Ministry of Education and Science of Ukraine
Taras Shevchenko National University of Kyiv
Educational and Scientific Institute of Philology
Department of English Philology and Intercultural Communication

Master's thesis

Artificial Intelligence Terminology

Daria Pytel
2nd year student of the Education Program
'English Communication Studies and Translation
and Two Western European Languages'
Field of science: 03 "Humanities"
Specialty: 035 "Philology"

Supervised by:

Irina Alexeyeva, PhD

«Допущено до захисту»
Протокол засідання кафедри англійської філології
та міжкультурної комунікації
Протокол № 8 від 18.04.2023
Зав. кафедри _____________д. філол. н., проф. Алла БЄЛОВА

KYIV 2023

Contents

Introduction

Language constantly develops in order to satisfy the demand created by social and historical changes as well as technological advancements. Whereas changes on grammatical and phonological level are gradual and slow, lexical level of language dynamically adapts to the new reality constantly updating and replenishing. Such processes as globalization, industrial and economic development impacted all spheres of our life including language.

The Lexical system of language is not homogeneous as it comprises different strata heterogeneous in their nature, purpose and domain of use, degree of formality, origin. These factors affect how quickly and by what means a certain stratum of language is replenished, and also how it is regulated. While words of daily use may appear and grow obsolete before ever being added to any dictionary, other words are consciously elaborated to name newly discovered concepts, novel creations and events. In this respect, terminology proves to be a unique stratum of language as it comprises highly specialized words and expressions that are specific to particular fields, disciplines, or professions; and are carefully regulated. Terms play a crucial role in facilitating communication, knowledge sharing, and collaboration within and across different fields and disciplines.

Each scientific field, discipline or technology has its own terminology. Depending on the sphere, terminologies vary in their volume and pace of growth. One of the new and dynamically developing fields today is Artificial Intelligence. This technology revolutionized the field of computer science by enabling machines to perform tasks that would normally require human intelligence, such as visual perception, speech recognition, natural language processing, decision-making, and problem-solving. This has led to the development of a wide range of applications and technologies, from autonomous vehicles and smart homes to personalized medicine and predictive analytics. All of these led to an emerging new niche of terminology. In this paper we are going to have a look at the AI terminology, history of its formation and means of the term creation.

The **topicality** is stipulated by the role of newly created AI terms in the enrichment and expanding of English terminology in the sphere of technology and by the lack of versatile research of the means of AI terminology formation in English.

The **objective** is ascertaining and analysis and the existing means of English terminology coinage in the field of artificial intelligence.

The achievement of the designated objective is will be fulfilling of the following **tasks**:

- determining the theoretical and methodological principles of studying English terminology in general and in the field of AI;
- establishing ways of AI terms creating in English;
- analysis of their linguistic and cultural specificity.

The **object** of study is English terminology of artificial intelligence.

The **subject** of research is the means of English terminology coinage, and their linguistic and cultural specifics.

The **researched material** is 310 terms in English that signify concepts related to AI selected and compiled from 9 English dictionaries [130; 131; 132; 134; 135; 140; 141; 144; 145], 2 encyclopedias [23; 98], 6 technical glossaries in English [133; 137; 138; 139; 142; 143;], 49 media sources [2; 6; 7; 8; 9; 11; 12; 13; 16; 17; 18; 20; 24; 27; 32; 34; 35; 36; 39; 40; 44; 47; 49; 50; 51; 52; 53; 54; 55; 68; 72; 73; 75; 81; 83; 86; 87; 88; 90; 92; 93; 94; 106; 108; 109; 114; 115; 116; 117]; 30 academic publications dedicated to various aspects and concepts within the field of AI [3; 10; 33; 37; 38; 45; 56; 57; 58; 60; 61; 62; 64; 66; 69; 70; 71; 74; 77; 78; 79; 82; 85; 102; 103; 105; 107; 112; 122] and 7 literary sources [15; 41; 43; 46; 59; 80; 101;].

In the process of research the following **methods** were used: critical analysis, description, generalization, giving examples - to highlight the theoretical and methodological foundations of the study of terms; word-formation, structural, semantic, analysis of part of speech - to establish methods of creation, structural features and characteristics according to the part of speech; semantic and linguistic and cultural analysis - to determine the linguistic and cultural specificity of the specified units; as well as elements of quantitative analysis - to provide quantitative data on the methods of their creation.

Scientific novelty of the achieved results rests on the fact that it is the first attempt to carry out complex analysis of artificial intelligence English terminology from the point of view of lexicology and culturology.

Theoretical significance of this academic research lies in the fact that it contains actual material about word coinage models of modern English and their productivity. It also broadens the knowledge and understanding of language processes related to cognition, perception, and culture. Thus, the work is a valuable contribution to practical and theoretical bases of lexicology, cognitive linguistics and culturology.

The **practical importance** of the results obtained is determined by the possibility to use the findings in the course of Practical English as well as in theoretical and practical courses in Lexicology, Cognitive Linguistics, Structural Linguistics, Discourse Analysis, Cultural Studies, Semantics, Sociolinguistics, Translation Studies. The findings may also contribute to developing thesauruses and glossaries.

Personal contribution is outlined by clarification of the methods of English terminology creation and linguistic and cultural features of AI terms.

The MA degree paper consists of **Introduction**, which briefly summarizes the topicality, novelty, object, subject-matter, the material of the research, its objective and tasks, practical value, and scientific novelty.

Chapter 1 deals with the general concept of terminology: it explores definitions, properties, and gives an overview of means of terminology coinage. The chapter also gives a short overview of the history of AI and historical perspective on terminology formation; approaches to classification of AI terms; and gives an outline of the current trends in the field of AI.

Chapter 2 dwells on the part of AI terminology coined by means of morphological derivation, and provides the data on their productivity and recurrent components.

Chapter 3 focuses on the part of AI terminology coined by means of morphological derivation, and provides the data on their productivity and recurrent components.

General conclusions contain the main theoretical and practical findings of the paper.

The **Appendix** contains the Glossary of AI terms used in the research.

A perspective of the current research may be a more profound analysis of AI terminology that would include terms that will appear in the future.

CHAPTER 1. THEORETICAL BASIS OF ARTIFICIAL INTELLIGENCE TERMINOLOGY FORMATION

1.1. Terminology as peculiar strata of language

Precision in language is crucial for effective communication, particularly in fields like technology and science where clarity is paramount. Naturally, scientific and technological advancement has made it necessary to elaborate unique pieces of vocabulary and extend the existing ones.

Naming in the area of natural sciences, namely chemistry, zoology, botany, medicine and mathematics was started in the late nineteenth century as mentioned by Sonneveld and Loening [100: p.2]. However, the true starting point was industrialization**.** It became the true leading force in terminology creation, which stirred scholars of the 18th and 19th centuries to attempt standardization of technical and scientific language. Emerging concerns regarding the nature of the new words resulted in terminological work having been organized into special fields. [29: p.7]

Terminology became a proper subject of linguistic studies in the 20th century. *Term* and *terminology* were put in the center of linguistic studies [113: p.11].

Term is traditionally defined as a single word or phrase, designating the concept of a specific subject field or discipline. According to expert estimates, the number of terms in specialized texts accounts for 30-80% of the vocabulary and the amount of scientific knowledge constantly expands. [22: p. 2254]. Thus, terminology serves as a tool for transferring and acquiring scientific knowledge.

Therefore, terminology represents the most dynamic and mobile lexical system of the language, therefore research in this field is often historically oriented, showing how terminology arises, develops and changes over time depending on the development of relevant sciences and the general thinking style of the era.

1.2. Definitions of term and terminology

Before outlining the place of terminology in the lexical stratum of language, it is necessary to establish proper definitions of *terminology* and *terms*.

Sonneveld and Loening (Sonneveld and Loening, 1993) defined terminology as systemized assembly of concepts on the basis of established principles and methods as well as activity associated with the process of accumulating and systematizing [100: p. 2]. Juan C. Sager put forward the following explanation: ‘Terminology is concerned with the study and use of the systems of symbols and linguistic signs employed for human communication in specialized areas of knowledge and activities.’ He also stated that terminology is a linguistic discipline [95: p. 3-4] Maria Teresa Cabré points out that terminology is not seen unanimously as a separate discipline, claiming that it is merely ‘a practice dealing with social needs…often related to political and/or commercial ends’, whereas others defend the idea that it is a true scientific discipline. [29: p.6]

A. V. Superanska, N. V. Podilska, and N. V. Vasylieva define several meanings of the term *terminology*:

1) a set of general scientific terms;

2) a set of terms (concepts and names) of any specific field of knowledge (construction terminology, medical terminology, etc.);

3) study about education, composition and functioning of general scientific terms;

4) teaching about the formation, composition and functioning of the terms of a certain field of knowledge, used in a certain language, and their equivalents in other languages;

5) general terminological teaching. [128: p. 94]

Based on these definitions, we can say that the first two definitions are associated with the system of concepts of a certain field of knowledge, and in the last three, the concept of terminology correlates with the concept of teaching – the science of a system of concepts. Thus, we can distinguish the existing definitions of terminology according to two aspects: 1) a set of term elements, concepts and names; 2) teaching about the creation of term elements, concepts and names.

In the 21st century, some scholars emphasized the importance of new technologies and the changing nature of language. The definitions became more multidimensional. Rodolfo Maslias (2009) described terminology as the science and practice that studies the principles and methods of term creation, acquisition, management, dissemination, and use in specialized communication in any medium, in any language, in any subject field, and in any subject area of human activity [76: p. 2-5].

Term also has already defined boundaries of its meaning, though the definitions of it may slightly vary in different dictionaries. Oxford English Dictionary defines *term* as 'a word or group of words expressing a distinct notion, especially one which recurs in different contexts or is used as an element in the construction of a larger expression' [142].

Jennifer Pearson defined the *term* following traditional terminologists, namely as 'labels for concepts which are abstract entities isolated from text. The term which is agreed upon may be a single word or a multiword unit'. [4: p.1034].

Eastern European linguists favored definitions that drew on its applied nature. In its historical development, the concept of *term* was understood as:

- 'a word that is a name of a strictly defined concept'(Volin and Ushakov 1940); 'a word or a collocation that expresses a concept of some special science, technology, art, social life, etc.' (Bazhan) [123: p.2];
- 'a word or a collocation of special (scientific, technical, etc.) language that is created (received or borrowed) for accurate expression of specific concepts and notations of specific objects' (Akhmanova) [123: p.2];

- ‘a specially cultivated word being artificially invented or taken from natural language’ (Superanskaja) [123: p.2];
- ‘a word or a collocation being the exact name of a special concept for any field of science, technology, production, social political life, culture, etc.’ (Zhovtobrjukh) [123: p.2].

Modern definition of *term* comes from the Merriam-Webster Dictionary, which defines it as ‘a word or expression that has a precise meaning in some uses or is peculiar to a science, art, profession, or subject’ [140].

A term may consist of one word as well as several. The process of establishing the concept of a term is long and diverse.

To avoid confusion between a *term* and a *word,* we will look at several definitions of the latter and point out the main differences.

Cambridge dictionary defines *word* as ‘a single unit of language that has meaning and can be spoken or written’ [130];

Collins dictionary stresses one other feature of it: ‘A word is a single unit of language that can be represented in writing or speech. In English, a word has a space on either side of it when it is written’ [131];

Dictionary.com offers a broader definition: ‘a unit of language, consisting of one or more spoken sounds or their written representation, that functions as a principal carrier of meaning. Words are composed of one or more morphemes and are either the smallest units susceptible of independent use or consist of two or three such units combined under certain linking conditions, as with the loss of primary accent that distinguishes the one-word *blackbird* (primary stress on ‘black’, and secondary stress on ‘bird’) from *black bird* (primary stress on both words). Words are usually separated by spaces in writing, and are distinguished phonologically, as by accent, in many languages’ [135].

So, whereas a term contains a word in most of its modern-day definitions, it is restricted to a specific field of knowledge, and may be structurally diverse and consist of a word combination that refers to a concept. Sager, an American computational linguist, highlighted the difference in the following way: '[terms are] the items which are characterized by special reference within a discipline, and collectively they form its 'terminology'; those which function in general reference over a variety of sublanguages are simply called *words* and their totality is the 'vocabulary' [95: p.19].

As terminologies in any language are the depiction of the underlying knowledge system, the designation of a term must reflect the concept it represents in a consistent and accepted manner to facilitate expert communications.

Terms serve a critical function in many fields, including science, technology, medicine, and law, among others. Its primary goal is to provide a standardized, precise, and consistent language that facilitates communication and understanding among professionals, researchers, and practitioners within a specific field.

Terms' functions are to provide help to experts and professionals within a field to communicate effectively and efficiently. This standardization helps reduce the likelihood of misunderstandings or errors, leading to more accurate and precise communication.

Moreover, terms help streamline communication, reducing the need for explanations or clarifications, leading to increased efficiency and productivity within a field. It provides a way to organize and classify information, making it easier to retrieve, analyze, and use in various applications.

Standardized lists of terms make it possible to maintain consistency and quality in research, publications, and other forms of documentation within a field. It also supports effective knowledge management, which is critical in fields where information is constantly evolving.

Terminology forms an autonomous section of the lexical level of language. In general, it belongs to the number of integrating factors that allow creating a single informational (scientific-technical, educational, economic) space, since it is terminology that ensures information mutual understanding and mutual exchange at the national and international levels, compatibility of legislative, legal and normative documents. Transmission of professional information, stylistically neutral layers of vocabulary are extremely necessary and have a different functional specificity within the framework of the language of science. [128: p. 23-95]

The analysis of the definitions of *terminology* and *term* in theoretical sources allows the following interpretation of these notions that will be used in the further research:

Term is a word or group of words used to express a specific idea or concept within a certain field or subject area.

Terminology is

1) a set of terms that are used in a particular field or subject area;
2) A branch of science that studies terms, their creation, specifics and use in a variety of domains.

1.2.1. Research methods in terminology

Every well-developed science has two basic methods that complement each other. For example, physics operates with the experimental and theoretical methods: experimental results are always compared with the theoretical predictions, and the theory is derived based on the empirical facts. Linguistic methods similarly can be of two kinds: experimental and theoretical ones. Terminology employs statistical and analytical research methods. The comprehensive use of the well-defined investigation methods in terminology manifesting the shift from terminology-teaching to terminology-science is equivalent to the transition from critical discourse to the puzzle solving inherent to science [65: 6-7].

The statistical method that determines what is customary to establish the presence of a linguistic fact regardless of its accuracy and correctness, and assumes accumulation of such facts. The principal components of this method are descriptive and observation method and statement stage in a number of linguistic methods (e. g., in the method of grammatical analogies and comparative method). Though this method does not provide a fault finding assessment of the results and transfer of acquired information to qualitatively new knowledge, and therefore needs further interpretational work.

The analytical method that determines 'what is right', provides critical methodical analysis and allows one to discover scientific validity and feasibility of a given linguistic unit (including lexemes and, particularly, terms) or operation mode of specific rules. The method's components are: induction and deduction, idealization and formalization, method of hypotheses, falsification method, taxonomization, transformation stage of the analysis and synthesis method, comparative historical method (reconstruction technique, relative chronology, historical and etymological analysis), structural method (opposition, distribution, transformation, component and string analysis, method of immediate constituents), functional method (lingual, pragmatic, conversation, contextual interpretive, discourse analysis, methods of functional semantic fields modeling), typological, comparative, lingual statistical method, method of acoustic invariants and others [124: p. 18].

The given overview of the field leads towards following conclusions:

- Progress in science and technology led to the emergence and growth of a peculiar stratum of language – terminology, an aggregation of terms to serve a specific field of knowledge. Terms are characterized by high precision and defined domain of use. In the 20th century terminology became an academic discipline and an object of studying. Today both term and terminology have been comprehensively defined that allows one to have a legit understanding to carry out academic research.

- *Term* is different from a *word* as it may differ structurally and is limited to a specific domain.
- To study terminology, we can use statistical and analytical methods. These methods contain a variety of components, i.e., observation, comparison, etymological analysis, discourse analysis, etc.

1.3. Word-building patterns of terminological systems

The swift advancement of scientific knowledge has a significant linguistic effect and given the need to elaborate new concepts and terms. Terminology related to science and technology is characterized by continuous development of new units of language due to the rapid development of these fields. Terms are often the most important words of the message, as they present scientific and technical concepts revealing the semantic meaning of the message. The process of coinage employs various methods in order to meet this need. In the Vocabulary of Science, Lancelot Hogben focuses on the traditional ways of coining terms, namely by using Greek and Latin roots. Hogben disapproves of modern scientists coming up with terms that he claims are not rational and threatens the past and future vocabulary of science. [48: 3-7] However, contemporary linguists do not rarely resort to using ancient languages.

The process of creation, defining and systemizing new terms should be based on certain rules. It is important to establish whether any term labeling a specific concept occurs in a specific subject field or. If there are multiple words that can be used to describe a concept, it's important to choose the word that best fits the definition and is widely accepted in the relevant field. [96: p.174]. Furthermore, terms should reflect the features of the concepts they refer to as much as possible as they are the linguistic representations of the concepts. Relationships between concepts and terms are based on linguistic appropriateness, linguistic economy, and derivability. In order to eliminate misunderstanding among scientists belonging to different cultures, terms should be clear, concise, unambiguous and consistent. [110: p. 4].

The process of new term formation is based on the existing vocabulary and is aimed to invent new concepts to serve scientific knowledge [22: p.1]. Term formation is influenced by many factors including the level of the subject field development, the nature of the persons involved in this process, the structure of the language, and the stimulus affecting term formation [110: p. 3]. It is determined by synonymic, hyponymic and antonymic relationships as well as by the process of conversion [126: p. 57]. Among the primary ways of term formation, the following ones can be determined:

– semantic narrowing of the meanings of the commonly used words, which have the varied meanings in different contexts or terminologization, e.g., *side* (of a triangle), *line* as mathematical terms;

– terminological derivation, that is, the use of word-formation means, e.g., *hyperfunction*, *microchip*;

– borrowings from other languages e.g., *morphography* [96: p. 58].

A large number of terms are created through reverse terminologisation [96: p. 173]. Some terms that have recently been known only to specialists turn into a category of commonly used words if this technology becomes the wide-spread. This process can be viewed as determinologisation. The process can be observed in many fields like medicine, law or technology, where people from non-professional backgrounds interact with professional terminology on the receiving end. These interactions result in terms like *infarct, processor, gigabits* being acquired by common knowledge.

Professional jargon sometimes also makes its way into terminology glossaries. Jargon words may be included into the subject field dictionaries and considered as professional terms because they have lost their eccentric meanings and can be used to convey the most precise meanings of certain specific concepts or procedures referring to the subject field. An example of this jargon-term transition may be words like *cookie,* which may have originally been a jargon word in the context of computer programming and engineering, but since has become a widely recognized term that is no longer exclusive to that field.

Terms are often neologisms, which means they are newly coined words or phrases that have not been widely used or recognized before. Neologisms are created to fill a gap in language where there is no existing term to describe a new concept or phenomenon. This is the reason for the use of a large number of neologisms by specialists in their subject field speech, which designate scientific and theoretical concepts and require for further scientific definitions at a certain stage of the subject field development. Consequently, lexical neologism is included into the subject field dictionary as a result of the discovery of a new concept. Neologisms can be divided into three groups as following ones:

1) proper neologisms (new lexical units are introduced in order to name new objects or processes);

2) lexical neologisms (new lexical units appear to denote already known objects or processes);

3) semantic neologisms (existing lexical units obtain new or change their initial meaning).

English terminology is formed by two primary means. They are either morphological (affixation, compounding, conversion, blending, back-formation, acronyms) or lexical-semantic ways.

Affixation is one the most common ways of creating neologisms [129: p. 90]. New words are created by adding prefixes or suffixes to their roots. Among the most popular prefixes, which are used to form new terms in English for Technologies are the following ones: *mini-; macro-; inter-; super-; multi-; hyper-* (e.g., *hypertext*, *hypermedia*); non- (e.g., *nonvolatile* , *nonbuffered*). Prefixes *dis-*, *mis-*, *over-*, *out-* , *re-*, *un-* are also widely used. The most productive suffixes are *-ing; -tion; -er; -or*.

Compounding is another tool to coin new word-formation. In this case, two or more already existing words are combined in a new word. In terms of its semantic motivation, these kinds of words are related to the phrases on the basis of which they have been created [130]. Its constituents have mainly denotative meanings, and due to their combination, the meaning of the whole phrase resides in one word. Compounds can be distinguished as endocentric (which have a head, e.g., database) and exocentric (without a head, e.g., upgrade). [129: p. 90]

A certain number of terms are introduced into the language by means of **borrowing**. This occurs when technological development advances in certain countries, other languages prefer adopting a foreign term instead of coining their own counterparts. Some of the borrowings have successfully come into use in the target language while others have been replaced at a later stage by more compliant words with the linguistic rules of the target language [129: p. 90], for example, *algorithm* is borrowed from the Arabic word "al-Khwarizmi"; avatar - borrowed from the Sanskrit word ‘avatāra’. [136]

Conversion or **zero derivation** is a way of word-formation when a word is created without any change in its form. The most productive forms of conversion are verb to noun conversion (e.g., to plug in – plug-in) and noun to verb conversion (e.g., access – to access). One more way of formation is reversion or back-formation, that is, derivation of terms by removing affixes (e.g., to multitask from multitasking) [129: p. 90].

Blending is another way of word-formation. A new word is created by combining different parts of different words. These new words have the features of both words and can be easily recognized. (e.g., netiquette is networked + etiquette; Fortran is formula + translation) [118].

There are many terms, which consist of two or more words. Therefore, **acronyms** are widely used to simplify the process of professional communication or reading, for example, *PDF* [118].

Typical for terminology is also the use of multi-word terms consisting of symbols and acronyms or abbreviations (e.g., *ChatGPT-3*).

Another way to shorten an originally complex term is **clipping**. A long word is cut to be one or two-syllable words. While it may often make the new word appear less formal, many terms created like this found their place in formal glossaries (e.g., *demonstration* = *demo*) [118].

Portmanteaus is a method that combines both clipping and blending. It can be described as compounding with a twist. Clipping of one word is combined with another full word or another clipped one. This way terms like *paratroops*, *internet* entered the lexical strata of language [21].

Some neologisms come to a language as a result of emerging new **semantic meanings** of the already existing language units. A word from one context and is applied to another. Thus, the crane, meaning lifting machine, got its name from the long-necked bird, and the computer mouse was named after the long-tailed animal. Such repurposing happens due to **metaphors** and **metonymies**. Thus, such commonly used words as bug, wallpaper, to hang, etc. have acquired new meanings in professional speech of various specialists. Metonymy is a tool that allows one to express the sense of an object by means of the use of the word designating the other object, which is closely interrelated with the first one. Such a relationship can be seen between the subject and the material, which the subject is made from (plasma/plasma screen;); between the process and its result (to interpret/interpretation); between the action and the tool of the action (browse/browser). There are also wide-spread examples when a term denoting a part of something refers to the whole of something or vice versa. [129: p. 90]

Conceptual metaphor is another tool in word-creation. Observation of correspondences between conceptual domains result in cross domain mappings – seeing and speaking of one thing in terms of another, e.g., speaking of machines as sentient beings. Such metaphors play an important role in the creation of AI terminology, for example, we often use the metaphor of a *brain* to describe work of a computer, or the metaphor of a *neural network* to describe the way that machine learning algorithms work. Conceptual metaphors interact with each other and can give rise to relatively complex metaphor systems. These systems are collections of more schematic metaphorical mappings that structure a range of more specific metaphors [42: p. 286-288].

Sub-atomic physics provides fascinating evidence of how new terms can be formed by traditional means of coinage, e.g., *lepton* – derived from a Greek root; and also employ semitraditional approaches like in *boson* – the term was created to honor S.N. Bose. The root of the term is a proper name with the attached suffix -on. This is an example of **eponyms** - words named after a person or place (e.g., Alzheimer's, atlas, cheddar) The issue of whether, and for how long, to retain the capital letters on eponyms remains unresolved [21].

Different methods prevailed in different times. Speaking of English, replenishment of dictionaries was in direct connection with its geopolitical and activity and international position. For long periods (1100-1500 and 1650-1900), there was a trend on loanwords from French. In the 19th century, borrowings from Indian languages started to inflow the dictionaries. There was also a brief exchange with Dutch and Flemish [21].

In the 20th century, a significant number of terms were generated by derivation, and suffixation. Abbreviations was another method that thrived due to the necessity in wartime of delivering messages as soon as possible. This passion for initialisms seems to be in decline. Partly, this can be attributed to further difficulty to avoid confusions, for example, *PC*, can now mean po*litically correct, police constable, per cent, personal computer, parsec, post cibum, peace corps, postcard, professional corporation or printed circuit* [21].

Summarizing the above-mentioned, we can conclude that terminology systems are continuously and dynamically developing and new neologisms appear to denote new objects and processes. These terms are coined according to the existing rules of formation in each language. The most productive ways of term formation in the English language can be distinguished as follows: affixation, compounding, conversion, acronyms. Lexical-semantic way of word-formation is used in both languages. So, it is a promising direction for further research.

1.4. Definition and brief history of AI

Artificial intelligence is the general name of the technology for the development of machines, which are created entirely by artificial means and can exhibit behaviors and behaviors like human beings, without taking advantage of any living organism.

AI is generally assumed to be a computer and is now ubiquitous in spheres like economics, business, medicine, education, engineering, and the military, they also can be found in common PC software applications, traditional strategy games like chess and other video games [14: p. 2].

The creation of AI was anticipated and envisioned long before it was introduced. Stanford University historian Dr. Adrienne Mayor, in her book 'Gods and Robots' argues that Greek myths speculated on the future existence of artificial intelligence, robots, driverless cars, and other modern technologies. Ancient Greece gave origins to some words like *automaton*, that can be regarded as proto-terms of modern AI [5].

In 1884 Charles Babbage, English mathematician, worked on a mechanical machine that could exhibit intelligent behavior**.** He coined and popularized some terms relevant not only to math but also to computation, e.g., *analytical Engine* – his famous invention [12].

In the XX century, the world saw the appearance of the first computer which gave birth to a whole new sphere and thus, new terminology to serve it. In the early 1950s, there were various names for the field of 'thinking processing' [91: p. 28-31]. One of the earliest terms used was *cybernetics*, which was coined by Norbert Wiener in the 1940s [117: p. 7].

The variety of names suggests the variety of conceptual orientations. The name *Artificial Intelligence* was picked partly for its neutrality; avoiding a focus on narrow automata theory, and avoiding cybernetics which was heavily focused on analog feedback [87: p.115]. It emerged in the process of preparation for Dartmouth conference (1956) – known as the first academic conference on the subject, proposed by American Computer and cognitive scientist John McCarthy and his fellow colleagues: Alan Turing, Marvin Minsky, Allen Newell, and Herbert Simon. The ten scientists gathered at the conference in an attempt to glimpse into how machines can be taught to use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves [77].

John McCarthy can be fairly regarded as the father of the term Artificial Intelligence. He also coined several other names and terms related to the field of AI. In 1958, he published an article Programs with Common Sense, where he described the *Advice Taker* – program for solving problems by manipulating sentences in formal languages [78: p. 5].

- *Advice taker* - a hypothetical computer program, proposed by John McCarthy in his 1958 paper "Programs with Common Sense". It was probably the first proposal to use logic to represent information in a computer and not just as the subject matter of another program [139].

McCarthy was the person behind the programming language *Lisp* (List Processing), which is still used today for AI research and development. He came up with a term of *time-sharing*, a technique that enables multiple users to access a computer simultaneously by dividing the CPU time among them; and *Garbage Collection*, a technique for automatic memory management in computer programs [40]. McCarthy contributed to the development of functions, and was the one who introduced the term *recursive function*. Another of his brainchild was the idea of *symbolic expressions* in programming languages, which are used for representing data and computations [40].

Another important visionary and theoretician was Alan Turing. In 1936, the British mathematician demonstrated that a universal calculator, also known as the Turing machine, is possible. Turing's central insight was that if any problem can be presented in a form of an algorithm it may be solved by a machine. If cognitive processes can be split into finite well-defined individual stages, they can be executed on a machine. He designed and developed *Colossus*, *one* of the first programmable computers. And only a few decades later, the first practical digital computers were actually built [14: p. 3].

In his paper 'On Computable Numbers, with an Application to the Entscheidungsproblem,' published in 1936, Turing introduced the concept of a *universal machine*, which is now known as a universal *Turing machine*. This theoretical machine is able to simulate the logic of any other computing machine, and is considered a foundational concept in computer science [107: p. 230-265].

In 1950, Alan Turing created a test to determine whether a machine was intelligent. He put forward an idea that a machine can hold a proper conversation without being suspected of its artificial nature. This test is now known as *Turing's test*.

Many names in the field of AI reveal that researchers endowed or saw human-like agency in computer programs. Such names as *the logic theorist* (Logic Theory – LT), introduced in 1956, may be an example of it. The developers behind it, Neweell, Shaw and Simon, made a program capable of solving mathematical problems. This system is now regarded as the first AI system [72]. Another example of terms that earlier belonged to the domain of solely human knowledge is *expert systems* (the term was coined by researchers at Stanford University) were designed to replicate the decision-making processes of human experts in specific fields. These systems were used in a variety of applications, including medical diagnosis and financial forecasting [24: p. 3-14].

Researchers like Margaret Masterman focused on creating schematic networks for machine translation. She made significant contributions to the field of computational linguistics, which involves using computers to process and analyze natural language. Margaret Masterman was a strong advocate for using both linguistic and non-linguistic information in machine translation, and argued that the best translation systems would be those that combined knowledge of the structure and meaning of language with information about the world [91: p. 274-278].

The invention of the Internet had a significant impact on the development of AI. The internet made vast amounts of data available to AI researchers and developers, which was crucial for training and testing machine learning algorithms. Large datasets such as ImageNet, which contains millions of labeled images, would not be possible without the Internet [53].

The internet has also had a significant impact on natural language processing (NLP), a subfield of AI that deals with the interaction between computers and human language. The availability of large amounts of text data on the Internet, as well as tools such as search engines and translation services, has enabled the development of sophisticated NLP algorithms that can understand and generate human language with increasing accuracy [53].

1993 became the year of creation of the first human-looking robot, *Cog* at MIT (the name implies the cognitive abilities the robot was designed to possess.). It was designed to mimic human movement and behavior in a way that was previously not possible with traditional robots [90: p. 45].

Toward the end of the centuries there were supercomputers capable of defeating human champions in chess. The computer’s name *Deep Blue* was chosen to evoke the image of the deep blue sea [51: p. 126-127], AI powered games entered the market. but it wasn't until the mid-2000s that computer programs began to surpass human experts in other games such as Go and poker, using techniques such as Monte Carlo *tree search*, called so because it involves exploring a tree-like structure to find a solution to a problem [33: p. 72-83].

In the early 2000s, machine learning techniques such as *support vector machines (SVMs)* and *decision trees* became popular for tasks such as image and speech recognition. In the latter half of the decade, deep learning methods such as *convolutional neural networks* (*CNNs*) and *recurrent neural networks* (*RNNs*) emerged, enabling breakthroughs in areas such as *natural language processing* (NLP) and *computer vision* [101].

Between 2010 and 2020, there were significant developments in the field driven by advancements in machine learning, deep learning, and neural networks. *Deep learning*, a subset of machine learning, became a major focus in AI research. With the development of powerful GPUs (graphics processing units), deep learning models became more feasible, and their accuracy in various applications, such as image and speech recognition, improved significantly [64]. *NLP* saw significant advancements during this period, leading to the development of *chatbots*, *virtual assistants*, and *machine translation systems* that could understand and generate human-like language [57]. *Reinforcement learning* became more prominent in the AI field, with the development of *AlphaGo*, an AI system that defeated the world champion in the board game Go. Reinforcement learning is a type of machine learning that focuses on agents learning through *trial and error* [38]. AI-driven robotics continued to advance, with the development of robots capable of performing more complex tasks, such as *autonomous driving* [3]. *Big Data* was introduced when the amount of data available for AI applications increased exponentially, leading to the development of large-scale AI models that could handle massive amounts of data [103].

- *Big data* – large or complicated for general data processing applications to handle [133].

This was also a period when concerns regarding ethics appeared. The ethical implications of AI started to worry many, with a growing recognition of the potential for AI systems to perpetuate biases and reinforce inequalities. This led to an increased focus on developing ethical AI systems [11]. One of the things that give way to such concerns were generative models, such as *GANs* (*Generative Adversarial Networks*) [45: p. 2672], that brought to life the generation of realistic images and other content [36].

These days AI includes various advanced systems such as *Neural Network and Fuzzy Systems*. These programs are used in typical problems such as *Pattern recognition*, *Natural language processing* and more. Spheres, that employ AI and benefit the most from it is computer games, expert systems (computers programmed to make decisions in real-life situations e.g., medicine); robotics and voice assistants. [14: p. 4]

- *Neural network* – is a computer system that functions as the brain of a human. Although researchers are still attempting to construct a computer model of the human brain, current neural networks can already accomplish many things regarding speech, vision, and board game strategy [133].
- *Natural language processing (NLP)* – is the capacity of computers to comprehend or extract meaning from natural human languages. NLP generally entails computer interpretation of text or speech recognition [133].
- *Fuzzy logic* – a logic operations method based on many-valued logic rather than binary logic (two-valued logic). Two-valued logic often considers 0 to be false and 1 to be true. However, fuzzy logic deals with truth values between 0 and 1, and these values are considered as intensity (degrees) of truth [144].
- *Pattern recognition* – a computer algorithm whose goal is to discover patterns and trends in data [133].

Every year new AI finds a new application and revolutionizes whole industries. Systems like Midjourney (2022) let the world see beautiful pictures, many would not recognize as non-human made. It stirred upheaval in the community of some artists by the threat of stealing their style and replacing them in their workplace [52].

ChatGPT and its several upgrades (ChatGPT-3, ChatGPT-4) have already become ubiquitous and assist in a variety of spheres: IT, science, journalism and business. It is a large language model developed by OpenAI based on the GPT (*Generative Pre-trained Transformer*) architecture. It was trained on a massive amount of text data using unsupervised learning methods to generate human-like responses to various prompts and questions [32].

- *OpenAI* - a for-profit technological company that conducts scientific research in artificial intelligence and machine learning. ChatGPT, a conversational AI platform, is their latest invention. Developed on the principle of reinforcement learning, ChatGPT is equipped with advanced AI capabilities to complete human-dependent tasks [144].

In conclusion, we can see that artificial intelligence was a long-awaited technology that only emerged in the previous century. It has a short but very dynamic history of development due to its revolutionary and omnipresent contributions in all spheres. The technology created an entirely new strata of terms that are continuing to be created right now. In contrast to other fields of mathematical study, the domain of artificial intelligence has engendered the creation of numerous technical terms that reflect human cognitive processes. These terms are used to label computer programs and the tasks they perform, and include memory, learning, and expertise, among others. This nomenclature mirrors the structure of human cognition and highlights the aspiration of AI research to emulate and expand upon human intelligence.

In chapters II and III we will look closer at the ways the new words appear and their incorporation in the already existing glossaries and dictionaries.

1.4.1. AI terminology in dictionaries

Major modern English dictionaries like OED, Cambridge, Collins, or Marriam-Webster have their own definitions of AI and the key terms that belong to the field.

The first dictionaries to include AI terminology were likely specialized technical dictionaries focused on computer science or artificial intelligence. Dictionaries began including AI terminology in the late 20th and early 21st centuries as the field of AI became more prominent and important in various industries. One of the earliest examples of a dictionary including AI terminology is the 'Handbook of Artificial Intelligence', published by Avron Barr and Edward A. Feigenbaum in 1981 [15: p. 3]. This book was a comprehensive overview of the field of AI and included a glossary of terms. Barr and Feigenbaum defined AI as 'the part of computer science concerned with designing intelligent computer systems, that is, systems that exhibit the characteristics we associate with intelligence in human behavior – understanding language, learning, reasoning, solving problems, and so on'[15: p. 3]. The handbook also explores terms like *Game tree search, Alpha-beta pruning, Semantic networks*, and many others.

Another early example of a dictionary including AI terminology is the 'Encyclopedia of Artificial Intelligence', edited by Stuart C. Shapiro and published in 1987. This book included over 500 entries on various AI topics, including machine learning, expert systems, and natural language processing. He described AI as 'a field of computer science and engineering concerned with the computational understanding of what is commonly called intelligent behavior, and with the creation of artifacts that exhibit such behavior. This definition may be examined more closely by considering the field from three points of view: computational psychology, computational philosophy, and machine intelligence [98: p. 89], the encyclopedia also contains entries on *artificial life*, *heuristic programming, subsymbolic ai*, and more [98: p. 91].

As AI has become more mainstream and widely used, AI terminology has been increasingly included in general dictionaries, for example Merriam-Webster, Oxford English Dictionary [141], Cambridge Dictionary include definitions for AI terminology such as *artificial intelligence*, *machine learning*, *neural network*, *algorithm*, *deep learning*, *natural language processing*, *robotics* and more.

The definitions may vary slightly in different dictionaries. For example, Mariam Webster has a twofold definition for AI. It is ‘a branch of computer science dealing with the simulation of intelligent behavior in computers’ and also ‘the capability of a machine to imitate intelligent human behavior’ [140];

Cambridge dictionary, describes AI as ‘the study of how to produce machines that have some of the qualities that the human mind has, such as the ability to understand language, recognize pictures, solve problems, and learn’ [130];

Oxford English Dictionary offers the following definition: ‘The capacity of computers or other machines to exhibit or simulate intelligent behavior; the field of study concerned with this. Abbreviated *AI*’ [141].

Collins English Dictionary: ‘AI (noun): the use of computer systems to perform tasks that would normally require human intelligence, such as visual perception, speech recognition, decision-making, and language translation’ [131].

There are also specific dictionaries and online resources that specialize in technology and computer-related terms including AI. Techopedia [144], TechTerms [145], and Computer Dictionary Online [132] are among them. Such dictionaries usually give more broad definitions and are likely to use other technological terms to explaining the concept, as well as linking them to other relevant terms, for example this is Computer Dictionary: ‘[AI is] the subfield of computer science concerned with the concepts and methods of symbolic inference by computer and symbolic knowledge representation for use in making inferences. AI can be seen as an attempt to model aspects of human thought on computers. It is also sometimes defined as trying to solve by computer any problem that a human can solve faster’.

It is difficult to determine exactly how many AI terms exist, as the field of artificial intelligence is constantly evolving and new terms are being created all the time. However, there are likely thousands of AI terms and concepts that have been defined and used in the field over the years.

The AI terms can be divided into more common terms that are widely used and are presented in big dictionaries (e.g., *neural networks, natural language processing*). However, there are many other more specialized terms that are used in different subfields of AI, (Reinforcement Learning, Conversational AI, Prompt Engineering), that have not yet entered dictionaries [84].

Also new additions are constantly added. There are some examples of recent additions:

- *ChatGPT-3* – a language model developed by OpenAI that uses deep learning techniques to generate human-like text [25: p.5];
- *Transformer Architecture* – a neural network architecture that has been used in many recent advances in natural language processing, including GPT-3. The Transformer architecture uses self-attention mechanisms to allow the network to focus on different parts of the input sequence [111: p. 9];
- *Generative Adversarial Networks (GANs)* – a type of neural network used in unsupervised learning, where two neural networks are trained together to generate realistic images or other data. One network generates synthetic data, while the other network tries to distinguish between the synthetic data and real data [45: p. 2672];
- *Federated Learning* – a machine learning technique that allows multiple devices to collaboratively train a machine learning model without sharing their data with each other. This approach is useful for applications where data privacy is a concern [56: p. 4-10].

Certain AI terms may become outdated and obsolete over time as new technologies and techniques emerge. However, it's important to note that many older

AI terms still hold value and are used in various contexts [18]. That being said, here are a few AI terms that may be considered less relevant or less commonly used today, like *expert systems, rule-based systems, symbolic AI* [59: p. 413]

To summarize, AI terminology has been included in dictionaries for several decades, with the earliest examples appearing in specialized technical dictionaries. Soon following the rapid development of AI its terminology appears in the general dictionaries, including Merriam-Webster, Oxford English Dictionary. There are thousands of AI terms and concepts that have been defined and used in the field over the years, including recent additions. There are also terms that were widely used in the previous decades, but currently may be considered obsolete. As AI research continues to progress, we can expect to see the development of new terms and concepts that will be added to these dictionaries and resources.

1.4.2. Classification of AI terms

Technical terms can be classified in various ways depending on the context and purpose of classification. Here are some common ways technical terms can be classified:

Scientific and technical terms can be classified according to their structure, meaning, usage and ways of formation. Terms can be divided into four main categories due to the fields of knowledge and activities, in particular, general science and common technical terms, subject field and field specific terms [129].

All these groups can be distinguished among the terms of AI: **subject field terms**, which are used to denote concepts referring only to AI (e.g., *cybernetics*, *operational memory*, *computation* etc.); **field specific terms** that designate concepts and processes related to the specific subfield (e.g. *chatbot, deep learning, data mining,*); **technical terms**, which are common to all technical fields (e.g.,), and general science terms that refer to all fields of Science and Technology and denote logical and philosophical categories (e.g., system, formulae, element, program, function, etc.). Moreover, AI terms can be related to a variety of fields in science, mathematics (*calculus, graph theory*), statistics (*regression analysis*), cognitive psychology (*perception, memory*), linguistics (*pragmatics, corpus linguistics*), philosophy(*ethical AI*), and neuroscience (*neural networks, brain imaging*).

Technical terms can be also classified based on their level of abstraction. Terms such as *variable* and *function* are abstract technical terms used in programming and mathematics, while terms such as *CPU* and are concrete technical terms used in computer hardware.

Another approach in classification is based on their complexity. Some technical terms are straightforward and easy to understand, while others are more complex and require a deeper understanding of the subject matter. For example, terms such as *database* and *encryption* are relatively straightforward technical terms used in computer science, while terms such as *Backpropagation* and *Reproducible analytical pipeline* are more complex terms.

In conclusion, technical terms can be classified in various ways, depending on the context and purpose of classification. One common approach is to categorize them based on their structure, meaning, usage, and ways of formation. Technical terms related to AI can be classified into subject field terms, field-specific terms, technical terms, and general science terms. Additionally, technical terms can be classified based on their level of abstraction and complexity. Some technical terms are straightforward, while others are more complex and require a deeper understanding of the subject matter.

1.4.3. Current trends in the field of AI

To analyze ongoing tendencies in AI terminology, we need to have a look at the trends within the general field first, and with it see what happens on the language level supporting the trends.

One of the main directions of technological advancement is *Generative AI* – technology that allows to make original content out of preexisting patterns the system has spotted before. This branch of AI aims to take tasks connected with text, audio, video, images, at a revolutionary new level, where it will achieve or even surpass human skills. Thus, the terminology to serve in the field will likely employ borrowings from the field of art and law [17]. Terms like *AI-generated content (AI-generated text, image, music,* etc.) are already in AI glossaries, however, now we can observe growth and broadening within this cluster of terminology. Collocations like *AI-generated installation, AI-assisted writing, human-assisted art* (depending on the ratio of the input) appear in articles, describing the new technology, social media and even in description of public events. All of them have potential to become recognized terms in the subfield, where AI overlaps with art and so does the terminology.

The technology, though, evokes a significant amount of concern among creators. Such programs as Midjourney, Stable Diffusion or DALL·E 2 threaten not only to oust humans from the field, but steal their already created works. This is where the subfield overlaps with the copyright law. This is reflected in opposition of terms AI-generated and *derivative works*, motivated by inability of machines to create something genuine, but derive new works from already existing ones. Terms like *derivative work right, open-source licenses,* are newly coined and now are mainly seen in media publications, though may serve as evidence of convergence of AI and legal spheres, while in the public discussions there are also speculations whether terms like *copyright*, *patent*, *trademark* can apply to AI creations [47].

- *Derivative work* – a new work derived from an original work protected under copyright law [47].
- *Derivative work right* – the legal permission to develop a new work derived from an original work protected under copyright law. Derivative work rights are only granted for derivative works with original content, versus duplicated copyrighted

material. The original author's permission to transform or adapt an original work by the original owner is the essence of a derivative work right [144].

- *Open-source license* – a permissive stance on what users can do with the original source code [47].

In 2021 Government bodies – including U.S. financial regulators, the U.S. Federal Trade Commission, and the European Commission – announced guidelines or proposals for regulating artificial intelligence. Companies were required to conduct assessments of AI risks and to document how such risks have been minimized. A host of regulatory frameworks refer to these types of risk assessments as '*algorithmic impact assessments*' – also sometimes called 'IA for AI' – which have become increasingly popular across a range of AI and data protection frameworks [6].

Another obvious trend within AI is its democratization. Thus, *No-Code* and *Low-Code* AI applications are becoming more popular and more accessible to non-experts. In fact, HPE (Hewlett Packard Enterprise) believes that no-code and low-code platforms mark the rise of the *citizen developer* [92].

- *Low-code/no-code (LC/NC)* development refers to an environment where visual drag-and-drop applications or similar tools allow individuals and teams to program applications without a lot of linear coding. These types of systems help the IT world to deal with a lack of skilled developers and streamline the emergence of new applications and interfaces [133].
- *Citizen developer* **-** a tech-savvy end user who has the ability to create a new software feature or application program from an approved corporate or cloud-based code base, system or structure [144].

This will inevitably lead to democratization and simplification in the terminology. At the same time, it is likely to increase the necessity of standardization. Robotic process automation is one example of the fast-growing categories of LC/NC systems. Some of the newly-coined collocations in this area are *small-scale* *automation/large-scale automation*, *shadow IT*, *storefronts*(platforms, where citizen developers, system developers and leaders can collaborate) [54].

The third overarching tendency regarding AI is its ethical side. The growing concerns about individual rights, privacy, non-discrimination, and non-manipulation make developers steer their work towards creation programs and systems that will not be capable of violating ethical norms existing in the society. AI ethics discussion places fundamental importance on considerations in determining legitimate and illegitimate uses of AI. The public discussions on the topic already operate with collocations like *ethical AI*, *responsible AI*, so, we are likely to see such words grow into proper terms [11].

- *Ethical AI* – can be defined as: an approach in AI programming that should deal with concerns regarding artificially intelligent systems, i.e., biases may significantly impact machine learning algorithms that are trained from data, ranging from gender to race to age to economic status and everything in between [133].
- *Responsible AI* – can be defined as: the development and use of artificial intelligence (AI) in a way that is ethically and socially trustworthy. Legal accountability is an important factor driving responsible AI initiatives [11].

In conclusion, we can point out three main trends in the development of artificial intelligence: programs are gaining capabilities of generating content indistinguishable from humans', also the field becomes less expert-reliant, as more non-expert learn basics, allowing them to do zero-coding. At the same time, growing concerns regarding ethical principles of AI work encourage attempts to implement ethical principles in work and use of AI. All these tendencies are reflected in terminology, which is expected to grow in number of terms, merge with fields like law, art, and ethical philosophy, and adopt for more general use.

Conclusions for Chapter 1

The function of terminology is to support effective communication, accurate description of concepts and ideas, and efficient and productive work within a particular field. By providing a standardized language, terminology helps professionals and experts collaborate and advance knowledge within their area of work.

Terminology is generated in accordance with the established rules of word formation within each respective language, for example, affixation, conversion and blending. Different methods of word coinage may dominate in certain fields depending on the novelty of the field and dynamics within it. Some approaches involve using words or phrases that clearly indicate the function or purpose of the technology or concept, making it easier to understand and remember. Another approach is to create terms that are more abstract and symbolic, which can be useful in situations where the technology or concept is difficult to describe using concrete language. One other way to create a term is to create terms that are based on existing words or phrases but are modified to reflect the unique characteristics of the technology or concept being described.

Artificial Intelligence is one of the new fields. It is closely linked to computation, but focuses on creation machines, capable of human cognitive abilities and above. In the process of its historical development, the AI field grew its own terminology to name computer programs, programming languages, and other processes and concepts.

At the present moment, key terms of AI terminology are covered in major English dictionaries, with more terms being added each year. It is impossible to define the exact number of terms in the field, since not all the lexical units that are used in the field are qualified as terminology and fixed in dictionaries. At the same time there are more comprehensive glossaries of AI terms within specialized technical dictionaries. The new terms are emerging, and there are neologisms that are already used in the field but are not represented in the dictionaries. Existing trends and tendencies in the field are also reflected in its terminology, qualitative and quantitative changes within it.

CHAPTER 2. TERMINOLOGY COINAGE BY MEANS OF MORPHOLOGICAL DERIVATION

The creation of new words and terms is an ongoing process that has been taking place throughout the history of language. Developments within the field of AI require constant replenishment of its terminology. Dynamic advancements in the field accelerate the emerging of new languages to serve this new subfield. Thus, a variety of means are employed to elaborate new vocabulary to name new programs, approaches, and operations within the field of AI.

This chapter aims to explore the derivative means of term creation where a new term emerges through the process of suffixation, clipping or blending. We will analyze the most productive suffixes, the nature of the terms created by it and the specifics of the words that are transformed by the means of derivation.

There is a variety of methods to coin a new term. The methods of coinage may be regarded as older and more conventional, like suffixation, derivation from ancient languages, as well as more novel and, thus, less conventional, such as acronymization, blending and portmanteau.

2.1. Prefixation and Suffixation

Suffixation is widely employed in creation of new AI terms and is one of the most productive means of coinage. Some of the most frequently used prefixes are multi-, pre-, over-, and under- . Overall, there are at least 10 prefixes which are actively used in the field of AI. They convey a variety of meanings, create a new word with a more specific or nuanced sense. In this paper's glossary 33 words out of 310 are created by prefixation, which is 11% of the total number.

multi-

In all the newly created AI terms, the prefix *multi-* conveys its primary meaning, namely plurality of the noun it is attached to. The prefix comes from the

Latin word *multus*, which means *much* or *many*. When combined with a word, *multi-* indicates that there are many of that thing or that it has multiple parts.

1. *Multimodal artificial intelligence* – an approach to AI which incorporates multiple types of data. For example, a speech-to-text model that is typically trained on audio and text data, could include image data of lip movements taken from video recordings [144].
2. *Multi-layer neural network* – a kind of neural network that contains more than one layer of artificial neurons or nodes. They differ widely in design. It is important to note that while single-layer neural networks were useful early in the evolution of AI, the vast majority of networks used today have a multi-layer model [144].
3. *Multi-agent system* – a loosely coupled network of software agents that interact to solve problems that are beyond the individual capacities or knowledge of each problem solver [85].
4. *Multi-class classification* – a classification task with more than two classes. Each sample can only be labeled as one class. For example, classification using features extracted from a set of images of fruit, where each image may either be of an orange, an apple, or a pear [144].
5. *Multi-layer perceptron* – a feed-forward neural network complement. It has three layers: an input layer, a hidden layer, and an output layer, as shown in Fig [86].
6. *Multi-task learning* – a single shared machine learning model that can perform multiple different (albeit related) tasks. Multi-Task Learning offers advantages like improved data efficiency, faster model convergence, and reduced model overfitting due to shared representations [66].
7. *Multi-view learning* – an emerging direction in machine learning which considers learning with multiple views to improve the generalization performance. Multi-view learning is also known as data fusion or data integration from multiple feature sets [59].

8. *Multi-class classification* – a classification task with more than two classes. Each sample can only be labeled as one class. For example, classification using features extracted from a set of images of fruit, where each image may either be of an orange, an apple, or a pear [144].
9. *Multi-agent system* – a loosely coupled network of software agents that interact to solve problems that are beyond the individual capacities or knowledge of each problem solver [144].

auto-

At least seven AI terms contain the prefix auto-, which comes from Greek word meaning *self* [136], the terms created with *auto*- express relation to own self or independent operation [130], for example: *automated machine learning, automatic speech recognition, autoencoder*. Another word that was derived from the prefix and later borrowed from Greek – *autonomous* (from Greek – 'having one's own laws'), is a term-formation unit that signifies independence of the concept, e.g., *autonomous vehicles, autonomous driving, autonomics.* It is the second most productive suffix in the AI glossary.

1. *Autoencoder* – a specific kind of unsupervised artificial neural network that provides compression and other functionality in the field of machine learning. The specific use of the autoencoder is to use a feedforward approach to reconstitute an output from an input. The input is compressed and then sent to be decompressed as output, which is often similar to the original input. That is the nature of an autoencoder – that the similar inputs and outputs get measured and compared for execution results [144].
2. *Automatic speech recognition (ASR)* – use of computer hardware and software-based techniques to identify and process human voice. It is used to identify the words a person has spoken or to authenticate the identity of the person speaking into the system [144].
3. *Autonomics* – a study of self-regulating systems for process control [131].

4. *Autonomous driving* – driving that does not need human input to operate [133].
5. *Autonomous vehicles* – an autonomous machine or vehicle does not need human input to operate [133].
6. *Autoregressive model (AR)* – model that can predict future behavior based on past behavior**.** It's used for forecasting when there is some correlation between values in a time series and the values that precede and succeed them [44].
7. *Automated machine learning* – is the process of applying machine learning models to real-world problems using automation. More specifically, it automates the selection, composition and parameterization of ML models. Automating the machine learning process makes it more user-friendly and often provides faster, more accurate outputs than hand-coded algorithms [145].

inter-

The prefix *inter*- is derived from the Latin word 'inter', which means 'between' or 'among' [133]. When added as a prefix to a word, it indicates that something is occurring or existing between or among multiple things or entities. The prefix is now commonly used in English to indicate a relationship or connection. In the field of AI, it may signify bridging work of several systems (*interoperability*), or encounter of two independent systems (*interface*).

1. *Interoperability* – the ability of a system or component to function effectively with other systems or components [131].
2. *Interface Agents* – intelligent tutoring systems capable of providing context-sensitive help [9].
3. *Brain-computer interface* – a technology that allows communication between a human or animal brain and an external technology [144].

mini- /nano-

Both prefixes signify things which are lesser in size or quantity, however, *mini-* and means smaller than a normal example, whereas *nano-* means ‘one-billionth’ of a unit or ‘extremely small’. For example, *Mini-batch* means a smaller sample of a dataset, while *Nanobots* robots act on a nanoscale level, which is one billionth of a meter.

- *Mini-batch* – a subset of a larger dataset that is used for training a model. Instead of using the entire dataset to train the model at once, the dataset is divided into smaller batches, and the model is trained on each of these batches in sequence [16].
- *Nanobots* – molecular-sized robots measured on a nanoscale and programmed to accomplish a specific task within the human body. The concept is used to create smart vaccines, cancer therapy, and immunotherapy through painless methods of medication [138].

pre-

The prefix *pre-* is a Latin prefix that means *before* or *in advance of* [136]. It is used to form words that describe something that comes earlier, or in preparation for something else. For example, the word *pre-processing* means preparing the input for its further use. In many cases, it implies that the action or event that comes after the prefix is somehow influenced or affected by what comes before it. For example, the word *precondition* refers to something that must be true or satisfied before a particular action can be taken. The prefix is mainly used to modify nouns (*precondition*) and verbs (*pre-training*), but among AI terms we also find adjectives coined this way, for example *preattentive processing* means primary analysis a program makes before focusing on the main one.

1. *Preattentive processing* – a kind of processing at which simple features are coded spatially in parallel and a later stage at which focused attention is required to conjoin the separate features into coherent objects [49].

2. *Pre-processing* – a technique of preparing (cleaning and organizing) the raw data to make it suitable for building and training machine learning models [46].
3. *Pre-training* – in AI refers to training a model with one task to help it form parameters that can be used in other tasks [68].
4. *Precondition* – a fact or set of facts that must be true before an action can be taken; e.g., before a robot can pick up an object, its hand must be empty [139].

Neur(o)-

The prefix is commonly used in medical terminology to refer to nerves, the nervous system. In the context of AI, the prefix *neur(o)-* is often used to refer to artificial neural networks, which are computational models that are inspired by the structure and function of biological neurons in the human brain.

1. *NeuroEvolution of Augmenting Topologies or NEAT* is often described as a genetic solution for improving neural networks. The NEAT concept can be used to provide a new model for selecting typologies for a neural network and for initializing weights [144].
2. *Neuroinformatics* – a research field that focuses on organizing neuroscience data through analytical tools and computational models. It combines data across all scales and levels of neuroscience in order to understand the complex functions of the brain and work toward treatments for brain-related illness [144].
3. *Neuromining* – a process of applying various behavioral intelligence and machine learning techniques in order to analyze human behavior. The goal is to understand human behavior in depth to manipulate and influence it at scale [144].
4. *Neuromorphic computing* – a kind of computing that utilizes an engineering approach or method based on the activity of the biological brain. This type of approach can make technologies more versatile and adaptable, and promote more vibrant results than other types of traditional architectures, for instance, the von Neumann architecture that is so useful in traditional hardware design [144].

5. *Neurotechnology* – a new prominent tech term describes any technology that helps us to understand brain function, or enables a direct connection of technology with the human nervous system [144].

un-

The prefix *un-* is used to indicate the opposite or the absence of the meaning of the root word that it is attached to. In the field of AI this prefix appears to make oppositional pairs: *supervised-unsupervised*. It also appears in adjectives converted from past participle, where *un-* specifies the absence of the action: *unstructured, unauthorized.*

1. *Unstructured data* – data that may have multiple sources, such as online digital files, text documents, SMS messages, video clips, photos, voices, sensors, pings, etc. The majority of the data created today is unstructured data, which is one of the keys to AI's growth [133].
2. *Unauthorized derivative works* – a new creative work that is based on or derived from an existing original work, such as a sequel, adaptation, or translation, created without the permission of the copyright owner of the original work. This type of work is considered to be a violation of copyright law, as it infringes on the exclusive rights of the copyright owner to control the use and distribution of their original work and any derivative works based on it [47].
3. *Unsupervised learning* – a form of machine learning technique that concludes datasets with unannotated data. Cluster analysis is the most frequent type of unsupervised learning [133].

Under-/Over-

The duo of oppositional prefixes *under-* and *over-* are also seen in AI terms creation, which are used to modify the meaning of words by indicating the degree or extent of something. *Under-* is used to indicate a lower or lesser degree or amount of something. The prefix usually conveys ideas of lacking (u*nder-specification*), insufficiency (*underfitting*), but may also signify a process of reduction, for example *undersampling* – is a technique which is used to balance data by deleting certain elements. *Over-* is used to indicate a higher or greater degree or amount of something. The prefix usually denotes excessiveness, however like in *oversampling* carries out the idea of necessity to balance data by duplicating certain elements.

1. *Underfitting* – the counterpart of overfitting, happens when a machine learning model is not complex enough to accurately capture relationships between a dataset's features and a target variable. An underfitted model results in problematic or erroneous outcomes on new data, or data that it wasn't trained on, and often performs poorly even on training data [108].
2. *Overfitting* – a kind of machine learning training in which the algorithm can only work on or identify specific examples from the training data. A functioning model should be able to generalize patterns seen in the data to tackle new instances [133].
3. *Undersampling* – is a technique to balance uneven datasets by keeping all of the data in the minority class and decreasing the size of the majority class. It is one of several techniques data scientists can use to extract more accurate information from originally imbalanced datasets. Though it has disadvantages, such as the loss of potentially important information, it remains a common and important skill for data scientists [116].
4. *Oversampling* – is the process of replicating data points in a dataset that is imbalanced. This is done in order to create a more balanced dataset that can be used to train a machine learning model [115].
5. *Under-specification* – is a failure to specify in enough detail. In machine learning and AI, the under-specification of training samples can result in vastly different predictions for edge cases, even when very similar models [109].
6. *Overparameterization* – a scenario where the number of parameters of the model exceed the size of the training dataset or a similar threshold [92].

Hyper-

Is a synonymic prefix to *over-*, however in general English there is a difference in connotation they bring to the word: *hyper-* usually refers to signifying unnaturally excessive, abnormal, whereas words created by *over-* refer to something that is gone over a certain limit. However, when it comes to technical terms the prefix manifests one more meaning, namely bridging points within an entity (such as a database or network) non-sequentially [140]. This can be seen in the examples below:

1. *Hyperautomation* – a process that combines artificial intelligence, machine learning, and robotic process automation to automate routine business tasks. It allows computer systems to perform simple, repetitive tasks that would otherwise take up a lot of time, freeing up human employees to give them more time for higher-value work [145].
2. *Hyperparameter* – is a parameter that impacts how a model learns. They're usually set manually outside of the model [133].
3. *Hypervisor* – a software program that manages one or more virtual machines (VMs). It is used to create, start, stop, and reset VMs. The hypervisor allows each VM or 'guest' to access the physical hardware, such as the CPU, RAM, and storage. It can also limit how many system resources each VM can use so that multiple VMs can run simultaneously on a single system [145].

Meta-

The prefix has either self-referential meaning, or connotate pertaining to a level above or beyond [134], correspondently with its original Greek semantics [136].

- *Metadata* – data about data, or data that provides information on one or more aspects of the data [145].
- *Metaverse* – an immersive, interactive environment generated by a computer. Although there is no unified agreement on what the metaverse will look like — or how individuals will interact with it. [144].

To speak about most productive suffixes, we need to distinguish between already existing words that we formed by a certain suffix before they became AI terms (e.g., *analytics, technology, sensitivity*), and terms that were specifically created by means of suffixation to name a new concept in the field of Intelligent machines, for example *robotics* or *over-parameterization.* In this part we will attempt to focus on the latter category.

Overall, there are at least 7 suffixes which are employed regularly in term-formation. The frequently observed are *-ics, -ive, -ation, -ization* and others.

-ics

Suffix *-ics* is used to denote a body of facts, knowledge, principles [135]. It is a common suffix in English that is used to form nouns. It is derived from the Greek word *ikos*, which means ‘pertaining to’ or ‘related to’ [136]. It can also be used to describe practices or systems (*robotics, cybernetics, neuroinformatics, predictive analytics, autonomics* [131]).

1. *Robotics* – engineering and operation of machines that can autonomously or semi-autonomously perform physical tasks on behalf of a human. Typically, robots perform tasks that are either highly repetitive or too dangerous for a human to carry out safely [144].
2. *Cybernetics* – a study of control and communication in living and man-made systems [132].
3. *Predictive analytics* – a type of analytics designed to forecast what will happen in a specified time frame based on previous data and patterns [133].

-ive

The suffix is used to signify adjectives that express doing or tending to do something specified [23]. It is another common suffix in English that is used to form adjectives. It is derived from the Latin suffix *-ivus*, which means ‘tending to’ or ‘having the nature of’. Added to a base word (usually a verb), it creates an adjective that describes something as having a particular quality or tendency. For example, *generative AI* signifies the ability of machines to generate visual, audio or textual content; and *autoregressive model* [44] would refer to regression the program is designed to do in order to analyze past behavior.

1. *Abductive Reasoning* – a form of logical reasoning which starts with single or multiple observations then seeks to find the most likely explanation or conclusion for the observation [93].
2. *Cognitive architecture* – a computer architecture involving non-deterministic, multiple inference processes, as found in neural networks. Cognitive architectures model the human brain and contrast with single processor computers [144].
3. *Cognitive computing* – the term is often referred to as artificial intelligence. It's utilized by marketing teams to minimize AI's mystique in some businesses [133].
4. *Generative adversarial network (GAN)* – a machine learning approach in which two neural networks compete to create new data with the same statistics as the training set. GANs, for example, are utilized in fashion, art, and marketing, but they are also becoming increasingly popular among malicious attackers to spread false news [133].
5. *Generative AI* – a broad label that's used to describe any type of artificial intelligence that can be used to create new text, images, video, audio, code or synthetic data [144].
6. *Predictive analysis* – this type of analytics is designed to forecast what will happen in a specified time frame based on previous data and patterns [133].

-ize/-ization

Greek suffix *-izein*, which means 'to make' or 'to cause to be' transformed into the English suffix *-ize* to form verbs of making or causing something to be [136]. The suffix *-ize* can also indicate the act of converting something into a certain form or state (*digitize*) and optimization formed with them When the suffix *-ize* is combined with the suffix -*ation*, it forms the noun *ization*, which refers to the act or process of making something into a certain state or condition [130] (*optimization*, *data normalization*). And although AI terminology contains many terms created by these suffixes, many of them are borrowings from other disciplines, and are not specifically coined within the AI field. Nonetheless, there are several examples where these suffixes were employed to make a unique AI term like *over-parametrization*.

1. *Digitize* – to convert the representation of an object in an analog signal into a series of discrete points or samples. This involves converting existing non-digital information or data into a digital form with the intention of storing, altering or sharing this data with electronic devices [144].
2. *Data normalization* – the method used whenever the attributes of the dataset have different ranges. It helps to enhance the performance and reliability of a machine learning model. In this article, we will discuss in brief various Normalization techniques in machine learning, why it is used, examples of normalization in an ML model, and much more [88].
3. *Optimization* – an act, process, or methodology of making something (such as a design, system, or decision) as fully perfect, functional, or effective as possible [140].
4. *Over-parameterization* – refers to the scenario where the number of parameters of the model exceeds the size of the training dataset or a similar threshold [92].

-ation

The suffix -*ation* is a versatile tool in English that can be used to create a wide range of nouns related to actions, processes, and conditions. It is derived from the Latin suffix -*atio*, which indicates a process or action [136].

When the suffix -*ation* is added to a base word, it creates a noun that refers to the act or process of doing something or the result of that action. For example, *augmentation* refers to the act of augmenting or changing, similar changes can be seen in the examples below.

1. *Approximation algorithm* – an algorithm for an optimization problem that generates feasible but not necessarily optimal solutions [132].
2. *Backpropagation* – a method of teaching neural networks based on a known, desired output for certain sample circumstances [133].
3. *Data augmentation* – adding value to base data by adding information derived from internal and external sources within an enterprise [144].

-able

- *Explainable AI (XAI)* – artificial intelligence that can document how specific outcomes were generated in such a way that ordinary humans can understand the process. The goal of XAI is to make sure that artificial intelligence programs are transparent regarding both the purpose they serve and how they work [144].

In conclusion, within the field of AI there are many terms, created by means of suffixation. Majority of them were already pre-coined upon entering the AI field, but a decent number of concepts and processes were modified with the above-mentioned prefixes and suffixes specifically. Most of the word-formation suffixes are of Greek and Latin origin, which are in common use in general English. The most productive units are prefixes like *multi*-, *over*-, *pre*-, and suffixes like -*ics*, -*ive* and -*ation*.

2.2. Conversion or zero derivation

Conversion is a way of creating a new lexical unit by taking an already existing word and changing its properties as a part of speech. The core meaning of the word is often preserved, but the function in a sentence shifts. In AI terminology we can find examples of substantivisation, verbalization, and adjectivization.

Substantivisation – when a word which is not a noun enters the group of nouns. This form of conversion can be widely represented by terminalized gerund forms, e.g., *machine* ***learning***, or *abductive* ***reasoning*** [93].

1. *Machine* ***learning*** – ability of computers to learn without being explicitly programmed. Computers 'learn' via patterns they detect and adapt their behavior as a result [133].
2. ***Backtracking* -** an algorithm for capturing some or all solutions to given computational issues, especially for constraint satisfaction issues [144].
3. *Backward* ***chaining*** – a reverse technique is in which machines work forward from the intended objective or output to see whether there is any evidence to support those aims or outputs [133].

There are also cases when a non-gerund verb is converted into a noun, initially this substantivation occurred as a name for the game of Go, later a computer program capable of playing this game was named *AlphaGo.* The name was created by means of blending the verb '*go'* with the substativized Greek letter *Alpha*, which refers to the artificial intelligence program's use of 'deep learning' techniques. The term *Alpha* implies that the program is at the forefront of AI research and development [2].

4. *AlphaGo* **-** is a narrow AI, a computer program developed by Google DeepMind to play Go, a Chinese strategy board game for two players similar to chess [144].

There are also examples of substsubstantivation ativization of adjectives, like in *deepfake* or computer's name *Deep Blue.* The degree of substantivation may be full or partial. Fully substantivized adjectives share all the characteristics of nouns: can be used in the singular and in the plural, in the common and possessive cases, with the indefinite, definite or zero articles. *Deepfake, false negatives* and *false positives* are examples of full substantivation, as they adopted all features of a noun. *Deep Blue*, though, is a proper name, thus does not manifest typical noun features, unless in a specific context.

5. *Deep Blue* – a supercomputer developed by IBM specifically for playing chess and was best known for being the first artificial intelligence construct to ever win a chess match against a reigning world champion, Grandmaster Garry Kasparov, under regular time controls [144].
6. *Deepfake* – a term for videos and presentations enhanced by artificial intelligence and other modern technology to present falsified results. One of the best examples of deepfakes involves the use of image processing to produce video of celebrities, politicians or others saying or doing things that they never actually said or did [144].
7. *False **negatives*** – an incorrect prediction where a model mistakenly assumes an input does not have a required result when one actually exists. (Actual Yes, Predicted No) [133].
8. *False **positives*** – an error in a model's prediction of the presence of the desired result in input when it is not present (Actual No, Predicted Yes) [133].

There are also instances when a whole phrase substantivized, e.g., *Did You Mean* and *Divide-and-conquer*.

9. *Did You Mean (DYM)* **–** an NLP function used in search applications to identify typos in a query or suggest similar queries that could produce results in the search database being used [137].
10. *Divide-and-conquer* – a problem-solving algorithm that breaks a problem into subproblems that are similar to the original problem, recursively solves the subproblems, and finally combines the solutions to the subproblems to solve the original problem [39].

Adjectivization, when a non-adjective enters a group of adjectives, is also represented among AI terms. The most frequent cases of adjectivization are converted verbs in gerund and past participle forms (*bounding*, *embedded*, *unstructured, unsupervised)*, but we can also find examples of adjectivized verbs (*AI-complete*). Most of such adjectives cannot function independently but in original colocations or similar ones.

11. *AI-complete* – a term that is used to describe problems or outcomes that would rely on having a strong AI system in place – in other words, being able to put together a computer system that functions at as high a level as a human being. IT pros describe problems as 'AI-complete' if they are too difficult to be achieved by the use of conventional algorithms [144].
12. *Bounding box* **–** a made-up square that serves as a guideline point for object recognition and creates a collision box for that element [143].
13. *Embedded intelligence* **–** a term for a self-referential process in technology where a given system or program has the ability to analyze its own operations [144].
14. *Unstructured data* **–** data that may have multiple sources, such as online digital files, text documents, SMS messages, video clips, photos, voices, sensors, pings, etc. The majority of the data created today is unstructured data, which is one of the keys to AI's growth [133].
15. *Unsupervised learning* **–** a form of machine learning technique that concludes datasets with unannotated data. Cluster analysis is the most frequent type of unsupervised learning [133].
16. *Augmented Reality* **–** a type of interactive, reality-based display environment that takes the capabilities of computer-generated display, sound, text and effects to enhance the user's real-world experience [144].

There are also occasional cases of verbalization, when a non-verb enters a group of verbs, e.g., *to debug.* The noun bug got a new semantic meaning and was terminalized in the field of AI, the word verbalized to signify the process of fixing errors.

17. *Debugging* **–** is the task of finding and fixing bugs (or errors) in a software program. Bugs can range from small inconveniences (like ignoring user input in certain circumstances) or significant problems that can

cause memory leaks or crashes. Several methods are available for software developers to debug a program, including using a debugger or analyzing crash reports [145].

To sum up, conversion is a unique method of derivation with at least 17 words which accounts for 5% of the total glossary.

2.3. Blending

Blending , which combine parts of different words to create a word with a new meaning, unlike compounding, where the words preserve their initial form, terms combined by these means contain elements of source-words which are sometimes not easily recognized.

1. *Lisp* (list + processing) **–** a family of computer programming languages that originated in 1958 and has since undergone a number of changes and dialects. It is considered the second-oldest high-level programming language in use today, after Fortran [144].
2. *Prolog* (*Programmation en Logique*) **–** a high-level programming language that has its roots in first-order logic or first-order predicate calculus. The language was conceived in Marseilles, France in the early 1970s by a group led by Alain Colmerauer [144]. It is one of the first logic programming languages and it remains popular today. It is a programming language commonly associated with computational linguistics and artificial intelligence and is used in expert systems, theorem proving and pattern matching over natural language parse trees and natural language processing [63].
3. *Chatbot (Chat+robot)* **–** a software application that imitates human-to-human conversation through text or voice commands [133].
4. *Perceptron* (Perception + neuron) **–** a machine learning algorithm that helps provide classified outcomes for computing. It dates back to the 1950s and

represents a fundamental example of how machine learning algorithms work to develop data [144].

5. *AdaBoost* (*adaptive boosting*) – a type of algorithm that uses an ensemble learning approach to weight various inputs. It was designed by Yoav Freund and Robert Schapire in the early 21st century. It has now become somewhat of a go-to method for different kinds of boosting in machine learning paradigms [144].
6. *ImageNet* – an extensive visual database that is intended to be used in computer vision software development. ImageNet has hand-annotated over 14 million URLs of images to reveal what objects are shown [133].

Blending can be considered a minor term-formation method as the number of found blends amount to 2% of all terminology.

2.4. Clipping

Clipping is a process of word formation in which a new word is created by removing one or more syllables from an existing word. The use of this method of word coinage is fairly rare in the field of AI. Developers might use some terms in their clipped form(e.g., *tech* for *technology*); however, those remain to be regarded as slang and are not recognized as terms, still there are at least two examples of clippings that entered formal terminology:

1. *Bot* (clipping of *robot*) - an automated software program that digitally replicates some type of human activity [144].

   It is an example for front clipping as it has the first part of the initial word removed.
2. *Cog* (clipping for *cognitive*) – a humanoid robot designed by Rodney Brooks's group at MIT as a platform to study robot cognition. It could track faces, grasp objects, and, perhaps most famously, play with a Slinky [101]. The name was

chosen to imply cognitive abilities the robot was designed to demonstrate. This is an example of back clipping, where the end part of the word is clipped.

2.5. Acronyms

Acronyms and abbreviations have become an essential part of AI terminology. Since an overwhelming number of terms consist of more than one word, the use of acronyms and abbreviations not only helps to simplify complex technical jargon but also enables efficient communication among experts, researchers, and practitioners in the field of AI. There are at least 40 (13%) terms that are used as acronyms. Below there are some of the most frequently used ones:

1. *Algorithmic impact assessments* (often abbreviated as 'AIAs') **–** tools that set out frameworks and processes for assessing possible societal impacts, both beneficial or adverse, of AI systems before the systems are in use (with ongoing monitoring often advised) [27].
2. *AI (Artificial intelligence)* **–** a computer program that simulates aspects of human intelligence but focuses on a single, specific function [133].
3. *ASR (Automatic speech recognition)* **–** use of computer hardware and software-based techniques to identify and process human voice. It is used to identify the words a person has spoken or to authenticate the identity of the person speaking into the system [144].
4. *BNF (Backus-Naur form)* **–** a formal metasyntax used to express context-free grammars [13].
5. *BERT (Bidirectional encoder representations from transformers)* **–** a deep learning strategy for natural language processing (NLP) that helps artificial intelligence programs understand the context of ambiguous words in text [144].
6. *CPU (Central Processing Unit)* **–** an electronic circuitry in a computer that executes the commands of a computer program by executing basic arithmetic, logical, control, and input/output operations specified by the instructions [133].

7. *CNN (convolutional neural network)* **–** are deep artificial neural networks used to classify pictures (e.g., identify what they see), group them by similarity (photo search), and recognize objects in scenes [133].
8. *DC-IGN (Deep convolutional inverse graphics network)* **–** a particular type of convolutional neural network that is aimed at relating graphics representations to images. Experts explain that a deep convolutional inverse graphics network uses a 'vision as inverse graphics' paradigm that uses elements like lighting, object location, texture and other aspects of image design for very sophisticated image processing [144].
9. *ENS (Enterprise nervous system)* **–** the connectivity of network elements and components to create an intelligent whole, a comprehensive system serving enterprise goals. The idea is that parts of an IT architecture can be connected in much the same way that the human central nervous system is connected — to serve common goals through smart collaborative processes [144].
10. *GAN (Generative adversarial network)* **–** a machine learning approach in which two neural networks compete to create new data with the same statistics as the training set. GANs, for example, are utilized in fashion, art, and marketing, but they are also becoming increasingly popular among malicious attackers to spread false news [133].
11. *Graph Neural Network (GNN)* **–** a class of deep learning methods designed to make predictions on data described by graphs.

    Graphs are a way of representing data, relationships and their complexity. GNNs are neural networks that can be directly applied to graphs, and provide a way to generate node-level, edge-level, and graph-level predictions [144].
12. *NLP (Natural language processing)* **–** capacity of computers to comprehend or extract meaning from natural human languages. NLP generally entails computer interpretation of text or speech recognition [133].

Abbreviations, although being very common in the field of AI, still the majority of abbreviations cannot be considered new terms. Since almost all of them are two/three letter acronyms, they do not have any distinctive features, and so may be highly confusing. Thus, they are either used interchangeably, or, for example in media, and literature proper abbreviations are established in the introduction to the main text. Nonetheless, there are acronyms which are rather or no worse recognized in their abbreviated form, e.g., *ChatGPT, AI, NLP*.

2.6. Eponyms

Names of prominent researchers, developers and contributors are sometimes reflected in terminology, this way another coinage approach appears – eponyms, they are words or phrases that are derived from the name of a person, used to honor or pay tribute to the individual. Here are a few examples:

1. *Turing test* **–** a test that assesses the capacity of a machine to mimic human behavior. The evaluation consists of a real-world conversation between a person and another individual and a computer, in which the participants are assessed on their understanding [133].

The test got the name of its creator the British mathematician and computer scientist Alan Turing, who proposed the test as a measure of a machine's ability to exhibit intelligent behavior equivalent to, or indistinguishable from, the human’s [14: p. 4];

2. *Backus-Naur Form* (BNF) **–** a metalanguage used to describe the syntax of programming languages, as well as other formal languages. It was developed by John Backus and Peter Naur in the 1960s and got its name after both of its creators [13];
3. *Boltzmann machine* - a type of recurrent neural network in which nodes make binary decisions with some bias. Boltzmann machines can be strung together to make more sophisticated systems such as deep belief networks [144].
4. *Restricted Boltzmann machine* **–** a type of artificial neural network invented by Geoff Hinton, a pioneer in machine learning and neural network design. It is named after Ludwig Boltzmann, Austrian physicist, who developed the

Boltzmann distribution, which is a statistical distribution that describes the dispersal of energy in a system. The RBM is called *restricted* because the connections between the neurons in the same layer are not allowed. In other words, each neuron in the visible layer is only connected to neurons in the hidden layer, and vice versa. This allows the RBM to learn a compressed representation of the input data by reducing the dimensionality of the input [94].

5. *Naive Bayes classifier* **–** an algorithm that uses Bayes' theorem to classify objects. Naive Bayes classifiers assume strong, or naive, independence between attributes of data points. Popular uses of naive Bayes classifiers include spam filters, text analysis and medical diagnosis. These classifiers are widely used for machine learning because they are simple to implement [144]. This Classification is named after Thomas Bayes (1702-1761), who proposed the Bayes Theorem [112].
6. *Hopfield network* **–** a specific type of recurrent artificial neural network based on the research of American physicist John Hopfield in the 1980s on associative neural network models. Hopfield networks are associated with the concept of simulating human memory through pattern recognition and storage. The network was named after its inventor [35].

In the examples above the names of the creators and scientists, after whom the inventions were named, are used unchanged and form a compound noun with the follow-up word. There are also cases, when a proper noun obtain feature of an adjective due to attached suffixes like *-ian*:

7. *Hebbian theory* **–** a theoretical type of cell activation model in artificial neural networks that assesses the concept of ‘synaptic plasticity’ or dynamic strengthening or weakening of synapses over time according to input factors. The terms was named after the Canadian psychologist Donald Hebb, who proposed the theory of Hebbian learning, which states that neurons that fire together, wire together [81];
8. *Approximate Bayesian Computation* (*ABC*) **–** a statistical method used for estimating the likelihood of complex models. It was first introduced by Andrew

Beaumont and colleagues in the late 1990s, the term ‘Bayesian’ comes from the name of the 18th century mathematician and theologian Thomas Bayes, who first developed the mathematical theory underlying Bayesian inference [140].

9. *Boolean logic* – is a type of algebraic system used in computer science and mathematics. It was developed by George Boole in the mid-19th century and is based on the use of binary values (true/false or 1/0) to represent logical statements and operations [114].

Another kind of eponyms which may appear less obvious examples, are using prominent names from myths, legends or history, which do not have direct connection to the field. Often, such terms involve a degree of metaphorical reinventing, projecting the properties associated with the source domain onto the thing from the recipient domain – the sphere of AI. An example of such can be the computer from Bletchley Park named *Colossus* to evoke associations with great size, influence and ability.

10. *Colossus* – computer used by Alan Turing at Bletchley Park, UK during the Second World War to crack the ‘Tunny’ cipher produced by the Lorenz SZ 40 and SZ 42 machines. Colossus was a semi-fixed-program vacuum tube calculator [132].

Overall, there are at least 10 eponyms (3% of all terms), the five of which bear the names of their creators. In other cases, AI terms may contain names of mathematicians, whose contributions in mathematics were later used in the development of AI or even be derived from ancient names. Such terms can also be regarded as borrowings from other fields.

2.7. Borrowings

AI terminology has one more source of replenishment, namely loanwords, also known as borrowings. Borrowings can be of two kinds: from other languages and from other fields and disciplines. In this chapter, we are going to cover foreign loanwords and analyze the means of adaptation of such words in English. In chapter three, we will look closely at borrowings from adjacent disciplines.

The development of AI began in the onset of the era of globalization. English became one of the primary languages of international communication. At the same time American universities like Massachusetts Institute of Technology became the center of technological progress in the areas of computing and artificial intelligence. Thus, the primary language of AI development and, thus, terminology was English. Nonetheless, we can find terms that were coined with use of other languages, mainly Greek. There is another peculiarity of AI terms, since the first developers of AI (McCarthy, Turing, Minsky) have a background in mathematics, they expectedly followed the tradition of natural sciences to derive newly-coined terms from ancient languages. In the beginning of the chapter, we studied the examples of Greek and Latin words that developed into prefixes and suffixes. Here, we will analyze terms that retained the form of a full word, and so can be regarded as borrowings.

1. *Lambda calculus* – a type of formal system from mathematical logic used in computer science for function definition, application and recursion [144]. The was coined by Alonzo Church after the use of the Greek letter lambda (λ) as the basic abstraction operator in the calculus [134].
2. *Stochastic* – the term refers to data which has a random probability that may be analyzed via statistics. Although individual events cannot be predicted, analyzing the distribution of random stochastic variables may result in a pattern [144]. The adjective comes from the Greek word στόχος, which means 'aim' [134].
3. *Automata theory* – one of the alternative names for AI [132]; the word *automata* comes from the Greek word αὐτόματος, which means 'self-acting, self-willed, self-moving' [136]. The word also transformed into a prefix *auto*- many aforementioned in part 2.2. terms contain. [134]. Also, the Greek word gave name to *autonomics* - the study of self-regulating systems for process control [131].

4. *Heuristic* – rules or methods used to solve problems that are too complex or too large for brute-force approaches [134]. The word ‘heuristic’ comes from the Greek word *heuriskein*, meaning ‘to find’ or ‘to discover’ [136].
5. *Alpha-beta pruning* – an optimization of the minimax algorithm for choosing the next move in a two-player game. The position after each move is assigned a value [132]. The term consists of two Greek letters following the traditions of other natural sciences, which adopted using Greek letters to name abstract concepts and explain dynamics long before AI. *Alpha-beta* algorithm reflects the nature of the process: the algorithm passes 2 extra parameters in the minimax function, namely alpha and beta. Alpha stands for the best (highest- value) choice we have found so far at any point along the path of Maximizer. The initial value of alpha is - ∞. Beta stands for the best (lowest- value) choice we have found so far at any point along the path of Minimizer. The initial value of beta is +∞ [1].
6. *Algorithm* – a pattern of procedures or instructions given to an AI, neural network, or other devices to assist them in learning on their own; classification, clustering, suggestion, and regression are four of the most common sorts [133]. The term comes from the name of the Persian mathematician Al-Khwarizmi (which makes the term partly an eponym). The word ‘algorithm’ is a Latinized version of his name, and it refers to a set of rules or procedures used to solve a problem [136]. Later the term came to old French and from there entered modern English with first known use in 1926 [140].
7. Previously we mentioned *Prolog* as an example of coining new terms by blending. The name of the program also can be regarded as a borrowing from French as it stands for *Programmation en Logique* [63].

The number of borrowings is relatively scarce (<2%), since English is the primary language used in development of the technology. Nonetheless, deep connection of AI and computing with mathematics, resulted in terms that were coined following the tradition of older sciences, namely derivation and borrowing words from Greek and Latin. These words can also form new compound terms, e.g., *heuristic programming*. Apart from those, there are also examples of loanwords from French and Arabic languages.

- *Heuristic programming* – an approach to the idea of artificial intelligence by solving problems using experience-based rules or protocols [144].

Conclusions for chapter two

AI terminology is constantly being replenished as new technologies, techniques, and applications are developed. The field of AI has already elaborated a complex system of terminology, which already allows us to analyze, which means of coinage are more productive than others.

Morphological derivation is a versatile tool for creating new terms. The most productive approaches in derivation are suffixation and prefixation with more than 35% of terms created this way. There are newly-coined terms are formed by suffixes and prefixes of Latin and Greek origin, many of which are already ubiquitous in general English (*auto-, multi-, inter-, -ive, -ics, -ation*).

Examples of zero derivation can also be observed in the field, with cases of adjective and verb conversion to nouns, some of which can be regarded as fully substantivized; there are also examples of verb forms of gerund and past participle taking on forms of adjectives. There are also terms, made by a substantivized verb phrase.

Only 2% of terms are eponyms. Whereas there are examples of adjectivization of some proper names, the majority of them remained unchanged.

There is a tendency for shortening of the many-word terms by means of acronymization, blending and clipping. Whereas clippings remain in the domain of the AI slang, abbreviations are more widely represented in AI glossary, with 13% of terms, often used in abbreviated form, however, only some of the abbreviations can be regarded as new term, that replaces its full form. Most of the abbreviations are used accompanied with the initial compound term.

The field of AI is mainly replenished by derivations from already existing English words, and does not require replenishment by the means of borrowings from other languages. Nonetheless, we can observe borrowings from Greek and French that are specifically elaborated.

The complex nature of the majority of AI terminology led to multiple means of coinage being employed in creation of a new term. In this chapter we have seen examples, where an acronym is composed of several words, which in their turn were created by means of suffixation, eponymizing or have a new semantic meaning.

Thus, we can see that AI terminology coinage involves all existing methods of derivation, with suffixation and prefixation being the most dominant ones. Overall, at least 35% of terms were created by different means described in this chapter.

CHAPTER 3. TERMINOLOGY COINAGE BY MEANS OF SEMANTIC DERIVATION

Along with the morphological method, lexical-semantic, or semantic derivation plays a significant role, the essence of which is to use it as a ready-made term, a word borrowed from another lexical subsystem. This process is called terminologization and is accompanied with reterminologization – when already existing terms are borrowed from a different scientific field or a discipline. In this chapter we are going to look into the complex and manyfold process of terminologization and reterminologization, analyze specifics of compounding and see how metaphors enrich the terminological vocabulary of AI.

3.1. Role of compounding in semantic shift

A short glance at any AI glossary is enough to notice that the majority of terms rarely consist of a single word. To be precise, only 40 out of 310 terms are one-word, which means that 87% of terms are compound nouns. In the field of AI key words like *artificial*, *machine*, *data*, *learning* generate numerous compounds. The most productive units are *learning* (21 terms), *machine* (15 terms), *artificial* (17 terms), *processing* (8 terms), and *data* (24 terms). We will look into definitions and semantic peculiarities of these terms in the following sections.

We can differentiate three kinds of compounding: **Endocentric** if it contains a 'head' morpheme that determines its morphosyntactic features and general semantic type. In *activation function*, for example, the head is a function. Hence the compound is a noun, and names a type of *function*.

- *Activation function* **–** the main calculative layer of a neural network. The activation function triggers the right decision node within the neural network and displays the node as an output. It converts a series of inputs into singular or multiple output classes [138].

Alternatively, a compound will be **exocentric** if there is an understood head that is not pronounced. For example, in the exocentric compound *deepfake* [144], which means AI-generated video, the understood 'head' would be 'video'.

Still another possibility is that a compound is 'doubly headed', where both parts are equally important [99]. Usually, such compounds contain a significant change in meaning. For example, *random forest* – is a type of algorithm, so *forest* cannot be a 'head' word, as it does not name a special kind of forest.

- *Random forest* – a consensus algorithm used in supervised machine learning to solve regression and classification problems. Each random forest is comprised of multiple decision trees that work together as an ensemble to produce one prediction [137].

Compounding is a process when a word from one domain is added to a word from a different domain, creating a new lexical unit. This creates conducive (though not exclusive) conditions for a shift in semantic meaning. According to Bloomfield [19: p.426-427] there are nine classes of semantic change:

1) narrowing
2) widening
3) hyperbole
4) litotes
5) degeneration
6) elevation
7) metaphor
8) metonymy
9) synecdoche

The nature of terminology is to be neutral, precise and consistent, which affect the principles terms are created. For instance, it means excluding connotations as degeneration and elevation, as terms have to be stylistically neutral. Nonetheless, Bloomfield's classes are relevant in analyzing semantic derivation as a terminology coinage means. According to Tomilenko [127: p. 126], there are two ways a word of general use becomes a term:

1) preservation or modification of primary seme and generalizations, and introducing a new seme;

2) change of the primary seme, actualization of hidden, functional meaning.

These two processes correspond with Bloomfield's semantic widening or narrowing and reinventing of meaning through means of metaphor and metonymy. We will successively explore both methods in the next part.

3.2. Terminologization through means of semantic narrowing and widening

Semantic narrowing is a process in which the meaning of a word becomes more specific or limited over time. It occurs when a word that was originally used to refer to a broad category of things or concepts gradually comes to be associated with a narrower and more specific meaning. Narrowing and widening involve shifts along the spectrum of narrower to broader meaning and vice versa.

In AI terminology we can observe semantic narrowing of the term 'machine'. The first known use of the word is 1545 in the meaning of 'a constructed thing whether material or immaterial' [140]. This meaning is now considered archaic, instead Merriam-Webster Dictionaries defines it as 'a mechanically, electrically, or electronically operated device for performing a task' [140]. However, in AI terminology, 'machine' is one of the most productive building blocks, where it is solely thought of as a 'computer', 'software, or 'application' in terms like *machine learning* [133], *machine translation*, *machine vision.*

- *Machine translation* – an application of NLP for language translation (human-to-human) in text- and speech-based conversations [144].
- *Machine vision* (MV) – an integrated mechanical-optical-electronic-software technology which makes use of optical instrumentation, digital video,

electromagnetic sensing, mechanics and image processing technology. The technology's goal is optical and non-contact sensing to receive and analyze a real image in order to provide more information. Machine vision technology is widely used in monitoring and controlling a wide range of applications [145].

The opposite process can be observed with the term 'computer'. Whereas dictionaries define it as 'a programmable usually electronic device that can store, retrieve, and process data', within the field of AI, the term can be used interchangeably with terms like *software, artificial*, *program, technology* and more apart from the hardware device. In other words, a term like *computer vision* does not imply that an electronic device is capable of seeing, but rather a software program that can process an image with the help of hardware devices. The term computer here signifies the artificial nature of the process as opposed to biological vision.

- *Computer vision* – a multidisciplinary scientific discipline that investigates how computers can be programmed to understand digital images or movies at a high level. It focuses on automating activities that the human visual system can perform [133].

Similarly, *cognitive architecture* [144] is defined as *computer architecture*, but is not applied to architecture of the hardware, but the system within.

- *Brain-computer interface* **-** a technology that allows communication between a human or animal brain and an external technology. This term can refer to an interface that takes signals from the brain to an external piece of hardware, or a technology that sends signals to the brain [144].
- *Cognitive computing* – often referred to as artificial intelligence, but it's more accurately described as AI. It's utilized by marketing teams to minimize AI's mystique in some businesses [133].

In conclusion, the processes of semantic narrowing and widening is natural in the course of language development. In the context of AI terminology, it plays a role in term formation, shifting the initial semantic meaning towards broader or more specialized use. Understanding these processes is essential for anyone seeking to engage with the world of AI, as it allows for a better understanding of the underlying principles and concepts at play.

3.3. Terminologization through means of metaphorical reinvention

The appearance of scientific terminology is preceded by a period of spontaneous term formation, when a sufficient number of thematically united words is accumulated for description of a system of special concepts. It is unavoidable that in this period words of general use are applied to describe newly discovered or invented things. In the initial period of formation of a scientific field, it requires a comparison of certain concepts with corresponding coexisting objects of the surrounding reality by analogy, which became one of the external stimuli for the terminology of words in general use. The language of science, using commonly used words, fills them with a special scientific content, fixed in the definitions [127: p. 124-127]. One way this process occurs is through metaphorical reinvention. Existing words are repurposed to describe new concepts or technologies. Metaphors provide a way to make complex or abstract ideas more accessible to a wider audience, and they can also help experts to understand and conceptualize new technologies.

George Lakoff and Mark Johnsen describe several different kinds of metaphors: **conceptual metaphors** (where more simple concepts are used to explain more abstract and sophisticated) [67: p. 155]; **orientational metaphors** (that exploit attitudes in space and motion) [67: p. 16]; **structural metaphors** (where abstract ideas are mapped through structures of more common and familiar things) [67: p. 62]; **ontological metaphors** (where the concepts are explained through physical experiences) [67: p. 23].

There are also metaphorical blends, and given that AI terminology is mostly many-word, it is a regular occurrence in this field, which complicates the task of classification and calculation of the exact percentage. Nevertheless, in the following sections we will attempt to see, what are the dominant metaphors that impacted term-formation in the studied field.

3.3.1 Conceptual metaphors

Conceptual metaphor is the most fundamental type of metaphor, where one concept is understood in terms of another. One of the first metaphors introduced in the field of AI was, obviously, the conceptual metaphor of intelligence. The term not only became the name of the field but significantly impacted the terminology coinage and shaped the development of AI itself. It predetermined the overall perception of the technology and brought about other concepts we associate with human (or biological) cleverness to name creations of artificial intelligence. The key concepts associated with intelligence, like brains, neurons, memory, and learning, guided the development and helped to make complex and abstract concepts more accessible and understandable to a broader audience.

Intelligence itself appears in more than six terms (and in thirteen more terms being abbreviated as AI). Most of them are extensions of the term *artificial intelligence*, for example, *artificial superintelligence*, *artificial general intelligence (AGI)*, *multimodal artificial intelligence*.

- *Artificial superintelligence* **–** a term referring to the time when the capability of computers will surpass humans. The term *artificial intelligence*, which has been much used since the 1970s, refers to the ability of computers to mimic human thought. Artificial superintelligence goes a step beyond, and posits a world in which a computer's cognitive ability is superior to a human's [144].
- *Artificial general intelligence (AGI)* **–** a computational system that may execute any intellectual function that a human can. Also known as *Strong AI*. At the moment, AGI is just a concept [133].

- *Multimodal artificial intelligence* **–** an approach to AI which incorporates multiple types of data. For example, a speech-to-text model that is typically trained on audio and text data, could include image data of lip movements taken from video recordings.

However, there are also other compound terms where the '*artificial'* part is omitted, though is understood, e.g., *actionable Intelligence*, *embedded intelligence.*

- *Actionable Intelligence* **–** information you can leverage to support decision making [137].
- *Embedded intelligence* **–** a term for a self-referential process in technology where a given system or program has the ability to analyze its own operations [144].

The concept of intelligence is mainly delivered through the noun, but may occasionally appear as an adjective – to describe the nature of certain processes (*intelligent document processing (IDP)* or *intelligent document extraction and processing (IDEP)*)

- *Intelligent Document Processing (IDP)* or *Intelligent Document Extraction and Processing (IDEP)* **–** ability to autonomically read and convert unstructured and semi-structured data, identify usable data and extract it, then leveraged it via automated processes [137].

Another metaphorically reinvented concept that emerged together with intelligence was the concept of neural network (first use 1947 according to Mer Web). The idea was borrowed from the field of neuroscience – the study of brains. Computer engineers attempted to create a system where processors are interconnected in a way similar to neurons in a human brain. This resulted in appearance of such technology as *Neural Turing machine (NTM), NeuroEvolution of Augmenting Topologies (NEAT), Neuromorphic computing,* and *Neurotechnology* as an umbrella term for all of them.

- *Neural network* **–** a computer system that functions as the brain of a human. Although researchers are still attempting to construct a computer model of the human brain, current neural networks can already accomplish many things regarding speech, vision, and board game strategy [133].

- *Neural Turing machine (NTM)* – a technology that uses neural network methodologies to achieve the capability to verify algorithms and do other computational work. It is based on the mid-20th century work of renowned data scientist Alan Turing [144].
- *NeuroEvolution of Augmenting Topologies or NEAT* – a genetic solution for improving neural networks. The NEAT concept can be used to provide a new model for selecting typologies for a neural network and for initializing weights [144].
- *Neuroinformatics* – a research field that focuses on organizing neuroscience data through analytical tools and computational models. It combines data across all scales and levels of neuroscience in order to understand the complex functions of the brain and work toward treatments for brain-related illness. Neuroinformatics involves the techniques and tools for acquiring, sharing, storing, publishing, analyzing, modeling, visualizing and simulating data [144].
- *Neurotechnology* – a new prominent tech term describes any technology that helps us to understand brain function, or enables a direct connection of technology with the human nervous system [144].

Furthermore, processes like *Neuromining* and a variety of networks emerged inspired by the concept and structure of neural system (*artificial neural network, deep neural network, graph neural network, multi-layer neural network*)

- *Neuromining* – a process of applying various behavioral intelligence and machine learning techniques in order to analyze human behavior. The goal is to understand human behavior in depth to manipulate and influence it at scale [144].
- *Neuro Symbolic Artificial Intelligence (neurosymbolic AI)* – an advanced version of artificial intelligence that improves how a neural network arrives at a decision by adding classical rules-based (symbolic) AI to the process. This hybrid

approach requires less training data and makes it possible for humans to track how AI programming made a decision [144].

- *Artificial neural network (ANN)* – artificially created network based on the human brain's neural network designs, particularly the brain [133].
- *Graph Neural Network (GNN)* – A class of deep learning methods designed to make predictions on data described by graphs [133].

  Graphs are a way of representing data, relationships and their complexity. GNNs are neural networks that can be directly applied to graphs, and provide a way to generate node-level, edge-level, and graph-level predictions [144].
- *Multi-layer neural network* contains more than one layer of artificial neurons or nodes. They differ widely in design. It is important to note that while single-layer neural networks were useful early in the evolution of AI, the vast majority of networks used today have a multi-layer model [144].

The concepts closely associated with brain activity – cognition, knowledge, perception, learning and memory also found their reflection in AI terms. *Cognitive* mainly reiterates the term 'neural' (in *cognitive architecture*), 'intelligent' (in *cognitive computing*), or 'knowledge' (in *cognitive map*), and does not imply 'conscious intellectual activity', as its definition suggests [140; 130]. These terms point out that machines are designed to perform tasks that were traditionally thought to require human cognition, such as language understanding, decision making, and problem-solving.

- *Cognitive computing* – refers to the use of reasoning, language processing, machine learning, and human capabilities that help regular computing better solve problems and analyze data. By learning patterns and behaviors and becoming more intelligent, a computer system can tackle complex decision-making processes [142].
- *Cognitive Map* **-** a mental representation (otherwise known as a mental palace) which serves an individual to acquire, code, store, recall, and decode information

about the relative locations and attributes of phenomena in their environment [137].

Knowledge as an AI term echoes data as both of them refer to the information put into a computer system. The key difference lies in the attributes of knowledge, namely ‘awareness’ of something that was ‘learned’. So, in terms of artificial system – *knowledge* refers to *data*, that was processed by the computer, enabling it to generate new information, or simulate decision making, which can be seen from the following terms:

- *Knowledge based systems (KBS)* – the term refers to a variety of systems. The one common element is the objectives to express knowledge explicitly and a reasoning mechanism that enables it to generate new information [133].
- *Knowledge engineer* – a professional engaged in the science of building advanced logic into computer systems in order to try to simulate human decision-making and high-level cognitive tasks. A knowledge engineer supplies some or all of the ‘knowledge’ that is eventually built into the technology [144].
- *Knowledge extractions* – the extraction of knowledge from technical documentation, XML, unstructured datasets, or relational databases. Knowledge elements are extracted by running specific queries that represent the data best [138].
- *Knowledge Graph* – a graph of concepts whose value resides in its ability to meaningfully represent a portion of reality, specialized or otherwise. Every concept is linked to at least one other concept, and the quality of this connection can belong to different classes [137].
- *Knowledge Model* – a process of creating a computer interpretable model of knowledge or standards about a language, domain, or process(es). It is expressed in a data structure that enables the knowledge to be stored in a database and be interpreted by software [137].

- *Knowledge representation* – a field that involves considering artificial intelligence and how it presents some sort of knowledge, usually regarding a closed system [144].
- *Machine perception* – the capacity for a system to acquire and comprehend information from the environment in the same manner that humans do with their senses [133].

Learning is one of the paramount processes and with those terms of AI. The word itself is also (and expectedly) one of the most fruitful components of compound nouns making up more than twenty terms. The primary term is *machine learning*. The term organically fits the system of conceptual metaphors, where machines appear sentient creatures, capable of processing and analyzing input patterns and adapt their output accordingly. *Machine learning* is accompanied with array of supportive terms associated with it, for example:

- *Machine learning* – the term refers to the ability of computers to learn without being explicitly programmed. Computers 'learn' via patterns they detect and adapt their behavior as a result [133].
- *Machine learning operations (MLOps)* – an approach to managing the entire lifecycle of a machine learning model — including its training, tuning, everyday use in a production environment and retirement [133].
- *Machine learning as a service (MLaaS)* – a range of machine learning (ML) tools offered by cloud service providers. As of this writing, popular MLaaS offerings include Amazon SageMaker, Microsoft Azure ML, IBM Watson Machine Learning and Google Cloud ML [145].
- *Machine learning workflow* – processes involved in machine learning work. Various stages help to universalize the process of building and maintaining machine learning networks [145].
- *Learning algorithm* – a set of instructions used in machine learning that allows a computer program to imitate the way a human gets better at characterizing some types of information. The math and logic that supports a learning algorithm can

update itself over time (without human intervention) as the programming becomes exposed to more data [144].

There are also different approaches and types of machine learning, many of them are extended by introducing of concept of depth, that reflects the level of complexity (*deep learning*, *deep reinforcement learning*, *deep Q-learning*); supervision and reinforcement, indicating the degree of human engagement in the process (*self-supervised learning*, *semi-supervised learning*, s*upervised learning*, *unsupervised learning*) and more. These terms also come from the sphere of education, and human learning, and behavioral psychology or at least strongly associated with the presence of a teacher.

- *Deep learning* – an artificial intelligence technique that mimics the human brain by learning from how data is structured rather than a pre-programmed algorithm [133].
- *Deep reinforcement learning (Deep RL)* – an approach to machine learning that blends reinforcement learning techniques with strategies for deep learning [144].
- *Deep Q-learning* – an algorithm that breaks the chain in order to find the optimal Q-value function. It determines this by combining Q-learning and a neural network. The uses of the deep Q-learning algorithm can be stated as finding the input and the optimal Q-value for all possible actions as the output [50].
- *Self-supervised learning* - an approach to machine learning allows machine learning algorithms to use observed inputs to predict unknown inputs [144].
- *Semi-supervised learning* - a method used to enable machines to classify both tangible and intangible objects. The objects the machines need to classify or identify could be as varied as inferring the learning patterns of students from classroom videos to drawing inferences from data theft attempts on servers. [144].

- *Supervised learning* – a type of machine learning in which output data trains the machine to produce the correct algorithms, such as a teacher guiding a student. It’s more prevalent than unsupervised learning [133].
- *Unsupervised learning* – a form of machine learning technique that concludes datasets with unannotated data. Cluster analysis is the most frequent type of unsupervised learning [133].
- *Reinforcement learning (RL)* – a type of machine learning that involves an algorithm that learns by interacting with its surroundings and is then penalized or rewarded depending on how it acts [133].

The idea of unconventionality and advancement is conveyed through the term ‘extreme’ in *extreme learning machine.* The advanced complexity is attributed to randomization used in the setup.

- *Extreme learning machine (ELM)* is a particular kind of machine learning setup in which a single layer or multiple layers apply. ELM includes numbers of hidden neurons where the input weights are assigned randomly. Extreme learning machines use the concept of random projection and early perceptron models to do specific kinds of problem-solving [144].

Another example of structural metaphors complimenting concepts is *federated learning* and *transfer learning*. Borrowed political term ‘federation’, meaning a ‘union’ served to name the process of machine learning that uses data stored in different locations:

- *Federated learning* – an approach to machine learning where a model is trained on data where the data exists (in multiple locations), rather than the traditional approach of moving the data to a central location for model training [56].
- *Transfer learning* – a process that allows a pre-trained machine learning model to be used as a starting point for training a new model. Transfer learning reduces the cost of building the new model from scratch and speeds up the training process [144].

The final core metaphor in this system of conceptual metaphors is *memory*. The ability to memorize and store information human brains have been artificially recreated in machines. Artificial memory refers to the use of technology or computational systems to create, store, and retrieve memories. This can take many forms, from simple databases that store information to complex artificial intelligence systems that can learn and recognize patterns in data.

- *Memory* – the term refers to any information or data, often in binary format, that a machine or technology can recall and use. There are many different kinds of memory in conventional computers and other devices, and they differ based on the complex design of the hardware in which they're stored [144].
- *Long short-term memory (LSTM)* – units or blocks that are parts of a recurrent neural network structure. Recurrent neural networks are made to utilize certain types of artificial memory processes that can help these artificial intelligence programs to more effectively imitate human thought [144].
- *Limited memory* – systems with short-term memory limited to a given timeframe. Limited memory AI derives knowledge from real-time experiences or events and stores it in the database. When a problem occurs, it gives out redundant results [138].

A less expected element in this system of conceptual metaphors is mappings from the domain of human biological immune system and genetics. Computational models used to detect and eliminate malware and viruses from computer systems were named *artificial immune systems*. However, the term does not only convey the idea of embedded self-protection and self-treatment, but also reflects the structure of the biological immune system.

- *Artificial immune system* – is a system that utilizes some of the engineering of biological immune systems to put together algorithms or technologies that address systemic goals. This may involve mathematical and computer modeling of immune systems, or the abstraction of some immunology-related principles into algorithms [144].

- *Genetic Algorithm* – a method of searching for the best solution based on a defined scoring method. It is used to find optimized solutions for complex problems based on the theory of natural selection [47].

Finally, a noticeable metaphorical personification is seen in the terms like: *expert system and expert automation and augmentation software*, which reveals that not only separate computing activities or capabilities were able to be recreated artificially, but the entire system can perform on the complex level of a human expert.

- *Expert Automation and Augmentation Software (EAAS)* – a type of software resource that works on the process of automating highly cognitive behaviors or tasks. These types of software programs and systems are said to be taking over the work of semi-skilled white-collar workers in the knowledge economy [144].
- *Expert system* – a computer program that is designed to emulate and mimic human intelligence, skills or behavior [144].

Finally, it is also worth noting that the use of conceptual metaphors in terminology also has its drawbacks and pitfalls. Certain concepts may prove to be misleading or oversimplified, which may lead to wrong expectations and perceptions (like some people believing that AI programs like chatbots have human-like level of consciousness, or disappointing in inability of AI to perform fully fledged intelligence in specific contexts). Nonetheless, no serious attempts of terminology revision have been made so far.

3.3.2. Orientational metaphors

These metaphors use spatial relationships to understand more abstract concepts. According to Lakoff, orientational metaphors are not just linguistic expressions; they are deeply embedded in our conceptual systems and influence the way we think and act in the world. They are a powerful tool for shaping our understanding of abstract concepts and for communicating those concepts to others.

The idea of moving back in AI software implies a) processing of some already known output or solutions (*backpropagation* [144], *backtracking*, *backward chaining*); b) ability to analyze past behavior in order to predict future behavior (*autoregressive model* [44]).

- *Backtracking* – an algorithm for capturing some or all solutions to given computational issues, especially for constraint satisfaction issues. The algorithm can only be used for problems which can accept the concept of a 'partial candidate solution' and allows a quick test to see if the candidate solution can be a complete solution. Backtracking is considered an important technique to solve constraint satisfaction issues and puzzles. It is also considered a great technique for parsing and also forms the basis of many logic programming languages [144].
- *Backward chaining* – the reverse technique is in which machines work forward from the intended objective or output to see whether there is any evidence to support those aims or outputs [133].

The opposite motion – forward in AI indicates a) motion from a problem to the solution (*forward chaining*); b) one direction of information flow, which does not create a loop (*feedforward neural network*).

- *Forward chaining* – a method in which a machine must go from an issue to solving it. The AI must evaluate many options to determine which hypotheses are relevant to the problem [144].
- *Feedforward neural network* – is the simplest artificial neural network where information exits through the output node and doesn't come back for analysis. Data only flows in the forward direction and doesn't form a loop [144].

Vertical orientation can be observed in terms like *bottom-up testing,* where *bottom* indicates the simplest components and *up* – the entire system. This also brings the idea of a multilayer structure, where structural elements are hidden below the surface of the construction.

- *Bottom-up testing* – a specific type of integration testing that tests the lowest components of a code base first. More generally, it refers to a middle phase in software testing that involves taking integrated code units and testing them together, before testing an entire system or code base [144].

BERT is an example of bidirectional spatial motion. It is the name of a process when a computer analyzes a sentence simultaneously in two directions from left to right (conventional reading) and from right to left. This way, the system is capable of understanding ambiguous words and more complex contexts. Traditional language models, such as unidirectional LSTMs, only consider the previous context when predicting the next word.

- *Bidirectional encoder representations from transformers (BERT)* – a deep learning strategy for natural language processing (NLP) that helps artificial intelligence programs understand the context of ambiguous words in text. By analyzing them from left to right as well as from right to left [144].

Another duality of pr processes is reflected in the term *cross validation*. The idea of *cross*, as an intersection of two lines is reflected in two simultaneous processes, the approach uses. At the same time, the idea of opposition that is also often expressed through the metaphor of cross, does not manifest in the term.

- *Cross validation* – an approach to reducing overfitting during model development, by iteratively selecting different portions of the data to train and validate a predictive (supervised) machine learning model. It involves splitting a dataset into subsets, where one subset is used for testing the model's performance, and the other subset is used for training the model [34].

One more process conveyed through orientational metaphors is approximation. Whereas the literal meaning indicates quality or state of being close and near, (Mer Web) from *proximus* 'nearest', superlative of *prope* 'near' [133]. The term is actively reinvented in other domains to signify the process of attaining similarity without becoming exactly the same. AI term *approximation algorithm* [132] is when the system finds an acceptable solution, which is near to the optimal one but does not necessarily have to be so. In *ABC*, *approximate* conveys the idea of likelihood, the method aims to estimate.

· *Approximate Bayesian computation (ABC)* – is a statistical method used for estimating the likelihood of complex models [8].

In many expressions of general English, the idea of importance is conveyed through the concept of centrality (e.g., the central figure, the central idea). Similar shift in meaning we can observe in AI terms, even if they name non-abstract things, like *central processing unit* which is hardware unit, is not located in physical center of computer, but the importance of the functions it carries out is expressed in its name through the metaphorical reinvention of the word 'central'.

· *Central Processing Unit (CPU)* – the electronic circuitry in a computer executes the commands of a computer program by executing basic arithmetic, logical, control, and I/O operations specified by the instructions [133].

Reinvention of the idea of narrowness as something of 1) small width; 2) limited size and scope [140] in the way that it is something niche and specific was already on common use to describe a narrow specialization and terminologised in AI terminology:

· *Narrow AI* – AI focussed on solving a specific problem.
For example, an AI built to identify cancerous tumors in breast scans, would not automatically be able to detect tumors in other parts of the body without significant rework [144].

· *Low-code* [133] similarly referred to the spatial concept of fighting to signify low level of entry skills required for coding.

A whole cluster of AI terminology revolves around the spatial concept of *depth* verbalized through the adjective form 'deep'. The idea of depth explicit in the terms has a triple meaning: knowledgeable, layered, and impenetrable, the inherent lack of interpretability and understanding (such as in the 'deep sea' or 'deep space') Depth is intrinsic to the nature of tasks in domains where *deep learning* models are applied (*Deep learning* [133], *deep reinforcement learning* [144], *Deep Q-learning* [50]).

The original name of the first computer program beating a human champion in chess, Deep Blue [144], was initially named Deep Thought, but later was renamed to Deep Blue, merging it with Big Blue – IBM’s nickname [37]. (The nickname comes from the pervasiveness of IBM's blue computers [105]. The word choice may give an idea of dominance of metaphor of depth over the metaphor of thinking.

One of newly coined AI terms is already mentioned *deepfake*. Its first recorded use was in 2018 [140], and since then it has become recognized beyond the field of AI. Although it may appear that ‘deep-’ part is an adverbialized adjective to enhance the meaning of ‘*-fake*’, ‘*deep*-’ actually comes from deep learning technology, that is behind the creation of deepfake content.

3.3.3. Structural metaphors

These metaphors use one concept to understand another that is more complex. They use the structure or organization of one domain to understand or represent another domain.

The concept of depth, explored in the previous part, can be extended into a complex of metaphors that also revolve around the idea of depth but use structural domains to express their meaning. One instance of such is the concept of foundation, the primary idea of which is a ‘basis upon which something stands or is supported’ [140]. In AI (foundation model) foundation represents preliminary training of an algorithm which serves as a basis for future training.

- *Foundation model* – a deep learning algorithm that has been pre-trained with extremely large data sets scraped from the public internet [144].

Deep-terms are often complimented with structural metaphors of networks, and conceptual metaphors of neural networks. Altogether they convey ideas of complex, multilayer, processes within interconnected systems through such terms as, *deep neural network, deep residual network (deep ResNet), deep Q networks (DQN).*

- *Deep residual network (deep ResNet)* – a type of specialized neural network that helps to handle more sophisticated deep learning tasks and models. It has received quite a bit of attention at recent IT conventions, and is being considered for helping with the training of deep networks [144].
- *Deep convolutional inverse graphics network (DC-IGN)* – a particular type of convolutional neural network that is aimed at relating graphics representations to images. Experts explain that a deep convolutional inverse graphics network uses a 'vision as inverse graphics' paradigm that uses elements like lighting, object location, texture and other aspects of image design for very sophisticated image processing [144].

  The term *Convolutional* comes from the mathematical operation of convolution, which is used to process signals in digital signal processing [7].
- *Deep neural network* – a neural network with a certain level of complexity, a neural network with more than two layers. Deep neural networks use sophisticated mathematical modeling to process data in complex ways [144].
- *Deep Q-Networks (DQN)* – neural networks (and/or related tools) that utilize deep Q learning in order to provide models such as the simulation of intelligent video game play. Rather than being a specific name for a specific neural network build, Deep Q-Networks may be composed of convolutional neural networks and other structures that use specific methods to learn about various processes [144].

There is a complex of structural metaphors that use and explore the concept of a tree and its attributes: branches, roots, and forest. Tree is a simple yet accurate way to represent organization of data as a flow-chart in order to program algorithm of machine decision making.

- *Decision tree* – a flowchart-like representation of data that graphically resembles a tree that has been drawn upside down. In this analogy, the root of the tree is a decision that has to be made, the tree's branches are actions that can be taken and the tree's leaves are potential decision outcomes [144].

- *Junction tree algorithm* – a technique for extracting marginalization in general graphs. In other words, it entails performing belief propagation on a modified graph known as a junction tree. Because the graph divides into several parts of data and variable nodes are the branches, it is referred to as a tree [133].

Separate components of the tree-like flowcharts are called *branches*, each of them contains a specific instruction for the program.

- *Branch* – an instruction that tells a computer to begin executing different instructions rather than simply executing the instructions in order. In high-level languages, these are typically referred to as flow control procedures and are built into the language. In assembly programming, branch instructions are built into a CPU [144].

Root cause analysis is not a component of 'decision-trees'. The term employs an already established metaphor of common English root – as a basis or main cause of something.

- *Root cause analysis* – a problem-solving method which is used to pinpoint the exact cause of a problem or event [144].

Another conceptual metaphor, widely used beyond the field of AI, is *architecture*. In AI it is verbalized through nouns – names of the creation and sphere (*cognitive architecture*), and person in charge (*data architect*).

- *Data architect* – a person who practices data architecture, a field of study concerned with designing, generating, deploying, and managing an organization's data architecture. Data architects are frequently involved in AI initiatives [133].

Metaphors – names of places and objects – are sometimes used in term formation to highlight some specific feature or purpose of the thing they name. Concept of a warehouse – place of storage – was helpful to convey the idea of data aggregation and consolidation processes in a *data warehouse*. Similar idea of place, where various inputs are gathered and merged was carried out through the metaphor of lake, like in *data lake*. Finally, a special sort of transformation dictionary that congregates and converts words in NLP was named *Bag of Words.*

- *Data warehouse* – a collection of corporate information and data derived from operational systems and external data sources. A data warehouse is designed to support business decisions by allowing data consolidation, analysis and reporting at different aggregate levels. Data is populated into the DW through the processes of extraction, transformation and loading [144].
- *Data lake* – the process of merging all of the information (both structured and unstructured) into a central repository [133].
- *Bag of Words (BoW)* – an algorithm used for document classification and information retrieval. It extracts the text from a document and stores it in a bag of words without the grammar and sentence order. The frequency of words is used as a feature to train the algorithm and classify the document [138].

  BoW is often implemented as a Python dictionary. Each key in the dictionary is set to a word, and each value is set to the number of times the word appears [144].

3.3.4 Ontological metaphors

These metaphors use concepts of physical experiences to understand more abstract or intangible processes. Given its nature, these metaphors are mainly verbalized through verbs modified into gerund forms or noun -er forms.

Aforementioned concept of depth is nicely extended by a fruitful metaphorical reinvention of ‘mining’. The process of extracting valuable resources from the ground was qualitatively reinvented to refer to the abstract idea of obtaining certain information from raw data. This way terms like *data mining, association rule mining, web mining.* The metaphor is usually verbalized through the gerund form of the verb *to mine*, however it also may form a noun to name the tool carrying out the task (*data miner*).

- *Data mining* – the process of analyzing hidden patterns of data according to different perspectives in order to turn that data into useful and often actionable information. Data is collected and assembled in common areas, such as data warehouses, and data mining algorithms look for patterns that businesses can use to make better decisions, such as decisions that help cut costs, increase revenue, or better serve customers or clients [144].
- *Association rule mining* – a procedure which is meant to find frequent patterns, correlations, associations, or causal structures from data sets found in various kinds of databases such as relational databases, transactional databases, and other forms of data repositories [144].
- *Web mining* – the process of using data mining techniques and algorithms to extract information directly from the Web by extracting it from Web documents and services, Web content, hyperlinks and server logs. The goal of Web mining is to look for patterns in Web data by collecting and analyzing information in order to gain insight into trends, the industry and users in general [144].
- *Data miner* – a class of database applications that discovers previously unknown relationships among data, reveals hidden data for a specific purpose or demonstrates common patterns within data sets. While data miner software is widely used within the math and science fields, it has gained widespread popularity among online marketers in the form of spyware and to gather data that will help companies increase online sales [144].

By framing our experiences in the context of objects and substances, we are able to identify specific components of our experiences and categorize them as distinct entities or uniform substances.

Among the AI terms we see ones that use the 'motion' component of 'driving' to express the idea of motivating or basing a decision (*data-driven decision making*).

- *Data-driven decision making* – the process that involves making decisions that are backed up by hard data rather than making decisions that are intuitive or based on observation alone [144].

Data, as can be seen from previous compound terms, acquires properties of a matter that can be mined, stored, reduced in size or fused together (*data fusion*). Processes like compression (as a process of pressing or squeezing in order to reduce something in size or volume); cleansing (as getting rid of impurities), and collection of garbage became names for an array of processes of data and memory management associated with optimization: *data compression* (data is modified to consume less operational memory*); data cleansing (*getting rid of flawed bits of data*); garbage collection (*garbage stands for unused elements consuming space*).*

- *Data compression* – the process of modifying, encoding or converting the bits structure of data in such a way that it consumes less space on disk [144].
- *Data cleansing* – the process of altering data in a given storage resource to make sure that it is accurate and correct. There are many ways to pursue data cleansing in various software and data storage architectures; most of them center on the careful review of data sets and the protocols associated with any particular data storage technology [144].
- *Garbage collection* – a type of memory management. It automatically cleans up unused objects and pointers in memory, allowing the resources to be used again. Some programming languages have built-in garbage collection, while others require custom functions to manage unused memory [145].
- *Data fusion* – process of getting data from multiple sources in order to build more sophisticated models and understand more about a project. It often means getting combined data on a single subject and combining it for central analysis [144].

Similar way data is seen as a substance one feeds a computer in the process of uploading information into a program.

- *Feed* (12, verb) – to feed information into a computer means to gradually put it into it [131].

More abstract connotations of data are seen in the terms like data *exploration/integration/modeling/managing/normalization and preprocessing*. The activities are already widely used in application to the concepts of different levels of abstraction. However, when positioned together words tend to develop new connotations within the term. This way, searching through big volumes of data in a free and unrestricted way is compared to exploration and so.

- *Data exploration* – an informative search used by data consumers to form true analysis from the information gathered. Often, data is gathered in a non-rigid or controlled manner in large bulks. For true analysis, this unorganized bulk of data needs to be narrowed down. This is where data exploration is used to analyze the data and information from the data to form further analysis [144].

Integration (in *Data integration*), unlike fusion signifies the diverse nature of data sources, hence data is seen as something that may come in different forms and ranges (*raw data*), and thus may require *data normalization [101]* or *data preprocessing.*

- *Data integration* – a process in which heterogeneous data is retrieved and combined as an incorporated form and structure. Data integration allows different data types (such as data sets, documents and tables) to be merged by users, organizations and applications [144].
- *Raw data* – any data object that hasn't undergone thorough processing, either manually or through automated computer software [144].
- *Data preprocessing* – involves transforming raw data to well-formed data sets so that data mining analytics can be applied. Raw data is often incomplete and has inconsistent formatting. The adequacy or inadequacy of data preparation has a direct correlation with the success of any project that involve data analytics [144].

Certain ontological metaphors are born from experiences with objects. For example, inability to see what is happening inside a black box may have resulted in the name of an AI subset, where programmers cannot fully understand the nature of decision-making made by a program: *black box AI*. The opposite to it is the concept of *white box,* where processes happening within the program are clear.

- *Black box AI* – any type of artificial intelligence that is so complex that its decision-making process cannot be explained in a way that can be easily understood by humans [144].
- *White box testing* – a type of software testing in which the person understands the application design and source code. Using these insights, the person can run tests that reveal the most likely bugs or performance issues [145].

Whereas in the previous terms the nature of the concept was revealed through color of the box, the box itself was also metaphorically reintroduced as something imposing limitations, e.g., *bounding box* [143].

The term *bootstrap* (*bootstrapping*) refers to the load button that was used to initiate a hardwired bootstrap program, or smaller program that executed a larger program such as the OS. The term is believed to be derived from the expression ‘pulling yourself up by your own bootstraps’, starting small and loading programs one at a time while each program is ‘laced’ or connected to the subsequent program to be executed next [144].

- *Bootstrap* – the program that initializes the operating system (OS) during startup [144].

Bugs in the meaning of errors entered computing terminology in its early stages. It is considered to be slang that underwent the process of terminologization. And while the metaphorical bond between bug and error may be not so obvious, the process of debugging (removing pests as something destructive) clearly corresponds to eliminating malfunctions.

- *Debug* – a task of finding and fixing bugs (or errors) in a software program. Bugs can range from small inconveniences (like ignoring user input in certain circumstances) or significant problems that can cause memory leaks or crashes [145].

Certain terms employ semantics of adjectives of strength spectrum to map onto qualities of concepts they name. Thus, the idea of basic and blunt approach is conveyed in the term *brute force search*, as it exhaustively searches the entire search space. This is a straightforward and comprehensive approach of trying all possible solutions in a given search space without any particular strategy or optimization.

- *Brute force search* – a query that is not limited by clustering/approximations; it explores all inputs. Typically, more time-consuming and costly, but more thorough [133]. *Brute force search* is a useful technique for carrying out certain types of tasks, but it is not always the most efficient approach. Other techniques such as *divide-and-conquer*[39], may be more suitable for problems with larger search spaces or more complex constraints. The term reveals two main stages: dividing (breaking a problem into subproblems) and conquering (solving sub-problems by calling recursively until solved) [132].

Strength and weakness are also used in new meanings in terms of strong and *weak AI*, where they reveal capabilities of different AI creations.

- *Strong AI* – The main objective of AI technology is to produce systems that are as intelligent and skilled as the human mind [133].
- *Weak AI* – The term "weak AI" refers to a non-sentient computer system that operates under set parameters and focuses on a single activity or a small number of activities. Weak AI is the most prevalent sort of AI currently in use [133].

Some ontological metaphors can be also classified as animating or personifying [122: p. 56-58]. Such metaphors project typically human experiences or features over non-animate or non-human things. An example of such can be the concept of enmity in terms like *deep stubborn network* and related to it *adversarial machine learning* and *generative adversarial network (GAN)* [133].

- *Deep stubborn network* – network in which components work against themselves to produce more refined results [144].

The deep stubborn network is based on a generative adversarial network that consists of two components: a generator and a discriminatory engine. The role of the generator is to try to fool the discriminatory engine when it opts between legitimate and synthetic results.

- *Adversarial machine learning* – is a type of machine learning that seeks to toughen models through adversarial input. It seeks to make it easier for machine-learning methods in adversarial settings such as spam filtering, malware detection, and biometric identification [133].

Conclusions to Chapter 3

Semantic derivation is essential in the process of terminology of artificial intelligence. We can see multiple ways of shift in meaning, suggested by Bloomfield, i.e., semantic narrowing and widening. But the most efficient method of semantic shift is metaphor. In this chapter we studied only 105 terms which already account for 30% of all terms. However, given that two thirds of all terms are compounds, which foster semantic shift, the percent can turn out significantly higher after a more profound analysis. It is also worth noting that the prevalence of compounding is one of the definitive features of AI terminology.

Source domains of the metaphors we encountered are often words signifying processes and concepts from daily life, that share familiar features with the nature of processes and concepts within AI. Such metaphors simplify understanding and use of the technology and enhance communication within the field. The other category of metaphors are borrowings from other fields like psychology, neuroscience, mathematics and biology. The process of reterminologization of the borrowings aims to give the most accurate description to the AI concepts and processes.

The variety of metaphors, described by Lakoff and Johnsen, like conceptual, structural, ontological and orientational kinds – all of them can be found in the AI terms with regular blends when one kind of metaphor compliments another. The dominant kind of metaphor is conceptual (48 out of 102 terms explored in this chapter). Second after conceptual are structural metaphors (24 out of 102) and mainly draw on structures like the human brain or a tree.

Certain metaphors tend to form clusters of complex metaphors, for example the concept of intelligence is extended through concepts of knowledge, memory, learning, cognitive abilities and neural networks. These metaphors impact perception of the technology and may affect further terminology coinage.

Orientational metaphors were mainly used to point out backward-forward directions of processes executed by a program, for example, *backpropagation* or *feedforward learning*. The orientational concept of depth is one of the most productive units of term-formation. The concept of depth is extended by the ontological metaphor of mining and foundation. Other ontological metaphors employ many physical activities like cleansing, compression or feeding to express similar processes occurring in computation.

Overall semantic derivation is a manifold approach of term creation, that in different ways changes or shifts the meaning of already existing verbs or terms and can be considered the primary means of terminology coinage within the field of AI.

General conclusions

Technology is constantly developing, bringing about significant changes and advancements in various aspects of human life, including communication, transportation, healthcare, education, entertainment, and many more. One of the recently emerged technologies that has an impact on all of those aspects is AI. AI, or artificial intelligence, is a rapidly developing field. It refers to the ability of machines and computer systems to perform tasks that typically require human intelligence, such as learning, reasoning, decision-making, and problem-solving.

Terminology evolves as the field advances and new techniques, and applications are developed. Additionally, as AI is applied to new domains and industries, new terminology may emerge to describe specific applications or use cases. For example, in the field of natural language processing (NLP), terminology such as *sentiment analysis* [133], *entity recognition* [137], and *machine translation* [133] have emerged to describe specific techniques and applications for analyzing and generating human language.

The terminology presents interest for researchers, translators and linguists. As it is important to establish precise definitions, analyze the specifics and nature of the terms and give foundation for future lexicographical work.

In order to carry out this research, 310 terms were assembled from a variety of sources including dictionaries of general English, technical dictionaries, and special glossaries dedicated to AI. We also selected terms that are not defined yet, but are already widely used in media resources. The accumulated terms were analyzed in terms of their formation and intrinsic peculiarities of the terminology. We have divided all means of coinage into two parts: morphological and semantic derivation, so we could scrutinize the terminology from both perspectives. Morphological analysis revealed that a significant number of terms are created by means of affixation (11%) mainly employing Greek suffixes and prefixes. Blends clippings are also observed among well-established AI terms (4%). There were found attempts to elaborate AI terms by derivation from Greek words following the traditions of math and physics (3%); and eponyms – derivations from proper names (<2%).

One of the dominant features of AI terminology is the use of compound terms (>87%). This leads to ubiquitous acronymization with more than 13% being well-recognized abbreviations. However, most of them are used interchangeably with full terms and, thus, only few abbreviations can be regarded as proper examples of new terms.

Compounding also creates conducive conditions for semantic shifts. We examined which semantic shifts are typical for AI terms. It was found that AI terminology significantly benefits from metaphorical reinventing of words of general English as well as borrows terms from other fields. Different kinds of conceptual metaphors simplify understanding of the sophisticated concepts of AI, and at the same time more precise understanding of the nature of the things they name. For example, implementation of terminology borrowed from neuroscience gave very clear comprehension of how artificial networks work.

The terms were analyzed for the present in them metaphorical reinvention and its nature. It resulted into the following conclusions: conceptual metaphors are omnipresent with more than 30% of terms introduced by semantic change, though they may vary in degree of distancing from the initial meaning; some metaphors may form a system, for example concepts regarding human intelligence: intelligence-knowledge-learning-memory-cognition. Such complexes of metaphors potentially define future patterns of terminology coinage.

By carrying out the analysis of both kinds of derivation, we can conclude that it is typical for AI terminology to form compound nouns, where the components may undergo a semantic change. AI terminology employs a significant number of lexical units from general use and other disciplines, often giving them new meaning. At the same time, many-word terms are likely to be created by multiple means, where different morphological and semantic methods of word-formation would complement each other.

Current processes within AI, such as democratization of the field with *low-code* [133] technology, AI ethics, and significant advancement of *generative AI* [144] are likely to be reflected in the terminology evolving along. It is expected that AI will adopt terminology of such spheres as art, copyright law, and ethics. The inflow of non-professionals is likely to be accompanied with a shift toward more simplistic terminology. And finally, terms like *artificial superintelligence* [144] signify the fact that some of the key concepts may be reintroduced and modified to reflect the overall reevaluation of capacities of the technology.

The ongoing progress in the field does not allow to see this research as exhaustive, and will certainly require further updates. Nonetheless, it provides a valuable foundation for future studies, defines the approaches of analysis and develops guidelines for sequential lexicographical works.

Summary

The **objective** of the MA degree paper was to analyze terminology of artificial intelligence from the point of view of its creation. In order to fulfill the tasks of the research we assembled a glossary accounting for 310 terms, where compiled terms from already existing dictionaries, glossaries, media and academic publications. The terms were analyzed in terms of their morphological, etymological and semantical nature to define the ways they were formed. The applied **methods** were critical analysis, description, generalization, giving examples - to highlight the theoretical and methodological foundations of the study of terms; word-formation, structural, semantic, analysis of part of speech. The findings were carefully calculated, so that we can quantitatively analyze the most productive methods.

According to the requirements of academic research the paper is threefold. In the **first chapter** we established the theoretical basis of AI terminology formation by outlining specifics of terminology as a stratum of vocabulary; defined the concepts of *term* and *terminology*; and provided an overview of term-building patterns typical for terminology. We also explored the history of AI terminology juxtaposing it with the historical development of the field; compared the ways different dictionaries present AI terms; attempted to classify AI terms; and had a brief look at the modern-day processes in the field and terminological system of AI.

The **second chapter** was dedicated to morphological derivation. We analyzed existing means of word-formation through suffixation, prefixation, conversion, acronymization and more.

In the **third chapter**, we focused on semantic derivation, focusing on different kinds of metaphor as a means of semantic shift and what contributes to it.

Therefore, the tasks of the research may be regarded as fulfilled as we have established the unique features of AI terminology and analyzed its primary means of term formation.

List of references

1. Alpha-Beta Pruning. Java T Point. URL: https://www.javatpoint.com/ai-alpha-beta-pruning
2. AlphaGo. Deepmind.com URL: https://www.deepmind.com/research/highlighted-research/alphago
3. Althoff, D., Zilberstein, S. (2017). Planning in autonomous driving: Taking traffic rules into account. Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence.
4. Ananiadou, S. (1994). A methodology for automatic term recognition, Published in Proceeding COLING '94 Proceedings of the 15th conference on Computational linguistics , 2, 1034-1038.
5. Ancient Greeks Predicted Robots, Claims Stanford Historian. Geek Reporter. URL: https://greekreporter.com/2022/12/21/ancient-greeks-predicted-robots/
6. Appel, G, Neelbauer J., Schweidel D.A. Generative AI Has an Intellectual Property Problem. Harward Business Review. URL: https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem
7. Apply convolution to image processing, signal processing, and deep learning. Mathworks. URL: https://www.mathworks.com/discovery/convolution.html
8. Approximate Bayesian Computation. Journals.plos. URL: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1002803
9. Autonomous Interface Agents. Media Laboratory. URL: https://web.media.mit.edu/~lieber/Lieberary/Letizia/AIA/AIA.html#20
10. Arnold, T. and Tilton L. Depth in Deep Learning: Knowledgeable, Layered, and Impenetrable. Redrobe, K. and Scheible, J. (Eds.) Deep Mediations. University of Minnesota Press. [pdf][html]
11. Artificial Intelligence: examples of ethical dilemmas. UNESCO. URL: https://en.unesco.org/artificial-intelligence/ethics/cases
12. Babbage. BBC History. URL: https://www.bbc.co.uk/history/historic_figures/babbage_charles.shtml

13. Backus-Naur Form. Isaac Computer Science. URL: https://isaaccomputerscience.org/concepts/dsa_toc_bnf?examBoard=all&stage=all
14. Bansla, A., Bansla, N. Artificial Intelligence. International Journal of Applied Engineering Research, Vol. 7 No.11 (2012) PDF: file:///C:/Users/xiaom/Desktop/%D0%BC%D0%B0%D0%B3%D1%96%D1%81%D1%82%D1%80%D0%B0%D1%82%D1%83%D1%80%D0%B0/4%20semester/%D0%BC%D0%B0%D0%B3%D1%96%D1%81%D1%82%D0%B5%D1%80%D1%81%D1%8C%D0%BA%D0%B0/paperonArtificialIntelligence.pdf
15. Barr, A., & Feigenbaum, E. A. (1981). Handbook of artificial intelligence. William Kaufmann, Inc.
16. Batch, Mini Batch and Stochastic gradient descent. Medium. URL: https://sweta-nit.medium.com/batch-mini-batch-and-stochastic-gradient-descent-e9bc4cacd461
17. Beliaeva-Baran, E. 3 AI Trends We'll Be Talking About in 2022. Medium. URL: https://medium.com/technology-hits/3-ai-trends-well-be-talking-about-in-2022-ea6d9d6db142
18. Bernard, M. The 7 Biggest Artificial Intelligence (AI) Trends In 2022. Forbes. URL: https://www.forbes.com/sites/bernardmarr/2021/09/24/the-7-biggest-artificial-intelligence-ai-trends-in-2022/?sh=215f95f12015
19. Bloomfield, Leonard. (1933) Language. New York: Henry Holt and Company.
20. Boden, M. A. (2006). Mind as Machine: A History of Cognitive Science. Oxford University Press.
21. Bodle, A. How new words are born. The Gardian. URL: https://www.theguardian.com/media/mind-your-language/2016/feb/04/english-neologisms-new-words
22. Bozdechova, Ivana. Word-formation and technical languages. URL: https://sites.ff.cuni.cz/ucjtk/wp-content/uploads/ sites/57/2015/11/Word-formation-and-technical-languages.pdf

23. Britannica Encyclopedia URL: https://www.britannica.com/
24. Brock, D.C. (2018) Learning from Artificial Intelligence's Previous Awakenings: The History of Expert Systems. AI MAGAZINE. Association for the Advancement of Artificial Intelligence. PDF: file:///C:/Users/xiaom/Desktop/%D0%BC%D0%B0%D0%B3%D1%96%D1%81%D1%82%D1%80%D0%B0%D1%82%D1%83%D1%80%D0%B0/4%20semester/%D0%BC%D0%B0%D0%B3%D1%96%D1%81%D1%82%D0%B5%D1%80%D1%81%D1%8C%D0%BA%D0%B0/2809-Article%20Text-5631-2-10-20181012.pdf
25. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. URL: https://arxiv.org/pdf/2005.14165.pdf
26. Budin, G. (2019). Terminology and Knowledge Engineering: State of the Art and Future Perspectives.
27. Burt, A. New AI Regulations Are Coming. Is Your Organization Ready? Harward Business Review. URL: https://hbr.org/2021/04/new-ai-regulations-are-coming-is-your-organization-ready?registration=success
28. Bush, Vannevar. 1945. As We May Think. The Atlantic Monthly. July 1945.
29. Cabré, M. T. (1999). Terminology: Theory, methods and applications. John Benjamins Publishing.
30. Caso, A. L. The Production of New Scientific Terms. American Speech, Vol. 55, No. 2 (Summer, 1980), pp. 101-111.
31. ChatGBT URL: https://chat.openai.com/chat
32. Cog. Robots. URL: https://robots.ieee.org/robots/cog/
33. Coulom, R. (2007). "Efficient Selectivity and Backup Operators in Monte-Carlo Tree Search". Computers and Games, 5th International Conference, CG 2006, Turin, Italy, May 29–31, 2006. Revised Papers. H. Jaap van den Herik, Paolo Ciancarini, H. H. L. M. Donkers (eds.). Springer. pp. 72–83.
34. Cross-Validation in Machine Learning: How to Do It Right. Neptune.ai. URL: https://neptune.ai/blog/cross-validation-in-machine-learning-how-to-do-it-right

35. Crouse, E. Hopfield Networks: Neural Memory Machines. Towards Data Science. URL: https://towardsdatascience.com/hopfield-networks-neural-memory-machines-4c94be821073
36. Curley, A., The Ethical Concerns of GANs. Medium. URL: https://medium.datadriveninvestor.com/the-ethical-concerns-of-gans-9ef6b88f79db
37. Deep Blue. Chess Programming. URL: https://www.chessprogramming.org/Deep_Blue
38. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers) (pp. 4171–4186).
39. Divide and conquer algorithms. Khan Academy. URL: https://www.khanacademy.org/computing/computer-science/algorithms/merge-sort/a/divide-and-conquer-algorithms
40. Early Artificial Intelligence Projects. MIT. URL:https://projects.csail.mit.edu/films/aifilms/AIFilms.html
41. Edwin, D. Reilly (2003). Milestones in computer science and information technology. Greenwood Publishing Group. pp. 156–157.
42. Evans, V. and Green M. Cognitive Linguistics an Introduction. 2006. Edinburgh University Press. 857p.
43. Forbus, K. and Usher J. Analogical Reasoning in Artificial Intelligence: Integrating Symbolic and Neural Approaches. AI Magazine, Vol. 21, No. 3 (Fall 2000), pp. 19-32.
44. Ghosh, A. Intro to the Autoregressive Model. Medium. URL: https://medium.com/@adityaghoshsbg/intro-to-the-autoregressive-model-3c4efd4b6559

45. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. In Advances in neural information processing systems. URL: https://arxiv.org/pdf/1406.2661.pdf
46. Goyal, K. Data Preprocessing in Machine Learning: 7 Easy Steps To Follow. UpGrad. URL: https://www.upgrad.com/blog/data-preprocessing-in-machine-learning/
47. Haman. What are derivative works under copyright law? Legalzoom. URL: https://www.legalzoom.com/articles/what-are-derivative-works-under-copyright-law#:~:text=The%20creator%20of%20the%20derivative,copyright%20to%20the%20original%20work.
48. Hogben, L. (1969) The Vocabulary of Science. London. Heineman. 192p.
49. Hossain, P. Preattentive Processing and How to Use it in Design. Medium. URL: https://uxplanet.org/preattentive-processing-and-design-e59eba74373e
50. How Are Neural Networks Used in Deep Q-Learning? Turing. URL: https://www.turing.com/kb/how-are-neural-networks-used-in-deep-q-learning#what-is-deep-q-learning?
51. Hsu, Feng-hsiung (2002). Behind Deep Blue: Building the Computer that Defeated the World Chess Champion (1st ed.). Princeton University
52. Huge 'foundation models' are turbo-charging AI progress. The Economist. URL: https://www.economist.com/interactive/briefing/2022/06/11/huge-foundation-models-are-turbo-charging-ai-progress
53. ImageNet: A Pioneering Vision for Computers. History of Data Science. URL: https://www.historyofdatascience.com/imagenet-a-pioneering-vision-for-computers/
54. Johannessen, C., Davenport T. When Low-Code/No-Code Development Works — and When It Doesn't. Harward Business Review URL: https://hbr.org/2021/06/when-low-code-no-code-development-works-and-when-it-doesnt?autocomplete=true

55. Kageura, K. (1996). Issues in Knowledge Representation for Natural Language Processing. Berlin: Springer-Verlag
56. Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., ... & Zhang, Z. (2019). Advances and open problems in federated learning. URL: https://arxiv.org/pdf/1912.04977.pdf
57. Karimi, S., Behrouz H. A survey of chatbot systems through a human-like conversational approach. Published in Expert Systems with Applications, 2020.
58. Karthikeyan, J. Learning Outcomes Of Classroom Research. Department of English - Vellore Institute of Technology, Vellore, India Russell, S., & Norvig, P. (2010). Artificial Intelligence: A Modern Approach. Prentice Hall.
59. Keith, J. H. and Paul Thagard. AI Approaches to Analogical Reasoning. The Cambridge Handbook of Artificial Intelligence, edited by Keith Frankish and William M. Ramsey, Cambridge University Press, 2014, pp. 413-436.
60. Keith, H. and Thagard P. The Future of Analogy: A Synthetic Review. Frontiers in Psychology, Vol. 10 (2019), Article 1642.
61. Kluza, K. and Katarzyna Klessa. The Role of Technical Terminology in the Digital Age. Proceedings of the 10th Conference on Terminology and Artificial Intelligence
62. Koehn, P. Statistical machine translation, published in Cambridge Handbook of Natural Language Processing, 2010.
63. Kowalski, R. THE EARLY YEARS OF LOGIC PROGRAMMING. Communications of the ACM . January 1988 Volume 31 Number I. PDF: http://www.doc.ic.ac.uk/~rak/papers/the%20early%20years.pdf
64. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), Advances in Neural Information Processing Systems 25 (pp. 1097–1105). Curran Associates, Inc.
65. Kuhn, Thomas S. 1970a. The Structure of Scientific Revolutions. Chicago: University of Chicago Press. Kuhn, Thomas S. 1970b. “Logic of Discovery or

Psychology of Research?”. In Criticism and the Growth of Knowledge, edited by I. Lakatos and A. Musgrave: 1–23. London: Cambridge University Press.

66. Kundu, R. Multi-Task Learning in ML: Optimization & Use Cases [Overview]. V7 Lab. URL: https://www.v7labs.com/blog/multi-task-learning-guide
67. Lakoff, G. and Johnsen M. (2003) Metaphors we live by. London: The university of Chicago press. 193p.
68. Lee, B. What Is Pre-Training in NLP? Introducing 5 Key Technologies. Medium. URL: https://medium.com/ai%C2%B3-theory-practice-business/what-is-pre-training-in-nlp-introducing-5-key-technologies-455c54933054
69. Li, Deng and Dong, Yu (2014). Deep Learning: Methods and Applications. Foundations and Trends® in Signal Processing: Vol. 7: No. 3–4, pp 197-387.
70. Liu, J., Cao, Y., Wang, X., & Zhu, Y. (2020). How AI is changing the landscape of finance: evidence from literature. Journal of Business Research
71. Logic Theorist. A history of Artificial Intelligence.URL: https://ahistoryofai.com/logic-theorist/
72. Logic Theorist Explained. History Computer. URL: https://history-computer.com/logic-theorist/
73. Lutkevich, B. Automated machine learning TechTarget. URL: https://www.techtarget.com/searchenterpriseai/definition/automated-machine-learning-AutoML
74. Maad, M. Mijwel (2015). History of Artificial Intelligence. Baghdad College of Economic Sciences University. PDF: file:///C:/Users/xiaom/Desktop/%D0%BC%D0%B0%D0%B3%D1%96%D1%81%D1%82%D1%80%D0%B0%D1%82%D1%83%D1%80%D0%B0/4%20semester/%D0%BC%D0%B0%D0%B3%D1%96%D1%81%D1%82%D0%B5%D1%80%D1%81%D1%8C%D0%BA%D0%B0/HistoryofArtificialIntelligence.pdf
75. Maia, P. An Introduction to Multiple Instance Learning. Nilg.ai. URL: https://nilg.ai/202105/an-introduction-to-multiple-instance-learning/
76. Maslias, R. (2009)Terminology in the changing world of communication.

77. McCarthy, J., Minsky, M., Rochester, N., Shannon, C.E., A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. PDF: http://raysolomonoff.com/dartmouth/boxa/dart564props.pdf
78. McCarthy, J. (1958) Programs with common sense, Symposium on Mechanization of Thought Processes. National Physical Laboratory, Teddington, England,. PDF: http://jmc.stanford.edu/articles/mcc59/mcc59.pdf
79. McCarthy, J. (1960). Recursive Functions of Symbolic Expressions and Their Computation by Machine". Communications of the ACM. 3 (4): 184–195.
80. McCorduck, P. Machines Who Think. A.K. Peters, Ltd, Second Edition, 2004
81. Mishra, P. What is Hebbian Learning. Medium. URL: https://medium.datadriveninvestor.com/what-is-hebbian-learning-3a027e8e4bbb
82. Mitchell, M. The Rise of Analogy in AI. Communications of the ACM, Vol. 59, No. 12 (December 2016), pp. 48-55.
83. Metz, C. (2018, April 2). The race to develop the best AI translator. The New York Times. URL: https://www.nytimes.com/2018/04/02/technology/ai-translator-apps.html
84. Moveworks: AI terms Glossary. URL: https://www.moveworks.com/insights/ai-terms-glossary
85. Multi-Agent Systems. Carnegie Mellon Robotics Institute. URL: https://www.cs.cmu.edu/~softagents/multi.html
86. Multilayer Perceptron. Science Direct. URL: https://www.sciencedirect.com/topics/computer-science/multilayer-perceptron
87. Nilsson, Nils J. The Quest for Artificial Intelligence: A History of Ideas and Achievements. Cambridge University Press, 2010
88. Normalization in Machine Learning. Java T Point. URL: https://www.javatpoint.com/normalization-in-machine-learning
89. Pearson, J. (1998). Term in Context. John Benjamin Publishing. 246p.

90. Pfeifer, Rolf, and Bongard, Josh C. "How the Body Shapes the Way We Think: A New View of Intelligence." MIT Press, 2007.
91. Proceedings of the British Academy. The British Academy 2010. PDF: https://www.thebritishacademy.ac.uk/documents/1705/166p273.pdf
92. Ray, J. Over-parameterized models. Medium. URL: https://medium.com/log-loss/what-are-over-parameterized-models-40c53c304367
93. Reasoning in Artificial intelligence. Java T Point. URL: https://www.javatpoint.com/reasoning-in-artificial-intelligence#:~:text=Abductive%20reasoning%20is%20a%20form,do%20not%20guarantee%20the%20conclusion.
94. Restricted Boltzman Machine. Geeks For Geeks. URL: https://www.geeksforgeeks.org/restricted-boltzmann-machine/
95. Sager, J. C. (1990). A Practical Course in Terminology Processing. Amsterdam: John Benjamins Publishing Company 258p.
96. Sasu, Laura. Terminology dynamics-conceptual patterns of term formation. Bulletin of the Transilvania University. Vol. 2 (51) - 2009 Series IV: Philology and Cultural Studies PDF: http://webbut2.unitbv.ro/bu2009/buletin2009/Series%20IV/BULETIN%20IV%20PDF/28_Sasu.pdf
97. Shair, Ali Khan (2016). The Distinction between Term and Word: A Translator and Interpreter Problem and the Role of Teaching Terminology. Procedia - Social and Behavioral Sciences. PDF: https://www.sciencedirect.com/science/article/pii/S1877042816313283/pdf?md5=f0217967a531b7d383b6e7470911c8bd&pid=1-s2.0-S1877042816313283-main.pdf&_valck=1
98. Shapiro, S. C. (Ed.). (1987). Encyclopedia of artificial intelligence (Vol. 1). John Wiley & Sons.
99. Snyder, William. Compound Word Formation.The Oxford Handbook of Developmental Linguistics. Oxford: Oxford University Press.PDF: file:///C:/Users/xiaom/Desktop/%D0%BC%D0%B0%D0%B3%D1%96%D1%

81%D1%82%D1%80%D0%B0%D1%82%D1%83%D1%80%D0%B0/4%20semester/%D0%BC%D0%B0%D0%B3%D1%96%D1%81%D1%82%D0%B5%D1%80%D1%81%D1%8C%D0%BA%D0%B0/Compound_Word_Formation.pdf

100. Sonneveld, H. and Loenning, K. (1993). Terminology: Application in Interdisciplinary Communication. John Benjamin Publishing. 228p.

101. Stone, P., Brooks R, Autonomous robots in the real world. published in Communications of the ACM, 2015

102. Support Vector Machines. Scikit Learn. URL: https://scikit-learn.org/stable/modules/svm.html

103. Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction. MIT Press.

104. Temmerman, R. (2000). Towards New Ways of Terminology Description. The sociocognitive approach. Amsterdam/ Philadelphia: John Benjamins. 258p

105. The Making of International Business Machines. IBM History. URL: https://www.ibm.com/ibm/history/ibm100/us/en/icons/makingibm/impacts/

106. Tree Search. Brilliant. URL: https://brilliant.org/wiki/tree-search/

107. Turing, A. M. (1936). "On Computable Numbers, with an Application to the Entscheidungsproblem." Proceedings of the London Mathematical Society, Series 2, 42

108. Underfiting. DataRobot. URL: https://www.datarobot.com/wiki/underfitting/

109. Underspecification in AI. Data Science Examples. URL: https://www.datascienceexamples.com/underspecification-in-ai/

110. Valeontis, K., Mantzari E. The linguistic dimension of terminology: principles and methods of term formation. Proceedings of the 1-st Athens International Conference on Translation and Interpretation Translation: Between Art and Social Science, 13–14 October 2006. URL: http://www.eleto.gr/download/BooksAndArticles/HAU-Conference2006-ValeontisMantzari_EN.pdf

111. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. URL: https://arxiv.org/pdf/1706.03762.pdf

112. Vikram, K. Bayes and Naive Bayes Classifier. Cornell University archive. URL: https://arxiv.org/abs/1404.0933

113. Warburton, K. The Corporate Terminologist. 2021. John Benjamins Publishing Company. 249p

114. What is Boolean Logic? Isaac Computer Science. URL: https://isaaccomputerscience.org/concepts/sys_bool_what_is_boolean_logic?examBoard=all&stage=all&topic=boolean_logic

115. What is oversampling in data mining? AI blog. URL: https://aiblog.co.za/ai-faq/what-is-oversampling-in-data-mining

116. What Is Undersampling? Masters in Data Science. URL: https://www.mastersindatascience.org/learning/statistics-data-science/undersampling/

117. Wiene, N. (1985) Cybernetics or control and communication in the animal and the machine. MIT Press. 239p. PDF: https://uberty.org/wp-content/uploads/2015/07/Norbert_Wiener_Cybernetics.pdf

118. Word Formation Processes in English: 10 Types with Examples. KSE Academy. URL: https://kseacademy.com/word-formation-processes/

119. Wright, S.E. and Gerhard Budin. Terminology Management: Basic Aspects of Terminology Management. Published by John Benjamins Publishing Company in 2010. (Volume 1. 372 pp)

120. Wüster, E. (1979). Einführung in die Allgemeine Terminologielehre und Terminologische Lexikographie. Wien: Wilhelm Braumüller

121. Yaseen, A. H. (2013). The commonest types of metaphor in English. Journal of Al-Frahedis Arts, 3(15), 52-69 URL: https://www.iasj.net/iasj?func=fulltext&aId=84981

122. Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent Trends in Deep Learning Based Natural Language Processing. IEEE Computational Intelligence Magazine, 13(3), 55–75.

123. Вакуленко М.О. Ukrainian Lingua-Information Fund Of Nasu Term And Terminology: Basic Approaches, Definitions, and Investigation Methods. IITF Journal Vol. 24 (2013-2014)

124. Вакуленко М. О. 2013. "Методологічні засади вивчення наукової термінології" Термінологічний вісник № 2(2): 16-21.

125. Канделаки Т.Л. Значение терминов и системы значений научно-технических терминологий. Проблемы языка науки и техники: логические, лингвистические и историко-научные аспекты терминологии. Москва: Наука, 1970. С. 19–92.

126. Мацько Л.І., Кравець Л.В. Культура української фахової мови: навчальний посібник. Київ : Академія, 2007. 360 с.

127. Томіленко Л. М. Термінологічна лексика в сучасній тлумачній лексикографії української літературної мови: [монографія] / Jl. М. Томіленко. - Івано-Франківськ: Фоліант, 2015.- 160 с.

128. Ханкішиєва Юлія. Місце Термінології в Лексичній Системі Мови. Матеріали Всеукраїнської наукової конференції. Вінниця. 2016. Ст. 93-96

129. Шилінська І. Ways of Information Technology Terminology Formation. «Південний архів» (Збірник наукових праць. Філологічні науки) URL: http://dspace.wunu.edu.ua/bitstream/316497/35192/1/ways%20of%20IT%20terminology%20formation.pdf

List of lexicographical sources:

130. Cambridge Dictionary. URL: https://dictionary.cambridge.org/dictionary/

131. Collins Dictionary. URL: https://www.collinsdictionary.com/

132. Computer Dictionary Online. URL: https://www.computer-dictionary-online.org/

133. Dataconomy: AI terms Glossary. URL: https://dataconomy.com/2022/04/23/artificial-intelligence-terms-ai-glossary/#Artificial_Intelligence_glossary_2022

134. Definitions.net Dictionary. URL: https://www.definitions.net/

135. Dictionary.com URL:https://www.dictionary.com/

136. Etymology online. URL: https://www.etymonline.com/

137. Expert.ai: AI Glossary. URL: https://www.expert.ai/glossary-of-ai-terms/

138. G2: AI Glossary. URL: https://www.g2.com/articles/artificial-intelligence-terms

139. MIMI: Artificial Intelligence glossary. URL: https://en.mimi.hu/artificial_intelligence/advice_taker.html

140. Merriam-Webster.com dictionary. URL: https://www.merriam-webster.com/

141. Oxford English Dictionary. URL: https://www.oed.com/

142. Sisense: Glossary. URL: https://www.sisense.com/glossary/cognitive-computing/

143. Tasq.ai. Glossary https://www.tasq.ai/glossary/bounding-box/

144. Technopedia Dictionary. URL: https://www.techopedia.com/dictionary

145. TechTerms.com Dictionary. URL: https://techterms.com/

Appendix

Glossary of AI terms

A

1. *Abductive Reasoning* – a form of logical reasoning which starts with single or multiple observations then seeks to find the most likely explanation or conclusion for the observation [93].
2. *Accuracy* – a statistic that indicates how successful your AI model is at predicting outcomes [133].
3. *Actionable Intelligence* – information you can leverage to support decision making [137].
4. *Activation function* – it is the main calculative layer of a neural network. The activation function triggers the right decision node within the neural network and displays the node as an output. It converts a series of inputs into singular or multiple output classes [138].
5. *AdaBoost* (*adaptive boosting*) – a type of algorithm that uses an ensemble learning approach to weight various inputs. It was designed by Yoav Freund and Robert Schapire in the early 21st century. It has now become somewhat of a go-to method for different kinds of boosting in machine learning paradigms [144].
6. *Adversarial machine learning* – a type of machine learning that seeks to toughen models through adversarial input. It seeks to make it easier for machine-learning methods in adversarial settings such as spam filtering, malware detection, and biometric identification [133].
7. *Advice taker* – a hypothetical computer program, proposed by John McCarthy in his 1958 paper 'Programs with Common Sense'. It was probably the first proposal to use logic to represent information in a computer and not just as the subject matter of another program [139].
8. *AI-complete* – a term that is used to describe problems or outcomes that would rely on having a strong AI system in place – in other words, being able to put together a computer system that functions at as high a level as a human being. IT

pros describe problems as ‘AI-complete’ if they are too difficult to be achieved by the use of conventional algorithms [144].

9. *Algorithm* – a pattern of procedures or instructions given to an AI, neural network, or other devices to assist them in learning on their own; classification, clustering, suggestion, and regression are four of the most common sorts [133].
10. *Algorithmic impact assessments* (often abbreviated as ‘AIAs’) – tools that set out frameworks and processes for assessing possible societal impacts, both beneficial or adverse, of AI systems before the systems are in use (with ongoing monitoring often advised) [47].
11. *Alpha-beta pruning* – an optimization of the minimax algorithm for choosing the next move in a two-player game. The position after each move is assigned a value [132].
12. *AlphaGo* – a narrow AI, a computer program developed by Google DeepMind to play Go, a Chinese strategy board game for two players similar to chess [144].
13. *Approximation algorithm* – an algorithm for an optimization problem that generates feasible but not necessarily optimal solutions [132].
14. *Approximate Bayesian computation (ABC)* – a statistical method used for estimating the likelihood of complex models [8].
15. *Boolean logic* – a type of algebraic system used in computer science and mathematics. It was developed by George Boole in the mid-19th century and is based on the use of binary values (true/false or 1/0) to represent logical statements and operations [132].
16. *Artificial general intelligence (AGI)* – a computational system that may execute any intellectual function that a human can. Also known as *Strong AI*. At the moment, AGI is just a concept [133].
17. *Artificial immune system* – a system that utilizes some of the engineering of biological immune systems to put together algorithms or technologies that address systemic goals. This may involve mathematical and computer modeling of immune systems, or the abstraction of some immunology-related principles into algorithms [144].

18. *Artificial intelligence* – a computer program that simulates aspects of human intelligence but focuses on a single, specific function [133].
19. *Artificial intelligence engineer* is someone whose job is to identify the right approach to using AI to solve a specific business problem. AI engineers work to develop and deploy learning algorithms that can use historical and real-time data to predict future events [144].
20. *Artificial life* – the study of synthetic systems which behave like natural living systems in some way. Artificial Life complements the traditional biological sciences concerned with the analysis of living organisms by attempting to create lifelike behaviors within computers and other artificial media. Artificial Life can contribute to theoretical biology by modeling forms of life other than those which exist in nature. It has applications in environmental and financial modeling and network communications [144].
21. *Artificial neural network (ANN)* – artificially created network based on the human brain's neural network designs, particularly the brain [133].
22. *Artificial superintelligence* – a term referring to the time when the capability of computers will surpass humans. The term *artificial intelligence,* which has been much used since the 1970s, refers to the ability of computers to mimic human thought. Artificial superintelligence goes a step beyond, and posits a world in which a computer's cognitive ability is superior to a human's [144].
23. *Association rule mining* – a procedure which is meant to find frequent patterns, correlations, associations, or causal structures from data sets found in various kinds of databases such as relational databases, transactional databases, and other forms of data repositories [144].
24. *Augmented Reality* – a type of interactive, reality-based display environment that takes the capabilities of computer-generated display, sound, text and effects to enhance the user's real-world experience [144].
25. *Autoencoder* – a specific kind of unsupervised artificial neural network that provides compression and other functionality in the field of machine learning. The specific use of the autoencoder is to use a feedforward approach to

reconstitute an output from an input. The input is compressed and then sent to be decompressed as output, which is often similar to the original input. That is the nature of an autoencoder – that the similar inputs and outputs get measured and compared for execution results [144].

26. *Automata theory* - a theoretical branch of computer science and mathematics. It is the study of abstract machines and the computation problems that can be solved using these machines. The abstract machine is called the automata. The main motivation behind developing the automata theory was to develop methods to describe and analyze the dynamic behavior of discrete systems. [132]
27. *Automated machine learning* – the process of applying machine learning models to real-world problems using automation. More specifically, it automates the selection, composition and parameterization of ML models. Automating the machine learning process makes it more user-friendly and often provides faster, more accurate outputs than hand-coded algorithms [73].
28. *Automatic speech recognition (ASR)* – use of computer hardware and software-based techniques to identify and process human voice. It is used to identify the words a person has spoken or to authenticate the identity of the person speaking into the system [144].
29. *Autonomics* – a study of self-regulating systems for process control [132].
30. *Autonomous driving* – driving that does not need human input to operate [133].
31. *Autonomous vehicles* – an autonomous machine or vehicle does not need human input to operate [133].
32. *Autoregressive model (AR)* – model that can predict future behavior based on past behavior**.** It's used for forecasting when there is some correlation between values in a time series and the values that precede and succeed them [44].

B

33. *Backpropagation* – a method of teaching neural networks based on a known, desired output for certain sample circumstances [133].
34. *Backtracking* – an algorithm for capturing some or all solutions to given computational issues, especially for constraint satisfaction issues. The algorithm

can only be used for problems which can accept the concept of a ‘partial candidate solution’ and allows a quick test to see if the candidate solution can be a complete solution. Backtracking is considered an important technique to solve constraint satisfaction issues and puzzles. It is also considered a great technique for parsing and also forms the basis of many logic programming languages [144].

35. *Backus-Naur Form* (BNF) – a metalanguage used to describe the syntax of programming languages, as well as other formal languages. It was developed by John Backus and Peter Naur in the 1960s and got its name after both of its creators [68];
36. *Backward chaining* – the reverse technique is in which machines work forward from the intended objective or output to see whether there is any evidence to support those aims or outputs [133].
37. *Bag of Words (BoW)* – an algorithm used for document classification and information retrieval. It extracts the text from a document and stores it in a bag of words without the grammar and sentence order. The frequency of words is used as a feature to train the algorithm and classify the document [138].
38. *BERT (Bidirectional encoder representations from transformers)* – a deep learning strategy for natural language processing (NLP) that helps artificial intelligence programs understand the context of ambiguous words in text [144].
39. *Big data* – large or complicated for general data processing applications to handle [133].
40. *Black box AI* – any type of artificial intelligence that is so complex that its decision-making process cannot be explained in a way that can be easily understood by humans. Black box AI is the opposite of explainable AI [144].
41. *Boltzmann machine* - a type of recurrent neural network in which nodes make binary decisions with some bias. Boltzmann machines can be strung together to make more sophisticated systems such as deep belief networks [144].
42. *Bootstrap* - the program that initializes the operating system (OS) during startup [144].

43. *Bounding box* – a made-up square that serves as a guideline point for object recognition and creates a collision box for that element [132].
44. *Bot* - an automated software program that digitally replicates some type of human activity [144].
    It is an example for front clipping as it has the first part of the initial word removed.
45. *Bottom-up testing* – a specific type of integration testing that tests the lowest components of a code base first. More generally, it refers to a middle phase in software testing that involves taking integrated code units and testing them together, before testing an entire system or code base [144].
46. *Brain-computer interface* – a technology that allows communication between a human or animal brain and an external technology [144].
47. *Branch* – an instruction that tells a computer to begin executing different instructions rather than simply executing the instructions in order. In high-level languages, these are typically referred to as flow control procedures and are built into the language. In assembly programming, branch instructions are built into a CPU [144].
48. *Brute force search* - a query that is not limited by clustering/approximations; it explores all inputs. Typically, more time-consuming and costly, but more thorough [133].

C

49. *Central Processing Unit (CPU)* – the electronic circuitry in a computer executes the commands of a computer program by executing basic arithmetic, logical, control, and I/O operations specified by the instructions [133].
50. *Chatbot* – a software application that imitates human-to-human conversation through text or voice commands [133].
51. *ChatGPT-3* – a language model developed by OpenAI that uses deep learning techniques to generate human-like text [144];

52. *Citizen developer* – a tech-savvy end user who has the ability to create a new software feature or application program from an approved corporate or cloud-based code base, system or structure [144].
53. *Classification* – an algorithm technique that allows machines to assign categories to data points [144].
54. *Cloud robotics* – technology that enables machines or robots to access cloud storage. These machines are equipped with a cloud-hosted provider over a converged infrastructure to access content, respond faster, and display accurate results [144].
55. *Clustering* - an application of unsupervised machine learning that seeks to 'cluster' data into groups by identifying similar patterns [132].
56. *Cog* – a humanoid robot designed by Rodney Brooks's group at MIT as a platform to study robot cognition. It could track faces, grasp objects, and, perhaps most famously, play with a Slinky [132].
57. *Cognitive architecture* – a computer architecture involving non-deterministic, multiple inference processes, as found in neural networks. Cognitive architectures model the human brain and contrast with single processor computers [144].
58. *Cognitive computing* – the term is often referred to as artificial intelligence. It's utilized by marketing teams to minimize AI's mystique in some businesses [133].
59. *Cognitive Map* - a mental representation (otherwise known as a mental palace) which serves an individual to acquire, code, store, recall, and decode information about the relative locations and attributes of phenomena in their environment [137].
60. *Colossus* – computer used by Alan Turing at Bletchley Park, UK during the Second World War to crack the 'Tunny' cipher produced by the Lorenz SZ 40 and SZ 42 machines. Colossus was a semi-fixed-program vacuum tube calculator [132].
61. *Computer vision* – a multidisciplinary scientific discipline that investigates how computers can be programmed to understand digital images or movies at a high

level. It focuses on automating activities that the human visual system can perform [133].

62. *Computational Linguistics (Text Analytics, Text Mining)* – an interdisciplinary field concerned with the computational modeling of natural language [137].
63. *Computational Semantics (Semantic Technology)* – the study of how to automate the construction and reasoning of meaning representations of natural language expressions [137].
64. *Content Enrichment* – the process of applying advanced techniques such as machine learning, artificial intelligence, and language processing to automatically extract meaningful information from your text-based documents [137].
65. *Controlled Vocabulary* – A controlled vocabulary is a curated collection of words and phrases that are relevant to an application or a specific industry. These elements can come with additional properties that indicate both how they behave in common language and what meaning they carry, in terms of topic and more [137].
66. *Conversational AI* - Used by developers to build conversational user interfaces, chatbots and virtual assistants for a variety of use cases. They offer integration into chat interfaces such as messaging platforms, social media, SMS and websites. A conversational AI platform has a developer API so third parties can extend the platform with their own customizations [137].
67. *Convolutional neural networks (CNN)* – are deep artificial neural networks used to classify pictures (e.g., identify what they see), group them by similarity (photo search), and recognize objects in scenes [133].
68. *Co-occurrence* - commonly refers to the presence of different elements in the same document. It is often used in business intelligence to heuristically recognize patterns and guess associations between concepts that are not naturally connected (e.g., the name of an investor often mentioned in articles about startups successfully closing funding rounds could be interpreted as the investor is particularly good at picking his or her investments) [137].

69. *Cross validation* – an approach to reducing overfitting during model development, by iteratively selecting different portions of the data to train and validate a predictive (supervised) machine learning model. It involves splitting a dataset into subsets, where one subset is used for testing the model's performance, and the other subset is used for training the model [132].
70. *Cybernetics* – a study of control and communication in living and man-made systems [132].

D

71. *Data analytics* - qualitative and quantitative techniques and processes used to enhance productivity and business gain. Data is extracted and categorized to identify and analyze behavioral data and patterns, and techniques vary according to organizational requirements [144].
72. *Data augmentation* – the process of adding value to base data by adding information derived from internal and external sources within an enterprise. Data is one of the core assets for an enterprise, making data management essential. Data augmentation can be applied to any form of data, but may be especially useful for customer data, sales patterns, product sales, where additional information can help provide more in-depth insight [144].
73. *Data architect* – a person who practices data architecture, a field of study concerned with designing, generating, deploying, and managing an organization's data architecture. Data architects are frequently involved in AI initiatives [133].
74. *Data cleansing* – the process of altering data in a given storage resource to make sure that it is accurate and correct. There are many ways to pursue data cleansing in various software and data storage architectures; most of them center on the careful review of data sets and the protocols associated with any particular data storage technology [144].
75. *Data compression* – the process of modifying, encoding or converting the bits structure of data in such a way that it consumes less space on disk [144].

76. *Data Discovery* – the process of uncovering data insights and getting those insights to the users who need them, when they need them [137].
77. *Data Drift* – process that occurs when the distribution of the input data changes over time; this is also known as covariate shift [137].
78. *Data Extraction* – the process of collecting or retrieving disparate types of data from a variety of sources, many of which may be poorly organized or completely unstructured [137].
79. *Data exploration* – an informative search used by data consumers to form true analysis from the information gathered. Often, data is gathered in a non-rigid or controlled manner in large bulks. For true analysis, this unorganized bulk of data needs to be narrowed down. This is where data exploration is used to analyze the data and information from the data to form further analysis [144].
80. *Data fusion* – process of getting data from multiple sources in order to build more sophisticated models and understand more about a project. It often means getting combined data on a single subject and combining it for central analysis [144].
81. *Data Ingestion* – the process of obtaining disparate data from multiple sources, restructuring it, and importing it into a common format or repository to make it easy to utilize [137].
82. *Data integration* – a process in which heterogeneous data is retrieved and combined as an incorporated form and structure. Data integration allows different data types (such as data sets, documents and tables) to be merged by users, organizations and applications [144].
83. *Data Labeling* - A technique through which data is marked to make objects recognizable by machines. Information is added to various data types (text, audio, image and video) to create metadata used to train AI models [137].
84. *Data lake* – the process of merging all of the information (both structured and unstructured) into a central repository [133].
85. *Data manager* – a person who is responsible for lawfully obtaining the correct data type for training AI systems by working with data scientists. A data manager

works with data architects to make sure that obtained information is correctly versioned and stored so that it may be analyzed and audited. A data manager must also ensure that the life cycle of the data is managed under legal and organizational standards while also ensuring that appropriate governance is provided over the collection, usage, and disposal of this information [133].

86. *Data miner* – a class of database applications that discovers previously unknown relationships among data, reveals hidden data for a specific purpose or demonstrates common patterns within data sets. While data miner software is widely used within the math and science fields, it has gained widespread popularity among online marketers in the form of spyware and to gather data that will help companies increase online sales [144].
87. *Data mining* – the process of analyzing hidden patterns of data according to different perspectives in order to turn that data into useful and often actionable information. Data is collected and assembled in common areas, such as data warehouses, and data mining algorithms look for patterns that businesses can use to make better decisions, such as decisions that help cut costs, increase revenue, or better serve customers or clients [144].
88. *Data modeling* - a representation of the data structures in a table for a company's database and is a very powerful expression of the company's business requirements. This data model is the guide used by functional and technical analysts in the design and implementation of a database [144].
89. *Data normalization* – the method used whenever the attributes of the dataset have different ranges. It helps to enhance the performance and reliability of a machine learning model. In this article, we will discuss in brief various Normalization techniques in machine learning, why it is used, examples of normalization in an ML model, and much more [88].
90. *Data preprocessing* – the process that involves transforming raw data to well-formed data sets so that data mining analytics can be applied. Raw data is often incomplete and has inconsistent formatting. The adequacy or inadequacy of data

preparation has a direct correlation with the success of any project that involve data analytics [46].

91. *Data processor* – a person who processes data on behalf of a data controller. A data controller decides the purpose and manner to be followed to process the data, while data processors hold and process data, but do not have any responsibility or control over that data [144].
92. *Data Scarcity* – the lack of data that could possibly satisfy the need of the system to increase the accuracy of predictive analytics [137].
93. *Data science* – interdisciplinary subject, data science, which draws from statistics, computer science, and information science, seeks to apply various scientific approaches, methods, and systems to address data issues [133].
94. *Data transformation* - process of converting data or information from one format to another, usually from the format of a source system into the required format of a new destination system. The usual process involves converting documents, but data conversions sometimes involve the conversion of a program from one computer language to another to enable the program to run on a different platform. The usual reason for this data migration is the adoption of a new system that's totally different from the previous one [144].
95. *Data warehouse* – a collection of corporate information and data derived from operational systems and external data sources. A data warehouse is designed to support business decisions by allowing data consolidation, analysis and reporting at different aggregate levels. Data is populated into the DW through the processes of extraction, transformation and loading [144].
96. *Data-driven decision making* – the process that involves making decisions that are backed up by hard data rather than making decisions that are intuitive or based on observation alone [144].
97. *Database management system* - a software package designed to store, retrieve, query and manage data. User interfaces (UIs) allow data to be created, read, updated and deleted by authorized entities [144].

98. *Debugging* – is the task of finding and fixing bugs (or errors) in a software program. Bugs can range from small inconveniences (like ignoring user input in certain circumstances) or significant problems that can cause memory leaks or crashes. Several methods are available for software developers to debug a program, including using a debugger or analyzing crash reports [145].
99. *Decision tree* – a flowchart-like representation of data that graphically resembles a tree that has been drawn upside down. In this analogy, the root of the tree is a decision that has to be made, the tree's branches are actions that can be taken and the tree's leaves are potential decision outcomes [106].
100. *Deep Blue* – a supercomputer developed by IBM specifically for playing chess and was best known for being the first artificial intelligence construct to ever win a chess match against a reigning world champion, Grandmaster Garry Kasparov, under regular time controls [144].
101. *Deep convolutional inverse graphics network* – a particular type of convolutional neural network that is aimed at relating graphics representations to images. Experts explain that a deep convolutional inverse graphics network uses a 'vision as inverse graphics' paradigm that uses elements like lighting, object location, texture and other aspects of image design for very sophisticated image processing [144].
102. *Deepfake* – a term for videos and presentations enhanced by artificial intelligence and other modern technology to present falsified results. One of the best examples of deepfakes involves the use of image processing to produce video of celebrities, politicians or others saying or doing things that they never actually said or did [144].
103. *Deep learning* – an artificial intelligence technique that mimics the human brain by learning from how data is structured rather than a pre-programmed algorithm [133].

104. *Deep neural network* – a neural network with a certain level of complexity, a neural network with more than two layers. Deep neural networks use sophisticated mathematical modeling to process data in complex ways [144].

105. *Deep Q-Networks (DQN)* – neural networks (and/or related tools) that utilize deep Q learning in order to provide models such as the simulation of intelligent video game play. Rather than being a specific name for a specific neural network build, Deep Q-Networks may be composed of convolutional neural networks and other structures that use specific methods to learn about various processes [144].

106. *Deep reinforcement learning (Deep RL)* – an approach to machine learning that blends reinforcement learning techniques with strategies for deep learning [144].

107. *Deep residual network (deep ResNet)* – a type of specialized neural network that helps to handle more sophisticated deep learning tasks and models. It has received quite a bit of attention at recent IT conventions, and is being considered for helping with the training of deep networks [144].

108. *Deep stubborn network* - network in which components work against themselves to produce more refined results [144].

109. *Deep Q-learning* – an algorithm that breaks the chain in order to find the optimal Q-value function. It determines this by combining Q-learning and a neural network. The uses of the deep Q-learning algorithm can be stated as finding the input and the optimal Q-value for all possible actions as the output [50].

110. *Derivative work* – a new work derived from an original work protected under copyright law [47].

111. Derivative work right – the legal permission to develop a new work derived from an original work protected under copyright law. Derivative work rights are only granted for derivative works with original content, versus duplicated copyrighted material. The original author’s permission to transform or adapt an original work by the original owner is the essence of a derivative work right [144].

112. *Did You Mean (DYM)* – an NLP function used in search applications to identify typos in a query or suggest similar queries that could produce results in the search database being used [137].

113. *Digital ecosystem* – several software platforms or cloud services that work in tandem across a network [137].

114. *Digital signal processing* – process of analyzing and modifying a signal to optimize or improve its efficiency or performance. It involves applying various mathematical and computational algorithms to analog and digital signals to produce a signal that's of higher quality than the original signal [144].

115. *Digital twin* – a virtual representation of an entity or system that exists in the physical world [144].

116. *Digitize* – to convert the representation of an object in an analog signal into a series of discrete points or samples. This involves converting existing non-digital information or data into a digital form with the intention of storing, altering or sharing this data with electronic devices [144].

117. *Disambiguation* – the process of removing confusion around terms that express more than one meaning and can lead to different interpretations of the same string of text [137].

118. *Divide-and-conquer* – a problem-solving algorithm that breaks a problem into subproblems that are similar to the original problem, recursively solves the subproblems, and finally combines the solutions to the subproblems to solve the original problem [137].

E

119. *Embedded hypervisor* – a type of virtualization hypervisor that is natively installed, programmed or embedded in a computing device or system. It is a pre-integrated hypervisor that is delivered as a component of a resident computer, server or device [144].

120. *Embedded intelligence* – a term for a self-referential process in technology where a given system or program has the ability to analyze its own operations [144].

121. *Emotion AI (aka Affective Computing)* – AI that can analyze the emotional state of a user (via computer vision, audio/voice input, sensors and/or software logic). It can initiate responses by performing specific, personalized actions to fit the mood of the customer [137].

122. *Emotion recognition* – a technique used in software that allows a program to 'read' the emotions on a human face using advanced image processing. Companies have been experimenting with combining sophisticated algorithms with image processing techniques that have emerged in the past ten years to understand more about what an image or a video of a person's face tells us about how he/she is feeling [144].

123. *Enterprise nervous system (ENS)* – the connection of network elements and components to create an intelligent whole, a comprehensive system serving enterprise goals. The idea is that parts of an IT architecture can be connected in much the same way that the human central nervous system is connected — to serve common goals through smart collaborative processes [144].

124. *Entity* – any noun, word or phrase in a document that refers to a concept, person, object, abstract or otherwise (e.g., car, Microsoft, New York City). Measurable elements are also included in this group (e.g., 200 pounds, 14 fl. oz.) [137].

125. *Entity annotation* – labeling unstructured phrases with data so that a computer can comprehend them is known as information extraction. This might include labeling all persons, organizations, and locations in a document [137].

126. *ETL (Entity Recognition, Extraction)* – entity extraction is an NLP function that serves to identify relevant entities in a document [137].

127. *Error-correcting code (ECC)* – a type of computer data storage specifically designed to detect, correct and monitor most common kinds of interior data corruption. As data is processed, ECC memory equipped with a special algorithm constantly scans and corrects single-bit memory errors. This ensures that no erroneous or corrupt data is accidentally stored in memory. It is typically found

and used in systems with high-value data such as scientific and financial computing systems [144].

128. *Ethical AI* – can be defined as: an approach in AI programming that should deal with concerns regarding artificially intelligent systems, i.e., biases may significantly impact machine learning algorithms that are trained from data, ranging from gender to race to age to economic status and everything in between [133].

129. *Event-driven architecture (EDA)* – a software architecture pattern that promotes the production, detection and consumption of, and reaction to, significant changes in a system's state (known as events). This is applied through the design and implementation of applications and systems that transmit events among loosely coupled software components and services [144].

130. *Expert Automation and Augmentation Software (EAAS)* – a type of software resource that works on the process of automating highly cognitive behaviors or tasks. These types of software programs and systems are said to be taking over the work of semi-skilled white-collar workers in the knowledge economy [144].

131. *Expert system* – a computer program that is designed to emulate and mimic human intelligence, skills or behavior [144].

132. *Extreme learning machine (ELM)* is a particular kind of machine learning setup in which a single layer or multiple layers apply. ELM includes numbers of hidden neurons where the input weights are assigned randomly. Extreme learning machines use the concept of random projection and early perceptron models to do specific kinds of problem-solving [144].

133. *Explainable AI (XAI)* – artificial intelligence that can document how specific outcomes were generated in such a way that ordinary humans can understand the process. The goal of XAI is to make sure that artificial intelligence programs are transparent regarding both the purpose they serve and how they work [144].

F

134. Facial recognition is a type of biometric technology that uses data to verify the presence of a human being's face in a digital capture. There are two main uses for facial recognition software: recognition and authentication [144].

135. *False negatives* – an incorrect prediction where a model mistakenly assumes an input does not have a required result when one actually exists. (Actual Yes, Predicted No) [133].

136. *False positives* – an error in a model's prediction of the presence of the desired result in input when it is not present (Actual No, Predicted Yes) [133].

137. *Feature engineering* – the process of assigning attribute-value pairs to a dataset that's stored as a table. Attribute-value pairs may also be referred to as features or descriptive properties [144].

138. *Feature extraction* – in machine learning, computer vision, or pattern recognition, feature extraction starts with dividing the image or data into bounding boxes and extracting one single feature out of the boxes. The features are extracted, pooled, and fed to a supervised vector machine to predict the output [137].

139. *Federated learning* – an approach to machine learning where a model is trained on data where the data exists (in multiple locations), rather than the traditional approach of moving the data to a central location for model training [144].

140. *Feed* (12, verb) – to feed information into a computer means to gradually put it into it [144].

141. *Feedforward neural network* – is the simplest artificial neural network where information exits through the output node and doesn't come back for analysis. Data only flows in the forward direction and doesn't form a loop [144].

142. *Foundation model* – a deep learning algorithm that has been pre-trained with extremely large data sets scraped from the public internet [144].

143. *Forward chaining* – a method in which a machine must go from an issue to solving it. The AI must evaluate many options to determine which hypotheses are relevant to the problem [144].

144. *Frame* – a frame is a unit of data. A frame works to help identify data packets used in networking and telecommunications structures. Frames also help to determine how data receivers interpret a stream of data from a source.

145. *Function* – a unit of code that is often defined by its role within a greater code structure. Specifically, a function contains a unit of code that works on various inputs, many of which are variables, and produces concrete results involving changes to variable values or actual operations based on the inputs.

146. *Fuzzy logic* – a logic operations method based on many-valued logic rather than binary logic (two-valued logic). Two-valued logic often considers 0 to be false and 1 to be true. However, fuzzy logic deals with truth values between 0 and 1, and these values are considered as intensity (degrees) of truth [144].

G

147. *Game AI* – a type of AI that uses an algorithm to replace randomness in video games. It's a computational behavior used by non-player characters to generate humanlike intelligence and reactive behaviors taken by the player in games. It is one of the most searched Artificial intelligence terms [133].

148. *Game theory* - the study of using mathematical models to assess interactive systems. Many experts describe it as the analysis of interplay between independent rational decision-makers or actors. Game theory is useful in many different kinds of research, such as projects that take into account elements of human psychology [133].

149. *Garbage collection* – a type of memory management. It automatically cleans up unused objects and pointers in memory, allowing the resources to be used again. Some programming languages have built-in garbage collection, while others require custom functions to manage unused memory [145].

150. *Generative adversarial network (GAN)* – a machine learning approach in which two neural networks compete to create new data with the same statistics as the training set. GANs, for example, are utilized in fashion, art, and marketing, but they are also becoming increasingly popular among malicious attackers to spread false news [133].

151. *Generative AI* – a broad label that’s used to describe any type of artificial intelligence that can be used to create new text, images, video, audio, code or synthetic data [144].

152. *Genetic Algorithm* – a method of searching for the best solution based on a defined scoring method. It is used to find optimized solutions for complex problems based on the theory of natural selection [47].

153. *Gradient descent* – a common approach used in supervised machine learning, where models are trained to fit the training data, by minimizing the error in predictions in an iterative fashion.
In gradient descent, small 'steps' are taken in different 'directions' in search of the smallest error [137].

154. *Gradient descent algorithm* – a strategy that helps to refine machine learning operations. The gradient descent algorithm works toward adjusting the input weights of neurons in artificial neural networks and finding local minima or global minima in order to optimize a problem [137].

155. *Graph Neural Network (GNN)* – a class of deep learning methods designed to make predictions on data described by graphs [132].
Graphs are a way of representing data, relationships and their complexity. GNNs are neural networks that can be directly applied to graphs, and provide a way to generate node-level, edge-level, and graph-level predictions [144].

H

156. *Hebbian theory* – a theoretical type of cell activation model in artificial neural networks that assesses the concept of ‘synaptic plasticity’ or dynamic strengthening or weakening of synapses over time according to input factors. The term was named after the Canadian psychologist Donald Hebb, who proposed the theory of Hebbian learning, which states that neurons that fire together, wire together [132].

157. *Heuristic* – rules or methods used to solve problems that are too complex or too large for brute-force approaches [133].

158. *Heuristic programming* – an approach to the idea of artificial intelligence by solving problems using experience-based rules or protocols [144].

159. *Hopfield network* – a specific type of recurrent artificial neural network based on the research of American physicist John Hopfield in the 1980s on associative neural network models. Hopfield networks are associated with the concept of simulating human memory through pattern recognition and storage. The network was named after its inventor [144].

160. *Hybrid AI* - any artificial intelligence technology that combines multiple AI methodologies. In NLP, this often means that a workflow will leverage both symbolic and machine learning techniques [145].

161. *Hyperautomation* – a process that combines artificial intelligence, machine learning, and robotic process automation to automate routine business tasks. It allows computer systems to perform simple, repetitive tasks that would otherwise take up a lot of time, freeing up human employees to give them more time for higher-value work [145].

162. *Hyperparameter* – is a parameter that impacts how a model learns. They're usually set manually outside of the model [133].

163. *Hypervisor* – a software program that manages one or more virtual machines (VMs). It is used to create, start, stop, and reset VMs. The hypervisor allows each VM or 'guest' to access the physical hardware, such as the CPU, RAM, and storage. It can also limit how many system resources each VM can use so that multiple VMs can run simultaneously on a single system [145].

I

164. *Image recognition* – the capacity of software to recognize things, places, people, text, and activities in photographs[133].

165. *Image segmentation* – the process of dividing a digital image into several parts/fragments to make the representation of an image simpler to analyze. Segmentation separates whole images into pixel categories, which may be

labeled and classified. Segmentation is the act of putting a bounding box around the target object in a picture and drawing a pixel-by-pixel outline of that object, leaving the background intact[133].

166. *ImageNet* – an extensive visual database that is intended to be used in computer vision software development. ImageNet has hand-annotated over 14 million URLs of images to reveal what objects are shown. Bounding boxes are also provided in at least one million of the pictures. It is one of the most searched Artificial intelligence terms [133].

167. *ImageNet Challenge* – the first phase of the competition is a research project. Teams evaluate their algorithms on the supplied data set and compete to achieve higher accuracy on several visual recognition problems [133].

168. *Inference* – in AI, inference is the process of making a prediction from a model that has already been trained [133].

169. *Intelligent Document Processing (IDP)* or *Intelligent Document Extraction* and Processing (IDEP) – the ability to automically read and convert unstructured and semi-structured data, identify usable data and extract it, then leverage it via automated processes. IDP is often an enabling technology for Robotic Process Automation (RPA) tasks [137].

170. *Interface agents* – are intelligent tutoring systems and context-sensitive help systems [144].

171. *Interoperability* – the ability of a system or component to function effectively with other systems or components [133].

J

172. *Junction tree algorithm* – a technique for extracting marginalization in general graphs. In other words, it entails performing belief propagation on a modified graph known as a junction tree. Because the graph divides into several parts of data and variable nodes are the branches, it is referred to as a tree [133].

K

173. *K-nearest neighbors (KNN)* - a k-nearest-neighbor algorithm, often abbreviated k-nn, is an approach to data classification that estimates how likely

a data point is to be a member of one group or the other depending on what group the data points nearest to it are in [144].

174. *Kernel* – is a mathematical algorithm that performs transformations on data so it can be further analyzed, adjusted and addressed. It essentially functions as the brain of the machine learning model, allowing for more accurate predictions and analyses from complex data sets [132].

175. *Keyphrase Extraction* – mutiple words that describe the main ideas and essence of text in documents [137].

176. *Knowledge based systems (KBS)* – is an application that utilizes a knowledge base to tackle complex issues. The phrase is broad and refers to a variety of systems. The one common element that all knowledge-based systems have in common is the desire to express knowledge explicitly and a reasoning mechanism that enables it to generate new information. As a result, a knowledge-based system has two characteristics: a knowledge base and an inference engine. It is one of the most searched Artificial Intelligence terms [133].

177. *Knowledge engineer* – a professional engaged in the science of building advanced logic into computer systems in order to try to simulate human decision-making and high-level cognitive tasks. A knowledge engineer supplies some or all of the 'knowledge' that is eventually built into the technology [144].

178. *Knowledge extractions* – the extraction of knowledge from technical documentation, XML, unstructured datasets, or relational databases. Knowledge elements are extracted by running specific queries that represent the data best [138].

179. *Knowledge Graph* – a graph of concepts whose value resides in its ability to meaningfully represent a portion of reality, specialized or otherwise. Every concept is linked to at least one other concept, and the quality of this connection can belong to different classes [137].

180. *Knowledge Model* – A process of creating a computer interpretable model of knowledge or standards about a language, domain, or process(es). It is expressed

in a data structure that enables the knowledge to be stored in a database and be interpreted by software [144].

181. *Knowledge representation* – field that involves considering artificial intelligence and how it presents some sort of knowledge, usually regarding a closed system. IT professionals and others may monitor and evaluate an artificial intelligence system to get a better idea of its simulation of human knowledge, or its role in presenting the data about focus input [144].

L

182. *Lambda calculus* – is a type of formal system from mathematical logic used in computer science for function definition, application and recursion [144].

183. *Large language model (LLM)* – a type of machine learning model that can perform a variety of natural language processing (NLP) tasks, including generating and classifying text, answering questions in a conversational manner and translating text from one language to another [133].

184. *Learning algorithm* – a set of instructions used in machine learning that allows a computer program to imitate the way a human gets better at characterizing some types of information. The math and logic that supports a learning algorithm can update itself over time (without human intervention) as the programming becomes exposed to more data [144].

185. *Least mean square (LMS)* algorithm is a type of filter used in machine learning that uses stochastic gradient descent in sophisticated ways – professionals describe it as an adaptive filter that helps to deal with signal processing in various ways [145].

186. *Lisp* - a family of computer programming languages that originated in 1958 and has since undergone a number of changes and dialects. It is considered the second-oldest high-level programming language in use today, after Fortran [144].

187. *Limited memory* – systems with short-term memory limited to a given timeframe. Limited memory AI derives knowledge from real-time experiences

or events and stores it in the database. When a problem occurs, it gives out redundant results [138].

188. *Location intelligence* – the ability a technology has to provide other software applications with the geographical location of a specific computing device or person. This capability works by actively or passively determining the target's geographical coordinates in respect to a reference point [144].

189. *Logic programming* – programming paradigm that is primarily based on formal logic. A set of statements in logical form, expressing facts and rules about a problem domain, is any program written in a logic programming language. Answer set programming (ASP), also known as solution-oriented programming (SOP), and Datalog are examples of significant logic programming languages [133].

190. *Logic theorist* – the first running artificial intelligence program [71].

191. *LangOps (Language Operations*) - The workflows and practices that support the training, creation, testing, production deployment and ongoing curation of language models and natural language solutions [144].

192. *Long short-term memory (LSTM)* – units or blocks are part of a recurrent neural network structure. Recurrent neural networks are made to utilize certain types of artificial memory processes that can help these artificial intelligence programs to more effectively imitate human thought [137].

193. *Low-code/no-code (LC/NC)* development – an environment where visual drag-and-drop applications or similar tools allow individuals and teams to program applications without a lot of linear coding. These types of systems help the IT world to deal with a lack of skilled developers and streamline the emergence of new applications and interfaces [144].

M

194. *Machine bias* – is the tendency of a machine learning model to make inaccurate or unfair predictions because there are systematic errors in the ML model or the data used to train the model [144].

195. *Machine consciousness* – is an awareness state attained by machines after a specific period of time to decipher human emotions and expressions. It's being experimented on in the robotics industry [137].

196. *Machine learning* – the term refers to the ability of computers to learn without being explicitly programmed. Computers 'learn' via patterns they detect and adapt their behavior as a result. It is one of the most searched Artificial Intelligence terms. On the other hand, machine learning engineering is a great career opportunity nowadays [133].

197. *Machine learning operations (MLOps*) is an approach to managing the entire lifecycle of a machine learning model — including its training, tuning, everyday use in a production environment and retirement [133].

198. *Machine learning as a service (MLaaS)* is a range of machine learning (ML) tools offered by cloud service providers. As of this writing, popular MLaaS offerings include Amazon SageMaker, Microsoft Azure ML, IBM Watson Machine Learning and Google Cloud ML [145].

199. *Machine learning workflow* describes the processes involved in machine learning work. Various stages help to universalize the process of building and maintaining machine learning networks [145].

200. *Machine perception* – the capacity for a system to acquire and comprehend information from the environment in the same manner that humans do with their senses [133].

201. *Machine translation* – an application of NLP for language translation (human-to-human) in text- and speech-based conversations [133].

202. *Machine vision* (MV) – an integrated mechanical-optical-electronic-software technology which makes use of optical instrumentation, digital video, electromagnetic sensing, mechanics and image processing technology. The technology's goal is optical and non-contact sensing to receive and analyze a real image in order to provide more information. Machine vision technology is widely used in monitoring and controlling a wide range of applications [145].

203. *Memory* – the term refers to any information or data, often in binary format, that a machine or technology can recall and use. There are many different kinds of memory in conventional computers and other devices, and they differ based on the complex design of the hardware in which they're stored [144].

204. *Metadata* – data about data, or data that provides information on one or more aspects of the data [145].

205. *Metaverse* – an immersive, interactive environment generated by a computer. Although there is no unified agreement on what the metaverse will look like — or how individuals will interact with it. [144].

206. *Mini-batch* – a subset of a larger dataset that is used for training a model. Instead of using the entire dataset to train the model at once, the dataset is divided into smaller batches, and the model is trained on each of these batches in sequence [16].

207. *Mini-batch training* is commonly used in deep learning, where large datasets are used to train complex models with many parameters. By using mini-batches, the model can be updated more frequently, which can speed up the training process and improve the accuracy of the final model [16].

208. *Multi-agent system* – a loosely coupled network of software agents that interact to solve problems that are beyond the individual capacities or knowledge of each problem solver [85].

209. *Multi-class classification* – a classification task with more than two classes. Each sample can only be labeled as one class. For example, classification using features extracted from a set of images of fruit, where each image may either be of an orange, an apple, or a pear [144].

210. *Multiple instance learning* – a form of weakly supervised learning where training instances are arranged in sets, called bags, and a label is provided for the entire bag, opposedly to the instances themselves. This allows to leverage weakly labeled data, which is present in many business problems as labeling data is often costly [75].

211. *Multi-layer neural network* – a kind of neural network that contains more than one layer of artificial neurons or nodes. They differ widely in design. It is important to note that while single-layer neural networks were useful early in the evolution of AI, the vast majority of networks used today have a multi-layer model [144].

212. Multi-layer perceptron – a feed-forward neural network complement. It has three layers: an input layer, a hidden layer, and an output layer, as shown in Fig [86].

213. *Multimodal artificial intelligence* – an approach to AI which incorporates multiple types of data. For example, a speech-to-text model that is typically trained on audio and text data, could include image data of lip movements taken from video recordings [144].

214. *Multi-task learning* – a single shared machine learning model that can perform multiple different (albeit related) tasks. Multi-Task Learning offers advantages like improved data efficiency, faster model convergence, and reduced model overfitting due to shared representations [66].

215. *Multi-view learning* – an emerging direction in machine learning which considers learning with multiple views to improve the generalization performance. Multi-view learning is also known as data fusion or data integration from multiple feature sets [69].

N

216. *Naive Bayes classifier* – an algorithm that uses Bayes' theorem to classify objects. Naive Bayes classifiers assume strong, or naive, independence between attributes of data points. Popular uses of naive Bayes classifiers include spam filters, text analysis and medical diagnosis. These classifiers are widely used for machine learning because they are simple to implement [144].

217. *Nanobots* – molecular-sized robots measured on a nanoscale and programmed to accomplish a specific task within the human body. The concept is used to create smart vaccines, cancer therapy, and immunotherapy through painless methods of medication [138].

218. *Narrow AI* – AI focussed on solving a specific problem. For example, an AI built to identify cancerous tumours in breast scans, would not automatically be able to detect tumours in other parts of the body without significant rework [133].

219. *Natural language generation (NLG)* – process which converts structured data into text or voice that humans can comprehend. NLG is concerned with what a machine writes or says as the conclusion of the communication process [133].

220. *Natural language processing (NLP)* – is the capacity of computers to comprehend or extract meaning from natural human languages. NLP generally entails computer interpretation of text or speech recognition [133].

221. *Natural Language Toolkit (NLTK)* is a platform used for building Python programs that work with human language data for application in statistical natural language processing (NLP)[144].

222. *Neuralink* is an American company using neurotechnology to research the possibility of future innovations in high-tech neurology applications. Neuralink has gained visibility as a project backed by Elon Musk, and as a ground-breaking way of looking at neurotech in general [144].

223. *Neural network* – is a computer system that functions as the brain of a human. Although researchers are still attempting to construct a computer model of the human brain, current neural networks can already accomplish many things regarding speech, vision, and board game strategy [133].

224. *Neural Turing machine (NTM)* is a technology that uses neural network methodologies to achieve the capability to verify algorithms and do other computational work. It is based on the mid-20th century work of renowned data scientist Alan Turing [133].

225. *NeuroEvolution of Augmenting Topologies or NEAT* is often described as a genetic solution for improving neural networks. The NEAT concept can be used to provide a new model for selecting typologies for a neural network and for initializing weights [144].

226. *Neuroinformatics* refers to a research field that focuses on organizing neuroscience data through analytical tools and computational models. It combines data across all scales and levels of neuroscience in order to understand the complex functions of the brain and work toward treatments for brain-related illness. Neuroinformatics involves the techniques and tools for acquiring, sharing, storing, publishing, analyzing, modeling, visualizing and simulating data [144].

227. *Neuromining* – the process of applying various behavioral intelligence and machine learning (ML) techniques in order to analyze human behavior. The goal is to understand human behavior in depth to manipulate and influence it at scale [144].

228. *Neuromorphic computing* utilizes an engineering approach or method based on the activity of the biological brain. This type of approach can make technologies more versatile and adaptable, and promote more vibrant results than other types of traditional architectures, for instance, the von Neumann architecture that is so useful in traditional hardware design [144].

229. *Neuro Symbolic Artificial Intelligence* – an advanced version of artificial intelligence (AI) that improves how a neural network arrives at a decision by adding classical rules-based (symbolic) AI to the process. This hybrid approach requires less training data and makes it possible for humans to track how AI programming made a decision [144].

230. *Neurotechnology* as a new prominent tech term describes any technology that helps us to understand brain function, or enables a direct connection of technology with the human nervous system [144].

231. *Node* – basic unit of a data structure, like a stack or a linked list, or a queue, that represents an underlying value or a variable [144].

O

232. *OpenAI* - a for-profit technological company that conducts scientific research in artificial intelligence and machine learning. ChatGPT, a conversational AI platform, is their latest invention. Developed on the principle of reinforcement

learning, ChatGPT is equipped with advanced AI capabilities to complete human-dependent tasks [144].

233. *Open-source license* – a permissive stance on what users can do with the original source code [47].

234. *Operating point* – the point chosen to define when a case is positive or not. The location of this point will determine the model performance measured by true positives, false positives, true negatives and false negatives [132].

235. *Optical character recognition (OCR)* – the process of converting text images (typed, handwritten, or printed) either electronically or manually into machine-encoded text is called OCR [133].

236. *Optimization* - an act, process, or methodology of making something (such as a design, system, or decision) as fully perfect, functional, or effective as possible [132].

237. *Over-parameterization* - refers to the scenario where the number of parameters of the model exceeds the size of the training dataset or a similar threshold [92].

238. *Overfitting* - a term frequently used in AI. Overfitting is a sign of machine learning training in which the algorithm can only work on or identify specific examples from the training data. A functioning model should be able to generalize patterns seen in the data to tackle new instances. It is one of the most searched Artificial intelligence terms [133].

239. *Oversampling* is the process of replicating data points in a dataset that is imbalanced. This is done in order to create a more balanced dataset that can be used to train a machine learning model [115].

P

240. *Pattern recognition* - the difference between pattern recognition and machine learning is sometimes blurred, but the goal of this discipline is to discover patterns and trends in data [133].

241. *Perceptron* is a machine learning algorithm that helps provide classified outcomes for computing. It dates back to the 1950s and represents a fundamental example of how machine learning algorithms work to develop data [86].

242. *Perceptual computing* is a new and somewhat confusing term in IT. The common definition of perceptual computing is a general advancement in technology where computers are better able to sense or analyze the environment around them and respond accordingly. Perceptual computing has a lot of potential to change the end-user interfaces through which humans interact with computers [137].

243. *Preattentive processing* - level of processing at which simple features are coded spatially in parallel and a later stage at which focused attention is required to conjoin the separate features into coherent objects [49].

244. *Pre-processing* refers to the technique of preparing (cleaning and organizing) the raw data to make it suitable for building and training Machine Learning models [68].

245. *Pre-training* in AI refers to training a model with one task to help it form parameters that can be used in other tasks [68].

246. *Precondition* – a fact or set of facts that must be true before an action can be taken; e.g., before a robot can pick up an object, its hand must be empty [139].

247. *Predictive analytics* - this type of analytics is designed to forecast what will happen in a specified time frame based on previous data and patterns. We prepared article about preventing bias in Predictive Analytics before. You might want to check that out [133].

248. *Predictive modeling* is a process through which a future outcome or behavior is predicted based on the past and current data at hand [137].

249. *Predictive text* is a technology used in text messaging that suggests words to a user based on the letters being entered and the overall context of the phrase being written. This technology is mainly used in smartphones and tablets to make writing and text messaging easier and more efficient [137].

250. *Principal component analysis (PCA)* is a technique used for identification of a smaller number of uncorrelated variables known as principal components from a larger set of data [137].

251. *Prolog – Programmation en Logique* (Programming in Logic) or Prolog is a high-level programming language that has its roots in first-order logic or first-order predicate calculus [132].

252. *Prompt-based learning* is a strategy that machine learning engineers can use to train large language models (LLMs) so the same model can be used for different tasks without re-training [133].

253. *Pruning* is the practice of limiting undesirable answers to a problem in an AI system using a search algorithm. The number of alternatives available to the AI system is reduced [133].

Q

254. *Q Learning* is a reinforcement learning technique that offers valuable insight into how computers can learn from experience. In essence, it is a way for machine learning models to utilize trial and error in order to gain understanding of environments in an autonomous fashion. More specifically, Q Learning involves creating an agent (computer program) with a specific goal or task. This agent interacts with its environment while receiving feedback based on the actions it performs. This feedback is used to develop a function known as the Q-Function, which maps out the ideal set of actions from any given state that will give the highest reward [50].

255. *Quantum computing* – Quantum computing is applied to compute using quantum-mechanical phenomena such as superposition and entanglement. This calculation may be done theoretically or in reality, and a quantum computer is used to perform it. It is one of the most searched Artificial intelligence terms. You can also learn about quantum machine learning in this article [133].

R

256. *Radial basis function network* – a type of supervised artificial neural network that uses supervised machine learning (ML) to function as a nonlinear classifier.

Nonlinear classifiers use sophisticated functions to go further in analysis than simple linear classifiers that work on lower-dimensional vectors [144].

257. *Random forest* – a consensus algorithm used in supervised machine learning (ML) to solve regression and classification problems. Each random forest is comprised of multiple decision trees that work together as an ensemble to produce one prediction [137].

258. *Random walk* – a somewhat popular mathematical construct that is used in computer science, and now in machine learning. It is described as a "stochastic" process because it works through the application of random variables. The random walk essentially tracks incremental steps by a particular modeled intelligence or digital rational actor [137].

259. *Raw data* refers to any data object that hasn't undergone thorough processing, either manually or through automated computer software. Raw data may be gathered from various processes and IT resources [144].

260. *'Rational agent'* is a concept that guides the use of game theory and decision theory in applying artificial intelligence to various real-world scenarios. The rational agent is a theoretical entity based on a realistic model, that has preferences for advantageous outcomes, and will seek to achieve them in a learning scenario [137].

261. *Reactive machines:* can analyze, perceive, and make predictions about experiences, but do not store data; they react to situations and act based on the given moment [137].

262. *Recommendation engine* is a machine learning (ML) system that uses explicit and implicit end user feedback to make predictions about what digital content — including ads — an end user might be interested in viewing [145].

263. *Recurrent neural network (RNN)* - a recurrent neural network (RNN) is a sort of neural network that understands sequential data and patterns, generates outputs resulting from those computations, and learns [133].

264. *Reinforcement learning (RL)* – a type of machine learning that involves an algorithm that learns by interacting with its surroundings and is then penalized or rewarded depending on how it acts [133].

265. *Responsible AI* – can be defined as: the development and use of artificial intelligence (AI) in a way that is ethically and socially trustworthy. Legal accountability is an important factor driving responsible AI initiatives [50].

266. *Restricted Boltzmann machine* – a type of artificial neural network invented by Geoff Hinton, a pioneer in machine learning and neural network design [132].

267. *Reward path* is a path that an agent takes in order to obtain cumulative rewards. This terminology isn't really used very much on its own in machine learning, but the concept of reward is central to many machine learning algorithms and Markov decision process models [137].

268. *Robotics* – the engineering and operation of machines that can autonomously or semi-autonomously perform physical tasks on behalf of a human. Typically, robots perform tasks that are either highly repetitive or too dangerous for a human to carry out safely [144].

269. *Robotic process automation (RPA)* is a technology that uses software agents (bots) to carry out routine clerical tasks without human assistance. RPA is useful for automating business processes that are rules-based and repetitive [144].

270. *Root cause analysis* – is a problem-solving method which is used to pinpoint the exact cause of a problem or event [137].

S

271. *Self-organizing map* – a type of artificial neural network that uses unsupervised learning to build a two-dimensional map of a problem space. The key difference between a self-organizing map and other approaches to problem solving is that a self-organizing map uses competitive learning rather than error-correction learning such as backpropagation with gradient descent [137].

272. *Self-supervised learning* (SSL) is an approach to machine learning allows machine learning algorithms to use observed inputs to predict unknown inputs [144].
273. *Semi-supervised learning* – a method used to enable machines to classify both tangible and intangible objects. The objects the machines need to classify or identify could be as varied as inferring the learning patterns of students from classroom videos to drawing inferences from data theft attempts on servers. To learn and infer about objects, machines are provided with labeled, shallow information about various types of data based on which the machines need to learn from large, structured and unstructured data they receive regularly. [144]
274. *Sensitivity* – the ability of a classification model to correctly identify individuals *with* a condition. This is also known as the true positive rate and recall. It is defined as the number of true positives over true positives and false negatives [137].
275. *Sentiment analysis* – the study of the viewpoints and opinions within a piece of text is generally to determine the writer's attitude toward something. It is one of the most searched Artificial intelligence terms [133].
276. *Statistical model* is a mathematical model that embodies a set of statistical assumptions concerning the generation of sample data (and similar data from a larger population). A statistical model represents, often in considerably idealized form, the data-generating process [132].
277. *Stochastic* – the term refers to data which has a random probability that may be analyzed via statistics. Although individual events cannot be predicted, analyzing the distribution of random stochastic variables may result in a pattern [133].
278. *Strong AI* – the main objective of AI technology is to produce systems that are as intelligent and skilled as the human mind. It is one of the most searched Artificial intelligence terms [133].
279. *Structured data* – clearly defined data with easily searchable patterns [145].

280. *Supervised learning* - Supervised Machine Learning is a type of machine learning in which output data trains the machine to produce the correct algorithms, such as a teacher guiding a student. It's more prevalent than unsupervised learning [133].

281. *Support vector machine* (SVM) are a set of supervised learning methods used for classification, regression and outliers detection [102].

282. *Swarm behavior* – from the perspective of the mathematical modeler, it is an emergent behavior based on basic rules that people follow without any central coordination [133].

T

283. *Tensorflow* - is an open-source machine learning library that's used for various purposes, including neural networks. It was released under the Apache 2.0 open-source license in 2015 and is utilized for research and production at Google. It is one of the most searched Artificial intelligence terms [133].

284. *Text processing* - text processing is the automated mechanization of the creation or modification of electronic text. Computer commands are usually involved in text processing, which help in creating new content or bringing changes to content, searching or replacing content, formatting the content or generating a refined report of the content [137].

285. *Time-sharing* - is the distribution of a computing resource to many users via multiprogramming or multitasking. This was introduced in the 1960s when computers were still too expensive to be prolific, so the solution was to allow many users to make use of one computer by affording each one time-shares, a specific amount of time that a user could access the computer. This allowed many people to use a computer, which most people could not afford, without actually owning one. This is now only a historical way of using computers as there is no need to queue up users since modern computers, even the smallest ones, are able to cater to multiple users because of fast processors and multi-tasking operating systems [144].

286. *Transfer learning* is a process that allows a pre-trained machine learning (ML) model to be used as a starting point for training a new model. Transfer learning reduces the cost of building the new model from scratch and speeds up the training process [144].

287. *Transformer model* is a type of deep learning architecture commonly used in machine learning (ML) and artificial intelligence (AI) for natural language processing (NLP) tasks [132].

288. *Tree search* - is a model used to represent hierarchical data. In correspondence to natural trees, it has nodes, leaves and branches. A commonly mentioned tree is a binary tree, in which each node contains leads to two other nodes [132].

289. *Turing test* - test that assesses the capacity of a machine to mimic human behavior. The evaluation consists of a real-world conversation between a person and another individual and a computer, in which the participants are assessed on their understanding [133].

U

290. *Unauthorized derivative works* – a new creative work that is based on or derived from an existing original work, such as a sequel, adaptation, or translation, created without the permission of the copyright owner of the original work. This type of work is considered to be a violation of copyright law, as it infringes on the exclusive rights of the copyright owner to control the use and distribution of their original work and any derivative works based on it [47].

291. *Underspecification* is the failure to specify in enough detail. In machine learning and AI, the underspecification of training samples can result in vastly different predictions for edge cases, even when very similar models [109].

292. *Underfitting* - the counterpart of overfitting, happens when a machine learning model is not complex enough to accurately capture relationships between a dataset's features and a target variable. An underfitted model results in problematic or erroneous outcomes on new data, or data that it wasn't trained on, and often performs poorly even on training data [108].

293. *Undersampling* is a technique to balance uneven datasets by keeping all of the data in the minority class and decreasing the size of the majority class. It is one of several techniques data scientists can use to extract more accurate information from originally imbalanced datasets. Though it has disadvantages, such as the loss of potentially important information, it remains a common and important skill for data scientists [116].

294. *Unix* - a portable, multitasking, multiuser, time-sharing operating system (OS) originally developed in 1969 by a group of employees at AT&T. Unix was first programmed in assembly language but was reprogrammed in C in 1973 [133].

295. *Unstructured data* - refers to data that may have multiple sources, such as online digital files, text documents, SMS messages, video clips, photos, voices, sensors, pings, etc. The majority of the data created today is unstructured data, which is one of the keys to AI's growth [133].

296. *Unsupervised learning* - Unsupervised learning is a form of machine learning technique that concludes datasets with unannotated data. Cluster analysis is the most frequent type of unsupervised learning [133].

V

297. *Validation data* – data structured similarly to training data, with input and labels, and it's used to evaluate a recently trained model against new data and assess performance, with a particular emphasis on detecting overfitting [133].

298. *Variable* - a symbolic name given to an unknown quantity that permits the name to be used independent of the information it represents. Variables are associated with data storage locations, and values of a variable are normally changed during the course of program execution. Variables represent all kinds of data, including booleans, names, integers, arrays, pictures, sounds, scalars, strings, or any object or class of objects depending on the programming language that supports them. The symbolic names of variables are replaced with the actual data location by compilers and interpreters. Data in locations changes during execution while locations and names are fixed [137].

299. *Variance* – the degree to which a machine learning model's intended purpose varies as it is being educated. Despite their flexibility, models with considerable variance are vulnerable to overfitting and low predictive success since they are dependent on their training data [133].

300. *Variation* - queries, also known as sentences or utterances, are artificial language processing that works in tandem with intents. The distinction is what someone might say to achieve a specific aim or goal [133].

301. *Validation Data Set* – validation Data Set is a sample of data used to evaluate the accuracy of a model fitted on the training set while tuning model hyperparameters. As skill on the validation dataset is incorporated into the model configuration, the evaluation becomes more biased [133].

302. *Value learning problem* is a specific fundamental issue in the development of machine learning and artificial intelligence technologies that addresses the difference between humans and computers, and the ways that they think.

303. *Vision Processing Unit (VPU)* – a new type of microprocessor and an AI accelerator created to speed up machine vision activities [133].

304. *Visual Recognition* is the capacity of computer software to recognize items, locations, people, text, and activities in photographs and videos [133].

305. *Voice recognition* – a human-computer interaction technique that allows computers to understand, interpret human dictation, and produce written output in accordance with speech commands [137].

W

306. *Weak AI* - The term "weak AI" refers to a non-sentient computer system that operates under set parameters and focuses on a single activity or a small number of activities. Weak AI is the most prevalent sort of AI currently in use. It is one of the most searched Artificial intelligence terms [133].

307. *Web Crawler (Spider)* - An automated Web surfer goes from one site to another, usually seeking to index the Internet for a search engine [133].

308. *Web mining* - process of using data mining techniques and algorithms to extract information directly from the Web by extracting it from Web documents

and services, Web content, hyperlinks and server logs. The goal of Web mining is to look for patterns in Web data by collecting and analyzing information in order to gain insight into trends, the industry and users in general [137].

309. *Web Scraper*– a bot or web crawler is used to automate operations. It’s a type of replication in which data from the Internet is collected and copied into a central local database or spreadsheet for future retrieval or analysis [133].

310. *White box testing* is a type of software testing in which the person understands the application design and source code. Using these insights, the person can run tests that reveal the most likely bugs or performance issues [145].