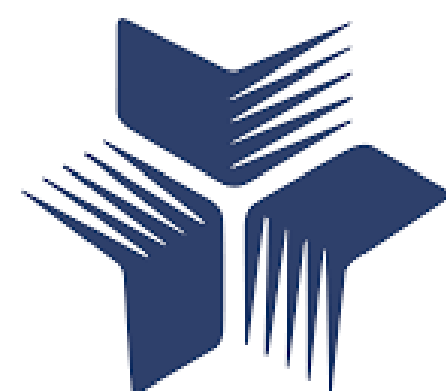


# Audiência Pública: Direitos Autorais e o PL 2338/23

Luca Schirru, *DSc., MSc., LL.M.*



**ibict**

Instituto Brasileiro de Informação  
em Ciência e Tecnologia



CENTRE FOR IT & IP LAW

# Mineração de Textos e Dados

**Art. 63. Não constitui ofensa aos direitos de autor e conexos a utilização automatizada de conteúdos protegidos em processos de mineração de textos e dados para os fins de pesquisa e desenvolvimento de sistemas de IA por organizações e instituições científicas, de pesquisa e educacionais, museus, arquivos públicos e bibliotecas, desde que observadas as seguintes condições: (e o jornalismo?)**

**I – o acesso tenha se dado de forma lícita;**

**II – não tenha fins comerciais;**

**III – a utilização de conteúdos protegidos por direitos de autor e conexos seja feita na medida necessária para o objetivo a ser alcançado, sem prejuízo dos interesses econômicos dos titulares e sem concorrência com a exploração normal das obras e conteúdos protegidos. [...]**

# TDM é sinônimo de treinamento de IA (Gen)?



**Projetos que não envolvem IA (Gen)**  
**Universidades/ Jornalistas**  
**Finalidades: pesquisa; educação; jornalismo**  
**Resultado esperado: padrões e correlações**  
**Sem riscos de substituição**

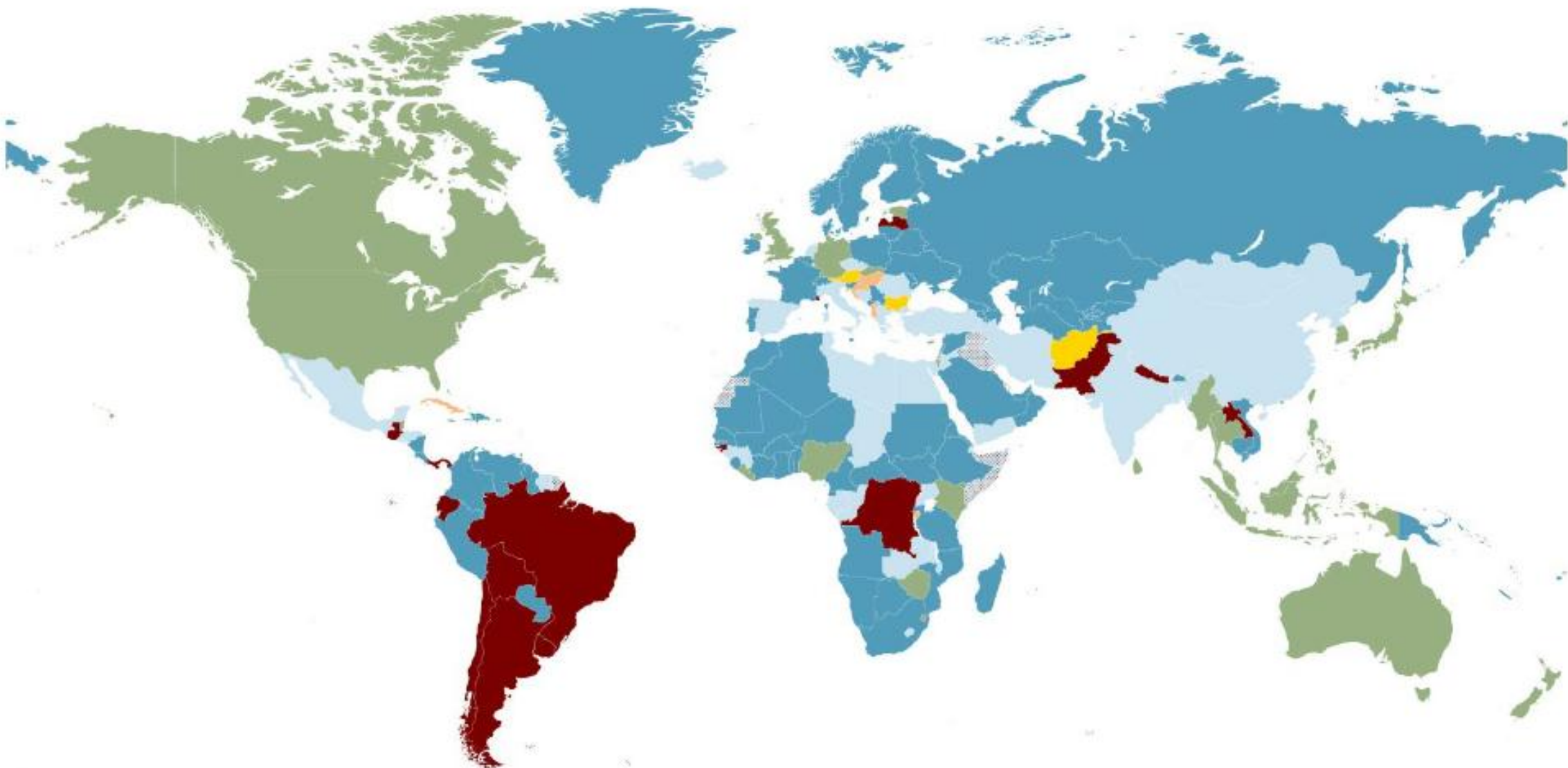
**Treinamento de IA Generativa**  
**Grandes empresas de tecnologia**  
**Finalidade comercial**  
**Resultado esperado: mesma natureza dos dados de treinamento**  
**Potencial substituição**

**TDM e Treinamento de  
sistemas de IA (Gen)  
envolvem (múltiplas)  
reproduções.**

(Carroll, 2019, 895; Gervais, 2024; Bartz v. Anthropic PBC,  
3:24-cv-05417, (N.D. Cal. Jun 23, 2025) ECF No. 231)

**Toda** reprodução deve ser  
autorizada ou remunerada?

Research exceptions in copyright laws around the world



- **Open research exception:** permits all research uses (including sharing) of all works by all users
- **Restrictions on sharing:** permit reproductions of whole works for research but do not extend to communications with others for purposes such as collaboration or validation
- **Restrictions to private reproduction:** provide a research exception only for use by individuals for a “private” or “personal” use
- **TDM restricted:** confine research exceptions to short excerpts, effectively prohibiting TDM uses
- Not mapped

| GII rank | Economy                      | Score |
|----------|------------------------------|-------|
| 1        | Switzerland                  | 67.5  |
| 2        | Sweden                       | 64.5  |
| 3        | United States of America     | 62.4  |
| 4        | Singapore                    | 61.2  |
| 5        | United Kingdom               | 61.0  |
| 6        | Republic of Korea            | 60.9  |
| 7        | Finland                      | 59.4  |
| 8        | Netherlands (Kingdom of the) | 58.8  |
| 9        | Germany                      | 58.1  |
| 10       | Denmark                      | 57.1  |
| 11       | China                        | 56.3  |
| 12       | France                       | 55.4  |

Number of highly cited publications in top 100 by select geographic areas, 2021–23

Source: AI Index, 2025 | Chart: 2025 AI Index report

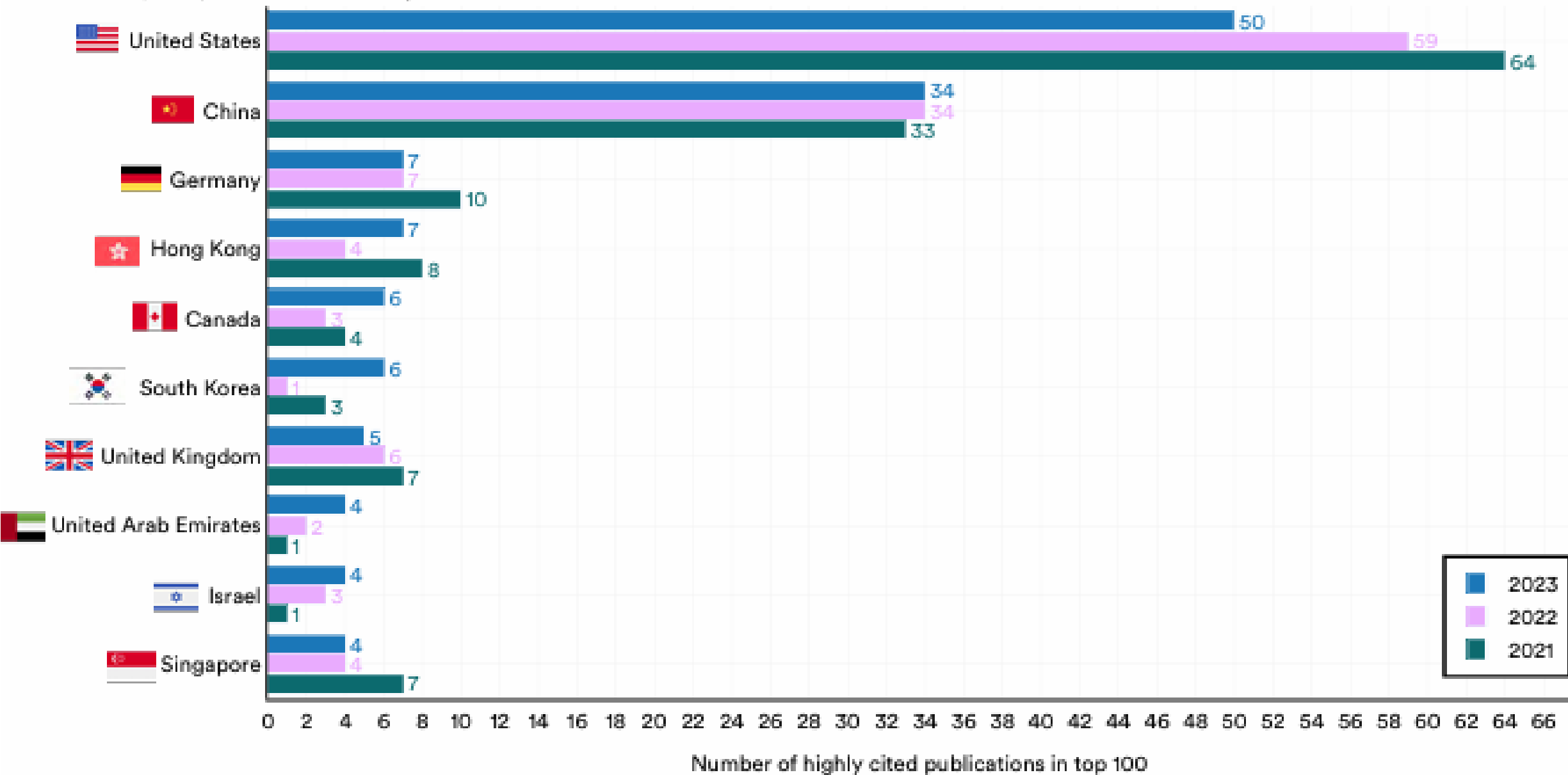


Figure 1.1.11

Lutando contra a desinformação



**Handke, Guibault and Vallbé (2021): “countries in which academic researchers must acquire the express consent of rights holders to conduct lawful DM exhibit a lower share of DM research output in their total research output [...] This implies that an application of copyright exceptions or limitations that establish the right to mine for academic researchers—if they have lawful access to input works and irrespective of explicit rights holder consent—boosts DM research.”**

**Soberania informacional e digital:** o Projeto **SAVITS (Sistema de Avaliação e Validação de Impactos em Tecnologias Sociais)** busca reduzir a dependência tecnológica externa e tem como **objetivo central prover ao Estado brasileiro uma ferramenta de apoio à decisão humana, transformando dados em ativos estratégicos para a otimização de políticas públicas e a redução de desigualdades sociais;**

**Infraestrutura de inteligência analítica:** o sistema integra Ciência de Dados, Aprendizado de Máquina (ML) e Grandes Modelos de Linguagem (LLM) para validar os impactos de investimentos em Tecnologias Sociais no âmbito do Distrito Federal;

**Treinamento com dados estratégicos:** a IA poderá ser treinada com políticas públicas, textos de patentes, artigos acadêmicos, e projetos de ciência cidadã, por exemplo.

# Remuneração

**Art. 65. O agente de IA que utilizar conteúdos protegidos por direitos de autor e conexos em processos de mineração, treinamento ou desenvolvimento de sistemas de IA deve remunerar os titulares desses conteúdos em virtude dessa utilização, devendo-se assegurar:**

**I – que os titulares de direitos de autor e conexos tenham condições efetivas de negociar coletivamente, nos termos do Título VI da Lei nº 9.610, de 19 de fevereiro de 1998 (Lei dos Direitos Autorais), ou diretamente a utilização dos conteúdos dos quais são titulares, podendo fazê-lo de forma gratuita ou onerosa; [...]**

**“A mesma tecnologia que concorre comigo (ou está me substituindo) foi treinada utilizando minhas obras, e eu não autorizei isso nem recebi nenhuma remuneração por tal uso.”**

**“A obrigação de remunerar vai sufocar as nossas startups e comprometer o desenvolvimento da indústria nacional de IA.”**

**A proteção a autores, artistas, e demais titulares e a pesquisa e inovação não precisam ser tratadas como escolhas excludentes.**

**Regimes diferenciados (PL da IA, AIA, DSM, DSA, DMA)**

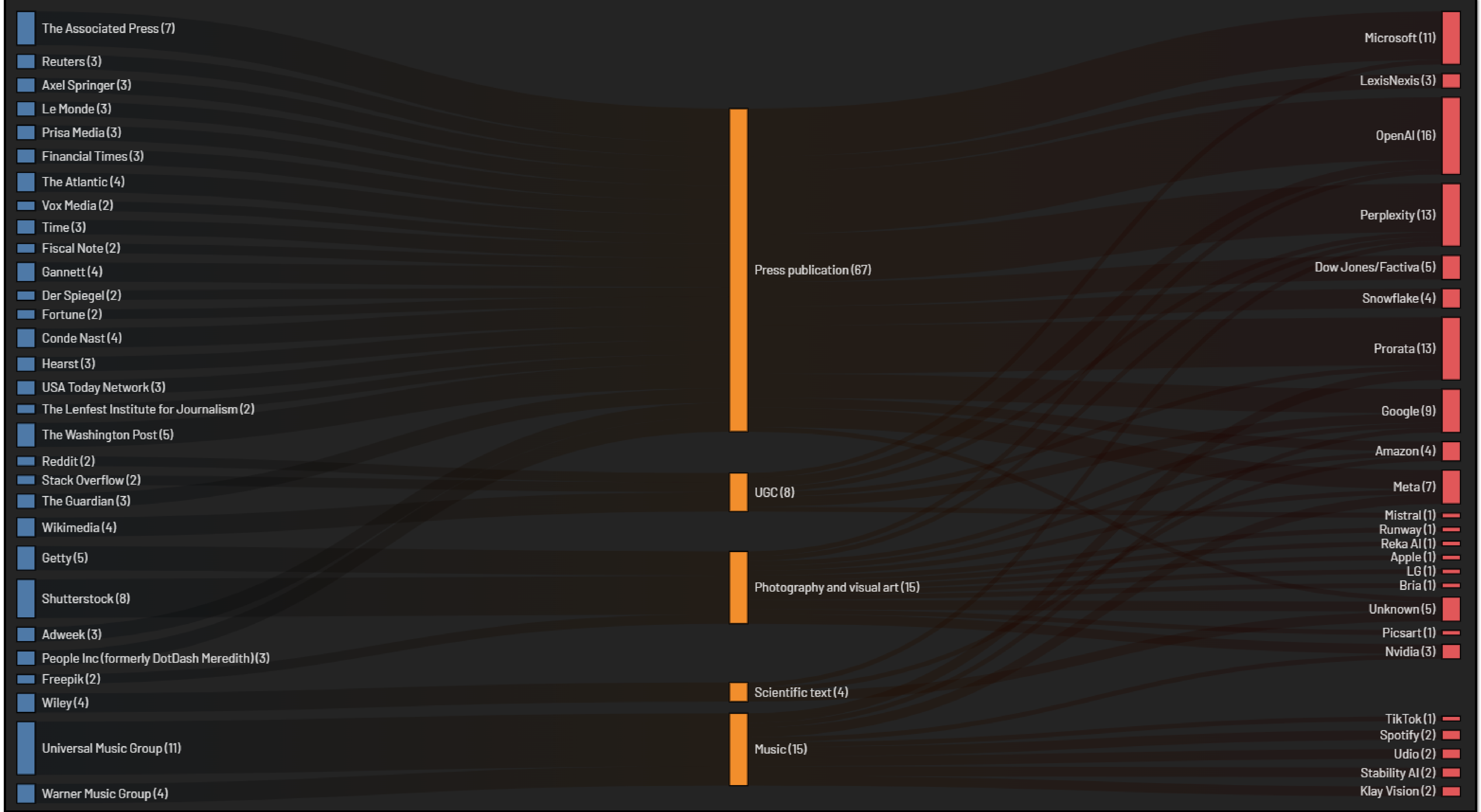
**IA x IA Generativa**

**Diferentes modelos de remuneração**

# Modelos de remuneração:

- **Licenças diretas;**
- **licenças coletivas;**
- **licenças coletivas estendidas;**
- **licenças estatutárias;**
- ***levies*;**
- **tributos.**

*PL da IA (2023); Senftleben (2023); Projeto de lei espanhol (2024); Craig (2024); Geiger & Iaia (2024); Jacques & Flynn (2024); Pasquale & Sun (2024); IBDAutoral (2024); Jenkins (2024); Wang et al (2024); Durantaye (2025); Dermawan & Mezei (2024); Moraitis (2024); Schinzler (2025); Zhang (2025); Chen et al (2025); Keller (2025); Lucchi (2025); Thomas & Kretschmer (2025); Schirru (2025; 2026)*



## E o ecossistema da informação?

- "Our content is free, our infrastructure is not" (Wikimedia, 2025)
- Percepção de aumento de tráfego de 90%. Quase 70% dos respondentes atribuem o aumento ao acesso por bots (GLAM Elab).
- 52% dos respondentes passaram a usar chatbots e TikTok como fonte de informação em vez de buscadores como o Google (The Verge).
- Diminuição de acesso à fonte original quando um AI summary é disponibilizado (Seer Interactive).
- "The data Commons is rapidly shrinking" (Stanford, 2025)

## Limitações dos modelos baseados em direito autoral

- Deslocamento do trabalho criativo x uso de obras protegidas;
- Remuneração baseada no *input*: treinamento no exterior; dados sintéticos e destilação; transparência e distribuição;
- O licenciamento é insuficiente para alcançar todas as partes afetadas;
- Assimetria de poder nos contratos de licenciamento;
- Custos de transação;
- O licenciamento coletivo e as entidades de gestão coletiva não representam todos os autores;
- Enquanto empresas de IA generativa e grandes detentores de conteúdo celebram licenças, muitos autores ficam de fora.

# Retorno para a Sociedade e para os Criadores\*



## Por que este modelo pode ajudar

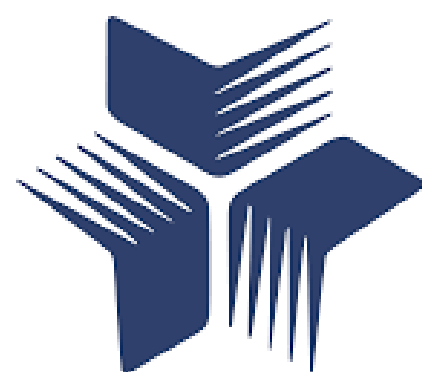
- Foca em algo que o direito autoral não alcança: o impacto sistêmico no trabalho criativo, o incentivo à criação e a soberania;
- Alcança os criadores como um todo (a longa cauda, os inéditos, os contribuidores dos commons);
- Sustentabilidade dos commons e da criação humana como um todo;
- Retorno em dois níveis: à sociedade e aos commons, e ao criador humano;
- Capacidade de IA soberana: converte extração em capacidade doméstica (a pilha de IA do Brasil);
- Modelos existentes interessantes;
- Desacoplamento dos direitos autorais: pode avançar independentemente de como as questões autorais sejam resolvidas.

Sem regulação que sustente as plataformas abertas, incentive novos autores e mantenha o ecossistema criativo, não há *commons* nem criação humana. **E sem criação humana, não há IA Generativa.**

# Obrigado!

[schirru@schirru.adv.br](mailto:schirru@schirru.adv.br)

CENTRE ON  
KNOWLEDGE  
GOVERNANCE



**ibict**

Instituto Brasileiro de Informação  
em Ciência e Tecnologia

**KU LEUVEN**

**CiTiP**

CENTRE FOR IT & IP LAW