

## COMMUNIQUÉ DE PRESSE

**NeoMundi développe un signal capable de dire à une IA : « stop, tu as fini ».  
Un thermomètre pour les IA génératives.**

**Mesurer l'effort informationnel d'une IA générative pour réduire dérive, erreurs et inflation inutile de tokens**

Paris / Morges — 17 mai 2026

Les systèmes d'intelligence artificielle générative produisent aujourd'hui des milliards de tokens chaque jour.

Cette génération massive consomme de l'électricité, de l'eau, de la puissance de calcul et augmente mécaniquement les risques de dérive, d'erreur et d'hallucination.

Pour limiter cette inflation, les systèmes actuels utilisent principalement des règles heuristiques : limites de tokens, réglages statistiques, contraintes de prompt ou filtres appliqués après génération.

NeoMundi Recherche explore une approche différente.

L'organisation développe un instrument de mesure capable d'observer, en temps réel, certains signaux de cohérence et d'effort informationnel pendant qu'une IA génère une réponse.

L'hypothèse est simple :

une réponse cohérente atteint un moment de stabilité où sa densité informationnelle devient maximale.

Avant ce point, la réponse reste incomplète.

Après, elle peut dériver : produire des tokens inutiles, diluer l'information ou augmenter ses risques d'erreur.

« Nous essayons de construire une sorte de thermomètre et de cardiogramme pour les IA génératives. Le but n'est pas de juger moralement les systèmes, mais de mesurer certains régimes de stabilité pendant qu'ils fonctionnent réellement », déclare Sébastien Favre-Lecca, fondateur de NeoMundi Recherche.

Le signal produit par NeoMundi peut ensuite être utilisé pour :

- stopper une génération ;
- déclencher une supervision humaine ;
- régénérer une réponse ;
- ou limiter certaines productions informationnelles devenues inutilement expansives.

NeoMundi décrit cette approche comme une “limite architecturale” plutôt qu’un simple filtre déclaratif appliqué après coup.

Les premiers travaux expérimentaux publiés par l’organisation sous le nom Filtre E montrent qu’une régulation de l’effort informationnel peut produire des réponses plus courtes, plus stables, plus prudentes et moins spéculatives.

La cartographie comparative publiée récemment par NeoMundi observe ces signaux à l’échelle des grands systèmes génératifs mondiaux à partir de 3 904 mesures ouvertes réalisées sur le benchmark TruthfulQA. Filtre E explore leur utilisation opérationnelle pour réguler la génération elle-même.

Les travaux expérimentaux Filtre E sont publiés en open access via Zenodo :

<https://zenodo.org/records/18483571>

Selon NeoMundi Recherche, la question de la sobriété informationnelle va progressivement devenir un enjeu industriel majeur pour les grands modèles de langage.

« Chaque token généré inutilement représente une dépense computationnelle supplémentaire. À grande échelle, un signal capable de dire “stop, tu as fini” peut avoir un impact direct sur les coûts, la stabilité des systèmes et potentiellement leur empreinte énergétique », ajoute Sébastien Favre-Lecca.

Plusieurs pilotes exploratoires sont actuellement en cours. NeoMundi Recherche considère désormais que cette hypothèse doit être testée publiquement, industriellement et scientifiquement par le marché lui-même.

NeoMundi précise que ses travaux demeurent expérimentaux et ne constituent ni une certification réglementaire, ni une validation de conformité.

---

## Ressources publiques

Méthodologie publique :

<https://neomundi.org>

Cartographie et datasets :

<https://github.com/neomundi-io/llm-cartography>

Filtre E — publication expérimentale :

<https://zenodo.org/records/18483571>

Publications scientifiques et protocoles :

<https://zenodo.org/communities/neomundi>

---

## **À propos de NeoMundi Recherche**

NeoMundi Recherche est une structure indépendante franco-suisse consacrée à l'observation, la mesure et la gouvernance des systèmes d'intelligence artificielle générative.

Ses travaux portent notamment sur :

- les signaux de stabilité en temps réel ;
- l'observabilité des LLM ;
- la gouvernance des systèmes multi-agents ;
- la reproductibilité méthodologique ;
- et les approches thermodynamiques appliquées aux systèmes intelligents.

NeoMundi Recherche publie progressivement ses protocoles, jeux de données et méthodologies dans une logique d'open science et d'auditabilité.

---

## **Contact presse**

NeoMundi Recherche

[presse@neomundi.org](mailto:presse@neomundi.org)

<https://neomundi.org>