

## A CORPUS-BASED APPROACH TO IDENTIFYING TOP-LEVEL TERMS FOR DEVELOPING A SKOS VOCABULARY FOR UKRAINIAN EPIGRAPHY

**Tamrazyan Hamest,**

Associate professor, PhD,

Department of Germanic Languages, Faculty of Foreign Languages,

Dragomanov Ukrainian State University

ORCID ID: 0000-0001-9630-0195

Scopus ID: 5914771920

Scientific collaborator, Collège des humanités,

École polytechnique fédérale de Lausanne

INN 140 (Bâtiment INN), Station 14, CH-1015,

Lausanne, Switzerland

*This study aims to identify and classify top-level terms in Ukrainian epigraphy to develop a SKOS vocabulary and facilitate the categorization, organization, and retrieval of epigraphic inscriptions. This research fills a significant gap in the digital humanities, where Ukrainian epigraphic heritage has been underrepresented. The foundation for describing epigraphic artifacts in Ukrainian academia is the «Corpus of Inscriptions of Saint Sophia Cathedral Kyiv» by V. Korniyenko, which represents a structured format of Ukrainian epigraphic heritage. The study is based on comparative methods and detailed corpus-based linguistic analysis of the contextual application of terms in academic discourse. The use of the "Corpus of Inscriptions of Saint Sophia Cathedral Kyiv" is a justified choice due to its structured format and systematic approach to the description of epigraphic artefacts. An important element of the study is the analysis of the structure and content of Korniyenko's works for developing a standardized SKOS vocabulary for Ukrainian epigraphic inscriptions. The research proposes a systematic approach to developing a SKOS vocabulary for Ukrainian epigraphy, integrated with existing frameworks such as the EAGLE vocabulary for Greco-Roman artifacts. This is not only a technical step but also a strategic one, ensuring broader applicability and interoperability, allowing Ukrainian inscriptions to be studied alongside those of other cultures. The findings contribute to a more integrated and accessible digital representation of epigraphic heritage, enriching the global digital humanities landscape and ensuring proper attention and scholarly recognition of the rich epigraphic heritage of Ukraine. The importance of digital tools and corpus analysis for the development of digital humanities, particularly digital epigraphy, is emphasized in this study. Developing a comprehensive SKOS vocabulary for Ukrainian epigraphy will enable the integration of these vocabularies with existing frameworks such as the EAGLE vocabulary for Greco-Roman artifacts, ensuring broader applicability and interoperability, allowing Ukrainian inscriptions to be studied alongside those of other cultures and enhancing cross-cultural and cross-linguistic research.*

**Key words:** SKOS, digital epigraphic heritage, EAGLE vocabulary, Ukrainian epigraphy

### **ТАМРАЗЯН Амест Сержиківна. Корпусний підхід до визначення термінів верхнього рівня для створення SKOS словника української епіграфіки**

*Метою дослідження є ідентифікація та класифікація термінів верхнього рівня в українській епіграфіці для розробки словника SKOS, що сприятиме категоризації, організації та пошуку епіграфічних написів. Дане дослідження заповнює значну прогалину в цифрових гуманітарних науках, де українська епіграфічна спадщина була недостатньо представлена. Основою для опису епіграфічних артефактів в українській академії обрано «Корпус графімі Софії Київської» В. Карнієнка, який представляє структурований формат української епіграфічної спадщини. Дослідження ґрунтується на порівняльних методах та детальному корпусному лінгвістичному аналізі контекстного застосування термінів у науковому дискурсі. Корпус графімі Софії Київської є обґрунтованим вибором завдяки його структурованості та систематичному підходу до опису епіграфічних пам'яток. Важливим елементом дослідження є аналіз структури та змісту робіт Карнієнка для розвитку стандартизованого словника SKOS для українських епіграфічних написів. Дослідження пропонує систематизований підхід до розвитку словника SKOS для української епіграфіки, інтегрованого з існуючими рамками, такими як словник EAGLE для греко-римських артефактів. Це не лише технічний, а й стратегічний крок, який забезпечує ширшу застосовність та інтероперабельність, дозволяючи вивчати українські написи поряд із написами інших культур. Отримані результати сприяють більш інтегрованому та доступному цифровому уявленню епіграфічної спадщини, що не лише збагачує світовий ландшафт цифрових гуманітарних наук, а й забезпечує належну увагу та наукове визнання багатой епіграфічної спадщини України. Важливість цифрових інструментів та корпусного аналізу розвитку цифрових гуманітарних наук, зокрема цифрової епіграфіки, наголошується у цьому дослідженні. Розробка комплексного словника SKOS для української епіграфіки дозволить інтегрувати ці словники з існуючими рамками, такими як словник EAGLE для греко-римських артефактів, що забезпечить ширшу застосовність та інтероперабельність, дозволяючи українським написам вивчати поряд з написами між мовами та культурами.*

**Ключові слова:** SKOS, цифрова епіграфічна спадщина, лексика EAGLE, українська епіграфіка

**Introduction.** Ukrainian epigraphy, with its decades-long academic traditions and centuries-long history, boasts a well-developed terminological system to describe its primary objects of study. However, despite its rich

and invaluable collections, this field is underrepresented in the digital domain and has not received the attention it deserves from global digital epigraphic, archaeological, and historical communities [1, p. 8–14; 2, p. 41]. To

decipher and share these assets with the global academic community, it is essential to make these collections Linked Open Data (LOD) [3; 4, p. 57–61]. LOD is a way of publishing data on the web so that different pieces of information can be connected to each other. This makes it easier for people and computers to find, share, and use data across different websites and platforms. LOD principles ensure that data is interconnected and can be used across different systems and platforms, facilitating a richer and more integrated understanding of the data. The Simple Knowledge Organization System (SKOS) [14] plays a vital role in this by providing a framework for organizing knowledge in a way that is both machine-readable and human-understandable [4, p. 1–55; 5, p. 611–630]. SKOS is a standard used to organize knowledge in a structured way. It allows concepts to be represented in a machine-readable format, making it easier to share, link, and integrate information across different systems and platforms.

Our research strives to develop a SKOS vocabulary specifically dedicated to the Ukrainian epigraphic collections to facilitate the categorization, organization, and retrieval of epigraphic inscriptions [15]. The initial step in this process is the compilation of a sample structured vocabulary. This vocabulary, which includes a diverse range of terms and concepts related to Ukrainian epigraphic inscriptions, serves as a foundation for developing a comprehensive SKOS framework. This sample vocabulary can then be expanded and enriched over time, accommodating the evolving needs of the epigraphic domain and integrating feedback from ongoing research and usage.

The objectives of our research are multifaceted and aim to address several key areas. Firstly, the project seeks to enhance digital epigraphy by developing comprehensive SKOS vocabularies for Ukrainian epigraphic inscriptions. This effort fills a significant gap in the digital humanities where Ukrainian epigraphic heritage has been underrepresented. By creating these vocabularies, the research aims to integrate them with existing frameworks, such as the EAGLE vocabulary [16] for Greco-Roman artifacts. The integration with the EAGLE vocabulary is not only a technical step but also a strategic one, as it ensures broader applicability and interoperability, allowing Ukrainian inscriptions to be studied alongside those of other cultures and enhancing cross-cultural and cross-linguistic research.

*The aim of this research* is to identify a set of category labeling terms within Ukrainian epigraphic terminology, analyze the key terms that denote the main subjects and categories in Ukrainian epigraphy, and compare these terms with their English counterparts found in the EAGLE vocabularies. This will help us identify the top elements of our SKOS-based vocabulary and establish the relationships between these terms and the conceptual schema.

This study involves a detailed examination of terms in Ukrainian, focusing on their usage patterns within our corpus. The detailed analysis of the frequency, context, and variations of these terms will help to comprehend their roles and nuances in Ukrainian academic discourse and their overall significance in the field. Moreover, it will clarify how equivalent concepts are labeled in EAGLE

vocabularies. Additionally, we seek to understand how linguistic and cultural contexts shape the usage of these terms in epigraphic studies and identify the reasons behind the preference for certain terms over others. By examining these aspects, we aim to uncover the linguistic, cultural, and contextual factors that influence terminology choices in Ukrainian epigraphy.

The insights gained from this analysis will enable us to propose clear usage patterns for the selected terms, thereby reducing ambiguity and enhancing the clarity and coherence of epigraphic discourse. We also aim to underscore the importance of digital tools and corpus-based analysis, demonstrating their effectiveness and potential in digital humanities and specifically in digital epigraphy.

**Materials and methods.** For our research, we investigate the terms used as labels for referring to the main objects of epigraphic collections in both digital and non-digital domains in Ukrainian and English academia. To fully understand the peculiarities of the usage patterns of these terms in academic discourse, we employ both comparative methods and detailed corpus-based linguistic analysis of their contextual application.

As the basis for describing epigraphic artifacts in Ukrainian academia, we have chosen the «Корпус графіті Софії Київської» by V. Korniyenko, published across 12 books [6; 7; 8; 9; 10; 11; 12]. Our decision to use Korniyenko's work is based on several key factors:

**Structured Format:** The series represents one of the few published versions of Ukrainian epigraphic heritage in a structured format in Ukrainian. This makes it an invaluable resource for our work.

**Systematic Approach:** Korniyenko's detailed and systematic approach to medieval epigraphy, his use of the Ukrainian language, comprehensive methodology, clear classification system, focus on authenticity, and inclusion of new findings provide a robust foundation for our research.

**Linguistic and Cultural Context:** Previous works were mainly published in Russian. Korniyenko's corpus, being published in Ukrainian, aligns better with the linguistic and cultural context of the inscriptions and our goal of compiling an epigraphic vocabulary in Ukrainian.

By analyzing the structure and content of Korniyenko's work, we aim to develop a comprehensive understanding of the classification principles used in Ukrainian epigraphy and how they compare to those in the EAGLE vocabulary. This analysis will provide the foundation for developing a standardized SKOS vocabulary for Ukrainian epigraphic inscriptions.

*The language data* for our research is the specialized sample corpus containing Ukrainian files specifically related to epigraphic inscriptions. Our corpus, the "Ukrainian Epigraphic Corpus," consists of two sub corpora: «Web Epigraphy» and «Academic.» The sizes of these sub corpora are as follows: Academic publications: 1,038,140 tokens (74.751%), Web Epigraphy: 350,654 tokens (25.249%)

A significant portion of the Academic sub corpus was derived from scholarly books and monographs dedicated to the study of epigraphy and dictionary and encyclopedia

entries. These texts provided in-depth analyses, historical context, and detailed descriptions of inscriptions. The sub corpus consists of 18 files, including 9 books/monographs, 5 articles, 1 report, and 3 dictionary or encyclopedia entries. Most of the monographs presented published versions of Ukrainian epigraphic heritage in a structured format. Articles from academic journals offered current and specialized insights into specific aspects of epigraphy. These publications often included the latest research findings, methodologies, and theoretical discussions. Various scientific publications and conference proceedings provided additional valuable texts, often including papers presented at epigraphic and archaeological conferences, offering insights into current trends, challenges, and innovations in the field.

The web sub corpus was compiled with the help of Sketch Engine [17], employing custom tuples like «епіграфіка графіті корпус» to systematically scrape the web for relevant texts. The goal was to survey the range of epigraphic terms circulating on the internet. The rationale behind the choice of these terms is as follows: First, the term *епіграфіка* represents the academic discipline of interest. Second, the term *графіті* is the most frequently used term in published editions dedicated to Ukrainian epigraphic heritage. This term represents the object of research in the epigraphic discipline. Third, the term *корпус* was selected to include not only scientific articles but also special published editions representing the corpus of Ukrainian epigraphy.

As a result, the compiled corpus comprises academic publications, including articles, dissertations, and web resources dedicated primarily to Ukrainian epigraphic heritage, providing a rich dataset for analysis. As expected, the majority of the resources included in our corpus turned out to be primarily academic publications, providing valuable insights into the academic discourse surrounding these terms within the field of epigraphy. A key limitation of our sample corpus is the substantial inclusion of texts by Korniyenko V. V. on the epigraphic heritage of Ukraine. This led to an over-representation of what could be termed «the personal academic vocabulary» of a single author. However, it is noteworthy that Korniyenko V. V. acknowledges the collaborative nature of the corpus on Sofiiski graffiti, particularly under the guidance of Dr. N. Nikitenko, who contributed significantly to the study of epigraphic monuments from the old Russian period [6, p. 6].

Relevant websites and digital archives were also scoured for texts, with 96% of the texts extracted from the wordpress.com website. These online resources included digital repositories of epigraphic data, scholarly articles, and project websites that document ongoing epigraphic research. Various scientific publications and conference proceedings provided additional valuable texts, often including papers presented at epigraphic and archaeological conferences, offering insights into current trends, challenges, and innovations in the field.

The selected sources contributed to the comprehensive understanding of epigraphic studies in several ways. Firstly, the inclusion of diverse types of texts ensured that

the corpus covered a wide range of topics, from detailed descriptions of individual inscriptions to broader theoretical discussions about epigraphy as a discipline. By integrating sources that provided historical, cultural, and linguistic contexts, the corpus offered a multidimensional view of epigraphic studies. This richness was crucial for developing a vocabulary that accurately reflects the complexities of Ukrainian inscriptions. Additionally, the use of reputable scholarly works, especially those by recognized experts like Professor Korniyenko, Nikitenko, and others, added credibility and depth to the corpus. These authoritative sources ensured that the extracted terms and concepts were grounded in rigorous scholarship.

**Discussion.** In the first digital corpus, the broader term used is *Hanus*. According to the Ukrainian literary encyclopedia, "Напис" encompasses various forms of inscriptions differentiated by their execution technique, such as carved inscriptions (інскульпт), painted inscriptions (дипінті), and scratched inscriptions (графіті). This implies that *Hanus* encompasses various forms of inscriptions differentiated by their execution technique. However, in non-digital collections, we observed that *Hanus* is not used as a labeling term; rather, it is used synonymously with графіті.

In Ukrainian academia, digital and non-digital epigraphic artifacts are differentiated at the top level according to their execution technique. As seen in the Greco-Roman digital collections and those that use EAGLE vocabulary, they generally do not differentiate inscriptions according to the technique at the top level. They use the term inscription with a unique identifier, providing information about the execution technique later in the file dedicated to each inscription using the terms collected in the EAGLE Execution Technique vocabulary.

In all the resources where we have a structured corpus of inscriptions in Ukrainian, *графіті* along with *інскульпт* and *дипінті* are the main terms to denote each item in the corpus, usually followed by a unique number assigned to that particular inscription. Moreover, in Ukrainian academic discourse, these three terms are used as the main objects of any epigraphic collections as three types of *епіграфічних пам'яток*.

Understanding linguistic, cultural, and contextual factors is crucial as they shape how epigraphic terms are classified and interpreted within local scholarship. These factors also influence the integration of such terms into global frameworks like EAGLE, ensuring their accurate representation and cross-cultural comparison.

**Results.** *Comparative analysis of the terms графіті, інскульпт, and дипінті in non-digital publications:* As mentioned above, we have chosen for describing epigraphic artifacts in Ukrainian academia, we have chosen the «Корпус графіті Софії Київської» (The Corpus of Inscriptions of Saint Sophia Cathedral, Kyiv) by V. Korniyenko, published across 12 books [6; 7; 8; 9; 10; 11; 12].

To fully understand the hierarchical structure of the future SKOS vocabulary, it was essential to comprehend how Korniyenko categorizes inscriptions at the top level. He states: «Графіті поділяються на дві великі групи – написи

та малюнки а кожна з них розбивається за зовнішніми характеристиками графіті на менші категорії» [10, p. 9]. This indicates that Korniyenko first differentiates graffiti based on their form into *inscriptions* (написи) and *drawings* (малюнки). However, this differentiation is not reflected in the top-level labels. Both groups of inscriptions are then further subdivided into smaller categories based on external characteristics and content. All these categories are reflected in Категорія, demonstrating a hierarchical structure where Графіті is a top-level label while Категорія includes different types of написи та малюнки according to their content and form. For написи-графіті, the categories include молитовні, пам'ятні, написи календарного змісту, окремі слова та словосполучення, абетки склади та окремі літери and others. For малюнки-графіті, the categories include хрести, монограми та теонімограми, зображення святих, символи пов'язані з частинами людського тіла, предмети релігійного життя та артефакти, and others.

As the focus of our study was the set of terms used to label the main objects of epigraphic collections, we specifically concentrated on the section that includes the unique number given to each artifact in the collection:

«Графіті № 2097 (табл. IV 3) На поверхні стіни прокреслений напис, що складається з чотирьох літер. Дві перші впевнено визначаються як АБ, дві останні менш впевнено як ОВ. Датування: За формою написання в з округлою петлею графіті датується в межах другої чверті XI – кінця XII ст. Категорія: абетки склади та окремі літери» [6, p. 10].

As observed in the structured corpora, each entry is given a unique identifier. To include only the terms used for labeling each item in the collection, we restricted our search to [word="№"] [word="\d+"]. Later, we examined the most frequent collocates preceding this expression, which were in upper case, as the labels of the categories are written in upper case across the published corpora. Out of 53 collocates, we chose five common nouns in singular in uppercase – **Графіті, Інскульпт, Напис, Запис, and Дипінті**.

The most frequently used collocate is the term **Графіті**, which is also the most evenly distributed term across the whole corpus, as you can observe in Figure 1.

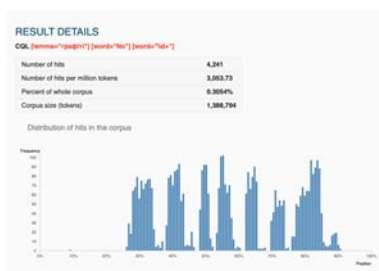


Fig. 1. Quantitative data for the term **Графіті**

Examining the context of usage of the term **Напис**, we found that it is not a label for the entry of the collection. Though it is used to refer to specific inscriptions with unique identifiers, they are used in the text to refer to the inscription labeled as Графіті No:

<Графіті No 140 (табл. XXX 4) Під час публікації напису No 365 С. Висоцький вказав малюнок на фотографії та прорисовці [Висоцький 1985 табл. XXXIX 1 2] однак у коментарі зовсім його не згадував. По личкувальному розчину прокреслений простий чотириконтинентний хрест. 42 Розділ I. >

(Token number 877118, Document number 409, doc. FT Published corpus File ID file 32708432, File name Korpus\_hrafiti\_tserkvy\_Spasa\_na\_Berestovi.pdf)

We observe a similar situation with the term **Запис**:

< Графіті No 1362б (табл. LXXXI) Праворуч від шостого рядка запису No 1362 при повторному дослідженні було виявлено дві невеликі літери І та Ї. Тиньк за ними пошкоджений, тож вірогідно тут знаходились інші літери, а все слово варто реконструювати як ії[іізè] 'поможі'. Датування: За формою написання І з виступом зверху графіті >

(Token number 692528, Document number 409, doc. FT Published corpus File ID file 32708431, File name Korpus\_hrafiti\_Sofii\_Kyivskoi\_XI\_pochatok\_XVIII\_st\_Chastyna\_II.pdf)

Korniyenko in his collection of Sofiiski graffiti [6, p. 12], mentions that he uses the terms написи and графіті as synonymous terms to avoid the repetition of the same term. During our research, we established that the terms *графіті*, *напис*, and *запис* referring to the epigraphic inscriptions share common collocates including numerals and functional words which show some overlap in their usage contexts. Yet, each term's unique collocates direct to different semantic fields or usage contexts. Each term, through its specific usage and collocates, highlights distinct facets of epigraphic work, ranging from the documentation of ancient writings to the organization of numerical data and the academic discourse surrounding these inscriptions. The term *графіті* is frequently associated with the tangible part of the inscription characterized by its historical and archaeological significance. *Напис* typically accentuates the textual composition and structure of the inscription, whereas *запис* is often associated with the role of inscriptions as documented records, incorporating numerical information or historical records [13, p. 92–110].

The term **Інскульпт** occurs only 8 times in our corpus as a labeling term and is not evenly distributed (Figure 2). To fully understand the usage pattern of the term, we investigated its usage across our corpus without restricting our search.

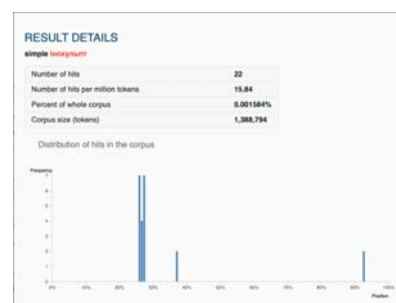


Fig. 2. Quantitative data for the term **Інскульпт**

Focusing on *інскульпт*, we observe that it occurs 22 times in our corpus with uneven distribution. This term is used in specific contexts often referring to carved or engraved inscriptions. Its top collocates include terms such as *№, табл., написи, вирізані, and вигравійовані*. In 8 out of 22 cases, *інскульпт* serves as a category label containing information about the ID of each inscription, emphasizing its role in cataloguing and classification. In 4 cases, it is used to refer to a specific entry in the collection. For example, in Korniienko's work, one of the labels reads:

*<або В. Датування. Середина XII – третя чверть XVI ст.; за формою прорізів. Категорія. Абетки склади та окремі літери. Інскульпт No 13 (табл. VIII) Публікації. В. Вуйцик30 В. Корнієнко31. Розташування. Знаходиться на фасадній стіні>*

(Token number 370355, Document number 407, file 31952920, File name – Korniienko – 2010 – Korpus hrafiti Sofii Kyivs'koї (XI – pochatok XVII.pdf)

This label provides detailed information about the specific inscription, including its placement on a facade.

In seven cases, the term appears in the context of *графіті та дипінті*, referring to epigraphic artifacts. For instance:

*<у науковий обіг матеріалів. Виходом з такої ситуації покликана стати публікація максимально повного корпусу інскульптів, графіті та дипінті, що надасть до рук дослідників необхідний масив інформації залучення якого в подальшому>*

(Token number 359057, Document number 407, File ID file 31952920, File name Korniienko – 2010 – Korpus hrafiti Sofii Kyivs'koї (XI – pochatok XVII.pdf)

Additionally, in two instances, it refers to specific artifacts, highlighting its application in identifying particular carved or engraved inscriptions, such as:

*<з Крилоса налічує сорок п'ять одиниць. Специфікою Успенської церкви можна назвати наявність в її стінах низки інскульптів, що відносяться як до часів зведення храму у 1579–1584 рр. (плити на східному фасаді зі згадками Романа Івановича та>*

(Token number 513865, Document number 407, File ID file 31952920, File name Korniienko – 2010 – Korpus hrafiti Sofii Kyivs'koї (XI – pochatok XVII.pdf)

The term *Дипінті* is used in contexts involving painted inscriptions, often referenced in descriptions of artistic or decorative elements such as *емблеми* and *фреска* (Figure 3). The top collocates for *дипінті* include *графіті, табл., категорія, напис*, and *емблеми*, indicating its association with painted and decorative inscriptions.

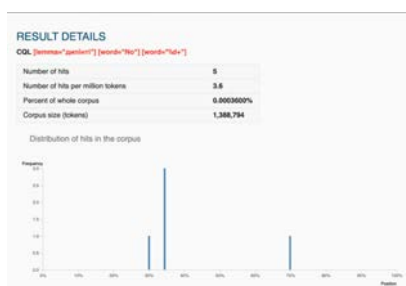


Fig. 3. Quantitative data for the term *Дипінті*

Similar to *інскульпт*, *дипінті* is used as a category label in five cases, such as:

*<1 2 Таблиця ССХVІ Фотографія та прорисовка: 1 – графіті No 756; 2 – дипінті No 759. 414 Розділ III. Епіграфічні пам'ятки церкви св. Пантелеймона>*

(Token number 512319, Document number 407, File ID file 31952920, File name Korniienko – 2010 – Korpus hrafiti Sofii Kyivs'koї (XI – pochatok XVII.pdf)

In seven instances, it appears in the context of *графіті та дипінті*, as shown in:

*<у науковий обіг матеріалів. Виходом з такої ситуації покликана стати публікація максимально повного корпусу інскульптів, графіті та дипінті, що надасть до рук дослідників необхідний масив інформації залучення якого в подальшому>*

(Token number 359057, Document number 407, File ID file 31952920, File name Korniienko – 2010 – Korpus hrafiti Sofii Kyivs'koї (XI – pochatok XVII.pdf)

In comparing these terms, *графіті* emerges as the dominant term in the corpus, reflecting its general use for various types of inscriptions. *Інскульпт*, on the other hand, is specifically used for carved or engraved inscriptions, indicating a narrower scope. Meanwhile, *дипінті* is associated with painted inscriptions often linked to artistic or decorative contexts. Both *інскульпт* and *дипінті* refer to specific methods of inscription, each with distinct contextual uses and collocates, highlighting their specialized applications within the field of epigraphy.

*Labeling terms in Digital Publication:* To fully understand and precisely establish the hierarchy of the terms labeling the objects of research in epigraphy as an academic domain and the labels used for identifying entries in epigraphic corpora, we also researched how these entries were labeled in the digital domain. Unfortunately, there exists only one corpus of epigraphic heritage from Ukraine, «Inscriptions of the Northern Black Sea» (IOSPE) [18], which is represented in English and Russian. The Ukrainian version is not fully ready, but the top-level pages and nomenclature have been translated. In this corpus, the top-level label is *Напис* and *Inscription* in English. Later, it is divided into *Серія 1* (*Лapidарні написи*). We also come across «Готується до публікації: II. 1 Графіті та дипінті Борисфену/Березані» and the English version «Series 2 of the corpus will contain graffiti and dipinti». As can be seen, in the digital domain, the preferred term to refer to inscriptions is *Напис*, which is translated into *Inscription* in English. However, on the top level, they are differentiated according to the execution technique in Ukrainian academia in the digital domain, just like in the non-digital domain.

We searched the top-level labels used in most known corpora of Greco-Roman epigraphic heritage. There are many published editions, corpora, and books dedicated to epigraphic heritage that differentiate the epigraphic artifacts based on the technique of execution from the start. However, digital collections do not differentiate inscriptions from graffiti or dipinti. They are included under the same top-level term *Inscriptions* with a unique ID. These inscriptions are indexed according to the execution technique as

the difference between them is established with the help of the EpiDoc schema usage [19]. EpiDoc is a set of guidelines and an XML schema used for encoding and sharing scholarly editions of ancient texts, particularly inscriptions. It provides a standardized framework for the digital representation of epigraphic data, ensuring that information is consistent, interoperable, and accessible across different digital platforms. In most cases, the <physDisc> is used to encode these using the terms included in the EAGLE <EXECUTION TECHNIQUE> vocabulary [20]. In the digital corpora, this information is encoded according to EpiDoc guidelines in the <physDesc> section using one term. This vocabulary collects all writing techniques and types as attested in the EAGLE BPN [21], a framework developed under the EAGLE project that sets guidelines and standards for cataloging and describing ancient Greek and Latin inscriptions. It aims to harmonize epigraphic data across different collections, making it easier to access, share, and compare information within the field of digital epigraphy. A major classification problem is faced in this respect by epigraphists. In some cases, the distinction between the method used to produce a text is used as a principle for a definition; in others, the name of the writing technique is based on the result obtained, no matter the method used. For example, Punctum litt., Scariphatis, and litt. Eminentibus are definitions of this second type.

According to the EAGLE BPN a first distinction that should be made when defining a writing technique should be whether: the text has been obtained by adding material; the text has been obtained by subtracting material; the text has been obtained by modifying the material (as in cases of impression with a matrix which does not always involve subtraction of material).

For instance, in the Epigraphic Database Heidelberg (EDH) [22], there is an Extended Search option where, in the drop-down window, you can refine your search according to the execution technique (Figure 4).



Fig. 4. The Epigraphic Database Heidelberg (EDH)

On the contrary, in Ukrainian academia, the differentiation of inscriptions based is not based on the writing technique: «Залежно від техніки письма виділяють такі різновиди епіграфем (написів): інскульпт – написи вирізані чи вигравійовані на твердому матеріалі; дипінті – зроблені пензлем; графіті – видряпані гострим предметом. Сукупність написів наз. епіграфією.» (Українська літературна енциклопедія) [23]. Here the emphasis is on the execution technique.

*A sample of SKOS vocabulary for ukrainian epigraphy:* As part of this effort, we developed a sample SKOS vocab-

ulary that categorizes key terms within Ukrainian epigraphy, such as *графіті*, *інскульпт*, and *дипінті*. This SKOS vocabulary provides a structured framework for organizing these terms, ensuring consistency and interoperability with global frameworks like EAGLE. Below is an example of how these terms are represented within the SKOS framework:

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-
rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:skos="http://www.w3.org/2004/02/skos/core#">
```

```
<!-- Concept Scheme -->
<skos:ConceptScheme rdf:about="http://example.org/
skos/ukrainian_epigraphy">
<rdfs:label>Ukrainian Epigraphy SKOS Vocabulary</
rdfs:label>
<rdfs:comment>A SKOS vocabulary for categorizing
Ukrainian epigraphic terms.</rdfs:comment>
</skos:ConceptScheme>
```

```
<!-- Concept: Графіті -->
<skos:Concept rdf:about="http://example.org/skos/
ukrainian_epigraphy/graffiti">
<rdfs:label>Графіті</rdfs:label>
<skos:inScheme rdf:resource="http://example.org/skos/
ukrainian_epigraphy"/>
<skos:broader rdf:resource="http://example.org/skos/
ukrainian_epigraphy/inscription"/>
<skos:definition>Scratched inscriptions, typically
carved into a surface using a sharp object.</skos:definition>
<skos:related rdf:resource="http://example.org/skos/
ukrainian_epigraphy/insculpt"/>
<skos:related rdf:resource="http://example.org/skos/
ukrainian_epigraphy/dipinti"/>
</skos:Concept>
```

```
<!-- Concept: Інскульпт -->
<skos:Concept rdf:about="http://example.org/skos/
ukrainian_epigraphy/insculpt">
<rdfs:label>Інскульпт</rdfs:label>
<skos:inScheme rdf:resource="http://example.org/skos/
ukrainian_epigraphy"/>
<skos:broader rdf:resource="http://example.org/skos/
ukrainian_epigraphy/inscription"/>
<skos:definition>Carved or engraved inscriptions on
a hard material.</skos:definition>
<skos:related rdf:resource="http://example.org/skos/
ukrainian_epigraphy/graffiti"/>
<skos:related rdf:resource="http://example.org/skos/
ukrainian_epigraphy/dipinti"/>
</skos:Concept>
```

```
<!-- Concept: Дипінті -->
<skos:Concept rdf:about="http://example.org/skos/
ukrainian_epigraphy/dipinti">
<rdfs:label>Дипінті</rdfs:label>
<skos:inScheme rdf:resource="http://example.org/skos/
ukrainian_epigraphy"/>
```

```

<skos:broader rdf:resource="http://example.org/skos/ukrainian_epigraphy/inscription"/>
<skos:definition>Painted inscriptions, often associated with decorative or artistic elements.</skos:definition>
<skos:related rdf:resource="http://example.org/skos/ukrainian_epigraphy/graffiti"/>
<skos:related rdf:resource="http://example.org/skos/ukrainian_epigraphy/insculpt"/>
</skos:Concept>

<!-- Broader Concept: Inscription -->
<skos:Concept rdf:about="http://example.org/skos/ukrainian_epigraphy/inscription">
<rdfs:label>Inscription</rdfs:label>
<skos:inScheme rdf:resource="http://example.org/skos/ukrainian_epigraphy"/>
<skos:narrower rdf:resource="http://example.org/skos/ukrainian_epigraphy/graffiti"/>
<skos:narrower rdf:resource="http://example.org/skos/ukrainian_epigraphy/insculpt"/>
<skos:narrower rdf:resource="http://example.org/skos/ukrainian_epigraphy/dipinti"/>
<skos:definition>A general term for any text or image inscribed on a surface.</skos:definition>
</skos:Concept>

</rdf:RDF>

```

**Fig. 6. sample SKOS ConceptScheme**

*Explanation (Figure 6):*

- **ConceptScheme:** Represents the overall SKOS vocabulary, titled "Ukrainian Epigraphy SKOS Vocabulary."
- **Concepts:** Each <skos:Concept> represents a specific term within the epigraphic domain:
  - **Графіті (Graffiti):** Defined as scratched inscriptions and related to other terms like "інскульпт" and "дипінті."
  - **Інскульпт (Insculpt):** Defined as carved or engraved inscriptions, also related to "графіті" and "дипінті."
  - **Дипінті (Dipinti):** Defined as painted inscriptions, with relationships to "графіті" and "інскульпт."
  - **Broader Concept:** "Inscription" is the broader term that encompasses "графіті," "інскульпт," and "дипінті."

This sample demonstrates how key terms within Ukrainian epigraphy can be structured and defined using SKOS, facilitating their integration into larger digital humanities frameworks like EAGLE.

This research successfully identified and analyzed the key terminological categories within Ukrainian epigraphy, focusing on their usage in both digital and non-digital contexts. A significant finding is the difference in top-level labels between EAGLE and Ukrainian epigraphic discourse. Ukrainian epigraphy categorizes inscriptions primarily based on their execution techniques, using terms such as "графіті," "інскульпт," and "дипінті" as top-level labels within collections. These terms reflect the historical and archaeological significance of the inscriptions within Ukrainian heritage. In contrast, the EAGLE framework, widely used for Greco-Roman epigraphic collections, typically employs a broader top-level term "inscription," with detailed information about the execution technique provided at a later stage within each entry. This underscores the importance of understanding the linguistic, cultural, and contextual factors that influence terminology choices in this domain to compile truly effective SKOS vocabularies that integrate Ukrainian digital epigraphy into the global network.

Integrating Ukrainian epigraphic vocabularies with the EAGLE framework is a strategic step towards enhancing cross-cultural and cross-linguistic research. By developing a SKOS vocabulary tailored to Ukrainian inscriptions and aligning it with established frameworks like EAGLE, we can promote a more integrated and accessible digital representation of epigraphic heritage. This effort not only enriches the global digital humanities landscape but also ensures that the rich epigraphic heritage of Ukraine receives the recognition and scholarly attention it deserves.

Looking ahead, we plan to compile a similar SKOS for Armenian epigraphy. Additionally, we will test our developed SKOS vocabulary on a sample trilingual corpus of Armenian epigraphic heritage from Ukraine. This approach will allow us to evaluate the adaptability and effectiveness of the SKOS model across diverse cultural and linguistic contexts. By doing so, we aim to foster cross-cultural integration and contribute to the preservation and digitization of Armenian epigraphic heritage alongside Ukrainian inscriptions, further enriching global digital humanities initiatives.

### Bibliography:

1. Німчук В. Історія української мови. Хрестоматія X–XIII ст. Житомир: Полісся, 2015. 384 с.
2. Скржинська М. В., Котляр М. Ф. Епіграфіка [Електронний ресурс] // Енциклопедія історії України: Т. 3: Е-Й / Редкол.: В. А. Смолій (голова) та ін. НАН України. Інститут історії України. Київ: Наукова думка, 2005. 672 с.: іл. URL: <http://www.history.org.ua/?termin=Епіграфіка> (останній перегляд: 18.06.2024).
3. Pattuelli M. C., Miller M., Lange L., Fizzell S., Li-Madeo C. Crafting linked open data for cultural heritage: Mapping and curation tools for the linked jazz project. Code 4 Lib, 21, July 2013.
4. Oomen J., Baltussen L., van Erp M. Sharing cultural heritage the linked open data way: Why you should sign up. Paper presented at Museums and the Web 2012, San Diego, CA, 2012.
5. Moraitou E., Aliprantis J., Christodoulou Y., Teneketzis A., Caridakis G. Semantic Bridging of Cultural Heritage Disciplines and Tasks. Heritage. 2019; 2(1):611-630. DOI: <https://doi.org/10.3390/heritage2010040>.
6. Корнієнко В. В. Корпус графіті Софії Київської (XI – початок XVIII ст.) [у 3 ч.]. Київ, 2009–2011. Ч. 1: Приділ св. Георгія Великомученика. – 2009. 464 с. Ч. 2: Приділ свв. апостолів Петра і Павла. – 2010. 272 с. Ч. 3: Центральна нава. – 2011. 400 с.
7. Корнієнко В. В. Корпус графіті Софії Київської (XI – початок XVIII ст.), Ч. V: Приділ свв. Іоакима та Анни (південна сторона). Київ, 2015. 440 с.

8. Корнієнко В. В. Корпус графіті Софії Київської (XI – початок XVIII ст.), Ч. V: Приділ свв. Іоакима та Анни (північна сторона). Київ, 2013. 656 с.
9. Корнієнко В. В. Корпус графіті Софії Київської (XI – початок XVIII ст.), Ч. II: Приділ свв. апостолів Петра і Павла. Київ, 2010. 272 с.
10. Корнієнко В. В. Корпус графіті Софії Київської (XI – початок XVIII ст.), Ч. VII, Т. 1: Західна зовнішня галерея. Київ, 2016. 240 с.
11. Корнієнко В. В. Епіграфіка сакральних пам'яток Галича (XII–XIX ст.). Галич, 2017. 448 с.
12. Корнієнко В. В. Корпус графіті церкви Спаса на Берестові (остання третина XI – перша третина XVIII ст.). Київ, 2013. 304 с.
13. Тамразян А. С. Exploring Descriptive Terminology in Ukrainian Epigraphy: A Corpus-Based Study of Synonymous Triplet «графіті», «напис» and «запис» // *Науковий часопис Національного педагогічного університету імені М. П. Драгоманова. Серія 9. Сучасні тенденції розвитку мов.* 2024. № 27. С. 23-35.
14. <https://www.w3.org/2004/02/skos/>
15. [https://epigraphy.info/ontologies\\_wg/](https://epigraphy.info/ontologies_wg/) Accessed 10 May 2024
16. <https://www.eagle-network.eu> Accessed 10 May 2024
17. <https://www.sketchengine.eu> Accessed 10 May 2024
18. <https://iospe.kcl.ac.uk/index.html> Accessed 10 May 2024
19. <https://epidoc.stoa.org/gl/latest/> Accessed 10 May 2024
20. <https://www.eagle-network.eu/resources/vocabularies/writing/> Accessed 22 May, 2024
21. <https://www.eagle-network.eu/wp-content/upload>
22. <https://edh.ub.uni-heidelberg.de> Accessed 26 Aprile 2024
23. <http://izbornyk.org.ua/ulencycl/ule.htm> Accessed 26 Aprile 2024

#### References:

1. Nimchuk, V. (2015). *Istoriia ukrainskoi movy* [The history of the Ukrainian Language]. Zhytomyr: Polissia. [In Ukrainian].
2. Skryzhynska, M.V., & Kotliar, M.F. (2005). EPIGRAFIKA [Elektronnyi resurs]. In V. A. Smolii (Ed.), *Entsyklopediia istorii Ukrainy: T. 3: E-I* (pp. 672). K.: V-vo "Naukova dumka". Retrieved from <http://www.history.org.ua/?termin=Epigrafika> [In Ukrainian] (last accessed: 18.06.2024).
3. Pattuelli, M. C., Miller, M., Lange, L., Fitzell, S., & Li-Madeo, C. (2013). Crafting linked open data for cultural heritage: Mapping and curation tools for the linked jazz project. *Code 4 Lib*, 21.
4. Oomen, J., Baltussen, L., & van Erp, M. (2012). Sharing cultural heritage the linked open data way: Why you should sign up. Paper presented at Museums and the Web 2012, San Diego, CA.
5. Moraitou, E., Aliprantis, J., Christodoulou, Y., Teneketzis, A., & Caridakis, G. (2019). Semantic Bridging of Cultural Heritage Disciplines and Tasks. *Heritage*, 2(1), 611-630. <https://doi.org/10.3390/heritage2010040>
6. Korniienko, V. V. (2009-2011). *Korpus hrafiti Sofii Kyivskoi (XI – pochatok XVIII st.)* [Corpus of graffiti of St. Sophia of Kyiv (XI – early XVIII century)]. [In 3 parts]. Kyiv [In Ukrainian].
7. Korniienko, V. V. (2015). *Korpus hrafiti Sofii Kyivskoi (XI – pochatok XVIII st.), Chastyna V: Pridil svv. Ioakima ta Anny (pivdenna storona)* [Corpus of graffiti of St. Sophia of Kyiv (XI – early XVIII century), Part V: Chapel of Sts. Joachim and Anna (south side)]. Kyiv [In Ukrainian].
8. Korniienko, V. V. (2013). *Korpus hrafiti Sofii Kyivskoi (XI – pochatok XVIII st.), Chastyna V: Pridil svv. Ioakima ta Anny (pivnichna storona)* [Corpus of graffiti of St. Sophia of Kyiv (XI – early XVIII century), Part V: Chapel of Sts. Joachim and Anna (north side)]. Kyiv [In Ukrainian].
9. Korniienko, V. V. (2010). *Korpus hrafiti Sofii Kyivskoi (XI – pochatok XVIII st.), Chastyna II: Pridil svv. apostoliv Petra i Pavla* [Corpus of graffiti of St. Sophia of Kyiv (XI – early XVIII century), Part II: Chapel of Sts. Peter and Paul]. Kyiv [In Ukrainian].
10. Korniienko, V. V. (2016). *Korpus hrafiti Sofii Kyivskoi (XI – pochatok XVIII st.), Chastyna VII, tom 1: zakhidna zovnishnia halereia* [Corpus of graffiti of St. Sophia of Kyiv (XI – early XVIII century), Part VII, Vol. 1: Western outer gallery]. Kyiv [In Ukrainian].
11. Korniienko, V. V. (2017). *Epihrafika Sakralnykh Pamiatok Halycha (XII–XIX st.)* [Epigraphy of Sacred Monuments of Halych (XII–XIX centuries)]. Halych [In Ukrainian].
12. Korniienko, V. V. (2013). *Korpus hrafiti Tserkvy Spasa na Berestovi (ostannia tretyna XI – persha tretyna XVIII st.)* [Corpus of graffiti of the Church of the Saviour on Berestove (last third of the XI – first third of the XVIII century)]. Kyiv [In Ukrainian].
13. Tamrazyan, A.C. (2024). Exploring Descriptive Terminology in Ukrainian Epigraphy: A Corpus-Based Study of Synonymous Triplet «графіті», «напис» and «запис» [Exploring Descriptive Terminology in Ukrainian Epigraphy: A Corpus-Based Study of Synonymous Triplet "graffiti", "inscription" and "entry"]. *Naukovyi Chasopys Natsionalnoho pedahohichnoho universytetu imeni M. P. Drahomanova. Serii 9. Suchasni tendentsii rozvytku mov*, 27.
14. <https://www.w3.org/2004/02/skos/>
15. [https://epigraphy.info/ontologies\\_wg/](https://epigraphy.info/ontologies_wg/) Accessed 10 May 2024
16. <https://www.eagle-network.eu> Accessed 10 May 2024
17. <https://www.sketchengine.eu> Accessed 10 May 2024
18. <https://iospe.kcl.ac.uk/index.html> Accessed 10 May 2024
19. <https://epidoc.stoa.org/gl/latest/> Accessed 10 May 2024
20. <https://www.eagle-network.eu/resources/vocabularies/writing/> Accessed 22 May, 2024
21. <https://www.eagle-network.eu/wp-content/upload>
22. <https://edh.ub.uni-heidelberg.de> Accessed 26 Aprile 2024
23. <http://izbornyk.org.ua/ulencycl/ule.htm> Accessed 26 Aprile 2024