

## Understanding Bias in AI Responses about Colonization

Ce scénario pédagogique invite les élèves de Première en DNL Histoire à adopter une posture critique face aux productions de l'intelligence artificielle appliquées à l'histoire coloniale britannique. À partir de la question *Does the AI tell the whole story of the British Empire?*, les élèves comparent des textes générés par IA avec des sources historiques authentiques afin d'identifier omissions, biais lexicaux et points de vue implicites. Ils conçoivent ensuite un contre-prompt pour obtenir une réponse plus nuancée et contextualisée. Le travail se conclut par une restitution orale en langue étrangère et par l'élaboration collective de règles de prudence face à l'usage de l'IA.

**Niveau :** Première générale et/ou technologique

**Dispositif :** DNL Histoire (section européenne / enseignement en anglais)

**Autres niveaux possibles :** Terminale (décolonisation, mémoire de la colonisation), EMI, projet interdisciplinaire histoire-anglais, SNT ou EMC, avec adaptation des documents et des exigences langagières.

### Thème du programme :

En DNL, il n'existe pas de programme spécifique ; l'enseignement prend appui sur les contenus disciplinaires du tronc commun. Ce scénario s'inscrit ainsi dans le thème 3 du programme d'Histoire de la classe de Première : La Troisième République avant 1914 : un régime politique, un empire colonial. Chapitre 3 : Métropole et colonies.

### Objectifs notionnels / capacités – compétences :

- Comprendre qu'une IA générative n'est pas une source neutre mais un modèle statistique susceptible de reproduire des biais
- Identifier des biais dans une production d'IA : discours édulcoré, univoque, eurocentré, simplificateur
- Comparer un texte généré par une IA à une source historique authentique
- Repérer les omissions, les implicites et le point de vue adopté
- Rédiger un contre-prompt pour orienter la réponse de l'IA
- Argumenter à l'oral en s'appuyant sur des preuves tirées des documents
- Coopérer au sein d'un groupe et mutualiser les observations

### Documents / outils mobilisés :

- Support introductif : fiche Understanding Bias in AI responses about Colonization.
- 3 études de cas (*The Great Exhibition and the Koh-i-Noor ; The White Man's Burden ; The Berlin Conference*) avec textes générés par IA à partir de prompts historiques et sources historiques authentiques fournies par l'enseignant

### Activité des élèves :

- Comprendre pourquoi une IA peut être biaisée grâce à l'analogie de la soupe et repérer les quatre "red flags" à rechercher dans les réponses générées (5-10 min)
- Analyser la narration produite par l'IA en relevant les formulations neutres, les euphémisations et le point de vue adopté (10 min)
- Confronter le texte généré à une source historique authentique afin d'identifier les faits oubliés, les écarts de ton et les biais de traitement (20 min)
- Rédiger un contre-prompt pour contraindre l'IA à intégrer des voix marginalisées et observer l'évolution éventuelle de la réponse (15 min)
- Préparer puis présenter un audit oral démontrant les biais repérés, les preuves tirées de la source et la reformulation du prompt (préparation 10 min + oral 2 min par groupe)
- Formaliser collectivement des règles de prudence face à l'usage de l'IA en histoire à partir d'un exit ticket (5 min)

### Différenciation :

Le scénario prend en compte l'hétérogénéité du groupe grâce à la conception de **trois études de cas de difficulté différente**. Cette progressivité permet d'adapter la complexité des textes, des implicites à repérer et des exigences d'analyse selon les besoins des élèves.

### Evaluation :

- La production orale permet d'évaluer à la fois la compréhension des biais, l'exploitation des documents, la capacité à justifier son analyse par des preuves précises et la clarté de l'expression en langue étrangère.
- Un exit ticket final permet de vérifier l'appropriation des règles de prudence élaborées collectivement.

### Bibliographie / sitographie :

- **UNESCO.** *L'IA générative menace l'histoire de l'Holocauste (rapport UNESCO)* [en ligne]. 18 juin 2024. Mis à jour le 16 août 2024. [Consulté le 19 avril 2026].
- **HAUT-COMMISSARIAT DES NATIONS UNIES AUX DROITS DE L'HOMME.** *Racisme et IA : « les biais du passé entraînent des biais pour l'avenir »* [en ligne]. 30 juillet 2024. [Consulté le 19 avril 2026].

- **MOHAMED, Shakir ; PNG, Marie-Therese ; ISAAC, William.** « Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence ». *Philosophy & Technology*, 2020.