CORRIGÉ

Culture Générale et Expression — 2e année

CHAPITRE 09 — APPLIQUER

Intelligence artificielle, deepfakes et crise de confiance

1. Corpus documentaire

DOCUMENT 1 — REPÈRES SUR LES MÉDITATIONS MÉTAPHYSIQUES , RENÉ DESCARTES (1641)
ŒUVRE PHILOSOPHIQUE

L'hypothèse du malin génie

Dans ses Méditations métaphysiques , Descartes pousse le doute méthodique à son point extrême en imaginant l'hypothèse d'un « malin génie », un être aussi puissant que trompeur qui emploierait toute son adresse à le tromper systématiquement — au point de lui faire prendre pour réels des sens, des images et des perceptions entièrement fabriqués. Cette hypothèse extrême sert un objectif précis : identifier ce qui, malgré une tromperie parfaite, résisterait encore au doute.

Descartes ne conclut pas que tout est nécessairement faux, ni que toute connaissance est impossible : il cherche un point d'appui suffisamment solide pour résister même à l'hypothèse d'une illusion parfaite. Cette expérience de pensée, vieille de presque quatre siècles, anticipe étrangement une question contemporaine : comment distinguer le réel d'une reconstruction suffisamment parfaite pour l'imiter ?

Notions mobilisées : doute hyperbolique · illusion parfaite · recherche d'un point d'appui certain

DOCUMENT 2 — UN SCANDALE À DOUBLE DÉTENTE (CAS FICTIF À BUT PÉDAGOGIQUE)
DOCUMENT NON LITTÉRAIRE

Quand la preuve elle-même devient incertaine

Une vidéo circule, montrant un responsable d'entreprise tenant des propos compromettants. L'enregistrement, fluide et de bonne qualité, se révèle après analyse être un deepfake : les mouvements des lèvres ne correspondaient pas exactement aux paroles, et aucune source officielle ne confirmait l'authenticité de la scène.

Quelques semaines plus tard, un enregistrement authentique montrant ce même responsable dans une situation embarrassante est diffusé. Celui-ci affirme immédiatement qu'il s'agit « encore d'un deepfake », sans apporter aucune preuve de cette affirmation. Une partie du public, échaudée par le premier épisode, choisit de le croire sur parole — bien que cette seconde vidéo soit, cette fois, authentique.

Notions mobilisées : deepfake · dividende du menteur · crise de confiance généralisée

Questions sur le corpus

1. En quoi l'hypothèse du malin génie (document 1) permet-elle d'éclairer la situation décrite dans le document 2 ?

✓ ÉLÉMENTS DE RÉPONSE

- Descartes imaginait une hypothèse extrême, purement philosophique, pour tester la solidité de la connaissance face à une tromperie parfaite.
- Dans le document 2, cette hypothèse devient une réalité technique concrète : un deepfake peut effectivement imiter le réel de façon presque indétectable.
- ■ La question cartésienne, autrefois purement spéculative, prend une actualité très concrète avec l'IA générative.

2. Identifie, dans le document 2, un indice qui a permis de détecter le premier deepfake, puis explique le mécanisme à l'œuvre dans le second épisode.

✓ ÉLÉMENTS DE RÉPONSE

- Indice de détection : le désajustement entre les mouvements des lèvres et les paroles, et l'absence de confirmation par une source officielle.
- Second épisode : le responsable nie une preuve pourtant authentique en l'accusant d'être un deepfake, sans apporter aucune preuve de cette affirmation — c'est le mécanisme du dividende du menteur.
- ■ Le doute installé par le premier épisode profite directement au responsable lors du second.

3. En confrontant les deux documents, montre pourquoi rechercher « un point d'appui certain » (Descartes) reste une démarche pertinente face au dividende du menteur.

✓ ÉLÉMENTS DE RÉPONSE

- Descartes ne conclut pas que tout est faux : il cherche ce qui résiste à l'illusion la plus parfaite.
- Face au dividende du menteur, la solution n'est pas non plus de renoncer à toute preuve, mais de rechercher des points d'appui fiables : détection technique, certification, source officielle.
- ■ Dans les deux cas, le doute n'est pas une fin en soi : il sert à identifier ce qui reste solide malgré l'illusion.

2. Essai argumenté

À l'épreuve E1, l'essai est noté sur 10 points : tu dois traiter le sujet de manière argumentée et nuancée , en t'appuyant sur tes connaissances personnelles, sur le corpus étudié et sur les notions vues en cours (contenus plausibles, hallucinations, deepfakes, dividende du menteur, détection, certification, éducation aux médias). Pour cet entraînement, vise 180 à 220 mots — un format resserré pour t'exercer à la construction d'une réflexion nuancée.

SUJET 1

Sujet 1 « L'intelligence artificielle menace-t-elle la vérité ? » Tu répondras à cette question de façon argumentée et nuancée, en t'appuyant sur le corpus étudié, sur tes lectures et sur des exemples précis.

✓ PLAN INDICATIF

- I. L'IA fragilise l'authentification des contenus — hallucinations, deepfakes, dividende du menteur rendent preuve et illusion plus difficiles à distinguer.
- II. Mais elle amplifie des tensions déjà anciennes plutôt qu'elle n'en crée de nouvelles — retouche photographique, propagande, viralité posaient déjà ces questions (Descartes imaginait déjà une illusion parfaite).
- III. L'IA ne rend donc pas la vérité impossible — elle exige de renforcer détection, certification et éducation aux médias, plutôt que de renoncer à toute vérification.

SUJET 2

Sujet 2 « Le dividende du menteur rend-il toute preuve suspecte ? » Tu répondras à cette question de façon argumentée et nuancée, en t'appuyant sur le corpus étudié, sur tes lectures et sur des exemples précis.

✓ PLAN INDICATIF

- I. Le dividende du menteur installe un doute qui peut sembler généralisé — si tout peut être fabriqué, toute preuve devient contestable en apparence.
- II. Mais ce doute généralisé n'est pas la seule issue possible — détection technique, certification institutionnelle et recoupement des sources permettent de distinguer les preuves fiables des affirmations non étayées.
- III. Toute preuve n'est donc pas suspecte — seule une preuve non vérifiée ou non recoupée le devient légitimement ; la vigilance doit rester proportionnée, sans céder au relativisme intégral.