

УДК 004:78.08

О. В. Сілагін, Д. О. Зелінська, В. А. Гірдвайніс

## ОНТОЛОГІЧНЕ МОДЕЛЮВАННЯ БАЗИ ЗНАТЬ З МУЗИЧНОГО НАДЖАНРУ «МЕТАЛ»

Вінницький національний технічний університет, Вінниця

**Анотація.** У роботі проведено онтологічне моделювання предметної області визначного культурного і музичного феномену, а саме сучасного музичного наджанру «метал». Створено термінологічний словник з даної предметної області із використанням поняття терміносистеми. Проаналізовано можливості середовища розробки онтологій Protégé в контексті заявленої предметної області. Вибрано базовий принцип моделювання онтології у вигляді семантичної мережі, яку можна поступово розширяти, поглиблюючи знання про предметну область. Онтологічну модель бази знань реалізовано засобами середовища Protégé, з використанням знань про відношення функціональності, транзитивності, рефлексивності для структуризації інформації та дотримання принципів її повноти, достовірності та несуперечливості. Вибрано критерій оцінювання коректності онтологічної моделі бази знань. Проведено тестування розробленої онтологічної бази знань та підтверджено високий рівень її коректності в процесі пошуку інформації. Вказано можливі напрямки подальшого розширення онтологічної моделі.

**Ключові слова:** онтологічна модель, база знань, метал, жанр, Protégé, семантика.

**Анотация.** В работе проведено онтологическое моделирование предметной области выдающегося культурного и музыкального феномена, а именно современного музыкального супержанра «метал». Создан терминологический словарь по данной предметной области с использованием понятия терминосистемы. Проанализированы возможности среды разработки онтологий Protégé в контексте заявленной предметной области. Выбран базовый принцип моделирования онтологии в виде семантической сети, которую можно постепенно расширять, углубляя знания о предметной области. Онтологическую модель базы знаний реализовано средствами среды Protégé, с использованием знаний об отношении функциональности, транзитивности, рефлексивности для структурирования информации и соблюдения принципов ее полноты, достоверности и непротиворечивости. Выбран критерий оценки корректности онтологической модели базы знаний. Проведено тестирование разработанной онтологической базы знаний и подтвержден высокий уровень ее корректности в процессе поиска информации. Указаны возможные направления дальнейшего расширения онтологической модели.

**Ключевые слова:** онтологическая модель, база знаний, метал, жанр, Protégé, семантика.

**Abstract.** The paper presents an ontological modeling of the subject area of a prominent cultural and musical phenomenon, namely the modern musical supergenre "metal". A terminological dictionary on this subject area was created using the concept of terminological system. The possibilities of the Protégé ontology development environment in the context of the stated subject area were analyzed. The basic principle of ontology modeling in the form of a semantic network is chosen, which can be gradually expanded, deepening the knowledge about the subject area. The ontological model of the knowledge base is implemented by means of the Protégé environment, using knowledge of the relationship of functionality, transitivity, reflexivity to structure information and adhere to the principles of its completeness, reliability and consistency. The criterion for assessing the correctness of the ontological model of the knowledge base was chosen. Testing of the developed ontological knowledge base was carried out and the high level of its correctness in the process of information retrieval was confirmed. Possible directions of further expansion of the ontological model are indicated.

**Keywords:** ontology model, knowledge base, metal, genre, Protégé, semantics.

**DOI:** <https://doi.org/10.31649/1999-9941-2021-51-2-45-50>.

### Вступ

Поняття онтології знань набуло важливого практичного значення останнім часом у сферах штучного інтелекту, управління знаннями та розробки складних спеціалізованих інформаційних ресурсів. Для фахівців з найрізноманітніших сфер нагальною потребою стає структуризація знань у межах предметної області їх дослідження. Всеосяжна та детальна формалізація знань має на увазі побудову концептуальної схеми предметної області. Така схема будується на основі двох базових термінів, а саме набору понять та інформації про дані поняття [1].

Побудови детальних карт набору понять на сьогоднішній день успішно реалізуються більшістю автоматизованих систем, проте набагато складнішим завданням є правильна подача усієї можливої інформації про поняття. Значимими блоками інформації є властивості об'єктів, відношення між ними, обмеження усіх відомих видів (від обмежень на тип даних і до обмежень на взаємодію певних об'єктів між собою), аксіоми тощо. Редактор онтологій Protégé [2] є одним з найпотужніших засобів формалізації знань, що може реалізовувати більшість взаємодій об'єктів. Окрім того, Protégé є системою з можливістю розширення функціональності, а це означає, що вона постійно перебуває у процесі розвитку та доповнюється зусиллями наукової спільноти.

### Актуальність

Тематика предметної області створення онтології, а саме музичне мистецтво, що належить до різновиду «метал» у базовій класифікації музичних жанрів [3], є доволі актуальною на сьогоднішній день як культурний феномен ХХ і ХХІ століття і як базис окремої філософської течії, що знайшла багато прихильників по усьому світу. Розроблена онтологічна модель бази знань даного напрямку може бути впроваджена на тематичних веб-ресурсах і значно полегшувати семантичний пошук інформації в межах предметної області. Разом із тим, дана тематика ідеально підходить для проведення дослідження можливостей редактору онтологій Protégé, проектування або встановлення додаткових плагінів у майбутньому в разі необхідності та вирішення питання доцільності розширення функціональності середовища.

### Мета

Метою даного дослідження є підвищення коректності семантичного пошуку інформації зацікавленими користувачами, що планується досягти за рахунок створення нової онтологічної моделі бази знань про музичний наджанр «метал».

### Задачі

Для досягнення цієї мети і створення нової онтологічної моделі бази знань потрібно вирішити наступні задачі:

1. Дослідити предметну область для створення онтологічної моделі бази знань і відповідно до неї скласти мінімальний словник термінів і сформулювати терміносистему.
2. Створити базову графічну модель у вигляді фрагменту семантичної мережі.
3. Реалізувати онтологічну модель бази знань засобами Protégé, зафіксувавши ключові концепти і терми, після чого закодувати їх.
4. Провести тестування коректності пошуку інформації в реалізованій засобами Protégé онтологічній базі знань та оцінити коректність пошуку за обраним критерієм.

### Створення мінімального термінологічного словника для предметної області музичного наджанру «метал»

Початковим етапом для створення онтології предметної області метал-музики є побудова так званої терміносистеми. Терміносистему потрібно розглядати як систематизовану сукупність термінів певної предметної області, що створюється експертами на основі теорії та володіє властивостями системності, повноти, несуперечливості, відносної стабільності, відкритості та динамічності [4]. Важливим є знання того факту, що при наявності ряду теорій в одній предметній області можуть мати місце одразу декілька терміносистем. У подальшому розвитку онтології потрібно розвивати кожен з них окремо, проте у будь-якому випадку основа повинна бути єдиною.

Таким чином, на даному етапі потрібно створити структуризовану мінімальну таблицю з основними дефініціями майбутньої онтології. Необхідно не лише врахувати основні поняття і відокремити їх від другорядних, а і обрати найбільш яскраві приклади певних термінів у вигляді класів та створити їх ієрархію. Ієрархія буде доповнюватись по мірі поглиблення і розширення онтології, проте її базис повинен мати такий вигляд, щоб дати достатнє уявлення про структуру предметної області користувачеві або розробнику, яким би нескінченно малим не був його набір знань у даному напрямку. Таблиця 1 ілюструє саме такий базовий набір знань та уявлень для глибшого розуміння предметної області з музичного наджанру «метал». Терміни подані одночасно українською та англійською мовами, так як дана предметна область значною мірою інтернаціоналізована.

Таблиця 1 – Мінімальний термінологічний словник для уявлення про наповнення онтології метал-музики

Термін	Дефініція	Приклади властивостей відношення	Приклади властивостей даних	Тип даних	Приклади індивідів
Band (гурт)	Колектив людей, об'єднаних якоюсь спільною метою або роботою, які так чи інакше пов'язані з музикою.	Have_motherland (мати країну походження), have_number_of_musicians (мати кількість музикантів у складі), have_rate (мати рейтинг).	Genre (жанр), frontman (фронтмен/лідер), age (вік/кількість років активної діяльності).	String	Gehenna, Immortal, Type O Negative, Olshano, Lucifugum, Drudkh, Electric Wizard, Deeds of Flesh.
Genre (жанр)	Сукупність стильових музичних характеристик, що дозволяє об'єднати виконавців наджанру «метал» у менші класифікаційні групи.	Have_origin (мати походження), have_relative_genres (мати споріднені жанри), have_founder (мати засновника, виконавця або гурт).	Age (приблизний вік існування), heyday (роки розквіту), derivative forms (похідні жанри), origins (походження).	String	Black Metal, Death Metal, Doom Metal, Avangard Metal, Gothic Metal, Folk Metal, Gothic Metal, Power Metal, Thrash Metal, Heavy Metal.

Country of origin (країна походження)	Країна, де було засновано гурт або з якої походять більшість членів гурту.	Have_capital (мати столицю), have_cultural_origins (мати культурні основи для виникнення даного жанру).	Region (частина світу, до якої належить дана країна).	String	Norway, Sweden, Ukraine, Poland, USA, France, Germany.
Label (лейбл звукозапису)	Бренд, створений компаніями, що займаються виробництвом, поширенням і просуванням аудіо- та іноді відеозаписів на носіях різних відповідних форматів.	Have_motherland (країна лейблу), have_founder (мати засновника).	Specialties (жанри спеціалізації), status (статус активності), contacts (контакти для зв'язку).	String	Propaganda, Supernal Music, Nuclear Blast, Drakkar Productions, Grieghallen Studio, Шум Деревьев Records, Hammermark Art.
Status (статус гурту)	Стан гурту, його позиція відносно даного моменту часу.	Be_unknown (бути невідомим).	Present_status (теперішній статус), past_status (минулий статус).	Може бути string або bool у спрощеному варіанті.	Active, Unactive, Changed name, Split up.

#### Створення базової графічної моделі у вигляді фрагменту семантичної мережі

Після виявлення ключових концептів доцільним є побудування ієрархії предметної області, в корені якої буде знаходитись терміносистема, а у вузлах – номенклатура. У якості моделі представлення знань номенклатури використовується семантична мережа знаків-фреймів.

Семантична мережа  $S$  будується як об'єднання знаків-фреймів  $\Phi_i$ :  $S = U_i \Phi_j$ . Будемо вважати, що загалом сформовано множину термінів Term. Це означає, що потужність множини  $\Phi = \{\Phi_i\}$  більша потужності множини Term [5]. Графічне зображення фрагменту семантичної мережі  $S$  для побудови онтології метал-музики можна побачити на рисунку 1.

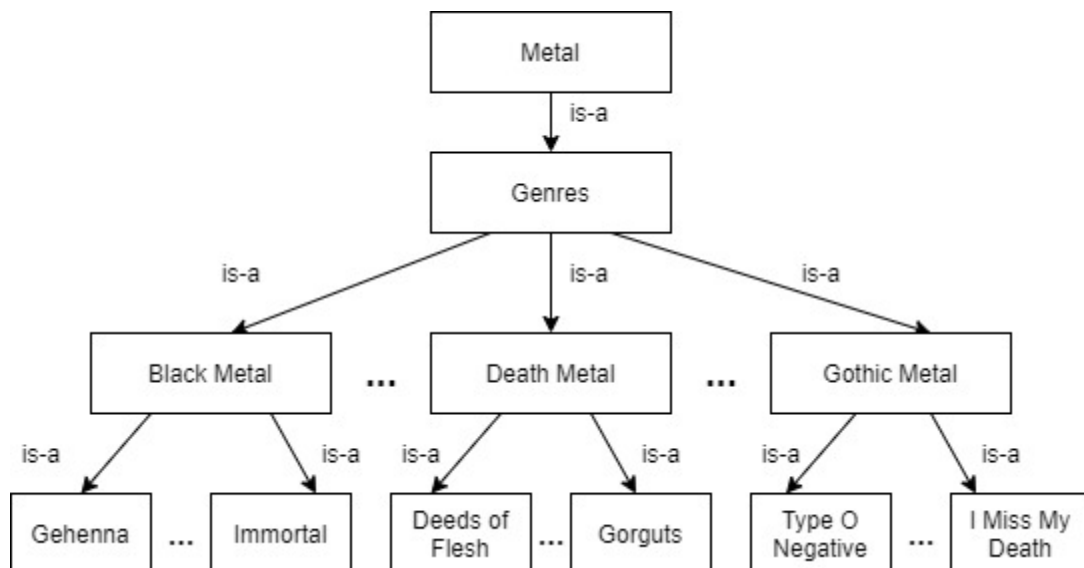


Рисунок 1 – Фрагмент семантичної мережі для побудови онтології метал-музики

Надалі проведемо кодування базису онтології метал-музики. Розпочинати потрібно із налаштування ієрархії класів та підкласів, потім надати їм властивості відношень та властивості даних (визначити усі

необхідні типи даних зі списку або створити нові за потреби), а після цього перейти до створення окремих індивідів у межах кожного класу.

### Реалізація онтологічної моделі бази знань засобами Protégé

У Protégé наявні дуже масштабні можливості для роботи з відношеннями, а надання їм властивостей транзитивності, рефлексивності, функціональності та інших дозволяє значно спростити проблеми, що часто постають у межах предметної області. Розглянемо конкретні приклади щоб довести практичне значення встановлення подібних властивостей.

Гурт, що умовно належить до індивідів класу жанру Black Metal (хоча за жанровою класифікацією також належить і до класів Post Metal, Avangard Metal), має назву Olshanoє з 2019 року, хоча до цього мав назву Isa. Проте користувач може не володіти даною інформацією і сприймати відповідно дані індивіди як два абсолютно різних гурти. Як відомо, інформація повинна володіти властивістю повноти, тому важливим буде повідомити користувача про даний факт. Здійснити це можна засобами встановлення відношення еквівалентності. У Protégé даний факт можна реалізувати засобами функції «Same Individual As».

Також слід звернути увагу на функцію «Disjoint With» – її необхідно застосовувати навпаки при забороні використання суперечливих понять або термінів. Даний оператор не можна застосовувати як розмежування жанрів, адже різниця між ними досить нечітка, і один гурт може поєднувати декілька з них. А при визначенні статусу того чи іншого гурту «Disjoint With» є нагальною потребою для забезпечення вимоги несуперечливості – один і той самий гурт не може бути одночасно активним і неактивним. Проте варто врахувати, що він може містити одночасно статус «Active» та «Changed\_name» як у попередньому прикладі, тобто бути активним, але змінити назву, і ставити функцію «Disjoint With» між цими двома категоріями не тільки непотрібно, а й категорично заборонено.

При розробці онтології метал-музики варто врахувати реалізацію пошуку індивідів за певними критеріями. Для цього існує Class Expression Editor – зручний інструмент для написання виразів у межах певного класу. Наприклад, користувач хоче знайти музиканта у межах жанру Black Metal, який при цьому мав би вік 38 років. Тому у межах даного жанру потрібно скласти наступну умову: «Musician and have\_age value 38». Далі здійснюється пошук усіх індивідів-музикантів даного жанру, яким 38 років. Варто зазначити, що перед цим у всіх індивідів були прописані дані про їх вік і жанр, у якому вони працюють, притому тип даних віку був позначений як «int». При наявності іншого типу даних система видала б помилку і не працювала б з цілочисельним значенням «38».

Проте найпотужнішим засобом для створення запитів на пошук інформації у Protégé є вбудований конструктор мови запитів SPARQL [6]. Основними видами таких запитів є SELECT запит, CONSTRUCT запит, ASK запит та DESCRIBE запит. Найбільш поширеним з даних різновидів є SELECT запит, результати виведення запиту подаються у зручному для користувача табличному вигляді. Коректність виведення запитів є однією з ознак правильності побудови онтології.

Окрім того, графічним відображенням онтології є інструмент Ontograp, що виводить усю семантику предметної області в найбільш зрозумілому та простому поданні.

На рисунку 2 зображено декотрі моменти з процесу створення онтології музичного наджанру «метал» у вигляді колажу, а на рисунку 3 показано власне онтологічний граф в результаті побудови базису онтології.

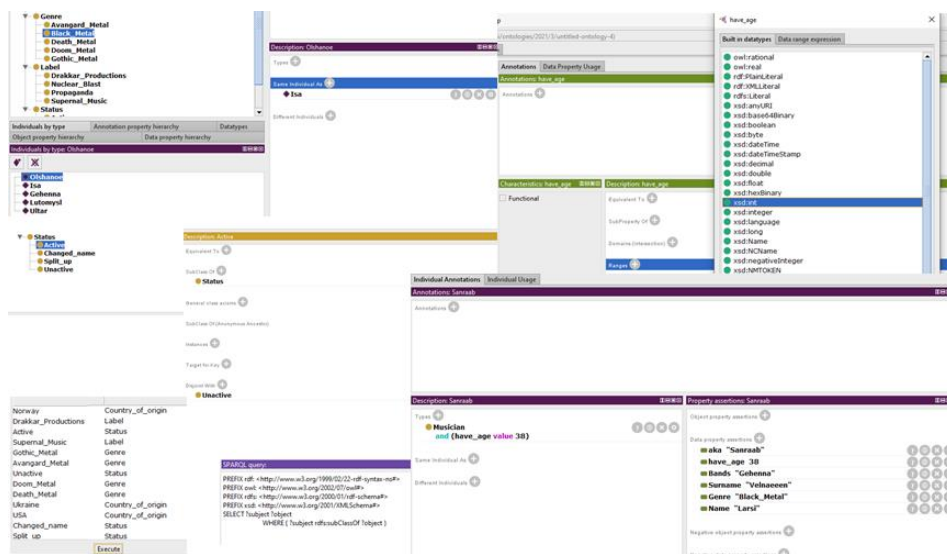


Рисунок 2 – Основні моменти кодування онтології наджанру «метал»

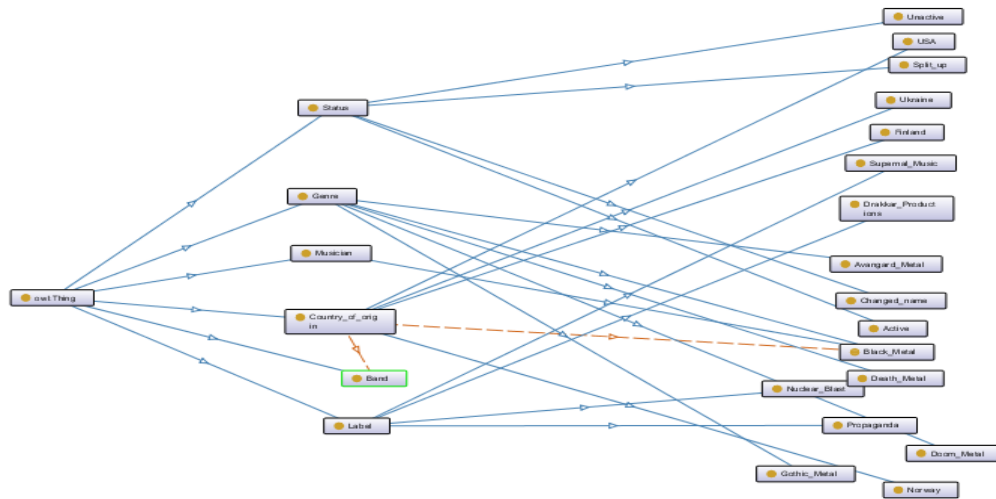


Рисунок 3 – Базовий граф онтології наджанру «метал»

### Тестування та оцінювання коректності семантичного пошуку інформації в онтологічній базі знань музичного наджанру «метал»

Гомез-Перез запропонував наступне визначення оцінки онтології у контексті технологій сумісного використання знань: «для складання технічного думки про онтологію, асоційованого з нею програмного середовища і документації, щодо критеріїв. Критеріями можуть виступати вимоги до специфікації, питання компетенції, і / або реальний світ» [7].

Серед сучасних метрик не так багато таких, які враховують усі задані критерії, тому було вирішено скористатись метрикою Single Usability Metrics (SUM), що розроблена аналітиком Джефом Соро [8].

Дана метрика визначається за формулою  $SUM = \frac{SM_{comp} + SM_{sat} + SM_{time} + SM_{er}}{4}$ . У контексті розробки онтологій показник *Completion* є булевим значенням і позначає досягнення мети здобуття інформації користувачем, *Satisfaction* показує рівень задоволення від структуризації інформації у даній онтології, *Errors* показує виникнення певних суперечливостей або інших помилок при отриманні інформації та *Times* показує середню кількість запитів для отримання тієї чи іншої інформації.

Для визначення коректності пошуку інформації у реалізованій онтологічній базі знань, вона була надана 10 користувачам, кожен з яких реалізував на ній по 10 запитів.

Результати опитування за даною метрикою представлені у таблиці 2.

Таблиця 2 – Результати опитування щодо коректності розробленої онтології

User	Satisfaction	Completion	Errors	Times	Average
1	5	1	0	5	90%
2	5	1	0	7	85,5%
3	5	1	2	3	87,5%
4	5	0	1	4	62,5%
5	4	1	0	4	88,75%
6	4	0	1	3	63,75%
7	5	1	0	3	100%
8	4	1	0	11	76,75%
9	5	1	0	3	100%
10	5	1	1	5	83,75%

Середній показник за метрикою SUM по всім користувачам відповідно становитиме 83,85%, що є досить непоганим показником для онтологічних баз знань. В той же час проведено за аналогічною методикою тестування бази даних наджанру «метал» сайту «Encyclopaedia Metallum», де використовувалась класична реляційна модель організації баз даних, показала набагато нижчі результати. Так середній показник за метрикою SUM по 10 користувачам відповідно становив 75,32%.

### Висновки

1. У статті запропоновано нова онтологічна модель організації бази знань (онтологія) предметної області: музичний наджанр «метал».

2. Інструментальними засобами середовища Protégé онтологія предметної області музичний наджанр «метал» реалізована у вигляді бази знань.

3. Проведено тестування реалізованої бази знань та оцінювання коректності семантичного пошуку за метрикою SUM.

4. Мета дослідження – підвищення коректності семантичного пошуку інформації досягається за рахунок застосування в реалізації бази знань нової онтологічної моделі (онтології) про музичний наджанр «метал». У порівнянні з класичною, реляційною моделлю організації баз даних застосування нової онтології дозволяє підняти коректність пошуку на 8%.

#### Список літератури

- [1] А. О. Никоненко, "Огляд баз знань онтологічного типу," *Штучний інтелект*, № 4, с. 208-219. 2009.
- [2] Protégé. [Online]. Available: <https://protege.stanford.edu/>. Accessed on: May 18, 2021.
- [3] Home – Encyclopaedia Metallum: The Metal Archives. [Online]. Available: <https://www.metal-archives.com/>. Accessed on: May 21, 2021.
- [4] Принципи побудови банків даних. [Електронний ресурс]. Режим доступу: [https://web.posibnyky.vntu.edu.ua/fitki/10savchuk\\_organizaciya\\_bazdanih\\_znan/gl\\_14.html](https://web.posibnyky.vntu.edu.ua/fitki/10savchuk_organizaciya_bazdanih_znan/gl_14.html). Дата звернення: May 17, 2021.
- [5] SPARQL. [Online]. Available: <https://znaio.com.ua/SPARQL>. Accessed on: May 21, 2021.
- [6] A. Gomez-Perez, N. Juristo, J. Pazos; in N.J. Mars, editor, "Evaluation and assessment of knowledge sharing technology," *Towards Very Large Knowledge Bases*, Amsterdam: IOS Press, 1995, p. 289-296.
- [7] Расчет SUM. Подход от Джефа Соро. [Електронний ресурс]. Режим доступу: <https://v-shliachkov.medium.com/расчет-sum-подход-от-джефа-соро-c1ecf796f1b2>. Дата звернення: May 25, 2021.

Стаття надійшла: 13.07.2021.

#### References

- [1] A. O. Nykonenko, "Ohliad baz znan ontolohichnoho typu," *Shtuchnyi intelekt*, № 4, s. 208-219. 2009.
- [2] Protégé. [Online]. Available: <https://protege.stanford.edu/>. Accessed on: May 18, 2021.
- [3] Home – Encyclopaedia Metallum: The Metal Archives. [Online]. Available: <https://www.metal-archives.com/>. Accessed on: May 21, 2021.
- [4] Pryntsyvy pobudovy bankiv danykh. [Elektronnyi resurs]. Rezhym dostupu: [https://web.posibnyky.vntu.edu.ua/fitki/10savchuk\\_organizaciya\\_bazdanih\\_znan/gl\\_14.html](https://web.posibnyky.vntu.edu.ua/fitki/10savchuk_organizaciya_bazdanih_znan/gl_14.html). Data zvernennia: May 17, 2021.
- [5] SPARQL. [Online]. Available: <https://znaio.com.ua/SPARQL>. Accessed on: May 21, 2021.
- [6] A. Gomez-Perez, N. Juristo, J. Pazos; in N.J. Mars, editor, "Evaluation and assessment of knowledge sharing technology," *Towards Very Large Knowledge Bases*, Amsterdam: IOS Press, 1995, p. 289-296.
- [7] Raschet SUM. Podkhod ot Dzhefa Soro. [Online]. Available: <https://v-shliachkov.medium.com/raschet-sum-podkhod-ot-dzhefa-soro-c1ecf796f1b2>. Accessed on: May 25, 2021.

#### Відомості про авторів

Сілагін Олексій Віталійович – кандидат технічних наук, доцент кафедри комп'ютерних наук.

Зелінська Дарія Олегівна – магістрантка групи ІКН-20м, кафедра комп'ютерних наук.

Гірдвайніс Владислав Аудрісович – магістрант групи ІКН-20м, кафедра комп'ютерних наук.

О. В. Силагин, Д. О. Зелинская, В. А. Гирдвайнис

## ОНТОЛОГИЧЕСКОЕ МОДЕЛИРОВАНИЕ БАЗЫ ЗНАНИЙ МУЗЫКАЛЬНОГО НАДЖАНРА «МЕТАЛ»

Винницкий национальный технический университет, Винница

O. V. Silagin, D. O. Zelinska, V. A. Girdvainis

## ONTOLOGICAL MODELING OF THE KNOWLEDGE BASE OF THE MUSICAL SUPERGENRE "METAL"

Vinnitsia National Technical University, Vinnitsia