



Evaluation d'une IA générative invitée au Comité de programme du Lambda Mu 24

Assessment of a generative AI invited to the Lambda Mu 24 Program Committee

CLAUDE Francis
Enseignant-chercheur
ESTP

Grande école d'ingénieurs de la
construction
Cachan, France
fclaude@estp.fr

THING LEO Gilles
Enseignant-chercheur
ESTP

Grande école d'ingénieurs de la
construction
Cachan, France
gilles.thingleo@estp.fr

VICTORRI Bernard
Directeur de la Recherche
Risk Intelligence & Decisions
bv@rid-lab.com

Résumé — Cet article est issu d'une expérimentation qui a consisté à inviter une IA générative au Comité de Programme du Congrès Lambda Mu 24 à Bourges. Après un survol historique rapide des avancées de l'IA, la problématique à résoudre est présentée ainsi que le verrou de cette recherche et l'hypothèse de travail retenue. Les travaux, résultats et tests sont présentés avant de les discuter et d'ouvrir des perspectives d'approfondissement complémentaires.

Mots-clefs — *IA générative, comité de programme, congrès, catégorisation,*

Abstract — This article is the result of an experiment which consisted of inviting a generative AI to the Program Committee of the Lambda Mu 24 Congress in Bourges. After a quick historical overview of advances in AI, the problem to be resolved is presented as well as the key to this research and the working hypothesis retained. The work, results and tests are presented before discussing them and opening up additional perspectives for in-depth study.

Keywords — *Generative AI, program committee, congress, categorization,*

1

INTRODUCTION

2

Depuis la première interrogation au milieu du siècle dernier de Alan Turing “ Les machines peuvent-elles penser ?” que de chemin parcouru par l'Intelligence Artificielle (IA). L'IA est devenue un domaine de recherche et développement de plus en plus important dans divers secteurs, des soins de santé à la finance, en passant par les transports et l'éducation. Un sous-domaine particulier de l'IA qui a retenu beaucoup d'attention est l'IA générative, qui peut créer de nouveaux contenus, tels que du texte, des images ou de la musique, sur la base de données. Alors que l'utilisation de l'IA générative continue de croître, il nous est apparu opportun de nous interroger sur l'invitation d'une IA au Bureau du Comité de Programme (BCP) du Congrès Lambda Mu 24.

3

4

Pour le BCP, en sus de la définition du contenu des conférences et tables rondes, ateliers, tutoriels et de la sélection des communications pour les différents Prix, l'activité la plus conséquente consiste à superviser la sélection des communications pour les sessions conférences et interactives du Congrès et à les répartir dans des sessions homogènes. Pour cela le BCP peut s'appuyer sur les Groupes d'Évaluation (GeVAL) formés à partir d'une centaine de relecteurs impliqués et experts de thématiques préalablement déterminées. Le processus se déroule en quatre phases. La première permet de recevoir et évaluer les propositions de communications par l'intermédiaire de « résumés » reçus sur le site des auteurs. La deuxième est dédiée à la relecture des textes complets. La troisième permet de vérifier la prise en compte des ultimes commentaires des GeVAL par les auteurs avant l'envoi d'une notification favorable ou pas à une présentation lors des sessions et publications ultérieures. Enfin, la quatrième concerne davantage le BCP et le Comité d'organisation afin que les sessions puissent se dérouler harmonieusement et soient présidées par des personnes qualifiées.

5

6

La première section, comme un état de l'art est consacrée à un bref survol historique du développement de l'IA. Dans la deuxième, nous précisons notre problématique à résoudre de catégorisation des résumés au sein des différents métiers du risque sachant qu'à ce jour “un” article, publié lors du Lambda Mu 22 en 2022, est disponible sur le World Wide Web. Le verrou de cette recherche, l'hypothèse retenue et la démarche suivie sont ensuite présentées compte tenu de l'absence de solution

7

8

23 exploitable issue de notre état de l'art. L'analyse des connaissances existantes porte particulièrement sur l'IA générative et
24 notamment des points de vues de son architecture et de son fonctionnement. La réalisation des travaux, les résultats et les tests
25 sont décrits dans une troisième section. Avant de conclure, une quatrième section discute des résultats obtenus et aborde des
26 perspectives d'approfondissements potentiels.

27 I. IA ET LANGAGE, UN SURVOL HISTORIQUE RAPIDE

28 A. Du programme de recherche initié par Alan Turing aux approches symbolique et connexionniste

29 Le langage a été un enjeu crucial pour l'IA dès ses débuts au milieu du siècle dernier. En témoigne l'article fondateur d'Alan
30 Turing (Turing, 1950), qui, pour répondre à la question *Can machines think?* (Les machines peuvent-elles penser ?), propose le
31 fameux jeu de l'imitation : une machine et un humain dialoguent par écrit avec un expérimentateur, et si celui-ci ne peut détecter
32 lequel de ses interlocuteurs est une machine, alors la machine sera dite « intelligente ». Nul doute que les systèmes d'IA
33 générative, apparus ces dernières années, passeraient haut la main le test de Turing. Il aura donc fallu quelques 70 ans, bien des
34 tâtonnements, et d'énormes progrès technologiques, pour réaliser entièrement le programme de recherche inspiré par Turing.

35 Dès les premiers développements, deux approches concurrentes se sont dégagées. La première, que l'on a appelée l'**approche**
36 **symbolique**, a cherché à modéliser la pensée humaine, en donnant de manière explicite à la machine les règles de fonctionnement
37 du raisonnement humain : cela a conduit notamment dans les années 70-80 au développement des systèmes experts qui ont
38 suscité un véritable engouement, mais l'enthousiasme est retombé par la suite quand sont apparues les difficultés de rendre
39 compte par ces méthodes de la complexité des processus à l'œuvre chez les experts humains, pour lesquels l'intuition et
40 l'expérience acquises inconsciemment jouent un rôle au moins aussi important que le raisonnement logique conscient. La
41 deuxième approche, que l'on a appelée l'**approche connexionniste**, a consisté à modéliser directement l'activité du cerveau
42 humain, à l'aide de réseaux de neurones formels imitant (en le simplifiant) le fonctionnement des neurones biologiques, les poids
43 des interactions entre neurones étant modifiés par apprentissage jusqu'à l'obtention d'un taux de réussite convenable pour une
44 tâche donnée. Cette approche a connu ses premiers vrais succès dans les années 80, pour des tâches de reconnaissance de forme,
45 après que soit mise au point une méthode très efficace d'apprentissage (basée sur l'algorithme de rétropropagation du gradient)
46 (Le Cun, 1985).

47 En ce qui concerne le traitement du langage, c'est l'approche symbolique qui a dominé durant ces années. Les recherches se
48 sont focalisées sur la manière de représenter le sens, ce qui impliquait de trouver un moyen de structurer les connaissances
49 humaines à l'œuvre lors de la compréhension d'un texte par exemple avec les graphes conceptuels de Schank (Schank, 1969),
50 ou encore la théorie des frames de Minsky (Minsky, 1974). Dans l'approche symbolique, les théories linguistiques sont fortement
51 sollicitées, notamment pour réaliser l'analyse syntaxique des textes en réception et en production. Mais malgré tous ces efforts,
52 les résultats sont peu convaincants. Le sens des phrases et des textes se révèle très difficile à formaliser, en dehors de quelques
53 domaines fermés très restreints. Quant à l'approche connexionniste, elle semble encore moins efficace, malgré la mise au point
54 d'un type de réseaux de neurones bien adaptés au traitement de données séquentielles, les réseaux récurrents (Elman, 1990).

55 B. La naissance de l'apprentissage profond

56 Au cours des années suivantes (1990-2005) apparaissent de nouvelles approches, à mi-chemin entre les approches symbolique
57 et connexionniste, caractérisées par une utilisation de plus en plus grande de méthodes quantitatives fondées sur l'analyse
58 statistique de volumes très importants de données (big data). Cette évolution est favorisée par les progrès de la technologie,
59 permettant des traitements massifs qui étaient impensables auparavant. Pour nous en tenir au traitement du langage, c'est à cette
60 époque que la traduction automatique prend tout son essor, avec l'exploitation de grands corpus bilingues : on renonce à essayer
61 de représenter le sens des textes que l'on veut traduire, mais pour chaque mot ou groupe de mots (syntagme), on prend en compte
62 le contexte dans lequel il est inséré pour obtenir une traduction correcte, grâce aux exemples proches que l'on extrait des corpus
63 bilingues dont on dispose (Koehn 2003). Il semble donc acquis à cette époque que les deux approches traditionnelles ont en partie
64 échoué : pour réaliser les objectifs de l'IA, plutôt que de chercher à imiter la pensée ou le cerveau humain, mieux vaut s'attaquer
65 à chaque problème indépendamment de manière pragmatique, sans a priori théorique, en utilisant toute la puissance de calcul et
66 les masses de données offertes par les avancées technologiques.

67 Mais les choses vont radicalement changer à partir de 2007. En effet, cette année-là trois spécialistes des réseaux de neurones,
68 Yann Le Cun, Geoffrey Hinton et Yoshua Bengio, présentent des résultats inégalés avec des réseaux de très grande taille : c'est
69 l'acte de naissance de l'**apprentissage profond** (deep learning), qui est d'abord appliqué à la reconnaissance de formes (Bengio
70 & LeCun 2007). L'architecture des réseaux utilisés rappelle par bien des aspects celle du système visuel des mammifères : un
71 grand nombre de couches successives, analysant les images d'un niveau très local au niveau global, grâce à des opérations de
72 même type (des convolutions) appliquées à différentes échelles. Les performances sont impressionnantes : pour la première fois,
73 elles égalent les performances humaines sur des tâches spécifiques, à commencer par la reconnaissance de chiffres manuscrits.

74 Les techniques de l'apprentissage profond se sont alors très vite répandues dans tous les domaines de l'IA. Une étape
75 remarquable a été franchie avec la maîtrise totale du jeu de go, réputée impossible pour une machine ; à noter que le meilleur
76 programme fonctionne sans aucune connaissance au départ hormis les règles du jeu, ni aucun exemple de partie jouée par des
77 humains, mais uniquement par de l'auto-apprentissage en jouant contre lui-même (Silver et al. 2017).

78 Dans le domaine du langage, l'apprentissage profond a aussi opéré une révolution, d'abord avec des architectures assez
79 complexes, à base de modules récurrents, puis, à partir de 2017, avec une architecture beaucoup plus simple mais des tailles de
80 réseau toujours plus gigantesques, à la suite d'un article fondateur d'une équipe de recherche de Google (*Attention is all you*
81 *need*) Vaswani A. et al. 2017). C'est le début des systèmes d'**IA générative**, qui vont très vite s'imposer par leurs performances

82 spectaculaires. Comme nous l'aborderons ci-dessous, ces systèmes, dont la taille commence à se rapprocher de celle du cerveau
83 humain, réalisent sans aucun doute l'objectif de l'approche connexionniste de l'IA : grâce à leur apprentissage intensif, ils sont
84 capables de s'adapter à différentes tâches langagières ; tout se passe comme s'ils avaient acquis la capacité de comprendre les
85 demandes de leur interlocuteur et d'y répondre de façon pertinente, et donc comme s'ils maîtrisaient le sens des textes auxquels
86 ils sont soumis. Mais nous n'avons pas accès, du moins pas encore, à la manière dont s'organise ces représentations internes du
87 sens, comme cela est d'ailleurs aussi le cas pour le cerveau humain...

88 Pour conclure cette première partie consacrée à l'historique des modèles d'IA jusqu'à 2017, nous pouvons retenir trois points
89 intéressants. Premièrement un développement des approches qui a été animé par des périodes d'enthousiasme et de déception,
90 un peu comme, avec une analogie météorologique, se succèderaient des saisons chaudes et des saisons froides. Deuxièmement,
91 nous pouvons identifier deux courants principaux : les approches symboliques et les approches connexionnistes. L'IA symbolique
92 exige une définition humaine des règles et symboles, tandis que l'IA connexionniste apprend de façon autonome à partir des
93 données. Et, troisièmement, on observe une grande hétérogénéité des modèles du fait de la diversité des domaines d'application
94 de l'IA.

95 II. PROBLEMATIQUE A RESOUDRE ET ANALYSE DES CONNAISSANCES EXISTANTES

96 A. Problématique à résoudre

97 Traditionnellement, pour organiser un Congrès Lambda Mu, les communications sont réparties par thématiques par le Bureau
98 du Comité de Programme (BCP). Le thème du Congrès du LM24 étant « les métiers du risque, clés de la réindustrialisation et de
99 la transition écologique », il est apparu opportun d'expérimenter une IA afin de catégoriser les communications au sein de ces
100 différents métiers. Concrètement, il s'agissait donc d'affecter à chacun des 170 résumés reçus, le ou les métiers du risque qui lui
101 correspondent le mieux. L'idée étant de faciliter la tâche du BCP pour former des sessions cohérentes dans la suite du processus
102 d'organisation du Congrès. Notons aussi que la notion de « métier du risque » est un nouveau concept et qu'il n'y a aucune
103 connaissance disponible sur le web à ce sujet et donc aucune possibilité qu'une IA ait déjà appris sur le sujet. Il existait lors de
104 l'écriture de ce papier, seul un article du LM22 publié sur HAL (Chang et al, 2020) et un ouvrage collectif en cours de rédaction
105 et non encore publié à ce jour sur le sujet.

106 Pour résoudre un problème de catégorisation, l'approche connexionniste nécessite une phase d'apprentissage sur un corpus
107 de texte volumineux que nous n'avons pas, et cette approche n'a jamais démontré son efficacité sur des données textuelles. Il
108 convient alors de faire appel à une IA de type symbolique spécialisée dans la catégorisation. Ce type de modèle fonctionne très
109 bien, mais ... pour qu'une IA de type symbolique puisse fonctionner, il est nécessaire de lui fournir une ontologie des concepts
110 employés dans le corpus de texte. Autrement dit, il est nécessaire de lui fournir la signification de chaque mot du corpus, ainsi
111 que tous les liens qu'il existe entre les différents concepts. Cette étape n'est alors pas réalisable en l'état pour faire l'exercice
112 demandé. En effet, tout initié de l'IMdR aura au moins une fois entendu la phrase suivante : « Si vous voulez passer 2h autour
113 d'une table à débattre avec vos collègues, essayez de définir ensemble le mot risque ». Question simple, réponse difficile à
114 obtenir. En effet, comme l'ont montré les travaux du groupe de travail GiRC (Gestion intégrée des Risques et de la Complexité)
115 qui a amené le projet d'ouvrage collectif sur les métiers du risque. En effet, il existe pour chaque terme de la communauté IMdR
116 des significations propres à chaque métier du risque. Ces travaux préliminaires ont d'ailleurs abouti à la publication d'un glossaire
117 sur le site internet de l'IMdR (IMdR, 2024a)

118 Le verrou de cette recherche pourrait alors être exprimé de la façon suivante : réaliser une tâche de classification d'un corpus
119 de textes sans disposer d'une ontologie préalable sur les concepts employés.

120 B. Hypothèses à vérifier et démarche suivie

121 Afin de résoudre ce verrou, nous posons l'hypothèse qu'une IA générative est capable de saisir le sens d'un texte sans disposer
122 au préalable de l'ontologie des termes employés. En effet ce nouveau courant de l'intelligence artificielle, dont la genèse est
123 décrite dans la sous-section C1 ci-dessous, offre des possibilités nouvelles car ces modèles se concentrent sur le sens d'un corpus
124 textuel plutôt que sur la signification individuelle de chaque symbole constituant le corpus. En d'autres termes, même si des mots
125 différents sont employés pour décrire une chose similaire, l'IA générative est capable de repérer le sens commun des deux
126 formulations. C'est cette faculté qui est mise à l'épreuve dans l'exercice consistant à catégoriser les résumés des communications
127 du congrès.

128 Pour ce faire, il a été décidé de faire appel à une IA générative opérant en local sur un poste déconnecté d'internet. En outre,
129 compte tenu de la matière première que sont les résumés des communications, il a été décidé d'éviter tout risque de fuite des
130 informations traitées. De plus, travailler en local permet d'éviter de solliciter l'énergie de fermes de serveurs distants et l'eau
131 pour les refroidir et en conséquence, d'avoir un coût environnemental marginal, en sus du poste de travail, nul. Nous nous
132 sommes tournés, lors de la réalisation de l'expérimentation, vers la solution Private GPT qui est un projet open source permettant
133 d'interroger localement des documents au format texte depuis un poste informatique disposant d'une capacité mémoire
134 suffisante. Cette solution peut être alimentée par un modèle d'IA générative dont le choix est à la discrétion de l'utilisateur. Nous
135 avons opté pour le modèle produit par Mistral AI disposant de 7 Milliards de paramètres. Plus le nombre de paramètres est élevé,
136 plus le modèle est pertinent. Il s'agit pourtant d'un « petit » modèle comparé aux 175 Milliards de paramètres des modèles GPT-
137 3 d'Open AI . En effet, pour des raisons matérielles il n'a pas été possible de réaliser l'exercice avec un modèle plus grand : dans
138 notre cas nous devions déjà avoir 16 GB de mémoire RAM au minimum pour lancer le modèle. De plus, la solution Private GPT
139 fournit une API codé en Python et facilite ainsi la manipulation de corpus de texte volumineux comme les 170 résumés que nous
140 avons. Une fois l'ingestion du corpus de texte terminée, Private GPT nous permet ensuite de soumettre un prompt, c'est-à-dire

141 une question, sur ce même corpus de texte. Par exemple, nous pouvons demander un résumé de texte, une extraction de mots
142 clés ou encore la traduction du texte. C'est donc en suivant cette démarche que notre expérimentation a été menée.

143 C. Analyse des connaissances existantes dans le domaine de l'IA générative

144 1) Architecture

145 Les systèmes d'IA générative utilisent un réseau de neurones appelé transformeur (Radford et al. 2018 ; Brown et al. 2020). Un
146 transformeur est un réseau composé d'un grand nombre de couches (de l'ordre de la centaine, voire plus...), chaque couche étant
147 elle-même composée de deux modules, un module attentionnel et un module unidirectionnel. À cela s'ajoutent un module de
148 codage en entrée et un module de décodage en sortie du réseau (cf. Figure 1 ci-dessous). L'entrée est un texte qui va être segmenté
149 en tokens (ce sont pour l'essentiel des mots, mais pas uniquement). L'ensemble des tokens utilisables constitue donc le
150 dictionnaire du système ; sa taille est de l'ordre de plusieurs dizaines de milliers de tokens. **La sortie est une distribution de**
151 **probabilités sur l'ensemble des tokens du dictionnaire.**

152 Sans entrer dans les détails ici, disons que le codage consiste à représenter chaque token du texte par un vecteur de très grande
153 dimension (de l'ordre de plusieurs milliers). C'est cette suite de vecteurs qui est présentée aux couches du réseau, qui vont les
154 transformer progressivement. Le décodage consiste à repasser des vecteurs transformés à des tokens : chaque token du
155 dictionnaire reçoit une valeur qui va fournir la probabilité que ce token soit choisi en sortie.

156 Comme dans tous les réseaux de neurones, les transformations (y compris le codage et le décodage) sont déterminées par les
157 poids des connexions entre neurones. Ces poids, qu'on appelle les paramètres du réseau, sont obtenus par apprentissage, comme
158 on va le voir ci-dessous. Le nombre de paramètres de ces réseaux est énorme : de l'ordre du billion (par comparaison, le cerveau
159 humain comporte seulement mille fois plus de connexions...).

160 La grande innovation des transformeurs, c'est la présence des **modules attentionnels** dans les couches intermédiaires. Ces
161 modules permettent de tenir compte des relations à distance entre différents mots du texte, par exemple entre un verbe et son
162 sujet, ou entre un pronom et le nom auquel il réfère. Ils sont donc très puissants, mais réclament beaucoup de connexions. Les
163 **modules unidirectionnels** sont beaucoup plus classiques dans les réseaux de neurones : ils transforment chaque vecteur de la
164 suite indépendamment, et dans un module donné les différents vecteurs subissent la même transformation. À noter enfin que
165 chaque module est muni d'un "court-circuit" qui consiste à recombinaison l'entrée et la sortie du module : c'est ce mécanisme qui
166 permet d'avoir un très grand nombre de couches sans que le signal d'entrée ne s'affaiblisse trop en cours de route.

167 2) Fonctionnement

168 Le fonctionnement du réseau est simple : on lui soumet un texte, que l'on appelle une *invite* (*prompt*). On obtient en sortie
169 un token, choisi suivant la distribution de probabilité obtenue en sortie du réseau. On ajoute ce token en queue de l'invite et on
170 soumet le texte ainsi augmenté : on obtient un second token, et l'on recommence autant que l'on veut ce processus, générant
171 ainsi tout un texte qui prolonge le texte initial. Si la taille maximale de l'entrée est atteinte (elle est de plusieurs milliers de tokens
172 au moins, tout de même), on supprime un token en tête du texte chaque fois que l'on rajoute en queue un nouveau token.

173 Pour que le texte ainsi produit ait du sens, il faut que le réseau ait appris à prédire correctement quels tokens sont susceptibles
174 de suivre le texte présenté en entrée. Pour cela, il faut passer par une phase d'apprentissage, longue et coûteuse (plusieurs dizaines
175 de jours, sur les plus gros ordinateurs actuels, avec des processeurs spécialisés et beaucoup de parallélisme...), en utilisant un
176 corpus de textes tout-venant multilingue lui aussi de taille impressionnante (des milliers de milliards de tokens). L'apprentissage
177 consiste à prélever au hasard un bout de texte dans cet immense corpus, à le présenter comme invite, à comparer le token produit
178 en sortie avec le token qui suivait le bout de texte dans le corpus, et à corriger les poids du réseau en conséquence. Cette correction
179 est donnée par l'algorithme de rétropropagation du gradient, un algorithme très efficace qui est à l'origine des succès de
180 l'approche neuronale depuis le début de l'IA. On réitère la procédure jusqu'à l'obtention d'un taux de succès satisfaisant.

181 Après cette phase d'apprentissage, on constate que le système produit des textes de bonne qualité sur le plan syntaxique,
182 sémantique et discursif : autrement dit, ces textes sont bien écrits, cohérents et pertinents par rapport au texte initial. Et cela pour
183 une grande variété de langues (sans parler de ses compétences dans une dizaine de langages informatiques). Ce résultat est déjà
184 en soi assez extraordinaire : il semblait complètement inaccessible il y a à peine dix ans. Cependant il faut encore une dernière
185 phase de traitements pour que le système soit vraiment opérationnel et qu'il réponde aux demandes précises des utilisateurs. En
186 effet, celui-ci a généralement un besoin spécifique, qui peut être de nature très diverse : résumer un texte, écrire une nouvelle,
187 rédiger un rapport, écrire un programme informatique, traduire un texte, répondre à une question, commenter une information,
188 ou encore, comme dans notre cas, classer des textes par thématique... Or, comme on l'a vu, le système ne sait faire qu'une chose
189 à l'issue de son apprentissage : prolonger le texte qu'on lui a soumis de manière pertinente, mais pas forcément dans le sens
190 attendu par l'utilisateur.

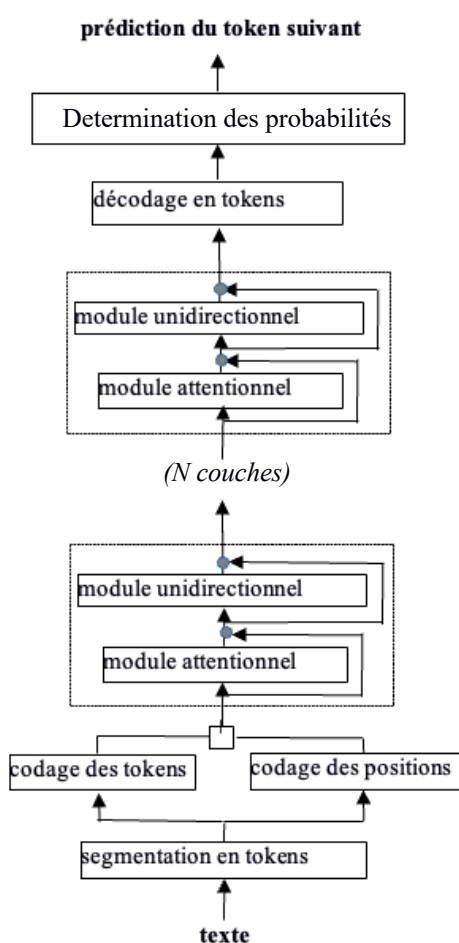
191 Pour surmonter cette difficulté, il faut une nouvelle phase d'apprentissage, beaucoup plus légère celle-là, qu'on appelle le
192 peaufinage (*fine-tuning*) (Ouyang et al. 2022). Cela consiste à noter les différentes réponses du système à une invite donnée
193 suivant qu'elles satisfont plus ou moins bien les attentes de l'utilisateur, et à modifier légèrement les paramètres de façon à
194 privilégier les réponses adaptées à sa demande. Après cette dernière étape, le système donne vraiment l'impression de
195 comprendre ce que lui dit son interlocuteur humain, quel que soit le type de demande qu'on lui adresse, comme tout un chacun
196 peut s'en convaincre en utilisant les systèmes grand public qui ont proliféré au cours de l'année 2023 : ChatGPT d'OpenAI,
197 Copilot de Microsoft, Gemini de Google, LLaMA de Meta, ou encore Le Chat de l'entreprise française Mistral. Pour une analyse

198 approfondie des performances de GPT-4, le plus récent système d'OpenAI, le lecteur intéressé pourra se référer à S. Bubeck
199 (Bubeck 2023).

200 Il faut souligner que la réussite de la dernière étape, le peaufinage, n'est possible que parce que l'apprentissage initial, centré
201 sur la tâche de prédiction du token suivant, a eu comme effet de bord, au-delà de la tâche demandée, de faire du réseau une
202 véritable machine de compréhension du langage : tout se passe comme si le système avait acquis la capacité de représenter le
203 sens d'un texte à l'intérieur du réseau, de la même manière que nos neurones sont effectivement capables de telles
204 représentations internes, ce que nous traduisons en disant que nous avons compris ce que nous lisons ou entendons.

205 Cela ne veut pas dire que ces systèmes soient sans défaut. Notamment ils peuvent très bien produire des informations fausses,
206 voire inventer complètement des événements qu'ils présentent comme s'étant réellement déroulés : c'est ce que l'on appelle le
207 phénomène d'*hallucination*. Il est très difficile de corriger ce défaut, qui est structurellement lié à la manière dont ces systèmes
208 ont été conçus. En effet, au moment où le système répond à un utilisateur, il n'a plus accès au corpus qui a servi à son
209 apprentissage, il est donc dans une situation analogue (toute proportion gardée) à celle un étudiant qui aurait essayé d'apprendre
210 par cœur une encyclopédie et à qui l'on demanderait lors d'un examen de dissertar sur tel phénomène physique ou telle période
211 historique : il est clair que l'on n'aurait pas intérêt à ne pas prendre pour argent comptant l'intégralité de ses réponses ! D'une
212 certaine manière on pourrait en conclure qu'en réussissant à modéliser (en partie) le fonctionnement du cerveau, on a développé
213 des systèmes qui ont aussi hérité de ses imperfections...

214 Fig. 1. Architecture d'un système d'IA générative : du texte à la prédiction du token suivant



215

216

III. TRAVAUX, RESULTATS ET TESTS

217 L'expérimentation s'est faite uniquement sur la base des résumés des communications. Tous les résumés comptaient à peu
218 près 500 mots en moyenne, ce qui représente une quantité d'information relativement limitée par rapport à une communication
219 complète comptant plus de 5 000 mots. Les résultats de l'expérimentation seront donc améliorables une fois les communications
220 finales récupérées.

221 Un des prérequis nécessaires avant de donner un prompt à une IA générative est de préciser le contexte dans lequel la question
222 est posée : la qualité de la réponse en dépend directement. Pour ce faire, à chaque prompt donné à l'IA générative, on fournit
223 l'en-tête suivante :

'Tu es un expert dans le domaine des risques et tu fais partie du comité d'organisation d'un Congrès dont le thème est : "Les métiers du risque : clés de la réindustrialisation et de la transition écologique". Tu parles le français. . Dans tes réponses, soit le plus concis possible.'

224

225 Demander à l'IA de donner des réponses concises est peut-être le prérequis le plus important dans notre expérimentation car,
226 comme nous l'avons vu plus haut, une IA générative produit une réponse mot à mot, de façon itérative, en tenant compte du
227 morceau de réponse imprimé à l'itération précédente. Autrement dit, au pire il peut potentiellement ne pas y avoir de fin à la
228 réponse donnée, au mieux la réponse peut être très longue : pour une simple tâche de catégorisation aucune des deux options
229 n'est judicieuse. De plus préciser le domaine de connaissance de l'IA, c'est-à-dire celui des risques, n'est a priori pas nécessaire
230 pour une tâche de catégorisation pure, mais comme nous le verrons plus bas, pour d'autres questions ouvertes ne se restreignant
231 pas à une catégorisation, cette précision peut être utile.

232 *A. Question à choix multiple : catégorisation par métier du risque*

233 La stratégie adoptée pour traiter les 170 résumés est de, non pas poser la question de catégorisation une fois sur l'ensemble
234 des textes concaténés en un seul bloc, mais plutôt de réitérer la question de catégorisation sur chacun des 170 textes. En effet,
235 avec la concaténation, l'IA ne peut connaître le début et la fin de chaque résumé. Ceci peut être résolu avec le peaufinage de
236 l'IA, notamment lui permettant de reconnaître la structure d'un template de résumé Lambda Mu, mais cela n'a pas été entrepris
237 compte tenu des délais impartis pour la rédaction de cet article. A la suite de l'en-tête présenté, voici la seconde partie du prompt
238 donné à l'IA afin de réaliser la tâche de catégorisation :

```
'Classe le document d'après son contenu dans une des catégories suivantes .'+  
  
'1. Risques, Ambigüité et Incertitude;'+  
'2. Fiabilité des composants mécaniques;'+  
'3. Fiabilité des structures;'+  
'4. Fiabilité des composants électroniques;'+  
'5. Sûreté de fonctionnement;'+  
'6. Sûreté industrielle;'+  
'7. Sécurité industrielle;'+  
'8. Facteurs humains et organisationnels;'+  
'9. Retour d\'expérience;'+  
'10. Risques professionnels;'+  
'11. Sûreté nucléaire;'+  
'12. Sécurité nucléaire;'+  
'13. Management des risques par les normes (Gestion d\'actifs);'+  
'14. Management intégré des risques (Sustainable Dependability);'+  
'15. Management intégré des risques (Méthodes et outils);'+  
'16. Enterprise Risk Management;'+  
'17. Gestion de crise.'  
  
'Au besoin, attribue une seconde catégorie si une seule catégorie ne suffit pas.'  
'Donne ta réponse adopte le format suivant : '+  
'[{fichier : {catégorie primaire , catégorie secondaire}}]'
```

239

240 Une fois l'ensemble des réponses données par l'IA générative il est possible d'en tirer un certain nombre de conclusions.
241 Premièrement, dans la forme, le formatage des réponses est respecté dans la majorité des cas. Cependant, l'IA propose parfois
242 plus de deux catégories. En inspectant les résumés correspondants on distingue deux cas : 1- le résumé mentionne effectivement
243 plusieurs thématiques présentes dans la liste des catégories; 2 – le résumé est plutôt vague et ne propose pas clairement un sujet
244 sur lequel l'IA peut statuer. Pour une petite partie des réponses on retrouve des intitulés de catégorie qui ne figurent même pas
245 dans la liste fournie. De la même manière on vérifie, en examinant les résumés concernés, que l'auteur ne s'appuie pas clairement
246 sur un corpus de connaissance identifié parmi les métiers du risque. Enfin, de manière qualitative, les propositions faites par l'IA
247 avec les catégories primaires et secondaires sont conformes dans l'extrême majorité à ce qu'un humain aurait pu en déduire. En
248 résumé les résultats obtenus sont honorables compte tenu d'une part, que la taille du modèle utilisé est très modeste et que,
249 d'autre part, il peut exister au sein des résumés l'usage d'un vocabulaire spécifique que l'IA n'arrive pas à pleinement comprendre
250 étant donné qu'aucun peaufinage n'a été préalablement entrepris, l'IA se basant uniquement sur des unités de sens de la langue
251 courante. Clairement il s'agit ici de deux points d'amélioration à mener pour de futures expérimentations de cette nature.

252 *B. Questions ouvertes*

253 *a) Quels sont les principaux thèmes abordés par les résumés ?*

254 En sus, de la tâche de catégorisation, nous avons demandé à l'IA générative d'identifier les vingt thématiques les plus
255 représentatives des résumés recueillis. Voici la liste des thématiques obtenues :

```
1. La maîtrise des risques dans la stratégie globale des organismes  
2. Références communes pour la maîtrise des risques  
3. Multiplicité et diversité des risques  
4. Identification d'un ensemble cohérent de leviers pour la maîtrise des risques  
5. Expérience des membres de l'IMdR et de ses partenaires  
6. Maîtrise des risques individuelle et collective
```

7. Résilience face aux risques
8. Assurabilité et prévention
9. Risques dans le contexte de la réindustrialisation
10. Transition écologique et maîtrise des risques
11. Stratégies de gestion des risques
12. Évaluation et analyse des risques
13. Gestion des risques dans les domaines d'activité divers
14. Ressources disponibles et capacités des organismes
15. Risques technologiques
16. Risques naturels
17. Risques financiers
18. Risques sanitaires
19. Risques environnementaux
20. Risques de cyber-sécurité

On constate premièrement que les thèmes sont globalement énoncés avec un vocabulaire connu de l'IMdR mais ne faisant pas référence explicitement à des domaines d'expertises particuliers, ou métiers du risque. Par exemple les termes "sûreté de fonctionnement", "facteurs humains et organisationnels", "ingénierie" n'apparaissent pas, alors qu'en nombre, ces trois "termes", constituent la part la plus importante en nombre de résumés. On peut expliquer cela de deux façons : 1- dans le vocabulaire commun, il s'agit d'une locution uniquement employée par des experts dans un contexte particulier ; 2 – dans les résumés, on retrouve souvent « SDF » plutôt que « Sûreté de Fonctionnement », ce qui ne facilite pas la tâche de l'IA. Il en ressort qu'il conviendrait de fournir à l'IA, en amont du traitement, une définition détaillée des termes spécifiques et la signification de tous les sigles s'y ramenant. Ce dernier point semble être important d'autant plus que les sigles sont très utilisés dans la communauté IMdR, par exemple : SDF, FOH, MBSA, ERM, etc.

b) *Quel est le thème du Congrès?*

Nous avons tenté dans un premier de fournir le prompt suivant à l'IA générative :

'En examinant le contenu de tous les résumés d'articles envoyés par les auteurs que je t'ai donné, quel pourrait être, selon toi, le thème du Congrès ?'

Les réponses n'ont pas été satisfaisantes dans la mesure où, l'IA ne considérait pas l'ensemble du corpus de texte pour fournir une réponse. En effet, en fonction de l'ordre d'ingestion des différents résumés, le thème proposé était directement extrait des quelques premiers résumés ingérés. Encore une fois, un peaufinage du modèle résoudrait ce problème.

Néanmoins nous avons proposé à l'IA d'extraire un intitulé de Congrès à partir de la liste des vingt sujets dominants fournis précédemment, sans tenir compte du corpus des 170 résumés. Après plusieurs tentatives, il en ressort deux formulations récurrentes :

'Gestion intégrée des risques : Maîtrise, résilience et prévention dans un monde complexe - Multiples risques technologiques, naturels, financiers, sanitaires, environnementaux et de cyber-sécurité'

'Maîtrise des Risques Globales : Identification, Évaluation et Gestion des Risques Technologiques, Naturels, Financiers, Sanitaires, Environnementaux et Cybernétiques dans le Contexte de la Réindustrialisation et de la Transition Écologique'

Ces deux formulations, assez complexes, intègrent bien l'étendue des divers sujets identifiés mais n'intègre pas la notion de "métiers du risque". Cela s'explique par le fait que le concept n'apparaît pas, ou très rarement dans les résumés de communications. En effet, la notion est nouvelle dans la communauté IMdR, et c'est d'ailleurs ce qui a motivé le choix d'intégrer "les métiers du risque" dans le titre effectif du Congrès pour faire ressortir l'étendue unique en France des expertises abordées sous un même toit.

Enfin, nous avons posé une dernière question à l'IA :

'Compte tenu des 20 sujets principaux du congrès que tu as identifiés, note sur une échelle de 0 à 100 le titre de congrès : « Les métiers du risque : clés de la réindustrialisation et de la transition écologique».'

Voici la réponse donnée par l'IA :

'Titre du Congrès : "Les métiers du risque : clés de la réindustrialisation et de la transition écologique"

Notes sur une échelle de 0 à 100 : 85

Ce titre est très approprié car il reflète bien le contenu des sujets abordés dans le congrès, qui traitent de manière exhaustive les différents types et aspects des risques dans un contexte industriel et écologique en évolution. Les mots clés tels que "maîtrise des risques", "réindustrialisation" et "transition écologique" permettent de capter l'attention d'un large public car ces sujets sont d'actualité. L'utilisation de la notion de "métiers du risque" est également intéressante car elle souligne l'importance de compétences

différentes dans le domaine des risques et leur rôle clé pour mener à bien une réindustrialisation durable et une transition écologique efficace.’

284

285 Il en ressort une évaluation plus que correcte de la part de l’IA concernant le titre actuel du Congrès. La remarque faite par
286 l’IA sur la pertinence de la notion de « métiers du risque » montre bien la nécessité d’identifier la diversité des apports propres à
287 chaque discipline et domaine d’expertise du risque : et cela commence nous semble-t-il par consolider le vocabulaire utilisé dans
288 chacun des métiers du risque.

289 En conséquence, au sens strict, notre hypothèse de recherche n’a pu être vérifiée en totalité à savoir, qu’une IA générative est
290 capable de saisir le sens d’un texte sans disposer au préalable de l’ontologie des termes employés. Néanmoins, en appliquant le
291 processus de décision des systèmes complexes proposé par J-L Le Moigne (Le Moigne, 1990, p131) et inspiré de celui de H. A.
292 Simon, par des actions de re-cogitation du type décision de réflexion et/ ou décision de s’informer les résultats pourraient être
293 améliorés.

294

IV. DISCUSSION ET PERSPECTIVES

295 L’intelligence artificielle générative présente certains biais : elle ne fournit notamment pas toujours les mêmes réponses à une
296 question posée. Cette particularité s’explique par le fonctionnement intrinsèque de ces outils : l’IA générative utilise des
297 algorithmes probabilistes qui analysent les données fournies et en déduisent une réponse la plus probable, mais pas forcément la
298 seule ou la correcte. Les résultats peuvent donc varier d’une requête à l’autre, même si les questions semblent similaires. Par
299 conséquent, il est nécessaire de reposer certaines questions plusieurs fois pour identifier des réponses stables et fiables. De plus,
300 une IA générative, même si celle-ci donne l’impression de pouvoir tout interpréter, ne peut pas connaître la signification de
301 termes ou de sigles issus du domaine spécifique et non utilisé dans le langage courant. En effet, toute IA générative non-peaufinée
302 n’a préalablement pas été entraînée que sur un corpus de texte généraliste et non spécifique. De ce fait, l’usage d’une IA
303 générative dans la communauté IMdR, compte tenu de ses domaines d’expertises, nécessite un peaufinage si l’on souhaite obtenir
304 des résultats plus précis. De manière générale, Il convient de ne pas se fier à 100% aux réponses de l’IA, car il existe des
305 phénomènes d’hallucination bien connus dans la littérature comme nous le mentionnions dans la revue des connaissances
306 existantes. C’est pourquoi, les réponses de l’IA générative sont une aide à la décision et non pas un substitut de l’être humain.
307 La catégorisation des résumés des communications du Congrès a d’ailleurs été réalisée initialement par des humains et après
308 notre expérimentation elle est repassée par le filtre humain.

309 L’exercice proposé ici a été entrepris en utilisant le projet Private GPT et le modèle d’IA générative entraîné par Mistral AI
310 à 7 Milliards de paramètres. Ce modèle est d’une taille relativement petite par rapport à ce qu’il existe sur le marché, mais les
311 résultats fournis sont tout à fait acceptables. Notre choix s’est porté sur ces solutions du fait qu’elles soient accessibles en source
312 ouverte, que leur mise en œuvre garantissait la confidentialité des données manipulées et permettait une utilisation en local. En
313 effet, l’IA générative fonctionnait en local sans avoir besoin de communiquer par le web. A l’avenir une piste de développement
314 serait d’étudier la mise en œuvre de la même architecture mais cette fois-ci dans le cloud avec notamment la location ou l’achat
315 d’une instance de calcul plus puissante capable d’embarquer un modèle d’IA de plus grande taille et assurant le cryptage des
316 données traitées. Une analyse coût – bénéfices devrait alors être préalablement menée de façon à juger du bénéfice au regard du
317 coût complet (financier et environnemental).

318 La tâche de catégorisation des résumés aurait été plus efficace avec une IA de type symbolique mais nous ne disposons pas
319 d’ontologie des domaines abordés dans le congrès. De plus, comme nous l’avons constaté, l’identification de la part de l’IA
320 générative de l’ontologie propre à un domaine d’expertise est satisfaisante mais non optimal. Dans des travaux ultérieurs il serait
321 pertinent de construire ces ontologies et de poursuivre les travaux qui ont mené à l’établissement du glossaire de l’IMdR sur le
322 vocabulaire utilisé par différents métiers du risque. Ces travaux permettront de mettre en valeur la richesse et la diversité des
323 métiers du risques présent à l’IMdR, mais aussi leurs similitudes. Il s’agit clairement d’une nouvelle voie de recherche pour la
324 communauté IMdR, par exemple, à travers un projet IMdR. En outre, dans l’expérimentation il apparaît également que certaines
325 thématiques, pourtant traditionnelles au sein de l’IMdR, ne sont pas ressorti souvent dans les réponses de l’IA générative. La
326 formalisation des ontologies pour chaque métier du risque, permettrait aussi d’identifier le vocable à privilégier afin d’améliorer
327 la visibilité des thématiques que l’on pourrait trouver en retrait par rapport à d’autres.

328 Enfin, même si l’expérimentation proposée avait pour but d’expérimenter l’automatisation d’une tâche dans le cadre de
329 l’organisation de ce Congrès, et de vérifier qu’une IA générative était capable, de manière satisfaisante, d’extraire l’ontologie
330 d’un domaine, celle-ci nous a permis d’en savoir davantage sur ce que l’IA peut dire de la façon de communiquer de la
331 communauté IMdR. En effet, bien que des améliorations certaines peuvent être apportées dans la configuration de l’IA utilisée,
332 nous pouvons garder en tête un point qui semble important de manière générale. Si une IA générative entraînée sur un corpus
333 du langage courant ne parvient pas à saisir la totalité du sens d’un résumé d’article, c’est aussi peut être que les auteurs, et plus
334 largement les membres de l’IMdR, experts de leur domaine qui souhaite partager leurs questionnement avec d’autres experts,
335 peuvent trouver dans la vulgarisation de leurs travaux une opportunité d’élargissement de leur public. Une communication écrite
336 à destination de la communauté devrait dans la mesure du possible reprendre des éléments de langages accessibles à tous, experts
337 ou non experts. Cela est d’autant plus pertinent si l’IMdR souhaite élargir son public cible, que cela soit en termes de nouveaux
338 adhérents ou en terme de sensibilisation envers le grand public et les institutions publiques. Cet aspect avait d’ailleurs été identifié
339 lors de l’analyse stratégique menée en 2023. Il s’agit d’un nouvel objectif potentiel qui pourrait participer au développement de
340 l’institut.

341 En termes de retour d'expérience et de perspectives, ce développement actuel de l'IA nous rappelle les travaux de H. A.
342 Simon, prix de la Banque de Suède en sciences économiques en 1978 pour ses recherches au carrefour de l'économie, de la
343 psychologie du risque, de l'informatique et de la complexité. H. A. Simon est un des pères de la systémique et il a contribué à
344 l'émergence des sciences de l'artificiel comme science de la conception. Il avait publié en 1960 un ouvrage : "the new science
345 of management décision" (Simon, 1960). Dans cet ouvrage, il discutait du rôle de l'informatique dans, les organisations, le
346 processus de décision pour des décisions programmables et non-programmables, l'évolution des fonctions des dirigeants et des
347 conséquences économiques et sociales et l'IA était déjà largement présente. Le monde d'aujourd'hui correspond assez bien à ce
348 qui était prévu. Il avait aussi envisagé que l'IA contraindra à une triple recherche : le développement d'une science de le homme,
349 la définition d'objectifs sociaux autre que le travail et la production (la Responsabilité Sociétale des entreprises et les objectifs
350 Environnementaux, Sociaux et de Gouvernance peuvent être considérés comme des exemples), une reformulation de la vision
351 que nous avons de notre place dans l'univers,... Précisons que le modèle de résolution de problèmes complexes sur lequel se base
352 le management intégré des risques (IMdR, 2024b) depuis les années 1990 a été décrit dans cet ouvrage et fonctionne encore
353 aujourd'hui plutôt bien, c'est-à-dire tout le temps, l'expérience le montre. De plus, comme il nous l'est souvent demandé : d'où
354 vient ce nom de l'Équipe N°5 de recherche de l'Institut de Recherche de l'ESTP « Architecture du risque et management des
355 décisions » ? Simple, c'est en référence à cet ouvrage. Le management des décisions n'est-il pas en fait au cœur des activités des
356 *risk managers* ? et pour l'entreprise qui souhaite faire du management intégré des risques un mode de gouvernance il est
357 nécessaire au préalable d'« architecturer » le risque d'entreprise sinon, compte tenu de la complexité des organisations, ce sera
358 in-manageable. En termes de mise en œuvre, techniquement, les approches du type *Risk Appetite Framework* vont en ce sens.

359 Nous avons vu l'IA générative est une machine à prendre des décisions sur la base d'un puissant et très technique calcul de
360 probabilités. Néanmoins nous ne connaissons toujours pas les propriétés mathématiques qui font que ça marche et nous sommes
361 toujours dans l'incapacité de savoir si le résultat d'un prompt est vrai ou pas. A l'inverse, pour les modélisations stochastiques
362 que ce soit déterministes ou multi-scénarios, avec matrices de corrélation ou pas, dépendances linéaires ou pas, nous connaissons
363 les propriétés mathématiques et nous sommes toujours, même si cela peut paraître compliqué, en mesure d'expliquer un résultat.
364 C'est le cas par exemple lorsque l'on tient compte dans les simulations de dimensions financières et extra-financières associées
365 au modèle d'affaires, au Capex/Opex ainsi qu'aux risques et opportunités. L'IA présente de grands avantages pour les
366 organisations, en matière de décisions programmables d'importants gains de productivités sont attendus. Néanmoins, pour ce qui
367 concerne les décisions non-programmables le développement d'une **IA de confiance** reste un enjeu majeur pour le management
368 intégré des risques.

369 CONCLUSION

370 Sur la base de l'état de l'art, de l'analyse des connaissances existantes fournies dans l'article ainsi que des résultats, on peut
371 conclure que les grands modèles de langage, conçus pour générer un texte de type humain basé sur des invites de saisie, présentent
372 des limites et des défis importants. Bien que ces modèles puissent apporter une aide précieuse dans diverses applications telles
373 que le service client ou la génération de contenu, ils ne sont pas infaillibles et peuvent produire des informations incorrectes ou
374 trompeuses. L'article met en lumière le problème des hallucinations, où le modèle génère des informations fausses ou fabriquées
375 qui peuvent être difficiles à distinguer de la réalité. Ce phénomène est un défaut structurel inhérent à la manière dont ces systèmes
376 ont été conçus, car une fois que le système répond à un utilisateur, il n'a plus accès aux données originales utilisées pour sa
377 formation. L'article suggère également que ce problème peut être comparé à un étudiant qui mémorise des informations mais ne
378 s'en souvient pas avec précision lors d'un examen. En conclusion, même si les grands modèles de langage offrent des possibilités
379 intéressantes pour automatiser la génération de textes et d'autres applications, il est essentiel de reconnaître leurs limites et
380 d'aborder leurs résultats avec prudence, en veillant à ce que les informations générées soient vérifiées, validées et publiées avant
381 d'être utilisées dans des contextes critiques.

382 REMERCIEMENTS

383 Les auteurs tiennent à remercier les membres du BCP du Congrès Lambda Mu 24 et le secrétariat du comité de programme
384 pour leur implication dans ce projet ainsi que la Société RISK INTELLIGENCE & DECISIONS (RiD) pour son soutien
385 scientifique, technique et logistique.

386 BIBLIOGRAPHIE

- 387 Bengio Y. & Le Cun Y. (2007). Scaling learning algorithms towards AI. Bottou L., Chapelle O., DeCoste D. & Weston J. (éds), *Large-Scale Kernel Machines*,
388 MIT Press, 321-359.
- 389 Bubeck S. et al. (2023). Sparks of Artificial General Intelligence: Early experiments with GPT-4, *arXiv:2303.12712v3*.
- 390 Brown T. et al. (2020). Language Models are Few-Shot Learners, *arXiv:2005.14165v1*.
- 391 Chang, L., Claude, F., Planchette, G., Signoret, J-P., Thing-Leo, G., et al. (2020). Risques et métiers du risque dans l'entreprise industrielle. In : Congrès
392 Lambda Mu 22 " *Les risques au cœur des transitions* " (e-congrès) - Congrès Lambda Mu 21 de l'institut de Maîtrise et de la Sécurité de Fonctionnement, Oct.
393 2020, Le Havre, France. hal-03483642.
- 394 Elman J.L (1990). *Finding Structure in Time*. Cognitive Science, 14(2), 179-211.
- 395 IMdR (2024a). *Glossaire*. <https://www.imdr.eu/>
- 396 IMdR (2024b). *Glossaire. Management intégré des risques*. https://www.imdr.eu/818_p_57387/management-integre-des-risques.html#:~:text=Le%20management%20int%C3%A9gr%C3%A9%20des%20risques,des%20cons%C3%A9quences%20et%20des%20probabilit%C3%A9s.
- 397 Koehn P., Och F.J. , Marcu D. (2003). *Statistical phrase based translation*. Proceedings of the Conference on Human Language Technologies, 127-133.
- 398 Le Cun Y. (1985). *Une procédure d'apprentissage pour réseau à seuil asymétrique*. Proceedings of Cognitiva 85, 599-604, Paris.
- 400 Le Moigne, J.-L. (1990). *La modélisation des systèmes complexes*. Paris: Dunod.
- 401 *Minsky M. (1974). A Framework for Representing Knowledge*. MIT-AI Laboratory Memo 306, MIT, Boston, Mass.

402 Ouyang L. et al. (2022). Training language models to follow instructions with human feedback, arXiv:2203.02155v1.
403 Radford A. et al. (2018). Improving Language Understanding by Generative Pre-Training. *OpenAI report*.
404 Schank, Roger (1969). *A conceptual dependency parser for natural language. Proceedings of the 1969 conference on Computational linguistics. Sång-Säby,*
405 *Sweden. 1–3.*
406 Silver D., Schrittwieser J. , Simonyan K. et al. (2017). *Mastering the game of Go without human knowledge, Nature*, 550 (7676), 354–359 .
407 Simon, H. A. (1960). *The New Science of Management Decision*. Prentice-Hall, Inc.
408 Turing A. M. (1950). *Computing machinery and intelligence. Mind*, 49, 433-460.
409 Vaswani A. et al. (2017). Attention Is All You Need, arXiv:1706.03762v5.
410