

Usages pédagogiques des systèmes générateurs de texte

Philippe Dessus, LaRAC, LIG, &
Inspé, Univ. Grenoble Alpes

Présentation à EFELIA-MIAI, LIG,
UGA, le 6 mars 2024



0. But de la présentation

- Partir du problème principal induit par l'usage pédagogique des systèmes générateurs de texte (dorénavant SGT), pour en dériver des problèmes alternatifs, tout aussi importants (si ce n'est plus)
- Sans revenir, ni sur les aspects techniques, ni sur leur fonctionnement, ni sur les avancées récentes, ... ni sur les aspects légaux ou environnementaux !

0. Le problème de départ (Luo 2024)

À l'université et à l'école, les usages des SGT compromettent *l'intégrité académique* et sont une *mauvaise conduite* : le contenu généré n'est pas l'œuvre originale de l'étudiant.e

- Formulé dans beaucoup de textes officiels universitaires (15 sur les universités du top 20 mondial enquêtées par Luo)
- Cette formulation induit plutôt des pratiques de suspicion, de "détection", et d'interdiction des usages

0. Questionner le problème

- Formuler le problème, c'est déjà prendre un certain parti, s'engager dans une certaine vue...
- Il peut y avoir d'autres manières de le présenter (l'arbre qui cache la forêt)

0. Représentations alternatives du problème (Luo 2024)

- **Conception de l'évaluation** : Les tâches d'évaluation devraient être reconsidérées
- **Limites** : Les SGT produisent de l'information frauduleuse et biaisée (**botshit**)
- **Équité** : L'usage des SGT amène problèmes d'équité et de justice
- **(In)formation** : Les enseignants et les étudiants devraient être formés
- **Politique** : Il manque une ligne de conduite claire des institutions

⚠ ce dernier point ne sera pas traité aujourd'hui

0. Plan de la présentation

0. Présentation et problème de départ
 1. Quelques définitions
 2. Conceptions de l'évaluation et SGT
 3. Les limites et biais des SGT (*botshit*)
 4. Former les enseignants
 5. L'équité et la justice dans l'utilisation des SGT
 6. Comment parler des SGT ?
 7. Discussion
 8. Références

0. Présentation sur ma page personnelle

- <http://pdessus.fr>
- en cliquant sur le premier item de la liste en haut à droite

1 Quelques définitions

1. Quelques définitions

- **Système générateur de texte (SGT)** : Système informatique générant du texte (e.g., résumés, explications, récritures) à la demande de prompts et fondés sur des grands modèles de langage (*large language models*)
- **Travail original** : Production écrite et/ou orale d'étudiant.e faite à des fins évaluatives, et produit par l'auteur.e
- **Intégrité académique** : "L'engagement des étudiants, du corps enseignant et du personnel à faire preuve d'un comportement honnête et moral dans leur vie universitaire" (ICAI 2021, cité par Coghlan et al. 2021)

1. Quelques définitions (2/2)

Usage d'un SGT

- Les frontières entre l'usage (ou non) d'un SGT ne sont pas claires
 - faire produire de 0 un document
 - utiliser un SGT pour avoir des idées (phase de brainstorming)
 - faire retoucher, polir un texte par un SGT
 - comme il y a des SGT partout, leur utilisation **non intentionnelle** va devenir de plus en plus fréquente (p. ex., des outils de correction de texte embarquent un SGT et peuvent être "détectés", voir [Leoffler 2024](#)))

1. Originale ? Vraiment ?

- Si on attend des étudiant.es qu'ils produisent des travaux originaux, on attend aussi qu'ils collaborent...
- Il y a un continuum dans la "sous-traitance" ou "l'externalisation" de travaux (Bretag et al. 2019), entre des comportements de partage et de tricherie

1. Où se place l'intégrité académique ?

- On indique souvent que détecter les productions créées par SGT permet de respecter l'intégrité académique
 - Accuser à tort un étudiant, est-ce intègre académiquement ? (**non**)
 - Utiliser des détecteurs de SGT dont le fonctionnement est opaque, le score non interprétable et les performances faibles, est-ce intègre académiquement ? (**non**)

2 Conceptions de l'évaluation : le licite et l'illicite

1. Comportements de sous-traitance ($N = 14$ k étudiants de 8 univ. d'Australie)

- Acheter, vendre ou échanger du matériel de cours (15 %)
- Donner un travail réalisé (pour quelque raison que ce soit) (27 %)
- Obtenir un travail réalisé de quelqu'un d'autre (pour le livrer à son nom propre) (2 %)
- Aider quelqu'un pendant un examen (3 %)
- Recevoir de l'aide de quelqu'un (2 %)
- Passer un examen à la place de quelqu'un d'autre (0,5 %)
- Demander à quelqu'un de passer un examen à sa place (0,2 %)

2. Les tâches évaluatives et leur relation à l'externalisation (1/2)

Les moins susceptibles d'externalisation (moins de 20 % de chances)

- Réflexion sur l'expérience de stage
- Soutenance orale
- Personnalisée et unique
- Réalisation en cours

 Bretag & Harper (2019)

2. Les tâches évaluatives et leur relation à l'externalisation (2/2)

Les plus susceptibles d'externalisation (de 20 à 55 % de chances)

- Évaluation ponctuelle (initiale)
- Évaluation ponctuelle (finale)
- Évaluations tout au long de l'enseignement
- Tâches à fort coefficient
- Délai de livraison court

 Bretag & Harper (2019)

2. Des tâches évaluatives plus authentiques ?

- Des tâches évaluatives plus authentiques ne sont pas liées à moins de tricherie...
- ...mais la tricherie est plus aisément détectable avec ce type d'évaluation, lorsqu'elle arrive

 Bretag & Harper (2019)

3 Les limites et biais des SGT

3. Les limites et biais des SGT : les botshits (Hannigan et al. 2023)

- Les SGT produisent des textes qui prédisent la suite de mots la plus probable, compte tenu - du prompt ; - du corpus d'entraînement ; - des étiquetages humains pour améliorer les réglages...
- La "botshit" c'est l'utilisation de ces textes par des humains, sans discernement

3. Quatre types de travail avec SGT (Hannigan et al. 2023)

Travail authentifié	Travail automatique
Soumettre des tâches et les vérifier méticuleusement leur exactitude. Ex. : tâches de budget, sécurité...	Soumettre des tâches routinières et utiliser leurs réponses pour une activité efficace. Ex. : traduire une réf. en BibTeX
Travail augmenté	Travail autonome
Utilisé pour générer des idées. Ex. : brainstorming	Déléguer des tâches où les SGT peuvent s'adapter. Ex. : tâches d'assistance

3. Les risques de chaque type de travail

- **Mauvaise calibration** pour le travail authentifié : sur/sous confiance envers les SGT
- **Routinisation** pour le travail automatisé : perdre de vue que les réponses proviennent du système
- **Ignorance** pour le travail augmenté : ne pas connaître les points positifs/négatifs de l'utilisation des SGT
- **Boîte noire** pour le travail autonome : ne pas savoir comment les SGT fonctionnent

4 L'équité et la justice dans l'utilisation des SGT

4. L'équité

- Variabilité des étudiant.es :
 - dans leur accès à des systèmes plus ou moins puissants
 - dans leur accès à des LLM plus ou moins entraînés et correspondant au domaine
 - dans leur type de travail
- Certain.es étudiant.es pourront se censurer dans leur utilisation des SGT pour ne pas risquer d'être "détectées"

4. La justice

1. L'avis des étudiant.es

- **53 %** des étudiant.es d'une université de GB ($N = 2\ 500$) pensent que c'est pertinent, juste que d'autres étudiants utilisent un SGT pour de l'aide grammaticale
- ils ne sont plus que **13 %** à trouver cela pertinent et juste pour rédiger une dissertation entière
- ... et **41 %** sont pour que les universités se dotent d'une politique d'utilisation des SGT spécifiant ce qui est licite ou pas

 Johnston et al. 2024

5 Former les enseignant.es (et les étudiant.es)

5. Former les enseignant.es (1/3)

- L'environnement universitaire joue un rôle dans les comportements de tricherie (Bretag et al. 2019)
- Les étudiant.es "tricheur.es" ont des scores moins importants que les non-tricheur.es dans les réponses aux questions :
 - comprendre les exigences de l'évaluation ; recevoir des rétroactions suffisantes ; pouvoir approcher les enseignants ; avoir des enseignements sur la pratique scientifique

5. Former les enseignant.es (2/3)

- De leur côté, enseignants et administrateur.es mentionnent que cette tricherie survient surtout lorsque :
 - les sanctions sont clémentes
 - les charges de travail sont inadéquates
 - le temps de contact est limité entre le personnel et les étudiants
 - les classes sont de grande taille et rendent difficile la connaissance des étudiants
 - l'environnement est centré sur l'examen des performances, de récompense et d'évaluation

 Bretag & Harper (2019)

5. Former les enseignant.es (3/3)

- L'arrivée des SGT a un impact sur les pratiques évaluatives des enseignants :
 - elle amplifie leur niveau de suspicion et de critique des productions
 - elle leur fait plus souvent percevoir certaines qualités des productions comme "non humaines", parfois à tort
 - elle les fait baisser leur notes

 Farazouli et al. (2023)

6 Comment parler des SGT ?

6. Comment parler des SGT ?

- On fait partie du problème (mais c'est impossible de ne pas y contribuer) en anthropomorphisant les processus et performances des SGT

6. Parler informatique sans recourir à des métaphores neuropsych ?

- C'est impossible...
- ...mais, à tout le moins, cela devrait être reconnu explicitement
- (à l'inverse, mais c'est une autre histoire, la neuropsych regorge de métaphores empruntées à l'informatique : architecture, canal, circuit, codage, filtrage, traitement de l'information) (Floridi & Nобре 2024)

6. Les métaphores neuropsychy dans l'informatique

- Dès le début, l'informatique utilise de nombreuses métaphores neuropsychologiques (Floridi & Nobre 2024) :
 - "*machine learning*", "vision informatique", "hallucination", "robot conversationnel"
 - les autres termes concurrents d'IA étaient moins tape-à-l'œil, moins vendeurs : "*automata studies*", "*complex information processing*"

6. Le problème de l'anthropomorphisation

- C'était déjà le cas avant, mais depuis les SGT, leurs usagers lui prêtent aisément des intentions ou des comportements humains
- Cela empêche de comprendre les processus et les produits des SGT, notamment leurs performances réelles

6. Un exemple

ChatGPT just can't resist pushing its own agenda. I asked for questions for a 7th grade unit on the solar system, the seasons, and the moon's phases, and it sneaks in a question about climate change:

1. How do the movements of Earth and the moon create the phases of the moon we observe from Earth?
2. What causes the seasons on Earth, and how are they related to the tilt of Earth's axis?
3. How does the position of Earth in its orbit around the sun affect the length of day and night?
4. Why do different planets in our solar system have different lengths of day and night and different seasons?
5. How do the characteristics of the moon's surface provide clues about its history and formation?
6. What role does gravity play in keeping planets and moons in orbit around the sun?
7. How do eclipses occur, and what can they teach us about the relationships between Earth, the moon, and the sun?
8. What are the potential effects of human activities on Earth's environment, and how can we protect our planet for future generations?

7 Discussion

7. Discussion

- L'usage des SGT à des fins de tricherie est un problème
- Mais il faut essayer de le formuler de manière non biaisée, en évaluant l'ensemble du problème et pas seulement les symptômes (la tricherie, la mauvaise conduite)
- Parvenir à garder un œil critique sur les processus et performances des SGT, en formant les enseignants et étudiants aux différents points ci-dessus (évaluation, botshit, équité), et en désanthropomorphisant leurs processus et performances

Pour finir

Why isn't AI doing the tedious shit for creative people instead of the creative shit for tedious people

[Isaac Io Schankler, 2024, source](#)

Merci de votre attention !

Des questions ?

✉ philippe.dessus[at]univ-grenoble-alpes.fr

8. Ressources proposées

- Utilisation de systèmes de génération de textes
- Ressources sur les systèmes de génération de textes

8. Références (1/3)

- Bin-Nashwan, S. A., Sadallah, M., & Bouteraa, M. (2023). Use of ChatGPT in academia: Academic integrity hangs in the balance. *Technology in Society*, 75. <https://doi.org/10.1016/j.techsoc.2023.102370>
- Bretag, T., & Harper, R. (2019). Contract cheating and assessment design: exploring the connection. https://ltr.edu.au/resources/SP16-5383_BretagandHarper_FinalReport_2019.pdf
- Bretag, T., Harper, R., Burton, M., Ellis, C., Newton, P., Rozenberg, P., Saddiqui, S., & van Haeringen, K. (2019). Contract cheating: a survey of Australian university students. *Studies in Higher Education*, 44(11), 1837-1856. <https://doi.org/10.1080/03075079.2018.1462788>
- Coghlan, S., Miller, T., & Paterson, J. (2021). Good Proctor or "Big Brother"? Ethics of

8. Références (2/3)

- Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K., & McGrath, C. (2023). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. *Assessment & Evaluation in Higher Education*, 1-13. <https://doi.org/10.1080/02602938.2023.2241676>
- Farzaneh, F., & Boyer, A. (2021). L'être humain face à l'IA : soumis ou dominant ? *Gestion 2000*, 38(2), 157-179. <https://doi.org/10.3917/g2000.382.0157>
- Floridi, L., & Nobre, A. C. (2024). Anthropomorphising machines and computerising minds: the crosswiring of languages between Artificial Intelligence and Brain & Cognitive Sciences. *SSRN Electronic Journal*.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4738331

8. Références (3/3)

- Hannigan, T. R., McCarthy, I. P., & Spicer, A. (2023). [Beware of botshit: How to manage the epistemic risks of generative chatbots](#). *SSRN Electronic Journal*.
- ICAI (2021). [The fundamental values of academic integrity](#) (3rd ed.).
- Johnston, H., et al. (2024). Student perspectives on the use of generative artificial intelligence technologies in higher education. *Int. J. Educ. Integrity*, 20(1). <https://doi.org/10.1007/s40979-024-00149-4>
- Leoffler, K. (2024). [Georgia college student used Grammarly, now she is on academic probation](#). Fox5 News.
- Lodge, J. M. et al. (2023). It's not like a calculator, so what is the relationship between learners and generative artificial intelligence? *Learning: Research and Practice*, 9(2), 117-124. <https://doi.org/10.1080/23735082.2023.2261106>