

Banque X/ENS - Session 2024
RAPPORT SUR L'ÉPREUVE ÉCRITE D'ANGLAIS
ENS PARIS-SACLAY

L'épreuve écrite de langue vivante en anglais portait cette année sur les avancées récentes de l'intelligence artificielle, et plus spécifiquement sur les risques réels et imaginaires qu'elle fait peser sur nos sociétés.

L'épreuve se divise en deux parties. Pour la première partie de l'épreuve (A), les candidats doivent exploiter quatre documents : trois articles et un graphique. Il fallait utiliser ce dernier pour confronter données et arguments présentés dans les textes. La synthèse devait comprendre de 600 à 675 mots. La seconde partie de l'épreuve (B) consistait à réagir à un texte d'opinion de 500 à 600 mots, sans hésiter à énoncer et défendre son propre point de vue.

Ces deux exercices font appel à l'esprit de synthèse et d'analyse des candidats ainsi qu'à leur capacité à s'exprimer et à argumenter dans une langue écrite correcte et riche, suivant une forme et une méthodologie soignée.

Le sujet pouvait sembler assez prévisible, ce qui explique pourquoi une majorité de candidats semble avoir opté pour une problématique assez convenue tant dans la synthèse que dans le texte d'opinion, avec assez peu de prise de risques.

LES DOCUMENTS

(A) Le dossier de la première partie, la synthèse, comportait quatre documents.

1. Synthèse

Texte 1: Fears about AI's existential risk are overdone, says a group of experts

Collectif d'auteurs autour de Blaise Agüera y Arcas, *The Economist*, 1 July 2023

L'article met en lumière les préoccupations relatives aux risques liés à l'intelligence artificielle (IA), notamment la possibilité qu'une IA surpuissante puisse menacer l'existence même de l'humanité (x-risk). Bien que la plupart des chercheurs qui signalent ces risques soient sincèrement motivés par leur préoccupation face aux dangers actuels et futurs liés à l'IA, les appels à l'action pour atténuer ces risques peuvent entraver le développement d'utilisations bénéfiques de l'IA et détourner l'attention des risques immédiats.

L'article souligne que les dommages causés par l'IA ne nécessitent pas une superintelligence ; ils sont déjà présents dans des domaines tels que les violations de la vie privée, la désinformation, les menaces sur la cybersécurité et les biais dans les systèmes de prise de décision. De plus, il souligne le risque imminent posé par l'utilisation militaire de l'IA, ainsi que sa mise en œuvre par des régimes répressifs pour la surveillance de masse et le contrôle social.

Bien que l'IA actuelle ne soit pas encore superintelligente, elle présente déjà des risques. Les modèles actuels d'IA sont puissants mais sujets à des erreurs, à des biais et à des limites de raisonnement. L'article examine également la question de savoir si l'IA, superintelligente ou non, devrait être considérée comme une entité autonome ou comme une technologie façonnée par l'humanité.

Certains craignent que cette IA ne cause l'extinction de l'humanité. Cependant, l'article remet en question cette idée en soulignant que les extinctions ne sont pas simplement le résultat de luttes pour la suprématie entre des organismes supérieurs et inférieurs. Il met en avant l'idée que la vie forme un réseau interconnecté, où la symbiose et le mutualisme entre différentes espèces sont courants. Il souligne ainsi que les IA dépendent complètement des humains pour leur existence, et que l'automatisation du travail de traitement de l'information rendu possible par l'IA est à la fois une opportunité majeure et un défi socio-économique primordial.

Enfin l'article souligne que les régulateurs devraient se concentrer sur les problèmes immédiats posés par l'IA et chercher à rendre les modèles plus sûrs et leurs opérations plus prévisibles, plutôt que de se concentrer exclusivement sur les risques existentiels futurs. Il suggère que la régulation devrait se concentrer sur la prévention des déploiements inappropriés de l'IA plutôt que ceux que poseraient une IA superintelligente dont on ne sait pas s'il est réaliste de prédire son émergence.

Texte 2: Artists may make AI firms pay a high price for their software's 'creativity'

John Naughton, *The Guardian*, 28 October 2023

L'article traite de l'utilisation croissante d'IA génératives comme Midjourney, un outil très à la mode, qui convertit des textes en images. Les utilisateurs peuvent donner des descriptions textuelles et faire générer les images correspondantes par l'IA. Cependant, le processus repose en grande partie sur le fait que ces IA sont entraînées sur des milliards d'images disponibles en ligne, dont beaucoup peuvent être des œuvres protégées par le droit d'auteur, ce qui soulève des inquiétudes concernant la propriété intellectuelle.

Les entreprises propriétaires de ces IA pourraient en effet être confrontées à des poursuites judiciaires pour utilisation non autorisée d'œuvres protégées par le droit d'auteur. Des actions en justice ont déjà été intentées par des artistes affirmant que leurs travaux ont été utilisés pour entraîner ces IA sans autorisation, ce qui pourrait entraîner le paiement de dommages et intérêts considérables – par exemple Getty a entamé une action contre Stability AI.

Pour permettre aux artistes moins fortunés de lutter contre les risques de vol de propriété intellectuelle, des chercheurs de l'Université de Chicago ont développé un outil appelé Nightshade, qui permet aux artistes d'ajouter des modifications invisibles à leurs œuvres avant de les publier en ligne. Cet outil peut perturber les modèles d'IA et rendre les résultats imprévisibles (l'article suggère des transformations humoristiques de chats en chien mais aussi de Boris Johnson en cochon, faisant référence implicite aux caricatures souvent publiées dans le *Guardian*). Une telle technique pourrait avoir des effets destructeurs sur les entreprises d'IA (ici l'auteur fait jeu de mot implicite sur « Deadly nightshade », la belladone, une fleur au poison très puissant).

Texte 3: What would humans do in a world of super-AI?

The Economist, 23 May 2023

L'article explore les conséquences possibles des progrès croissants de l'intelligence artificielle générale (AGI), qui pourrait surpasser les compétences humaines dans de nombreux domaines d'ici la prochaine décennie. Inspiré par des œuvres de fiction comme le dessin animé *WALL-E* et la série *Culture* de Iain M. Banks, l'article examine les conséquences potentielles d'une telle avancée sur les individus et la société.

Il souligne que si l'AGI devait automatiser entièrement la production, cela pourrait conduire à une croissance économique explosive, mais uniquement si l'IA pouvait également automatiser le processus de recherche et d'auto-amélioration. En revanche, la croissance économique pourrait être entravée si l'AGI ne parvenait pas à automatiser tous les aspects de la production.

En ce qui concerne la politique, l'article suggère que les humains seraient peu susceptibles de céder le contrôle aux robots. Au contraire, il est probable que les humains continueront à accorder une importance primordiale à leur influence politique, même dans un monde dominé par l'AGI.

Enfin, l'article explore les domaines où les humains pourraient conserver un avantage par rapport à l'AGI, notamment dans la fourniture de services qui impliquent un niveau d'interaction humaine et émotionnelle que les machines ne peuvent pas reproduire. Cela inclut des domaines tels que les soins et la thérapie, ainsi que la capacité des humains à donner un sens et une valeur aux produits et services par leur histoire et leur origine humaine (l'article donne comme exemple la valeur supérieure donnée aux diamants authentiques par rapport aux diamants de synthèse pourtant identiques d'un point de vue chimique).

Document 4: AI Astrology

L'intérêt du graphique, qui était associé au texte 3, est de montrer l'évolution des prédictions des spécialistes concernant l'arrivée potentielle d'une IA générale sur la période 2022-23. Il montre que nombre de spécialistes ont révisé leurs prédictions et voient venir cette IA dans un avenir beaucoup plus proche (autour de 2040) qu'il y a 2 ans (autour de 2070). Cependant le titre du graphique indique bien qu'il ne s'agit pas de prévisions scientifiques, et les variations fréquentes observables sur la ligne du 75^{ème} percentile montrent bien que ces prévisions sont très influencées par des événements ou informations tels que la publication de Chat-GPT etc. Enfin il était utile de faire le lien avec le texte 3, puisqu'il explicitait l'origine du graphique et l'identité probable des sondés : a priori il s'agissait de personnes compétentes ou bien informées, puisque la plateforme Metaculus est désignée comme « a favorite of many techies ». Il ne s'agissait donc pas d'un sondage effectué auprès du grand public, comme l'ont supposé un certain nombre de candidats. Cela dit, dans l'ensemble le graphique a été mieux compris et exploité que les années précédentes – presque trop dans certains cas.

2. Opinion

Governments must not rush into policing AI

The Economist, 26 October 2023

L'article examine les préoccupations concernant les risques potentiels associés à l'intelligence artificielle (IA), notamment la possibilité que l'IA devienne suffisamment puissante pour menacer l'humanité (*reprenant certains arguments énoncés dans le texte 1*). Les scénarios dystopiques incluent

des IA qui deviennent autonomes et hostiles, ainsi que des modèles de langage de grande taille (LLMs) alimentant des armes cybernétiques dangereuses.

Les gouvernements et les régulateurs sont confrontés au défi de prendre des mesures pour atténuer ces risques sans entraver l'innovation. Alors que la réglementation est nécessaire pour répondre aux préoccupations liées à la vie privée, à la désinformation et à la propriété intellectuelle, une réglementation passée trop rapidement pourrait également étouffer la concurrence et l'innovation.

Une proposition est la création d'un organisme similaire au Groupe intergouvernemental d'experts sur l'évolution du climat (GIEC), chargé d'étudier les risques associés à l'IA. De plus, des codes de conduite volontaires pour les fabricants de modèles d'IA pourraient aider à promouvoir la transparence et à éviter un environnement non réglementé.

Selon l'auteur, il est important de ne pas créer d'antagonisme entre gouvernements et entreprises créatrices d'IA pour encourager la transparence et le partage d'informations, et éviter l'écueil des boîtes noires. Mais il faut agir vite pour éviter l'erreur commise par les gouvernements qui n'ont pas vu arriver la montée en puissance des réseaux sociaux autour de 2010.

Les candidats ont donc été amenés entre autres à prendre position pour ou contre les propositions de « GIEC de l'IA », ou en faveur d'un modèle volontariste contre un modèle de régulation forte par les gouvernements.

OBSERVATIONS DU JURY

Forme et méthodologie

Dans l'ensemble, les objectifs de l'exercice ont été bien compris, ce qui atteste de la bonne préparation des candidats. Cependant, les consignes n'ont pas toujours été respectées, et les correcteurs et correctrices déplorent encore des faiblesses dans la méthodologie et la forme des productions.

Nous rappelons par exemple que le nombre de mots doit être indiqué (sans tricher) à la fin de chaque exercice par les candidats, sous peine de malus. A noter que les barres de compte de mots de la même couleur que l'écrit rendent les copies difficilement lisibles – il est conseillé d'utiliser un crayon ou une autre couleur. De plus en plus de copies sont rédigées de manière indéchiffrable – un effort de lisibilité serait appréciable.

Si certains titres ont semblé un peu prévisibles, d'autres étaient courts et percutants, avec parfois des allusions cinématographiques ou littéraires pertinentes ou des jeux de mots astucieux :

BrAIve new world - A(m) I a threat? - Of AI and Men - Humpty Dumpty sat on AI - Will cAI be Abel to murder his brother humanity? - Are you in the Open AI or Open Your Eyes team? - I am not in danger? AI am the danger - Error 404: Humans not found - the Botfather: AI makes an offer you can't refuse - Open (to) AI?- In Super-AI we trust - An AI for an eye - Artificial intelligence: real

threat? - (A)I will survive! - The development of super AIs is only Midjourney - Should AI stay or should AI go? - The Good, the Bad and the AI - 2030: An AI Odyssey - Are we s-AI-f from our creation? - Please wait, a title will soon be generated by an AI - AI-POCALIPS NOW! - AI: a weapon of math destruction - this text was not AI generated, open your AIs - Prompt: are you going to kill us all? Answer: for the moment I do not have enough information to answer your question.

Ces suggestions pertinentes ont fait l'objet de bonus pour les candidats.

(A) La synthèse

La synthèse ne doit pas inclure de remarques personnelles ou d'éléments extérieurs aux documents donnés, même pour l'accroche. Elle doit être concise, mais complète, n'omettant pas d'informations importantes, mais ne se perdant pas non plus dans les détails. Tous les documents — graphiques inclus — doivent être traités équitablement et nommés avec précision, sans pour autant que l'introduction se transforme en pure énumération. Trop de problématiques se reposent sur l'expression « to what extent » : or celle-ci marque trop souvent une absence de problématisation. Peu de candidats parviennent à formuler des problématiques qui rendent compte de l'ensemble des aspects du dossier. A contrario, les bonnes copies ont articulé une présentation des documents déjà problématisée en introduction (et non une simple liste descriptive des documents), ou mentionné le fait que la source du document 4 est donnée dans le document 3.

La synthèse est un exercice exigeant où les candidats doivent examiner des points de vue divergents pour parvenir à une conclusion équilibrée. Il est essentiel de reformuler l'essentiel des documents avec ses mots propres, en utilisant un vocabulaire précis pour capturer la finesse des arguments. Les opinions variées des auteurs ou des spécialistes cités dans les textes doivent être fidèlement et impartialement présentées, en respectant les nuances des déclarations, sans simplifications, jugements ou caricatures. Les mots de liaison doivent être maîtrisés et refléter la logique du texte, plutôt que d'être utilisés systématiquement.

Le sujet permettait une synthèse riche en détails, or souvent la fin du texte 3 a été occultée (le lien entre développement technique et valeur économique, tout particulièrement), une place trop grande a été faite à une description de la technique présentée dans le texte 2 mais il manquait les références légales qui émaillent ce texte, et la liste intéressante des méfaits actuels de l'IA citée dans le texte 1 est à peine abordée. Les meilleures copies font la distinction entre le sujet du dossier (opportunités et menaces liées au développement rapide et peu contrôlé de l'IA) et la problématique qu'ils développent, en proposant par exemple une synthèse qui repose sur la notion d'incertitude / instabilité face au développement de l'IA. De nombreux ponts entre les textes étaient possibles (labour and the economy, the law and regulatory issues, politics and human organisation/civilisation) pour construire une problématique originale. On regrette donc que la problématique choisie par la majorité des copies pour la synthèse soit le sujet du dossier et que l'approche prise soit convenue (situation, problème, solution / avantages, inconvénients, futur possible) ou se concentre uniquement sur les aspects techniques de l'IA. Notons l'importance de prendre en compte les définitions fournies par les textes. Dans certaines synthèses, ce qui caractérise une AI surpuissante ou générale n'a pas été clarifié. Une fois n'est pas coutume, les candidats donnent parfois trop de poids au document 4 dans la construction de la problématique. Cependant, de nombreux candidats comprennent mal ou n'exploitent pas assez ce graphique, donnant trop d'importance aux percentiles pour perdre de vue les lignes essentielles qui s'en dégagent. En revanche certains explicitent astucieusement l'utilisation du terme « astrology » dans le titre de ce document et soulignent à quel point les prédictions oscillent dans le temps. Si certains ont confondu la prédiction avec l'arrivée effective de l'IA générale, d'autres ont introduit de

la subjectivité quant à l'interprétation du graphique, suggérant que les "forecasters" sont inquiets ou optimistes.

Il est regrettable que les candidats fassent l'impasse sur le langage inclusif. Les auteurs des articles y sont systématiquement des hommes, même quand l'auteur n'est pas nommé.

Conseil aux futurs candidats : Ne pas multiplier les problématiques. Utiliser des guillemets pour les titres des articles (ne souligner que le titre du journal). Bien lire la légende du graphique présenté en document 4 pour en comprendre le sens. Inclure une introduction claire et une conclusion pertinente.

(B) Texte d'opinion

Si l'exercice de la seconde partie a été bien compris et effectivement traité par la majorité des candidats, ils n'ont pas toujours répondu de manière appropriée, et beaucoup ont encore manqué de temps pour la rédiger ou la terminer. Bien que l'exercice exige des candidats qu'ils prennent position sur la question, il convient en effet d'éviter toute partialité excessive. Que les candidats soient en accord ou en désaccord avec l'auteur du texte d'opinion, leur réaction doit rester mesurée. Il est tout à fait possible de ne pas être d'accord avec l'auteur, à condition d'avancer des arguments accompagnés par des exemples précis. Le jury a néanmoins constaté que beaucoup moins de candidats ont cédé à la tentation de l'invective – beaucoup ont souligné qu'ils ou elles étaient en accord avec le texte, mais qu'ils souhaitaient le nuancer, et l'ont fait avec des exemples pertinents et des arguments solides. Des formats originaux ont été proposés : billet de blog, format « article review », « letter to the editor » etc.

Certains candidats font encore l'impasse sur le texte d'opinion et partent directement sur une dissertation déconnectée du sujet. A l'inverse, d'autres collent trop au texte et le décortiquent point par point sans apporter de références nouvelles. Le jury rappelle que le texte d'opinion ne peut pas se limiter à un résumé de l'éditorial. Il est également conseillé de rédiger un plan de son argumentation, l'écriture au fil de l'eau donnant souvent des copies maladroites, voire contradictoires.

Il est regrettable que la majorité des candidats ne donnent que très peu d'exemples non cités dans le texte, ou ne se réfèrent même pas à l'actualité, pourtant en lien direct avec la problématique soulevée. En revanche il était tout à fait possible de s'appuyer sur les éléments présents dans les documents de la première partie, sans pour autant répéter le contenu de la synthèse.

Les bonnes copies font par exemple référence aux craintes de Geoffrey Hinton, au travail de Nick Bostrom sur le "X-risk", ou donnent des exemples précis tels que des références littéraires venant appuyer une argumentation sur les peurs et les fantasmes sur le développement de technologies et ses conséquences sur le futur (par exemple les trois lois de la robotique formulées par Isaac Asimov, *Frankenstein* et d'autres romans de science-fiction, récits utopistes ou spéculatifs). Pour appuyer l'argumentation sur la régulation, les bonnes copies incorporent des références précises à des événements ou faits récents comme l'AI Act de l'Union Européenne ; l'interdiction de ChatGPT en Italie en avril 2023, le AI Safety Summit tenu à Bletchley Park à l'automne 2023. Elles mentionnent l'exploitation des travailleurs cachés de l'IA (ceux qui taggent les images dans les pays pauvres) ou encore l'effet sur l'environnement des data-centers et de l'augmentation exponentielle du recours à Internet. Une copie souligne avec raison la contradiction entre les 2 articles de la synthèse et celui du texte d'opinion (tous émanant de *The Economist*).

Qualité de langue et d'écriture

Les structures simples sont en général bien maîtrisées. En revanche, beaucoup de copies ne savent pas introduire correctement une problématique et ne maîtrisent pas la syntaxe de la question (style indirect ou direct). On note également des erreurs liées à la thématique proposée, en particulier pour ce qui concerne la manière de se référer aux articles, plusieurs d'entre eux provenant du magazine *The Economist* : on trouve ainsi « the 2 The Economist articles » ou encore « The Economist's graph », voire « the economist's article », transformant ainsi le nom d'une publication en un métier de l'auteur. A cet égard, la manière de se référer aux articles anonymes tend parfois à personnaliser le texte lui-même (the article tries to understand* / the article seems to think*), ou à réduire le collectif au seul auteur nommé (Blaise Agüera y Arcas). On trouve encore quelques copies qui utilisent IA au lieu de AI, ou l'emploi de pronoms masculins ou féminins pour des entités non vivantes (AI, ChatGPT...). Beaucoup de candidats ont eu des difficultés avec la construction de to steal (qu'ils ont parfois même confondu avec stealth). Le prétérit et le participe passé de to steal posent ainsi très souvent problème, même quand la phrase n'est pas construite de façon erronée. Enfin les termes nécessaires pour parler du graphique sont souvent approximatifs.

Les correcteurs et correctrices ont noté des efforts réels d'apprentissage de tournures et structures idiomatiques, mais regrettent de retrouver en parallèle des erreurs récurrentes telles que « to what extend* » (au lieu de « to what extent ») ou des défauts dans l'usage de la ponctuation (absence de virgules, absence de points, phrase qui commencent après une virgule ou après rien). Les confusions sont fréquentes entre few / a few ; stop to be / stop being; don't have to / mustn't ; feel / feel like ; as / such as ; actual / current ; problematic / issue ; arrive / arrival au lieu de advent ; apparition / appearance pour emergence ; prevent / warn ou stop from ; threat / threaten; scared / scary ; era / area ; politics / politicians ; actual / current ; ask / demand ; resume / summarize ; scrape / scrap ; *benefic ou benefical pour beneficial ; expose pour exposer. On rencontre également des difficultés de distinction entre humankind et humanity, human et humane, humans, human beings.

Les correcteurs et correctrices invitent les candidats à soigner l'orthographe, qui est souvent influencée par le français : *ressource pour resources ; *developpment ; *technologie ; *futur ; *responsability ; *engeneer ; to / too ; waist / waste ; it's / its ; were / where. Les préfixes et suffixes posent également problème (ex: *powerfull / carefull / usefull; *unecessary), de même que la formation des adverbes (*sincerly / technicaly / illegaly / precisly / extremly / finaly etc.) ou les doubles consonnes (*dropped).

Parmi les erreurs les plus courantes, on note des difficultés avec les s : le s final du pluriel ne nécessite pas d'être précédé d'une apostrophe ; these et those ne nécessitent pas de – s pluriels, tout comme les adjectifs en général. De même on trouve encore souvent des s de 3ème personne du singulier absents là où il en faudrait mais présent après des modaux (ex : it can highlights). Le s du génitif pose aussi souvent problème (the threat of AI's, people opinion...).

Conclusion

Le jury souhaite conclure sur une note positive en félicitant les nombreux candidats qui ont fait preuve d'un lexique riche et précis, d'une rédaction variée et élégamment tournée, d'une méthodologie

exemplaire, d'une pensée rigoureuse, ainsi que d'une bonne connaissance des problématiques soulevées par l'Intelligence Artificielle.