

Олексюк В. П.

доктор педагогічних наук, професор, старший дослідник, професор кафедри інформатики та методики її навчання ТНПУ імені Володимира Гнатюка, Тернопіль, Україна,
oleksyuk@fizmat.tnpu.edu.ua
ORCID ID: <https://orcid.org/0000-0003-2206-8447>

Спирін О. М.

доктор педагогічних наук, професор, академік НАПН України, професор кафедри комп'ютерних наук та інформаційних технологій Житомирського державного університету імені Івана Франка, Житомир, Україна, spirin@iitlt.gov.ua
ORCID ID: <https://orcid.org/0000-0002-9594-6602>

Осадча К. П.

доктор педагогічних наук, професор, провідний науковий співробітник Інституту цифровізації освіти НАПН України, Київ, Україна, k.osadcha@iitlt.gov.ua
ORCID ID: <https://orcid.org/0000-0003-0653-6423>

МОДЕЛЬ СТАНДАРТИЗАЦІЇ FAIR-ДАНИХ ДЛЯ ГАЛУЗІ ОСВІТНІХ НАУК

Анотація. Зростання обсягів цифрової інформації в науково-педагогічних дослідженнях актуалізує проблеми її збереження, індексування й обміну. Метою статті є розроблення та теоретичне обґрунтування моделі стандартизації дослідницьких даних у галузі освітніх наук відповідно до міжнародних принципів відшукованості, доступності, сумісності, багаторазовості, що нині описуються поняттям «FAIR-дані». У дослідженні коротко подано зарубіжний досвід реалізації проєктів відкритої науки, виділено специфічні риси дослідницьких даних для галузі освітніх наук. У результаті спроектовано авторську модель, яка містить цільовий, змістовий, процесуальний і результативний блоки. У межах змістового блоку розроблено таксономію освітніх даних, описано формати їх подання, а також запропоновано структуру типового пакета FAIR-даних. У межах змістового компонента моделі висвітлено етичні норми, дотримання яких є обов'язковим під час формування наборів даних, процедури їх формування та елементи інфраструктури для їх зберігання. Особливу увагу приділено результативному компоненту, для якого створено систему із критеріїв і показників для оцінювання відповідності дослідницьких даних у галузі освіти принципам FAIR. Запропоновано шкалу, що дає змогу класифікувати набори даних за чотирма рівнями: недостатнім, базовим, достатнім та високим. Доведено, що впровадження розробленої моделі й системи показників забезпечує перехід від декларативного застосування принципів відкритої науки до вимірюваного та технологічного життєвого циклу управління даними. Імплементация спроектованої моделі сприятиме підвищенню прозорості, якості й відтворюваності педагогічних досліджень.

Ключові слова: принципи FAIR, дослідницькі дані, освітні науки, таксономія даних, метадані, стандартизація, критерії та показники.

DOI: 10.32987/2617-8532-2026-2-91-103.

Розвиток цифрових технологій та зростання обсягів даних досліджень

висувають перед науковою спільнотою завдання забезпечити ефектив-

© Олексюк В. П., Спирін О. М., Осадча К. П., 2026

не управління, обмін ними та їх повторне використання. Дані, що відповідають таким вимогам, називають FAIR-даними. Аббревіатуру FAIR утворено від англійських слів «Findable» (відшукуваність), «Accessible» (доступність), «Interoperable» (сумісність, або інтероперабельність), «Reusable» (багаторазовість). За даними Європейської комісії, у 2018 р. щорічні втрати через відсутність FAIR-даних в ЄС становили понад 11 млрд євро [1]. Отже, неправильне управління даними наукових досліджень спричиняє фінансові, наукові й репутаційні втрати. Саме тому принципи FAIR, запропоновані 2016 р. групою науковців [2], є відповіддю на потребу у створенні універсальних, не залежних від конкретної дисципліни рекомендацій для управління науковими даними та цифровими об'єктами. Як зазначають автори, найбільшу цінність FAIR-дані становлять у контексті проектування відкритих наукових систем [3].

Однак значна частина дослідників у галузі освіти залишаються недостатньо обізнаними щодо концепції FAIR-даних та їх практичного застосування. На думку авторів публікації [4], принципи FAIR і практики відкритої науки слід систематично поширювати серед наукової спільноти. Проте такий підхід є проблематичним у гуманітарних та соціальних науках, в яких чимало даних є описовими чи категорійними. Тому деякі дослідники в галузі соціально-гуманітарних наук зауважують, що термін «дані» недостатньо відповідає їхній діяльності, і як більш релевантне поняття обирають «дослідницькі матеріали» [5]. Таке розмежування підкреслює необхідність адаптації

принципів FAIR до специфіки освітніх досліджень.

Метою статті є проектування моделі стандартизації FAIR-даних для галузі освітніх наук.

Задля уточнення вимог до FAIR-даних, що отримуються, опрацьовуються та публікуються в процесі виконання досліджень у галузі освітніх наук, нами було розроблено таксономію [6]. Вона систематизує їх набори за різними ознаками, а саме: за цільовим призначенням (фундаментальні, прикладні, практичні розробки); за спеціальностями / напрямками в галузі «Освіта»; за кваліфікаційними рівнями; за видом науково-дослідної роботи; за суб'єктом подання й виконання даних; за типами й методами досліджень; за видами, рівнями, формами здобуття освіти; за джерелами отримання даних, їх форматами, рівнем агрегованості, режимами доступу та ін. Додатково до зазначеної таксономії FAIR-дані можна класифікувати за режимом доступу (відкриті, з обмеженим доступом, закриті), за ступенем ідентифікації (з особистими даними, частково або повністю анонімізовані), а також за життєвим циклом (сирі, очищені, агреговані, похідні, архівовані).

Проектована модель має відобразити сучасні принципи роботи з експериментальними даними, специфіку саме науково-педагогічних досліджень, а також технологічні процеси, виконання яких призводить до отримання стандартизованих FAIR-даних. Результатом запровадження моделі є перетворення різнорідних даних на відшукуваний, доступний, сумісний і повторно використовуваний ресурс. Модель є системою складників – цільового, змістового, процесуального та результативного (рисунок). Вони

відтворюють логіку процесу стандартизації даних, отриманих у процесі виконання науково-педагогічних досліджень.

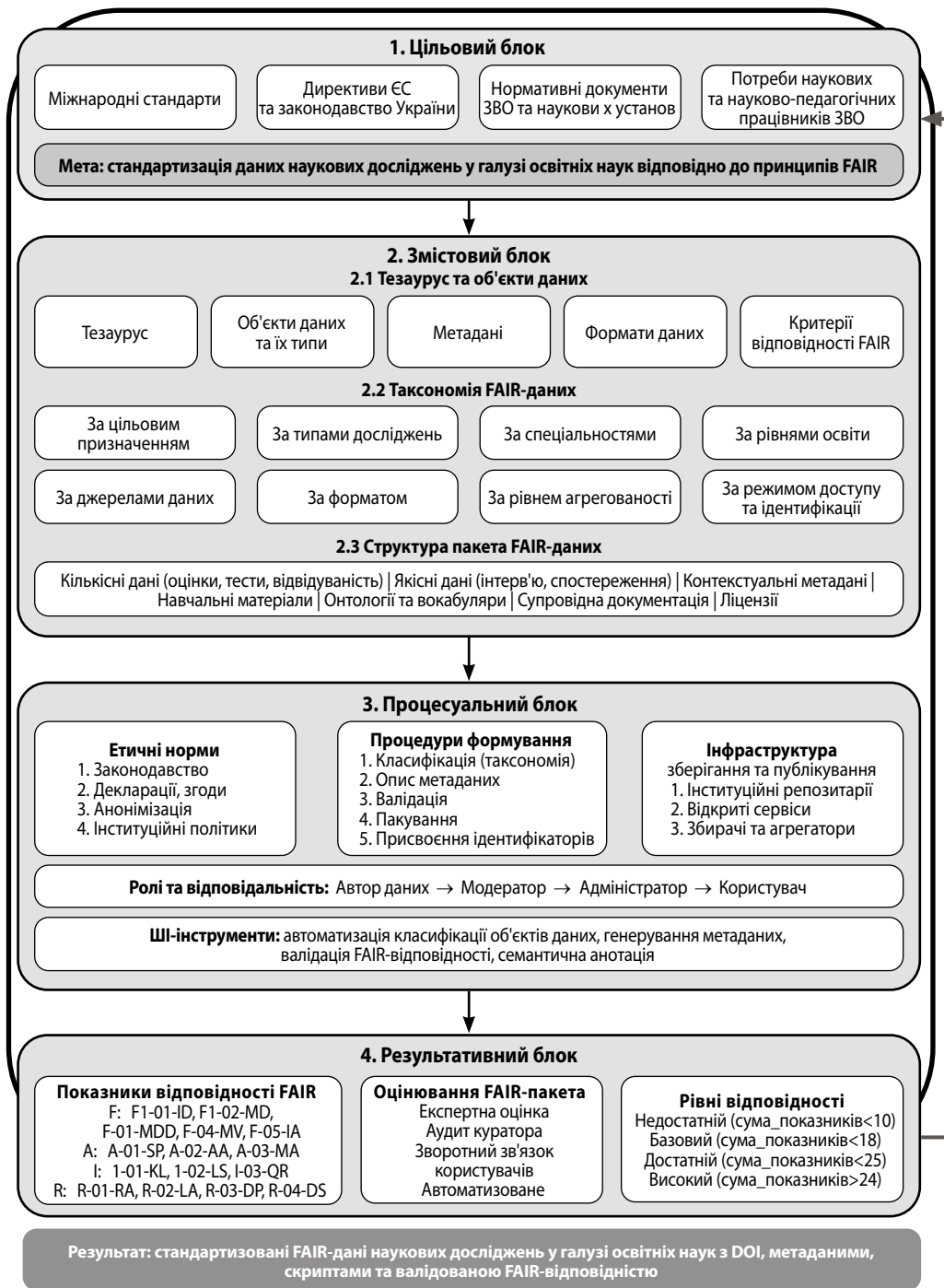


Рисунок. Модель стандартизації FAIR-даних для галузі освітніх наук

Побудовано авторами.

Цільовий компонент визначає функціональне призначення даних як наукового продукту. На основі таксономії FAIR-даних цей складник дає можливість відповісти, які типи даних (теоретичні, експериментальні, кореляційні, регресійні тощо) і для яких рівнів освіти будуть стандартизовані. З метою забезпечення сумісності та придатності даних до повторного використання в цільовому складнику моделі передбачено дотримання таких національних і міжнародних стандартів:

- ДСТУ ISO 25964 «Тезауруси та взаємосумісність» – регулює створення та використання тезаурусів і контрольованих словників, що, у свою чергу, має забезпечувати Findable (відшукуваність), тобто семантичну однозначність термінів.

- Дублінське ядро – базовий стандарт метаданих будь-якого цифрового об'єкта, забезпечує видимість FAIR-даних в інституційних репозитаріях, електронних бібліотеках, а також можливість обміну ними.

- ISCED – міжнародна класифікація освіти, розроблена ЮНЕСКО. У моделі вона дає можливість стандартизувати рівні освіти від 0 (дошкільна освіта) до 8 (третій (PhD) рівень вищої освіти).

- DCAT – словник моделі RDF (Resource Description Framework), що призначений для опису наборів даних і є стандартом для відкритих даних.

Цільовий компонент моделі передбачає дотримання регламенту GDPR (General Data Protection Regulation) [7], зокрема його положень щодо точності, цілісності, конфіденційності, підзвітності даних, а також, за потреби, цільове обмеження доступу до них. Стосовно наукових досліджень

у галузі освіти обов'язковим для виконання є Закон України «Про захист персональних даних» від 01.06.2010 № 2297-VI, який регламентує законність обробки даних, їх знеособлення, відповідність цілям дослідження, право суб'єктів на доступ [8].

Змістовий компонент містить тезаурус основних термінів, що забезпечує уніфікацію спеціальностей, рівнів освіти, загальних і спеціальних компетентностей, програмних результатів навчання тощо. Складник встановлює вимоги, щоб усі об'єкти даних були класифіковані відповідно до таксономії, а метадані документували їхній контекст, тобто описували джерело отримання, рівень агрегованості, методик збору й опрацювання даних.

Упродовж багатьох років базовим стандартом метаданих залишається Дублінське ядро. Як показують дослідження, якісне заповнення його стандартних полів забезпечує відшукуваність даних [9; 10]. Наступним кроком опису метаданих, зокрема для каталогів, є DCAT. Його використання забезпечує не лише пошук даних сервісами на кшталт Google Dataset Search, OpenAIRE, а й сумісність з іншими системами, що працюють із моделлю RDF.

Ідентифікація даних згідно зі спроектованою моделлю здійснюється за допомогою схеми метаданих DataCite. Її можна вважати своєрідним паспортом набору даних, адже саме ця схема присвоює ідентифікатор (найчастіше DOI), що забезпечує цитування даних у наукових публікаціях та пов'язує їх набір із відповідною статтею.

Згадана модель RDF і мова опису онтологій у веб (OWL) є засобами

забезпечення найвищого рівня сумісності метаданих – їх семантичної інтероперабельності. У результаті можливе отримання значення кожної змінної, зіставлення її з іншими наборами даних, а також об'єднання даних різних досліджень.

Змістовий складник моделі передбачає оперування багатьма даними в різних форматах, базовим рівнем яких є табличні структуровані масиви (CSV, TSV), що використовуються для статистичного аналізу кількісних показників успішності чи анкетування. Розширення дослідницького контексту можливе шляхом обробки напівструктурованих даних у форматах JSON, xAPI, log, отриманих із систем управління навчанням, масових відкритих онлайн-курсів, бібліометричних засобів тощо. Модель передбачає збереження неструктурованих текстових масивів, а також геопросторових і мультимедійних даних.

У змістовому компоненті розробленої моделі також присутні критерії та показники відповідності даних принципам FAIR. Під час їх визначення було враховано ідеї статті «Керівні принципи FAIR щодо управління науковими даними та їх використання» [2], підходи до адаптації принципів FAIR до освітньої галузі [11], а також рекомендації МОН України щодо FAIR для дослідницьких даних [12]. Модель використовує чотири критерії, які сумарно оцінюються за 14 показниками (табл. 1).

У межах змістового складника блок «Критерії відповідності FAIR» виконує функцію нормативно-теоретичної основи. Тобто дослідник ще на етапі планування свого набору даних має враховувати систему критеріїв і показників, зокрема розуміти, які саме метадані потрібно зібрати, який формат файлу обрати та яку ліцензію застосувати, щоб

Таблиця 1

Критерії та показники оцінювання FAIR-даних у галузі освітніх наук

Кри-терій	Показ-ник	Опис показника	Опис значень показника
F Відшукваність	F1-01-ID	Наявність унікального й постійного ідентифікатора	2: Присвоєно DOI, Handle або URN, який «розв'язується» через відповідні сервіси. 1: Лише звичайний URL. 0: або ідентифікатор відсутній.
	F1-02-MD	Дані описуються достатньою кількістю метаданих	2: Надано всі обов'язкові поля метаданих відповідно до моделі RDF (контекст дослідження, рівень освіти та методологію). 1: Частковий опис, що відповідає Дублінському ядру. 0: Опис відсутній.
	F-01-MDD	Зв'язок метаданих із файлами даних	2: Зв'язок між метаданими й даними прописаний у машинозчитуваному коді та підтверджений успішним автоматичним запитом до файлу. 1: Метадані містять прямі посилання на файли. 0: Файли даних не пов'язані з описом автоматично.
	F-04-MV	Машинна видимість	2: Дані сформовано відповідно до стандартів JSON-LD чи Microdata. 1: Дані сформовано за розмітками Microdata або Schema.org. 0: Дані не відшукуються пошуковими системами.
	F-05-IA	Індексація в агрегаторах	2: Метадані проіндексовано глобальними реєстрами (DataCite, OpenAIRE) за допомогою протоколу OAI-PMH. 1: Метадані індексуються лише одним агрегатором. 0: Метадані доступні лише локально (на вебсервері установи або особистому пристрої).

Кри-терій	Показ-ник	Опис показника	Опис значень показника
А Доступність	A-01-SP	Доступність через стан-дартизований прото-кол	2: Доступ здійснюється через відкритий протокол (HTTP/HTTPS, FTP) за прямим PID-посиланням. 1: Доступ можливий лише через пропріетарний інтерфейс конкретної платформи. 0: Дані неможливо отримати за ідентифікатором автоматично.
	A-02-AA	Наявність процедур автентифікації / автори-зації	2: Чітко визначені умови доступу (Open/Restricted) з автоматизованою процедурою авторизації. 1: Умови доступу згадані в тексті, але потребують ручного запиту (на-приклад, через e-mail). 0: Процедури доступу не визначено.
	A-03-MA	Доступність метаданих	2: Метадані зберігаються в репозитарії (сервісі), зокрема після вида-лення або закриття доступу до самих даних. 1: Метадані видаляються разом із даними, але є доступними локально. 0: Метадані недоступні, якщо файл даних було видалено.
І Інтероперабельність (сумісність)	I-01-KL	Використан-ня формаль-ної мови подання знань	2: Метадані подано в машиночитаних семантичних форматах відпо-відно до стандартів RDF, JSON-LD, OWL. 1: Дані подано в загальноживаних структурованих форматах (CSV, JSON, XML). 0: Дані подано в неструктурованих чи закритих форматах (PDF, DOCX, скановані копії).
	I-02-LS	Використан-ня словників, що відпові-дають прин-ципам FAIR	2: Використовуються стандартизовані контрольовані вокабуляри та онтології (ISCED, EuroVoc, спеціалізовані освітні тезауруси), які мають ідентифікатори. 1: Використовуються загальноприйняті терміни, але без посилання на глобальні класифікатори. 0: Використовується довільна термінологія й абрєвіатури.
	I-03-QR	Наявність кваліфікованих посилань на інші ре-сурси	2: Метадані містять чітко визначені зв'язки з іншими об'єктами із зазна-ченням їхніх ідентифікаторів (DOI, DOI попередніх етапів дослідження). 1: У метаданих згадуються пов'язані ресурси лише в текстовому ви-гляді без активних ідентифікаторів. 0: Посилання на контекст або пов'язані дослідження відсутні.
R Багаторазовість	R-01-RA	Повнота опи-су ознак	2: Надано вичерпний опис контексту (вибірка, методика, умови експе-рименту, інструментарій), достатній для повторного його проведення. 1: Опис обмежений базовими характеристиками, що ускладнює повне розуміння контексту збору даних. 0: Контекстуальний опис відсутній.
	R-02-LA	Наявність і чіткість ліцензії	2: Присвоєно стандартну ліцензію (наприклад, CC BY 4.0), яку містять метадані. 1: Умови використання згадані лише текстом у README-файлі без по-силання на стандартну ліцензію. 0: Ліцензія відсутня (дані захищені копірайтом за замовчуванням, ви-користання заборонено).
	R-03-DP	Походження даних	2: Зафіксовано всі етапи обробки даних: від сирих даних до фінальних, зокрема надано скрипти їх обробки. 1: Надано лише загальний опис етапів обробки даних без технічних деталей. 0: Шлях формування даних не описано.
	R-04-DS	Відпо-відність стандартам галузі	2: Структура та метадані повністю відповідають стандартам галузі освітніх наук (наприклад, xAPI – для журналів, ISCED – для рівнів осві-ти, LOM – для навчальних об'єктів). 1: Дані структуровані логічно, але не використовують специфічні для галузі стандарти даних (метаданих). 0: Формат і структура даних є довільними та специфічними лише для автора.

Складено авторами.

кінцевий продукт відповідав принципам FAIR.

Наступним блоком моделі є «2.2 Таксономія FAIR-даних», що визначає типи даних. Проте сам по собі файл певного типу (наприклад, таблиця у форматі CSV) не є повноцінним об'єктом наукового дослідження. У зв'язку з цим до цільового складника моделі введено окремий блок «2.3 Структура пакета FAIR-даних», призначення якого полягає в систематизації даних, що у свої сукупності утворюють цілісний пакет FAIR-даних. Тобто він описує пакет FAIR-даних як структуру таких елементів: даних, метаданих, словників, дослідницького інструментарію, дидактичних матеріалів, супровідної документації, ліцензій.

Процесуальний компонент моделі описує механізм перетворення окремих файлів, отриманих у процесі виконання дослідження, на пакет FAIR-даних із його подальшою публікацією та введенням у науковий обіг. Модель передбачає, що перед опрацюванням даних їх слід проаналізувати на відповідність етичним вимогам, які включають:

- визначення, чи не суперечитиме опублікований пакет законам України та ЄС, що вказані в цільовому блоці моделі;

- отримання інформованої згоди від учасників експериментальних досліджень (у разі відсутності такої згоди встановлюється обмежений режим доступу до даних);

- анонімізацію даних шляхом видалення безпосередніх і непрямих ідентифікаторів учасників дослідження;

- урахування інституційних політик ЗВО чи наукової установи, які регламентують, кому належать дані, а також як вони мають зберігатися.

Ядром процесуального складника моделі є процедури формування даних, до яких входить:

1. Класифікація даних відповідно до таксономії. Наприклад, дослідник визначає, що його дані належать до експериментального типу, виконані в межах дослідження за спеціальністю А1 «Освітні науки», на першому (бакалаврському) рівні вищої освіти. Дані мають інформовані згоди на використання й публікацію, а також є анонімізованими.

2. Опис метаданих, який передбачає використання тезаурусу і стандартів метаданих для створення опису набору FAIR-даних, зокрема зазначення авторів, часу отримання даних, методик їх опрацювання тощо.

3. Валідація як процедура перевірки даних на цілісність, відсутність у них помилок та їх відповідність критеріям FAIR.

4. Пакування з метою об'єднання окремих файлів (даних, інструментарію й супровідної документації) у FAIR-пакет, як це передбачено в блоці 2.3.

5. Присвоєння ідентифікаторів, для чого необхідно отримати DOI або інший постійний ідентифікатор. Ця процедура робить пакет FAIR-даних унікальним і відшукуваним.

Блок «Інфраструктура зберігання та публікування» визначає програмні засоби, які забезпечують довгострокове зберігання й надають доступ до FAIR-даних. Першим рівнем вказаного блоку є інституційні репозитарії (електронні бібліотеки), що забезпечують архівування локальних наукових матеріалів та їх опрацювання інформаційними системами інституційного рівня, до яких належать системи звітності,

формування рейтингів, модулі інтеграції з національними базами даних [13]. Для забезпечення доступу на міжнародному рівні слід використовувати спеціалізовані відкриті сервіси, що надають наборам даних унікальні постійні ідентифікатори та забезпечують їх відшукуваність. До них належать сервіси Zenodo, Open Science Framework, Figshare. Забезпечення сумісності виконують збирачі й агрегатори метаданих, прикладом яких є портали European Open Science Cloud та OpenAIRE. До прикладу підсистеми OpenAIRE Graph і OpenAIRE Explore, використовуючи протокол обміну метаданими OAI-PMH, збирають метадані з інституційних репозитаріїв, архівів даних та журналів. Додаючи в OpenAIRE дані власних досліджень, науковець створює вузол семантичної мережі, що пов'язує набір даних із профілем ORCID автора, опублікованою статтею та, опціонально, з грантом, за яким фінансувалося дослідження.

Розглядаючи життєвий цикл керування, зауважимо, що процес формування та публікації пакета FAIR-даних не є повністю автоматизованим, а отже, передбачає взаємодію фахівців із такими повноваженнями:

- Автор даних. Ним є сам дослідник, який отримує дані, анонімізує їх, формує пакет і завантажує до репозитарію або сервісу.

- Модератор. Зазвичай ним є користувач сервісу чи репозитарію, що має повноваження для редагування метаданих і погодження завантаження набору даних. Окрім цього, він перевіряє якість метаданих, коректність ліцензії та визначає відповідність усього пакета критеріям FAIR.

- Адміністратор як відповідальна особа, що має повний обсяг повноважень. Він підтримує роботу репозитарію, виконує оновлення платформи, резервне копіювання файлів, баз даних, ОС, а також налаштовує машинну інтероперабельність (додавання форматів даних, обмін метаданими, вмикання API, упровадження контрольованих словників).

- Користувач. Ним є науковець, який знайшов вказаний пакет FAIR-даних, цитує його й використовує у своєму дослідженні.

Наразі ефективним інструментом виконання багатьох завдань із підготовки та використання FAIR-даних є штучний інтелект (ШІ). Наведемо приклад використання сучасних чат-ботів генеративного ШІ для реалізації окремих процедур формування пакета FAIR-даних. До них належать операції анонімізації (текстові бланки інтерв'ю, табличні дані тестування), генерування чорнового варіанта опису набору у форматі файлу README, допомога у виборі правильних термінів із тезаурусу під час анотування пакета FAIR-даних.

Результативний компонент відповідає за верифікацію досягнення мети моделі як отримання кінцевого продукту, так й інструментів для вимірювання його якості. Тобто цей складник не просто вимагає констатувати факт публікації даних, а й підтвердження, що створений цифровий об'єкт справді відповідає міжнародним стандартам FAIR. Структурно він складається з таких блоків:

1. Показники відповідності FAIR. Блок дає змогу оцінити рівень відповідності даних із використанням індикаторів (див. табл. 1).

2. Методи оцінювання пакетів FAIR-даних, до яких належать як автоматизовані, так і людиноорієнтовані, зокрема експертна оцінка, аудит куратора, відгуки користувачів.

3. Рівні відповідності даних принципам FAIR. Як видно з табл. 1, загалом виокремлено 15 показників, кожний з яких оцінюється від 0 до 2 балів. Отож 30 балів є максимально можливою сумою балів за всіма показниками. Пропонуємо для визначення рівня відповідності використати шкалу, наведену в табл. 2.

Як видно з табл. 2, для оцінювання вказаних рівнів уведено поняття базових показників, для яких оцінка нуль є неприпустимою. До них слід зарахувати F1-01-ID (наявність ідентифікатора), F1-02-MD (наявність метаданих), A-01-SP (доступність через стандартизований протокол), R-02-LA (наявність ліцензії).

У результативному блоці, на відміну від змістового, критерії та показ-

ники відповідності даних принципам FAIR набувають вимірювальної й діагностичної функцій. Застосовуючи бальну шкалу за кожним показником розробленої таблиці, експерт або автоматизована система можуть на основі запропонованих критеріїв і показників оцінити рівень відповідності пакета даних принципам FAIR.

На основі аналізу наукової літератури й досвіду реалізації FAIR-принципів у різних дисциплінах можна зробити такі висновки:

1. Розвиток моделі стандартизації FAIR-даних для галузі освітніх наук є актуальним завданням у контексті цифрової трансформації та імплементації на практиці ідей відкритої науки. Обсяги даних освітніх досліджень стрімко зростають, але значна їх частина залишається неопублікованою або доступна в нестандартизованих форматах, що непридатні до автоматизованого індексування, пошуку й аналізу.

Таблиця 2

Рівні відповідності FAIR-даних

Рівень	Діапазон балів	Обов'язкова умова	Опис (характеристика набору)
Недостатній (Pre-FAIR)	0–10	Базові показники оцінено нулем	Дані не готові до наукового обміну. Не зберігаються взагалі або традиційно (як додаток до статті) без ліцензії, постійного посилання та повноцінного опису
Базовий (Min-FAIR)	11–18	Набрано мінімум по 1 балу за всіма критеріями (F, A, I, R); відсутні нульові значення за базовими показниками	Дані опубліковано в Інтернеті з URL, мають базові метадані, що відповідають стандарту Дублінського ядра, текстову ліцензію й умови доступу. Їх може знайти та прочитати людина
Достатній (Suff-FAIR)	19–25	Понад 50 % показників оцінено у 2 бали	Оптимальний рівень для більшості досліджень. Дані збережено у структурованих форматах (CSV/JSON), визначено їхній контекст, походження та присвоєно ідентифікатори
Високий (Max-FAIR)	26–30	Майже повна відповідність машиночитаним семантичним стандартам	Набір FAIR-даних найвищого гатунку у форматі RDF, JSON-LD, містить онтології ISCED і вузькогалузеві стандарти xAPI, LOM, придатний до автоматизованого машинного аналізу

Складено авторами.

2. Досвід природничих і медичних наук свідчить, що проектування моделей FAIR-даних можливе та забезпечує переваги як для дослідників, так і для суспільства загалом, а описані в статтях проекти доводять, що стандартизація FAIR-даних є практично досяжним завданням.

3. Модель стандартизації FAIR-даних для галузі освітніх наук враховує специфіку галузі, зокрема мультидисциплінарність досліджень, поєднання кількісних та якісних методів, необхідність захисту персональних даних учасників освітнього процесу. Важливим складником моделі є критерії та показники, відповідно до яких має здійснюватись оцінювання FAIR-даних у галузі освітніх наук. Використовуючи запропоновану шкалу,

а також базові показники, авторська модель дає можливість класифікувати набори FAIR-даних за одним із таких рівнів: недостатній, базовий, достатній, високий.

4. Апробація моделі стандартизації FAIR-даних для освітніх наук вимагає залучення і спільної роботи адміністрації ЗВО й наукових установ, дослідників у галузі освіти, розробників програмного забезпечення, фахівців у галузі бібліотечної справи та інших. Жоден із професіоналів перелічених категорій не може самостійно розв'язати зазначене завдання, проте їхня спільна діяльність дасть змогу також оцінити ефективність моделі та вдосконалити технології формування FAIR-даних.

Список використаних джерел

1. Cost-benefit analysis for FAIR research data – Cost of not having FAIR research data / Directorate-General for Research and Innovation, PwC EU Services. Publications Office, 2018. 48 p. DOI: <https://data.europa.eu/doi/10.2777/02999>.
2. The FAIR Guiding Principles for scientific data management and stewardship / M. D. Wilkinson et al. *Scientific Data*. 2016. Vol. 3, 160018. DOI: <https://doi.org/10.1038/sdata.2016.18> (дата звернення: 04.03.2026).
3. Мар'єнко М. В., Шишкіна М. П., Коновал О. А. Методологічні засади формування хмаро орієнтованих систем відкритої науки у закладах вищої педагогічної освіти. *Інформаційні технології і засоби навчання*. 2022. Т. 89. № 3. С. 209–232. DOI: <https://doi.org/10.33407/itlt.v89i3.4981> (дата звернення: 04.03.2026).
4. Open Science Training in TRIPLE / L. Provost et al. *Open Research Europe*. 2023. Vol. 3, 39. DOI: <https://doi.org/10.12688/openreseurope.15430.1> (дата звернення: 04.03.2026).
5. Gualandi B., Pareschi L., Peroni S. What do we mean by «data»? A proposed classification of data types in the arts and humanities. *Journal of Documentation*. 2023. Vol. 79, No. 7. P. 51–71. DOI: <https://doi.org/10.1108/jd-07-2022-0146> (дата звернення: 04.03.2026).
6. Снірін О. М., Олексюк В. П. Таксономія FAIR-даних для наукових досліджень у галузі освіти. *Цифрова трансформація науково-освітніх середовищ в умовах воєнного стану*: звітна наук. конф. Інституту цифровізації освіти НАПН України, м. Київ, 25 лют. 2026 р. DOI: <https://doi.org/10.5281/zenodo.18909544> (дата звернення: 08.03.2026).
7. General Data Protection Regulation (GDPR). *Intersoft consulting*. URL: <https://gdpr-info.eu/> (дата звернення: 08.03.2026).
8. Про захист персональних даних : Закон України від 01.06.2010 № 2297-VI. URL: <https://zakon.rada.gov.ua/laws/show/2297-17#Text> (дата звернення: 08.03.2026).

9. A Minimal Metadata Schema and Its Tool to Improve the Searchableness of Research Data in Bioinformatics / O. Brandt et al. *Journal of Library Metadata*. 2024. Vol. 24, Iss. 3. P. 165–188. DOI: <https://doi.org/10.1080/19386389.2024.2338314> (дата звернення: 08.03.2026).

10. AgroFIMS: A Tool to Enable Digital Collection of Standards-Compliant FAIR Data / M. Devare et al. *Frontiers in Sustainable Food Systems*. 2021. Vol. 5, 726646. DOI: <https://doi.org/10.3389/fsufs.2021.726646> (дата звернення: 08.03.2026).

11. Towards FAIRification of learning resources and catalogues – lessons learnt from research communities / L. Provost et al. *Frontiers in Education*. 2024. Vol. 9, 1390444. DOI: <https://doi.org/10.3389/feduc.2024.1390444> (дата звернення: 08.03.2026).

12. Принципи FAIR для дослідницьких даних. *НАУКА*. URL: <https://nauka.gov.ua/information/pryntsyru-fair-dlia-doslidnytskykh-danykh/> (дата звернення: 08.03.2026).

13. Олексюк В. П., Габрусев В. Ю., Балик А. В. Деякі аспекти інтеграції веб-сервісів вищого навчального закладу. *Наукові записки Тернопільського національного педагогічного університету імені Володимира Гнатюка. Сер. : Педагогіка*. 2011. № 1. С. 228–234.

Матеріал надійшов до редакції 10.03.2026 р.

Vasyl Oleksiuk

Dr. Sc. (Pedagogical), Professor, Senior Researcher, Ternopil Volodymyr Hnatyuk National Pedagogical University, Ternopil, Ukraine, oleksyuk@fizmat.tnpu.edu.ua
ORCID ID: <https://orcid.org/0000-0003-2206-8447>

Oleh Spirin

Dr. Sc. (Pedagogical), Professor, Academician of the NAES of Ukraine, Zhytomyr Ivan Franko State University, Zhytomyr, Ukraine, spirin@iitlt.gov.ua
ORCID ID: <https://orcid.org/0000-0002-9594-6602>

Kateryna Osadcha

Dr. Sc. (Pedagogical), Professor, Institute for Digitalisation of Education, Kyiv, Ukraine, k.osadcha@iitlt.gov.ua
ORCID ID: <https://orcid.org/0000-0003-0653-6423>

A STANDARDISATION MODEL OF FAIR DATA FOR EDUCATIONAL SCIENCES

Abstract. *The growing volume of digital information in scientific and pedagogical research underscores the urgent need for its preservation, indexing, and exchange. The aim of this article is to develop and theoretically substantiate a standardisation model for research data in educational sciences, in accordance with the international principles of findability, accessibility, interoperability, and reusability, collectively known as FAIR data. This study briefly examines international experience in implementing Open Science projects and identifies the specific features of research data within the field of educational sciences. Consequently, an original model was designed, comprising goal-oriented, content-related, procedural, and outcomes-based components. Within the content block, a taxonomy of educational data is developed, data presentation formats are described, and the structure of a standard FAIR data package is proposed. Furthermore, the procedural component of the model outlines the ethical standards mandatory for dataset compilation, the procedures for generating these data, and the infrastructural elements required for their storage. Particular attention is given to the outcomes-based component, for which a system of criteria and indicators has been established to evaluate the compliance of educational research data with FAIR principles. A scale is proposed to classify datasets into four levels of maturity: insufficient, basic, sufficient, and high. The conclusion demonstrates that implementing the developed model and indicator system ensures a transition from the declarative application of Open Science principles to a measurable and technologically integrated data management lifecycle. The implementation of the designed model will enhance the transparency, quality, and reproducibility of educational research.*

Keywords: *FAIR principles, research data, educational sciences, data taxonomy, metadata, standardisation, criteria and indicators.*

References

1. Directorate-General for Research and Innovation, PwC EU Services. (2018). *Cost-benefit analysis for FAIR research data – Cost of not having FAIR research data*. Publications Office. DOI: <https://data.europa.eu/doi/10.2777/02999>.
2. Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ...& Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. DOI: <https://doi.org/10.1038/sdata.2016.18>.
3. Marienko, M., Shyshkina, M., & Konoval, O. (2022). Methodological principles of formation of cloud-oriented systems of open science in institutions of higher pedagogical

education. *Information Technologies and Learning Tools*, 89(3), 209-232. DOI: <https://doi.org/10.33407/itlt.v89i3.4981> [in Ukrainian].

4. Provost, L., Di Donato, F., Tóth-Czifra, E., Dumouchel, S., Blotière, E., & Chen, Y. (2023). Open Science Training in TRIPLE. *Open Research Europe*, 3, 39. DOI: <https://doi.org/10.12688/openreseurope.15430.1>.

5. Gualandi, B., Pareschi, L., & Peroni, S. (2022). What do we mean by "data"? A proposed classification of data types in the arts and humanities. *Journal of Documentation*, 79(7), 51-71. DOI: <https://doi.org/10.1108/jd-07-2022-0146>.

6. Spirin, O., & Oleksiuk, V. (2026). Taxonomy of FAIR Data for Scientific Research in Education. *Digital transformation of scientific and educational environments in conditions of martial law*, Reporting scientific conference of the Institute of Digitalization of Education of the National Academy of Pedagogical Sciences of Ukraine. Kyiv. DOI: <https://doi.org/10.5281/zenodo.18909544> [in Ukrainian].

7. Intersoft consulting. (n. d.). *General Data Protection Regulation (GDPR)*. Retrieved from <https://gdpr-info.eu/>.

8. Verkhovna Rada of Ukraine. (2010). *On the protection of personal data* (Act No. 2297-VI, June 1). Retrieved from <https://zakon.rada.gov.ua/laws/show/2297-17#Text> [in Ukrainian]

9. Brandt, O., Gauza, H., Kaltenbach, J., Müller, M. E., Schneider, G., & Zinn, C. (2024). A Minimal Metadata Schema and Its Tool to Improve the Searchableness of Research Data in Bioinformatics. *Journal of Library Metadata*, 24(3), 165-188. DOI: <https://doi.org/10.1080/19386389.2024.2338314>.

10. Devare, M., Aubert, C., Benites Alfaro, O. E., Perez Masias, I. O., & Laporte, M.-A. (2021). AgroFIMS: A Tool to Enable Digital Collection of Standards-Compliant FAIR Data. *Frontiers in Sustainable Food Systems*, 5, 726646. DOI: <https://doi.org/10.3389/fsufs.2021.726646>.

11. Provost, L., Bezuidenhout, L., Venkataraman, S., van der Lek, I., van Gelder, C., Kuchma, I., ...& Braukmann, R. (2024). Towards FAIRification of learning resources and catalogues – lessons learnt from research communities. *Frontiers in Education*, 9, 1390444. DOI: <https://doi.org/10.3389/educ.2024.1390444>.

12. NAUKA (n. d). *FAIR principles for research data*. Retrieved from <https://nauka.gov.ua/information/pryntsypy-fair-dlia-doslidnytskykh-danykh> [in Ukrainian].

13. Oleksiuk, V., Habrusiev, V., & Balyk, A. (2011). Some aspects of integration of web services of a higher educational institution. *Scientific notes of the Volodymyr Hnatyuk Ternopil National Pedagogical University. Series: Pedagogy*, 1, 228-234 [in Ukrainian].



This work is licensed under Creative Commons Attribution 4.0 International License