



Open Discussion #3

VENICE
2025

Theme

**Beyond the Hype:
Confronting AI's Real-World Limitations**

IARIA Congress 2025 & DigiTech 2025



Stalling Progress?

VENICE
2025

LLM Research & Development Seems to Hit a Wall



Except for incremental improvements...

Issue 1: Algorithms

Issue 2: Return on Invest

Issue 3: Training Data

Issue n?



Erik Buchmann
*Center for Scalable
Data Analytics and
Artificial Intelligence
Dresden/Leipzig,
Germany*



Issue 1: Algorithms

VENICE
2025

- LLMs **stop reasoning** and make up results, **if the problem is too difficult**
 - For example, see this paper
- LLMs **degrade over time**
 - CHEN, Lingjiao; ZAHARIA, Matei; ZOU, James. How is ChatGPT's behavior changing over time?. *Harvard Data Science Review*, 2024, 6. Jg., Nr. 2.
 - ChenghaoZhu, et al. Is Your LLM Outdated? A Deep Look at Temporal Generalization. In: *Conference of the Association for Computational Linguistics*, 2025
- LLMs do **not benefit from scaling-up training**
 - SPRINGER, Jacob Mitchell, et al. Overtrained language models are harder to fine-tune. *arXiv preprint arXiv:2503.19206*, 2025.
- LLMs **collapse with minor changes**
 - MA, Ziyang, et al. Model Hemorrhage and the Robustness Limits of



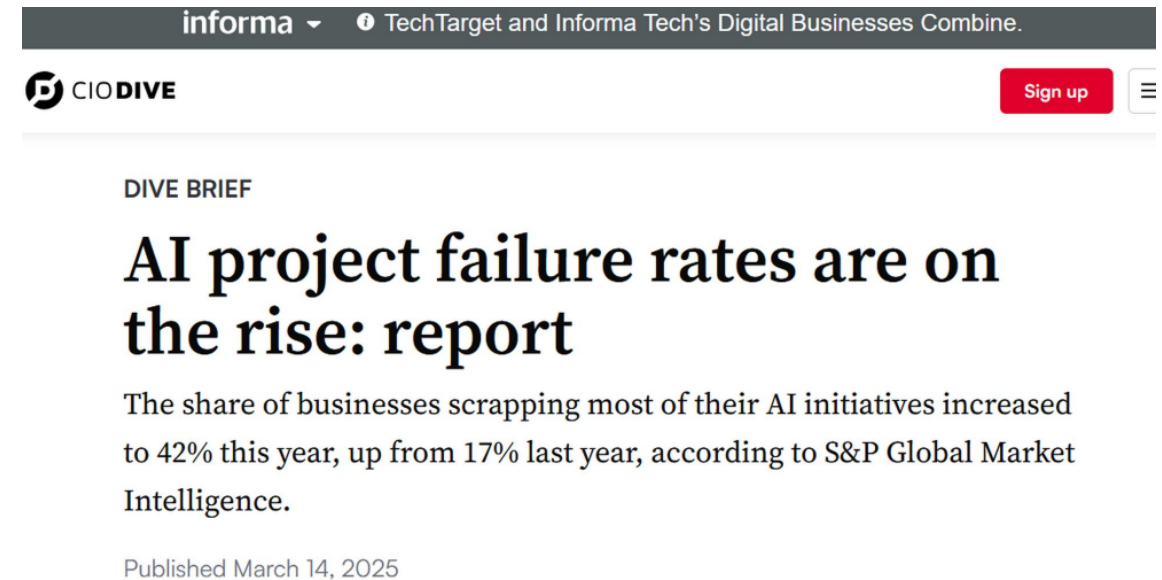


Issue 2: Return on Invest

VENICE
2025

- Many business reports in the media read like this
 - "failure": technically **working**, but do not **achieve business goals**
- Many services **bundle AI** with their products, **without options for removal**
 - Microsoft (Copilot in Office 365, Win11)
 - Google (Search, Gmail, Docs, Sheets)
 - Meta (Facebook, Instagram, WhatsApp)
 - ...

*Remember the Microsoft Antitrust Case,
when IE 4.0 was bundled with Win98?*





Issue 3: Training Data

VENICE
2025

- It seems that all available high-quality **training data has been used up**
 - *Elon Musk says all human data for AI training 'exhausted'. The Guardian, Jan. 2025*
- Companies **pushing ethical or data privacy boundaries** to use copyrighted material or customer data for AI training
- AI models **collapse, when trained with AI generated data**
 - *SHUMAILOV, Ilia, et al. AI models collapse when trained on recursively generated data. Nature, 2024, 631. Jg., Nr. 8022*

Meta AI: How Facebook & Instagram Data Powers AI Training

**Meta is training AI
Using Your Data**

heyData



Arthur

24.04.2025



Open Discussion #2

VENICE
2025

- 1. Which other "difficult" **limitations for current AI** systems exist?
 - *(Particularly, when using LLMs for high-stakes problems)*
- 2. Which alternatives exist to **circumvent such limitations**?
 - *(What are the new answers to problems that cannot be solved incrementally?)*