Quando o nome vira a verdade

Quando o nome vira a verdade

Relatório técnico sobre um erro de catalogação em banco de imagens, e o que ele demonstra sobre epistemicídio computacional

Autora: Nina da Hora Cientista da computação (PUC-Rio), mestre em Ciência da Computação pelo Instituto de Computação da Unicamp, fundadora do Instituto da Hora

Documento: relatório técnico de trabalho, versão 1 **Data:** julho de 2026 **Destinação:** documento de trabalho, base para desdobramento em artigo empírico **Nível:** técnico, escrito para leitores de fora da computação. As passagens formais estão sempre acompanhadas de tradução em linguagem corrente, e podem ser puladas sem perda do argumento.

1. OBJETIVO E ESCOPO

Este relatório formaliza um evento concreto e mostra que ele não é acidente operacional, e sim a manifestação, numa camada específica, de um mecanismo que venho descrevendo desde a dissertação sob o nome de epistemicídio computacional.

O evento tem uma propriedade rara para quem estuda sistemas de identificação: nele não participa nenhum modelo de aprendizado de máquina. Não há rede neural, não há dataset de treino, não há gradiente. Existe apenas um banco de imagens, um campo de metadado e uma cadeia editorial humana. Isso permite fazer algo que raramente é possível, que é isolar o mecanismo do seu substrato técnico e verificar se ele sobrevive à ausência de inteligência artificial. Ele sobrevive. É esse o resultado que sustenta o documento inteiro.

O escopo é deliberadamente estreito. Não trato aqui de intenção, de responsabilidade individual nem de conduta profissional. Trato da estrutura formal do sistema que produziu o erro, das razões pelas quais ele não podia ser detectado por nenhuma verificação existente, e das razões pelas quais a correção aplicada não corrige.

2. O EPISÓDIO

Em 26 de julho de 2026, um caderno comemorativo de circulação nacional publicou um perfil sobre mim, dentro de uma lista de cem e uma pessoas apresentadas como construtoras do futuro. A fotografia que ilustra o perfil não é minha. É de outra mulher negra, que não foi identificada em lugar nenhum da publicação e que, até onde sei, nunca foi consultada.

O texto está correto. A legenda me nomeia. O crédito fotográfico traz o nome do fotógrafo e a data, e está correto para aquela fotografia. O erro saiu simultaneamente na edição impressa e na versão digital.

Ao ser avisada, a redação substituiu a imagem no site em poucas horas, sem publicar nota de correção. A edição impressa já estava em circulação. A explicação recebida foi que a fotografia estava catalogada no banco de imagens sob o meu nome.

Essa última frase é o objeto deste relatório. Ela descreve, com precisão técnica e aparentemente sem perceber, exatamente o mecanismo que eu estudo.

3. NOTAÇÃO: O ACERVO COMO ESTRUTURA DE DADOS

Para poder raciocinar sobre o sistema, é preciso escrevê-lo. A formalização abaixo é mínima e não depende de nenhuma tecnologia específica. Vale para um banco digital moderno, para um arquivo de negativos com envelopes etiquetados ou para uma pasta de recortes.

Seja `I` o conjunto de imagens do acervo e `P` o conjunto de pessoas que podem ser retratadas. Definem-se duas funções sobre esse conjunto:

```
d : I → P     denotação, quem de fato está retratado na imagem
ℓ : I → P     rotulagem, quem o metadado do banco afirma que está retratado
```

A distinção entre as duas é o eixo do documento. A função `d` é um fato do mundo. A função `ℓ` é um registro produzido por alguém, em algum momento, e gravado numa tabela. São objetos de natureza diferente, e o acervo armazena apenas o segundo.

O índice do banco é a relação materializada `A = {(i, ℓ(i)) : i ∈ I}`. A busca por nome devolve o conjunto

```
R(p) = { i ∈ I : ℓ(i) = p }
```

isto é, todas as imagens cujo metadado corresponde ao nome consultado. Diz-se que a rotulagem é fiel na imagem `i` quando `ℓ(i) = d(i)`.

Em linguagem corrente: o banco não guarda quem está na foto. Guarda quem alguém escreveu que está na foto. Todo o resto do sistema conversa com o que foi escrito.

4. A CADEIA EDITORIAL, ETAPA POR ETAPA

O episódio pode ser descrito como uma composição de operações, cada uma com sua especificação. Vale percorrer todas, porque a conclusão é contraintuitiva.

A repórter consulta o banco pelo meu nome. O sistema devolve `R(p)`, o conjunto de imagens etiquetadas com aquele nome. A imagem publicada, chame-se `i*`, pertence a esse conjunto, porque `ℓ(i*) = p`. A recuperação fez o que deveria fazer.

A diagramação seleciona `i*` e gera a legenda. A legenda exibe o nome associado à imagem no banco, que é `ℓ(i*) = p`. A legendagem fez o que deveria fazer.

O crédito fotográfico exibe autoria e data do arquivo, informação que acompanha a imagem e que está correta para aquela fotografia. O crédito fez o que deveria fazer.

A impressão reproduz a página diagramada. A gráfica fez o que deveria fazer.

Chega-se assim ao primeiro resultado.

> **Proposição 1.** Todas as etapas da cadeia satisfizeram suas especificações. A publicação da imagem incorreta não viola nenhuma pós-condição existente no sistema.

A demonstração é imediata. A única condição capaz de detectar o erro é `ℓ(i*) = d(i*)`, isto é, a afirmação de que a etiqueta corresponde à pessoa realmente retratada. Essa condição não é pós-condição de nenhuma etapa da cadeia, porque `d` não é um campo do banco. Não existe coluna para ela, não existe consulta que a devolva, não existe rotina que a compare com `ℓ`. O sistema é internamente consistente com a rotulagem e estruturalmente cego à denotação.

A consequência prática merece ser dita com todas as letras. Um sistema formalmente verificado, com cobertura de teste completa e revisão editorial impecável, publicaria esse mesmo erro sem emitir nenhum sinal. Não porque alguém foi descuidado, mas porque o predicado que falha não pertence ao vocabulário do sistema. Não existe teste para uma condição que o sistema não sabe enunciar.

Trata-se, portanto, de erro de especificação, e não de erro de execução. Essa distinção importa porque muda inteiramente o tipo de correção que faz sentido pedir. Pedir mais cuidado a uma redação que já foi cuidadosa não altera nada. O que falta não é diligência, é um campo.

5. O DESLOCAMENTO DA AUTORIDADE

Depois que a imagem é indexada, tudo o que acontece a jusante consome `ℓ`. A busca consome `ℓ`, a legenda consome `ℓ`, o arquivo consome `ℓ`. A função `d` nunca mais é consultada por ninguém.

Dizer isso em linguagem técnica é dizer que `d` se torna epistemicamente inerte. O rosto real da pessoa perde poder causal dentro do sistema. Ele existe no mundo, mas deixou de ser aquilo que determina o que o sistema afirma sobre ela.

É aqui que a formalização encontra o marco teórico. Sueli Carneiro descreve o epistemicídio não apenas como destruição de conhecimentos, mas como sequestro da condição de sujeito do conhecimento, ou seja, como a retirada da pessoa do lugar de quem pode saber e afirmar. O que a estrutura acima descreve é exatamente isso, escrito como substituição de função. A pergunta “quem é essa pessoa” deixou de ser respondida pela pessoa e passou a ser respondida por um registro que ela não escreveu, que não pode consultar e sobre o qual não tem permissão de escrita.

Não se trata de uma analogia. É a mesma operação, num suporte diferente.

6. IRREVERSIBILIDADE

O resultado mais forte do documento é este, e ele explica por que a correção aplicada não repara o sistema.

Trate a identidade como variável aleatória. Seja `D` a identidade real da pessoa retratada e `N` o nome gravado no banco. A cadeia forma um processo markoviano, o que significa apenas que cada etapa recebe informação exclusivamente da etapa anterior:

```
D → N → legenda → página impressa → acervo
```

A desigualdade do processamento de dados, resultado clássico de teoria da informação, afirma que nenhuma transformação aplicada a um sinal pode aumentar a informação que ele carrega sobre a fonte original. Aplicada aqui:

```
I(D; acervo) ≤ I(D; página) ≤ I(D; legenda) ≤ I(D; N)
```

> **Proposição 2.** Nenhum processamento posterior à etiquetagem pode recuperar informação sobre a identidade real que tenha sido perdida na etiquetagem.

A tradução em linguagem corrente é simples e severa. Se a informação foi destruída no momento em que alguém escreveu o nome errado, ela não está mais em lugar nenhum da cadeia. Revisar não recupera. Diagramar melhor não recupera. Imprimir com mais qualidade não recupera. Todas essas operações estão à direita da desigualdade, isto é, todas trabalham sobre um sinal que já perdeu o que se quer resgatar.

Disso decorrem três consequências verificáveis no caso concreto.

A primeira é que revisão editorial não é reparo, e sim mais processamento de `N`. Uma redação pode revisar indefinidamente sem nunca tocar no problema.

A segunda é que a substituição da imagem no site não reparou o sistema, reparou uma cópia. A operação realizada foi atribuir o valor correto a uma instância renderizada, deixando intacto o par `(i*, p)` dentro do índice `A`. Como a origem permanece inalterada, a probabilidade de recorrência permanece inalterada. Em termos de estrutura de dados, corrigiu-se uma folha e não a raiz. Qualquer nova consulta ao banco pelo mesmo nome pode devolver a mesma imagem outra vez.

A terceira é que o suporte impresso funciona como ponto fixo. Sobre papel não existe operação de escrita depois da tiragem. O valor de ℓ registrado ali é imutável por construção. É por isso que o exemplar comprado tem estatuto probatório diferente de qualquer captura de tela: não é a melhor prova por ser mais nítida, é a única prova cujo estado não pode ser silenciosamente revisado.

7. O COROLÁRIO DO ARQUIVO

Há um efeito adicional, e é o que dá ao caso relevância para além dele mesmo.

Quando a versão digital é corrigida sem nota, o par ordenado formado pelo erro e pela sua correção deixa de existir no registro público. O sistema converge para um estado em que a informação mútua entre o erro e o registro é nula. O evento não fica apenas não reparado. Ele fica não observável.

Isso tem uma implicação metodológica que será retomada na seção 12. Um pesquisador que queira medir a frequência desse tipo de erro só consegue contar os casos que foram corrigidos publicamente. Correções silenciosas são, por definição, invisíveis ao levantamento. O mecanismo, ao se corrigir da maneira mais barata, apaga o rastro de si mesmo e reduz a base empírica de quem tenta estudá-lo.

E tem uma implicação política imediata. Se a única cópia imutável do erro está com a pessoa lesada, então a preservação da prova depende inteiramente de que ela tenha comprado o jornal. A assimetria é evidente: a instituição controla o registro oficial e o sujeito controla, na melhor das hipóteses, um exemplar. Quem controla o arquivo controla o que aconteceu.

8. A QUANTIDADE CENTRAL: RESOLUÇÃO DE IDENTIDADE POR GRUPO

Chega-se agora à definição que unifica as camadas e que proponho como núcleo formal do conceito.

Seja φ um mapa qualquer que resolva identidade, isto é, qualquer procedimento que receba uma pessoa ou uma imagem e produza algo que sirva para distinguir aquela pessoa das demais. O ponto importante é que φ não precisa ser uma rede neural. Pode ser o sistema perceptual de um revisor humano olhando um rosto. Pode ser a chave de indexação de um banco de imagens. Pode ser um encoder que produz vetores.

Defina a relação de indistinguibilidade sob tolerância ε:

```
x ~ε y   se e somente se   ‖φ(x) − φ(y)‖ < ε
```

Essa relação agrupa em uma mesma célula todos os elementos que o mapa não consegue separar. Defina então a resolução de identidade no grupo `g` como o número esperado de identidades distintas que caem dentro de uma mesma célula:

```
R_g(ε) = E[ número de identidades distintas contidas numa célula ]
```

Quando `R_g(ε) = 1`, o mapa é injetivo naquele grupo, isto é, pessoas diferentes ocupam células diferentes e o sistema consegue distingui-las. Quando `R_g(ε) > 1`, o mapa colapsa pessoas distintas na mesma célula, e o sistema não tem resolução suficiente para separá-las.

Com isso, a definição:

> **Epistemicídio computacional.** Um sistema de identificação exerce epistemicídio computacional quando a partição induzida pelo seu mapa de resolução de identidade é sistematicamente mais grosseira em uns grupos do que em outros, ou seja, quando `R_g(ε) > R_{g'}(ε)` de modo estruturado segundo marcadores sociais.

O que está sendo afirmado não é que o sistema erra mais em certos grupos. É algo mais específico e mais grave: dentro daqueles grupos, o sistema deixa de ser injetivo. Ele para de distinguir sujeitos. Duas pessoas passam a ocupar o mesmo lugar na estrutura, e qualquer uma delas pode ser devolvida quando se pergunta pela outra.

9. AS TRÊS CAMADAS

O que transforma isso em conceito, e não em metáfora, é que a mesma quantidade se instancia em camadas materialmente distintas, com implementações que não têm nada em comum entre si.

Na camada perceptual, φ é o reconhecimento facial humano do revisor. `R_g > 1` significa que ele não discrimina rostos daquele grupo com resolução suficiente, e portanto a verificação não dispara mesmo quando a imagem está trocada.

Na camada vetorial, φ é o encoder de uma rede neural que mapeia rostos em vetores. `R_g > 1` significa que identidades daquele grupo caem dentro da mesma bola de raio ε, e a busca por vizinho mais próximo devolve um resultado essencialmente arbitrário entre elas. É esse o objeto da segunda frente metodológica da dissertação, com FaceNet e ArcFace.

Na camada arquivística, φ é a própria chave de catalogação ℓ. `R_g > 1` significa que identidades daquele grupo colapsam sob o mesmo nome no índice, que é literalmente o que a redação descreveu ao dizer que a imagem estava catalogada com o meu nome.

O caso do jornal é a instância arquivística pura, e o seu valor está justamente aí. Ele demonstra que o mecanismo não depende de rede neural, de dataset nem de otimização. O que se repete de camada em camada é a estrutura, não a tecnologia. Quando o mecanismo reaparece dentro de um sistema de aprendizado de máquina, ele não foi criado ali. Foi herdado.

10. FORMULAÇÃO INFORMACIONAL

Existe uma maneira mais compacta de dizer a mesma coisa, útil para comunicação pública.

Condicione a identidade real ao nome consultado e observe a entropia residual:

```
H(D | N = p) > 0
```

Entropia, aqui, é a medida da incerteza que sobra. Quando ela é nula, saber o nome equivale a saber quem é a pessoa. Quando é positiva, saber o nome não basta, porque mais de uma pessoa continua compatível com aquele registro.

O epistemicídio computacional, nessa formulação, é a afirmação de que essa entropia residual é positiva para uns nomes e nula para outros, de forma correlacionada com raça e gênero. Dito de maneira direta, e é assim que eu recomendo dizer em público:

> Meu nome, dentro daquele banco, é um canal de banda mais estreita que o de outra pessoa. Consultar por ele devolve um conjunto, não um indivíduo.

11. CONTINUIDADE COM A CAMADA DE APRENDIZADO

Esta seção conecta o episódio com o eixo empírico da dissertação. Ela não é necessária para o argumento arquivístico, e o leitor não técnico pode passar direto para a seção 12. Está aqui porque mostra que o mesmo funcional atravessa as camadas, que é a tese central do trabalho.

11.1 O apagamento pela agregação

Treinar um modelo é minimizar uma perda média sobre a população amostrada. Escrevendo por grupo:

```
L(θ) = Σ_g π_g · L_g(θ)
∂L/∂θ = Σ_g π_g · ∂L_g/∂θ
```

O gradiente, que é a direção em que os parâmetros se movem a cada passo, é uma média ponderada pela massa amostral π_g de cada grupo. Sob capacidade finita, o modelo aloca representação proporcionalmente a essa massa.

O ponto que costuma escapar é que nenhuma intenção discriminatória é necessária para produzir o resultado. O apagamento está no operador de agregação, e o operador de agregação é a definição de aprender. Não existe versão do treinamento supervisionado que não faça isso.

11.2 Detecção é um evento de limiar

Detecção não é classificação. Detecção é uma função indicadora, que responde apenas se algo existe ou não:

```
existe(x) = 1[ s(x) > τ ]
recall_g   = P(s > τ | g) = 1 − F_g(τ)
```

A sensibilidade do recall a uma perturbação δ no score é a densidade avaliada no limiar:

```
d(recall_g)/dδ = f_g(τ)
```

O que essa derivada diz é que a fragilidade de um grupo a mudança de iluminação, de pose, de ângulo ou de compressão é literalmente a quantidade de massa que aquele grupo tem acumulada em cima do limiar. E o limiar `τ` é escolhido otimizando um critério global, o que significa, na prática, que o ponto de operação da distribuição majoritária é aplicado a todo mundo.

A diferença entre falhar na classificação e falhar na detecção também precisa ser dita para público amplo. Falhar na classificação é ser confundida com outra pessoa. Falhar na detecção é não estar lá. A primeira é erro de rótulo. A segunda é ausência de existência dentro do pipeline, e é por isso que a detecção é o caso paradigmático.

11.3 A camada geométrica

Depois de detectado, o rosto é convertido em vetor. A pergunta muda de “quem é legível” para “como quem foi detectado é representado”. A auditoria, nesse ponto, usa distância de Wasserstein entre medidas empíricas por grupo, transporte ótimo regularizado por entropia e avaliação da coerência geométrica das categorias demográficas no espaço de embeddings.

A pergunta que organiza essa frente é se as categorias usadas em auditorias de equidade correspondem a estruturas geometricamente coerentes no espaço vetorial, ou se são impostas de fora sobre uma geometria que não as reflete.

11.4 As quatro dimensões auditáveis

O conceito só se torna operacional quando vira estimador. As quatro dimensões que uso são a invisibilidade diferencial, medida como a distância entre o melhor e o pior recall por grupo; o gradiente de confiança, medido como distância entre as distribuições de score condicionais, verificando se o sistema detecta com menos certeza mesmo quando detecta; a falha interseccional, avaliada em células produto e não em marginais, porque disparidade pode desaparecer na marginal e reaparecer no cruzamento; e a consistência cross-arquitetural, que verifica se o ordenamento de grupos se mantém ao trocar a família de modelo.

A última é a que distingue artefato de estrutura. Se o padrão sobrevive à troca de CNN por transformer, a causa não está na arquitetura.

11.5 A escolha de norma

Vale registrar uma observação que costuma passar despercebida em papers. Otimizar acurácia agregada equivale a minimizar uma norma `L1` ponderada sobre o vetor de erros por grupo. Otimizar pelo pior grupo equivale a minimizar a norma do máximo. Toda a disputa entre critério utilitário e critério maximin, na prática de engenharia, se reduz a uma escolha de norma. Ela quase nunca é declarada no texto, porque vem embutida no valor padrão da biblioteca que se usou.

12. POR QUE O ERRO SOBREVIVE DE FORMA DESIGUAL

Esta seção transforma o episódio em hipótese empírica, e é a parte que pretendo desenvolver como artigo.

Erros de catalogação são gerados e depois filtrados. Em algum ponto entre o banco e a banca de jornal, alguém pode perceber que a imagem não corresponde ao nome. Modelando:

```
P(erro publicado | g) = P(erro gerado | g) · (1 − P(captura | g))
```

A captura ocorre quando a similaridade percebida entre o rosto na imagem e a pessoa esperada fica abaixo do limiar perceptual de quem revisa. Ou seja, a probabilidade de captura depende diretamente de `R_g` na camada perceptual, definida na seção 8.

Existe evidência experimental robusta sobre isso. O efeito de outra raça em reconhecimento facial humano está entre os achados mais replicados da psicologia experimental, com meta-análise consolidada desde o início dos anos 2000: a acurácia de discriminação de rostos de fora do próprio grupo racial é sistematicamente menor. Traduzido para a notação deste relatório, isso é `R_g > 1` na camada perceptual, para grupos fora do grupo majoritário da redação.

> **Proposição 3.** Ainda que a taxa de geração de erros seja uniforme entre grupos, a taxa de sobrevivência não é. A distribuição dos erros que efetivamente chegam ao público é enviesada por um filtro de verificação, mesmo na ausência de viés na origem.

Esse resultado tem três virtudes. Explica o fenômeno sem precisar atribuir má-fé a ninguém, o que o torna discutível com as próprias redações. É falsificável. E o dado necessário para testá-lo existe e nunca foi sistematicamente coletado no Brasil.

12.1 Desenho de teste

A hipótese a testar é que a proporção de correções e erratas de imagem referentes a pessoas negras é desproporcional à sua presença na cobertura.

A unidade amostral é a nota de correção de imagem publicada por veículo de circulação nacional, num intervalo definido. As variáveis a codificar são o grupo racial da pessoa retratada e da pessoa nomeada, o tipo de erro (troca de pessoa, erro de legenda, erro de crédito), a latência entre publicação e correção e o suporte em que a correção saiu, distinguindo digital de impresso.

O denominador é indispensável e é a parte que costuma ser feita errado. É preciso estimar a frequência de aparição por grupo no corpus de matérias do mesmo período. Sem esse denominador o resultado é ininterpretável, porque grupos sub-representados aparecem menos e portanto acumulam menos erros em termos absolutos. A quantidade de interesse é a razão entre taxa de erro por grupo e taxa de aparição por grupo, nunca a contagem bruta.

O viés de observação é a parte metodologicamente mais interessante, e conecta com a seção 7. Só entram na amostra os erros que foram corrigidos publicamente. Correções silenciosas, como a que ocorreu no caso que motiva este relatório, são invisíveis ao desenho. Logo qualquer estimativa produzida é um limite inferior. O mecanismo que se pretende medir suprime parte da evidência de si mesmo, e o desenho amostral herda a censura que está tentando quantificar. Isso não invalida o estudo, mas precisa ser declarado, e o valor estimado precisa ser lido como piso.

13. POR QUE OS REPAROS USUAIS NÃO REPARAM

Três abordagens correntes seriam propostas por qualquer equipe técnica diante desse problema. Nenhuma resolve, e as razões são distintas em cada caso.

A primeira é o reparo por transporte ótimo, que constrói um mapa `T` tal que a distribuição de scores de um grupo passe a coincidir com uma distribuição de referência. O obstáculo é estrutural: o push-forward preserva medida, isto é, `T` é essencialmente uma bijeção entre distribuições.

> **Proposição 4.** Não existe transporte que restaure injetividade perdida. Se φ colapsou duas identidades na mesma célula, a informação foi destruída dentro do encoder e não deslocada dentro da distribuição. Equalizar marginais deixa `R_g` inalterado.

O resultado é que transporte ótimo corrige o gradiente de confiança da seção 11.2 e não toca no colapso da seção 8. Esse é um limite teórico do reparo por calibração, e me parece um resultado publicável por si só.

A segunda abordagem é impor critérios de paridade estatística, do tipo paridade demográfica ou chances equalizadas. Esses critérios são condições sobre distribuições marginais de saída. Duas identidades colapsadas satisfazem qualquer um deles trivialmente, porque a saída agregada não distingue quem é quem. Um sistema pode estar perfeitamente equalizado segundo todas as métricas de fairness disponíveis e continuar incapaz de me separar de outra pessoa.

A terceira é aumentar a quantidade de dados do grupo sub-representado. Isso reduz a variância das estimativas e melhora a alocação de capacidade, mas não impõe injetividade. Injetividade é uma restrição sobre o mapa, e portanto precisa entrar na função objetivo. Não entra pela amostra.

14. REPARAÇÃO, EM TERMOS FORMAIS

O que eu chamo de reparação computacional não é etiquetar melhor. É uma mudança na estrutura do sistema, e cada requisito abaixo deriva de uma das proposições anteriores.

O primeiro requisito é tornar a denotação consultável, introduzindo no esquema do banco um campo de verificação de identidade com procedência registrada e responsável nomeado. Isso deriva da Proposição 1: enquanto `d` não estiver no vocabulário do sistema, o predicado que falha não é sequer expressável, e nenhuma quantidade de rigor produz detecção.

O segundo é dar ao sujeito autoridade de escrita sobre as linhas do índice que o nomeiam. Autodeclaração acima de anotação por terceiros. Isso deriva da seção 5, e é exatamente o mesmo critério que defendo na avaliação de datasets de detecção facial, onde a anotação demográfica feita por terceiros nega à pessoa a possibilidade de se declarar. Mesmo princípio, dois sistemas.

O terceiro é corrigir na origem e propagar, alterando o par `(i*, p)` no índice e não apenas nas cópias renderizadas. Isso deriva da Proposição 2. Sem ele, a probabilidade de recorrência permanece exatamente onde estava.

O quarto é documentar o erro em vez de apagá-lo, mantendo a correção anexada ao registro arquivado em lugar de substituir silenciosamente o estado anterior. Isso deriva do corolário da seção 7. Sem esse requisito, o sistema converge para um estado em que o

erro é não observável, e a série histórica necessária para estudar o fenômeno é destruída pelo próprio procedimento de correção.

15. GLOSSÁRIO DE NOTAÇÃO

`d(i)` é quem está de fato retratado na imagem `i`, e `ℓ(i)` é quem o banco afirma que está. `A` é o índice materializado, o conjunto dos pares imagem e nome. `R(p)` é o conjunto recuperado pela consulta ao nome `p`. `φ` é o mapa que resolve identidade, em qualquer camada, e `ε` é a tolerância abaixo da qual duas coisas são tratadas como indistinguíveis. `R_g(ε)` é a resolução de identidade no grupo `g`, isto é, quantas pessoas distintas o sistema mistura numa mesma célula. `π_g` é a massa amostral do grupo, `τ` é o limiar de detecção, `F_g` e `f_g` são a distribuição e a densidade do score condicionadas ao grupo. `I(X; Y)` é informação mútua, e `H(D | N)` é a entropia da identidade real dado o nome, ou seja, a incerteza que sobra depois de saber o nome.

16. REFERÊNCIAS

Trabalho próprio

da Hora, Nina. *Algoritmos, poder e subjetividade: epistemicídio computacional no reconhecimento facial*. Dissertação de mestrado, Instituto de Computação, Universidade Estadual de Campinas, defendida em 11 de maio de 2026, aprovada por unanimidade. Orientação de Sandra Avila, coorientação de Marisol Marini, banca composta por Virgílio Almeida (UFMG) e Angela Figueiredo (UFRB). Dados de pesquisa depositados no REDU com DOI. Fonte primária da definição de epistemicídio computacional, do framework de auditoria em quatro dimensões e das duas frentes metodológicas, detecção e geometria de embeddings.

da Hora, Nina. Resultados da dissertação aceitos para apresentação na ACM FAccT 2026, Atenas.

da Hora, Nina. *Imagens que a máquina lê, mundos que ela apaga: notas sobre epistemicídio computacional*. South Feminist Futures Knowledge Hub. Ensaio em que proponho a tipologia dos três regimes da imagem algoritmizada, de treino, operacionais e de plataforma, e a distinção entre imagem socializada e imagem algoritmizada. Base direta das seções 8 e 9 deste relatório.

da Hora, Nina. Texto sobre epistemicídio computacional e violência digital contra meninas e mulheres negras, publicado pelo Geledés. Desenvolve o argumento de que não ser contada é não ser protegida, retomado aqui na seção 12.

Instituto da Hora. Pesquisa e defesa de direitos digitais com abordagem antirracista. Contexto institucional do trabalho.

Base conceitual

Carneiro, Sueli. Sobre epistemicídio e o sequestro da condição de sujeito do conhecimento. Base da seção 5.

Santos, Boaventura de Sousa. Epistemologias do Sul.

Gonzalez, Lélia. Amefricanidade e crítica à fixidez das taxonomias raciais.

Figueiredo, Angela. Epistemologia insubmissa, e a crítica à anotação demográfica feita por terceiros.

Imagem e visualidade técnica

Farocki, Harun. Imagens operacionais, imagens feitas por máquinas para máquinas.

Beiguelman, Giselle. *Políticas da imagem*, sobre imagem, memória e vigilância na dadosfera.

Browne, Simone. *Dark Matters*, sobre vigilância racializada.

Crítica técnica de sistemas

Buolamwini, Joy; Gebru, Timnit. *Gender Shades*, sobre erro diferencial em classificação facial.

Benjamin, Ruha. *Race After Technology*, sobre design discriminatório.

Formalismo

Cover, Thomas; Thomas, Joy. *Elements of Information Theory*. Desigualdade do processamento de dados, usada na Proposição 2.

Villani, Cédric. *Optimal Transport: Old and New*. Preservação de medida sob push-forward, usada na Proposição 4.

Hardt, Moritz; Price, Eric; Srebro, Nathan. Equality of opportunity in supervised learning. Contraste com critérios de paridade, seção 13.

Sagawa, Shiori et al. Distributionally robust optimization por grupo. Formulação maximin, seção 11.5.

Meissner, Christian; Brigham, John. Meta-análise do efeito de outra raça em reconhecimento facial humano. Base empírica da Proposição 3.

17. OBSERVAÇÃO DE ENCERRAMENTO

O que dá força a este caso não é o erro. É a explicação do erro.

Uma instituição que responde “a imagem estava catalogada com o seu nome” está descrevendo, com precisão e sem perceber, uma arquitetura em que a resposta à pergunta sobre quem é uma pessoa foi transferida da pessoa para o banco. É esse deslocamento, e não a troca de fotografia, que o conceito nomeia.

E é por isso que o caso interessa a quem trabalha com sistemas automatizados. O mecanismo apareceu aqui na sua forma mais simples, sem modelo, sem dado de treino, sem otimização. Quando ele reaparece dentro de uma rede neural, o que se está vendo não é uma patologia nova da inteligência artificial. É uma estrutura antiga sendo executada mais rápido.