

## ПОБУДОВА ОНТОЛОГІЇ АРХІТЕКТУРНИХ ТЕРМІНІВ

**В. В. Литвин, О. І. Ремшило-Рибчинська, В. А. Висоцька**

**Національний університет “Львівська політехніка”**

**E-mail: Victoria.A.Vysotska@lpnu.ua**

Запропоновано метод побудови онтології архітектурних термінів для подання і розподілу знань загального словникового запасу і підтримки інтелектуальних запитів у базах даних з архітектури, а також автоматизованого обміну даними та інтеграції навчальних програм з архітектури. Апробовано її для пошуку релевантної інформації, а саме, для вибірки сторінок у мережі Інтернет із синтаксично різними, але семантично однаковими архітектурними термінами.

**Ключові слова:** *онтологія, архітектурний термін, поняття, відношення, множина.*

## ARCHITECTURAL TERMS ONTOLOGY CONSTRUCTION

**V. V. Lytvyn, O. I. Remeshylo-Rybchynska, V. A. Vysotska**

**National University “Lviv Polytechnics”**

The method of the architectural ontology terms development is proposed. It can be used as a mechanism for the provision and distribution of knowledge for definition of the general vocabulary and intellectual queries support in databases on architecture and support of data automated exchange and the architectural educational programs integration. It is proposed to use it for a relevant information search. For example, it is used for sampling pages from the Internet with syntactically different, but semantically identical architectural terms. The content of the research is to solve the application problem of ontology building for intelligent systems of information resources processing of the architectural structure (thesauruses, Web-resources, e-distance learning resources, educational programs, scientific publications, etc.) for textual data sets analyzing of architectural domains. The knowledge base ontology use in intelligent systems for information resources processing of architectural profile helps to solve a number of problems of methodological and technological character that arise during the development of such systems. In particular, Ukraine has a problem of the lack of conceptual integrity and coherence of individual techniques and knowledge on engineering methods for the development of appropriate systems in this area; lack of qualified specialists for software development of architectural direction; stiffness of developed software tools and their low adaptive capacity; complexity of the intelligent systems implementation for information resources processing of architectural profile, due to the psychological aspects. All this demonstrates and confirms the relevance of research problems using ontologies in intelligent systems building for information resources processing.

**Keywords:** *ontology, architecture, term, architectural terms, concepts, relations, set.*

**Загальне формулювання задачі.** Дослідження полягає у вирішенні прикладної задачі побудови онтології для інтелектуальних систем опрацювання інформаційних ресурсів архітектурного профілю (тезаурусів, Web-ресурсів, електронних ресурсів дистанційного навчання, навчальних програм, наукових публікацій тощо) для аналізу текстових масивів даних з архітектурних предметних областей (ПО). Використання онтологій у базі знань інтелектуальних систем опрацювання інформаційних ресурсів архітектурного профілю допомагає вирішити низку завдань методологічного та технологічного характеру, які виникають під час розроблення таких систем. Зокрема, в Україні відсутня концептуальна цілісність й узгодженість окремих прийомів та методів інженерії знань для розроблення відповідних систем у цій галузі; недостатньо кваліфікованих фахівців для розроблення програмного забезпечення архітектурного спрямування, жорсткі розроблені програмні засоби та низька їх адаптивна здатність; складно впроваджува-

ти інтелектуальні системи опрацювання інформаційних ресурсів архітектурного профілю, що зумовлено особливістю образного сприйняття дійсності (ІТ-спеціалісти сприймають світ аналітично через цифри, символи, знаки, формули та алгоритми, а фахівці з архітектурної галузі – емоційно через форми, кольоровий спектр, композиційні елементи). Все це підтверджує актуальність досліджень використання онтологій під час побудови інтелектуальних систем опрацювання інформаційних ресурсів.

Основною компонентою інтелектуальних систем опрацювання інформаційних ресурсів є база знань на основі профілюючого тематичного тезаурусу, що формується відповідно до предметної області, на яку зорієнтоване функціонування цієї системи. Традиційні методи інженерії знань (отримання знань від експерта, машинне навчання тощо) не ґрунтуються на системі вивіренних та загальноприйнятих стандартів, тому побудовані на їхній основі бази знань з часом втрачають свою функціональність через низьку ефективність під час експлуатації. Як стандарт інженерії знань використовують онтологічний інжиніринг, у результаті якого отримують онтологію бази знань. Онтологія – це детальна формалізація деякої області знань, подана за допомогою концептуальної схеми або концептуалізації. Така схема складається з ієрархічної структури термінів, зв'язків між ними, теорем та обмежень, які прийняті у певній предметній області. Пропонується зважувати елементи онтології для прийняття оптимальних рішень у певній предметній області на основі опрацювання інформаційних ресурсів та для оптимізації структури та змісту онтології. В області наукових досліджень з архітектури виникає проблема її опису та формалізації для опрацювання комп'ютерною технікою. Також виникає проблема формалізації опису навчальних дисциплін і планів навчання, формування бібліотечних каталогів [1–4]. Останнім часом широкого розповсюдження для побудови інформаційно-логічної моделі предметної області набуває онтологічний підхід. Концептуалізація – це опис множини об'єктів і понять, знань про них і зв'язків між ними. Формально онтологія складається з термінів (понять, концептів), організованих у таксономію, їхніх визначень і атрибутів, а також пов'язаних з ними аксіом і правил виведення.

**Аналіз останніх досліджень.** Під формальною моделлю онтології  $O$  розуміють трійку такого вигляду

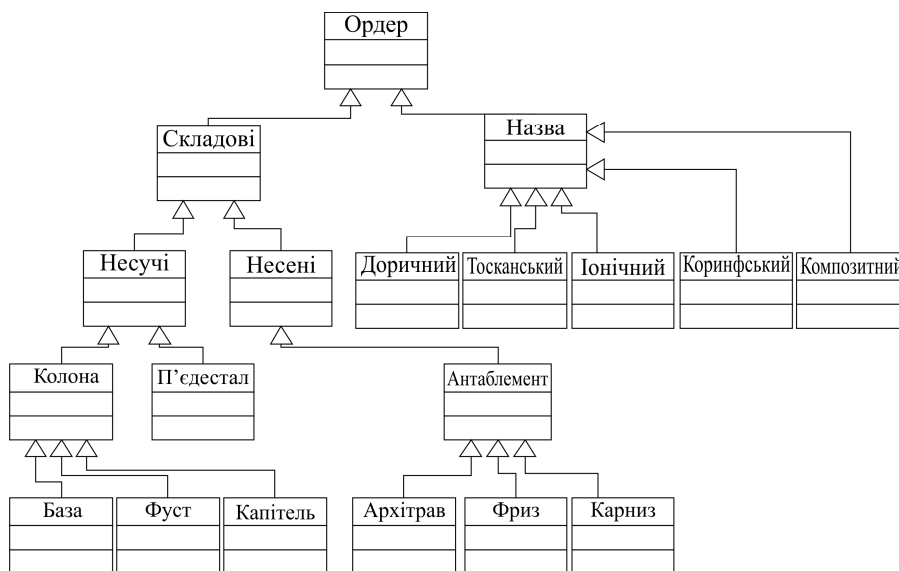
$$O = \langle C, R, F \rangle,$$

де  $C$  – скінченна множина понять (концептів, термінів) ПО, яку задає онтологія  $O$ ;  $R: C \rightarrow C$  – скінченна множина відношень між концептами (поняттями, термінами) заданої ПО;  $F$  – скінченна множина функцій інтерпретації (аксіоматизація, обмеження), заданих на концептах чи відношеннях онтології  $O$ . Зазначимо, що природним обмеженням, яке накладається на множину  $C$ , є її скінченність і непустота,  $F$  і  $R$  мають бути скінченними множинами [1, 5–7].

Онтологію можна подати у вигляді концептуального графа, де його вершини – концепти ПО, дуги вказують напрями відношень між концептами. Вершини можуть бути як інтерпретовані (визначені аксіоми понять), так й неінтерпретованими. Своєю чергою, відношення діляться на вертикальні (відношення ієрархії  $IS - A$  та агрегації) та горизонтальні. Вертикальні відношення задають таксономію понять ПО. Для задання горизонтальних відношень необхідно встановити їх область визначення (домен) та множину значень (ренг). У загальному вигляді структура онтології є набором елементів чотирьох категорій: поняття; відношення; аксіоми; окремі екземпляри. Поняття розглядаються як концептуалізації класу всіх представників певної сутності (наприклад, Басейн, Балкон, Будівля, Крамниця) чи явища (наприклад, Бароко, Ордер). Класи (або поняття) є загальними категоріями, які можна впорядкувати ієрархічно. Кожен клас описує групу інди-

відуальних сутностей (або екземплярів), які об'єднані загальними властивостями. Поняття можна поєднати відношеннями (наприклад, знаходиться, пропорція), які пов'язують класи і описують їх. Найпоширенішим типом відношення, яке використовують у всіх онтологіях, є категоризація, тобто віднесення деякого об'єкта до певної категорії [8]. Поряд із зазначеними елементами онтології в неї також входять т. зв. *екземпляри*. Це окремі представники класу сутностей або явищ, тобто конкретні елементи певної категорії (наприклад, екземпляром класу Будівля є Головний корпус Національного університету “Львівська політехніка” або Замок є Підгорецький замок).

Складові онтології підпорядковуються за своєю ієрархією. На нижньому її рівні – екземпляри, конкретні індивіди, вище йдуть поняття, тобто категорії. На рівень вище розташовані відношення між цими поняттями. Об'єднують ці всі елементи правила та аксіоми [9]. Наведемо приклад онтологічного підходу. Для побудови онтологічної моделі, насамперед необхідно визначити ієрархію понять (множину C). Приклад такої таксономії понять, яку подано за допомогою діаграми класів UML, наведено на рисунку.



Таксономія поняття Ордер у вигляді діаграми класів UML.

Під час побудови інфологічної моделі у вигляді онтології повинні залучатися фахівці з предметної області. Зазвичай вони використовують абстрагування і комбінування. Слід виділяти поняття як атомарні, які будуються з набору диференційних ознак. Основним методом побудови онтологій є використання класифікацій.

Класифікація – засіб впорядкування знань, який в об'єктно-орієнтовному аналізі визначення загальних властивостей об'єктів допомагає знайти загальні ключові абстракції й механізми, що значно спрощує архітектуру інфологічної моделі системи. На жаль, сьогодні не розроблені строгі методи класифікації й немає правил, що дають змогу виділяти класи й об'єкти, та таких понять, як “ідеальна структура класів”, “правильний вибір об'єктів”. Вибір класів є компромісним рішенням. Мета класифікації – визначити загальні властивості об'єктів. Класифікуючи, поєднуємо в одну групу об'єкти, які мають однакову будову або поведінку [10].

**Виділення проблем.** Під час побудови онтології може виникнути низка проблем. Під час класифікації існує можливість появи неоднозначностей залежно

від того, якій диференційній ознаці віддавати перевагу. Вибір того чи іншого рішення часом досить складно обґрунтувати. Найважливіше вибрати який-небудь підхід і дотримуватися його упродовж всієї роботи, тому варто задуматися над проблемами і можливими універсальними методами їх усунення ще до початку прийняття рішень. Під час створення концептів виникає *проблема примітивності*. Майже вся семантика та подання знань ґрунтуються на композиційній гіпотезі: можна визначити обмежений набір одиничних сутностей (атомів), а всі інші (молекули) подати як комбінацію (композицію) атомів. При цьому виникає питання: як багато таких атомів необхідно? У сучасних дослідженнях можна знайти два різних підходи: економний і неощадливий.

Економний підхід полягає у створенні малої кількості елементарних концептів, семантично простих, за допомогою яких можна пояснити складніші поняття. Тоді легко виявити зв'язність понять, однак складно утворювати нові сутності. Семантичні примітиви – це лексичні одиниці, які виражають елементарні, базові значення. Кількість таких одиниць не перевищує 100. Неощадливий підхід дає змогу створювати будь-яку кількість індивідуальних сутностей онтології. Ця кількість може варіюватися від 10 до 100 000 і більше. Тут важко визначати зв'язність понять. Тобто такий підхід супроводжується, по суті, відмовою від композиційної гіпотези, але він має й перевагу, а саме, легкість формування складних сутностей.

Важко сказати, який підхід кращий, все залежить від того, як багато сутностей або наскільки складна предметна область. Підсумовуючи, слід сказати, що економного підходу дотримуються в основному формалісти, а користувачі схильються до неекономного [4, 5].

**Аналіз отриманих наукових результатів.** Побудова онтології складний процес, оскільки незавжди легко виділити поняття та диференційні ознаки. Існує кілька варіантів дій, що залежать від конкретних завдань і вихідного матеріалу. Є можливість збору елементів для онтології безпосередньо. За такого підходу спочатку збирають і класифікують поняття, потім визначають відношення між ними. Універсальна вимога до онтології полягає в тому, що загальна її структура має бути зрозумілою і повинна існувати можливість її багаторазового використання. Інші вимоги такі: зрозумілість: має легко передавати суть предметної області та бути об'єктивною; послідовність: у ній повинні міститися твердження, які не суперечать одне одному, ієрархії понять (які зв'язані відношеннями), екземпляри; можливість розширення: введення нових елементів; мінімальна ступінь спеціалізації онтології: небажаність повного її підпорядкування конкретній задачі, що може ускладнити її подальше використання в інших задачах. Не можна стверджувати, що цей список вимог до онтології є вичерпним, але він може допомогти під час прийняття тих чи інших рішень для побудови інфологічної моделі ПО.

Існують формалізовані і докладні описи стандартів для онтологій. Наприклад, Expert Advisory Group on Language Engineering Standards, International Standard for Language Engineering, The Language Technology Resource Center. Використання стандартів має стати запорукою того, що створений інформаційний ресурс легко впроваджуватиметься у вже існуючі системи і враховуватиме всі особливості інформаційних технологій.

Основною проблемою такого підходу є правильна класифікація понять, тобто їхнє місцезоташування в онтології.

Збір знань, релевантних ПО архітектури, які містяться в онтології, здійснювали на основі короткого російсько-українського словника [11].

Ієрархію понять ПО архітектури почали із “сутність” (essence), яке поділяється на абстрактні (abstract) і фізичні (physical) поняття. Поняття Ордер належить до фізичних (див. рисунок). Так, Ордер складається з *Несучих (Колона, П'є-*

дестал) та Несених (Антаблемент). Колона (Пілястра), своєю чергою, це База, Фуст (Стовбур) та Капітель, Антаблемент містить Архітрав, Фриз та Карниз. Записано аксіоми термінів словника та атомарні висловлювання про екземпляри понять.

Допускається задавати класи понять, накладаючи обмеження на властивості їх екземплярів. Такі обмеження можна згрупувати у три основні категорії:

- 1) кванторні обмеження (існування, загальності);
- 2) обмеження кількості допустимих значень (мінімум  $\leq$ , якраз  $=$ , максимум  $\geq$ );
- 3) обмеження типу “може приймати значення з множини”.

Обмеження існування описує клас екземплярів, які мають принаймні один зв’язок вказаного семантичного значення з екземпляром вказаного класу. Тоді квантор існування застосовують до множини зв’язків екземпляра (а не до множини екземплярів класу). Квантор існування свідчить, що цей клас має лише ті екземпляри, множина зв’язків яких містить конкретний зв’язок:  $\{x | \exists r, r(x, y)\}$ . Квантор загальності свідчить, що в класі є лише ті екземпляри, множина всіх зв’язків яких містить вказаний тут явно винятковий перелік зв’язків:  $\{x | \forall r, R_i \in r, R_1(x, y_1) \wedge R_2(x, y_2) \wedge \dots \wedge R_n(x, y_n)\}$ . Для задання логічних правил онтології використано SWRL-правила, а саме:

поверх (?x)  $\wedge$  поверх (?y)  $\wedge$  містить.ордер (?x, Тосканський)  $\wedge$   
містить.ордер (?y, Іонічний)  $\rightarrow$  Поверх\_вищий (?y, ?x).

Це правило означає, що якщо на якомусь поверсі є *Тосканський* ордер, а на іншому у тій самій будівлі *Іонічний*, то поверх з *Іонічним* знаходиться вище. Аналогічні правила записані для інших ордерів:

поверх (?x)  $\wedge$  поверх (?y)  $\wedge$  містить.ордер (?x, Доричний)  $\wedge$  містить.ордер  
(?y, Іонічний)  $\rightarrow$  Поверх\_вищий (?y, ?x);

поверх (?x)  $\wedge$  поверх (?y)  $\wedge$  містить.ордер (?x, Іонічний)  $\wedge$  містить.ордер  
(?y, Коринфський)  $\rightarrow$  Поверх\_вищий (?y, ?x);

поверх (?x)  $\wedge$  поверх (?y)  $\wedge$  містить.ордер (?x, Доричний)  $\wedge$  містить.ордер  
(?y, Коринфський)  $\rightarrow$  Поверх\_вищий (?y, ?x);

поверх (?x)  $\wedge$  поверх (?y)  $\wedge$  містить.ордер (?x, Тосканський)  $\wedge$   
містить.ордер (?y, Коринфський)  $\rightarrow$  Поверх\_вищий (?y, ?x).

Для задання функцій інтерпретації використовують описову логіку (англ. Description logics, DL, іноді ще називають дескриптивна логіка). Це мова подання знань, яка дає змогу описувати поняття предметної області у формалізованому вигляді. Вона поєднує в собі, з одного боку, логічні вирази, а з іншого – обчислювальні властивості такі, як розв’язність і відносно невисока обчислювальна складність, що дає змогу легко її застосовувати на практиці. Тобто описова логіка є компромісом між записом логіки формалізованими виразами і їх розв’язністю (логічне виведення нових знань). Її можна розглядати як розв’язні фрагменти логіки предикатів, синтаксично ж вони близькі до модальних логік.

Описова логіка оперує поняттями концепт і ролей, одномісний предикат (або множина, клас) та двомісний предикат (або бінарне відношення). Інтуїтивно концепти використовують для опису класів деяких об’єктів, наприклад, Будівлі, Фасаду, Ордера. Ролі використовують для опису двомісних відношень між об’єктами, наприклад, на множині осіб є двомісне відношення “Ордер X є\_на\_будівлі Y”, де як X і Y можна підставляти конкретні об’єкти. За допомогою мови DL можна формулювати твердження загального вигляду та приватного виду – про конкретні об’єкти.

Мовою описової логіки набір тверджень загального вигляду називають TBox, набір тверджень приватного виду – ABox, а разом вони складають т. зв.

базу знань або онтологію. Як тільки онтологія побудована, виникають питання про те, як можна отримувати знання, що виводяться з наявних в онтології знань, чи можна це робити програмно і які відповідні алгоритми. Всі вони вирішуються теоретично в науці *описова логіка*, а практично вже реалізовані у прикладних програмних системах (reasoners), які дають змогу автоматизовано виводити знання з онтологій та виконувати інші операції з онтологіями.

Приклад аксіом (TBox) :

Ордер  $\equiv$  Тосканський  $\cup$  Доричний  $\cup$  Іонічний  $\cup$  Коринфський  $\cup$  Композитний  
Композитний ордер  $\equiv$  Елемент.Іонічний  $\cap$  Елемент.Коринфський  
Антаблемент  $\subseteq$  Ордер.

Приклад тверджень (ABox) :

Колізей: 1поверх.Тосканський  $\cap$  2поверх. Іонічний  $\cap$  3поверх. Коринфський  
Каплиця Воїмів: 1поверх. Іонічний  
Успенська церква: 1поверх.Тосканський  
Вежа Корнякта: 1поверх.Тосканський  $\cap$  2поверх.Тосканський  $\cap$   
3поверх.Тосканський  $\cap$  4поверх.Тосканський.

Проаналізувавши особливості архітектури міста Львова, побудуємо відповідні множини елементів порталів XVI–XVII ст. зі застосуванням архітектурних ордерів: тосканського, доричного, іонічного, коринфського та композитного [12].

Тосканський ордер = {Палаццо Корнякта; Монастир Бенедиктинок; фасади будинків на вул. Вірменській 30; фасади будинків на пл. Ринок 2 (боковий); Королівський арсенал};

Доричний ордер = {фасади будинків на пл. Ринок 2, 28; фасади будинків на вул. Руській, 10; портал будинку на пл.Ринок, 9; фасад на пл. Ринок 4; фасад неіснуючого будинку Піноццего};

Іонічний ордер = {Портали Чорної кам'яниці, фасади будинків на вул. Вірменській 20; фасади на пл. Ринок 28 (інтер'єр); фасади на пл. Ринок 6 (галерея); Портал галереї палацу Корнякта};

Коринфський ордер = {Фасади на пл. Ринок 6; каплиця Трьох Святителів; будинок С. Шольца};

Композитний = {}.

Приклад аналізу отриманих множин елементів порталів Львова XVI–XVII ст.:

Будинок на пл. Ринок 28: 1поверх.Доричний  $\cap$  2поверх.Іонічний.

Проаналізувавши особливості архітектурного ансамблю замку в с. Підгірці Львівської області XVII–XVIII ст., побудуємо твердження [13]:

Підгорецький замок: 2поверх.Доричний  $\cap$  3поверх.Коринфський;

Каплиця Підгорецького замку: 1поверх.Коринфський  $\cap$  2поверх.Композитний.

Ці два твердження констатують факти про наявність певних ордерів та коротко описують унікальні особливості двох пам'яток архітектури України XVII–XVIII ст., але не розкривають всю багатогранність та унікальність архітектурної композиції, згідно з працею [13]. Це обумовлює необхідність в подальшому науковому дослідженні основних принципів та особливостей побудови онтології для архітектурних термінів, яка б змогла передати всю глибину тектонічного ансамблю будь-якої існуючої архітектурної композиції.

Сьогодні онтологія архітектурних термінів перебуває на стадії розроблення. У майбутньому передбачені роботи з її верифікації й написання запитів.

## ВИСНОВКИ

Запропоновано метод побудови онтології архітектурних термінів для подання і розподілу знань загального словникового запасу і підтримки інтелектуальних запитів у базах даних з архітектури, а також автоматизованого обміну даними та інтеграції навчальних програм з архітектури. Апробовано її й для пошуку реле-

вантної інформації, а саме, для вибірки сторінок у мережі Інтернет із синтаксично різними, але семантично однаковими архітектурними термінами.

1. *Gruber T.* A translation approach to portable ontologies // *Knowledge Acquisition*. – 1993. – № 5 (2). – P. 199–220.
2. *Guarino N.* Formal Ontology, Conceptual Analysis and Knowledge Representation // *Int. J. of Human-Computer Studies*. – 1995. – № 43 (5–6). – P. 625–640.
3. *Sowa J.* Conceptual Graphs as a universal knowledge representation // *Semantic Networks in Artificial Intelligence, Spec. Issue of An International Journal Computers & Mathematics with Applications*. – Part 1. – 1992. – **23**, № 2–5. – P. 75–95.
4. *Montes-y-Gómez M., Gelbukh A., López-López A.* Comparison of Conceptual Graphs [Електронний ресурс]// *Lecture Notes in Artificial Intelligence*. – 2000. – Vol. 1793. – Springer-Verlag. – Режим доступу до ресурсу: <http://ccc.inaoep.mx/~mmontesg/publicaciones/2000/ComparisonCG>.
5. *Литвин В. В.* Бази знань інтелектуальних систем підтримки прийняття рішень. – Львів: Вид-во Львівської політехніки, 2011. – 240 с.
6. *Литвин В. В.* Мультиагентні системи підтримки прийняття рішень, що базуються на прецедентах та використовують адаптивні онтології // *Радіоелектроніка, інформатика, управління*. – 2009. – № 2 (21). – С. 120–126.
7. *Литвин В. В., Оборська О. В., Вовнянка Р. В.* Метод моделювання процесу підтримки прийняття рішень у конкурентному середовищі // *Математичні машини й системи*. – 2014. – № 1. – С. 50–57.
8. *Литвин В. В.* Підхід до побудови інтелектуальних систем підтримки прийняття рішень на основі онтологій // *Проблеми програмування*. – 2013. – № 4. – С. 43–52.
9. *Досин Д. Г., Литвин В. В., Вовнянка Р. В.* Комп'ютерна система автоматизованої розбудови базової онтології Crocus // *Електротехнічні та комп'ютерні системи*. – 2014. – № 13 (89). – С. 135–143.
10. *Lytvyn V., Horyak M., Oborska O.* Method of automated development and evaluation of ontologies' qualities of knowledge bases // *Applied computer science*. – 2014. – **10**, № 4. – P. 92–97.
11. *Безродний П. П.* Архітектурні терміни // *Короткий російсько-український словник: довідн. пос.* – К.: Вища школа, 1993. – 272 с.
12. *Білінська О. Б.* Особливості застосування ордерних членувань у порталах споруд Львова кін. XVI–XVII ст. [Електронний ресурс]. – Режим доступу: <http://ena.lp.edu.ua:8080/bitstream/ntb/9606/1/21.pdf>.
13. *Remeshylo-Rybychynska O.* Kształtowanie się i obecny stan spuszczony architektonicznej i artystycznej wsi Podhorce w odwodzie Lwowskim // *Monument t.II. Studia i materiały ośrodka badań dokumentacji zabytkow.* – Warszawa: KOBiDZ, 2005. – S. 443–567.

Одержано 18.02.2016