

## АРХЕОЛОГІЯ

УДК 902

DOI <https://doi.org/10.32838/2663-5984/2021/2.20>**Бочковський С.О.**

Донецький національний університет імені Василя Стуса

### МЕТОДИ ПОБУДОВИ БАЗИ ДАНИХ В АРХЕОЛОГІЧНИХ ДОСЛІДЖЕННЯХ НА ПРИКЛАДІ КУРГАННИХ ПОХОВАНЬ ЯМНОЇ КУЛЬТУРИ ПРАВОБЕРЕЖНОГО УКРАЇНСЬКОГО ЛІСОСТЕПУ

*Оскільки археологічні дослідження здебільшого пов'язані з дослідженням великих обсягів інформації, нагальною проблемою є організація та виокремлення нової інформації із цих масивів. Задля вирішення цієї проблеми найбільш доречним є використання реляційних баз даних, побудованих за архітектурою SQL. Для вітчизняних дослідників такі методи роботи виявляються досить інноваційними, хоча в зарубіжній практиці такий спосіб організації великих комплексів інформації розроблений ще 20 років тому та широко використовуються. Автором статті запропоновано варіант методики створення, наповнення та роботи з такими базами даних на прикладі власних наробітків, сформованих на базі роботи з археологічним матеріалом Ямної культури Правобережного Українського Лісостепу. За мету автор ставить окреслення методики роботи із системами керування базами даних, для реалізації якої пропонує передусім огляд найбільш удалого програмного забезпечення, окреслення його можливостей, постановку базових рекомендацій для створення бази даних, виявлення потенційних можливостей для впровадження запропонованої автором методики для майбутніх досліджень. Для наочності автор демонструє наведені методи формування бази даних на прикладі створення бази даних поховань Ямної культури Правобережного Українського Лісостепу, яку він поділив на три концептуальні модулі: курганні групи, поховання, артефакти, які, своєю чергою, пов'язані реляційними таблицями, котрі своїми зв'язками формують цілісну систему інформаційних кластерів. Підсумовуючи, автор окреслює інформаційний потенціал наведеної бази даних та конкретні потреби управління даними збору археологічних досліджень. На основі результатів дослідження була створена концептуальна інформаційна модель.*

**Ключові слова:** археологічний матеріал, системи керування базами даних, реляційні таблиці, Ямна культура, кургани.

**Постановка проблеми.** Сучасні археологічні дослідження, з одного боку, значно багатші в плані накопиченого матеріалу порівняно з тим, із чого починалася археологія як наука, а з іншого – така насиченість ставить перед дослідниками нові виклики, зокрема акумуляція, систематизація, та компарація археологічного матеріалу на даний момент є наріжним каменем сучасних досліджень кабінетної археології. На допомогу сучасним дослідникам у вирішенні подібних завдань за сучасного рівня діджиталізації приходять різноманітні міждисциплінарні підходи, зокрема математичні методи компаративістики та систематики, для використання яких необхідно ланкою стає дата-аналіз через системи керування базами даних (СКБД), основні архітектури яких було роз-

роблено ще в 70-ті роки минулого століття. Бази даних, створені на таких платформах, містять великі колекції інформації, яка може допомогти археологам приймати обґрунтовані рішення, інтерпретувати і перевіряти гіпотези на основі відновлених даних. Запровадження таких баз у кабінетну роботу археологів можна було б використовувати, щоб задати більш складні питання про археологічний матеріал.

Велика частина розроблення археологічної бази даних обертається навколо прийняття рішення про типи інформації, яку дослідник хоче зберегти, і атрибути, які він створив для збору й аналізу даних, а також відношення між цими атрибутами. Очевидно, що базові рішення про те, які характеристики записувати або не писати

для зібраних артефактів (наприклад, тип кераміки, розміри артефактів і походження), мають довгострокові наслідки для того, як набір даних може використовуватися в подальшому. Менш очевидний спосіб, яким артефакти повинні бути пов'язані один з одним абстрактно й як вони повинні бути записані по відношенню до інших археологічних аналітичних одиниць. Рішення щодо визначення цих відносин мають вирішальне значення для подальшого аналізу даних і створення успадкованих даних. Нарешті, дизайн бази даних повинен бути спрямований на мінімізацію потенційних помилок за введення і подальшої зміни даних, а також на максимальну аналітичну гнучкість, включаючи можливість використання даних іншими, більш пізніми дослідниками. З огляду на ці факти, планування архітектури бази даних має бути важливим аспектом загального дизайну і планування дослідження.

**Аналіз останніх досліджень і публікацій.** Метод розроблення та аналізу баз даних у нашій країні розвинутий слабо, хоча за останні роки починає дещо приростати новими публікаціями, розробленими на базі власного досвіду роботи з базами даних. Таку роботу в 2016 р. представили А.В. Борисов та О.В. Мандіга, співробітники Інституту археології АН України [7]. У своїй роботі вони на прикладі власних розробок у царині дата-аналітики привели вдалий приклад архітектури бази даних для археологічних досліджень.

Більш ґрунтовною є робота С.А. Васильєва [8], який у своїй статті, присвяченій СКБД «Археограф», описує загальні методи роботи з базами даних на базі вищезазначеного ПЗ та способи імплементації даних у картографуванні.

Але найбільших успіхів у формуванні, обґрунтуванні та впровадженні систем керування базами даних досягли дослідники в Європі та Америці, що очевидно у зв'язку зі значно вищим рівнем діджиталізації. Зокрема, Шенон МакПеррон, автор найбільш ґрунтовної праці з методик роботи з базами даних в археологічних дослідженнях [4]. Хоча його робота, урахувавши оптимізацію архітектур сучасних СКБД, є дещо застарілою.

Не менш корисною з методологічного погляду є робота Гаррі Локка, [3], яка у цілому присвячена загальним методам використання комп'ютерної техніки в історичних дослідженнях, але й базам даних у ній присвячено значний інформаційний кластер.

Але попри досягнення сучасної науки всі вищезазначені роботи не можуть запропонувати єдиного уніфікованого алгоритму побудови баз

даних, що, своєю чергою, сповільнює інформаційну мобільність та позбавляє наукову спільноту можливості створення грандіозних проєктів співпраці.

**Постановка завдання.** Зважаючи на вищезазначене, на базі концептуальних засад формування та використання баз даних в археологічних дослідженнях автор ставить за мету окреслити методику створення та використання реляційної бази даних, придатної для використання в археологічних дослідженнях. Для виконання поставленої мети проміжними завданнями є:

- характеристика основних принципів формування бази даних;
- розроблення лаконічної структури реляційних таблиць у рамках бази;
- виявлення потенційних можливостей створеної бази для отримання нової інформації.

**Виклад основного матеріалу дослідження.** У першу чергу для побудови бази даних для дослідника є головним завданням вибору найбільш удакої системи керування базою даних (DBMS) у даному разі автором вибрано найбільш розповсюджену систему для побудови реляційних баз даних MySQL від компанії Oracle, яка має такі переваги:

- Простота у використанні. MySQL досить легко інсталується, а наявність безлічі плагінів і допоміжних додатків спрощує роботу з базами даних.
- Великий функціонал. Система MySQL володіє майже всім необхідним інструментарієм, який може знадобитися в реалізації практично будь-якого проєкту.
- Безпека. Система спочатку створена так, що безліч вбудованих функцій безпеки в ній працюють за замовчуванням.
- Масштабованість. Будучи досить універсальною СКБД, MySQL рівною мірою легко може бути використана для роботи і з малими, і з великими обсягами даних.
- Швидкість. Висока продуктивність системи забезпечується за рахунок спрощення деяких використовуваних у ній стандартів.

Загалом система функціонує за принципом побудови реляційних таблиць, в яких дані акумулюються та систематизуються. Реляційна база даних – це набір даних із зумовленими зв'язками між ними. Ці дані організовані у вигляді набору таблиць, що складаються із стовпців і рядків. У таблицях зберігається інформація про об'єкти, представлені в базі даних. У кожному стовпчику таблиці зберігається певний тип даних, у кож-

ному осередку – значення атрибута. Кожна строка таблиці являє собою набір пов'язаних значень, що відносяться до одного об'єкта або сутності.

Інструменти моделювання даних, які підтримують поділ між концептуальною та логічною моделями, повинні забезпечувати логіку для перетворення конструкцій ER у реляційні таблиці та формування цих таблиць у реляційну модель. Результати прямого перетворення можуть бути чистою логічною моделлю, яка не включає таблиці для конкретної реляційної системи управління базами даних (СУБД).

Фізичне проектування бази даних починається із заданої реляційної моделі, яка є визначенням набору таблиць та відповідних стовпців. Завдання фізичного проектування бази даних полягає у виконанні вимог до продуктивності набору програм шляхом оптимізації використання систем управління базами даних. Ключові напрями включають оптимізацію конфігурації індексу, розміщення даних та розподіл пам'яті. Одним із найбільших впливів на ефективність програми є вибір індексів. Для операційних баз даних індекс, що поліпшує ефективність пошуку даних, може погіршити продуктивність для всіх видів оновлень, оскільки витрати на обслуговування індексу повинні оплачуватися за кожне оновлення. Оскільки у сховищі даних не так багато дій з оновлення, кількість індексів не викликає особливого занепокоєння. Вибір оптимальних індексів для даного набору таблиць (конфігурація індексу) є нетривіальною проблемою, яка вимагає компромісів між різними видами операцій із базами даних (операціями пошуку, транзакціями оновлення та утилітами). Ураховуючи індекс, потрібно визначити його властивості, наприклад упорядкування стовпців для багатостолбкових індексів, а також те, чи має індекс бути кластеризацією, секціонуванням чи упорядкуванням. Поки вибір індексу є найважливішою та найскладнішою проблемою у фізичному проектуванні бази даних, важливим є також розміщення даних та оптимізація простору. Потрібно розподілити оптимальний простір таблиці для даної таблиці та вибрати, чи є простий, сегментований або розділений табличний простір найбільш доречним. Виділіть достатньо місця для збереження властивостей кластеризації, не витрачаючи занадто багато місця. Вибір індексу, розміщення даних та оптимізація простору передбачають стабільність структури реляційної схеми. Однак фізичний дизайн бази даних також може змінити реляційну схему. Вам може знадобитися розділити таблиці на різні роз-

діли, щоб скористатися перевагами одночасних операцій, або об'єднати таблиці у разі, коли різні таблиці часто об'єднуються, так що продуктивність покращується, якщо одна об'єднана таблиця представляє збережене об'єднання. Також можуть бути прийняті рішення щодо налаштування параметрів часу роботи СУБД. Більшість додатків в археології використовують традиційне моделювання ER для створення та ведення оперативних баз даних. Ці бази даних, як правило, дуже спеціалізовані і розроблені для підтримки дуже конкретних вимог програми.

Останніми тенденціями є створення сховищ даних. Ці нові мегабази даних дають змогу кінцевим користувачам отримувати доступ до інформації на основі даних, які раніше були недоступні для них в одному місці. Для спеціаліста з обробки даних найновішим завданням є створення оптимізованої реляційної бази даних, яка задовольняє набагато інший набір вимог, колекцій та публікацій. Однак хоча колекція може бути згадана в декількох посиланнях, опубліковані деталі для цієї колекції повинні бути постійними, оскільки це якості, властиві колекції. Тому існує необов'язковий взаємозв'язок «один на один» між колекціями та опублікованими деталями, а також необов'язковий взаємозв'язок «багато-до-багатьох» між колекціями та посиланнями. У межах трьох категорій даних існує також безліч зв'язків між деякими категоріями, їхніми атрибутами. Наприклад, колекція, можливо, була зібрана в результаті низки різних методів відновлення, випадкового підбору, можливо, систематичних польових прогулянок та розкопок.

Таким чином, існує «один-до-багатьох» взаємозв'язок між колекціями та методами відновлення. Отже, всередині даних існує низка різних взаємозв'язків між елементами:

- а) необов'язкові зв'язки, де дані можуть бути присутніми, а можуть і не бути (опубліковані дані, парафія тощо);
- б) зв'язки «один-до-багатьох»;
- в) відносини «багато-до-багатьох».

Взаємозв'язки між елементами даних, що описують колекцію, можуть бути формалізовані як модель даних, що деталізує ці зв'язки. Зазвичай створення моделі супроводжується реалізацією структури даних, розробленням інтерфейсу та тестуванням. Ця блок-схема наголошує на ранньому плануванні, методичному завершенні етапів та тестуванні як невід'ємній частині процесу. Однак цей суворий лінійний процес рідко трапляється на практиці.

Проектування археологічних баз даних повинно відповідати таким вимогам:

а) високий ступінь перевірки даних із використанням посиальної цілісності;

б) контроль користувачем перевірки достовірності даних;

в) ретельне представлення всіх можливих взаємозв'язків у системі запису. Вирішення, які зміни впровадити, є непростим завданням для того, щоб створити надійний дизайн перед упровадженням. Однак у деяких випадках зміни у структурі даних повинні бути внесені після впровадження. Такі зміни витрачають час розробників через високий ступінь цілісності посилань. Проста зміна ключового поля може багато разів відбиватися у всій структурі бази даних. Після створення ядра бази даних необхідний дружній інтерфейс, щоб надати користувачам доступ та можливість уведення даних. Основною проблемою створення інтерфейсу є створення простого чистого набору форм, які дають доступ до складних набір таблиць. Уведення даних могло бути досягнуте лише завдяки «високому рівню» знань системи баз даних та більш глибоким знанням систем археологічного запису. Це обмежує випадкового користувача, але означає, що введення даних стає більш розумним процесом, ідеальним для навчання базових навичок, необхідних для створення археологічних архівів.

У базі, наведеній як приклад, ця система реалізована за трьома основними модулями.

У перший модуль входять таблиці, в яких подано інформацію про локалізацію археологічних комплексів, які, своєю чергою диференціюються на окремі кургани з їх детальною характеристикою. Він, своєю чергою, поділений на три основні таблиці, які пов'язані між собою лінійно. У першій таблиці сконцентровано інформацію про локалізацію курганної групи, її розташування відносно сучасних адміністративно територіальних одиниць (населений пункт, район, область). Друга таблиця містить інформацію про кількість курганів у групі, дату та мету розкопу. Третя таблиця містить основну інформацію про курган: його висоту та діаметр, форму, стан збереженості та кількість поховань, наявність досіпок. Перший інформаційний модуль у цьому разі є відправною точкою для подальшої роботи з даними, оскільки надає перші кількісні дані, інформацію про локалізацію та концентрацію археологічних пам'яток у регіоні. І попри низьке інформаційне навантаження формує основний контекст для подальшої роботи з базою.

Другий фундаментальний модуль формується таблицею характеристики поховань, в якій подано найбільш широкий інформаційний зріз:

- форма поховальної камери;
- її ширина, довжина та глибина;
- інформація стосовно первинності поховання;
- орієнтація небіжчика стосовно сторін світу;
- якщо можливо, статевікові дані небіжчика;
- поза небіжчика;
- наявність поховальної конструкції;
- рештки поховального ритуалу;
- наявність супровідного інвентарю.

Дана таблиця має найбільшу кількість стовпців у зв'язку з тим, що розглядає найбільш насичений інформацією масив даних, згрупований за єдиним принципом, а саме характеристикою поховання. Дана таблиця займає проміжне положення у системі реляцій та несе найбільший інформаційний потенціал для подальшого дослідження

Третій модуль формують таблиці, пов'язані з другим модулем, та являють собою таблиці, де представлена інформація про поховальний інвентар, конструкції та елементи поховального обряду. Усього в модулі вісім таблиць, які можна умовно поділити на дві групи: п'ять таблиць відображають поховальний інвентар, згрупований за матеріалом походження (камінь, кераміка, кістка, метал, рештки тканини), та три таблиці, що відображають елементи поховального обряду (органічна підстилка, природні барвники, поховальна конструкція). У кожній таблиці детально відображено параметри знахідки, її позицію в поховальній ямі та форму. Даний модуль поза контекстом надає значно менше інформації, ніж попередні, але суттєво доповнює їх найбільш репрезентативною інформацією та завершує картину археологічного комплексу.

**Висновки.** Таким чином, представлений набір критеріїв, алгоритмів, та принципів формування баз даних, дизайн реляційних баз даних порушують багато питань щодо модельованої ситуації у «реальному світі». Будь-які незначні зміни в системі запису на паперовій основі незабаром виявляються в базі даних, яка забезпечує цілісність даних. Однак цілісність даних – потужний інструмент дослідження. У процесі проектування археологічних баз даних важливо мати повне розуміння існуючої археологічної системи перед будь-якою спробою створити більш досконалу модель.

## Список літератури:

1. Banning, E.B. *The Archaeologist's Laboratory: The Analysis of Archaeological Data*. New York: Kluwer Academic, 2000. 245 p
2. Gardin, Jean-Claude. "Archaeological Discourse, Conceptual Modeling and Digitalization: An Interim Report of the Logicist Program." Report to the CAA, 2002. 137 p
3. Lock, Gary. *Using Computers in Archaeology: Towards Virtual Pasts*. New York: Routledge, 2003. 98 p
4. McPherron, Shannon and Harold Dibble. ed. *Using Computers in Archaeology*. Mayfield: McGraw-Hill, 2002. 233 p
5. Richards, Julian and Damian Robinson. ed. *Digital Archives from Excavation and Fieldwork: A Guide to Good Practice*. Oxford: Oxbow Books, 2000. 347p
6. Ryan, Nick. "Databases." *Internet Archaeology*. Oxford: Oxbow Books. 2004. 124 p
7. Борисов А.В., Мандіга О.В. Досвід розробки і реалізації бази даних археологічних пам'яток для дослідження систем розселення. *Археологія і давня історія України*. 2016. Вип. 3(20).
8. Васильев С.А. «Археограф»: система описания археологических памятников и вывода данных в ГИС. *Археология и компьютерные технологии: представление и анализ археологических материалов*. Ижевск, 2005. 127 с
9. Дейт К.Дж. Введение в системы баз данных / пер. с англ. ; 8-е изд. Москва, 2005. 1328 с.
10. Янишевский Б.Е., Киселев Д.И. База данных памятников археологии GRAVE. *Компьютеры в археологии* : материалы конф. «Опыт компьютерной обработки археологических материалов». Москва, 1996. С. 109–113.

**Bochkovskiy S.O. METHODS TO CONSTRUCT DATABASE FOR THE ARCHAEOLOGICAL STUDY ON THE CASE OF BURIAL MOUNDS OF PIT CULTURE OF THE RIGHT-BANK FOREST-STEPPE OF UKRAINE**

*Archaeological research is mostly associated with the study of large amounts of information by an urgent problem in the organization and the extraction of new information from these arrays. To solve this problem, it is most appropriate to use relational databases built on the SQL architecture. For domestic researchers, such methods are quite innovative, although in foreign practice, this method of organizing large sets of information was developed 20 years ago, and is widely used. The author of the article proposes a variant of the method of creating, filling and working with such databases on the example of own developments, formed on the basis of work with archaeological material of the Pit culture of the Right-Bank Ukrainian Forest-Steppe. The author aims to outline the methodology of working with database management systems, for the implementation of which offers first of all, an overview of the most successful software, outlining its capabilities, setting basic recommendations for creating a database, identifying potential opportunities for implementing the author's methodology for future research. For clarity, the author demonstrates the above methods of database formation on the example of creating a database of burials of the Pit Culture of the Right-Bank Ukrainian Forest-Steppe, which the author divided into three conceptual modules: burial mounds, burials, artifacts. Which in turn are connected by relational tables, which by their connections form a holistic system of information clusters. Summing up, the author outlines the information potential of the database, and outlines the specific needs of data management data collection of archaeological research. Based on the results of the study, a conceptual information model was created.*

**Key words:** *archaeological material, database management systems, relational tables, pit culture, mounds.*