CONSISTENCY IN CROATIAN IT TERMINOLOGY: CURRENT STATE AND PROFESSIONALS' OPINIONS

Anamarija Miličević

Abstract

The importance of standardized terminology for a field of knowledge is well known. In fast-developing fields such as IT, the standardization of terms created in the process of secondary term formation is often one step behind new developments. This paper focuses on the Croatian terminological system in the field of IT, describing its current state and soliciting professionals' opinions about it. A total of 45 English terms and their corresponding Croatian terms are analysed in respect of Sager's (1990) classification of procedures used in secondary term formation. A survey was conducted among IT students and professionals as primary users of this terminology in order to find out about their preferred procedures for term formation and their opinions on some of the currently available set of terms in Croatia. Both the term analysis and the survey have produced mixed results, showing a lack of consistency within the terminological system, as well as among the respondents' preferences. This suggests that the IT terminological system in Croatia, which includes the currently available terms, as well as the regulations and principles relating to them, is not yet defined or completely unified, and that professionals in this field neither use the terms consistently nor perceive the system as consistent.

1. Introduction

With the new industrial revolution and advancing technological trends, the field of IT[1] has been rapidly growing in the 21st century. This growth brings about new things, methods and concepts, resulting in the creation of new terms. Just like

[1] For the purposes of this paper, the term "IT" is used in a broad sense to encompass similar, as well as related scientific and technological disciplines and fields, such as computer science, software and hardware engineering, computing etc.

the field of IT, the number of new terms that need to be created in languages during the process of language transfer to other language communities is growing quickly. This makes fulfilling the task of developing terminology difficult.

This is especially true of situations in which standard language and standardized terminology needs to be used, such as in the case of academic and scientific texts. As opposed to informal contexts, where jargon can be used (*komp* as opposed to *računalo* in Croatian everyday nonstandard-language communication), standard language requires the use of standardized terminology. However, Sager (1990: 84) states that "[i]mportation of terminology occurs at such a pace that planned assimilation cannot cope." In languages where terminology is created mostly by secondary term formation, such as Croatian, not developing terminology quickly and consistently enough can bring about a number of problems. In practice, this means that experts as the main standardized terminology users have to revert to existing foreign terms, or that there exist several competing terms with no standardization. The goal of this paper is, therefore, to take a closer look at how English IT-related terms have been translated into or developed in Croatian and to find out what views IT experts and professionals have on this issue.

This paper[2] is organized as follows: Section 2 deals with IT terminology in Croatia, the choice of the topic and the focus of the study. Next, an overview of previous research is provided. Key concepts needed for the analysis are defined, along with the research questions, hypotheses and aims. The methodology used to analyse terms and to conduct the survey is explained in Section 6, with the results of the term analysis and survey reported and discussed in the subsequent section.

2. IT terminology in Croatia

The beginning of the 21st century is the start of the so-called Fourth Industrial Revolution (Encyclopaedia Britannica 2018) or Industry 4.0. The recent rapid development of computer and IT technologies, such as the internet, artificial

[2] This paper was originally written as the author's M.A. thesis at the Department of English of the Faculty of Humanities and Social Sciences, University of Zagreb, Croatia.

intelligence, increasing computer processing speeds etc., has made expansive technological advancements possible. Consequently, three of the five of the fastest growing sectors of global economy are directly connected to IT and related technological disciplines, while the remaining two are heavily influenced by them (World Finance 2018). Additionally, this field has a strong influence on our everyday lives. The same is true of Croatia – the IT-related industrial sector is one of the fastest growing and more important sectors in the country[3].

With the growth of this sector comes an increase in the number of concepts, inventions and, consequently, names for them – terms. Most of these terms are created in English, as that is the most productive language in this field. These terms are often "imported" and used in their original form by speakers of other languages:

> Societies which depend on importing technological and scientific knowledge need to designate the new concepts and therefore tend to use a large number of terms from other languages which, once a part of usage, are very difficult to displace. Scientific and technological transfer is the most frequent cause behind the high number of borrowed terms from a language in which the product or idea was created. (Cabré, 1999: 89)

This is particularly relevant for languages with a relatively small number of speakers, such as Croatian (Halonja and Mihaljević 2012: 11). As it is impossible to develop all new terms instantly, some users rely on their original (usually English) form, particularly with more recent and less frequently used terms. Another issue is the fact that there has been a tendency to preserve the purity of the Croatian language by giving priority to Croatian words and terms (2012: 13). However, in the field of IT, there seems to be a discrepancy between what is standard ("pure" Croatian words) and what is used by experts (both Croatian and foreign terms). Not even national institutions provide a developed terminological system for the IT sector; e.g. the Institute of Croatian Language and Linguistics (IHJJ) does not currently have any data sets available for this field on the terminological portal Struna and the field of IT is not listed as one of its projects. Entering the English terms used for this research into the Struna search engine

[3] Retrieved June 02, 2019, from https://tockanai.hr/tehnologija/hrvatski-ict-sektor-13934/

does not return any results marked as IT-related. In other words, there is no one authority providing official Croatian equivalents of relevant English terms. This leaves IT experts to fend for themselves, either by choosing to use the original English terms, or trying to find their own ways to translate or form the terms they need, particularly in formal contexts. This is why, in this paper, formal and standardized texts are used – specifically informative texts written for scientific and academic purposes (Mihaljević 1993: 7-8). Ideally, every English term should have its standard and widely-accepted Croatian equivalent, at least in formal texts. This uniformity and level of standardization make this type of text a good basis for linguistic analysis.

The terms used in this paper were compared to translation options provided in a number of resources an IT expert might use when looking for the Croatian equivalent of an English term. They are further explained in Section 6 of this paper, while a full list of sources and the term candidates found in them are provided in Appendix 1 and Appendix 2, respectively.

Overall, IT terminology in Croatia is still an undefined field. Even though there are a number of normative sources to consult, there is no one primary and widely-accepted resource. This leads to a somewhat unstable situation, resulting in synonymy, varying views and opinions among experts, multiple signifiers for one signified concept, and a lack of uniformity and precision (Mihaljević 1998: 10). As the terminology is not clearly defined and limited to one term per concept, the "semantic clarity" (Cabré 1999: 111) sought after by terminologists cannot be reached. These issues will be further explored in the term analysis and survey, as described in Sections 7 and 8.

3. Previous research

At the moment of writing this paper, the topic of IT-related terminology in Croatia has been researched by a relatively small number of authors and has not gained much formal attention. This is perhaps due to the fact that the field is still relatively new and has only recently become the subject of research. Because of this lack in variety of research perspectives, the existing research is somewhat one-sided and prescriptive in part.

Research was mostly conducted by, but is not limited to, members of the Institute of Croatian Language and Linguistics (IHJJ), primarily by Mihaljević and her colleagues. In Mihaljević (1993, 1998 and 2003), the author sets the groundwork for future research done in this field by researching how Croatian (IT) terms are created, what the most common translation methods are, and which issues came up during the translation process. The author is against using foreign words, especially in the field of IT, where English is becoming the dominant language worldwide. These works offer the perspective of a linguist specializing in Croatian and at times prescribe solutions that could today be considered obsolete, such as recommending the Croatian term *strojevina* for the term hardware. The problem is also that many concepts used today were only created years after these books were published, showing a need for more up-to-date research.

This need was in part fulfilled in the author's more recent work (Mihaljević, 2006, 2007 and 2009). The focus of this research lies in highlighting the notion that Croatian IT terminology (still) needs to be standardized, on explaining how this could and should be done, and on dispelling some usual misconceptions when it comes to this topic, such as that English terms are more precise than Croatian ones. These papers may be seen as an expansion of Mihaljević's earlier articles in that they include certain new terms, but the results and conclusions remain largely the same. In Mihaljević (2009), the author also describes the tools that were being developed at the time to help solve the issues of Croatian IT terminology, such as the previously mentioned Struna database. However, as of the completion of the present paper, this specific resource could not yet be utilized for this purpose. In Mihaljević (2007), there is also mention of a Croatian IT terminology portal (hrana.ffzg.hr), but it was not accessible at the time of writing this paper.

Halonja and Mihaljević (2012) offer a further elaboration of Mihaljević's previous work by distinguishing between IT jargon and standardized terminology. That is an important distinction for the present paper, as this research analyses only formal and standard language. Halonja and Mihaljević (2012) also include a dictionary of the Croatian computer jargon, with standard English and Croatian

variants provided where possible. The list of terms analysed by Halonja and Mihaljević is updated to an extent, compared to the previously described works, albeit still insisting on a number of not widely-accepted Croatian term candidates.

Other authors who have dealt with this field include Škifić and Mustapić (2012). They describe the state of IT terminology, as well as related anglicisms, in respect to the currently dominant language ideology (language purism) in Croatia. Furthermore, the question of whether or not that stance is always viable or necessary is discussed. The authors also conducted a survey with elementary school students and found that Croatian generally is not threatened by English, but that English terms are chosen where it is easier to apply them.

Of note is also the work by Miščančuk and Vučković (2011). The authors conducted a study that is similar to the present one. They analysed a number of random English terms and their proposed Croatian equivalents from three normative sources, finding that these sources often did not agree or provide a single agreed-upon translation for a given term. The authors also found that there could be many possible Croatian equivalents of a single English term.

Overall, while there has been some research on this topic, a majority of the available resources are quite similar with regard to the conclusions reached. Most of the available papers and books were written by the same author or group of authors, which could mean that there is little diversity in research perspective and methods. Despite existing research, there are many aspects of Croatian IT terminology left to explore.

4. Key concepts

This section outlines the most important concepts for the present research, primarily the principles, standards and norms of developing and translating terminology from English into Croatian.

4.1 Croatian terminological principles

When it comes to term formation or the development of a terminological system in Croatian, there are certain widely-accepted principles which, if adhered to, should ensure that the new terms and terminologies fit into the currently existing terminological systems and standard Croatian. These guidelines are given in the updated terminological handbook (Hudeček, Mihaljević, and Nahod, 2009: 69-78):

1. Croatian words should be prioritized over foreign words.
2. Terms of Greek or Latin origin should be prioritized over terms of English, German, French etc. origin.
3. The more widely used and accepted term should be prioritized over the less used term.
4. Terms need to comply with the Croatian standard language system.
5. Shorter terms should be prioritized over longer ones.
6. Terms which are easier to derive new words and terms from should be prioritized over those with few possible derivations.
7. Terminological polysemy within the same terminological system should be avoided.
8. Existing terms should not be modified without a valid reason.
9. One term should be prioritized over another if it fits the concept it is associated with and if it reflects its position in the conceptual system.

These principles correlate with those proposed by the International Organization of Standardization (ISO), as described by Sager (1990: 88-89). This is particularly true of principles number four, seven, eight and nine. Other principles were added specifically for Croatian, in relation to the current language policies. According to Halonja and Mihaljević (2012: 101), principles number one, four and nine hold the highest hierarchical position of the nine and should take precedence over the others when creating new terms in standard Croatian.

4.2 Most common Croatian IT term formation procedures

Terms can be created in two main ways: through primary or secondary term formation. Primary term formation follows the creation of a new concept and results in the coinage of a term for this new concept, while secondary term formation refers to the creation of a new term for an already existing concept (Sager 1990: 80), as is often the case when taking over terminology from another language. Since terms in the field of IT in Croatia are mostly developed through secondary term formation, Sager's classification (1990: 90) will be used for this paper.[4] There are four main ways of introducing new IT-related terms from English into Croatian, illustrated here with examples from the table of researched terms (Appendix 2).

1. Borrowing – an existing term is borrowed to be used in a new environment and terminological field. For this paper, Sager's classification is further divided into two groups: borrowing from a foreign language (here: English) and borrowing within the same language, i.e. from the general language or another terminological field (here: Croatian)[5]. In time, a foreign borrowed term can (but does not have to) be adapted to the Croatian morphological, phonetic and orthographic system to varying degrees. According to some authors, this should not be the primary way of introducing new terms into Croatian, but it often is (see Halonja and Mihaljević, 2012: 83). The terms that are borrowed from other Croatian terminological fields or the general language are more adapted than those taken from a different language as they are already part of the language system. Their meaning in the target language is usually similar or related to that in the source, but they do not share their literal meaning (as in the *handshake* example below).

[4] While the aforementioned Croatian authors (see Mihaljević 2003: 96-105 and Halonja and Mihaljević 2012: 83-84) provide a classification of terminology development procedures, they do not differentiate between primary and secondary term formation.
[5] It can be argued that some terms borrowed from Croatian are actually semantic borrowings or extensions; however, in many cases the terms were literally translated from English, with the processes of semantic extension or semantic borrowing already occurring in English prior to their translation.

software > software or *softver* (varying levels of adaptation to standard Croatian)

handshake > *dogovaranje* (from general Croatian – arrangement or agreement on something)

2. Literal translation – the lexical bases of a word or phrase are translated literally in order to form a new word or phrase in the target language. This means that, unlike borrowings from the same language, the English and Croatian terms do have the same connotative meaning. These are usually well-adapted to the target language phonetically, orthographically, and morphologically (which adheres to the Croatian terminological principles). This term formation method can result in completely new words or add new meaning to existing words.

computer > *računalo* (from *računati* – to compute; a new word in Croatian, but translated literally from English)

stack > *stog* (a literal translation and an existing general-language word in both languages)

3. Neologism – the creation of an entirely new word with its own meaning in the target language. Many originally English IT terms are neologisms, which are sometimes literally translated into Croatian, but there are also Croatian neologisms. This type of term formation procedure is reflected in developing previously non-existent words, deriving new words from existing ones, joining words or lexical bases together etc. For example:

object program > *odredišnik* (derived from *odredište* – destination, goal)

software > *napudbina* (derived from *naputak* – instruction, direction)

4. Paraphrase – IT terms are often introduced into Croatian as multi-word units which together refer to a single term/concept. While this solution is often well-adjusted to the standard language, deriving new terms from paraphrases can be problematic, as they contain multiple words. This is why single-word terms could be used, even if they are not the preferred or standard option (Škifić and Mustapić, 2012):

software > *programska podrška* (more adapted to standard Croatian than the mentioned software and *softver*, but deriving related terms from it is difficult)

Of note are also acronyms and initialisms, usually kept in their original, English form, which can be seen as borrowings – the term Local Area Network will be translated as *lokalna mreža* or *područna mreža*, to name but two options; if used as LAN, the Croatian abbreviation will not be translated as *LM* or *PM*, but will be kept as LAN, such as in the case of *LAN-kartica*.

The terminological principles and term formation procedures listed here will be important in the upcoming sections as part of an analysis of terms and their translation candidates found in various sources, as well as a description of survey results. These principles and procedures will also serve as a reference point to compare Croatian term candidates (where there are multiple options) to see which are "better" or "more suitable" than others, i.e. which are in accordance with the described principles. This does not mean that the term candidates described as "better" are objectively the superior choice, but rather, they will we used to compare what is prescribed as "good" and what is chosen by users as such.

5. Research questions, hypotheses and aims

The main questions this research works towards answering are the following:

1. Is the Croatian IT terminology system standardized, i.e. is there a resource which would provide IT experts as the primary terminology users with definitive, formal and widely-accepted Croatian terms for this field?
2. Do scientists, academics and other IT experts use these terms consistently and uniformly in formal communication?

The hypotheses are negative answers to these questions – it is assumed that there is no institution or resource on which one could rely to find definitive terminological answers; thus, IT professionals do not (and cannot) use this terminology consistently and uniformly country-wide or even within a single institution. They have to find their own translations for English IT-related terms. This has been explained in Sections 2 and 3, and will be further shown on specific examples in Section 7.

Therefore, the aim of this research is to determine the current state of IT terminology in Croatia by taking a closer look at the existing term candidates (selected for this purpose from a number of sources) and by conducting a survey among IT experts who are the users of this terminology in order to get their opinions directly.

6. Methodology

The study consists of two parts: an analysis of terms, i.e. term candidates found in diverse sources, and a survey conducted among members of the academic and scientific community in Croatia. The aim of the survey is to find out which the participants' preferred translation solutions are and what they think of the state of IT terminology in Croatia in general.

6.1 Term analysis

In order to decide which terms would be used in the survey, both the English terms and their Croatian terminological equivalents had to be found and researched in various sources. The analysed English terms, provided in Appendix 2 (along with their Croatian term candidates), were first randomly selected from a list composed during a longer period of time by reading IT-related literature from relevant scientific sources[6] and collecting the terms used in them. This larger list of terms was then narrowed down based on how often they could be found in the used Croatian resources. A number of normative and conventional sources were used to find Croatian equivalents, all of which are listed with their respective codes in Appendix 1. For example, if a term is too new or too specialized to be present in any of the listed Croatian sources, it was not further researched. While the lack of a term's terminological equivalent(s) is an indication of the state of a language's terminological system, it was not the topic of this research, as translation options were required. Additionally, the narrowed-down list of terms was further adjusted to include only those terms that are neither too complex nor too specialized so as to ensure that the participants are

[6] These include, but are not limited to, various articles on websites such as computer.org and acm.org.

familiar with them. Familiarity with a term should increase the possibility that the participants had already come across it in a Croatian context, irrespective of what form the term was used in. This resulted in 45 English terms. Ultimately, the final list of terms used in the survey was further narrowed down to 20 terms by selecting terms with at least two translation candidates and not more than five, as well as translation candidates that were used in more than just one Croatian resource.

The Croatian sources used for this research are both digital and printed, including printed and online dictionaries, glossaries, linguistic literature, papers written by students, experts and academics etc. This list can therefore serve as a general overview of what resources are available for this purpose in Croatia. Corpora were not used for this research, as there are currently no corpora of standard Croatian which include scientific texts or other types of texts requiring formal language in the field of IT.

The sources used in this study were divided into larger groups, with individual resources listed with their codes (see Appendix 1 for a list with complete references). The following groups of resources were identified:

- Linguistic literature: Books and papers written on the topic of terminology in general and about the field of IT. The authors offer an overview of available options or suggestions on how a term can be translated or formed. This includes online style guides for standard, formal communication. This group includes the following: Mihaljević 1, Mihaljević 2, Halonja, Jezični savjetnik, Bolje je hrvatski.
- Printed dictionaries and lexicons: These are general language English-to-Croatian dictionaries or specialized terminological dictionaries and lexicons. Only the resources providing explicit translations from English into Croatian were used. These are the most standardized resources, but some of them are obsolete, e.g. one of the dictionaries was published in 1991, before the internet (and all related terms) as we know it today existed. Overall, there were few recent dictionaries or lexicons (with English terms) available in Croatian at the time of writing this paper. This

group encompasses the following: Bujas, Kiš 1, Kiš 2, Microsoft Press, Šijak, Babić, Školska knjiga, Štambuk, Jakobović.

- Online scientific databases: Primarily the CROSBI scientific database (code: CROSBI, with all the individual authors listed in this category), and the HRČAK open-source journal portal (code: hrčak, with all the individual authors listed under this category); again, only papers providing explicit translations from English into Croatian were used, such as in keywords or abstracts of scientific or academic publications.
- Online dictionaries, glossaries and translation tools: These are the most up-to-date, but the least normative and standardized, as they include personal glossaries compiled by university professors, online dictionaries relying on crowd-sourcing, and tools such as Google Translate, suggested by IT experts as a place they look for translations. The following resources belong in this group: Nazivlje, Microsoft, Groš, Google Translate, EUdict, Begušić, Muljević.

Based on these resources, the translation options i.e. term candidates were organized in a table: the first column contains the English terms, the columns to the right are suggested terminological equivalents. All the resources a term candidate appears in are listed with their respective codes. The analysis will be discussed in Section 7.

6.2 Survey

The participants in the survey were Croatian IT professionals, the main users of this terminological system. The survey can be found in its entirety in Appendix 3. In short, it consisted of these three parts:

In the first part, respondents had to choose from among the 2 to 5 offered term candidates. They also had the option of suggesting their own terminological equivalent for a given English term. After narrowing the list of 45 terms down, as described in Section 6.1, in total there were 20 questions, i.e. English terms, with their respective translation options. Having completed this part, participants could not go back to it or redo it.

In the second part, participants were asked to choose the best method for secondary term formation in their opinion; furthermore, they offered their opinions on questions relating to IT terminology in Croatia and terminology in Croatia in general. They rated statements on subjects such as the current state of this terminological system in Croatia, its level of standardization, etc.

The third part of the survey gathered demographic information on the participants, including their age and level and field of education. A field for optional comments regarding the survey was also provided.

The first two parts of the survey were expected to showcase which solutions are preferred in actual use and whether the participants' choices correspond with the principles described in Section 4. The individual choices made in the first and second part could be compared for consistency, for example, to see if a respondent tended to choose English (borrowed) terms in the first part, while expressing a general preference for Croatian terms in the second part.

The survey was made using the LimeSurvey[7] platform and was sent to university staff and student groups of the following institutions: The University of Zagreb's Faculty of Electrical Engineering and Computing, The University of Dubrovnik, The Zagreb University of Applied Sciences, The University of Split's Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, and social media groups (Facebook) with members related to this field. The survey was anonymous in its entirety, aside from the information the participants entered about themselves. The results are described in detail in the next Section of this paper.

7. Results

The next two subsections report on the results of the two parts of the present study, the first one being the term analysis, the second the survey with IT professionals.

[7] The platform can be used free of charge by users with a valid AAI@EduHr account, including students of the University of Zagreb and the Faculty of Humanities and Social Sciences.

7.1 Results of term analysis

A total of 45 IT-related terms were analysed. Their 223 translation options (61 borrowings, of which 35 from English and 26 from Croatian, 92 literal translations, 16 neologisms, 47 paraphrases, 2 descriptive and 5 mixed translations) were found in various resources. The resources were assigned codes for easier use (see Appendices 1 and 2) and added to a table for a clear overview of the researched English terms and their possible translations. The English terms each have at least one and up to 11 Croatian term candidates, with an average of 5 possible term equivalents per term (mean: 4.9, median: 5), meaning that most English terms do not have an agreed-upon Croatian counterpart, or that there are many synonyms for one term, all of which should be avoided in a scientific and terminological context. Considering that the majority of the resources used are provided by relevant normative institutions and authors, it can be concluded that the Croatian IT terminological system, i.e. its existing terms and principles, has not yet been thoroughly standardized, at least for the needs of formal communication. Even terms or concepts which could be considered basic, such as *hardware*, are problematic when it comes to Croatian terminology. For example, there are eight possible translations for the term *hardware*, some of which are only used by one to two authors (for a full list of references see Appendix 2), as shown in Table 1.

Table 1 Translation options for the term hardware

hardware	**sklopovska oprema** (Školska knjiga, Muljević)	**hardver** (Mihaljević 2, Bujas, Microsoft, Nazivlje, Microsoft Press, Šijak, Školska knjiga, Vrhovski…)	**strojna oprema** (Kiš 1, Halonja 95, Microsoft Press, Bolje je, Jezični savjetnik, EUdict, Begušić)	**sklopovlje** (Kiš 1, Halonja 95, Šimunac, Kiš 2, Microsoft Press, EUdict, Begušić, Muljević…)	• **tehnička oprema** (Bujas, Muljević) • **strojevina** (Kiš 2) • **računalna oprema** (Muljević) • **sklopnjak** (Mihaljević 2)

In terms of formation procedures, there are paraphrases (*sklopovska oprema, strojna oprema, tehnička oprema, računalna oprema)*, neologisms (*strojevina, sklopnjak, sklopovlje)* derived from Croatian words *stroj* (machine) and *sklop* (circuit, construction), and a borrowing from English (*hardver*), which

is semi-adapted to Croatian. *Hardver* was listed in the largest number of sources. None of the sources providing multiple term options singled out any one option as the most adequate in any way. Considering the Croatian terminological principles (Section 4.1), according to principles number 1 and 2, terms derived from Croatian (Slavic) words, such as *sklop* and *stroj (sklopovska oprema, strojevina,* but also *računalna oprema*) should be given priority over borrowed words (here: *hardver*). Even the internationalism *tehnička* in *tehnička oprema* should be prioritized over *hardver*. The actual use by professionals and their preferences will be discussed as part of the survey results; nevertheless, the number of authors using *hardver* shows that this term should be foregrounded, as stipulated by terminological principle number three. All of the terms comply with standard Croatian (principle 4). If shorter terms should be prioritized (principle 5), only *strojevina, sklopnjak, sklopovlje* and *hardver* remain as options. All of the shorter terms can have new words derived from them (principle 6). Principle 7 (polysemy/synonymy) has already been mentioned; principle 8 is not relevant in this case. Finally, all of the suggested one-word term candidates could reflect their position in the conceptual and terminological system (principle 9) if their respective related terms were adjusted – e.g. *hardver* should be used with *softver*. Overall, even if principles number one, four and nine should be prioritized, there is still no one definitive translation to be judged and used as "the best" option. Some Croatian authors, such as Halonja and Mihaljević (2012: 83-85), suggest using terms derived from Croatian words, in this case and others, but these suggestions do not always align with the professionals' opinions and preferences (see Section 7.2).

A similar example is that of the terms *assembler*, *compiler* and *interpreter*. These three concepts are similar, but also have specific individual qualities. This is important for the Croatian terms, considering that all three have *prevodilac, prevodioc* or *prevoditelj* as Croatian options. Nouns ending in -oc and signifying a subject performing an action are not part of standard Croatian; the suffix -telj is preferred over -ac for the same nouns (Institut za hrvatski jezik i jezikoslovlje 2019). If only Croatian words are considered (as recommended), this would mean that for all three terms the Croatian equivalents could be *(programski) prevoditelj*, which is a problematic case of polysemy and can be confusing even

with context. While the terms *assembler* and *interpreter* each have a clear Croatian option (*zbirnik* and *tumač*, respectively), there is no standard Croatian term suggested for the term *compiler*. This goes against terminological principle number nine. This points to the practicality of using borrowings (listed in many sources), adapted to Croatian: *asembler*, *kompajler* and *interpreter*. More on this in Section 7.2. It is interesting that the term candidate *zbirnik* was also proposed for the English term *bus* in two resources, creating another example of polysemy (*assembler* and *bus* are both *zbirnik*). Another example of polysemy is the Croatian term candidate *pretraživač* being suggested for two terms (*browser* and *crawler*) in two different sources.

For the term encryption there are four options, three of which are provided by multiple authors (*šifriranje, kriptiranje, enkripcija*), one by a single author (*zakrivanje*). While the former two are literal translations, their lexical bases are borrowed words (from French and Greek, albeit already accepted in Croatian). The third option, *enkripcija*, is an adapted English borrowing, while the last one is derived from a Croatian word, but it is not given in any other resource. They are all short terms, new words can be derived from them (maybe even more so from the foreign ones), and they could all work well within the conceptual and terminological system. Still, it seems that the foreign options are preferred to the domestic one, despite the terminological principles.

Pleonasms and paronyms can also be found among the analysed term candidates. For example, the term *LAN-mreža* (engl. LAN) is redundant since the word *mreža* is included in the English acronym (Local Area *Network*). Similarly, there are paronyms among the translation options: *e-mail* is translated as both *elektronska* and *elektronička pošta*. In this case, *elektronička* is deemed correct[8], as this adjective refers to electronics, as opposed to *elektronska,* which describes something relating to electrons. The term *operating system* also resulted in paronyms *operacijski* or *operativni* sustav. *Operativni* is derived from the internationalism *operativa* (capability or possibility of action). *Operacijski* stems from *operacija*, which in Croatian primarily refers to medical operations. Therefore, the term that should be given preference, according to the

[8] Retrieved 02 June, 2019, from http://jezicni-savjetnik.hr/?page=7

terminological principles, should be *operativni sustav,* even though it was listed in fewer sources.

The possible equivalents of the term *thread* provide an example of dialects entering the terminological system. The proposed terms are the general language borrowing *dretva*, which is a northern Croatian dialectal word (and also a Germanism) for shoemaker's thread[9], and *nit*, a standard Croatian word and a literal translation. Despite being dialectal, more resources suggested using *dretva* than *nit*. The IT terminology development in Croatia is further influenced by other terminological systems: for example, the Croatian borrowing *osmak* comes from agricultural terminology; it is corn that has eight rows of kernels on one cob. The English original, *byte*, is a neologism and has no relation to corn.

The term *boilerplate* is an example of a term relatively newly introduced into this field. While it exists in other fields, in the context of IT, it has only recently become popular. Therefore, there were virtually no Croatian translations in any of the used resources, aside from *standard* and *ponavljajući tekst*, each in a single resource only, one of which is from 1995. This is an instance of new terms not being translated or developed quickly enough as they enter Croatian, so authors use them in their original form or translate them on their own.

However, not all terms are necessarily problematic – the terms *database* and *programming language* are examples of there only being one term candidate provided by a large number of sources. These are *baza podataka* and *programski jezik,* respectively. While they are both literal translations consisting of two words (and shorter terms are preferred), these translations were consistently the only ones provided in the resources. They show that standardization and agreeing upon a single term is possible and could be achieved in the future.

To sum up, the analysed terms point towards the Croatian IT terminology system not yet being thoroughly standardized and unified. There is still a large number of English terms which each have numerous Croatian counterparts. This is not a situation that is normally sought after for formal texts. There are many synonyms, as well as polysemous and unclear (translations of) terms. This can

[9] Retrieved 02 June, 2019 from http://hjp.znanje.hr/index.php?show=search_by_id&id=fF9jWhA%3D&keyword=dretva

lead to a number of issues, primarily a lack of clarity and precision. While there are examples of well-translated terms that experts agree upon, the majority of the analysed terms are nonetheless problematic.

This is only a small segment of an ever-growing terminological and technological field, and these results cannot be applied to all terms; however, considering that most of the concepts and terms selected for this analysis are relatively simple and well-known, it can be assumed that the situation is similar with more complex and less frequently used terms and concepts.

7.2 Survey results

The survey was taken by 75 participants: 48 (60%) of the participants are students of IT or related scientific fields, with the average age of around 23. Twenty (27%) participants are engineers of the same fields, aging from 22 to 51, and seven (10%) participants, aged 30 to 63, have a Ph.D. degree in IT-related fields. There were no noticeable differences between the age groups when it comes to the preferred term formation procedure. However, in order to reach a definite conclusion about the preferences of each age, gender or level of education group, such a survey would have to include a larger number of participants in each of these groups. Therefore, the focus of this research was on its global results, rather than on specific segments or groups of the participants.

7.2.1 Secondary term formation choices

The first part of the survey produced mixed results when looking at the consistency of choice based on term formation procedures. For example, in Question 1, 46% of participants chose *asembler* (a foreign borrowing) instead of the other options, all of which are Croatian words. However, in Question 5, 51% chose *sklopovlje*, a neologism and derivation of the Croatian word *sklop*, making it a "better" choice according to terminological principle number one (see 4.1). Furthermore, the second choice in Question 1 was *programski prevoditelj*, the previously explained example of polysemy, relating to Questions 8 (*interpreter*) and 16 (*compiler*), both of which had *prevoditelj, programski prevoditelj* or *prevodilac* as options. Since they are very similar, issues could occur when e.g.

two or all three terms are used, as readers would likely not be able to differentiate between *programski prevoditelj*, *prevoditelj* or *prevodilac.* However, there are other recommended solutions, such as *zbirnik* (assembler), *prevodnik* (compiler) and *tumačnik* (interpreter) (Mihaljević, 1993: 165). This set of terms would be more fitting in light of the terminological principles, and the terms themselves are relatively clear. However, only one of them (*zbirnik)* was actually found in the various sources as a confirmed translation (and this is problematic since *zbirnik* is also a term candidate for *bus*). Therefore, this is a case when borrowed terms are useful and, accordingly, *asembler* (46%), *interpreter* (47%) and *kompajler* (39%) were all the first choice in their respective questions.

In Question 2, 46% of participants chose *dubinska analiza podataka* for data mining, which is a paraphrase, while 35% opted for *rudarenje podataka*, a literal translation. Two terminological principles clash: the former option was chosen by more participants, but the latter is shorter and should therefore be preferred. The English borrowing *data mining* was chosen by 17% of the participants. *Majnanje* and *sakupljanje podataka* were also added as options.

In Question 3, *elektronička pošta* was the translation of choice for over half of the participants (52%), and another 36% said they would use *e-mail*, a borrowing from English. Only 11% of the participants chose the "incorrect" term *elektronska pošta*, and one participant suggested they would use just *pošta*, which is not necessarily the clearest option.

When asked to choose an equivalent for *framework* in Question 4, the preferred term candidate (50%) was *razvojni okvir*, a paraphrase of Croatian origin, which is a good solution according to terminological principle number one. *Radni okvir* was also a choice (25%), but since *razvojni okvir* was chosen by more participants, this can be seen as the go-to translation of the term framework. Another solution was *radno okruženje*, but this term is inconveniently similar to potential translations of the term *environment* (or *development environment*) within this field. Participants also suggested using the borrowed term *framework*, just *okvir*, or *razvojno okruženje*. One participant said that their choice was context-dependent.

Questions 5 (hardware) and 10 (software) are discussed together due to their relation as defined by principle number nine. For hardware, s*klopovlje* was the first choice (51%), and *hardver*, an English borrowing, the second (40%). Other options include the borrowing in its original form, *hardware*, followed by the paraphrase *strojna oprema* and the term *tvrdi disk*, which is not a correct translation in this case. However, *softver* was chosen by most participants (40%) for the term software, meaning that 10 participants would combine *sklopovlje* and *softver*, going against terminological principle number nine. The second choice was *programska podrška* (34%), followed by the English borrowing in its original form, *software* (17%). Four participants from this last group combined a borrowing from English (*software*) with a Croatian neologism (*sklopovlje*). This brings the percentage of participants who combined software/*softver* and *sklopovlje* to 14, or 18%. The participants also chose *programska oprema* and suggested the term *aplikacija*. The terms hardware and software are therefore problematic, as they are seen as a pair and their translations should be developed accordingly, e.g. by combining *hardver* and *softver* (20 participants or 27%), or *strojna oprema* and *programska oprema* (only two participants or 2.3%), as discussed in the previous section of this paper.

Question 6 focused on the *operacijski – operativni sustav* paronyms: 73% of the participants chose the term that does not adhere to terminological principles, *operacijski sustav*. The more accurate term *operativni sustav* was the preferred option for only a quarter of the participants. Additionally, one participant suggested they would just use the English abbreviation OS.

An example of synonymy can be found in Question 7 (*save*). *Spremiti* (56%) and *pohraniti* (39%) are both Croatian words and literal translations which do not go against any of the terminological principles. The deciding principle is number three, as *spremiti* was chosen by a larger number of participants. Only four participants altogether opted for other solutions – *sačuvati*, *sejvati* and save.

The mixed solution (borrowing from English and literal translation) *web-stranica* was chosen by a majority of the participants (85%) in Question 9 (*website*), although the literal translation *mrežna stranica* would be a better choice according to the terminological principles. One participant added *internet*

stranica as their reply, which is not grammatically correct and should instead be *internetska stranica* or *internet-stranica*. None of the participants chose the proposed mixed candidate *web-mjesto.*

Question 12 (*thread*) is particularly interesting, as it is a clear example of a dialectal term entering a specialized field's terminological system; 69% chose *dretva* as their preferred translation. That is perhaps due to the fact that some participants (as they explained after filling in the questionnaire) did not know this was a dialectal word and perceived it as a neologism. The literal translation *nit* was chosen by 23%, while another 7% added the English borrowing thread. *Dretva* is another example of the majority of participants choosing a non-standard term.

The same can be concluded for Question 13 (*update*) and 14 (*widget*). *Ažurirati* was the choice of an overwhelming majority of the participants (95%), displayed in Table 3, even though this is a literal translation of French etymology, which goes against a number of terminological principles outlined in 4.1. Similarly, 76% chose the term *widget*, which is a non-adapted borrowing from English. Even though the participants could also choose *posuvremeniti* or *dopuniti*, which are more in line with the terminological principles as they are of Croatian origin, none of the participants chose these term candidates.

The term *dongle* (Question 15) was problematic for some participants (5%) as they did not know what it refers to. The preferred translation here was *hardverski ključ* (45%), followed by *privjesak* (24%) and *ključić* (16%). It should be noted that 15 participants (20%) who chose *sklopovlje* in Question 5 selected *hardverski ključić* in this question, which is an example of inconsistency within the terminological system (they had the option of adding their own answer to be consistent – e.g. *sklopovski ključić* if they chose *sklopovlje*).

Priručna memorija was the most favoured choice (52%) for the term *cache* (Question 17), followed by the non-adapted borrowing *cache* (29%) and neologism *predmemorija* (19%). Once again, multiple terminological principles are in conflict: *priručna memorija* was selected by most participants; *predmemorija* is shorter and better suited for derivations; more participants chose the non-adapted borrowed term *cache* than *predmemorija*.

Question 18 (*random access memory (RAM)*) was also interesting in that a number of participants (25%) said they would use the abbreviation *RAM*, even though it is originally in English, or that they would use the pleonasm *RAM-memorija* (5%). Still, 68% chose *radna memorija*, which is an acceptable choice in light of the terminological principles in 4.1. However, it is debatable whether this (descriptive) paraphrase can or should be applied to only this type of computer memory.

The most favoured option for *locale* (Question 19) was *regionalne postavke* (79%), followed by *regionalna shema* (15%). Of note are also the replied the participants added themselves: *lokal* (colloquially used as the equivalent of the English word *bar*), *lokalitet* (locality or site) and the adjective or adverb *lokalno* (local(ly)), presumably due to their etymological similarity. None of the last three options were found in the sources consulted for this study.

Finally, the majority of participants (65%) chose the English borrowing *enkripcija* in Question 20 (*encryption*), even though *kriptiranje* (20%) would be more fitting according to terminological principles number 2, as its lexical base is borrowed from Greek. The third option, *šifriranje* (15%) was chosen by the fewest participants, although the words it derives from and their French lexical base are already accepted in Croatian.

It can be concluded at this point that there is a lack of consistency when it comes to preferred choices for the use of IT-related terms in Croatian formal texts. The respondents sometimes chose borrowed English terms with varying degrees of adaptation to standard Croatian; other times, they chose terms which can be considered Croatian in their entirety and are well-adapted to the terminological principles which are seen as normative in Croatia (4.1). This can be related to the next part of the survey: professionals' opinions.

7.2.2 Professionals' opinions

When asked about their general opinion on the best method of term formation in this field, 45% of the participants opted for mixed solutions (such as *web stranica* and *LAN-kartica*). Borrowing of foreign terms with their adaptation (grammatical, orthographic, etc.) to the Croatian language came in second

(23%), and neology third (17%). Unadapted borrowing and literal translation had considerably fewer proponents (9% and 5% respectively). These replies mostly correspond with those provided in the first part of the survey.

In the next set of questions, the respondents were asked to rate their agreement with several statements on a scale of 1 to 5 (1=strongly disagree; 5=strongly agree). The first of these questions was whether all terms should be as adapted to standard Croatian as possible, with participants choosing the middle ground (average rating: 3.24), which also corresponds with their varying choices in the first part of the survey. They mostly disagree with the statement that Croatian IT terminology should by no means contain any untranslated or inadequately adapted foreign words (average rating: 2.2), as confirmed by the choice of e.g. *widget* as the term they would use in a formal context.

Responses are almost equally distributed (average rating: 3.13) when it comes to the claim that every country should have a standardized terminological system for each scientific and academic field within their own language. They agree slightly more with that claim when asked about standard Croatian specifically, the average rating being 3.56. However, with an average of 3.89, participants tend to agree that the Croatian IT terminology is not developing quickly enough compared to the rest of the world, implying that they are dissatisfied with its current state. This is confirmed in the subsequent question, where the respondents overwhelmingly disagree (with the average rating of only 1.88) with the statement that Croatian scientific and professional communities using IT terminology are in agreement about said terminology. Participants agree (average rating: 4.32) that terms should be translated and formed through collaboration between IT experts and linguists, as opposed to just one of the two professions working alone. Finally, with an average answer of 2.02, the participants disagree with the statement that the Croatian IT terminology system is well-developed and translated.

A small number of participants made additional comments about the survey. On the one hand, some said that the research topic is relevant due to the current state of standard Croatian and IT terminology (one even stating that Croatian is "dying out"); on the other hand, one respondent said there is no point in learning

Croatian terms, as only the English terms will ever be used for practical purposes (i.e. in a workplace).

Overall, the results of the term analysis and survey for the most part confirm the hypotheses stated in Section 5: Although some Croatian terms found in resources and preferred by respondents seem consistent and abide by the prescribed principles, there are numerous examples of synonymy (*spremiti* and *pohraniti)*, polysemy (*programski prevoditelj*), pleonasms (*RAM-memorija*) and paronyms (*elektronska* vs *elektronička pošta*). Such cases are typically undesirable within a well-defined terminological field and should especially not be used within the same text by experts in formal situations. While the results of this study are based on only a fraction of the entire terminological field and a small number of survey participants, they still provide clear examples of problematic areas. A number of opportunities for further research on this topic are available, such as analysing more recent terms and concepts, researching improvised translations created by IT experts, or discussing the motivation behind choosing a specific term candidate (e.g. in the case of *dretva*).

8. Conclusion

IT terminology is a growing field influencing not only the formal discourse of related scientific and academic texts, but also our everyday lives. Nevertheless, the Croatian IT terminological system is not keeping pace with these developments. Most IT-related terms are imported into Croatian from English and are adapted to varying degrees. While there are tendencies to remove all foreign and unadapted borrowings, the actual use and preferences paint a different picture.

In this paper, a list of 45 IT-related terms in English was composed (Appendix 2). Their possible Croatian counterparts were researched, found in various resources and listed so as to give an overview of the available Croatian term candidates. Some of these options were analysed in more detail to be assessed according to the prescribed terminological principles in Croatia. Twenty of the 45 analysed terms were used in a survey conducted among professional users of this terminological field. They also gave their opinions on which

translation options they found best and what they thought about the current status of Croatian IT terminology. The results have shown that there are situations in which there is consistency and clarity; however, there are also cases showing that the Croatian IT terminological system lacks systematization and consistent implementation of the accepted standards. There is also a need for relevant institutions to focus on this field and to strive towards working on developing and maintaining this system. This could be done by closer cooperation with professionals from the field of IT, by conducting research on this topic to see what options are spontaneously used or created, as opposed to trying to enforce ones created by linguists, or by providing clear and concise guidelines for professionals to rely on when unsure how to translate, develop or create a term in the field of IT (and other fast-developing fields). Considering the influence of this scientific and technological domain, it can be expected that the needs for terminological research will grow, as well.

References

Babić, Stjepan and Žic-Fuchs, Milena. 2007. *Rječnik kratica.* Zagreb: Nakladni zavod Globus.

Bujas, Željko. 2008. *Veliki hrvatsko-engleski rječnik.* Zagreb: Nakladni zavod Globus.

Cabré, M. Theresa. 1999. *Terminology - Theory, methods and applications.* [Translated by Janet Ann DeCesaris]. Vol. 1. Amsterdam/Philadelphia: John Benjamins.

Čemerin, Vedrana, Valenčić, Davorin and Rubić, Ivana. 2014. "Hrvatska i engleska računalna terminologija s osvrtom na upravljanje u kriznim uvjetima." *Dani kriznog upravljanja.* Velika Gorica: Veleučilište Velika Gorica. 349-362. Accessed June 02, 2019. https://dku.hr/wp-content/uploads/2016/09/DKU-zbornik-radova-2014.pdf.

Encyclopaedia Britannica. 2018. Accessed June 02, 2019. https://www.britannica.com/topic/The-Fourth-Industrial-Revolution-2119734.

EUdict. 2019. eudict.com.

Google Translate. 2019. translate.google.com.

Groš, Stjepan. 2019. *Englesko-hrvatski rječnik.* http://www.zemris.fer.hr/~sgros/stuff/rjecnik.shtml.

Halonja, Antun and Mihaljević, Milica. 2012. *Od računalnog žargona do računalnog nazivlja.* Zagreb: Hrvatski institut za jezik i jezikoslovlje.

Hrvatski terminološki portal. 2019. nazivlje.hr.

Hudeček, Lana, Mihaljević, Milica and Nahod, Bruno. 2009. *Hrvatski terminološki priručnik.* Zagreb: Institut za hrvatski jezik i jezikoslovlje.

Institut za hrvatski jezik i jezikoslovlje. 2019. *Bolje je hrvatski.* http://bolje.hr.

Institut za hrvatski jezik i jezikoslovlje. 2019. *Jezični savjetnik.* jezicni-savjetnik.hr.

Jakobović, Zvonimir (ed.). 2007. *Tehnički leksikon.* Zagreb: Leksikografski zavod Miroslav Krleža.

Kiš, Miroslav et al. 1993. *Englesko - hrvatski informatički rječnik s računalnim nazivljem.* Zagreb: Školska knjiga.

Kiš, Miroslav. 2000. *Englesko-hrvatski i hrvatsko-engleski informatički rječnik.* Zagreb: Naklada Ljevak.

Microsoft Language Portal. 2019. https://www.microsoft.com/en-us/language.

Microsoft Press. 1995. *Informatički rječnik. Cjelovit i detaljan priručnik za posao, školu, knjižnicu i dom.* Zagreb: Znak.

Mihaljević, Milica and Nahod, Bruno. 2009. "Croatian Terminology in a Time of Globalization." In: Ledinek, Nina, Žagar Karer, Mojca, Humar, Marjeta (eds.), *Terminologija in sodobna terminografija.* 17-26.

Mihaljević, Milica. 1993. *Hrvatsko računalno nazivlje.* Zagreb: Hrvatska sveučilišna naklada.

Mihaljević, Milica. 1998. *Terminološki priručnik.* Zagreb: Hrvatska sveučilišna naklada.

Mihaljević, Milica. 2003. *Kako se na hrvatskome kaže WWW?* Zagreb: Hrvatska sveučilišna naknada.

Mihaljević, Milica. 2006. "Hrvatsko i englesko računalno nazivlje." *Jezik* (Hrvatsko filološko društvo) 53 (2): 41-50. Accessed June 02, 2019. https://hrcak.srce.hr/16635.

Mihaljević, Milica. 2007. "Problemi hrvatskoga računalnoga nazivlja (s jezikoslovnog motrišta)." *Studia lexicographica* (1): 61-79. Accessed June 02, 2019. https://hrcak.srce.hr/110771.

Miščančuk, Marija and Vučković, Mate. 2011. "Terminologija računarstva u hrvatskom jeziku." *Zbornik radova Međimurskog sveučilišta u Čakovcu 2(2).* Accessed June 02, 2019. https://hrcak.srce.hr/74960.

Seger, Juan C. 1990. *Practical Course in Terminology Processing.* Amsterdam/Philadelphia: John Benjamins Publishing.

Šijak, Andrija and Lončarek, Darko. 1993. *Informatički rječnik.* Varaždin: Varteks.

Škifić, Sanja and Mustapić, Emilija. 2012. "Anglizmi i hrvatsko računalno nazivlje kroz prizmu jezičnoga konflikta i jezične ideologije." *Jezikoslovlje* 13 (3): 809-839. Accessed June 02, 2019. https://hrcak.srce.hr/95181.

Školska knjiga. 2017. *Veliki englesko-hrvatski rječnik.* Zagreb: Školska knjiga.

STRUNA. 2019. struna.ihjj.hr.

World Finance. 2018. Accessed June 02, 2019. https://www.worldfinance.com/markets/top-5-of-the-fastest-growing-industries-in-the-world.