

УДК 001.103:004.451.5
DOI: 10.31866/2616-7654.8.2021.247590

**ПОНЯТТЯ «ДАНИ
ДОСЛІДЖЕНЬ»:
ВИДИ ТА ТИПИ ДАНИХ
ДОСЛІДЖЕНЬ
У КОНТЕКСТІ ПРАКТИКИ
УПРАВЛІННЯ ДАНИМИ**

*Світлана Чуканова,
завідувачка сектору Наукової бібліотеки
Національного університету
«Києво-Могилянська академія»,
кандидатка педагогічних наук (Київ, Україна)
e-mail: chukanovas@ukma.edu.ua
ORCID: 0000-0002-5717-5050*

В умовах стрімкого розвитку концепції «Відкритої науки», кількісного зростання отриманих під час досліджень даних посилюється наукова увага до практики управління даними досліджень (research data management), що актуалізує завдання визначення поняття «дані досліджень», а також виявлення видів і типів даних досліджень у рамках практики управління ними, обґрунтування і висвітлення специфіки таких даних.

Основу методологічного інструментарію дослідження становлять термінологічний метод, застосування якого було обумовлене потребою виявлення актуальних тлумачень поняття «дані досліджень», а також аналіз репозитаріїв для даних із різних галузей науки, індексованих реєстром [re3data.org.](https://www.re3data.org/), за загальними представленими у реєстрі напрямками, а саме: описи репозитаріїв включно з інформацією про типи даних, які депонуються науковцями та кураторами даних.

Проведений аналіз дав змогу визначити дані дослідження як матеріали, отримані і зібрані для обґрунтування наукових результатів дослідження у будь-якій галузі та у будь-якій формі: числовій, текстовій, комп'ютерного коду тощо, а також виявити типи даних, характерні для різних галузей науки, що, зі свого боку, надало змогу зробити висновки стосовно наявних форматів даних, найбільш розповсюджених серед як природничих, так і гуманітарних дисциплін: текстові, числові та графічні формати. В результаті проведеного аналізу з'ясовано, що даними досліджень можуть вважатись текстові, числові, програмні, архівні, графічні та інші об'єкти (файли), що слугують підґрунтям дослідження і фактографічною базою для наукових висновків у конкретній галузі науки. Виявлено, що тип даних напряму залежить від характеру дослідження та особливостей дисципліни чи галузі здійсненого дослідження.

Ключові слова: дані досліджень, репозитарії для даних, типи даних, формати даних, депонування даних.

ВСТУП

Сучасні наукові дослідження неможливо уявити без застосування даних. На сьогодні дані є найбільш вагомою аргументацією будь-якого дослідження. Наявність сучасних технологій та інструментів збору й обробки даних сприяє тому, що дані накопичуються в значних обсягах. Це явище американська дослідниця Крістін Боргман (Borgman, 2007) слідом за своїми колегами – фахівцями з управління даними досліджень – називає «data deluge» (тобто «потоп із даних») (с. 6). За таких умов очевидно, що потрібно удосконалювати наявну практику управління даними. Для того, щоб практика управління даними розвивалась відповідно до вимог часу, передусім необхідно окреслити, що таке дані дослідження та які типи і види цих даних існують, адже кожній галузі науки та кожній окремій спеціальності притаманні різні типи даних.

ТЕОРЕТИЧНЕ ПІДГРУНТЯ

З метою вивчення та удосконалення практики управління даними досліджень необхідно дати чітке визначення поняттю «дані досліджень».

Відповідно до глосарію CASRAI, поняття «дані» трактується як «факти, обчислення, записи, нотатки чи спостереження про світ, зібрані вченими та іншими особами, з мінімумом контекстної інтерпретації» ("Research data management glossary", n.d.) (*тут та далі переклад з англійської мови наш – С. Ч.*).

Відповідно до **Конкордату Великобританії щодо даних відкритих досліджень (UK Concordat on Open Research Data)** даними досліджень є докази, які слугують відповіддю на питання дослідження та застосовуються для верифікації результатів. Дані можуть існувати у різних формах, а також бути кількісними або якісними, зібрані дослідниками в ході експериментів, спостережень, моделювання, інтерв'ю тощо ("Concordat on Open Research Data", 2016).

Дослідницька рада з інженерних і фізичних наук (EPSRC) зазначає, що «дані дослідження визначаються як зафіксований фактичний матеріал, який зазвичай зберігається і приймається в науковому співтоваристві як необхідний для підтвердження результатів дослідження; хоча більшість таких даних створено в цифровому форматі, всі дані дослідження включені незалежно від формату, в якому вони створені» ("Scope and benefits", n.d.).

Глосарій з управління даними досліджень CASRAI трактує поняття «дані досліджень», як дані, які використовуються як первинні джерела для підтримки технічних або наукових досліджень, наукової або творчої діяльності, і які використовуються як докази у процесі дослідження та/або загальноприйняті в дослідницькому співтоваристві як необхідні для підтвердження результатів дослідження. Весь інший цифровий і нецифровий контент може стати дослідницькими даними. Дані досліджень можуть бути експериментальними даними, даними спостережень, оперативними даними, даними третьої сторони, даними державного сектору, даними моніторингу, обробленими даними або переробленими даними ("Research data", n.d.).

Крістін Боргман (Borgman, 2007) зазначає, що наукові дані та документи мають найбільшу цінність, коли вони взаємопов'язані, а не розрізнені. Результати дослідницького проекту найкраще можна зрозуміти, коли можна простежити шлях від створення грантової пропозиції до збору даних, створення набору даних, їх публікації, подальшого огляду та коментування (р. 6–11).

На думку фахівців Wageningen University & Research Library (Нідерланди), термін «дослідницькі дані» стосується усіх даних, що були зібрані, спродуковані та проаналізовані з дослідницькою метою. Наприклад:

- таблиці із замірами;
- аудіо- та відео-записи;
- транскрипти аудіозаписів;
- зразки;
- моделі та скрипти.

Файли з даними, що використовуються для певного дослідження, зазвичай групуються у пакет даних. Дані досліджень можуть бути частиною заявки на патент, захищеного авторським правом твору або бази даних, на яку поширюється право на базу даних ("Data management", n.d.).

У політиці даних Університету Теннессі ми бачимо таке трактування: «Дані досліджень включають усі записи, необхідні для реконструкції та оцінки повідомлених результатів досліджень та подій і процесів, що призводять до цих результатів, незалежно від форми чи носія інформації» ("The University of Tennessee Research Data Policy", n.d.). Зі свого боку, фахівці, **Единбурзького університету** додають, що дані досліджень, на відміну від інших типів інформації, обробляються та збираються з метою здійснення аналізу для подальшого отримання оригінальних результатів дослідження ("Information Services", 2011).

Альянс наукових організацій у Німеччині деталізує, що «дані досліджень можуть включати дані вимірювань, лабораторні значення, аудіовізуальну інформацію, тексти, дані обстеження, предмети з колекцій або зразки, які були створені, розроблені або оцінені під час наукової роботи. Методичні форми тестування, такі як анкети, програмне забезпечення та моделювання, також можуть дати важливі результати для наукових досліджень, і тому їх також слід класифікувати як дані дослідження» ("DFG Guidelines on the Handling of Research Data", 2015).

Отже, підсумовуючи проаналізовані тлумачення поняття «дані досліджень» або «дослідницькі дані», можемо сказати, що дані дослідження – це основа, яка забезпечує науковий результат. Даніми дослідження можна вважати матеріали, зібрані для обґрунтування результатів дослідження у будь-якій галузі у будь-якій формі: числовій, текстовій, комп'ютерного коду тощо. Дані можуть бути цифровими (born-digital) та оцифрованими (digitized).

ФОРМУЛЮВАННЯ МЕТИ

Метою нашого дослідження є прослідкувати та проаналізувати, які типи даних депонуються науковцями в репозитарії для даних. Для здійснення такого аналізу ми оберемо репозитарії, що індексуються реєстром re3data.org відповідно до наукових напрямів, представлених у цьому реєстрі. Таким чином, ми зможемо на реальних прикладах описів репозитаріїв із зазначеного реєстру побачити найбільш розповсюджені типи даних наукових досліджень ("Re3data.org – registry of research data repositories", n.d.).

МЕТОДИ ТА МАТЕРІАЛИ ДОСЛІДЖЕННЯ

Для того, щоб перевірити, які типи даних частіше за інші використовуються у сучасних дослідженнях, проведемо аналіз репозитаріїв, що індексуються реє-

стром re3data.org. Відповідно до цього реєстру, репозитарії розподілено за такими науковими напрямками:

1. Соціогуманітарні.

Гуманітарні: історія; мистецтво; мовознавство; літературознавство; соціально-культурна антропологія; теологія; філософія.

Соціально-поведінкові: педагогіка; психологія; соціологія; економіка; юриспруденція.

2. Про життя.

Біологія: загальна; рослин; тварин.

Медицині: мікробіологія; медицина; неврологія.

Сільське господарство та лісництво: сільське господарство; лісове господарство; садівництво; ветеринарна медицина.

3. Природничі.

Хімія: молекулярна хімія; хімічні дослідження твердого тіла та поверхні; фізична та теоретична хімія; аналітична хімія, розробка методів (хімія); біологічна хімія та харчова хімія; дослідження полімерів.

Фізика: фізика конденсованих речовин; оптика, квантова оптика та фізика атомів, молекул і плазми; частинки, ядра та поля; статистична фізика, м'яка речовина, біологічна фізика, нелінійна динаміка; астрофізика та астрономія.

Математика: математика.

Науки про Землю (включно з географією): наука про атмосферу та океанографія; геологія та палеонтологія; геофізика і геодезія; картографія; геохімія, мінералогія та кристалографія; географія; дослідження води.

4. Технічні.

Машинобудування та промислове машинобудування: технологія виробництва; механіка та конструктивне машинобудування.

Теплотехніка / Технологія: технологія, технічна хімія; теплоенергетичні технології, теплові машини, механіка рідини.

Матеріалознавство та інженерія: інженерія матеріалів; матеріалознавство.

Комп'ютерні науки, електротехніка та системна інженерія: системна інженерія; електротехніка; комп'ютерні науки.

Будівельна інженерія та архітектура: архітектура; історія будівництва; технологія сталого будівництва; проектування будівель; урбанізм; територіальне планування; транспортне та інфраструктурне планування; ландшафтне планування; будівельні матеріали; хімія, фізика; інженерія; інформатика; будівельна експлуатація; прикладна механіка; статика і динаміка; геотехніка; гідротехніка.

Варто зауважити, що кожна підгалузь також може поділятися на інші відгалуження, тому ми будемо аналізувати загальні напрями, представлені у реєстрі. З кожної галузі оберемо по одному репозитарію та проаналізуємо контент, який депонується дослідниками, таким чином, досліджуючи типи контенту, притаманного тематичними репозитаріям для даних. Варто зауважити, що у згаданому нами реєстрі наявний розподіл не лише за темами, а й за країнами та за типом контенту, аналіз якого сприятиме кращому розумінню поняття «дані дослідження».

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Соціогуманітарні науки.

Гуманітарні науки (а саме історію) представляє репозитарій Data base of Religious History (DRH) – кількісна та якісна енциклопедія релігійної історії. Репозитарій складається з різних типів внесень, включаючи релігійні групи та релігійні місця. Науковці вносять записи відповідно до своїх тем досліджень, відповідаючи на запитання у стандартизованих опитуваннях. Відповіді спочатку кодуються у двійковому форматі «так/ні» або «категорично», з полями для коментарів, посилань тощо. В репозитарії представлені такі типи даних: структурований текст, зображення, бази даних, звичайний текст, архівні дані тощо ("The world's first comprehensive", n.d.).

Соціально-поведінкові науки, зокрема соціологія, представлені, для прикладу, репозитарієм GOVDATA (Das Daten portal für Deutschland), що є порталом даних Німеччини і пропонує послідовний централізований доступ до адміністративних даних на федеральному, земельному та місцевому рівнях. Мета порталу полягає в тому, щоб зробити дані доступнішими та простішими у використанні і зібрати їх в одному місці. База покликана полегшити використання відкритих ліцензій і збільшити кількість необроблених даних для машинного зчитування. На порталі представлені такі типи даних: стандартні офісні документи, зображення, наукові та статистичні формати даних, необроблені дані, структурований текст тощо ("Daten portal für Deutschland", n.d.).

Науки про життя.

Біологія (рослин) представлена архівом VegBank (The Vegetation Plot Archive Project). VegBank – це база даних про рослинність, розроблена Екологічним товариством Америки із класифікації рослинності. VegBank складається з трьох пов'язаних баз даних, які містять фактичні записи ділянок, типи рослинності, визнані в Національній класифікації рослинності США, та інші типи рослинності, подані користувачами, і всі таксони рослин, визнані ITIS/USDA. В архіві розміщено такі типи даних: формати наукових та статистичних даних, необроблені дані, архівовані дані, програмні продукти, структурована графіка, структурований текст (VegBank, б. д.).

Медицинні науки (медицина) представлені платформою Brain-CODE (Brain-Centre for Ontario Data Exploration). Brain-CODE – це інформаційна платформа, яка керує збором і зберіганням багатовимірних даних, зібраних від учасників із різними захворюваннями мозку. Типи даних, притаманні цій базі: програмні продукти, стандартні офісні документи, наукові та статистичні формати даних, зображення тощо ("Brain-CODE", n.d.).

Сільське господарство та лісництво (лісове господарство) представлено базою даних Copernicus (The European Earth Observation Programme). Copernicus – європейська система моніторингу Землі. База складається зі складного набору систем, які збирають дані з кількох джерел: супутників спостереження за Землею і датчиків, таких як наземні станції, бортові та морські датчики. Система обробляє ці дані та надає користувачам надійну й актуальну інформацію щодо проблем навколишнього середовища і безпеки. Дані охоплюють шість тематичних напрямків: моніторинг суші, морський моніторинг, моніторинг атмосфери, зміна клімату, управління надзвичайними ситуаціями та безпека. В системі представ-

лено такі типи даних: бази даних, зображення, структурована графіка, наукові та статистичні формати даних, необроблені дані, структурований текст, звичайний текст тощо ("Home page", n.d.).

Природничі науки.

Хімію, а саме – дослідження полімерів, представлено репозитарієм iChem-VD, що є електронним репозитарієм для управління цифровими пакетами даних з обчислювальної хімії та матеріалознавства. Платформа дотримується філософії кристалографічних баз даних, принципів і потреб відкритого доступу, а також покликана вирішувати проблеми відтворюваності та відстежуваності. Типи даних, що депонуються у цей архів, такі: наукові та статистичні формати даних, структурований текст, зображення, звичайний текст ("Home", n.d.).

Фізика, а саме інформація стосовно частинок, ядра та поля, представлена у базі CHIANTI An Atomic Database for Spectroscopic Diagnostics of Astrophysical Plasmas. CHIANTI складається з критично оціненого набору актуальних даних разом зі зручними для користувача програмами, написаними інтерактивною мовою даних (IDL) і Python для обчислення спектрів астрофізичної плазми. Типи даних: архівовані дані, бази даних, програмні продукти, формати наукових і статистичних даних, структурована графіка, необроблені дані, звичайний текст, структурований текст, стандартні офісні документи ("CHIANTI. An Atomic Database", n.d.).

Математика представлена архівом DAIS Digitalni arhiv izdanja SANU. DAIS – Цифровий архів Сербської академії наук і мистецтв – є спільним цифровим сховищем Сербської академії наук і мистецтв (SASA) та науково-дослідних інститутів під егідою SASA. Метою репозитарію є надання відкритого доступу до публікацій та інших дослідницьких результатів проєктів, реалізованих SASA і його інститутами. Репозитарій використовує програмну платформу на базі DSpace, розроблену та підтримувану Комп'ютерним центром Белградського університету (RCUB). Типи даних: інші формати наукових і статистичних даних, зображення, аудіовізуальні дані, архівні дані, структурований текст, звичайний текст ("Digital Archive", n.d.).

Науки про Землю (включно з географією), зокрема наука про атмосферу та океанографія представлені системою GAW SIS GAW Station Information System. GAW SIS розробляється і підтримується Федеральним управлінням метеорології та кліматології Meteo Swiss у співпраці із Секретаріатом WMO GAW, Світовими центрами даних GAW та іншими представниками GAW для покращення управління інформацією про мережу наземних станцій GAW. GAW SIS надає спільноті GAW та іншим зацікавленим особам оновлену базу даних із можливістю пошуку з описами сайтів, програмами вимірювань і наявними даними, контактними особами, бібліографічними посиланнями. Типи даних: стандартні офісні документи, формати наукових та статистичних даних, звичайний текст, архівні дані ("GAW SIS Station Information System", n.d.).

Технічні науки.

Машинобудування та промислове машинобудування, а саме механіка і конструктивне машинобудування, представлені, зокрема, репозитарієм PTB-OAR PTB Open Access Repository. PTB – Національний інститут метрології Федеративної Республіки Німеччина. Репозитарій відкритого доступу Physikalisch-Technische Bundesanstalt надає безкоштовний доступ до низки фактичних наборів даних і документів, які були розроблені в PTB. Сюди входять такі публікації, як «PTB-

Mitteilungen», метрологічний експертний журнал РТВ тощо. Типи даних: наукові та статистичні формати даних, необроблені дані, архівні дані, структурований текст ("Willkommen im Open Access Repository", n.d.).

Теплотехніка/Технологія, в тому числі технологія, технічна хімія представлені архівом Alternative Fuels Data Center AFDC. Центр даних альтернативного палива (AFDC) – центр інформації про передові транспортні технології. AFDC пропонує особам, які приймають рішення щодо транспортування, неупереджену інформацію, дані та інструменти, пов'язані з використанням альтернативних видів палива і сучасних транспортних засобів. Типи даних: зображення, необроблені дані, звичайний текст, формати наукових і статистичних даних, стандартні офісні документи, структурована графіка, програмні продукти, аудіовізуальні дані ("Alternative Fuels Data Center", n.d.).

Матеріалознавство та інженерія представлені у базі MPDS Materials Platform for Data Science. База даних (відома як проєкт PAULING FILE) містить онлайн-матеріали з майже 2 мільйонами записів: фізичні властивості, кристалічні структури, фазові діаграми. База доступна через API, сумісна з більшістю сучасних додатків і оперує великим обсягом даних. Джерелом цих записів є близько 300 000 рецензованих публікацій у галузі матеріалознавства, опрацьованих протягом останніх 16 років міжнародною командою науковців. Результати представлені в інтернеті з інтерфейсом швидкого пошуку. Базовий доступ надається безкоштовно. Типи даних: стандартні офісні документи, зображення, структурована графіка, наукові та статистичні формати даних, структурований текст ("Materials Platform for Data Science", n.d.).

Комп'ютерні науки, електротехніка та системна інженерія (особливо електротехніка) представлені архівом CaltechLabNotes, California Institute of Technology LabNotes. LabNotes Online репрезентує історичні наукові дані з колекцій архівів CaltechLabNotes Калтеху у форматі цифрового факсиміле. Перші публікації в серії, починаючи з осені 2008 року, – це записні книжки Роберта А. Міллікена для його експериментів із вимірювання заряду електрона, що датуються з жовтня 1911 року по квітень 1912 року. Інші лабораторні, польові чи дослідницькі записи із часом додаватимуться до архіву. Типи даних: стандартні офісні документи, звичайний текст, інше ("Caltech Archives LabNotes Online", n.d.).

Будівельна інженерія та архітектура представлені репозитарієм Digital Pantheon BDPP. База містить результати повного сканування Пантеону. Типи даних: зображення, бази даних, інші, структурована графіка ("Digital Pantheon", n.d.).

Варто зауважити, що для аналізу нами були відібрані лише деякі репозитарії з окремих дисциплін або на перетині дисциплін (мультидисциплінарні). Повний перелік галузей та суміжних галузей представлено у реєстрі re3data.org, але на прикладі соціо-гуманітарних наук можемо прослідкувати шлях пакета даних або конкретного типу даних дослідження (див. рис. 1).

Повертаючись до питання типів даних, які приймаються для депонування у репозитарії для даних, варто зауважити, що в реєстрі re3data.org відображено наступний перелік можливих варіантів:

- архівні дані;
- аудіовізуальні дані;
- дані конфігурації;
- бази даних;

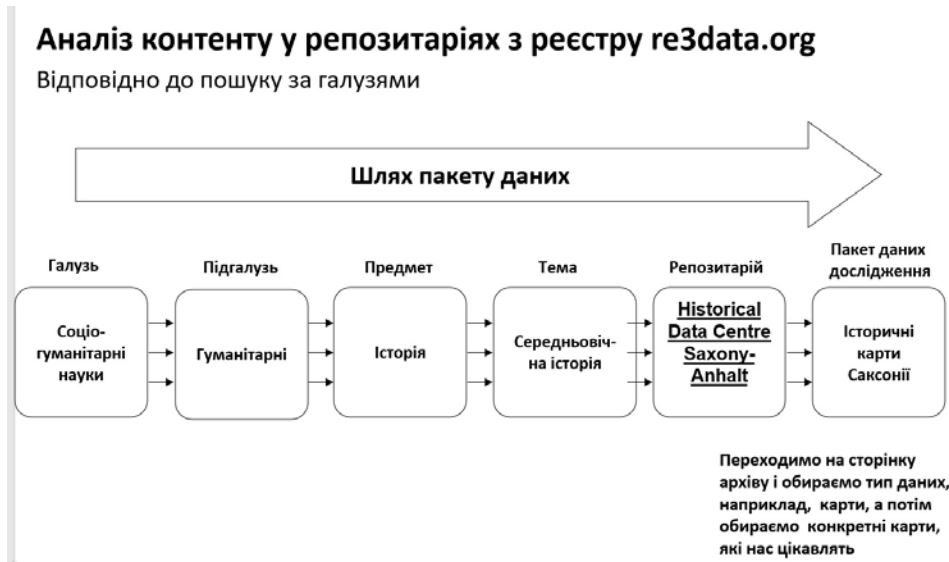


Рис. 1. Шлях пакета даних у репозитарії.

- зображення;
- мережні дані;
- простий текст;
- необроблені дані;
- науково-статистичні формати даних;
- програмні додатки;
- вихідний код;
- стандартні офісні документи;
- структурована графіка;
- структурований текст;
- інші типи.

Дані повинні бути підготовлені до депонування за одним із таких можливих стандартів метаданих: ABCD – Access to Biological Collection Data; AVM – Astronomy Visualization Metadata; CF (Climate and Forecast) Metadata Conventions; CIF – Crystallographic Information Framework; CIM – Common Information Model; CSMD-CCLRC Core Scientific Metadata Model; DCAT – Data Catalog Vocabulary; DDI – Data Documentation Initiative; DIF – Directory Interchange Format; Darwin Core; Data Cite Metadata Schema; Dublin Core; EML – Ecological Metadata Language; FGDC/CSDGM – Federal Geographic Data Committee Content Standard for Digital Geospatial Metadata; FITS – Flexible Image Transport System; Genome Metadata; ISA-Tab; ISO 19115; International Virtual Observatory Alliance Technical Specifications; MIBBI – Minimum Information for Biological and Biomedical Investigations; MIDAS-Heritage; OAI-ORE – Open Archives Initiative Object Reuse and Exchange; PROV; RDF Data Cube Vocabulary; Repository-Developed Metadata Schemas; SDMX – Statistical Data and Metadata Exchange; SPASE Data Model та інші ("Re3data.org – registry of research data repositories", n.d.).

ВИСНОВКИ

Проаналізований матеріал дає нам змогу зробити висновок, що найчастіше науковці депонують дані в текстових, числових, графічних форматах, проте також використовуються і коди, програмні продукти, бази даних. Звісно ж, типи депонованих даних залежать від специфіки репозитарію та особливостей галузі, в якій проводиться дослідження, але, як видно з обраних для аналізу описів репозитаріїв, деякі типи даних є спільними для різних дисциплін.

Отже, з огляду на отримані результати, можна дати таке визначення поняттю «дані досліджень»: дані досліджень – це текстові, числові, програмні, архівні, графічні тощо об'єкти, що слугують підґрунтям дослідження в конкретній галузі науки, тип яких напряму залежить від характеру дослідження та особливостей дисципліни чи галузі, в якій це дослідження здійснюється. Зі свого боку, не можемо не зазначити, що пакети даних супроводжуються описовою документацією в текстових форматах, тому в переліку типів матеріалів, розміщених у обраних репозитаріях, ми можемо спостерігати цей тип даних найчастіше. Дані досліджень готують до розповсюдження через репозитарії для даних (тематичні чи інституційні) відповідно до стандартів метаданих, визначених для кожної конкретної галузі науки. При обробці даних та підготовці їх до депонування у відповідні репозитарії перевага надається форматам, що сприяють розвитку Відкритої науки (nonproprietary).

СПИСОК ПОСИЛАНЬ

- Alternative Fuels Data Center. *Energy Efficiency & Renewable Energy*. URL: <https://afdc.energy.gov/> (accessed: 23.11.2021).
- Borgman C. L. *Scholarship in the Digital Age: Information, infrastructure, and the Internet*. The MIT Press, 2007. 360 p.
- Brain-CODE. *Ontario Brain Institute*. URL: <https://www.braincode.ca/> (accessed: 23.11.2021).
- Caltech Archives Lab Notes Online. *CaltechLabNotes*. URL: <http://caltechln.library.caltech.edu/> (accessed: 20.11.2021).
- CHIANTI. An Atomic Database for Spectroscopic Diagnostics of Astrophysical Plasmas. *CHIANTI Database*. URL: <https://www.chiantidatabase.org> (accessed: 25.11.2021).
- Concordat on Open Research Data. *UKRI – UK Research and Innovation*. 2016, July 28. URL: <https://cutt.ly/pTBWCXl> (accessed: 20.11.2021).
- Data management. *Wageningen. University & Research*. URL: <https://www.wur.nl/en/value-creation-cooperation/collaborating-with-wur-1/wdcc/data-management-wdcc.htm> (accessed: 23.11.2021).
- Datenportal für Deutschland. *GovData*. URL: <https://www.govdata.de/> (accessed: 23.11.2021).
- DFG Guidelines on the Handling of Research Data. (2015, September 30). *Deutsche Forschungsgemeinschaft*. URL: <https://cutt.ly/FTBQuHg> (accessed: 23.11.2021).
- Digital Archive of the Serbian Academy of Sciences and Arts. *DAIS*. <https://dais.sanu.ac.rs/?locale-attribute=en> (accessed: 20.11.2021).
- Digital Pantheon. *Edition Topoi*. <http://repository.edition-topoi.org/collection/BDPP> (accessed: 23.11.2021).
- GAWSIS Station Information System. *World Meteorological Organization*. URL: <https://gawsis.meteoswiss.ch/GAWSIS/index.html#/> (accessed: 23.11.2021).
- Home page. *Copernicus*. URL: <https://www.copernicus.eu/en> (accessed: 23.11.2021).
- Home. *ioChem-Find*. URL: <https://www.iochem-bd.org/> (accessed: 23.11.2021).

- Information Services. Defining research data. How to manage research data. *The University of Edinburgh*. 2011, Juni 20. URL: <https://cutt.ly/CTBWzx4> (accessed: 25.11.2021).
- Materials Platform for Data Science. *mpds.io*. URL: <https://mpds.io/#start> (accessed: 23.11.2021).
- Re3data.org – registry of research data repositories. Home. *re3data.org*. URL: <https://www.re3data.org/> (accessed: 19.11.2021).
- Research data. *CASRAI*. URL: <https://casrai.org/term/research-data/> (accessed: 19.11.2021).
- Research data management glossary. *CASRAI*. URL: <https://casrai.org/rdm-glossary/> (дата звернення: 19.11.2021).
- Scope and benefits. *Engineering and Physical Sciences Research Council (EPSRC)*. URL: <https://epsrc.ukri.org/about/standards/researchdata/scope/> (accessed: 25.11.2021).
- The University of Tennessee Research Data Policy. *Wayback Machine*. URL: <https://cutt.ly/uTBWEpJ> (accessed: 25.11.2021).
- The world's first comprehensive online quantitative and qualitative encyclopaedia of religious cultural history. *The Database of Religious History*. URL: <https://religiondatabase.org/landing/> (accessed: 25.11.2021).
- VegBank. URL: <http://vegbank.org/vegbank/index.jsp> (accessed: 25.11.2021).
- Willkommen im Open Access Repository der Physikalisch-Technischen Bundesanstalt (PTB-OAR). *Physikalisch-Technische Bundesanstalt*. URL: <https://oar.ptb.de> (accessed: 23.11.2021).

REFERENCES

- Energy Efficiency & Renewable Energy. (n.d.). *Alternative Fuels Data Center*. <https://afdc.energy.gov/> [in English].
- Borgman, C. L. (2007). *Scholarship in the digital age: Information, infrastructure, and the Internet*. MIT Press [in English].
- Ontario Brain Institute. (n.d.). *Brain-CODE*. <https://www.braincode.ca/> [in English].
- CaltechLabNotes. The Caltech Institute Archives. (n.d.) *Caltech Archives Lab Notes Online*. <http://caltechln.library.caltech.edu/> [in English].
- CHIANTI Database. (n.d.). *CHIANTI. An Atomic Database for Spectroscopic Diagnostics of Astrophysical Plasmas*. <https://www.chiantidatabase.org> [in English].
- UKRI – UK Research and Innovation. (2016, July 28). *Concordat on Open Research Data*. <https://www.ukri.org/wp-content/uploads/2020/10/UKRI-020920-ConcordatonOpenResearchData.pdf> [in English].
- Wageningen. University & Research. (n.d.). *Data management*. WUR. <https://www.wur.nl/en/value-creation-cooperation/collaborating-with-wur-1/wdcc/data-management-wdcc.htm> [in English].
- GovData. (n.d.). *Datenportal für Deutschland [Data portal for Germany]*. <https://www.govdata.de/> [in German].
- Deutsche Forschungsgemeinschaft. (2015, September 30). *DFG Guidelines on the Handling of Research Data*. <https://cutt.ly/FTBQuHg> [in English].
- DAIS. (n.d.). *Digital Archive of the Serbian Academy of Sciences and Arts*. <https://dais.sanu.ac.rs/?locale-attribute=en> [in English].
- Edition Topoi. (n.d.). *Digital Pantheon* <http://repository.edition-topoi.org/collection/BDPP> [in English].
- World Meteorological Organization. (n.d.). *GAWSIS Station Information System*. <https://gawsis.meteoswiss.ch/GAWSIS//index.html#/> [in English].
- Copernicus. (n.d.). *Home page*. <https://www.copernicus.eu/en> [in English].
- ioChem-Find. (n.d.). *Home*. <https://www.iochem-bd.org/> [in English].
- The University of Edinburgh. (2011, Juni 20). *Information Services. Defining research data. How to manage research data*. <https://cutt.ly/CTBWzx4> [in English].
- mpds.io. (n.d.). *Materials Platform for Data Science*. <https://mpds.io/#start> [in English].

- re3data.org. (n.d.). *Re3data.org – registry of research data repositories. Home*. <https://www.re3data.org/> [in English].
- CASRAI. (n.d.). *Research data*. <https://casrai.org/term/research-data/> [in English].
- CASRAI. (n.d.). *Research data management glossary*. <https://casrai.org/rdm-glossary/> [in English].
- Engineering and Physical Sciences Research Council. (n.d.). *Scope and benefits*. <https://epsrc.ukri.org/about/standards/researchdata/scope/> [in English].
- Wayback Machine. (n.d.). *The University of Tennessee research data policy*. <http://surl.li/avenl> [in English].
- The Database of Religious History. (n.d.). *The world's first comprehensive online quantitative and qualitative encyclopedia of religious cultural history*. <https://religiondatabase.org/landing/> [in English].
- VegBank. (n.d.). <http://vegbank.org/vegbank/index.jsp> [in English].
- Physikalisch-Technische Bundesanstalt. (n.d.). *Willkommen im Open Access Repository der Physikalisch-Technischen Bundesanstalt (PTB-OAR) [Welcome to the Open Access Repository of the Physikalisch-Technische Bundesanstalt (PTB-OAR)]*. <https://oar.ptb.de> [in German].

UDC 001.103:004.451.5

Svitlana Chukanova,
Head of the Sector, National University
“Kyiv-Mohyla Academy” Research Library,
PhD in Education (Kyiv, Ukraine)
e-mail: chukanovaso@ukma.edu.ua
ORCID: 0000-0002-5717-5050

THE NOTION OF "RESEARCH DATA": TYPES AND KINDS OF RESEARCH DATA IN THE CONTEXT OF DATA MANAGEMENT PRACTICE

With the rapid development of the concept of Open Science, the quantitative growth of data obtained during the research, scientific attention to the practice of research data management (research data management) increases, which actualizes the definition of “research data” and identifying types of research data within the practice of their management, justification and coverage of the specifics of such data.

The methodological tools of the study are based on the terminological method, the use of which was due to the need to identify relevant interpretations of the concept of “research data”, as well as analysis of repositories for data from various fields of science, indexed by re3data.org., in the general areas presented in the register, namely: descriptions of repositories, including information on the types of data deposited by scientists and data curators.

The analysis made it possible to define research data as materials obtained and collected to substantiate the scientific results of research in any field and in any form: numerical, textual, computer code, etc., as well as to identify types of data specific to different branches of science, which, in turn, allowed us to conclude the existing data formats, the most common among both natural and human sciences: text, numerical and graphic formats. As a result of the analysis, it was found that research data can be considered textual, numerical, software, archival, graphic and other objects (files) that serve as the basis of the study and the factual basis for scientific conclusions in a particular field of science. It was found that the type of data directly depends on the nature of the study and the characteristics of the discipline or field of research.

Keywords: research data, data repositories, data types, data formats, data deposit.

Стаття надійшла до редакції 17.11.2021 р.