УДК: 001.103:021.61]:004.424.6

**Чуканова Світлана Олександрівна**

ORCID http://orcid.org/0000-0002-5717-5050

Кандидат педагогічних наук, завідувач сектору

Наукової бібліотеки, старший викладач

кафедри психології та педагогіки Національного університету

“Києво-Могилянська академія”

Сертифікований тренер Google for Education,

Тренер ТОВ “Академія Цифрового Розвитку”

Київ, Україна

email: [chukanovaso@ukma.edu.ua](mailto:chukanovaso@ukma.edu.ua)

**ЗДІЙСНЕННЯ ПОШУКУ ДАНИХ ДОСЛІДЖЕНЬ У GOOGLE DATASET SEARCH ДЛЯ ПОВТОРНОГО ВИКОРИСТАННЯ**

У доповіді розглянуто пошуковий механізм Google Dataset Search як ефективний інструмент пошуку наборів даних для повторного їх використання. Доведено, що цей пошуковий механізм індексує репозитарії для даних, що працюють на основі відкритого стандарту schema.org, що сприяє обміну даними досліджень та розвитку Відкритої науки.

*Ключові слова*: дані досліджень, Google Dataset Search, репозитарії даних, повторне використання даних, Відкрита наука.

**Постановка проблеми.** Накопичення даних у сучасному світі є актуальною проблемою, оскільки дані збираються швидше, ніж опрацьовуються. Саме тому важливо депонувати дані досліджень у репозитарії, адже такий підхід значно скоротить час на проведення повторних досліджень. Депонування даних досліджень у репозитарії має ряд переваг. По-перше, це дозволяє дослідникам заявити про своє дослідження і зробити його більш видимим для інших науковців. По-друге, це надає доступ іншим науковцям до даних для повторного їх використання. По-третє, це може значно скоротити час на проведення повторних досліджень, оскільки інші науковці зможуть використовувати вже наявні дані, а не збирати їх заново. По-четверте, це виключає дублювання одних і тих самих даних, що може призвести до отримання більш точних і надійних результатів досліджень. По-п’яте, це може зробити наукову комунікацію більш ефективною, оскільки науковці зможуть обмінюватися даними та результатами досліджень більш швидко і легко. Існує ряд різних репозитаріїв даних, у яких дослідники можуть депонувати свої дані. Деякі з найпопулярніших репозитаріїв: Dataverse Network, Dryad Digital Repository, Figshare, Zenodo.

**Виклад основного матеріалу**. Google Dataset Search – це пошуковий механізм, що використовується для пошуку наборів даних для можливого повторного використання. Пошук наборів даних можна здійснювати за ключовими словами, тегами та іншими критеріями. Google Dataset Search містить набори даних з усього світу, у тому числі з університетів, урядових установ, компаній та інших організацій. Набори даних доступні в різних форматах, у тому числі в CSV, JSON, XML та інших. Цей механізм шукає дані з репозитаріїв, які використовують відкритий стандарт Shema.org. Наразі існує можливість фільтрувати результати на основі типів наборів даних, які потрібні (наприклад, таблиці, зображення, текст). Крім того, продукт тепер доступний на мобільних пристроях, а розробники пошукового механізму значно покращили якість описів наборів даних.

Основні особливості Google Dataset Search включають:

1. Пошук даних: Google Dataset Search надає можливість шукати дані з різних джерел, таких як наукові публікації, громадські набори даних, статистичні звіти та інші джерела. Інструмент використовує пошукові запити, щоб знайти найбільш відповідні дані до потреб користувача. Пошук здійснюється переважно за ключовими словами, а отриманий пошуковий результат можна відфільтрувати за потреби.

2. Фільтрація і сортування: Користувачі можуть застосовувати різні фільтри та параметри сортування для точнішого пошуку даних. Наприклад, можна фільтрувати дані за роком публікації, автором, типом джерела та іншими параметрами. Фільтри можна також і скидати, якщо пошуковий результат після використання фільтрів стає менш релевантним, ніж до їх застосування.

3. Результати пошуку: Після виконання пошуку Google Data Search відображає результати з відповідними даними, включаючи метадані, відомості про джерело, посилання на деталізовані дані та іншу корисну інформацію. Ці результати можна зберегти, отримати бібліографічне посилання для використання їх у публікаціях або ж поділитись ними через різні канали комунікації, в тому числі й через соціальні мережі.

**Висновки**. Через значне накопичення даних в умовах сучасної Відкритої науки, управління даними досліджень і пошук даних для повторного використання стає досить складним завданням. Для того, щоб дані досліджень були доступні для повторного використання, вони мають бути підготовлені належним чином та депоновані у репозитарій для даних. Проте репозитарій ще має легко знаходитись для того, щоб дані теж можна було знайти легко і безперешкодно. З цією метою багато репозитаріїв наукових даних використовують стандарт schema.org, який своєю чергою використовується такими пошуковими механізмами як Google Dataset Search. Використання цього стандарту забезпечує більш ефективну індексацію таких репозитаріїв і як результат - кращу видимість даних досліджень.

Google Dataset Search - надійний пошуковий механізм, що ефективно індексує набори дослідницьких даних. Пошук у цьому інструменті зручний та простий з можливістю використовувати різноманітні фільтри: дата останнього оновлення, тип файлу, права використання, тема, вартість тощо. Інтерфейс зрозумілий та непереобтяжений зайвими функціями, що економить час на пошуки й не відволікає від результатів. Пошукові результати можна зберігати - для цього варто увійти у свій обліковий запис Google. Якщо в пошуковому результаті разом з даними згадуються статті або інші матеріали, де ці дані були опубліковані, то пошуковик надасть перепосилання на опис матеріалу в таких пошукових механізмах як Google Scholar. Отже, Google Dataset Searh є сучасним зручним пошуковим механізмом, що індексує перевірені джерела дослідницьких даних для більш ефективної наукової комунікації в контексті Відкритої Науки.

UDC: 001.103:021.61]:004.424.6

**Svitlana Chukanova**

ORCID http://orcid.org/0000-0002-5717-5050

Ph.D in Education, head of the Research library sector,

senior lecturer at the Department of

Psychology and Pedagogy of the

National University of Kyiv-Mohyla Academy

Google for Education Certified Trainer,

Trainer of "Academy of Digital Development" LLC

Kyiv, Ukraine

email: chukanovaso@ukma.edu.ua

**SEARCHING RESEARCH DATA IN GOOGLE DATASET SEARCH FOR REUSE**

The report examines the Google Dataset Search engine as an effective tool for finding datasets for reuse. This search engine has been proven to index data repositories based on the schema.org open standard, which facilitates the sharing of research data and the development of Open Science.

*Keywords:* research data, Google Dataset Search, data repositories, data reuse, Open Science.