

[Мовець 1]

Доброго дня. Я сподіваюся, ви бачите мій екран і мене чути. Доброго дня.

Я представляю Інститут теоретичної фізики. Ми фактично репозитарі відкритих даних Академії наук. Він розміщений в нашому інституті і ми забезпечуємо його супровід і промислову вже експлуатацію.

Але я хочу розпочати з інфраструктури відкритої науки. Ще один слайд. Згідно дорожньої карти, яка була прийнята в ЕОС Асоціації, інфраструктура відкритої науки будується на трьох рівнях.

Перший рівень – це європейський рівень. Він фінансується або програмою «Горизонт Європа», або безпосередньо європейською комісією і формує загальну доступність сервісів. Тобто ті сервіси, які об'єднують практично всю інформацію, яка є на нижніх рівнях.

Це і пошукові сервіси, і сервіси доступу до інфраструктури обчислення. Наступний рівень – це національний рівень. Тобто це відповідальність країни, яка повинна розробити, вести і підтримувати сервіси відкритої науки.

Це теж саме репозиторії, тематичні репозиторії, це пошукові сервіси, так звані харвестери, це журнали відкритого доступу. І наступний рівень – це інституційний, це елементи інфраструктури для представлення наукових результатів діяльності установ. Тобто, фактично, це частина, за яку відповідає установа.

Що це може бути? Це журнали відкритого доступу, які мають відповідне і програмне забезпечення, і протоколи обміну інформації. Це і репозиторії текстів, і харвестери.

Це приклад Фінляндії. Маленька країна, вона має настільки розголошену систему відкритої науки, яка пов'язана і є гарним прикладом побудови і експлуатації цієї системи. Із яких складових складається інфраструктура відкритої науки?

На слайді ви бачите, це практично всі частини. Основи журналів відкритого доступу або і репозиторії текстів, які містять метаданні про первоисточник цієї статті. Це архіви препринтів, це репозиторії даних, де зберігаються відкриті дані, відкриті освітлювальні ресурси, окрема інфраструктура для проведення обчислень і є відкрите програмне забезпечення.

Приклади на цьому слайді і фактично всі основні складові. Що стосується проекту Академії наук, створення і впровадження інфраструктури відкритої науки, то на початку, це 2023 рік, ми провели такий аналіз, що є в Академії наук, як це функціонує і чого недостатньо для побудови цієї інфраструктури. І фактично все, що виділене червоним кольором, цих елементів не було в Академії наук.

Перш за все, увага була приділена універсальній видавничій платформі, тобто перевод журналів, яких багато в Академії наук, на платформу Open Journal System, яка має відповідні протоколи обміну даними. Це побудова репозитарії відкритих даних, його теж у нас не було. Не було архіву при принті.

Був харвестер, але це була застаріла версія, і зараз це новий харвестер, який агрегує мета-дани і по публікаціям, і по відкритим даним. Я також удосконалював ті елементи, які були. Репозитарії наукових текстів, тобто бібліотека Вернадської Академії наук, ну і створення Центру компетенції управління даними як орган, будемо сказати, як структура, яка забезпечує навчання і консультацію по питанням.

Всі ці елементи дуже добре представлені у цій монографії, яка вийшла у цьому році. Це відкрита наука, візія Національної Академії наук. Там вона побудована по розділам, і фактично кожен розділ описує ту чи іншу складову частину інфраструктури відкритої науки, яка побудована за ці два роки.

Цей проєкт «Опенс» продовжується на наступні два роки, і основне завдання далі – це так званий технологічний розвиток. Також у нас є інформація, яка розповідає про розвиток технологічних розвитків, про розвиток інформації, про розвиток інформації, про розвиток інформації, про розвиток інформації, про розвиток інформації, про розвиток інформації. На даний час ви бачите першу сторінку – репозиторія відкритих даних.

Вона побудована на базі програмного забезпечення Dataverse. Він є частиною світової мережі репозиторій відкритих даних, їх більше сотні інсталяцій. На даний час у нас опубліковано 21 набір даних, ну і вже 9 установ зареєструвалися, частини ви бачите на слайді, частини там далі за прокритиками.

Як він працює? Ну це загальна схема роботи репозиторія, фактично дві частини – сам репозиторій Dataverse, це програмне забезпечення Dataverse, і сайт підтримки репозиторію. Ви бачите, що у нас є куратор даних репозиторії, тобто ви розумієте, що потрібен відповідний персонал, який забезпечує як функціонування чисто адміністрування, так і супровід користувачів, які розміщують дані.

Ну і є установа, в якій ми пропонуємо, щоб був куратор даних, який забезпечує підтримку тих, хто розміщує дані. Dataverse зареєстровано в організації DataSite, і дої ми отримуємо автоматично. Тобто, якщо набір даних розміщено в репозиторії, він відповідає всім вимогам FHIR, і він пройшов перевірку в DataSite, він отримує DOI.

Ми є учасником консорціуму, ми маємо поємний пул DOI, який ми надаємо на даний час безкоштовно для всіх установ Академії наук. Значить, що стосується тих функцій, які виконують ті, зазначені на слайде основно? Куратор даних.

Ну, перш за все, він забезпечує перевірку інформації в запиті. Користувач, який бажає розмістити дані в репозиторії, заповнює так званий запит на розміщення даних на сайті підтримки. Це окрема форма.

І куратор даних виконує перевірку. Це перевірка стосується як особи, так і тих інформацій, яку він надає. Далі він забезпечує консультації.

Ми називаємо це депозитором, тобто співробітником, який бажає розмістити дані. Він виконує перевірку описової інформації, тому що опис інформації по набору даних більше, ніж зазвичайні статті в науковому журналі. Це стосується, будемо казати, мінімального набору.

Якщо для статті це дублінкоре, певний набір поліг, то для даних це набір дата-сайт. Він більший, і він дає можливість описати не тільки автора, назву, а й додаткові мета-дані, як ці дані були отримані, на якій установці, які характеристики, як проводився експеримент. Це дає можливість формувати, будемо казати, великий набір мета-даних, які потім будуть використовуватися пошуковими системами.

І якщо перевірка завершилася і все добре, то куратор даних надає дозвіл на публікацію. Що стосується куратора даних установи, то це наша пропозиція, щоб у кожній установі був відповідний фахівець по управлінню даними, і він відповідав тільки за дослідників цієї установи. Частина у нас є, і в декількох установцях такі фахівці є.

Але якщо немає куратора даних в установі, то співробітники теж можуть розміщувати дані, і перевірку буде виконувати куратор Dataverse. Для кожної установи створюється відповідна колекція в термінах Dataverse, і всі набіри даних формуються в цій колекції, що є певне представлення наукової діяльності

установи. Але якщо цієї колекції немає, і немає куратора, то дані все одно можуть бути розміщені, вони будуть в верхній директорії.

Відповідно до того, що ми маємо куратор даних в установі, ми кажемо, що повинен призначати установу, а з боку репозиторії ми надаємо право на адміністрування наборів даних цієї установи. Він має можливість вносити зміни, звичайно, з консультацією автора, тому що він, автор, є відповідальним, і дає добро на публікацію цього набору даних. Як стати куратором даних?

На сайті Open Data BTP Kyiv, тобто сайт підтримки процесів, є реєстрація установи, окрема новка в частинній реєстрації, і якщо установа бажає, то вона має бути відповідальна. Вона має бути відповідальна, якщо вона має бути відповідальна, якщо вона має бути відповідальна, якщо вона має бути відповідальна. Вона має бути відповідальна, якщо вона має бути відповідальна, якщо вона має бути відповідальна, якщо вона має бути відповідальна, якщо вона має бути відповідальна.

Якщо вона має бути відповідальна, якщо вона має бути відповідальна, якщо вона має бути відповідальна, якщо вона має бути відповідальна, якщо вона має бути відповідальна, якщо вона має бути відповідальна, якщо вона має бути відповідальна. Що стосується розміщення даних? В час, який є, і згідно тих документів по відкритій науці, які були прийняті президією Академії науки, кожен дослідник, який бажає розмістити дані, повинен мати дозвіл від установи на розміщення даних.

Тому що майнові права на ці дані належать установі. Це, зазвичай, може бути звичайний експертний висновок, який ви робили на науковій статті, але без дозволу на розміщення даних ми не починаємо працювати. Якщо дослідник бажає розмістити дані, він отримує дозвіл від установи на розміщення даних.

В цьому дозвілі не є офіційна форма, але в ньому повинна бути вказана назва набору даних, перелік файлів, які він розміщує. Для того, щоб розмістити дані в репозиторії, кожен користувач повинен мати обліковий запис в репозиторії Dataverse. Цей обліковий запис надає можливість, перш за все, адміністратору змінювати права.

Зазвичайні права на перегляд змінюються на права депозитора і він має можливість потім завантажувати файл. В той версії, яка зараз є, яка доступна, реалізована реєстрація за допомогою власних облікових записів Google або RSEED, а для користувачів установи Академії наук реєстрація за допомогою облікового запису в системі RedNote, тобто в домені NASGOV-UA. І для цього можна вибрати посилання і обліковий запис створюється автоматично.

В поточній версії є ще дві можливості реєстрації на сайті бізнес-процесів, але ми закриємо цю можливість, тому що реєстрація з'явилася, що це шпаринка для спаму. Тобто, до кінця року ми, сподіваюся, внесемо зміни і, зазвичай, у кожного користувача є і RSEED, і Google, або є реєстрація в домені Академії наук, і це не є проблемою для реєстрації. Тобто, до кінця року ми, сподіваюся, внесемо зміни і у кожного користувача є і RSEED, і Google, або є реєстрація в домені Академії наук, і Google, або є реєстрація в домені Академії наук, і це не є проблемою для реєстрації.

Тобто, до кінця року ми, сподіваюся, внесемо зміни і у кожного користувача є і RSEED, і Google, або є реєстрація в домені Академії наук. І потім він надсилає цей договір адміністратору сайту. Наявність цих двох документів – це є початкова точка для розгляду того запиту і розміщення даних, які є в репозиторії. Далі депозитор, отримавши логін і пароль доступу до репозиторії і отримав відповідні права на розміщення, завантажує файли.

Як ці файли треба підготувати, буде розповідати Наталя Денисюк у своїй лекції, це буде 21-го. Завантажує описову інформацію, заповнює метадани. Після завантаження всіх цих файли надсилається повідомлення куратору репозиторія, що все готово і треба перевірити.

І тут куратор перевіряє, якщо необхідно, коригує, перевіряє відповідні консультації з депозитором. Як тільки куратор це перевіряє і дає згоду, що все, що там завантажено, всі необхідні поля заповнені, завантажені файли відповідають заявленому типу та формату, яка відповідає на ці файли. Витраплюю.

[Мовець 4]

Дякую, дякую. Мені дуже цікаво. От є питання...

[Мовець 3]

Дякую за таку докладну інструкцію Дякую за таку докладну інструкцію дуже гарно ілюстровано по-крокову. Єдине, що я не до кінця зрозуміла, можливо, не дочула, підкажіть, будь ласка, стосовно фінансування, для наукових установ відбувається ця реєстрація в репозиторії на безоплатній основі, чи ми маємо вносити якісь кошти за реєстрацію для того, щоб підтримувати...

[Мовець 1]

Ну, ви вже в реєстрацію, які кошти за реєстрацію? Ніяких коштів за реєстрацію немає, більше того. В цьому році немає коштів за тримання DOI, тому що, я вже казав, ми є учасниками світового консорту мандату BERCE, і ми сплатили як членський внесок, так і річний внесок, який дає можливість отримати певну кількість DOI передоплаченої.

І зараз нічого платити не треба. Да, коли ми виберемо ту кількість, яка нам надана за річний внесок, можливо, якась оплата буде, але я сподіваюся, що не буде.

[Мовець 3]

Дякую. Дозвольте ще одне запитання.

[Мовець 1]

Будь ласка.

[Мовець 3]

Якщо, скажімо, створений обліковий запис і користувач, ну, наприклад, так сталося, немає можливості у цей обліковий запис зайти, тому що там або втрачені паролі, або виникають якісь проблеми із входом, то до кого можна в таких випадках звертатися?

[Мовець 1]

Дуже просто. На сайті підтримки є контакти і є електронна адреса адміністратора DATEBERC. Ну, він пише лист, зазначає, що трапилось, а ми вже відновимо цей доступ.

Нема питань. Там є контакти.

[Мовець 3]

Дуже дякую.

[Мовець 4]

Так, ще запитання у когось є?

[Мовець 5]

Віктор Кузьменко, інститут кібернетики. Можна? Так, питання дотичне.

А от зараз за дої для статті ми платимо 90 гривень, а не можна чи буде через вас платити, або не платити взагалі, а платити менше?

[Мовець 1]

Ні, не можна. Тому згідно цієї угоди ми надаємо дої тільки для наборів даних. Ще для підприємців.

А, ще для підприємців. Це інша доповідь. Якщо ви дивилися, то на Zenodo, хоча він розроблявся як репозиторій відкритих даних, там є і наукові статті.

Але ми на даний час наукові статті не розміщуємо в репозиторії даних, тільки даних.

[Мовець 6]

Дякую.

[Мовець 4]

Ще питання, будь ласка. Бойко, Іслам Миколаївич.

[Мовець 6]

Важне питання, вийдемо з іншими пошуковими базами. Ви тестували, чи вийдемо із вашої бази з Google Академії?

[Мовець 1]

На даний час ви кажете про репозиторії. Репозиторії зареєстровано в OpenDoor, і вони там є. Репозиторії зареєстровано в OpenAIR.

[Мовець 6]

Ні, пошукові системи, наприклад, Google Академія, чи відомі дані для Google Академії, ваші дані?

[Мовець 1]

Я кажу про європейські сервіси. Як тільки ми зареєстровані в європейських сервісах, ми автоматично пройдемо реєстрацію і в інших сервісах. Ми зареєстровані в нашому академічному каравестері, і він вже надав інформацію про репозиторії пошукових баз.

Тобто це такий поступовий процес. Я сподіваюся, що ми там теж будемо.

[Мовець 4]

Ми перевіримо, В'ячеслав Миколаївич, ми не перевіряли, чи Google Scholar нас бачить чи ні. Наступне питання Скопненко, будь ласка, Олександр.

[Мовець 2]

Дякую. Дякую вам за таку докладну інформацію. У мене питання.

Я щойно почув, що ви статей не розміщуєте у репозиторії. Правильно?

[Мовець 1]

Правильно.

[Мовець 2]

Тоді в мене таке питання. Наукові установи гуманітарного профілю. Я речник Інституту мовознавства імені протидні Національної академії наук України.

Що можуть у вашому репозиторії розмістити установи гуманітарного профілю, якщо ви не розміщуєте статей? Дякую.

[Мовець 1]

Ну, якщо ви зайдете на репозиторії, ви там побачите Сумський державний університет, вони розміщують дані досліджень соціологічних. Це ми розміщуємо. Тому що є відповідні мета-дані в Dataverse.

Що стосується...

[Мовець 2]

Даруйте, але соціологія і лінгвістика — це різні галузі науки. У нас немає соціології. Що ж ми будемо розміщувати тоді у вас?

[Мовець 1]

Ну, ви можете розміщувати ваші статті в журналах відкритого доступу. Ви можете розміщувати ваші статті в архіві припринтів, які є. Але зараз ми не розміщуємо статті в репозиторії.

Це, будемо казати, вважайте тимчасове рішення, або, може, ні, тому що це тільки репозиторії відкритих дан.

[Мовець 2]

У мене тоді до вас прохання до організаторів і, може, до вас як керівника репозиторію Національної академії наук. Чи не були б ви такі люб'язні зробити інструкцію для установ гуманітарного профілю? Бо я думаю, що це проблема не тільки, от я помітив, стосовно нашого інституту, а стосовно всіх установ гуманітарного профілю.

Установи такого типу можуть у вас розміщуватися.

Тому що одна річ — це дані установ фізичного профілю, наприклад, а зовсім інше — мовознавці, літературознавці, мистецтвознавці, етнологи, тощо. Будь ласка, підготуйте якісь поради, щоб ми знали, як реагувати.

[Мовець 1]

Дякую. Я розумію вас, ми подумаємо над цим. Я не фахівець цього питання, я не можу відповісти, що ви можете розміщуватися, але ми подумаємо над цим.

[Мовець 2]

Може, я відновим, але гуманітарні дисципліни, вони ж часто використовують якісь експедиції, те ж саме етнографічні, де збирається якийсь матеріал. З приводу мовознавства теж, як би, як діалекти. Якщо в цих експедиціях ви не збирали якийсь матеріал, який потім будете обробляти, писати статті, то цей матеріал можна викласти в репозиторії дата-версії.

Ну, якось так, так.

[Мовець 4]

Пробачте, що знову питаю. Мовознавство – це цілий комплекс розгалужених відгалужень наук. Мовознавство – це не тільки діалектологія.

Це, наприклад, історичне мовознавство і таке інше, і таке інше, і багато різних особливостей є у цих дослідженнях мовознавчих. Тому треба, щоб фахівці дійшли згоди, що ми можемо розміщувати, бо ми хотіли, бо це вимога часу, зрештою, що ми були у репозиторії, але що ми можемо розмістити, якщо статей не можна, монографій не можна, а як ми можемо себе показати?

[Мовець 2]

Дуже дякую. Гуманітарна наука – це звичайна наука, де є експериментальні якісь дані, які збираються в експедиціях. Якщо Ваш інститут проводить експедиції, то дані, які отримуються в період експедиції, можна викласти в дата-версію, а аналіз цих даних, він вже буде опублікований в статті з посиланням на дої Ваших даних, які Ви розміщуєте в дата-версії.

Ще які питання? Вибачаюся.

[Мовець 1]

Там питання, що стосується, наскільки дані розміщені в репозиторії Академії наук захищені від неправомірного використання. Чи захищені такі дані від доступу з території країн-агресора? Перше, я відповів, що дані, які розміщені, ми кажемо про дані відкриті, але в репозиторії є певна кількість ліцензій, які обирають даних.

Це може бути і ембарго, це може бути і закриті дані, і в дата-версії є відповідний механізм, який захищає дані від завантаження, якщо ця ліцензія цього не дозволяє. Що стосується країн-агресорів, ну, перш за все, провайдер Академії наук, це його відповідальність, який блокує сайти і запити, які йдуть з боку країни-агресора. Але у нас є і певний файрвол, і ми теж, ну, будемо казати, менше, ніж провайдер, теж ці запити блокуємо.

[Мовець 2]

Добре, то була б Білочкіна питання від вас, чи я помиляюсь? Добре, я бачу.

[Мовець 1]

Ну, давайте, це є ще питання в чаті, а є, Юлія, давайте.

[Мовець 3]

Ще раз, дивіться, питання, в нас гуманітарний інститут, історики, мовознавці, літературознавці, і питання щодо розміщення, чи ми можемо, наприклад, розміщувати архіви, наприклад, оцифровані автором, чи інтерв'ю, але, наприклад, є специфіка інтерв'ю, які записані, які вже не можна попросити згоду на використання інтерв'ю, наприклад, там 80-х, 90-х роках наші працівники записували, і ми, наприклад, розміщуємо архіви, розміщуємо інтерв'ю, розміщуємо інтерв'ю, розміщуємо інтерв'ю,

розміщуємо інтерв'ю, розміщуємо інтерв'ю, розміщуємо інтерв'ю, розміщуємо інтерв'ю, розміщуємо інтерв'ю, розміщуємо інтерв'ю, розміщуємо інтерв'ю.

[Мовець 1]

Так, я розумію, є різні аспекти цього, давайте це залишимо вже поза цією вебінарою, ми подумаємо, як це може бути, спілкуючись з вами, з авторами, ви покажете, які дані, наші контакти є, напишіть на пошту, будемо спілкуватися. І ще питання пов'язане, якщо дані не експериментальні, отримані комп'ютером моделюванням чи теоретичними розрахунками, іншими методами синтезування, чи можна вважати такими, що розміщують? Так, можна.

Дані комп'ютером моделювання є відповідний набір метаданих, і якщо ви зайдете на сайт, ви побачите, що такий набір даних вже розміщений в репозиторії відкритих даних.

[Мовець 2]

Добре, ще питання є?

[Мовець 7]

Так, будь ласка, підкажіть, будь ласка, я побачив, може це випадковість, яка різниця між словами репозиторій і репозиторії? Навіть в програмці такі є, може це помилка, а може ні.

[Мовець 1]

Є така частина, ми використовували вперше репозиторії, але фахівці, академперіологи сказали, що треба використовувати репозиторії. Я не можу відповісти вам, чому. І всі ці роки ми використовуємо репозиторії.

Це те ж саме, але інша анотація, ну не знаю.

[Мовець 2]

Дякую, дякую.

[Мовець 6]

Я хотіла уточнити, чому не можна використовувати репозиторії. Чи я зрозуміла, в академії наразі буде два репозиторії відкритих даних і архів репринтів? Чи так?

Це два різних репозиторії, вони будуть під керуванням різних організацій, чи це все буде інститут теоретичної фізики, як це буде? Дякую.

[Мовець 1]

Я вам відповім. Так, це два різних репозиторії. Репозиторій відкритих даних, він розміщений в інституті теоретичної фізики, в нашому інституті Багалюбова, а архів репринтів, теж можна сказати, що репозиторій, це відповідальність інституту кібернетики.

Він розміщений там, і фахівці, будемо казати, відповідають за наповнення і за експлуатацію. Що стосується DOI, DOI надається дата-сайт для архіва репринтів і для відкритих даних однієї установи тому, що так прописано в той угоді, яку ми уклали.

[Мовець 6]

Будуть різні DOI? Чи одне це те саме?

[Мовець 1]

Ні, організація, яка видає DOI, одна, а DOI це різний, звичайно, це різні префікси, це різні репозиторії, там у архіві свій префікс, у відкритих даних свій префікс.

[Мовець 2]

Добре, ще питання, якщо немає, то надаю слово Вадиму Тульчинському, який повинен розповісти нам про архів репринтів. Я його не бачу, Сергій Якович, мабуть, може, у нього з інтернетом, і ми переносимо тоді на завтра його.

[Мовець 1]

Так, я його теж не бачу.

[Мовець 5]

Вибачте, а можна ще одне питання? Так, так, так, так. Владислав Сіденко, Український гідрометалургічний інститут.

Питання полягає в наступному. Ми в інституті створили свій репозиторій для кліматичних даних. Він також заснований на принципах FAIR та Open Science, які були описані в попередніх даних.

Не настільки він, можливо, має складну структуру сайту і так далі, але також ми там присвоїли і отримали дої для кожного з наборів даних. Наразі маємо чотири набори. Чи доцільно нам, наприклад, дублювати чи також подавати у ваш репозиторій?

[Мовець 1]

Ну, як на мене, я вважаю, що це недоцільно, тому що чим більше репозиторіїв у вашій репозиторії, тим більше це відноситься до тематичного. Там специфічні набори даних. Ми кажемо, що дата верси наші репозиторії відкритих даних ми позиціонуємо як мультидисциплінарні.

Але ваш репозиторій, щоб він був доступний для пошуку і запитів в мережі EOСКА, його треба зареєструвати. Перш за все зареєструвати в Харвестері Академії Наук, і він буде, відомо, зареєструвати його в OpenDoor. Процедура не складна, але потребує певний час.

І все. Якщо у вас є можливість технічно, організаційно і кадрово підтримувати ці репозиторії, не треба дублювати. Чим більше репозиторій в Академії Наук, тим краще.

Це мій погляд. Дякую за відповідь.