



**44/2024**  
 Research Journal  
 Studies about Languages  
 pp. 107–122  
 ISSN 1648-2824 (print)  
 ISSN 2029-7203 (online)  
 DOI 10.5755/j01.sal.1.44.36235

## LINGUISTICS / KALBOTYRA

### Preferences of Lithuanian cybersecurity synonymous terms in different user groups

Received 03/2024

Accepted 05/2024

**HOW TO CITE:** Rackevičienė, S., & Utkā, A. (2024). Preferences of Lithuanian cybersecurity synonymous terms in different user groups. *Studies about Languages / Kalbų studijos*, 44, 107–122. <https://doi.org/10.5755/j01.sal.1.44.36235>

# Preferences of Lithuanian cybersecurity synonymous terms in different user groups

Lietuvių kalbos kibernetinio saugumo sinoniminių terminų pasirinkimai skirtingose vartotojų grupėse

SIGITA RACKEVIČIENĖ, Mykolas Romeris University, Lithuania

ANDRIUS UTKA, Vytautas Magnus University, Lithuania

## Abstract

Cybersecurity is a rapidly developing domain, where emerging new concepts are usually first designated in English and then find their way into the usage of other languages. As the Lithuanian terminology in this domain develops, different types of synonymous terms appear in usage, which are treated differently by speakers. The article presents a terminology survey involving 593 respondents from various age groups, from different regions and expertise levels. In the survey, the respondents had to name the most suitable terms for 10 cybersecurity concepts: the respondents could choose the terms proposed in the questionnaire or they could propose their own terms and give the reasons why they made their choices. The concepts and their terminological designations were selected from the Lithuanian-English Cybersecurity Termbase, the dataset of which is based on bilingual parallel and comparable cybersecurity corpora. The quantitative and qualitative analysis of survey results reveals preferences for different types of terms, such as borrowings, metaphorical calques, and descriptive terms, and how these preferences differ across the two segments of respondents: students vs. graduates, and cybersecurity experts vs. general public. The results show that some terminological designations have been already established in the Lithuanian language, while most of them are still competing for their positions. The analysis of the reasons reveals that accuracy and clarity are the main factors for selecting a term. The research contributes to the standardisation of cybersecurity terms in Lithuania and provides insights into user preferences and the reasons behind them.

**KEYWORDS:** cybersecurity terms, term synonymy, borrowings, metaphorical terms, descriptive terms, terminology survey, term preferences.

## Introduction

The cybersecurity domain has gained special relevance in the 21<sup>st</sup> century, marked by digital transformation and global connectivity. Digital data and services have become invaluable assets that play an indispensable role in the functioning of society and life of every individual. However, this increasing reliance has made them vulnerable to malicious activities. Attacks on net-

works and information systems, data theft, and disruption of digital services have become a part of everyday life, and, consequently, the need to understand and use cybersecurity terminology has increased considerably. Cybersecurity terms are created primarily in English, and high dynamics of the domain makes it particularly difficult for other languages not to lag behind. Therefore, English terminology prevails in intra-disciplinary cybersecurity communication among experts. Development of national cybersecurity terminology poses various challenges as it has to speedily reflect the ever-changing realities and meet the needs of various groups of people with different levels of the domain knowledge.

The Lithuanian cybersecurity terminology is developed and used in various discourses and in different registers. It is characterised by huge variation and usage inconsistency. There is a great need for standardisation of the Lithuanian cybersecurity terminology, which is a pre-condition of term usage in legal acts and other official documents, as well as an important instrument in dissemination of cybersecurity knowledge and raising cyber awareness among the general public.

The aim of the research is to determine the preferences of term types among diverse Lithuanian language user groups and understand the reasons behind these priorities. To achieve the aim, we conducted a terminology survey on the synonymous cybersecurity terms in Lithuanian and performed the quantitative analysis of the obtained results. A total of 593 respondents, spanning different age groups and areas of expertise, participated in the survey and expressed their preferences for the most suitable terms to designate 10 cybersecurity concepts and provided reasons for their votes. The paper presents:

- the background of the research: insights of various scholars on reasons for synonymy in terminology, approaches to defining and managing synonymous terms, as well as a brief overview of studies on Lithuanian synonymous terms;
- the preparation for the terminology survey: investigation of terminological variation in the Lithuanian-English Cybersecurity Termbase based on the dataset from bilingual cybersecurity corpora;
- the survey form, variables, and their structuring;
- the survey sample and sociodemographic data of respondents;
- the comparative data analysis results revealing preferences of synonymous terms designating the same concepts in two respondent segmentations: Students vs. Graduates and General public vs. Experts (experts of IT and/or cybersecurity domains);
- the comparative data analysis results revealing reasons for preferences of the synonymous terms in the same two respondent segmentations.

We believe that the research results will provide an overview of the trends in the cybersecurity term usage and contribute to the discussion on their standardisation.

### **Research Background: Synonymy in Terminology**

#### *Synonymy in Terminology and its Causes*

Despite the firm advocacy of the principle of univocity in traditional terminology, development and usage of terms are subject to common phenomena and interactions in natural language. As terms are lexical units, they develop various lexical relations among themselves and other linguistic units and often trespass the traditional one concept – one term relationship requirements. One type of such lexical relations, which emerges among terms, is synonymy.

Terminology scholars present various reasons for appearance of synonymous terms which sometimes differ considerably in their usage. Among the multiple factors of different usage patterns of synonymous terms, L'Homme (2020, p. 152) enumerates the following ones: geographical areas in which designations appear, time frame, level of specialisation in communication (which depends on knowledge symmetry between the addressor and addressee: expert vs. expert, expert vs. layperson), communication channel (oral vs. written), theoretical framework.

Cabré (1999, p. 110) also points out the impact of formal/informal register on occurrence of synonymous units (e.g. scientific/popular terms or standard/dialect forms), while Schmitz (2015, p. 18) mentions commercial reasons: “company or product-specific efforts to use terminological differences as one means of positioning a product in the market”. Schmitz also indicates that synonymy may be especially extensive in evolving domains where concepts are still undergoing development. In evolving domains, “competing terms are used in parallel until unambiguous terms are gradually established, either through a natural selection process or by conscious standardization” (Schmitz, 2015, p. 17). But even after standardisation, synonymous terms may remain in use for long period of time (ibid.).

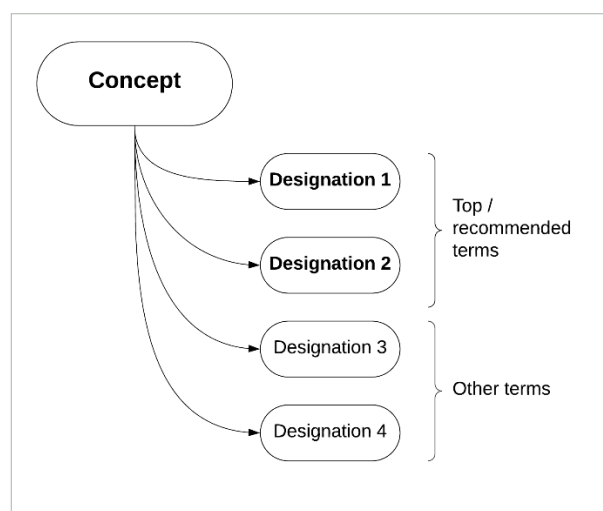
#### *Approaches to Defining and Managing Synonyms in Terminology*

Terminology focuses on exact synonymy, i.e. designations of the same concept: terminological designations, as well as other types of designations, such as symbols or icons (Schmitz, 2015).

ISO 704 highlights that such synonyms have to be “always interchangeable” (ISO 704, 2022, p. 57). Thus, they cannot possess any semantic or pragmatic differences. Designations which are assigned to concepts whose intensions are not absolutely identical and which are interchangeable only in some situations, are called quasi-synonyms (ibid.). The same position is advocated by Cabré (1999, pp. 109–110), who states that, though various approaches can be taken, but “strictly speaking, terminology only considers synonyms to be semantically equivalent units that belong to the same historical language and to the same formal register”. Meanwhile, L’Homme (2020, pp. 151–153; 161–163) distinguishes synonymy conceptions based on knowledge-driven approach and on lexicon-driven approach. The former deals mainly with exact-synonymy, whereas the latter considers also near-synonymy – designations which share many semantic components, but not all of them and are interchangeable in most, but not all contexts.

The onomasiological approach applied in most terminology resources, focuses on synonyms which designate the same concept and are organised around the concepts they pertain to in concept-oriented entries in terminology databases (see Fig. 1). In order to reduce confusion and ambiguities, which might have severe consequences in specialised communication, the included synonyms are usually graded, thus encouraging users to select the top one(s). Grading usually depends on the chosen approach and selected criteria by terminology resource compilers (c.f. Cabré, 1999, p. 144; IATE User’s Handbook, 2022).

In terminology studies, term synonymy is often contrasted to terminological variation. L’Homme (2020, p. 153) states that terminological variation is related, but distinct from terminological synonymy: terminological variation “includes a wider range of phenomena where the same concepts can be expressed differently in text”. According to the scholar, term variants include inflected forms, graphical variants, synonymic and near-synonymic expressions, as well as contextual variants. However, in Cabré’s study (1999, p. 142), synonymy is understood in a much broader way. It is stated that synonymy can exist between “a standard form and any formal, orthographical, or phonological variant; an abbreviation and its full form; a term and its shortened form; a term and its scientific name, a term and the symbol representing it”. Warburton (2021, p. 54) also presents a broad notion of synonymy, distinguishing between two types: variant synonyms (abbreviations, acronyms, short forms, rearranged version of multiword terms, spelling



**Fig. 1** Management of synonymous terms in concept-oriented terminological resources

variants, etc.) and lexical synonyms (having “no similarity with the surface form of the term of which it is a synonym”). Thus, positioning of terminological designations of the same concept as synonyms or variants, which do not have the status of synonyms, as well as their inclusion or non-inclusion into terminology resources depends on the chosen approach to term synonymy and its limits. The present study follows a broad approach to synonymy, considering all linguistic designations of the same concept as terminological synonyms.

### *Studies of Lithuanian Synonymous Terms*

The phenomenon of synonymy has been discussed in fundamental works on Lithuanian terminology by Gaivenis (2002) and Keinys (2005). A comprehensive study of Lithuanian term synonymy, its conception, types and identification problems, has been published by Mitkevičienė (2015). The scholar presents an exhaustive typology of Lithuanian synonyms, as well as discusses various issues which have to be taken into consideration when collecting terms for terminological resources (Mitkevičienė, 2015).

Term synonymy in various domains has been investigated in the Lithuanian terminology works: informatics and computer science (Aukšoriūtė, 2018), construction (Stunžinas, 2005, 2010), architecture (Kitkauskienė, 2005), law (Umbrasas, 2004), philosophy (Mikelionienė, 2004), medicine (Zemlevičiūtė, 2001), to mention just a few.

Special attention is given to synonyms of diverse origins within the computer domain, where Lithuanian designations are competing with English borrowings (e.g. Aukšoriūtė, 2018). Synonyms of different origin in other domains are also explored extensively: e.g. Stunžinas (2005, 2010) investigated synonymous terms of different origin in the construction domain, and Zemlevičiūtė (2001) analysed synonyms of different origin in the medicine domain.

Thus, synonymy of concept designations encompasses various dimensions which have to be taken in consideration in terminology management work. Appearance of competing synonymous terms also depends on various factors, the decisive of which is preferences of users. Namely, these preferences determine which of the competing terms get established, which of them become obsolete, or acquire different development tendencies in different environments. Analysis of user preferences enables to understand and predict these tendencies.

### **Preparation for the Terminology Survey**

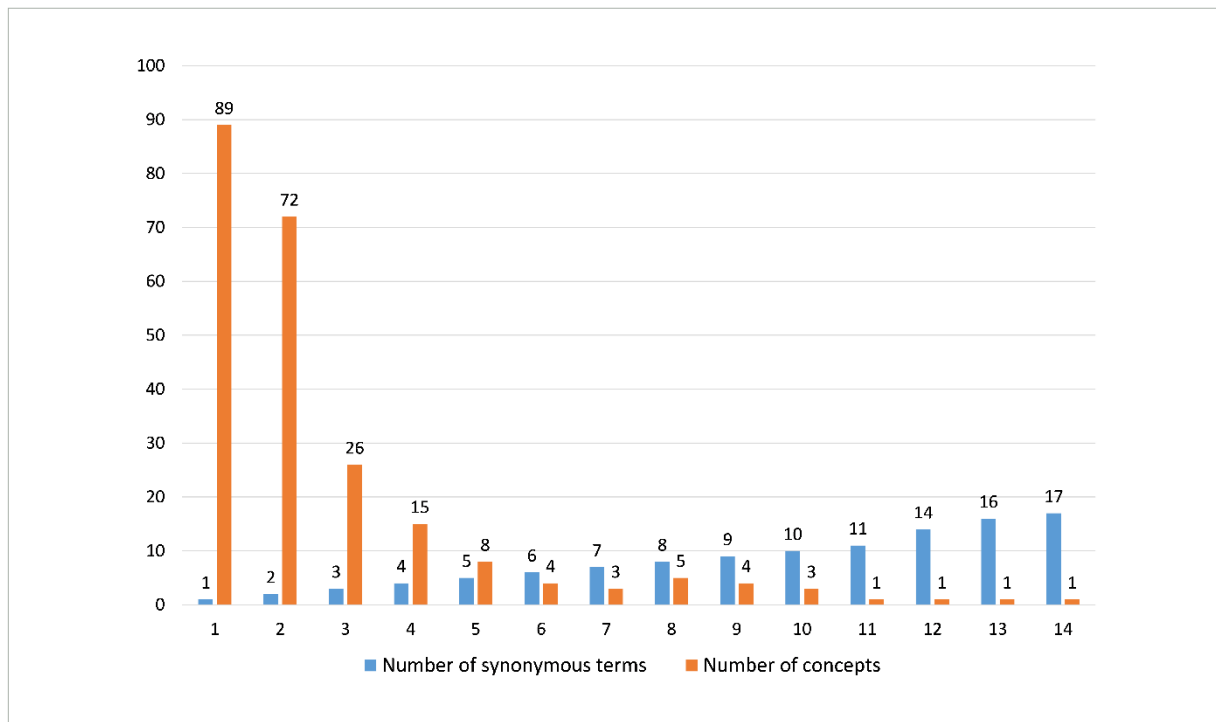
In order to select synonymous terms for their assessment by respondents of the terminology survey, the variation of cybersecurity terms in the Lithuanian-English Cybersecurity Termbase<sup>1</sup> was investigated (Rackevičienė et al., 2023). The terminological data for the termbase has been collected from the comparable and parallel cybersecurity corpora specifically developed for compiling the termbase. The corpora are freely accessible in the CLARIN-LT repository<sup>2</sup>.

The corpora have been composed of texts from various discourses: the parallel corpus has been composed mainly of the EU legislative acts and related documents (regulations, directives, communications, recommendations, etc.); meanwhile, the comparable corpus contains much more diverse texts produced in various discourses: legislative, administrative, informative, academic, and media (Utkā et al., 2022). The descriptive approach has been chosen for the compilation of the termbase and, following it, all synonymous terms detected in the corpora have been added to the termbase entries. The most frequent synonymous terms (up to 3) have been presented as the main designations of the concepts, other terms have been presented in section called “Other synonymous terms”.

In all, 233 concepts are included in the termbase. The term variation differs considerably depending on the concepts. 80% of the concepts have up to 3 synonymous designations. The number of synonymous designations of the remaining 20% differs from 4 to 17. Fig. 2 presents the number of concepts (columns in orange) which have the indicated number of synonymous terminological designations (columns in blue) in the

<sup>1</sup> Lithuanian-English Cybersecurity Termbase <https://www.terminologue.org/csterms/>; TBX file <https://clarin.vdu.lt/xmlui/handle/20.500.11821/55>.

<sup>2</sup> English-Lithuanian Parallel Cybersecurity Corpus – DVITAS <https://clarin.vdu.lt/xmlui/handle/20.500.11821/46>, English-Lithuanian Comparable Cybersecurity Corpus – DVITAS <https://clarin.vdu.lt/xmlui/handle/20.500.11821/47>.



**Fig. 2** Ratio of synonymous terms and concepts

termbase. The figure shows the continuum which ranges from 89 concepts having 1 designation to 1 concept having 17 designations.

10 concepts which have most diverse synonymous terminological designations in the cybersecurity termbase and corpora were selected for the survey in order to establish which term types are preferred by different user groups (see Table 1).

The Lithuanian synonymous terms designating the selected concepts differ in various aspects: the origin (native/foreign/hybrid), the denotation type (primary/figurative), the length and explicitness (the number of constituents and use/disuse of abbreviations), the lexical/morphological structure, etc. Considering the most evident variations in formation patterns, these terms can be classified into the following typology:

#### ◆ Borrowings

- Localised borrowings (words of foreign origin with added Lithuanian endings and spelling adapted to the Lithuanian pronunciation). Most of such borrowings come from the English language, e.g. *spamas* 'spam', *fišingas* 'phishing', *botnetas* 'botnet', *hakeris* 'hacker'. One of the selected multi-word terms contains a modifier of Greek origin: *kibernetinė ataka* (where *kibernetinis* is derived from *kibernetika* borrowed from Greek). However, its emergence and spread in the Lithuanian language is likely to be attributed to the influence of the English language, particularly of the English term *cyber attack*. The head *ataka* 'attack' is an old and fully localised borrowing from French.

**Table 1** Concepts selected for the terminology survey

Concept 1	'cyberattack'
Concept 2	'spam'
Concept 3	'denial-of-service attack'
Concept 4	'man-in-the-middle attack'
Concept 5	'brute force attack'
Concept 6	'phishing'
Concept 7	'botnet'
Concept 8	'hacker'
Concept 9	'honeypot method'
Concept 10	'zero-day vulnerability'

- Semi-localised borrowings (hybrid terms that along with fully localised heads include unlocalised modifiers which are English abbreviations or single/multi-words), e.g. *DoS ataka* ‘DoS attack’, *MitM ataka* ‘MitM attack’, „*man-in-the-middle*“ *ataka* ‘man-in-the-middle attack’, „*brute force*“ *ataka* ‘brute force attack’, „*phishing*“ *ataka* ‘phishing attack’.
- **Metaphorical calques** (literal translations of English metaphorical terms), e.g. „*žmogaus viduryje*“ *ataka* ‘man-in-the-middle attack’, *slaptažodžių žvejyba* ‘(literally) password fishing, (common EN term) phishing’, *medaus puodynės metodas* ‘honeypot method’, *nulinės dienos pažeidžiamumas* ‘zero-day vulnerability’.
- **Descriptive terms**
  - Terms which directly express the characteristics of the concepts. They often comprise constituents of the Lithuanian origin or fully localised borrowings, e.g. *nepageidaujami elektroniniai laiškai* ‘(literally) unsolicited electronic mail, (common EN term) spam’, *elektroninės paslaugos trikdyimo ataka* ‘(literally) electronic service disruption attack, (common EN term) DoS attack’, *slaptažodžių parinkimo ataka* ‘(literally) password selection attack, (common EN term) brute force attack’.
  - Terms which are transparent Lithuanian metaphors. They are created using Lithuanian metaphorical means (not calqued from English) and their meaning is readily apparent to Lithuanian language speakers: *tarpininko ataka* ‘(literally) middleman attack, (common EN term) man-in-the-middle attack’, *duomenų viliojimas* ‘(literally) data luring, (common EN term) phishing’, *masalo metodas* ‘(literally) bait method, (common EN term) honeypot method’.

Thus, the analysis of the termbase allowed selecting the concepts and synonymous terms for the survey and constructing the questionnaire in a way which would enable to reveal term preferences of respondents regarding the most apparent variations in the Lithuanian synonymy patterns within the cybersecurity terminology.

### Structure and Variables of the Terminology Survey

The conducted survey consisted of two main parts: 1) sociodemographic data collection part and 2) the main part with questions on term preferences and reasons driving the preferences (see Fig. 3).

The sociodemographic data collection part included the following variables:

- **Age:** respondents indicated their age as a numerical value.
- **Residence:** respondents selected a county in Lithuania or the option ‘Living outside Lithuania.’

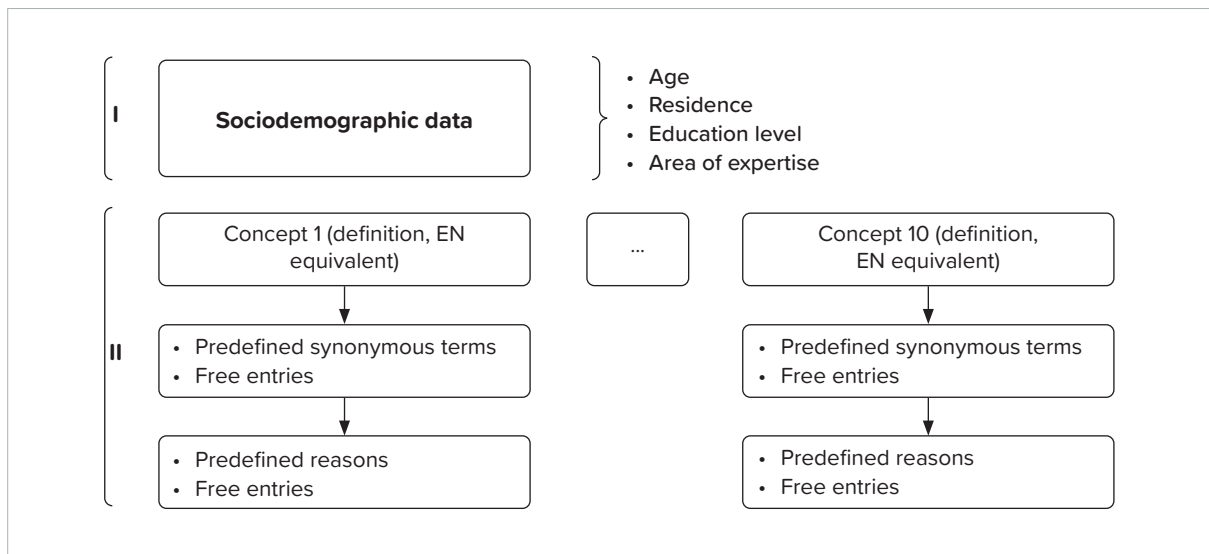


Fig. 3 Structure of the survey

- ◆ **Education level:** respondents selected one of the following options: Higher (Graduate of a university/college); Unfinished higher (Student of university/college); Upper secondary; Lower secondary, with an additional free-entry slot for other responses.
- ◆ **Area of expertise:** respondents selected one of the following options: Informatics/cybersecurity studies/work; Philology/translation studies/work; additional free-entry slot for other responses.

The main part was divided into 10 sections devoted for 10 cybersecurity concepts each. Each section included the following elements:

- ◆ **Concept definition:** A definition of the concept in Lithuanian.
- ◆ **Common English designation:** The most frequently used English term for the concept.
- ◆ **Question 1:** ‘Which term, in your opinion, is the most suitable to designate this concept?’  
Respondents were presented with several synonymous terms, alongside a free-entry slot allowing them to introduce their own proposals. Thus, the answers comprised a **categorical variable** (a term chosen from a list) and a **textual variable** (freely proposed term(s) provided in a free-entry slot labelled “Other”).
- ◆ **Question 2:** ‘Why have you chosen this term or proposed your own variant?’  
Respondents were offered several predefined reasons, in addition to a free-entry slot for personalised explanations. Thus, the answers comprised a **categorical variable** (one or several reasons chosen from the list) and a **textual variable** (freely indicated reason(s) in a free-entry slot labelled “Other”).

The paper focuses on the analysis of the categorical data collected in Question 1 and Question 2.

### Sociodemographic Data about the Respondents

The survey was sent to educational institutions (universities and colleges – faculties related to IT and cybersecurity studies, language and translation studies, as well as faculties of other study areas), state institutions responsible for regulation of electronic communication and IT application in various sectors, cybersecurity departments at the Ministry of Defence, the IT company association, IT companies, as well as state institutions and private companies of various other sectors.

We collected data from 593 respondents, who were segmented into diverse user groups based on four different variables:

- ◆ **Age of respondents:** The age of the participants spans from 17 to 69 years. 66% of the respondents fall within the age range of 17 to 29 years, while the remaining 34% are aged between 30 and 69 years.
- ◆ **Regions represented:** Respondents come from various regions across Lithuania, as well as beyond its borders. 71% of the respondents come from the Vilnius County, while 27% represent other Lithuanian counties: Kaunas (63 respondents), Telšiai (20), Panevėžys (18), Šiauliai (15), Klaipėda (13), Alytus (13), Utena (10), Marijampolė (6), Tauragė (3). Moreover, 2% of the respondents are emigrants, residing outside the boundaries of Lithuania.
- ◆ **Education level:** The data indicating respondents’ education allowed distinguishing two prominent respondent groups: university/college students (58.9%) and university/college graduates (41.1%).
- ◆ **Area of expertise:** The data pertaining to the respondents’ areas of expertise also served as the basis for categorising them into two distinct groups: experts (41%) and the general public (59%). The expert group comprises IT, cybersecurity and electronic communication students and IT professionals. On the other hand, the general public group comprises students and professionals engaged in a multitude of fields, including and related to language and translation, social sciences (law, political science, public administration, business management, economics, psychology, social work, education, communication, etc.), natural sciences (life sciences, physical sciences, medicine, agricultural sciences), and formal sciences (mathematics, civil engineering, financial technologies, etc.).

Thus, two data segmentations were formed to explore term preferences in different user groups (see Fig. 4).



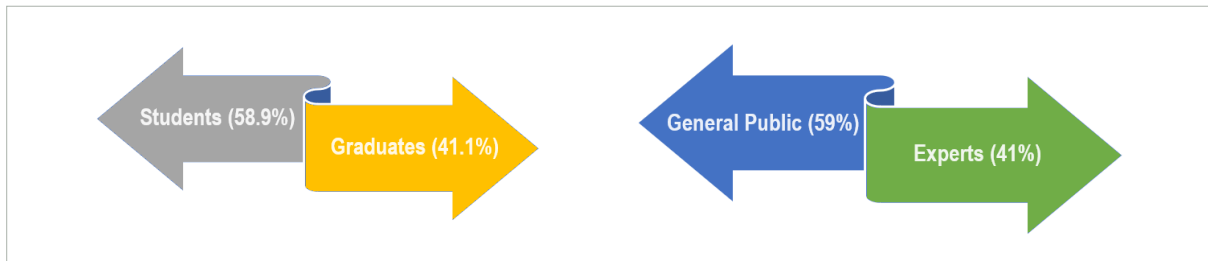


Fig. 4 Data segmentations and their proportions

## Survey Results

As explained in the section "Structure and Variables of the Terminology Survey", the respondents were instructed to choose the most preferable terms from lists of synonymous terms designating the same concepts (a total of 10 concepts) and to provide reasons for their selections. The outcomes of the first and the second questions are discussed separately in the following subsections, comparing respondent groups in two segmentations: Students vs. Graduates and General Public vs. Experts.

### Preferences of Synonymous Terms

#### *Established Terms vs. Competing Terms*

Some concepts have designations which are obvious leaders in all respondent groups, while designations of other concepts are still competing for dominant positions. The competing designations have similar scores, and their positions differ in different respondent groups. Based on the number of votes collected by the leading terms, the distribution is as follows: the leading term of one concept received 75% of the votes, the leading terms of six concepts received between 41% and 52% of the votes, and the leading terms of the remaining three concepts received between 29% and 35% of the votes. Some examples that illustrate these differences, are given below.

#### Example of an established term

For Concept 1 'cyberattack', the designation *kibernetinė ataka* was chosen by the absolute majority of respondents in all groups, surpassing other competing designations by a considerable margin (see Fig. 5).

In this case, a variety of modifiers was proposed in the survey, all selected from the cybersecurity corpora: *skaitmeninė ataka* 'digital attack', *kOMPIuterinė ataka* 'computer attack', *internetinė ataka* 'Internet attack',

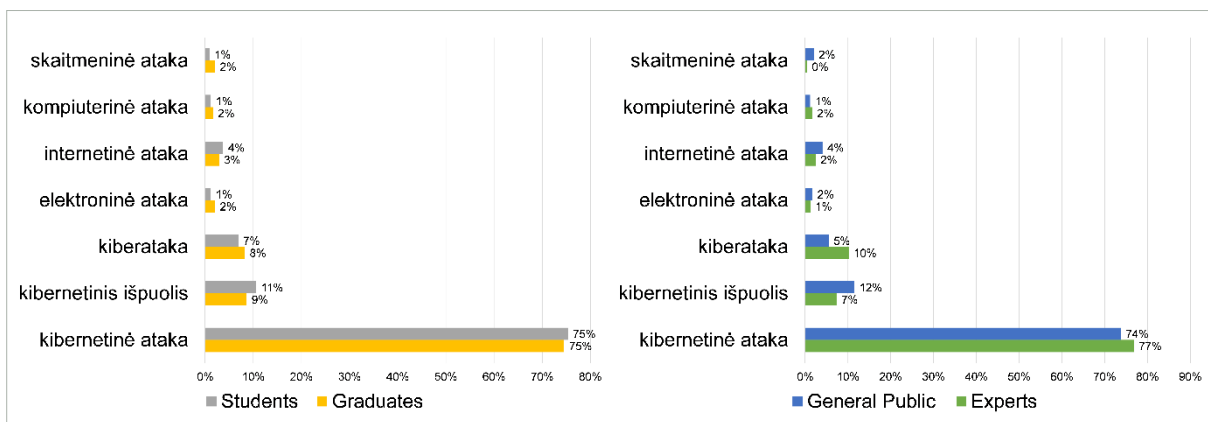


Fig. 5 Selection of terms for Concept 1 by different respondent groups



*elektroninė ataka* ‘electronic attack’. In addition, two variations, present in the corpora, of the term *kibernetinė ataka* were proposed: the designation with the modifier *kibernetinis* and the head of the Lithuanian origin *išpuolis* and the shortening *kiberataka* which is the literal translation of the English term, but whose formation is uncommon in Lithuanian (contains the lexical unit *kiber-*, which is the cutting of the root *kibernet-*).

The results of the survey presented in Fig. 5 show that the designations with the modifiers *kibernetinis*, *kiber-* dominate: *kibernetinė ataka* is the absolute leader collecting 75–77% of votes in different respondent groups. In the Student vs. Graduates segmentation, the second position is taken by *kibernetinis išpuolis* (9% and 11%, respectively), while in the General Public vs. Experts segmentation, the preferences in the second position differ: the General Public prefers the term *kibernetinis išpuolis* (11%) and Experts – the shortening *kiberataka* (10%). The results indicate that the modifiers *kibernet-*, *kiber-* have overwhelmingly surpassed other modifiers, and the term *kibernetinė ataka* can be considered as an established term among the respondents.

### Example of competing terms

Meanwhile, for Concept 6 ‘phishing’, the choices are much more scattered among the synonymous terms. This concept has numerous designations in the cybersecurity corpora, 10 of which were suggested for the respondents (see Fig. 6).

The suggested designations included the localised borrowing *fišingas*, the hybrid term “*phishing*” *ataka* composed of an original English modifier and a fully localised French head, the Lithuanian metaphorical calque *slaptažodžio žvejyba* formed following the formation principles of the English blending *phishing* (phishing = password + fishing). Additionally, various descriptive terms were suggested. They directly denote the characteristics of the concept, but point out their distinct aspects, particularly, the techniques of phishing (*duomenų klastojimas* ‘data forgery’, *klastočių ataka* ‘forgery attack’, *viliojimas į suklastotas internet svetaines* ‘luring to fake websites’; *elektroninis sukčiavimas* ‘electronic fraud’, *socialinės inžinerijos ataka* ‘social engineering attack’, *duomenų viliojimas* ‘data luring’) or the result of phishing (*duomenų vagystė* ‘data theft’).

The survey results reveal that that the term *fišingas* ‘phishing’ is the preferred designation for the majority of the respondents, gathering 28–32% votes across different respondent groups. However, several other synonyms also have collected a similar number of votes. In the Students vs. Graduates segmentation, *duomenų vagystė* ‘data theft’ claims the second position among Students (25%), while *duomenų viliojimas* ‘data luring’ secures the second position among Graduates (24%). The term „*phishing*” *ataka* ‘phishing attack’ takes the third position in both groups, collecting 11% and 10%, respectively.

In the General Public vs. Experts segmentation, among the General Public, the second position is claimed by *duomenų vagystė* ‘data theft’ (29%), leaving *duomenų viliojimas* ‘data luring’ and *elektroninis sukčiavimas* ‘electronic fraud’ in the third and fourth positions (18% and 8%). Among Experts, the second position is taken by

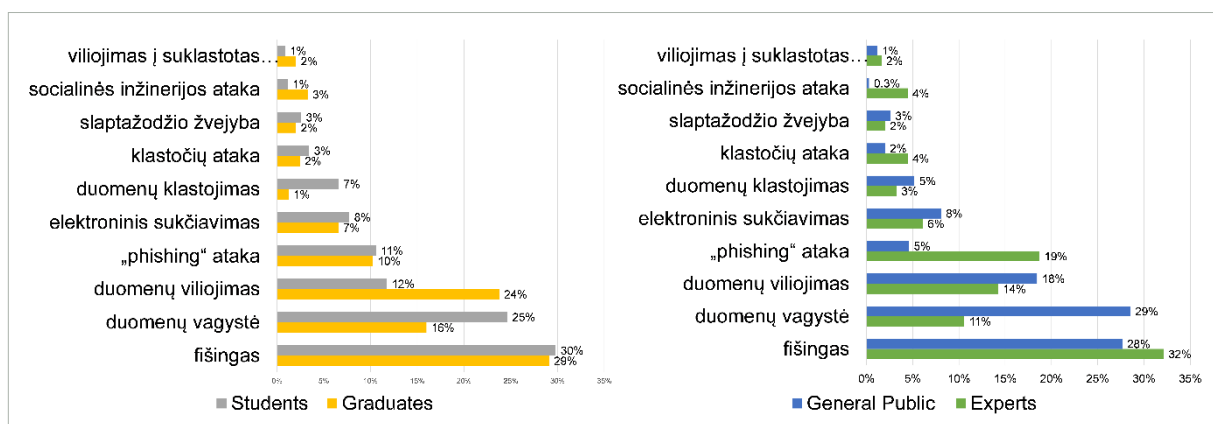


Fig. 6 Selection of terms for Concept 6 by different respondent groups

„phishing“ *ataka* ‘phishing attack’ (19%), leaving *duomenų viliojimas* ‘data luring’ and *duomenų vagystė* ‘data theft’ in the third and fourth positions (14% and 11%). Most other synonyms received a significantly lower number of votes, ranging from 1% to 8%. In most cases, the differences between respondent groups are minimal, with notable exceptions being *duomenų klastojimas* ‘data forgery’ (7% among students and 1% among graduates) and *socialinės inžinerijos ataka* ‘social engineering attack’ (4% among experts and 0.3% among general public). The presented results allow concluding that, though the term *fišingas* prevails across all respondent groups, other designations also enjoy popularity among respondents. However, the extent of their popularity varies across different groups. Thus, these designations are still competing for their positions within the Lithuanian terminology.

#### *The Preferred Term Types in Different Respondent Groups*

The data analysis uncovers distinct preferences for different term types across the concepts. In certain cases, localised or semi-localised borrowings are predominantly favoured in all respondent groups. However, for some concepts, term preferences vary among respondent groups: some groups predominantly prefer borrowings, while others opt for descriptive terms.

Based on the respondents’ preferences of term types, the concepts were grouped into three groups, each of which is discussed separately.

The 1<sup>st</sup> group consists of six concepts. Across these concepts, the majority of respondents in all groups prioritised borrowings, which include localised borrowings and those comprising unlocalised constituents. Many of these borrowings come from the English language or include English lexical units. Table 2 presents the designations of the concepts that collected most respondent votes.

**Table 2** Concepts and their leading designations

Concept	Leading designation in the overall dataset	Students	Graduates	General Public	Experts
Concept 1 ‘cyberattack’	<i>kibernetinė ataka</i> (75%)	75%	75%	74%	77%
Concept 2 ‘spam’	<i>spamas</i> (49%)	55%	39%	45%	54%
Concept 3 ‘denial-of-service attack’	<i>DoS ataka</i> (48%)	50%	44%	37%	64%
Concept 6 ‘phishing’	<i>fišingas</i> (30%)	30%	29%	28%	32%
Concept 7 ‘botnet’	<i>botnetas</i> (40%)	46%	31%	36%	44%
Concept 8 ‘hacker’	<i>hakeris</i> (52%)	63%	37%	52%	53%

The terms *spamas* ‘spam’, *fišingas* ‘phishing’, *botnetas* ‘botnet’, *hakeris* ‘hacker’ are localised English borrowings with added Lithuanian endings and, in two cases, spelling adapted to the Lithuanian pronunciation (*fišingas* and *hakeris*). The term *DoS ataka* ‘DoS attack’ is a hybrid composed of the unlocalised English abbreviation *DoS* and the fully localised borrowing from French *ataka*. The term *kibernetinė ataka* ‘cyberattack’ stands out as it does not incorporate elements of English origin (its constituents are of Greek and French origin). However, as it was mentioned in the section “Preparation for the Terminology Survey”, the emergence and spread of this term in the Lithuanian language is likely due to the influence of the English term of the same origin.

Four of these terms are leaders in all respondent groups (students and graduates, experts and general public): *kibernetinė ataka* ‘cyberattack’, *spamas* ‘spam’, *DoS ataka* ‘DoS attack’, *hakeris* ‘hacker’. The positions of other terms differ in different respondent groups. The term *fišingas* ‘phishing’ is the second priority by the General

Public whose most votes go to the descriptive term *duomenų vagystė* ‘data theft’. The designation *botnetas* ‘botnet’ is the second priority of Graduates whose most votes go to the descriptive term *užvaldytų kompiuterių tinklas* ‘network of compromised computers’.

The 2<sup>nd</sup> group comprises 3 concepts. The common English equivalents of these three concepts are metaphors: *brute force attack* (Concept 5), *zero-day vulnerability* (Concept 10) and *honeypot method* (Concept 9). While selecting the most appropriate terms for these concepts, Lithuanian respondents’ votes varied significantly across different groups:

- in the Students vs. Graduates segmentation, the majority of Students prioritise borrowings, while the majority of Graduates give priority to descriptive terms (see Fig. 7),
- in the General Public vs. Experts segmentation, the majority of respondents belonging to the General Public group give priority to descriptive terms, while the majority of Experts prioritise borrowings (see Fig. 8).

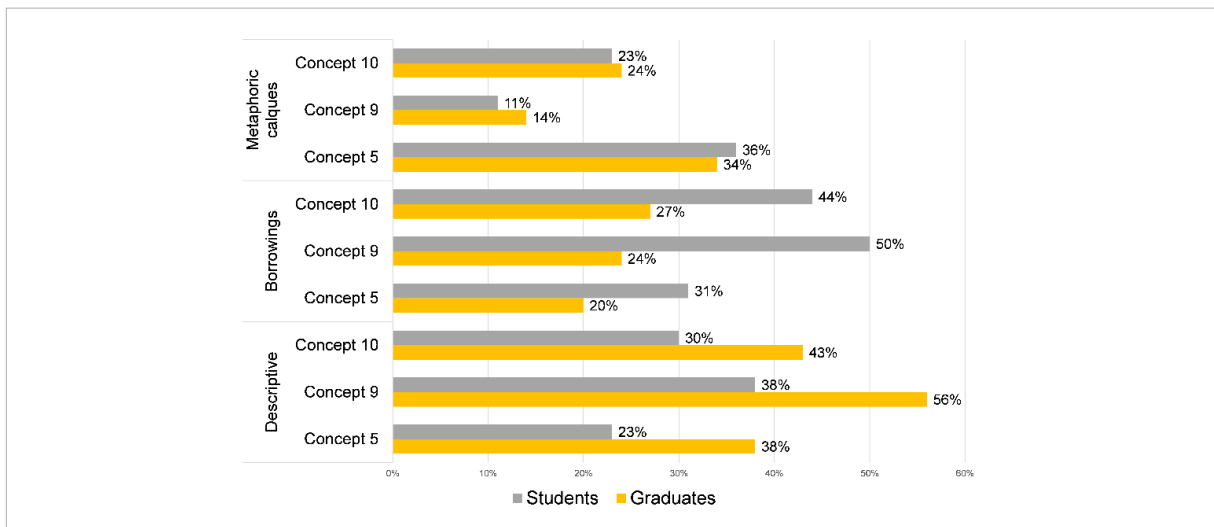


Fig. 7 Ratio of preferences of different term types between Students and Graduates

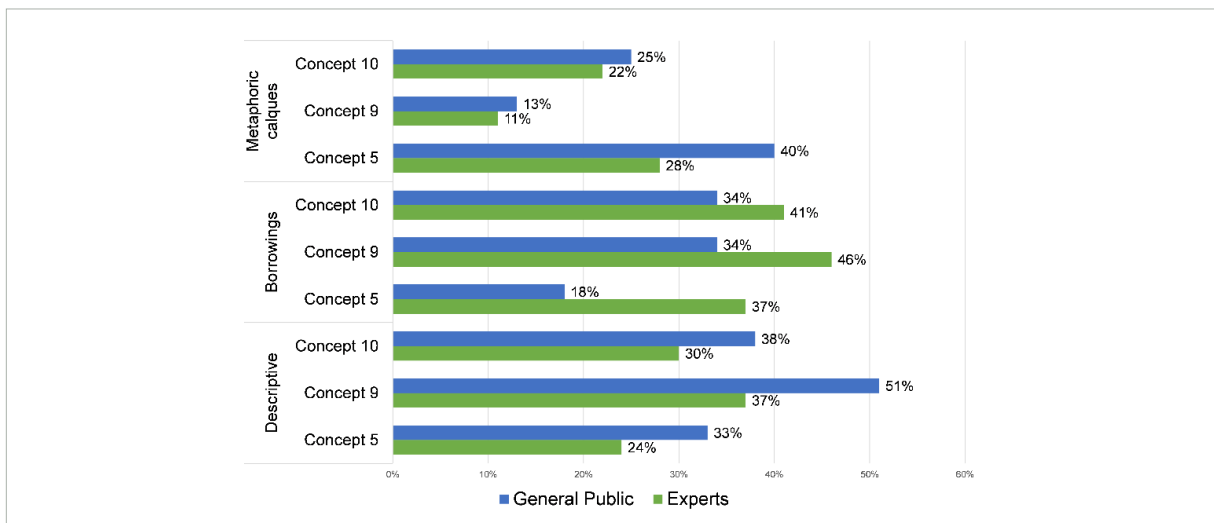


Fig. 8 Ratio of preferences of different term types between Experts and General Public

The data in the figures show clear differences between term preferences of different respondent groups. The majority of Graduates and the General Public prefer descriptive terms which include direct designations that explicitly express the characteristics of the concepts and a transparent metaphor that is created using Lithuanian metaphorical means and is easily discernible to Lithuanian language speakers. Direct designations are predominantly prioritised for Concept 5 (*slaptažodžių parinkimo ataka* 'password selection attack', common English equivalent is *brute force attack*) and Concept 10 (*ką tik nustatytas pažeidžiamumas* 'vulnerability that has just been identified', common English equivalent is *zero-day vulnerability*), while a transparent Lithuanian metaphor is predominantly prioritised for Concept 9 (*masalo metodas* 'bait method', common English term is *honeypot method*).

Meanwhile, the majority of Students and Experts give priority to hybrid terms which include original English constituents. The most popular designation for Concept 5 is „brute force“ *ataka* 'brute force attack', for Concept 9 – „honeypot“ *metodas* 'honeypot method', Concept 10 – „zero-day“ *pažeidžiamumas* / „zero-day“ *spraga* 'zero-day vulnerability'.

The data on metaphorical calques show that Concept 5 collected most votes among metaphorical calques of all three concepts. While the difference between the number of votes made by Students and Graduates is insignificant (36% and 34%, respectively), there is a considerable difference between the number of votes made by the General Public and Experts (40% and 28%, respectively). The metaphorical calques of Concept 5 include the following designations: *brutalios jėgos ataka* 'brute force attack' and *brutalioji ataka* 'brutal attack'. The metaphorical calques designating other two concepts were *medaus puodynės metodas* 'honeypot method' (Concept 9), *nulinės dienos pažeidžiamumas* / *nulinės dienos spraga* 'zero-day vulnerability' (Concept 10). The calque designations for Concept 10 were more popular than those for Concept 9; however, no significant differences between respondent groups were established.

The 3rd group is composed of designations of one concept: Concept 4, the common English equivalent of which is the metaphorical term *man-in-the middle attack*.

Across all respondent groups, the transparent Lithuanian metaphor *tarpininko ataka* 'middleman attack' was predominantly prioritised as the most suitable term for this Concept. The hybrids that include the original English abbreviation (*MitM ataka*) and the original English phrase („*man in the middle*“ *ataka*) take second positions in all respondent groups, while the metaphorical calque and its abbreviation („*žmogaus viduryje*“ *ataka*, *ŽV ataka*) remain in the last positions with considerably lower number of votes in all respondent groups (see Fig. 9). This is the only Concept for which the majority respondents prioritised a metaphorical term, and, notably, this metaphorical term was created using Lithuanian metaphorical means, rather than being calqued from English.

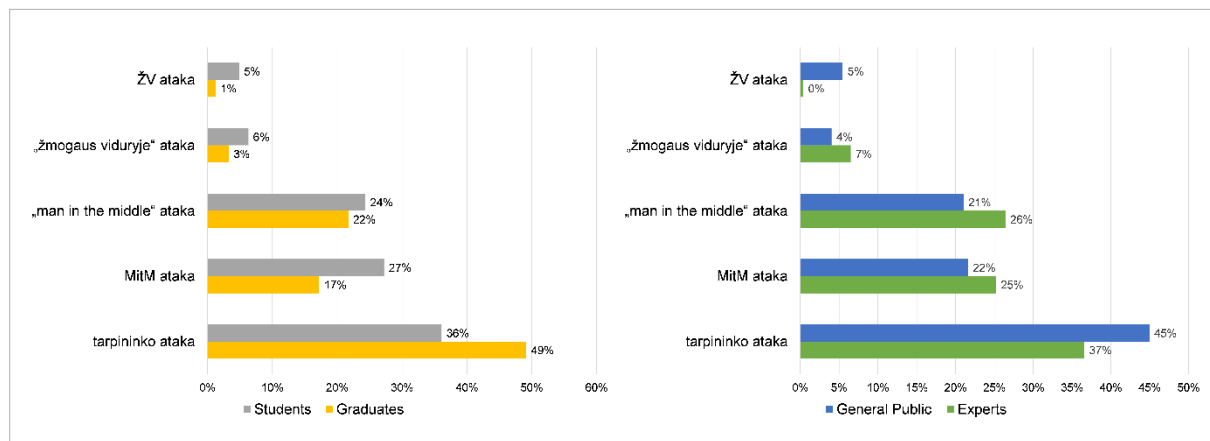


Fig. 9 Selection of terms for Concept 4 by different respondent groups

### Reasons for Choosing Terms

The second question in the survey was about the reasons that were important for respondents when they were selecting terms. Respondents had to choose from 5 predefined reasons, namely *Accuracy/Clarity*, *Brevity*, *Expressiveness*, *Correctness*, *Frequency* and *Other*. If they chose *Other*, they could also suggest their own reasons.

Fig. 10 presents the results for the two segmentations. Across all groups, most respondents indicated *Accuracy/Clarity* as the reason for their selection of a synonymous term. *Brevity* was chosen by the second-largest number of respondents, while *Expressiveness* received the fewest votes. A notable difference emerges within the Students vs. Graduates segmentation: whereas the Graduates' votes rank in the order *Accuracy/Clarity*, *Brevity*, *Correctness*, *Frequency*, the Students' votes reveal that *Frequency* is important to a higher number of students than *Correctness*.

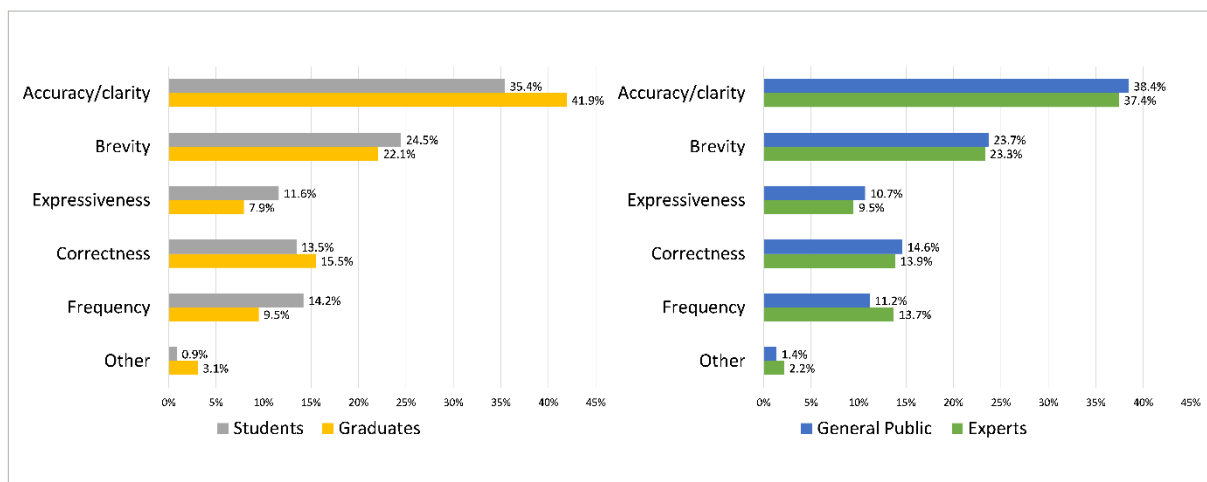


Fig. 10 Reasons for selecting terms by different respondent groups

As could be expected, Graduates and Experts suggested notably more additional reasons than Students and the General Public. The most frequently recurring additional reason given by respondents was the similarity of the term to the original English term. This reason, formulated in various ways, was indicated in free-entry slots for the designations of seven concepts. Notably, this reason was provided for the designations of Concept 1 *kibernetinė ataka*, *kiberataka* 'cyberattack' by eleven respondents. In their explanations, the respondents pointed out that their votes for Lithuanian terms were determined by their correspondence, equivalence or similarity to the English terms, literal translation of the terms, possibility to relate the Lithuanian terms with the English terms in Lithuanian texts, and the internationality of the terms.

- The conducted research allows drawing the following conclusions:
- Conclusions**
- 1 The analysis of the Lithuanian-English cybersecurity corpora and the Lithuanian-English cybersecurity termbase, based on the corpora dataset, proved that the Lithuania terminology of this domain is still very young and inconsistent. Most concepts have several Lithuanian designations which differ in origin and denotation type.
  - 2 The analysis of the terminology survey, incorporating data from 593 respondents of varying ages and expertise areas, showed that few concepts have designations that clearly lead in all respondent groups (e.g. *kibernetinė ataka* 'cyberattack' among all designations of Concept 1), while designations of most concepts

are still contending for their dominance and are prioritised differently by different respondent groups (e.g. the designations of Concept 6 'phishing': *duomenų vagystė* 'data theft', *duomenų viliojimas* 'data luring', „*phishing*“ *ataka* 'phishing attack', *elektroninis sukčiavimas* 'electronic fraud').

- 3 The types of terms (borrowing, metaphorical calque, descriptive) have an important impact on priorities of concept designations. For some concepts, localised or semi-localised borrowings are predominantly favoured in all respondent groups. However, for some concepts, term preferences vary among respondent groups: some groups predominantly prefer borrowings, while others opt for descriptive terms. The possible explanations might be that this distinction arises between concepts frequently discussed in the public space, with widely known designations, and those which are more specific and less commonly addressed in the public discourse.
- 4 In cases of the concepts which are well-known and their designations widely used, most respondents prioritise English borrowings, e.g. *spamas* 'spamas', *DoS ataka* 'DoS attack', *fišingas* 'phishing', *hakeris* 'hacker'. In cases of the concepts which are not widely known, differences between the respondent groups become apparent: most Students and Experts prioritise English borrowings („*brute force*“ *ataka* 'brute force attack', „*zero-day*“ *pažeidžiamumas* 'zero-day vulnerability'), while the majority of Graduates and the General Public give priority to descriptive terms (*slaptažodžių parinkimo ataka* 'password selection attack', *ką tik nustatytas pažeidžiamumas* 'vulnerability that has just been identified').
- 5 Metaphorical calques are less popular than other types of terms. However, some of them also collect a significant amount of votes (e.g. *brutalioji ataka* 'brutal attack', *brutalios jėgos ataka* 'brutal force attack') and compete with other types of terminological designations. Lithuanian metaphors, created using Lithuanian metaphorical means, not calqued from English, may also become popular if they are transparent and easily discernible to Lithuanian language speakers. In the survey dataset, one metaphor of this kind, namely *tarpininko ataka* 'middleman attack', received the highest number of votes among the synonymous terms designating Concept 4 'man-in-the middle attack' across all respondent groups.
- 6 The analysis of the reasons for selection of synonymous terms reveals a consistent pattern across all respondent groups, with most respondents indicating Accuracy/Clarity, followed by the second-largest number of respondents choosing Brevity, and Expressiveness being indicated by the lowest number of respondents. Notably, reasons indicated by Graduates can be ranked in order of prevalence as follows: Accuracy/Clarity, Brevity, Correctness, Frequency, and Expressiveness. Meanwhile, Students indicated that Frequency is more important reason than Correctness. Additionally, a frequently recurring reason mentioned by respondents in free-entry slots is the similarity of the term to the original English term.

### Conflict of Interest

The authors declare no conflict of interest regarding the publication of this article.

### References

- 1 Auksoriūtė, A. (2018). Informatikos ir kompiuterijos terminijos raida ir sinonimija. In Auksoriūtė, A., Gaivenytė-Butler, J., Labanauskienė, S., Mitkevičienė, A., Stunžinas, R., Umbrasas, A., Zemlevičiūtė, P., Lietuviškos informatikos ir kompiuterijos terminijos tyrimai (pp. 247-307). Vilnius: Lietuvių kalbos instituto leidykla.
- 2 Cabré M. T. (1999). Terminology. Theory, Methods and Applications. John Benjamins Publishing Company. <https://doi.org/10.1075/tlrp.1>
- 3 Gaivenis, K. (2002). Lietuvių terminologija: teorijos ir tvarkybos metmenys. Vilnius: Lietuvių kalbos instituto leidykla.
- 4 IATE User's Handbook: European Union (2023). IATE User's Handbook. <https://iate.europa.eu/assets/handbook.pdf>
- 5 ISO 704: International Organization for Standardization (2022). Terminology Work - Principles and Methods (ISO Standard No. 704:2022). <https://www.iso.org/standard/79077.html>

- 6 Keinys, St. (2005). Dabartinė lietuvių terminologija. Vilnius: Lietuvių kalbos instituto leidykla.
- 7 Kitkauskienė L. (2005). Architektūros terminų sinonimija. Lietuvių katalikų mokslo akademijos suvažiavimo darbai 19, 211-219.
- 8 L'Homme, M. C. (2020). Lexical Semantics for Terminology. An Introduction. John Benjamins Publishing Company. <https://doi.org/10.1075/tlrp.20>
- 9 Mikeliūnienė J. (2004). Dabartinės lietuvių filosofijos sinonimų ir darybos priemonių įvairovė. In Keinys, St. (Ed.), Terminologijos istorijos ir dabarties problemos (pp. 271-275). Vilnius: Lietuvių kalbos instituto leidykla.
- 10 Mitkevičienė, A. Kai kurie terminų sinonimijos aspektai. Terminologija 22, 39-60. <https://lki.lt/en/terminologija-22/>
- 11 Rackevičienė, S., Utkā, A. & Mockienė, L. (2023). Lithuanian-English Cybersecurity Termbase: Principles of Data Collection and Structuring. *Rasprave*, 2(49), 1-24. <https://doi.org/10.31724/rihjj.49.2.12>
- 12 Schmitz, K. D. (2015). Terms in Texts and the Challenge for Terminology Management. *Terminologija*, 22, 15-25. <https://lki.lt/en/terminologija-22/>
- 13 Stunžinas R. (2010). Gyvenamojo namo ir pagrindinių jo dalių liaudies terminijos sinonimija. *Terminologija*, 17, 130-142.
- 14 Stunžinas, R. (2005). Statybos terminijos sinonimija. *Terminologija*, 12, 7-26.
- 15 Umbrasas A. (2004). Bendrašakniai dvižodžių teisės terminų sinonimai ir variantai 1918-1940 m. Lietuvos kodeksuose. *Terminologija*, 11, 100-118
- 16 Utkā, A., Rackevičienė, S., Mockienė, L., Rokas, A., Laurinaitis, M., & Bielinskienė, A. (2022). Building of Parallel and Comparable Cybersecurity Corpora for Bilingual Terminology Extraction. Selected Papers from the CLARIN Annual Conference 2021, edited by Monica Monachini and Maria Eskevich, 126-138. <https://doi.org/10.3384/ecp18912>
- 17 Warburton, K. (2021). The Corporate Terminologist. John Benjamins Publishing Company. <https://doi.org/10.1075/tlrp.21>
- 18 Zemlevičiūtė P. (2001). Sinoniminiai "Medicinos enciklopedijos" terminai. *Terminologija*, 8, 34-49.

**Sigita Rackevičienė, Andrius Utkā**

## Santrauka

**Lietuvių kalbos kibernetinio saugumo sinoniminių terminų pasirinkimai skirtingose vartotojų grupėse**

Kibernetinis saugumas yra sparčiai besivystanti sritis, kurioje naujas sąvokas įvardijantys terminai paprastai pirmiausia įsitvirtina anglų kalboje ir tik vėliau jų atitikmenys sukuriama kitose kalbose. Formuojantis kibernetinio saugumo terminijai lietuvių kalboje, vartosenoje atsiranda daug sinoniminių terminų, įvardijančių tas pačias sąvokas. Šiuos sinoniminiuos terminus įvairios vartotojų grupės traktuoja ir vartoja skirtingai. Straipsnyje pristatoma apklausa, kurioje dalyvavo 593 respondentai iš įvairių amžiaus grupių, skirtingų regionų ir skirtingų profesinių sričių. Apklausoje respondentai turėjo pasirinkti, jų manymu, tinkamiausius terminus, įvardijančius 10 kibernetinio saugumo sąvokų: respondentai galėjo pasirinkti anketoje siūlomus terminus arba pasiūlyti savo terminus ir nurodyti savo pasirinkimų priežastis. Sąvokos ir jas įvardijantys terminai atrinkti iš lietuvių-anglų kalbų kibernetinio saugumo terminų bazės, kuri remiasi duomenimis, surinktais iš anglų-lietuvių kalbų lygiagrečiojo ir palyginamojo kibernetinio saugumo tekstynų. Kiekybinė ir kokybinė apklausos rezultatų analizė parodė, kaip respondentai vertina skirtingų tipų terminus, pvz., skolinius, metaforinius ir aprašomuosius terminus, ir kokiems terminų tipams atiduoda pirmenybę skirtingos respondentų grupės dviejuose segmentuose: studentų ir absolventų bei plačiosios visuomenės ir kibernetinio saugumo ekspertų. Apklausoje rezultatai parodė, kad nors kai kurie kibernetinio saugumo terminai lietuvių kalboje jau yra nusistovėję, dauguma jų vis dar tarpusavyje konkuruoja. Terminų pasirinkimo priežasčių analizė atskleidė, kad tikslumas ir aiškumas yra pagrindiniai termino parinkimo veiksniai. Tikimasi, kad šis tyrimas prisidės prie kibernetinio saugumo terminų standartizavimo Lietuvoje ir suteiks įžvalgų apie vartotojų pasirinkimus ir jų priežastis.



---

## About the Authors

### **SIGITA RACKEVIČIENĖ**

Professor at the Institute of Humanities at the Faculty of Human and Social Studies, Mykolas Romeris university, Vilnius, Lithuania

#### **Research interests**

Multilingual terminology (its conceptual, linguistic and pragmatic dimensions), terminography, corpus linguistics, computational linguistics

#### **Address**

Institute of Humanities, Faculty of Human and Social Studies, Mykolas Romeris University, Ateities st. 20, LT-08303 Vilnius, Lithuania

#### **E-mail**

sigita.rackeviciene@mruni.eu

#### **Orcid ID**

0000-0001-5794-0296

### **ANDRIUS UTKA**

Associate Professor at the department of Lithuanian studies and a senior researcher at the Institute of Digital Resources and Interdisciplinary Research (SITTI), Vytautas Magnus university, Kaunas, Lithuania

#### **Research interests**

Statistical text analysis, language resources, computer-assisted translation, automatic summarisation, terminology, terminology extraction, and the language of disinformation

#### **Address**

Institute of Digital Resources and Interdisciplinary Research, Vytautas Magnus University, V. Putvinskio st. 23-216, LT-44243 Kaunas, Lithuania

#### **E-mail**

andrius.utka@vdu.lt

#### **Orcid ID**

0000-0001-5212-4310

