

Human Intelligence vs Artificial Intelligence: Secondary Term Creation in Information and Communication Technology Field

Dace Šostaka^{*,†}, Juris Borzovs[†] and Jānis Zuters[†]

University of Latvia, Faculty of Computing, Raina bulvaris 19, Riga, LV-1586, Latvia

Abstract

Information and Communication Technology (ICT) terms are predominantly created in source language (English) and then secondary-created in other languages (Latvian, Lithuanian and others).

In first part we focus on insight into ICT secondary term creation (STC) process in Latvia in general and analyses of the process in particular, research done of the term creation in general and STC in ICT field in particular, and introduce to the possible steps of STC process to be optimized.

The second part is devoted to case study: ISO standard “Artificial intelligence concepts and terminology ISO ISO/IEC 22989:2022” in general and “dependable and dependability” in particular. We contrast and compare human approach to term creation vs the current possibilities of applying existing artificial intelligence (AI) tools to term creation.

In conclusion, we outline the most productive ways for automating parts of the STC process of ICT in Latvian and provide insight into possibilities for further development.

Keywords

Artificial intelligence, Information and Communication Technology, secondary term creation, Latvian, English

1. Introduction

By way of introduction, it can be said that if in a way the Three Laws of Robotics [7], applicable in the broadest sense of the word to what is called artificial intelligence (AI), were already described in the 1960s and 1970s by Isaac Asimov in his “I, robot” and Arthur C. Clarke in “2001: A Space Odyssey” [8] by personification of HAL 9000; if B.C. and A.D. stand for Before Christ and Anno Domini, respectively – a beginning of a new, different era then similarly, within the last few years, since the access to the open AI

Baltic DB&IS Conference Forum and Doctoral Consortium 2024

* Corresponding author.

† These authors contributed equally.

✉ dace.sostaka@lu.lv (D. Šostaka); juris.borzovs@lu.lv (J. Borzovs); janis.zuters@lu.lv (J. Zuters)

🆔 0000-0003-2066-937X (D. Šostaka); 0000-0001-7009-6384 (J. Borzovs), 0000-0002-3194-9142

(J. Zuters)



© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

technologies has been provided, in a way it can be considered that a new time counting system has been created within regard to gathering information, processing it and creating a new information: namely, before the release of various types and subtypes of AI and after the release.

General futurological speculations about the pros and cons of AI systems in human life in general and language in particular [14] in the broadest understanding of the concept were especially popular in the 1960s and 1970s, predominantly by science fiction writers [13], a detailed research article appeared already in 1980s, dealing with applications of AI in terminology work [12], describing the steps to be taken with a rather precise prediction.

When researching the development of Information and Communication Technology (further in the text – ICT) terms, the main aims and objectives of this article are as follows. First of all, provide a brief insight into the research in the term creation in the scientific community. Second, describe the process of secondary term creation in Latvian in general, in the ICT field. Third, the detailed process of term formation is described, looking for aspects that could be automated. Fourth, in the case study, the methods for parts to be automated were researched, compared and contrasted: efficiency of writing tailored programs vs using already existing AI tools. Fifth, we draw preliminary conclusions about the research and look at possible future research avenues.

1.1. State of Art: Automated Terminology Creation

While looking for relevant research regarding the automation of terminology, we looked into the latest publications of the leading academic journal “Terminology” and even if there could be found articles devoted to the automatic term recognition [32] and extraction [33]; also there is research, dedicated to the possibilities of automating the process of creating dictionaries [34]. There is also pertinent research on the possibilities of automated term creation [35] and of secondary term creation (also - localization) of medical terminology in the German language [36]. Nevertheless, there can't be found published research on the possibilities of automation of secondary term creation.

1.2. Research in AI and Terminology

Within the last few years, there has been an understandable explosion of specific research regarding the application of various AIs on a general, popular science level [11]. Specific research [24], published in 2024, is already devoted to linguistics and AI and to aiding AI in terminology in particular [25].

On the one hand, in the last decade, numerous researches have been carried out on ICT terms from various perspectives and in different languages: in 2014, Albanian on codification of Computer Terminology [3], Romanian on linguistic approaches when creating IT vocabulary [4], and Serbian on secondary term creation in 2016 [6].

Although there is research regarding the term formation per se in various languages available, the number of research narrows down with the specialisation of the field in secondary term creation for ICT terminology.

There is research in the 2020 on state of art, regarding Europe and languages [18]: in the article about the leaders of the European language technology in 2020, In the article it is emphasized that following leading trends can be observed: already developed language technologies for Latvian, Lithuanian, Estonian and other languages promotes visibility for languages on digital landscape and facilitate their sustainability and development.

1.3. Research in ICT Terminology and its Creation

There is available bilingual research about the translation of IT terms in Swedish [20], where the Swedish experience of forming English ICT terms into Swedish described - as the main method is outlined borrowing and loan translation. In Romanian, there has been carried out research focusing on hybrid terms in computer science [21], regularities and norms in secondary term formation [22] and phraseology in Romanian ICT texts [23]. Also, there is research from the Polish language perspective, devoted in general to linguistics and AI [24] and terminology work [26], terminology work and possibilities of generative AI for low-resource languages [25].

There has been carried out research on English computing terminology as a system, in 2010, as well as research on trilingual term creation in ICT field of correspondingly English, French and Igbo [27] and English, Russian and Kazakh [28]. There is also available research on the Computer science (ICT terms) in academic discourse, analysing 444 most often used words in Computer science publications (2023) [19].

Our previous research focused on Latvian ICT terminology creation principles and observation of the implementation of ICT terms in general [29] and, in particular, to a case study of Computer-Assisted Latvian ICT terminology development [30].

It is worth emphasizing that even though there are terms that have multiple meanings in the source language (English) we secondary recreate the concept. We work with ISO standards (ISO/IEC/IEEE 24765:2017(E), "Systems and software engineering – Vocabulary" and ISO/IEC 22989:2022(en) "Information technology – Artificial intelligence – Artificial intelligence concepts and terminology – ISO standard").

2. Case study: Human Intelligence vs Artificial Intelligence

In order to understand the actual process of the term creation and distinguish the places to be optimized, the following avenue of research was chosen in the following four steps.

Firstly, an analysis of the terms recently discussed and accepted in our commission was run through the AI tool (ChatGPT), comparing the answers provided by AI and the decision of the terminology Commission.

The percentage of correspondence was checked and creativity aspect for the terms that did not return identical answer with our commission's decision was checked.

Two ways were chosen. Prompting (phrasing of the question to the AI system) was used in order to devise a meaning for the cases when both the meaning of the term and

form of it was wrongly represented and B, a program for automated extraction of the terms from AkadTerm (Academic term database) was created for checking the results

Secondly, a detailed transcript [10] and analyses of the term discussion and accepting from actual term commission meeting was carried out, looking for sources consulted (that could be automated), for extended clarification moments of the meaning of the term and the concept it denotes.

Then AI tools were used in checking the possible solutions for the dependable and dependably and checking the obtained results for validity. Full prompting history is available online [9].

First, we explored the actual term formation process in Latvian ICT, analyzing the most typical scenarios. Determined the actions (steps) of the process that could be automated and distinguished two main parts of STC: technical aspect (searching for existing equivalent/s of the term) and creative aspect (looking for and/or creating new term) and focused upon optimizing the search for existing terms and/or term parts and case study: pilot project looking for ways of *optimizing* the search for existing terms and/or term parts.

The most often **scenario** of STC is searching in AkadTerm (Academic Terms database) [16] as a first resource, looking if term or part of term as been accepted by the Information and Communication Technology Terminology Subcommittee, the Latvian Academy of Sciences (further in the text – the Commission). The most often used **method** for STC is combining already accepted terms (if they reflect the meaning of the concept) in a new term.

2.1. Computer Program for Extracting Terms from the Academic Terms Database

When manually searching for the terms, the process looked as follows, when preparing for the Commission meeting. Term preparation consists of repetitive steps, four of them are mainly mechanical:

first, copy the source term, paste it into search field (of AkadTerm),

second, check returned results a) if the term has already been accepted by the Commission and b) check if the new term is part of already accepted terms.

third, if the term cannot be found as accepted, then the IATE, Microsoft Terminology and other sources are searched.

fourth, copy the results in the end document and continue the process in other sources.

Pilot project was created, comparing the time needed for manual work and with the aid of a computer program. For the manual implementation of the four abovementioned steps, when preparing 40 terms, the time invested was 4 hours.

The computer program for testing the idea was created, largely imitating the four steps taken manually, when looking for ways of reducing time for the term extraction and combining the returned results:

First, “search and return Commission’s accepted terms”:

Second, “search and return ALL terms, available via AkadTerm: Commission, Microsoft Terminology and others”

Third, “search and return: compound terms; partial matches.”

Forth, “search and return: compound terms; full matches.”

Summary. For preparing 40 terms with aid of program, the time invested was 1 hour. The preliminary results were presented in the 2023, in 5th international conference on terminology “Scientific, administrative and educational dimensions of terminology” [15]. The first results were promising and we decided to try out AI approach in facilitating the secondary term creation. The process will be reflected in the next subchapters.

2.2. Dependable and Dependability

In order to understand in detail the exact process of not-so-trivial case of STC, we analysed the discussion of two terms that have posed a challenge for some time because of difficulties in reaching consensus. Thus, a more detailed analyses of the discussion process is reflected.

When discussing the **dependable** and **dependability** (ISO/IEC/IEEE 24765:2017(E), Systems and software engineering – Vocabulary), first of all the definition and context is given for understanding of the terms.

2.2.1. Determining the Term Context

In the case of the terms, the main context was provided by commission member [9] and it was explained that *“historically, «reliability» was created as a concept long ago. There has been already discussions regarding offering a completely different word, with another root and term already in use, «drošums» (safety, security). Nevertheless in Latvian «drošums» gravitate in its meaning very much towards either English «safety» or something related to «security», so it is not a good choice for «dependability» and «dependable»”* and the commission shall look for different word.

Further, it was explained that *“if the meaning of the English word «reliable» is checked, then «uzticams» (trustworthy) is returned in Latvian. So, it shall be understood what this term means on the conceptual level in this field, namely, in the field of systems. Looking further into the meaning of the term, the answer is as follows: «It is the ability of a system to perform/implement defined functionality within a defined timeframe», and it is usually expressed as a percentage of the total time the system exists. It is a long-established method, and it is nothing [conceptually] new.”*

A chronological approach to the understanding of the terms is taken, and it is explained that *“In its turn, «dependability» came relatively later from various sources, primarily standards; I don't even know all of them. There is even more confusion and chaos with «dependability» because, in some sources, it is interpreted as an utterly synonymous term with «reliability». But mostly, there are some other attributes.”*

The etymology of the terms is briefly discussed and the documents they originated are emphasized, one coming from a political background and another from technical documentation *“[...] I assume «dependability» might be more like a political term. It could be originating from different political documents. «Reliability», on the other hand, has been strictly a technical term from the very beginning.”*

2.2.2. Checking the Meaning of Latvian Terms

Then, the problem with searching for an appropriate term in the target language is defined as following: both terms have the same translation into Latvian, namely “drošums”: *“Thus,*

the problem is very clear-cut: if we take a general language dictionary and check both of these terms, namely «dependability» and «reliability», then we find as a translation the same Latvian word «drošums»; it is unacceptable because it creates confusion in understanding and that is the problem.”

The actual reasoning for the need of clear term is clearly defined in the very beginning of the discussion as “[..] there should be a new term for «dependability» because it is a more recent term that appeared later. Something very similar, but not the same. And there we have the linguistic problem: we haven't found the word for the Latvian equivalent of «dependability».”

2.2.3. Consulting External Sources

The next step is a discussion among the terminology commission members about the individual understanding of the terms «dependability» and «reliability». It shall be noted that this is the most time-consuming part of the discussion where various approaches while looking for a clear understanding of the term, are used: distinction of the terms, aiming to determine dominating aspects “*distinction is to be made between «dependability» and «reliability», where the former includes the latter as one of its aspects, then I would rather say that «dependability» as a broader term should be translated as «uzticamība» (reliability) and «reliability» as «noturība» (stability, resilience).”*

Then the offered term “noturība” is back translated from Latvian into English, emphasizing that “*«noturība» meant something else and it wouldn't reflect the concept in this case.”*

The possible translations of the “noturība” in English are searched for and another commission member provides the answer that “*Latvian «noturība» in English can be translated as «persistence», «resilience»; «stability», and «immunity».*” [9]

The same process – offering words that could reflex the meaning of the term in Latvian and the backtranslation of them into possible English equivalents - is provided for other Latvian words, offered for discussion: *noturība, uzticība, paļāvība, drošums, pārliecinošs, vērtums, vērtība.*

2.2.4. Creating Compound Terms

Then, when it is not possible to find one word that would reflect the concept, the next step is taken, creating the compound terms: *vispāruzticams, vispāruzticamība, paļāvībvērts.*

It shall be mentioned that the offered words and compound words are checked in the language system: how they can be inflected and how they can be used in the sentences. It should be possible for a term to be used in sentences; thus the adjective form is checked. In the course of the discussion various sources are checked, in order to clarify the meaning and justify the offer: dictionaries, term bases, also AI system, looking for terms in different contexts. The decision commission members agreed upon, are «dependable» as «paļautiesvērts», «dependability» as «paļautiesvērtums».

To sum it up, it can be seen that there are a lot of searching, comparing and contrasting, confirming the provided ideas, looking for context. Thus the idea for AI fine-tuned to compare the results from AI and our accepted terms.

2.3. ISO Standard ISO/IEC 22989:202

For research purposes we choose an ISO/IEC 22989:2022(en) “*Information technology – Artificial intelligence – Artificial intelligence concepts and terminology – ISO standard*” that we have been discussing in our commission from the October, 2023. The total number of the terms in the ISO standard are 140. The further results in this subchapter were first represented in Baltic Digital Humanities Forum April 25-26, 2024 [31].

The experiment was carried out as follows. First we ran through all the terms that are in the ISO standard with definitions; looking from answers in AI system ChatGPT4.

For each English term (e.g. “genetic algorithm”), with its definition (“algorithm which simulates natural selection by creating and evolving a population of individuals (solutions) for optimization problems”):

- 1) ChatGPT was prompted to carry out secondary term formation (e.g. “Provide a Latvian translation for an ICT term “genetic algorithm”, with following definition (see above) and the result was recorded (“ģenētiskais algoritms”);
- 2) “manual” secondary term formation by the Commission was carried out (e.g. “ģenētiskais algoritms”);
- 3) results of the two processes were compared and categorized as exact match (Table 1) with 75 terms, partial match (Table 2) with 65 terms and no match (Table 3) with 5 terms.

Table 1

Both concept and grammatical forms are correct.

No.	Term (ISO, English)	Term (ChatGPT, Latvian)	Term (Commission, Latvian)
49.	data quality checking [...]	datu kvalitātes pārbaude [...]	datu kvalitātes pārbaude [...]
66.	decision tree [...]	lēmumu koks [...]	lēmumu koks [...]
24.	genetic algorithm [...]	ģenētiskais algoritms [...]	ģenētiskais algoritms [...]

Table 2

The concept is formed almost correctly. There are stylistic differences or/and minor grammatical mistakes.

No.	Term (ISO, English)	Term (ChatGPT, Latvian)	Term (Commission, Latvian)
65.	Bayesian network [...]	Bejas tīkls [...]	Beijesa tīkls [...]
48.	data annotation [...]	datu anotācija [...]	datu anotēšana [...]
39.	procedural knowledge [...]	procedurālas zināšanas [...]	procedurālās zināšanas [...]

Table 3

The concept is formed incorrectly, even if a part of the term is translated correctly.

No.	Term (ISO, English)	Term (ChatGPT, Latvian)	Term (Commission, Latvian)
8.	application specific integrated circuit [...]	konkrētai lietošanai pielāgots integrētais shēmviens [...]	lietojumam pielāgota integrētā shēma [...]
13.	cognitive computing [...]	kognitīvā datortehnika [...]	kognitīvā datošana [...]
54.	ground truth [...]	pamattiesība [...]	mērķa vērtība [...]

It can be concluded that:

1. *Exact* match usually refers to the verbatim (word-for-word) translated terms or terms accepted some time ago. *Partial* match requires post-editing, still it speeds up the term discussion process. *No match* requires manual research in the term databases and other sources.
2. No significant element of innovation has been detected so far.
3. Nevertheless, AI can find and mechanically combine terms found in its sources, thus facilitating the groundwork for discussing terms and freeing up more human resources for the creative part of secondary term formation.

AI tools can be used as a tool for the first checking of the new material and then, the remaining part of the terminology units are checked with the computer program from reliable sources instead of attempting to obtain all results from the AI. Thus, hybrid approach, combining the best from both worlds, works the best in order to speed up the term creation process. Generalized conclusions can be at the ending of the article.

2.4. AI Approach: Dependability and Dependable

As a concluding phase of the current research process, in order to find out what manual tasks can be delegated to the AI when preparing terms for discussion and during the discussion, we asked the MI system prompts that were, as far as possible, similar to the questions and clarifications that were made during the Commission meeting.

Of course, it was not possible to completely duplicate the discussion during the terminology commission meeting and the prompts to the AI system but we summarised the main points:

- 1) clarifying the conceptual understanding of the terms in source language (English),
- 2) searching for possible duplicating terms in source language, corresponding to the definition provided in the source language,
- 3) translating in the obtained terms in target language (Latvian),
- 4) looking in obtained results in the target language (Latvian) the possible useful propositions for the terms that could be used to reflect the meaning.

2.4.1. Definition of the Dependable and Dependability

First, taking into account the time in the Commission was devoted to the reach understanding of the concept of terms **dependable** and **dependability**, the prompt “Define the “dependability” in ICT” was created, looking for definition of the both terms from AI (“The Chat GPT4.0” and “BingCopilotNotepad” were chosen) perspective. Taking into account the limitations of the paper, the full prompting history is available at the link [10].

The ChatGPT4.0 returned rather inclusive definition “In the context of Information and Communication Technology (ICT), “dependability” is a broad term that encompasses several key aspects related to the reliability, availability, safety, and security of computer systems and networks.

Dependability is a measure of a system's ability to deliver a service that can justifiably be trusted. This concept is crucial in systems where failure or security breaches can lead to significant consequences, such as in financial transactions, healthcare records, critical infrastructure, and communication systems.”

In addition, there was provided additional information to the straightforward prompt, regarding six attributes of “dependability” and explaining each attribute:

“Dependability includes the following attributes:

1. Reliability: The ability of a system to perform its required functions under stated conditions for a specified period of time without failure. It is often quantified using metrics such as mean time between failures (MTBF).”

In the similar way next five attributes – availability, safety, security, maintainability and integrity – were enumerated and explained.

The Bing (CopilotNotepad) returned extensive definition of “dependability” as well: “In Information and Communication Technology (ICT), “Dependability” is a measure of a system's availability, reliability, maintainability, and in some cases, other characteristics such as durability, safety and security.

It refers to the ability of a system to provide services that can be trusted within a time-period. The service guarantees must hold even when the system is subject to attacks or natural failures. The International Electrotechnical Commission (IEC) provides systematic methods and tools for dependability assessment and management of equipment, services, and systems throughout their life cycles”.

There were six attributes listed as well, but they were not explained, yet there was something worth making prompt in another AI system: actually working list sources included, leading to corresponding articles about “dependability” in following sources: Wikipedia, Springer publishing, Yale University and Study portal.

2.4.2. Comparison of Definitions

When compared with the definition of “dependability” that is found originally in ISO standard “Dependability 3.1109

1. trustworthiness of a computer system such that reliance can be justifiably placed on the service it delivers [IEEE 982.1-2005 IEEE Standard Dictionary of Measures of the Software Aspects of Dependability, 2.2.]

2. availability performance and its influencing factors: reliability performance, maintainability performance and maintenance support performance [ISO /IEC 150256-1:2013 Systems and software engineering - Systems and software assurance - Part 1: Concepts and vocabulary, 3.134]”.

When all three definitions are compared (see Table 4), it can be seen that definitions, provided by ChatGPT4 and Bing clarify the meaning, provided in the ISO standard. In addition, explanation and sources are given, that helps in understanding of the concept of the term. The same results were returned for the definition of the term “dependable”.

Table 4
Definitions of “dependability”

ISO	ChatGPT	BingCopilotNotepad
<p>Dependability 3.1109</p> <p>1. trustworthiness of a computer system such that reliance can be justifiable placed on the service it delivers [IEEE 982.1-2005 IEEE Standard Dictionary of Measures of the Software Aspects of Dependability, 2.2.]</p> <p>2. availability performance and its influencing factors: reliability performance, maintainability performance and maintenance support performance [ISO /IEC 150256-1:2013 Systems and software engineering - Systems and software assurance - Part 1: Concepts and vocabulary, 3.134]</p>	<p>Dependability is a measure of a system's availability, reliability, maintainability, and in some cases, other characteristics such as durability, safety and security. It refers to the ability of a system to provide services that can be trusted within a time-period. The service guarantees must hold even when the system is subject to attacks or natural failures. The International Electrotechnical Commission (IEC) provides systematic methods and tools for dependability assessment and management of equipment, services, and systems throughout their life cycles</p>	<p>In Information and Communication Technology (ICT), "Dependability" is a measure of a system's availability, reliability, maintainability, and in some cases, other characteristics such as durability, safety and security. It refers to the ability of a system to provide services that can be trusted within a time-period. The service guarantees must hold even when the system is subject to attacks or natural failures. The International Electrotechnical Commission (IEC) provides systematic methods and tools for dependability assessment and management of equipment, services, and systems throughout their life cycles.</p>

Thus, for the function of better comprehension of the concept of the term, clarifying the definition both AI systems serves well, at least in the examples examined in this article.

Two next prompts were created with an intention to receive possible terms in source language that corresponded to the definition: “Using your own definition, please provide ten most often used terms in ICT, reflecting the meaning, described in the definition” and “create terms, using given ISO definition”. For further experiment the terms returned to the ISO definition were chosen (see Table 5).

Table 5

Terms returned for ISO definition of “dependability”

ChatGPT4 offered term	ChatGPT4 translation	Bing offered term	Bing translation
Service Level Agreement (SLA)	Uzticamība	Reliability	Uzticamība
Reliability:	Pieejamība	Availability	Pieejamība
Availability:	Uzturējamība	Maintainability	Uzturēšana
Maintainability:	Uzturējamība	Serviceability	Servisējamība
Mean Time to Repair (MTTR)	Vidējais remonta laiks (VRL)	Mean Time Between Failures (MTBF)	Vidējais laiks starp kļūmēm (MTBF)
Change Management	Izmaiņu pārvaldība	Fault-Tolerant Systems	Kļūdu tolerējošas sistēmas:
Incident Management	Incidentu pārvaldība	Data Integrity	Datu integritāte
Disaster Recovery Planning	Katastrofu atveseļošanas plānošana	Downtime	Neaktīvs laiks
Failover Mechanism	Pārslēgšanās mehānisms	Uptime	Darbības laiks
Load Balancing	Slodzes izlīdzināšana	Failover	Pārslēgšanās

It can be seen that even though there are words that reflect the meaning of the concept into Latvian, namely “uzticamība”, “darbības laiks”, “pieejamība” and these terms certainly can serve as an inspiration when looking for a possible terms to be created into Latvian, there is no ready, sure and reliable answer provided.

The preliminary conclusions for this part of research are as follows: on one hand, AI systems seem right now useful, when tested in the aspect of providing extensive definitions, comparing and contrasting different aspects of the terms, and usage examples in collocations and sentences; on another hand, there is no significant creative aspect in translation of terms and, what is even more important, there is no guarantee of repeatability and reliability of the results received when prompting after some time or different user.

AI works well as a sort of sorting tool for terminology units; yet it is relatively useless in creating reliable results. Still it can be used as an aide for comparing and contrasting looking for examples in sentences; looking for synonyms.

Thus, it can be said that the computer program written for sorting the terminology for meetings is the best short-term solution for speeding up the term creation process. Sources are indicated, it is reliable, and it reduces the manual labour and provides ground for fruitful discussion.

3. Conclusion

Based on the research work carried out in this article, the following can be preliminarily concluded.

In the short term, the most efficient way to increase productivity and reduce the time needed for term pre-processing (preparation for the discussion) is the computer program created specifically for collecting and exporting already accepted terms in Latvian from AkadTerm (Academic Term database).

In the long term, there is a possibility that AI systems can be used successfully for professionals in the STC field for the following specific tasks: a) devising a definition for a term; b) searching for terms corresponding to the definitions; c) comparing and contrasting terms and their definitions.

The main drawbacks of AI systems analysed, in order of significance for secondary term creation, are the unreliability of results, the lack of credibility of sources indicated, the absence of confidentiality, and the limited amount of term processing.

Further research avenues are a) developing further programs for traceable and confidential term pre-processing b) specific application of already existing AI systems for secondary term creation and c) looking into possibilities to train natural language processing model(s) for term generation purposes.

Acknowledgments

The work described in this paper was supported by the project “Innovative information technologies” at the University of Latvia. Any errors or omissions are the faults of the authors.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

- [1] Rusko, T.: English computing terminology as a system, pp. 107–112, (2010), URL: https://www.researchgate.net/publication/266886183_English_Computing_Terminology_as_a_System
- [2] Breivik, T.: Action Plan For Norwegian Language And ICT And The Norwegian Human Language Technology (HLT) Resource Collection, URL: <https://www.sprakradet.no/localfiles/12399/ifip2005.doc>
- [3] Cerpja, A., Cepani, A.: Computer Terminology in Albanian Versus other Languages and the Benefits of Its Codification. *Journal of Education and Practice*. ISSN 2222-288X. Vol.5, No.11, 2014, 42-47 (2014). URL: https://www.researchgate.net/publication/342751546_Computer_Terminology_in_Albanian_Versus_other_Languages_and_the_Benefits_of_Its_Codification
- [4] Diaconu, D.C.: Linguistic Mechanisms Of IT (Computers) Vocabulary Formation In Romanian Language (2014), URL: <https://www.academia.edu/>

- 87984436/Linguistic_Mechanisms_of_It_Computers_Vocabulary_Formation_in_Romanian_Language
- [5] Darmois, E., Böcker, M.: A Harmonized Multi-lingual Terminology for ICT Devices and Services with a User-Centric View (2023), URL: https://www.researchgate.net/publication/372227342_A_Harmonized_Multi-lingual_Terminology_for_ICT_Devices_and_Services_with_a_User-Centric_View
- [6] Kavgić, A., Kavgić O.P.: Computer-Related Terminology In Serbian: A Diachronic Analysis Of Borrowing Trends, Translation And Adaptation Strategies (2011), URL: https://www.academia.edu/2071665/computer_related_terminology_in_serbian_a_diachronic_analysis_of_borrowing_trends_translation_and_adaptation, last accessed 2024/02/27
- [7] Asimov, I.: I, robot. Bantam books, New York (2004)
- [8] Clark, A.C.: 2001: A Space Odyssey. ACE, New York (2018)
- [9] Sostaka_Borzovs_Zuters_AI_dependable-dependability_DBIS_2024_.pdf, URL: <https://1drv.ms/b/s!AqrEVf8PaU0Gg4EgSqoCP6UXp1kL7Q?e=JzqZaH>, last accessed 2024/03/20
- [10] Sostaka_Borzovs_Zuters_ITTEA_dependable-dependability_DBIS_2024.pdf, URL: <https://1drv.ms/b/s!AqrEVf8PaU0Gg4EhDdMJ74mW1W-XPQ?e=QG4Py5>, last accessed 2024/03/20
- [11] Maier, E.: In the Shadows of the ChatGPT Spotlight, URL: <http://lt-innovate.org/content/shadows-chatgpt-spotlight>
- [12] Nedobity, W.: Terminology and Artificial Intelligence, URL: https://www.researchgate.net/publication/327456080_Terminology_and_artificial_intelligence
- [13] Rossis, N.C.: The 1960s: A Golden Age for Science Fiction, URL: <https://nicholasrossis.me/2023/08/17/the-1960s-a-golden-age-for-science-fiction/>
- [14] Science Fiction, URL: <https://www.britannica.com/art/science-fiction>
- [15] Šostaka D., Borzovs J., Zuters J.: How Do We Think? The Case Study of Latvian Information and Communication Technology Secondary Terms Development, URL: https://lki.lt/wp-content/uploads/2023/09/term_Vilnius_2023_programme.pdf
- [16] AkadTerm, Academic terms database, <http://www.akadterm.lv/term.php>
- [17] Šostaka D., Borzovs J., Zuters J.: Would It Be Possible to Optimise the Secondary ICT Term Creation in Latvian? 14th Conference on DATA ANALYSIS METHODS for Software Systems, Vilnius University Proceedings, 89, pp. 1–104.
- [18] Rehm, G. et al “The European Language Technology Landscape in 2020: Language-Centric and Human-Centric AI for Cross-Cultural Communication in Multilingual Europe” URL: <https://www.semanticscholar.org/paper/The-European-Language-Technology-Landscape-in-2020%3A-Rehm-Marheinecke/02e3c00a7375e4cdd7e6545b48782aecc0de3420>
- [19] In 2023, Uba, S.,Y. et all, “A New Computer Science Academic Word List”, 319-330, URL: https://www.researchgate.net/publication/369877950_A_New_Computer_Science_Academic_Word_List Sciedu Press, World Journal of English Language

- [20] Izwaini, Sattari in 2005, (ESP Across Cultures 2 (2005) 76-93) in “A corpus-based study of the translation of IT terms”, URL: https://www.academia.edu/34913337/Corpus_based_Study_of_IT_Terms.
- [21] Postolea, Sorina- 2011 - Hybrid Terms in the Romanian Language of Computer Science URL: https://www.academia.edu/5200765/Hybrid_Terms_in_the_Romanian_Language_of_Computer_Science
- [22] Postolea, Sorina- 2012, from regularities to norms in the secondary formation of romanian ict terms, *linguaculture*, 2, 2012, 101-115., secondary formation of Romanian ICT terms URL: https://www.researchgate.net/publication/276172400_From_Regularities_to_Norms_in_the_Secondary_Formation_of_Romanian_ICT_Terms
- [23] Postolea, Sorina- 2014, The Phraseology of General Use ICT Texts and its Translation from English into Romanian Volumul X, Nr. 2, iunie 2014 URL: https://www.researchgate.net/publication/281589550_The_Phraseology_of_General_Use_ICT_Texts_and_its_Translation_from_English_into_Romanian_-_A_Corpus-based_Analysis
- [24] Artificial Intelligence for the Humanities. Fundamental Concepts Explained (2024) URL: https://www.researchgate.net/publication/378335852_Artificial_Intelligence_for_the_Humanities_Fundamental_Concepts_Explained
- [25] Corpus linguistics and generative AI tools in term extraction: a case of Kashubian - a low-resource language (2023) URL: https://www.researchgate.net/publication/378234423_Corpus_linguistics_and_generative_AI_tools_in_term_extraction_a_case_of_Kashubian_-_a_low-resource_language
- [26] Terminology work in the era of AI-generated content. A case of ChatGPT (an end-user perspective) URL: https://www.researchgate.net/publication/378297882_Terminology_work_in_the_era_of_AI-generated_content_A_case_of_ChatGPT_an_end-user_perspective
- [27] A Trilingual Creation of Terms of the Computer Language in English, French and Igbo, *LALIGENS*, VOL. 7(2), S/N 16, OCTOBER, 2018, 37-48. URL: https://www.researchgate.net/publication/329098902_A_trilingual_creation_of_terms_of_the_computer_language_in_English_French_and_Igbo
- [28] Meiramova Saltanat Akimovnaa*, Mussagozhina Kuanysh Kanatovna, The investigation of English, Russian and Kazakh computer terms borrowings to be acquired at English class, 94-102, URL: <https://www.sciencedirect.com/science/article/pii/S187704281504495X>
- [29] Šostaka D., Borzovs, J, Cauna. E., Keiša I., Pinnis M., Vasiljevs A.: The Semi-Algorithmic Approach to Formation of Latvian Information and Communication Technology Terms *Baltic J. Modern Computing*, Vol. 11 (2023), No. 1, 67-89 URL: <https://doi.org/10.22364/bjmc.2023.11.1.05>
- [30] Šostaka D., Borzovs, J.: Towards Computer-Assisted Latvian ICT Terminology Development: from Theoretical Guidelines to Case Studies, *Baltic J. Modern Computing*, Vol. 11 (2023), No. 1, 134-154, URL: <https://doi.org/10.22364/bjmc.2023.11.1.08>

- [31] Šostaka D., Borzovs J., Zuters J.: Preliminary Results of the Use of Artificial Intelligence in Secondary Term Formation. Digital Humanities, 2024-04-25, URL: <https://drive.google.com/file/d/12a3lZpvv7SxzBo2B03zu1fhdePinCy9n/view>
- [32] Kageura K., and Umino B.: Methods of automatic term recognition: Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication, Volume 3, Issue 2, Jan 1996, p. 259 – 289 DOI: <https://doi.org/10.1075/term.3.2.03kag>
- [33] Fulford H.: Exploring terms and their linguistic environment in text: A domain-independent approach to automated term extraction. Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication, Volume 7, Issue 2, Jan 2001, p. 259 – 279, DOI: <https://doi.org/10.1075/term.7.2.08ful>
- [34] Serrec A. L., L'Homme Marie-Claude, Drouin P. and Kraif O.: Automating the compilation of specialized dictionaries: Use and analysis of term extraction and lexical alignment. Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication, Volume 16, Issue 1, Jan 2010, p. 77 – 106 DOI: <https://doi.org/10.1075/term.16.1.04les>
- [35] Gurbanova A. M.: Analysis of automation methods of the term creation activity. 2018. Problems of Information Technology. URL: <https://jpit.az/en/journals/219>
- [36] Schulz S., Hammer L., Hashemian-Nik D., Kreuzthaler M.: Localising the Clinical Terminology SNOMED CT by Semi-automated Creation of a German Interface Vocabulary. Proceedings of MultilingualBIO 2020, pages 15–20. URL: <https://aclanthology.org/2020.multilingualbio-1.3.pdf>