

Прилуцький Олег Владиславович

*Харківський національний університет імені В.Н. Каразіна
біологічний факультет
61022, Україна, Харків, майдан Свободи, 4;
oleh.prylutskyi@gmail.com*

ВІДКРИТІ ДАНІ З БІОРИЗНОМАНІТТЯ В УХВАЛЕННІ РІШЕНЬ: ПЕРСПЕКТИВИ ВПРОВАДЖЕННЯ В УКРАЇНІ

Відкриті дані з біорізноманіття (далі ВДБР) сьогодні складають одну з основ ухвалення рішень в галузі охорони природи та природокористування.

Збір, накопичення, публікування та обмін ВДБР забезпечують як глобальні проекти (GBIF: Global Biodiversity Information Facility), так і локальні або спеціалізовані ресурси та бази даних. Також стрімко зростає роль citizen-science-проектів (iNaturalist, eBirds тощо), що, фактично, вже вийшли за межі сервісів для любительських спостережень і є потужними платформами накопичення ВДБР (iNaturalist; eBird: An Online Database of Bird Distribution and Abundance).

Оскільки навряд чи можливо уявити ситуацію, коли всі ВДБР у світі агрегуватимуться одним ресурсом, актуальною стає розробка спільних стандартів даних, що дало б змогу безшовно суміщати дані з різних ресурсів. Під спільними стандартами ми розуміємо а) узгоджені принципи організації даних; б) узгоджені термінологічні словники та цифрові формати даних; в) використання стандартованих технологій публікування та обміну даними. Розглянемо ці пункти детальніше.

Узгоджені принципи організації даних – це принципів домовленості щодо характеру даних та способів їх використання, без технічної деталізації. На сьогоднішній день основні принципи організації ВДБР детально викладені у тезах FAIR (Wilkinson et al., 2016), що є абрєвіатурою від слів Findable (дані надаються до пошуку), Accessible (доступні), Interoperable (зчитувані як людьми, так і машинами) та Reusable (мають чітке і прозоре ліцензування, що дає змогу використовувати їх повторно). Лише ті дані з біорізноманіття, що відповідають принципам FAIR data, можуть вважатися дійсно відкритими.

Узгоджені термінологічні словники та цифрові формати даних є необхідною передумовою для реалізації принципу Interoperable. Для ВДБР базовим вважаються словники термінології та форматів даних Darwin Core, що розроблені і підтримуються міжнародним об'єднанням Biodiversity Information Standards (Biodiversity Information Standards (TDWG)). Ці словники об'єднують рекомендації з уживання термінів і способів представлення інформації, що складає ВДБР. Наприклад, існує чимало способів записати одні й ті самі географічні координати знахідки живого організму: N43°38'19.39", 43 38 19.39, 43.63871944444445, 11T 561191mE. Використання різних форматів ускладнює обмін даними, адже потребуватиме не лише залучення конверторів координат, а й інструментів розпізнавання вжитих форматів. Таким чином, зростає ймовірність помилок геоприв'язки. Серед цього багатого вибору Словник Darwin Core рекомендує використовувати форму десяткового градуса у да-

тумі WGS84 (43.6387194) що є найпоширенішою у професійному геоінформаційному програмному забезпеченні, легко інтерпретується машинами. Подібним чином Словник визначає кращі практики й в інших типах даних.

Використання стандартованих технологій публікування та обміну даними (<https://sftp>, відкриті API) дає змогу безперешкодно зчитувати дані, а також використовувати дані, опубліковані на одних ресурсах, в роботі інших ресурсів, обмінюватися даними між ресурсами.

Наявність та доступність ВДБР самі по собі є важливою підосною в ухваленні рішень з охорони природи чи природокористування. Втім, все більшого поширення набувають практики використання інформації, похідної від точкових даних з біорізноманіття, наприклад, предиктивних моделей поширення видів. Свіжим прикладом може слугувати Map of Biodiversity Importance від проєкту NatureServe – спільна розробка корпорацій Esri© та Microsoft© (Map of Biodiversity Importance). На основі точкових ВДБР, наборів цифрових даних про умови середовища (детальні карти землекористування, біотопів, опадів, температур, цифрові моделі рельєфу, карти «коридорів» та «бар'єрів» на шляху міграцій тощо), із задіянням методів машинного навчання моделюються карти територій/акваторій, важливих для підтримання сталих популяцій рідкісних та загрожуваних видів. Подібні розробки, засновані на поєднанні відкритих даних, відтворюваних наукових методів та потужних обчислень відкривають нові обрії у царині практичної природоохорони.

Використання ВДБР в ухваленні рішень є необхідною і неминучою умовою наближення України до міжнародних практик управління природними ресурсами. Однак, можна окреслити низку перешкод та застережень, що наразі стоять на шляху імплементації цих практик. Їхнє розв'язання є першочерговим завданням.

Невнормованість правового статусу ВДБР

Нормативні документи України передбачають ведення державних кадастрів тваринного та рослинного світу (Кабінет міністрів України, 1994, 2006), однак жодної офіційної публічної бази даних біорізноманіття в Україні досі не існує. Мобілізація та оприлюднення ВДБР досі залишаються в нашій державі справою академічних інститутів та громадських об'єднань. Очевидно, що «неофіційні» дані, отримані з джерел без формалізованого статусу, можуть мати лише рекомендаційний характер у процесі ухвалення рішень.

Вразливі до браконьєрства види

Для багатьох вразливих видів розкриття точних координат знахідок може потягнути за собою збільшення браконьєрського добування. З одного боку, уникнути цього можна, задіявши «розмиття» координат – зумисне пониження точності геоприв'язки до рівня, що унеможливає знайдення локальної популяції за координатами. Однак, подібний метод ускладнює використання даних, наприклад, під час планування природоохоронних територій малого розміру або моделювання поширення видів. Можливим виходом з цієї колізії може бути дублювання даних, коли точні координати знахідок вразливих видів залишаються інформацією для службового користування, тоді як «розмиті» координати стають відкритими даними. Однак, це повертає нас до першого застереження – для зберігання точних координат потрібен офіційний ресурс, що підтримується державою, а не спільнотою.

Низька культура публікування даних у академічних колах

Відстаючи від передових світових наукових практик, українська академічна спільнота в більшості своїй тільки зараз дізнається про публікування первинних даних з біорізноманіття. В Україні немає жодного наукового журналу у форматі Data Paper, наявні форми звітності для наукових та науково-педагогічних працівників не заохочують публікувати набори даних та статті про дані.

Неоцифрованість біологічних колекцій

Колосальний пласт даних з біорізноманіття, що був зібраний поколіннями українських науковців і продовжує поповнюватися, залишається недоступним широкому загалові через неоцифрованість біологічних колекцій. Навіть та незначна частина матеріалів, що оцифрована, переважно не відповідає критеріям FAIR data.

Невизначений статус даних аматорів

Значну частину ВДБР складають дані, зібрані в рамках т.зв. «citizen-science» проєктів, тобто дані любителів природи та натуралістів, що не мають відношення до офіційних науково-дослідних установ. При цьому, перехресна перевірка та широке залучення професійних модераторів підносять якість значної частини цих даних на один рівень з даними «професійних» дослідників. Однак, лишається відкритим питання відповідальності за достовірність даних. Адже за достовірність даних, зібраних професійними науковцем в рамках виконання його посадових обов'язків, несе формальну відповідальність його установа. За достовірність даних «любительської науки» не несе формальної відповідальності ніхто, що ускладнює залучення цих даних у процес ухвалення рішень на рівні глибшому, ніж рекомендаційний.

Низька стандартованість наявних даних

Останньою з переліку, але не за вагомістю, перешкодою до впровадження ВДБР у природоохоронну практику України є їхнє розпорощення між різними ресурсами та базами даних, частина з яких не відповідають вимогам FAIR data частково або повністю, що унеможлиблює на теперішньому етапі консолідацію всіх наявних даних.

Які ж рекомендації можна сформулювати для подолання вищезгаданих перешкод? На нашу думку, першочерговими кроками з боку держави, академічної спільноти та громадськості мають бути:

1. Інтеграція України у світовий простір ВДБР, через офіційне приєднання України до ініціативи GBIF: Global Biodiversity Information Facility та запуску роботи Національного осередку GBIF, як центру координації подальшої роботи на основі кращих світових практик у цій царині;
2. Приведення наявних українських ресурсів з ВДБР до спільних стандартів, на основі критеріїв FAIR data й кращих міжнародних практик;
3. Розробка нормативної бази мобілізації, зберігання та використання ВДБР в ухваленні рішень, пов'язаних з охороною природи та природокористуванням;
4. Створення та запуск офіційного ресурсу ВДБР України, за принципом Державних кадастрів, що слугуватиме агрегатором всіх наявних достовірних ВДБР з території України;

5. Запуск державних програм з оцифрування біологічних колекцій;
6. Переведення наукової звітності для установ, завданням яких є вивчення біорізноманіття, з паперової форми у цифрову, відповідно до провідних практик роботи з ВДБР (електронний «Літопис природи» для установ ПЗФ тощо).
7. Запуск українських Data Paper – журналів; прирівнювання публікації набору ВДБР до фахової публікації.

Список використаних джерел.

1. Biodiversity Information Standards (TDWG). Available at: <https://www.tdwg.org/> (Accessed February 10, 2020).
2. eBird: An Online Database of Bird Distribution and Abundance. Cornell Lab of Ornithology. Available at: <http://www.ebird.org> (Accessed February 10, 2020).
3. GBIF: The Global Biodiversity Information Facility. What is GBIF? Available from <https://www.gbif.org/what-is-gbif> (Accessed February 10 2020).
4. iNaturalist. Available at: <https://www.inaturalist.org> (Accessed February 10, 2020).
5. Map of Biodiversity Importance. NatureServe. Available at: <https://www.natureserve.org/conservation-tools/projects/map-biodiversity-importance>. (Accessed February 10, 2020).
6. Wilkinson, M., Dumontier M., Aalbersberg I.J., Appleton G., Axton M., Baak A., Blomberg N. et al. 2016. The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Scientific Data*, 3(1): 1–9. <https://doi.org/10.1038/sdata.2016.18>.
7. Кабінет міністрів України. 1994. Постанова № 772 від 15 листопада 1994 р. «Про порядок ведення державного кадастру тваринного світу.»
8. Кабінет міністрів України. 2006. Постанова № 195 від 22 лютого 2006 р. «Про Затвердження Порядку Ведення Державного Обліку і Кадастру Рослинного Світу.»