

LA TERMINOLOGIE DU DOMAINE DE L'INTELLIGENCE ARTIFICIELLE : NÉOLOGIE ET PLURIDISCIPLINARITÉ

DOI : 10.15452/SR.2024.24.0012

ORCID : 0009-0000-0078-5475

Veronika Resslerová

Université d'Ostrava

République tchèque

veronika.resslerova@osu.cz

Résumé. L'évolution rapide de la société suscite constamment l'émergence de nouveaux domaines spécialisés, parmi lesquels l'intelligence artificielle (IA) qui est un sujet omniprésent dans le monde actuel. Le présent article explore la terminologie de l'IA à partir d'une analyse du glossaire bilingue *Une intelligence artificielle bien réelle : les termes de l'IA*, disponible sur le site du *Grand dictionnaire terminologique* (en tant que ressource terminologique officielle du gouvernement québécois). L'article vise à illustrer, à travers de nombreux exemples, la manière dont les termes de l'IA sont formés et utilisés. L'accent est mis sur l'aspect pluridisciplinaire des termes ainsi que sur leur motivation linguistique (comprenant la motivation sémantique et morphologique). L'essence de la motivation sémantique des termes de l'IA repose principalement sur les métaphores et les métonymies, tandis que la motivation morphologique consiste surtout dans l'affixation qui représente un moyen essentiel en ce qui concerne la formation de nouveaux termes. Dans le cas des termes complexes, qui prédominent sur les termes simples, l'article développe le rôle du moyen le plus fréquemment utilisé pour abréger les termes dans le domaine de l'IA, à savoir la siglaison. L'article souligne également l'influence de la langue anglaise sur la terminologie de l'IA qui contient un grand nombre de calques, reflétant la tendance à l'internationalisation des terminologies techniques et l'impact des échanges interculturels sur l'évolution des termes spécialisés.

Mots-clés. Terminologie. Intelligence artificielle. Glossaire terminologique. Néologie terminologique. Motivation linguistique.

Abstract. Terminology in the field of artificial intelligence: neology and multidisciplinary. The rapid evolution of society continually leads to the emergence of new specialized fields, among which artificial intelligence (AI) stands out as an omnipresent topic in today's world. This article explores the terminology of AI through an analysis of the bilingual glossary *Une intelligence artificielle bien réelle : les termes de l'IA*, available on the *Grand dictionnaire terminologique* website (as an official terminological resource of the Quebec government). This article aims to illustrate, through numerous examples, how AI-related terms are formed and used. It emphasizes the multidisciplinary nature of the terms as well as their linguistic motivation (including both semantic and morphological motivation). The essence of the semantic motivation of AI terms mainly lies in metaphors and metonymies, while morphological motivation primarily involves affixation, which is a crucial method in the creation of new terms. In the case of complex terms, which are more prevalent than simple terms, the article discusses the role of the most frequently used method for abbreviating terms in the field of AI, namely acronymization. The article also highlights the influence of the English language on AI terminology, which includes a significant number of literal translations (*calques*), reflecting the trend toward the internationalization of technical terminologies and the impact of intercultural exchanges on the evolution of specialized terms.

Keywords. Terminology. Artificial Intelligence. Terminological glossary. Terminological neology. Linguistic motivation.

1. Introduction

La constante évolution de la société mène à une multiplication de nouveaux concepts et domaines de spécialité, ce qui va de pair avec la nécessité de créer continuellement des nouvelles terminologies. Comme le souligne Dubuc (2002, p. 4), la terminologie sert à repérer systématiquement, à analyser et à normaliser le vocabulaire pour une technique donnée. Cette caractéristique s'applique non seulement à des domaines spécialisés longtemps établis, tels que la médecine ou le droit, mais également à des domaines plus récents, y compris le domaine de l'intelligence artificielle (ci-après l'IA) que nous avons choisi comme sujet d'analyse. Le domaine de l'IA est l'un des résultats du développement technologique sans précédent et sa terminologie reflète la diversité de ses conceptions et techniques.

Le terme anglais *artificial intelligence* est né en 1956 à l'Université de Dartmouth dans le New Hampshire, grâce au mathématicien John Mac Carthy qui a émis l'hypothèse selon laquelle toute activité intelligente pouvait être décrite avec suffisamment de précision pour être simulée par une machine, et qui a proposé de désigner cette nouvelle discipline par le nom *artificial intelligence* (Otman, 2000, p. 373). Peu à peu, cette désignation anthropomorphique a commencé à se répandre dans d'autres langues. Otman (2000, p. 373) précise que l'introduction de la notion en français date de 1962.

Étant donné qu'il s'agit d'un sujet qui attire de plus en plus d'attention et qui commence à devenir partie intégrante de la vie quotidienne, l'IA peut être caractérisée de diverses façons. Nous avons choisi comme point de départ la définition de l'Organisation internationale de normalisation (2024) qui est l'une des institutions normalisatrices essentielles assurant la stabilité terminologique et qui définit l'intelligence artificielle ainsi :

Dans l'absolu, l'IA désigne la capacité d'une machine ou d'un système informatique à exécuter des tâches qui, d'ordinaire, nécessitent l'intervention de l'intelligence humaine. Cette technologie s'appuie sur des systèmes de programmation permettant d'analyser les données, d'apprendre de l'expérience et de prendre des décisions intelligentes, le tout à partir d'informations d'origine humaine.

Le présent article aborde la problématique de la transterminologisation (transfert interdisciplinaire de termes), les principaux modes de formation des nouveaux termes, la motivation linguistique reflétant la relation entre les concepts et leurs désignations (surtout la métaphore et la métonymie terminologiques, y compris l'emploi de noms propres dans la terminologie), la forme graphique des termes, le phénomène de l'abréviation dans la terminologie, ainsi que l'influence d'autres langues qui se manifeste sous forme d'emprunts et calques ; le tout à travers la terminologie de l'IA.

2. Corpus analysé

Afin d'analyser la terminologie de l'IA dans son ensemble, nous avons choisi comme corpus principal le glossaire terminologique bilingue intitulé *Une intelligence artificielle bien réelle : les termes de l'IA* et accessible sur le site officiel du *Grand dictionnaire terminologique* (ci-après *GDT*), ce dernier représentant une ressource linguistique gérée par l'Office qué-

bécois de la langue française (autorité normalisatrice canadienne instituée par la Charte de la langue française).¹ Comme le mentionne Gouadec (1993, p. 44), la Charte a confié à l'Office québécois de la langue française une double mission, à savoir la francisation des milieux de travail et la recherche linguistique et terminologique.

Dans le but de rédiger le glossaire de l'intelligence artificielle, les spécialistes de l'Office ont collaboré avec l'Institut des algorithmes d'apprentissage de Montréal, le Département d'informatique de l'Université de Sherbrooke et le Département de génie électrique et de génie informatique de l'Université Laval ; cette collaboration professionnelle attestant la conformité et la précision du lexique. En outre, dans la présentation du glossaire, l'Office québécois de la langue française précise que « le Québec constitue l'un des pôles d'expertise de la recherche en intelligence artificielle ». Le glossaire choisi compte 85 termes/concepts liés à l'intelligence artificielle.² Il convient de noter que cet échantillon de termes est actuel (sa dernière mise à jour a eu lieu en 2023) et assez représentatif, car des autres glossaires de ce type sont généralement moins exhaustifs, par exemple le *Glossaire* disponible sur le site du Conseil de l'Europe (Conseil de l'Europe, 2024) qui ne compte que 16 termes.

Nous tenons à préciser que toutes les définitions citées dans le présent article et se référant aux termes de l'IA sont tirées exclusivement de ce glossaire, à moins d'indication contraire.

3. Pluridisciplinarité des termes

Bien qu'il s'agisse d'un glossaire de l'intelligence artificielle, un grand nombre de termes qu'il contient peuvent être classés dans plusieurs domaines de spécialité. On parle ainsi de *pluridisciplinarité des termes* ou bien de *transterminologisation* au cours de laquelle des termes particuliers passent d'un domaine de spécialité vers un autre. Ce concept est souvent associé à l'homonymie ou la polysémie interdisciplinaires, puisque des termes sont interprétés différemment à travers divers domaines scientifiques (Horecký, 1977, p. 324). La notion de *transterminologisation* en tant que transfert de termes entre plusieurs domaines de spécialité est utilisée p. ex. par Bozděchová (2009, p. 56). Honová (2020, p. 83) mentionne encore d'autres notions utilisées pour désigner ce processus, telles que : *la migration des termes*, *le partage des termes*, *la notion de concepts nomades*, etc. Ce phénomène peut aussi être qualifié d'*interdisciplinarité*, ce terme étant utilisé p. ex. par Gostkowska (2010, p. 47) :

La science contemporaine est fortement marquée par le phénomène de l'interdisciplinarité. Les recherches sont de plus en plus détaillées, les spécialisations se multiplient, les disciplines traditionnelles voient sans cesse se renouveler leurs champs de recherche.

1 Il existe également d'autres glossaires en ligne consacrés à la terminologie de l'IA, mais nous ne les avons pas retenus car ils sont soit obsolètes, soit peu exhaustifs, soit proviennent de sources non officielles. De plus, certains ont pour langue d'origine une autre langue que le français, par exemple le *Vocabulário panlatín da Inteligência Artificial* (Realiter, 2013) qui est en portugais.

2 Situation au 1^{er} mai 2024. L'un des termes y est présent deux fois : *représentation des connaissances* (1) et *représentation des connaissances* (2) puisqu'il peut représenter soit le procédé qui consiste à encoder et à stocker des connaissances, soit l'ensemble des études et des techniques concernant l'encodage et le stockage des connaissances.

Le glossaire compte 85 termes dont 49 sont classés exclusivement dans le domaine de l'intelligence artificielle (p. ex. *agent cognitif*, *coefficient synaptique*, *environnement multiagent*, *perceptron*, *rétropropagation d'erreurs*, *seuil d'activation*), les autres appartenant à d'autres domaines et sous-domaines. Certains termes reflètent des sous-domaines de l'IA, comme le traitement de la parole (p. ex. *voix de synthèse*), le traitement de texte (p. ex. *reconnaissance optique de caractères*) ou le système expert (p. ex. *système expert*).

En plus du domaine de l'intelligence artificielle, de nombreux termes pluridisciplinaires appartiennent également au domaine de l'informatique (discipline très proche) : c'est le cas de 10 termes (p. ex. *analyse des mégadonnées*, *apprentissage automatique*, *interaction personne-machine*). En outre, le domaine de l'informatique compte, lui aussi, d'autres sous-domaines : informatique linguistique (p. ex. *assistant virtuel*), représentation des données (p. ex. *langage de représentation des connaissances*) ou logiciel d'application (p. ex. *assistant vocal personnel*). D'autre part, certains termes se classent aussi dans les domaines suivants : mathématiques (p. ex. *algorithme*) ; cybernétique et ses sous-domaines robotique (p. ex. *vision par ordinateur*) et type de robot (p. ex. *animat*) ; science de l'information et son sous-domaine archivistique (p. ex. *reconnaissance optique de caractères*).

À part ces domaines de caractère technique, on trouve également des termes rattachés à une discipline des sciences humaines, à savoir la philosophie (p. ex. *raisonnement abductif*) et son sous-domaine théorie de la connaissance (p. ex. *raisonnement*).

4. Motivation linguistique des termes

La motivation linguistique, comprenant la motivation sémantique et morphologique, est l'une des caractéristiques essentielles d'un terme. Dubuc (2002, p. 36) souligne qu'un terme se dit motivé au cas où à travers sa forme, il est possible d'entrevoir la notion qu'il recouvre. La motivation sémantique consiste dans une sorte d'analogie, tandis que la motivation morphologique concerne la relation entre le sens du terme et ses éléments formels (affixes, racines et d'autres morphèmes). Les procédés de composition et de dérivation facilitent la compréhension et l'homogénéité des termes. À propos de la dérivation, Holeš (2023, p. 46) postule que c'est surtout la combinaison morphologique des éléments gréco-latins qui augmente la motivation terminologique. Selon Holeš (2023, p. 46), la motivation est souhaitable dans la terminologie puisqu'elle facilite la compréhension des termes.

4.1 Motivation sémantique

Les termes utilisés dans le domaine de l'intelligence artificielle sont souvent basés sur les métaphores. Tandis que la terminologie technique est nécessaire pour décrire des concepts de manière précise, les métaphores permettent de rendre les concepts plus accessibles. En référence au *Dictionnaire de linguistique*, la métaphore est « l'emploi de tout terme auquel on en substitue un autre qui lui est assimilé après la suppression des mots introduisant la comparaison [...] ». Comme la métaphore est généralement basée sur un rapport de similitude ou d'analogie, dans le cas de l'IA et de sa terminologie, une telle analogie est souvent établie entre les capacités des machines et les capacités humaines. Parmi les

termes métaphoriques qui se répètent dans la terminologie de l'IA, on peut mentionner : *intelligence* (p. ex. *intelligence artificielle distribuée*, *intelligence artificielle faible*, *intelligence artificielle forte*), *apprentissage* (p. ex. *apprentissage automatique*, *apprentissage non supervisé*, *apprentissage par renforcement*), *raisonnement* (p. ex. *raisonnement abductif*, *raisonnement déductif*, *raisonnement inductif*), *connaissance* (p. ex. *base de connaissances*, *langage de représentation des connaissances*) ; tous ces mots se référant – avant leur métaphorisation – exclusivement à l'intelligence humaine. Dans ce cas, on parle d'*analogie de fonction* à propos de laquelle Dubuc (2002, p. 120) explique que « la comparaison ne s'établit plus sur l'aspect, mais sur l'usage de l'objet ». En ce qui concerne la terminologie de l'IA, ce type d'analogie permet d'exprimer la ressemblance entre les machines et les hommes.

En outre, Kocourek (1991, p. 170) mentionne les *métaphores anthropomorphiques* qui sont basées sur la ressemblance avec des parties du visage ou du corps humain. De telles analogies entre les capacités des machines et les capacités humaines peuvent être observées même à travers les notes et définitions de certains termes (dans le même corpus). C'est le cas du terme *assistant virtuel* : « on attribue généralement à l'assistant virtuel un nom, une voix ou un avatar pour lui donner un aspect plus humain », ou bien du terme *neurone artificiel* : « à l'origine, les inventeurs du neurone artificiel se sont inspirés du neurone biologique en tentant de lui donner un modèle mathématique ». Il est évident que cette dénomination a été motivée par le neurone biologique (humain) en tant que réalité déjà existante, le terme devenant assez transparent quant à son sens.

4.2 Noms propres dans la terminologie

Dans le contexte de la motivation des termes, il faut remarquer qu'il y a aussi des noms propres qui pénètrent dans la terminologie. Parmi les critères traditionnels déterminant un nom propre, Leroy (2004, p. 7) énumère les suivants : le nom propre s'écrit avec une majuscule initiale, ne se traduit pas, ne figure pas dans les dictionnaires, se construit sans déterminant, ne se fléchit pas, est dépourvu de sens, renvoie à un référent unique. Cependant, certains noms propres n'apparaissent pas toujours sous leur forme de base (voir exemple *algorithme* ci-après).

Les noms propres se présentent surtout sous forme d'éponymes qui sont souvent considérés comme un cas particulier des expressions métaphoriques et métonymiques. Selon Bernet (1990, p. 1258), un éponyme « donne son nom à un lieu, à une personne ou à une chose ». Généralement, les dénominations éponymiques sont basées sur des noms de personnes réelles – le plus souvent des auteurs qui ont été les premiers à décrire ou à découvrir le phénomène donné (Bozděchová, 2009, p. 77).

Parmi les termes éponymes dans le domaine de l'IA, mentionnons le *test de Turing* qui a été inventé par le mathématicien britannique Alan Turing. Le *GDT* précise qu'il s'agit d'un « test qui consiste à mettre en communication, à l'aveugle, un être humain et un ordinateur afin de vérifier s'ils sont capables d'atteindre les mêmes niveaux de performance ». L'origine de ce terme basé sur un nom propre est liée exclusivement au domaine de l'IA. Cependant, en dehors des termes formés au sein du domaine spécialisé donné, on distingue aussi

des termes existant dans la langue depuis longtemps mais qui pénètrent dans le domaine spécialisé au fil du temps. De tels termes sont aussi basés sur un nom propre, mais de manière plutôt implicite puisqu'ils ont généralement subi plusieurs changements orthographiques au fil des siècles. C'est le cas du terme *algorithme* qui est une altération d'un nom de personne. Sur le plan de son étymologie, le *Dictionnaire historique de la langue française* le décrit ainsi :

algorithme [...] désigne d'abord, sous la forme *augorisme* (v. 1230), puis *algorisme* (XIII^e s.), le calcul en chiffres, l'arithmétique. Ces formes viennent de l'ancien espagnol *algorismo*, alors que la forme moderne calque le latin médiéval *algorithmus*, altération sous l'influence de *arithmetica* du nom du grand mathématicien arabe *Hwarizmi* (= originaire du *Hwarizm*, au sud de la mer d'Aral), inventeur du terme qui a donné *algèbre*.

4.3 Motivation morphologique

En ce qui concerne la motivation morphologique, on y classe les termes dérivés et les termes formés par composition. Lors de la création de nouveaux termes, la dérivation par affixation représente un moyen important qui permet de former des nouveaux termes à partir de mots déjà existants dans la langue française. À ce propos, Rondeau (1984, p. 135) parle de la *dérivabilité* en tant que capacité d'un nouveau terme à « se présenter, à l'aide de la préfixation ou de la suffixation, sous différentes formes de catégories grammaticales ».

On peut constater que la plupart (52) des 85 termes sources sont dérivés ou contiennent un ou plusieurs mots dérivés. Les préfixes les plus représentés dans le glossaire de l'IA sont : *multi-* (p. ex. *perceptron multicouche*), *méga-* (*analyse des mégadonnées*), *inter-* (*interaction personne-machine*), *mono-* (*perceptron monocouche*), *co-* (*coefficient synaptique*), *rétro-* (*rétropropagation d'erreurs*) et *sur-* (*surapprentissage*). Ces préfixes sont d'origine savante (latine ou grecque) et s'intègrent facilement avec diverses racines, permettant ainsi de créer de nouveaux termes à travers différentes disciplines scientifiques, tout en conservant leur clarté. Grâce à leur caractère uniforme et universel (leurs significations sont largement reconnues et bien ancrées), ils facilitent la compréhension interlinguistique ainsi que le processus de transterminologisation. Dans la note relative à tous les termes construits avec *multi-* (préfixe le plus fréquent dans notre corpus), le *GDT* précise : « Bien que *multi-* possède le sens de « plusieurs », les mots construits avec ce préfixe suivent en principe la règle générale d'accord en nombre : ils s'écrivent sans *s* au singulier et avec un *s* au pluriel. L'adjectif *multiagent* est également invariable en genre. ». La fréquence élevée du préfixe *multi-* dans la formation des nouveaux termes, surtout sur le plan des sciences techniques, est aussi attestée par Holeš (2023, p. 50) qui le compare dans son étude avec les préfixes synonymiques moins fréquents *poly-* et *pluri-*³. D'un autre côté, parmi les suffixes les plus fréquents dans la formation des termes liés à l'IA, on compte : *-age* (p. ex. *apprentissage automatique*), *-ance* (*reconnaissance de la parole*), *-ation* (*seuil d'activation*), *-iciel* (*logiciel d'intelligence artificielle*), *-ment* (*raisonnement abductif*), *-tique* (*robotique*), *-eur* (*synthétiseur de la*

3 Aucun de ces deux préfixes ne figure dans notre corpus.

parole) et *-isme* (*connexionnisme*). Certains termes peuvent contenir à la fois un préfixe et un suffixe, par exemple *coefficient synaptique* (préfixe *co-* et suffixe *-tique*). Les termes formés avec ces affixes n'ont souvent pas été créés spécifiquement dans le domaine de l'IA, mais il s'agit de mots préexistants qui fonctionnent comme éléments de termes complexes spécifiques à l'IA.

Selon Honová (2020, p. 58), dans la langue française, la composition est un procédé terminogène moins fréquent que la dérivation, ce qui se manifeste également dans le corpus source. Notre corpus ne contient que 4 éléments formés par composition (qui font tous partie des syntagmes terminologiques) reconnus par le trait d'union, à savoir : *interaction personne-machine*, *interaction personne-robot*, *synthèse texte-parole*, *synthétiseur texte-parole*.

5. Forme graphique

5.1 Termes simples et complexes

Les termes (en tant qu'unités lexicales) peuvent coïncider dans leur forme avec des unités de langue générale et peuvent correspondre à des structures plus complexes, c'est-à-dire des syntagmes (L'Homme, 2005). En fonction du nombre de lexèmes dont un terme dispose, on distingue les termes simples (une seule entité graphique) et les termes complexes (plusieurs entités graphiques). Selon Honová (2020, p. 62), les termes complexes – ou bien les syntagmes terminologiques – sont beaucoup plus nombreux que les termes simples, ce qui est aussi le cas dans notre corpus. Nous pouvons constater que le glossaire de la terminologie de l'intelligence artificielle ne compte que 14 termes simples, les autres (71) appartenant à la catégorie de termes complexes.

Étant donné que la forme graphique d'un terme peut faciliter sa compréhension, la transparence sémantique relève (entre autres) de la complexité d'un terme. Honová (2020, p. 62) postule que « plus spécialisé est le terme, plus longue est sa forme ». Il est vrai que le glossaire de l'IA compte de nombreux termes dont la transparence sémantique est élevée grâce à leur forme complexe. À titre d'exemple, prenons les termes *agent* et *agent physique autonome* (tous deux sont présents dans le glossaire source). Le *GDT* enregistre le terme *agent* dans plusieurs domaines spécialisés, y compris philosophie/psychologie (« être ou entité qui est auteur d'un acte ») ou industrie minière (« tout corps qui produit un effet ») : il faudrait donc préciser le contexte afin d'utiliser et de comprendre ce terme de manière correcte ; tandis que le terme *agent physique autonome* n'est rattaché qu'au domaine de l'IA et sa définition est plus précise (« agent physique qui détermine lui-même ses actions et son état interne, sans l'intervention d'un tiers »).

À propos de la structure syntaxique, Bozděchová (2009, p. 66) parle du « caractère nominal » des termes complexes, précisant que ce sont généralement soit deux substantifs, soit un substantif et un adjectif, qui forment les termes complexes. Bien que cette constatation concerne la langue tchèque, les mêmes structures syntaxiques se manifestent également dans le glossaire français de l'IA, qui contient dans la plupart des cas des termes avec la structure substantif-substantif (p. ex. *base de connaissances*) ou substantif-adjectif

(p. ex. *neurone artificiel*). Les termes complexes dans le glossaire sont constitués de deux, trois, ou au maximum quatre éléments (sans compter les mots grammaticaux), p. ex. *environnement multiagent*, *intelligence artificielle faible*, *réseau de neurones à propagation avant*.

5.2 Abréviation (siglaison) des termes complexes

L'un des inconvénients des termes complexes consistant en plusieurs éléments graphiques est leur mauvaise maniabilité dans les textes. Par conséquent, le besoin de les réduire en utilisant des abréviations apparaît. Bien qu'un terme abrégé devienne opaque et que sa motivation soit réduite, la fréquence des abréviations dans la communication spécialisée est élevée par rapport à la communication courante (Bozděchová, 2009, p. 68). Il convient de constater que la forme d'abréviation la plus fréquente dans la terminologie de l'intelligence artificielle est la siglaison.

À propos des sigles, Humbley (2018, p. 151) explique que dans le cas de mots séparés, c'est normalement la première lettre de chaque mot qui est retenue, sans prendre en compte les mots grammaticaux. Fréquemment, un seul terme peut disposer de plusieurs sigles qui sont acceptables : p. ex. le terme *base de connaissances* peut être siglé soit sans initiale de la préposition comme *BC*, soit avec la préposition comme *BDC*. Il peut arriver qu'un terme ne soit siglé que partiellement : p. ex. le terme *intelligence artificielle distribuée* peut être abrégé non seulement comme *IAD*, mais aussi comme *IA distribuée*.

Au cas où il existe plusieurs variantes acceptables d'un seul terme, il peut arriver qu'il en existe également plusieurs sigles. P. ex., les équivalents du terme *traitement automatique des langues* sont *traitement automatique du langage naturel*, *traitement automatique des langues naturelles*, *traitement automatique de la langue naturelle*, *traitement du langage naturel* et *traitement des langues naturelles*. Étant donné que tous ces termes sont acceptables, il est possible de les remplacer par les sigles *TAL*, *TALN* ou *TLN*.

Dans certains cas, on trouve aussi des formes adaptées et substantivées de l'anglais, comme *OCR* (en anglais : *optical character recognition*) pour *reconnaissance optique de caractères*. Le *GDT* précise qu'il est possible en français de dériver un sigle à prononciation alphabétique et que ce procédé s'accompagne généralement d'une transcription graphique de la prononciation du sigle. Par conséquent, les sigles acceptables pour le terme *reconnaissance optique de caractères* sont *ROC*, *OCR*, ou même le terme dérivé du sigle *océrisation*.

6. Néologie externe : emprunts et calques

Après avoir déjà abordé l'influence de la langue anglaise, il faut constater à ce propos que les nouveaux termes ne sont pas créés de manière isolée et uniquement à partir des ressources linguistiques propres à la langue française. Grâce à de nombreux contacts entre les langues liés à la mondialisation, il faut mentionner aussi la néologie externe, c'est-à-dire les emprunts et les calques. À propos de l'emprunt terminologique, Humbley (2018, p. 322) parle d'un conflit constant entre le souci de la sauvegarde et de la défense de l'homogénéité des langues d'une part et la tendance à l'internationalisation et à l'uniformisation des terminologies d'autre part.

Bien que les institutions normalisatrices françaises luttent systématiquement contre les emprunts (notamment contre les anglicismes), il est évident que c'est surtout l'anglais qui a une influence majeure sur les termes de l'intelligence artificielle puisque la plupart des concepts liés à l'IA proviennent des États-Unis⁴ qui sont en quelque sorte le berceau du développement technique. Kocourek (2001, p. 167) mentionne dans le cadre de la formation des termes un parallélisme entre le français et l'anglais, qui est causé par le « caractère plutôt international des terminologies technoscientifiques d'aujourd'hui ».

Nous pouvons constater que ce sont des calques, définis par Cabré (1998, p. 166) comme « constructions inédites réalisées sur le modèle d'une construction existant dans la langue d'emprunt », qui constituent la plupart des termes sources présents dans le glossaire de l'IA. Il s'agit de termes qui ont été créés selon un modèle anglais⁵, mais en utilisant les éléments linguistiques français et en respectant les règles grammaticales françaises (comme l'ordre des mots dans le cas de syntagmes terminologiques) : *algorithme d'apprentissage* – *learning algorithm*, *approche multiagent* – *multiagent approach*, *couche cachée* – *hidden layer*, *perceptron monocouche* – *single-layer perceptron*, *raisonnement déductif* – *deductive reasoning*, *reconnaissance optique de caractères* – *optical character recognition*, *seuil d'activation* – *activation threshold*, etc.

Dans certains cas, le GDT indique aussi des équivalents dans d'autres langues romanes qui utilisent également des calques, s'appuyant ainsi sur les mêmes modèles anglais : *expert system* (anglais) – *système expert* (français) – *sistema experto* (espagnol) – *sistema expert* (catalan) – *sistema especialista* (portugais) ; *natural language* (anglais) – *langage naturel* (français) – *lenguaje natural* (espagnol) – *llenguatge natural* (catalan) – *linguagem natural* (portugais) ; *voice recognition* (anglais) – *reconnaissance de la parole* (français) – *reconocimiento del habla* (espagnol) – *reconeixement de la parla* (catalan) – *reconhecimento da fala* (portugais).

7. Conclusion

Bien que le domaine de l'intelligence artificielle semble être très récent (étant donné son essor dans la société actuelle), la naissance de l'IA remonte aux années soixante du XX^e siècle et dans le contexte de la terminologie, Rondeau (1984, p. 2) le mentionnait déjà dans son *Introduction à la terminologie* il y a une quarantaine d'années :

L'homme a inventé, principalement au cours des cinquante dernières années, un nombre extraordinaire de machines, produits et procédés destinés à l'accomplissement d'activités jusque-là inconnues [...]. De nouveaux champs de recherche et d'application de disciplines

4 Parmi les termes de l'IA d'origine américaine, on peut mentionner à titre d'exemple *perceptron*, inventé par l'informaticien Frank Rosenblatt qui a créé le concept ainsi que le terme qui le désigne. Notre corpus comprend aussi deux termes créés à partir du mot *perceptron* : *perceptron monocouche* et *perceptron multicouche*.

5 Outre les termes provenant de l'anglais et un terme d'origine arabe (*algorithme*), le corpus contient aussi un mot international qui a son origine dans la langue tchèque, à savoir *robot* : « [...] tiré de *robota*, < travail pénible, corvée >, et utilisé pour désigner un androïde capable de faire certains travaux à la place des hommes » (Académie française, 2024). Le glossaire de l'IA comprend également deux termes créés à partir du mot *robot* : *robotique* et *interaction personne-robot*. Néanmoins, les termes *algorithme* ou *robot* existaient déjà avant la création de l'IA, tandis que les calques anglais susmentionnés ont un lien direct avec l'IA.

traditionnelles se sont créés, comme la physique nucléaire, la biochimie, la chimie alimentaire, la transplantation d'organes, l'intelligence artificielle, etc. [...]

Par la suite, Rondeau (1984, p. 2) a précisé que tous ces nouveaux développements provoquent « une avalanche de notions nouvelles qui, dès le moment de leur conception, reçoivent une ou parfois plusieurs dénominations dans une langue ». L'un des objectifs du présent article a été d'observer comment on attribue les dénominations à ces « notions nouvelles » dans le cadre de ce domaine qui est moins récent qu'il ne semble, mais qui est certainement l'un des plus dynamiques dans la société contemporaine.

Malgré la spécialisation du glossaire, la classification des termes n'est pas catégorique puisqu'un grand nombre des termes sont classés dans plusieurs domaines (et sous-domaines) spécialisés. Plus précisément, 49 termes appartiennent exclusivement au domaine de l'IA et 36 termes sont classés dans d'autres domaines aussi. Parmi les disciplines qui partagent certains termes avec le domaine de l'IA, on compte l'informatique, les mathématiques, la cybernétique, la science de l'information et la philosophie (ainsi que leurs sous-domaines).

La motivation sémantique comprend l'emploi figuré (métaphores et métonymies), y compris le cas particulier d'éponymes. En plus de l'emploi figuré, la terminologie de l'intelligence artificielle se développe en profitant de différents procédés linguistiques, en particulier de la dérivation par affixation. Les affixes les plus représentés sont *multi-*, *méga-*, *inter-*, *mono-*, *co-*, *rétro-*, *sur-*, *-age*, *-ance*, *-ation*, *-ciel*, *-ment*, *-tique*, *-eur* et *-isme*. En tout, le corpus contient 52 termes qui sont dérivés ou contiennent un ou plusieurs mots dérivés.

Quant à la forme des termes, le nombre des termes complexes dépasse celui de termes simples, ce qui est logique puisque les termes complexes (ou bien les syntagmes terminologiques) assurent précision et transparence sémantique. Dans son ensemble, le corpus contient 71 termes complexes et 14 termes simples. Étant donné que les termes complexes comprennent plusieurs éléments graphiques, il convient souvent de les abrégier dans le texte. La forme d'abréviation la plus fréquente dans la terminologie de l'IA est la siglaison, qui utilise les initiales des mots dans le but de former un nouveau terme.

Outre la néologie se produisant au sein de la langue française, l'influence de la langue anglaise est aussi évidente. Cependant, au lieu d'anglicismes en tant que tels, dans la majorité des cas on rencontre des calques. Les calques ne sont pas en contradiction avec la politique linguistique française, puisque les mots anglais ne sont alors qu'une base pour former des nouveaux termes en utilisant les ressources linguistiques françaises. Comme le glossaire comprend également des équivalents éventuels dans d'autres langues romanes (espagnol, catalan, portugais), nous pouvons observer que des calques basés sur des modèles anglais s'y manifestent aussi.

Bien que les racines de la terminologie de l'IA remontent au passé, celle-ci constitue actuellement l'un des domaines les plus dynamiques en matière de développement terminologique. En effet, ce secteur connaît non seulement l'émergence de nouveaux

concepts et outils techniques (tels que de nouvelles applications logicielles), mais aussi un élargissement de son champ d'application (dans la médecine, l'éducation, etc.). Il est donc à prévoir que des ensembles de termes, tels qu'*Une intelligence bien réelle : les termes de l'IA*, puissent subir des transformations radicales dans les années à venir (certaines unités risquent de devenir obsolètes, tandis que d'autres émergeront). Pour cette raison, le présent article se propose de résumer le status quo à un moment donné de l'évolution de l'IA.

Bibliographie

- » Académie française (2024). *Dictionnaire de l'Académie française* (9^e éd.). <https://www.dictionnaire-academie.fr/>
- » BERNET, C. (1990). Le dictionnaire d'éponymes. Dans F. J. HAUSMANN, O. REICHMANN, H. E. WIEGAND, et L. ZGUSTA (Éds.), *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie. An International Encyclopedia of Lexicography. Encyclopédie internationale de lexicographie* (Vol. 2, pp. 1258-1260). De Gruyter.
- » BOZDĚCHOVÁ, I. (2009). *Současná terminologie (se zaměřením na kolokační termíny z lékařství)*. Karolinum.
- » CABRÉ, M. T. (1998). *La terminologie : théorie, méthode et applications*. Les Presses de l'Université d'Ottawa.
- » Conseil de l'Europe (2024). *Glossaire*. <https://www.coe.int/fr/web/artificial-intelligence/glossary>
- » DUBOIS, J. (1994). *Dictionnaire de linguistique*. Larousse. <https://ia904501.us.archive.org/12/items/DictionnaireDeLinguistiqueDubois/Dictionnaire-de-Linguistique-Dubois.pdf>
- » DUBUC, R. (2002). *Manuel pratique de terminologie* (4^e éd.). Linguattech. <https://doi.org/10.1515/9782920342804>
- » GOSTKOWSKA, K. (2010). Le transfert des termes entre les disciplines scientifiques : sur l'exemple du génie biomédical. *Romanica Wratislaviensia*, 3228, 47-58.
- » GOUADEC, D. (1993). *Terminologie et phraséologie : acteurs et aménageurs*. Maison du Dictionnaire.
- » HOLEŠ, J. (2023). Compétition entre *poly-*, *pluri-* et *multi-* dans les néologismes officiels français. *Kalbotyra*, 76, 42-53. <https://doi.org/10.15388/Kalbotyra.2023.76.3>
- » HONOVÁ, Z. (2020). *Le terme dans sa variabilité : dans la perspective contextuelle en français contemporain*. Université d'Ostrava.
- » HORECKÝ, J. (1977). Vzťah jazykovednej terminológie k terminológii iných vedných odborov. *Kultúra slova*, 11(10), 321-325.
- » HUMBLEY, J. (2018). *La néologie terminologique*. Lambert-Lucas. <https://doi.org/10.7202/1070546ar>
- » KOCOUREK, R. (1991). *La langue française de la technique et de la science. Vers une linguistique de la langue savante* (2^e éd.). Oscar Brandstetter Verlag.
- » KOCOUREK, R. (2001). *Essais de linguistique française et anglaise : mots et termes, sens et textes*. Éditions Peeters.
- » LEROY, S. (2004). *Le nom propre en français*. Ophrys.
- » L'HOMME, M.-C. (2005). Sur la notion de « terme ». *Meta*, 50(4), 1112-1132. <https://doi.org/10.7202/012064ar>
- » Office québécois de la langue française (2024). *Grand dictionnaire terminologique*. <https://vitrine-linguistique.oqlf.gouv.qc.ca/>

- Office québécois de la langue française (2024). *Une intelligence bien réelle : les termes de l'IA*. <https://www.oqlf.gouv.qc.ca/ressources/bibliotheque/dictionnaires/vocabulaire-intelligence-artificielle.aspx>
- Organisation internationale de normalisation (2024). *Intelligence artificielle : comprendre les bases et son importance*. <https://www.iso.org/fr/intelligence-artificielle>
- OTMAN, G. (2000). Le Vocabulaire de l'informatique et des technologies de l'information et de la communication. Dans G. ANTOINE et B. CERQUIGLINI (Éds.), *Histoire de la langue française 1945–2000*. CNRS, 371-396.
- Realiter (2024). *Vocabulário panlatín da Inteligência Artificial*. <https://www.realiter.net/wp-content/uploads/2013/06/InteligenciaArtificial.pdf>
- REY, A. (1998). *Dictionnaire historique de la langue française*. Dictionnaires Le Robert.
- RONDEAU, G. (1984). *Introduction à la terminologie* (2^e éd.). Gaëtan Morin.

Veronika Resslerová

Filozofická fakulta
Ostravská univerzita
Reální 5
701 03 OSTRAVA
République tchèque