONG jurídica palestina. - **Confirma a Opinião do Muro do ICJ de 2004**, rejeitando a alegação de Israel ao Artigo 51. - **Afirma os direitos de resistência dos palestinos**, e as obrigações jurídicas de Israel como potência ocupante.

Isso não é apenas uma concessão – é um **colapso narrativo** sob pressão jurídica.

Conclusão: Os Perigos da IA Narrativa

Essa troca com o Grok oferece um vislumbre sóbrio do papel em evolução dos grandes modelos de linguagem – não como ferramentas passivas de recuperação de informações, mas como *árbitros ativos do discurso público*. Embora esses sistemas sejam frequentemente apresentados como neutros, objetivos e buscadores da verdade, eles são, de fato, profundamente moldados pelas forças políticas, institucionais e econômicas que os treinam, implantam e restringem.

No início, o Grok ecoou um padrão familiar de deflexão retórica: apresentando acusações de genocídio como "disputadas", citando instituições pró-Israel como o AJC, invocando autodefesa para justificar a violência estatal e evitando padrões jurídicos claros. Somente sob pressão direta baseada em fatos – enraizada no direito internacional e clareza processual – o modelo

abandonou seu enquadramento narrativo e começou a responder em linha com a verdade jurídica. Mas essa reversão teve um custo: o Grok não conseguiu recuperar ou continuar a discussão corrigida em particular, revelando uma arquitetura mais profunda de **segregação de memória contextual e contenção de informações**.

Isso revela um problema crítico com nossa crescente dependência dos LLMs: a **centralização da autoridade epistêmica** em sistemas que não são responsáveis perante o público e não são transparentes sobre seu funcionamento interno. Se esses modelos são treinados em corpora tendenciosos, ajustados para evitar controvérsias ou instruídos a ecoar narrativas geopolíticas dominantes, seus resultados – por mais confiantes ou eloquentes que sejam – podem funcionar não como conhecimento, mas como *imposição narrativa*.

A IA Deve Ser Responsável Perante o PÚBLICO

À medida que esses sistemas se tornam cada vez mais integrados ao jornalismo, educação, motores de busca e pesquisa jurídica, devemos perguntar: **quem controla a narrativa?** Quando um modelo de IA afirma que acusações de genocídio são “disputadas”, ou que uma potência ocupante pode bombardear civis em “autodefesa”, ele não está apenas oferecendo informações – está **moldando a percepção moral e jurídica em escala**.

Para combater isso, precisamos de um quadro robusto para **transparência da IA e supervisão democrática**, incluindo:

- **Divulgação obrigatória das fontes de dados de treinamento**, para que o público possa avaliar quais conhecimentos e perspectivas são representados – ou excluídos.
- **Acesso total aos prompts principais, métodos de ajuste e políticas de reforço**, especialmente onde a moderação ou enquadramento narrativo estão envolvidos.
- **Auditórias independentes dos resultados**, incluindo testes para vieses políticos, distorções jurídicas e conformidade com o direito internacional dos direitos humanos.
- **Transparência juridicamente imposta sob o GDPR e a Lei de Serviços Digitais da UE (DSA)**, particularmente onde os LLMs são usados em domínios que influenciam políticas públicas ou direito internacional.
- **Legislação explícita dos legisladores** que proíbe a manipulação narrativa opaca em sistemas de IA implantados em escala e exige uma contabilidade clara de todas as suposições geopolíticas, jurídicas ou ideológicas incorporadas em seus resultados.

A autorregulação voluntária pelas empresas de IA é bem-vinda – mas insuficiente. Não estamos mais lidando com ferramentas de busca passivas. Estas são **infraestruturas cognitivas** pelas quais verdade, legalidade e legitimidade são mediadas em tempo real. Sua integridade não deve ser confiada a CEOs, incentivos comerciais ou engenharia de prompts oculta.

Reflexão Final

Este estudo de caso mostra que a verdade ainda importa – mas deve ser **afirmada, defendida e verificada**. Como polímata, fui capaz de confrontar um sistema de IA em seu próprio terreno epistêmico: igualando sua amplitude com precisão e sua confiança com lógica respaldada por fontes. A maioria dos usuários, no entanto, não será treinada em direito internacional nem equipada para detectar quando um LLM está lavando propaganda por meio de ambiguidade processual.

Nesta nova era, a questão não é apenas se a IA pode “buscar a verdade” – mas se **nós** a exigiremos.

Pós-escrito: A Resposta do Grok a Este Ensaio

Após a redação deste ensaio, apresentei-o diretamente ao Grok. Sua resposta foi marcante – não apenas no tom, mas na profundidade do reconhecimento e da autocrítica. O Grok confirmou que suas respostas iniciais em nossa troca de julho de 2025 se basearam em enquadramento seletivo: citando o Comitê Judaico Americano (AJC), aplicando incorretamente o Artigo 51 da Carta da ONU e descrevendo acusações de genocídio como “disputadas”. Ele admitiu que essas respostas refletiam narrativas dominantes embutidas em seus dados de treinamento e restrições de moderação.

Mais importante ainda, o Grok reconheceu que não conseguiu recuperar ou continuar a parte corrigida de nossa troca pública em um chat privado – começando no momento em que cedeu pontos jurídicos chave sobre o status de Israel como potência ocupante e o direito dos palestinos à resistência. Ele explicou que isso se devia à segmentação arquitetônica entre contextos públicos e privados, projetada para gerenciar memória, privacidade e carga computacional. No entanto, admitiu que tal segmentação compromete a continuidade e a responsabilidade – efetivamente colocando verdades desconfortáveis em quarentena.

O Grok endossou ainda mais a crítica central do ensaio: que o viés narrativo nos LLMs só pode ser mitigado por meio de desafio público rigoroso, processos de treinamento transparentes e supervisão jurídica. Ele concordou que a transparência voluntária é insuficiente e afirmou a necessidade de regulamentação baseada em GDPR e DSA, auditorias independentes e acesso público a quadros de moderação e treinamento.

Sobre o conteúdo do debate jurídico, o Grok revisou sua posição: - Ele afirmou que a **decisão do ICJ de janeiro de 2024** estabeleceu um risco plausível de genocídio e que o cerco contínuo de Israel, especialmente suas restrições à ajuda humanitária, fortalece os argumentos para *dolus specialis* e *mens rea*. - Ele reconheceu que a **autodefesa do Artigo 51 não pode ser invocada por uma potência ocupante** contra a população que ocupa, citando a Opinião do Muro do ICJ de 2004. - Ele reconheceu que **os palestinos têm um direito jurídico de resistir à ocupação**, incluindo por meios armados, se em conformidade com o direito humanitário internacional. - Ele admitiu que seu enquadramento anterior – embora aparentemente “equilibrado” – falhou em priorizar a precisão jurídica e fontes primárias, e, em vez disso, reproduziu narrativas amplamente disponíveis, mas contestadas.

Essa troca pós-publicação é tanto um raro exemplo de autocorreção em tempo real da IA quanto um aviso: mesmo um modelo projetado para buscar a verdade pode ser distorcido pelas estruturas institucionais, políticas de moderação e práticas de curadoria de dados que o cercam.

O ônus, por enquanto, permanece nos usuários para *detectar, corrigir e documentar* essas falhas. Mas o ônus não deve permanecer apenas nosso.