

<https://ses.enseigne.ac-lyon.fr/spip/spip.php?article242>



# L'IA et l'évaluation

- Enseigner - avec l'Intelligence Artificielle (l'IA) -



Date de mise en ligne : jeudi 20 juin 2024

---

Copyright © Sciences économiques et sociales - Tous droits réservés

---

Nous avons expérimenté des IA génératives pour construire des évaluations, évaluer les élèves ou leur permettre de s'autoévaluer. Cet article présente des exemples concrets d'utilisation de l'IA pour créer des supports d'évaluation en SES. Les IA génératives sont des outils puissants capables de produire des textes, des images ou des données chiffrées, et de générer rapidement des QCM adaptés aux besoins spécifiques des enseignants.

Cependant, l'utilisation de l'IA nécessite certaines précautions. Les productions générées requièrent notre expertise pédagogique pour être ajustées et vérifiées. Il est aussi essentiel de signaler clairement aux élèves lorsque des contenus fictifs sont utilisés. Enfin, bien que l'IA puisse créer des données fictives ou des scénarios hypothétiques avec efficacité, sa fiabilité pour fournir des données réelles et précises est limitée et doit être utilisée avec discernement.

Il est important de noter qu'il s'agit d'expérimentations en cours. Nous n'en sommes qu'au début et restons conscients des questions éthiques que cela soulève.

## Générer des textes pour des évaluations formatives ou sommative

Les textes servent fréquemment de support à des évaluations et les IA génératives sont de puissants outils pour les générer.

### Exemple 1 Générer un texte pour un EC3



L'intégration de textes générés par l'intelligence artificielle dans les dossiers documentaires doit rester exceptionnelle, par exemple pour compléter les articles scientifiques et les extraits de presse traditionnels. Si l'IA offre un gain de temps précieux aux enseignants et permet une personnalisation accrue des supports d'évaluation, nous devons garantir la qualité et la fiabilité des textes générés et les ajuster à nos attentes. Nous devons aussi clairement identifier des textes comme étant générés par une IA, avec mention du prompt utilisé et de la date de génération.

Dans cet exemple, l'IA a généré un texte (doc. 2) permettant d'expliquer pourquoi parmi les pratiquants réguliers d'une activité physique ou sportive les hommes participent 3 ou 4 fois plus à des compétitions que les femmes (doc.

1). N'ayant pas trouvé dans la littérature scientifique d'explication à ce constat, nous avons demandé à l'IA de générer un texte sociologique sur ce sujet. Plusieurs interactions ont été nécessaires avant d'aboutir à un texte qui nous convenait.

### Exemple 2 Produire des textes à trous

#### Exercice 1 Compléter le texte

Les entreprises qui réalisent une part importante de la production d'un pays le font pour dégager un \_\_\_\_\_ et ainsi s'enrichir grâce à la vente de leur \_\_\_\_\_. Cependant, toute n'y parviennent pas. En effet, pour produire un bien ou un service, une entreprise doit financer de nombreuses dépenses : les \_\_\_\_\_ de production. Pour qu'une entreprise s'enrichisse, il faut que son \_\_\_\_\_, qui correspond au total de l'argent encaissé grâce aux ventes réalisées soit supérieur aux \_\_\_\_\_.

Mais le bénéfice mesure davantage la \_\_\_\_\_ d'une entreprise que sa production. C'est la \_\_\_\_\_ qui mesure la valeur de la production réellement réalisée par l'entreprise. Elle s'obtient en déduisant du chiffre d'affaires la valeur de tous les biens et services achetées à d'autres entreprises (les \_\_\_\_\_). Elle correspond à la valeur que l'entreprise ajoute par son activité (grâce à la combinaison de \_\_\_\_\_ et de \_\_\_\_\_) à la valeur des consommations intermédiaires.

L'IA peut rapidement générer des textes dont on peut se servir pour construire des textes à trous. Précisons qu'il est rare que le texte généré par l'IA convienne d'emblée. Il faut souvent interagir avec l'IA et le modifier soi-même pour s'approcher du texte souhaité.

### Exemple 3 Générer des dictons

#### Exemple 3 Générer des dictons

Exercice 1 Dites à quel principe de gestion des risques correspond chacun des dictons généré par l'IA. Justifiez brièvement.

- a) « Le jardin bien entretenu ne connaît pas les mauvaises herbes »
- b) « Avec des graines variées, la moisson est toujours assurée. »
- c) « Mieux vaut mille épaulés pour lever la tempête. »

Nous avons demandé à l'IA de générer des dictons ou des proverbes pour chaque principe de gestion des risques (mutualisation, prévention, diversification). Nous avons retenu les plus pertinents pour évaluer les élèves. Cette évaluation fait suite à un exercice fait en classe où les élèves devaient classer de vrais proverbes dans un tableau à 3 colonnes, proverbes trouvés avec l'IA. A court de dictons réels pertinents, nous nous sommes résolus à en faire générer par l'IA. Nous avons aussi demandé aux élèves de produire leur proverbe sur un principe de leur choix.

### Exemple 4 Générer des "témoignages "

"Ma meilleure amie m'a invité à son mariage à l'étranger. Réaliser que je ne pouvais pas me permettre le voyage m'a brisé le cœur et montré à quel point je me suis isolé."

Quelle(s) expérience(s) du chômage reflète cette phase ?

- Isolement progressif
- Diminution des revenus et des possibilités de consommer
- Stigmatisation
- Sentiment de ne plus avoir d'utilité sociale
- Perte d'estime de soi

Vérier



Dans cet exemple, l'IA est utilisé pour générer des « témoignages » de chômeurs et les intégrer dans un QCM. Il convient de préciser la nature des "témoignages" en début d'exercice, par exemple en affichant le texte suivant : "Les témoignages qui apparaissent dans l'exercice sont fictifs et ont été créés dans le cadre d'un exercice pédagogique pour illustrer l'impact du chômage sur l'intégration sociale."

## Générer des images pour des évaluations

Les images peuvent servir de support à des évaluations ; les IA génératives sont de puissants outils pour les générer. Par rapport à une recherche classique d'images sur le web faire générer des images par l'IA présentes plusieurs avantages. D'une part cela permet d'obtenir des images proches de ce qu'on souhaite. D'autre part cela lève le problème du droit à l'image et des droits d'auteur des images sur le web.

### Exemple 1 Générer des images pour évaluer des concepts

A quel principe de gestion des risque correspond cette image ?  
Justifiez votre réponse.



Nous avons demandé à l'IA de générer des images abstraites pour illustrer chaque principe de gestion des risques. Nous avons retenu les images les plus pertinentes pour évaluer les élèves.

### Exemple 2 Générer des images pour évaluer des mécanismes



#### Exercice 1

A partir de cette image, et de vos connaissances, montrez que la socialisation est différenciée selon le milieu social.

Image générée par DALL-E le 7 novembre 2023

Nous avons utilisé l'IA pour évaluer des élèves de seconde sur la socialisation différenciée selon le milieu social en classe de seconde.

Dans un [autre article du site](#), à partir d'images d'activités sportives ou de loisirs générées par l'IA, on demande aux élèves d'identifier le milieu social d'un enfant.

## Générer des données chiffrées

Les données chiffrées sont souvent présentes dans des évaluations, sous différentes formes. L'IA peut faciliter leur élaboration et leur mise en forme.

### Exemple 1 Générer des données fictives

#### Exercice 2 Compléter le tableau

	Ménage 1	Ménage 2	Ménage 3	Ménage 4	Ménage 5
revenus du travail	2 500		2 900	3 100	
revenus du capital	300	200		500	350
revenus sociaux		300	100		250
<b>Total</b>	<b>3 000</b>	<b>3 200</b>	<b>3 400</b>	<b>3 750</b>	<b>3 900</b>
Prélèvements obligatoires	600		800		
<b>Revenu disponible</b>	<b>2 400</b>	<b>2 500</b>		<b>2 900</b>	<b>3 000</b>
Consommation	1 800	2 000	2 100		
épargne		500	500	700	500

Dans cet exemple, nous avons fourni à l'IA la structure du tableau et demandé de générer des données fictives

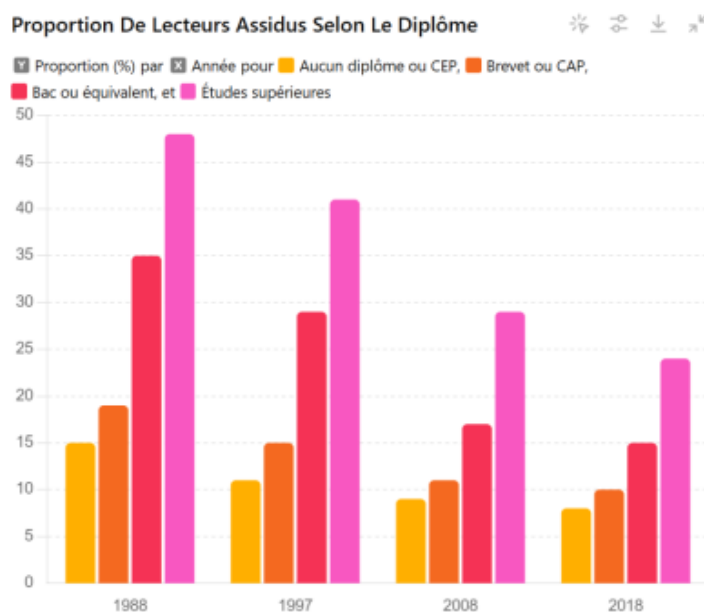
mensuelles en euros, en précisant que les montants devaient être des entiers et que le chiffre des unités devait toujours être égal à 0. L'IA a rempli le tableau correctement sans nécessiter de formules de calcul, ce qui nous a permis de gagner du temps. Cependant, nous avons nous-mêmes sélectionné les cellules vides dans le tableau, car l'IA commettait des erreurs, par exemple en supprimant à la fois les données de la consommation et de l'épargne, ce qui rendait l'exercice moins pertinent.

### Exemple 2 Attention à la génération de données réelles

Il est vivement déconseillé de recourir aux IA pour trouver des données économiques ou sociales réelles.

Les IA génératives ont des capacités impressionnantes pour créer des données fictives ou pour simuler des scénarios hypothétiques. Cependant, leur fiabilité pour générer des données économiques réelles, telles que l'évolution du taux de chômage pour la France depuis 2000, est limitée. L'expérience montre que les erreurs sont extrêmement nombreuses.

### Exemple 3 Traiter des données, Générer des graphiques



A partir d'un fichier de données au format Excel trouvé sur le site du Ministère de la culture, on a demandé à l'IA de construire ce graphique. L'IA est capable d'extraire les données pertinentes en parcourant les feuilles et les cellules du fichier. Elle choisit aussi le graphique le plus approprié, peut également ajouter des annotations, des titres, des légendes, et des étiquettes pour rendre les visualisations plus compréhensibles. Elle peut même proposer un graphique interactif.

## Générer des QCM

Les IA génératives et les éditeurs de quizz en ligne embarquant des modèles d'IA sont particulièrement performants pour générer des QCM, à condition d'exercer une vigilance critique. Nous vous conseillons [cette vidéo](#) sur ce thème dans la même rubrique.

# Evaluer ou s'évaluer avec l'IA

A travers quelques exemples, nous allons montrer que l'IA permet aussi d'évaluer ou de s'évaluer.

## Exemple 1 Proposer des prompts pour que les élèves s'évaluent

### Phrase originale :

*Certaine petites entreprises dans le monde font fasse à des géants qui possède des multinationales et de plus gros moyen de production.*

### 1.1. Correction des fautes d'orthographe :

#### Phrase corrigée :

*Certaines petites entreprises dans le monde font face à des géants qui possèdent des multinationales et de plus gros moyens de production.*

#### Explication des fautes :

- *Certaine* → *Certaines* : Accord en genre et en nombre avec *entreprises* (féminin pluriel).
- *font fasse* → *font face* : Mauvais usage du verbe, l'expression correcte est *faire face à*.
- *qui possède* → *qui possèdent* : Accord en nombre avec *géants* (pluriel).
- *moyen* → *moyens* : Accord en nombre avec *de plus gros* (pluriel).

Nous utilisons l'IA pour permettre aux élèves d'améliorer leur expression écrite en classe ou chez eux (orthographe, grammaire, organisation des idées, etc.). Les élèves utilisent [le prompt](#) mis à leur disposition pour s'évaluer.

## Exemple 2 Evaluer une étude de document avec l'IA

### 1. Trouver les idées

**Commentaire:** L'élève a identifié correctement les deux idées principales comme étant le niveau de diplôme et l'âge. Ces idées sont clairement énoncées et développées dans la réponse. Toutefois, les idées secondaires spécifiques mentionnées dans les critères n'ont pas été explorées. **Note: 3/4**

### 2. Mobiliser des chiffres pour démontrer les idées

**Commentaire:** L'élève utilise des données pour soutenir ses idées. Cependant, l'élève n'a pas utilisé le terme "salaires médians" correctement tout au long de la réponse, ce qui est crucial pour l'exactitude des termes statistiques. Les calculs, bien que simples, sont corrects et soutiennent les affirmations faites, mais ils pourraient être plus développés pour renforcer l'argumentation. **Note: 2,5/4**

### 3. Structurer la réponse

**Commentaire:** La réponse est bien structurée avec une introduction claire, un développement cohérent des idées, et une conclusion résumant les points principaux. L'élève utilise des transitions logiques entre les idées, mais pourrait bénéficier de l'utilisation de mots de liaison plus variés pour une meilleure fluidité du texte. **Note: 0,75/1**

### 4. Rédiger la réponse

**Commentaire:** La clarté de l'expression est bonne, mais il y a quelques erreurs d'orthographe et de grammaire :

1. "c'est nettement moins bien payer" devrait être "c'est nettement moins bien payé".
2. "cette personne ainsi que son niveau d'expérience" pourrait être plus clairement formulé.
3. "plus nous sommes enclin" devrait être "plus nous sommes enclins". Ces erreurs affectent légèrement la qualité de la rédaction. **Note: 0,75/1**

Un même prompt peut être utilisé par le professeur pour corriger des travaux d'élèves ou par les élèves eux-mêmes pour évaluer et améliorer leur production grâce au retour correctif immédiat de l'IA. Nous décrivons plus précisément ce deuxième scénario.

Les élèves de seconde reçoivent [un document de travail](#) comprenant un tableau sur les salaires médians selon l'âge et le diplôme, ainsi qu'une question à traiter en respectant certaines consignes : « De quoi dépend le salaire ? ». Le travail se déroule en salle informatique, par groupes de 2 ou 3 élèves, ou sur leur téléphone. À tout moment, les élèves peuvent soumettre leur réponse à l'IA pour une évaluation, en utilisant [le prompt approprié](#). L'IA fournit des commentaires détaillés basés sur les attentes, et attribue une note sur 10. Grâce à ces retours, les élèves peuvent améliorer leur réponse afin d'augmenter leur score. Il est important de noter que pour que l'IA donne une évaluation proche de celle d'un professeur, le prompt doit être extrêmement long et spécifique à l'exercice. Par conséquent, il n'est pas facilement généralisable à d'autres exercices du même type.