

[Мовець 1]

Доброго дня усім. Вітаю присутніх учасників вебінару. Те, що ми зараз з вами будемо робити, це розбиратися із загальними питаннями відкритої науки і FAIR-даних як один із ключових концептів цієї відкритої науки.

Моя задача дати загальний контекст для того, щоб всі подальші практичні питання, які будуть розглядатися наступними, вони були зрозумілими і ясно було, для чого це існує, навіщо це потрібно, в яку картину світу воно вписується. Отже, ми почнемо з загальних питань. Зараз вам має бути видно моя презентація.

Підтвердіть, будь ласка. Зараз переключаюся. Дякую за відгук.

Отже, почнемо з того, що відкрита наука це такий особливий спосіб, яким чином організувати і реалізувати науково-технічну діяльність так, щоб був доступ до об'єднаного дослідницької інфраструктури, до отриманих наукових результатів, науково-технічної інформації з можливістю багаторазового використання цих результатів і можливістю обміну та розповсюдження даних за допомогою інформаційно-комунікаційних технологій, для того, щоб розвивалася сфера науково-технічна і поглиблювалася співпраця і дотримувалися авторські суміжні права. Чому це виникло?

У зв'язку з чим? Взагалі це результат поступового еволюційного розвитку, до якого саме життя призвело. Тобто це не штучна якась концепція, придумана просто для вдосконалення управління і менеджменту, а були об'єктивні передумови і чинники, які призвели просто до необхідності саме такої побудови сфери наукової та науково-технічної діяльності.

Передусім це потужне зростання наукових знань, їх кількості, обсягів. Спеціалізація науки призвела до того, що необхідна міждисциплінарність і колаборація в дослідженнях. Мало того, оці обсяги знань, це величезна кількість потужностей вчених, можливостей займатися такою діяльністю, призвела до того, що спостерігається добування досліджень.

А це дуже значні витрати праці, грошей, часу і з цим треба було щось робити. Крім того, необхідно забезпечувати достовірність даних і якимось чином дивитися, наскільки вони відтворювані, наскільки ті ресурси, ті результати, які отримані, можуть бути відтворені в інших дослідженнях, чи вони дійсно так працюють. Також можливість використання вже отриманих результатів для того, щоб їх розвивати і проводити наступні дослідження.

Ну і, звісно, розвиток цифрових технологій дозволяв оперувати великими масовими даних, систематизувати, збирати, формувати цифрові каталоги, забезпечувати пошук, обмін, проводити обчислення даних. Це з одного боку. Тобто, технічні можливості, об'єктивні потреби, пов'язані з паціфікою взагалі самих досліджень.

А з іншого боку, з'явилися такі глобальні виклики, які стали ще більш масштабними, ніж, мабуть, раніше, які стосуються сталого розвитку, кліматичних змін, звісно, ризиків пандемії. І виклик потребує того, щоб були відповідні наукові відповіді. І ось відкрита наука – це також той концепт, який дозволяє проводити колаборацію і шукати спільно, і розбудовувати інфраструктури спільні, які будуть дозволяти проводити дослідження і шукати відповіді на виклики.

Сучасного розвитку. Крім того, спільстві, тому що потреба в управлінні, в державному управлінні, в корпоративному управлінні, підтримка рішень наукових, так само вона штовхає до того, щоб результати досліджень стали доступними. Ну і, звісно, формування простору довіри до науки, тому що складається враження, якщо немає довіри до науки, що це зайве витрачання, взагалі, ресурсів.

Наприклад, той же самий виклик пандемії. Він продемонстрував те, що виникла ситуація, виникла загроза людям на планеті взагалі, а в багатьох країнах світу науковці не змогли дати відповіді. А ті відповіді, які були дані, вони виявилися не такими, як очікувалось, і це дуже серйозно підбиває довіру до науки, до знань.

Відповідно, необхідність того, щоб громадськість розуміла і цінність науки, і її особливості, і була залучена до наукового знання, до наукових досліджень. Плюс іще питання про шуку талантів. Все це спонукає і формує отой простір, які передумови, які підштовхнули до ось цього потужного розвитку цього руху, який можна звати відкритою наукою.

Історичні видаки. Ми не будемо говорити про ті передумови, ті дискусії, розмови, які відбувалися до початку 2000-х років, тому що в колах наукових і в колах бібліотечних точилися дискусії про те, що робити з цими даними, з результатами досліджень і так далі. І ось інституційно, фактично, такою поворотною точкою можна вважати зустріч в Будапешті, результатом якої стала Будапештська ініціатива відкритого доступу, де було обговорено ключові питання відкритого доступу і офіційно визначений термін відкритий доступ і сформульовано ключові принципи і засади доступності результатів досліджень.

В 2003 році Берлінська декларація, вона перевела оцю дискусію і ці розмови до політики інституції, тобто це не просто рух асоціації, рух громадських активістів, науковців, а це вже політика інституційна. 2008 рік найкрупніший, найбільший науковий фонд світу National Institute of Health США приходить на обов'язковий відкритий доступ і це вже переміщує питання відкритого доступу на регуляторний рівень. 2011 рік створюється Research Data Alliance, це глобальна платформа, яка забезпечує інтероперабельність та відкритість дослідницьких даних і це вже означає, що починає формуватися інфраструктура, яка дозволяє втілити загальні принципи відкритого доступу.

В 2013 році в програмі Horizon 2020 в ЄС було запроваджено політику Open Access і це стало обов'язковою фактично вимогою до тих результатів, які отримані через виконання європейських грантів. І це вже цілий регіон, європейський регіон, який охоплений запровадженням цих практик відкритого доступу. В 2016 році ще один крок в розбудові інфраструктури відкритих даних пов'язаний з тим, що запроваджена стандартизація і основа того, щоб дані могли бути відкритими.

А ми будемо про феопринципи говорити детальніше от сьогодні. 2018 рік найбільші європейські фонди зібрались на з'їзд і вони вимагали негайного відкриття доступу до всіх публікацій. Тобто це вже жорстка вимога до переходу до повної відкритості.

І Коалішн-С вона достатньо відома, досить відома, дуже навіть відома ця ініціатива. І вона підштовхує до всезагального використання підходів відкритої науки. Дуже важлива позиція для України, тому що на засіданні ЕОСКУ, це Європейська хмара відкритої науки, Інститут теоретичної фізики Мені Боголюбова став мандатною організацією, яка представляє Україну.

21 рік рекомендації ЮНЕСКО щодо відкритої науки, це на рівні міжнародному, всесвітньому рівні, були прийняті одноголосно, спільні принципи і стандарти відкритої науки. І це фактично глобальне визнання цього року і цього підходу Open Science. В 2022 році Україна офіційно приєднується до цих рекомендацій ЮНЕСКО, і у нас починається сфера більш потужно розвиватись.

Ще недостатньо, але менше з тим, це закладає таку рамку, основу, міжнародні зобов'язання України щодо розвитку відкритої науки. В 2003-2004 році в Horizon Europe і відкритий простір європейської науки було запроваджено такий принцип, що дані, публікації, методи, вони мають бути відкриті за замовчуванням, медзя на увазі ті, які виконуються в рамках програм, грантових програм. Тобто фактично відкритість стає таким операційним стандартом для сучасних наукових проєктів.

Тобто ми бачимо поступовий розвиток від обговорення ініціативи до появи технічної можливостей, створення інфраструктури і закріплення в міжнародних документах ключових позицій, які розвивають простір відкритої науки, принципи відкритої науки. Ключові стовпи, на чому взагалі базується ця відкрита наука. З технічної точки зору, з інфраструктурної точки зору утворена європейська хмара відкритої науки, ЄОСК, це федерація різних європейських ініціатив, інфраструктури, сервісів.

Фактично це така екосистема, в якій відбуваються ці процеси, яка опосередковує їх. Для того, щоб дані і результати досліджень були доступними, у 2016 рік була запроваджена концепція FAIR даних, яка передбачає підходи такої побудови даних і такого їх управління ними, щоб вони були знаходжуваними, відкритими, інтероперабельними і можливість їх подальшого використання. Ще раз ми будемо про це говорити сьогодні.

Крім того, з'являється потреба розміщення результатів дослідження в відкритому доступі, а раз так, то оскільки процес рецензування достатньо тривалий, то дуже потужно починає використовуватись практика припринтів, відкриття припринтів, практика відкритого рецензування, коли в відкритому доступі рецензенти можуть залишати свої рецензії і автори можуть так само відкрито відповідати. І, звісно, ця відкритість результатів дослідження припринтів спонукає до необхідності доброчесності досліджень і дотримання етичних принципів, розробки етичних кодексів. Незважаючи на те, що наявні ліцензії, які рекламують основні правила доступу, менше з тим, питання доброчесності нікуди не дівається, оскільки немає тих органів, які слідкують за їх дотриманням і проводять моніторинги, і вживають якихось інших заходів.

Фактично, залишається інструмент етики і інструмент суду. І доброчесність нікуди не зникає, а навпаки посилюється надзвичайно потужно через ось оцей рух відкритої науки. Крім того, відкрита наука означає, що треба якимось чином оцінювати внесок науковців і думати, які мають бути моделі фінансування науки і стимулів, окрім того, що забезпечені фінансово, ще й стимулів для всіх дослідників та установників, які працюють в рамці відкритої науки.

Також нові покоління метрик, які фактично обліковують, систематизують, визначають індекси для наукових результатів, наукових досягнень. Вони отримали надзвичайний пошук завдяки тому, що ось є оця відкрита наука. Також з'являється потреба в освіті і навичках для того, щоб були ті фахівці, які здатні управляти даними і забезпечувати роботу цієї інфраструктури відкритої науки, а також користувачі, які можуть її використовувати і відповідним чином імплементувати всі ці підходи.

Також пропагування відкритої науки, тому що про це потрібно знати і розповідати, і формувати розуміння цього. Ну і, звісно, громадянська наука. Ми не можемо минути громадськості широкого кола людей, які можуть бути залучені навіть до наукових досліджень, наприклад, збір даних, спостереження в процесах, наприклад, дослідження космосу, аналіз знімків, які отримані, відеоматеріалів, спостереження за природою і так далі.

І, з іншого боку, не тільки залучення для досліджень, але і розмова з людьми доступною мовою, поширення того, що називається научпоп, що пропагує дані, що має пояснювати так, щоб людям було зрозуміло, що це означає. Це не тільки того, що наука показує, що ми ціні, ми маємо якісь результати, але це й можливість пошуку талантів, це можливість інтересу, це майбутнє, це люди, які потім прийдуть і будуть працювати, стануть дослідниками, науковцями і так далі. Тобто, громадянська наука, мені здається, у нас в Україні недостатньо розвинена, і я б воліла, щоб цьому приділяли значно більше уваги, з'явилося значно більше інфлюенсерів і наукових журналістів, які це просувають.

Йдемо далі. Європейська хмара відкритої науки. Це така екосистема, це інфраструктура і віртуальний простір, який створює це середовище, де відбувається, в чому, в цьому просторі, за допомогою

інструментів, якого можливі оці публікації відкриті, знаходження інформації про результати і про дані, повторне використання, де пропонуються відповідні інструменти та сервіси.

Фактично, завдяки ЄОСКО відбувається мобілізація згодження ресурсів європейських для того, щоб прискурити впровадження відкритої науки, забезпечити відтворюваність досліджень, аналізу, обробку даних. І ця інфраструктура, вона до кінця ще не вибудована, вона розвивається. Фактично, ЄОСКО будується як об'єднання, як така федерація, яка включає національні і регіональні системи, які об'єднуються в такі вузли, ноди, які ще тільки формуються, і, власне, вони мають працювати спільно за спільними правилами сумісності.

Європейський Союз визначає ЄОСКО як один із десяти загальних просторів європейських, цифрових. І він, фактично, така ключова екосистема для того, щоб все це втілювалося в життя. Значить, що дає ЄОСКО для окремих дослідників?

По-перше, можна бути контраб'ютером, той, хто робить внесок, і отримати видимість і довіру, сформувані довіру до отриманих результатів. Завдяки наявності репозиторів, завдяки присвоєнню цифрового ідентифікатора, який дозволяє і знайти ці дані, і далі використовувати, отримати відкрите рецензування, спілкуватися з рецензентом та додавати відповіді на них. Крім того, це колаборація з іншими дослідниками, також використовувати результати, які отримані іншими науковими установами, групами дослідників в рамках проєктів, наприклад, і доступ до масивів і даних, і публікацій, і сервісів.

Ну і завдяки спільним правилам ФЕР, які імплементуються, можна забезпечувати і знаходжуваність, і відтворюваність, і доступ, і збереження даних, і результатів досліджень. Україна. Що в Україні?

Я вже говорила, що Інститут юридичної фізики імені Боголюбова, він є так званою мандатною організацією. Мандатна організація – це уповноважена організація, вона уповноважена, ну, за визначенням вона має бути уповноважена урядом представляти інтереси країни. І вона, мандатна організація, координує участь дослідників і установ країни в ІОСК, і представляє в асоціації ІОСК.

Це одна з складових, один з елементів ІОСК – це асоціація. І також має можливість, має право підписувати угоду від імені держави. Звісно, така організація мандатна повинна відповідати всім європейським практикам, пов'язаним з ідентифікацією, з принципами інтероперабельності.

Що у нас ще є в Україні? У нас Міністерство освіти і науки має портал «Відкрита наука», де загальна рамка представлена. Ви отримаєте презентацію, там будуть посилання, зможете зайти подивитися.

Також у нас є концепція реалізації європейських принципів відкритої науки Національною академією наук до 30-го року, яка була затверджена в 2023 році, де було проголошено, що навіщо реалізовувати принципи відкритої науки. Основні задачі – це поліпшити видимість результатів досліджень науковців, розширити доступ у науковій спільноті і в Україні, і, звісно, на міжнародному рівні до результатів досліджень, проведених на Україні. Підтримка використання результатів досліджень і їх поширення, тому що оцей доступ до даних і можливість їх використання в наступних дослідженнях, вона сприяє розвитку цієї сфери діяльності.

Також доступ до даних, він забезпечує, уможлиблює їх перевірку і встановлення достовірності, надійності і відтворюваності отриманих наукових результатів. І ще одна завдання цієї концепції – це побудова інфраструктури відкритої науки, про що сьогодні також, я думаю, Сергій Якович вам більш детально розповість про ось цю інфраструктуру. Є проєкт «Оупенс», який був спрямований на реалізацію концепції європейських принципів відкритої науки в нанукраїні.

«Оупенс-1» був у 2003-2004 роках, зараз цей проєкт «Оупенс-2» на 2005-2006 роки, в рамках яких відбуваються ось ці семінари. Також можна зайти на портал відкритої науки нанукраїни. Значить, відкрита наука, ось так виглядає цей портал, але що найважливіше, найцікавіше?

Оце трошки вище, ніж надпис відкрита наука, портал відкритої науки. Якщо ми зайдемо туди, в цей розділ, то ми відкриємо відкрита наука в Україні, де ми побачимо репозитор відкритих даних Dataverse, архів припринтів і центр компетенцій з відкритої науки та FAIR-даних і ще інші інструменти. Я думаю, що це буде один з найбільш корисних інструментів для вас, щоб ви могли використовувати, аналізувати, знайомитися з даними і мати доступ до сервісів, які пропонують і тієї інфраструктури, яка вибудовується наразі в НАН України.

Взагалі, якщо ми говоримо про інфраструктуру відкритої науки в НАН України, то вона загалом повинна включати цілий набір досить великих ресурсів, які показують шлях автора, який має результати своїх досліджень у вигляді наборів даних, статті, яка може бути опублікована як припринт. Журнали і відповідно два види репозиторів, репозиторії відкритих даних і репозиторії наукових текстів. Фактично є службові елементи, електронні бібліотеки і Харверстер, як той інструмент, який, власне, збирає метаданні і створює таку єдину точку для пошуку оцих результатів досліджень і наборів даних і статей, і забезпечує зв'язок України з європейською хмарою відкритої науки.

Власне, Сергій Якович розповість вам більш детально про те, як це функціонує, і я думаю, що ви зможете поставити більш детальні запитання. Але важливо розуміти, що це велика інфраструктура, яка потребує і відповідного ресурсного забезпечення, і кадрового забезпечення тих, хто спроможний функціонувати і працювати в цьому середовищі, і також обізнаності науковців, які мають розуміти, яким чином рухаються дані, інформація, і забезпечується доступ до цих даних, і отримати для себе можливості використовувати відкриті дані від інших науковців, дослідників установ. Це була найбільш загальна рамка про відкриту науку. Тепер про відкриті дані, або так звані не відкриті дані, я помилуюся, це FAIR-дані.

І зараз поговоримо, що означає FAIR-дані і відкриті дані. FAIR-дані концепт був прийнятий в 2016 році, і це один з найважливіших інструментів, який повинен забезпечити можливість того, щоб дані в умовах відкритої науки належним чином могли використовуватись і управлятись. Але почну я, мабуть, з відкритих даних, тому що FAIR-дані і відкриті дані – це не одне і те саме.

Відкриті дані – це дані, які оприлюднені у цифровому вигляді, вони отримані в ході дослідження, і вони можливі для скачування, для завантаження, для обробки іншими користувачами, без серйозних фінансових, юридичних, технічних перешкод, лише на умовах етики належного цитування – це відкриті дані. А FAIR-дані – це дещо інше. FAIR-дані – це такі дані, які вони структуровані певним чином, описані, задокументовані, ліцензовані, і вони організовані таким чином, щоб їх можна було знайти, щоб був доступ до них, щоб можна було їх повторно використовувати як людьми, так і комп'ютерними системами.

Тобто, FAIR-дані – це дані, які відповідають оцім принципам – знаходжуваність, доступність, інтероперабельність, повторне використання. Власне, це аббревіатура і оці знаходжуваність, доступність і так далі – це, власне, перші літери ось цього слова FAIR. Що вони дають?

Ось така організація даних. Це значить, що їх можна легко знайти і багаторазово використовувати, якщо надано таке право. Такі дані, оскільки вони, якщо знаходжувані, то це підвищення рівенцитованості.

Також ці дані можна повторно використовувати, відтворювати в дослідженнях і формувати довіру до результатів. І FAIR-дані на сьогодні – це вимога і грантодавців, і журналів. І детальніше про кожне з оцих FAIR, що означає кожна позиція і навіщо вона потрібна, трошки детальніше.

Почнемо з того, що принцип FAIR, він потрібен, передусім, на етапі поширення результатів досліджень. Тому що етап до їх поширення не потребує, власне, ось цього FAIR. І ці позиції треба розділяти, тому що управління даними на принципи FAIR – це ж затрати часу і затрати праці.

І тому у зв'язку з цим доцільно розуміти, що на етапі поширення і виходу у відкритий світ, дані мають бути FAIR. До того вони можуть управлятися і на інших принципах. Отже, findability – це відшукованість даних, це можливість легко знайти ці дані і людьми, і комп'ютерними системами.

І для того, щоб вони могли бути такими доступними знаходжуваними, вони повинні мати унікальні ідентифікатори. А також спеціальний опис даних, тобто мета-дані. Це два ключових принципи, які забезпечують видимість даних і можливість їх пошуку.

Якщо цього не зроблено, то дані будуть невидимими. Далі, доступність, збереження і можливість завантаження. Якщо є доступність, то на яких правах і на яких правилах ця доступність додається.

Мають бути відповідні протоколи, стандартизовані протоколи для доступності даних. І також дуже важливий момент. Якщо навіть самі набори даних були видалені, то мета-дані повинні бути завжди доступними.

Навіть з позиції того, що встановити, що такі дослідження вже проводилися. Де вони проводилися, ким вони проводилися. І якщо дані не доступні, то можна звернутися до дослідників, дізнатися, чому вони були видалені? Можливо, там були допущені якісь помилки або з якихось інших причин, але можна відновити історію, можна бачити, як це відбувалося і уникати помилок, або використовувати вже знаючі ті наслідки для наступних своїх досліджень. Принцип інтероперабельності.

Досліного перекладу на нашу мову немає, тому найчастіше говорить інтероперабельність – це дані, готові до поєднання з іншими пакетами даних або інтеграції до інших досліджень для того, щоб можна було легко інтегрувати отримані результати, отримані дані з іншими даними. А інтероперабельність передбачає використання особливих форматів сумісних і також загально вживаних мов словників. І взагалі формування словників – один з досить різних викликів для України, для того, щоб можна було належним чином все ж таки поєднувати і інтегрувати дані досліджень.

Резуабіліті, тобто повторне використання, можливість використовувати в подальших дослідженнях з використанням обчислювальних методів. І повторне використання потребує, звісно, і ліцензування, але також і можливість скачування цих даних в таких форматах, які дозволяли їх подальшу обробку. Коли ми говоримо про відшукованість, мені б хотілося зараз зупинитись на декілька позицій, які найбільш важливими для взагалі розуміння оцих ФРД.

І перша з цих позицій – це перша літера «Ф» – відшукованість даних. Видимість даних – основа основ, тому що якщо дані невидимі, то їх неможливо знайти. Мова не йде про відкриття даних, мова йде про опис даних, метадані, які описують, хто робив дослідження, в який період, в якому регіоні, в який період часу, чого стосуються ці дані і так далі.

Тобто це опис самого набору цих даних, цих результатів, або опис наукової статті, яка публікується. І оці метадані, вони дозволяють пошук за різними позиціями – за автором, за організацією, тематичний пошук і так далі. І друга позиція – це ідентифікатор даних.

Не друга, але так само перша. Це унікальний код, який однозначно визначає дані. Найбільш відомий нам усім – це DOI Digital Object Identifier, який дозволяє нам досить швидко, надійно знаходити дані і результати дослідження їх цитувати.

Навіть якщо змінюється їх фізичне розміщення, навіть якщо сайту більше немає, то за DOI можна знайти ось ці результати. Також метадані. Ми трошки далі будемо про них говорити, тому зараз не зупиняюсь детально.

Метаданами можна описувати фізичні об'єкти і цифрові об'єкти. Збереження метаданих передусім в репозиторіях. Репозиторії вони можуть зберігати тільки як метадан, так і самі набори даних або результати досліджень текста.

Існують різні типи репозиторіїв. Основні з них тут задані – це загальні репозиторії, галузеві, тематичні. Інституційні репозиторії в Україні, здається, всі університети наразі мають свої інституційні репозиторії.

Ну і є репозиторії текстів і репозиторії даних. Мінімальний набір метаданих. Якогось 100% всіма схваленого і узгодженого стандарту цього набору опису даних його не існує.

Можуть бути різні підходи, можуть бути різні стандарти. Наукові організації можуть задавати свої параметри для дослідників, що має входити в мінімальний набір даних. Це те, що тут представлено, лише невеличка рекомендація того, як це може бути.

І найважливіше, чому це потрібно описувати? Звісно, будь-який набір метаданих повинен включати назву для того, щоб правильно знайти і процитувати дані. Якщо їх не назвати, то їх неможливо буде знайти.

Також треба сказати, хто є автором або відповідальною особою, або, можливо, власником цих даних. Для чого це потрібно, щоб можна було віднести, кому це належить, процитувати відповідно, а якщо якісь обмеження доступу, то отримати контакти для того, щоб зробити запит на ці дані. Звісно, короткий опис даних, щоб було зрозуміло, про що це воно.

Навіть не відкриваючи ці дані, можна зрозуміти, чи вони потрібні, чи вони не потрібні досліднику, який проводить цей пошук, чи потрібно їх використовувати, цитувати і так далі. Ключові слова. Ключові слова – це одна з найважливіших позицій для проведення пошуку.

Якщо немає ключових слів, то фактично це майже онеможливує їх пошук. Можна, звісно, знайти за організацією, за автором, але менше з тим ключові слова – найважливіша позиція для пошуку. Далі період створення або дата створення, або період, коли збиралися ці дані – це створює контекст для того, щоб розуміти, що це за дані, і дозволяє встановити актуальність, необхідність цих даних.

Також версія, тому що дані можуть бути різних версій, на різних стадіях дослідження, на інших етапах. Знову таки, це задає певний контекст. Формат даних і тип даних – це дозволяє іншим використовувати ці дані і обробляти ці дані, або в машинному форматі обробляти.

Звісно, якщо це не описано, то постають питання, чи можливо використовувати, чи ні. Методи збору та обробки даних – це дуже важливо для забезпечення відтворюваності дослідження і перевірки надійності цих даних. Дуже часто включають в обов'язковий мінімальний набір даних – це список змінних, для того, щоб можна було правильно працювати дані і давати їх інтерпретації.

Звісно, ліцензія. Ліцензія показує, що дозволено робити з цими даними, на яких правилах. Ну, і певним чином захищає дослідника від порушення авторських прав.

Звісно, як я вже казала, немає організації, які моніторять і проводять нагляд, але менше з тим, це є підставою потім звернення до суду, якщо встановлена ця ліцензія. Умови доступу. Умови доступу – це хто контролює дані, чи можна їх бачити, завантажувати.

І особливо важливо це для чутливих даних і дотримання законодавства про захист персональних даних і питань безпеки. Далі постійний ідентифікатор. Цей цифровий ідентифікатор, він гарантує те, що використані дані будуть правильно процитовані і буде правильно оформлено посилання.

Місце збереження, де зберігаються ці дані. І ще раз, кожна позиція, кожне поле в описі метаданих фактично несе з собою певне змістовне навантаження. І в залежності від того, що хоче дослідник і які вимоги визначає наукова організація, в рамках якої проводяться дослідження, цей набір відповідно буде уточнюватись.

Ну і ще й репозиторії, де буде розміщуватись, так само там можуть допомогти. Про план управління даними ми будемо говорити трошки пізніше. Доступ це accessibility.

Доступ до даних це те, що буде визначено в плані управління даних. І регламентує, хто може мати доступ, на яких умовах і коли може. Всі умови доступу фактично, загальні умови доступу регламентуються на підставі ліцензії.

Найбільш відомий сервіс надання ліцензії – це Creative Commons. Creative Commons – це вининпробудкова організація, яка розробила сімейство ліцензій на безкоштовній основі. І на сайті Creative Commons є інструмент, який дозволяє підібрати рецензію.

Він пропонує декілька запитань, на підставі якої можна визначити, яку саме ліцензію отримає дослідник. І варто використовувати досліднику за його умов, хоче він дозволити використовувати комерційно чи ні. Це буде відкритий доступ чи будуть якісь обмеження.

Це буде простий доступ чи буде доступ за запитом, наприклад. Відповідно до цього буде відповідна ліцензія і на Creative Commons можна її підібрати. Наприклад, найбільш розповсюджена ліцензія, яка дозволяє використовувати роботу за умови належного цитування – це CC BY 4.0. Її можна знайти на Creative Commons. Інтероперабельність – це готовність даних до обробки і поєднання з пакетами даних інтеграції до інших досліджень. Основа цієї інтероперабельності – це використання машиночитаних та відкритих форматів даних, завдяки чому різні програми можуть ефективно взаємодіяти з даними. Також принцип операбельності потребує спільного розуміння термінів і понять різними користувачами, для чого формуються словники, список термінів найпростіше, або гласарі, або таксономія у нас найбільших.

Також в рамках інтероперабельності дуже важлива позиція – це зв'язані дані. Тому що, наприклад, набір даних або публікація може бути пов'язана з іншими дослідниками і авторами, може виконуватися в певній організації, таким чином ці результати прив'язані до конкретної організації. Або результати дослідження або набір даних були отримані в рамках якогось грантового проєкту, або були застосовані якісь програмні інструменти.

Тобто, фактично, з цими даними, які оприлюднюються або які описуються, є зв'язки з іншими позиціями, з іншими установами, організаціями, з іншими параметрами, які також мають бути прописані. Чому це важливо? Оці зв'язані дані задають, по-перше, контекст, по-друге, забезпечують навіть знаходжуваність даних, тому що можна знайти за проєктом, які були результати отримані.

От в рамках грантового проєкту можна шукати так, в рамках наукової організації можна проводити такий аналіз і такий пошук. І оці зв'язки також дозволяють взаємодії різних сервісів і робочих

процесів, коли мова йде про роботу з даними. Повторне використання даних, подальшого дослідження.

Для того, щоб їх можна було використовувати, звісно, має бути ліцензія, яка говорить, на яких правах їх можна використовувати. Далі потрібно розуміння того, звідки взяли дані, походження даних і які були методи збору та обробки даних. І стандартизація використання загальнопритітих форматів, термінології, онтології.

Інформація про походження даних фактично відображає історію, методологію створення даних, а також їх контекст. Чому були зібрані дані, методи, методологія, де, коли, хто збирав, з якими даними, можливо, попередніми чи іншими дослідженнями це було пов'язано. Хто є автором, яке програмне забезпечення.

Фактично, ці позиції забезпечують можливість і контекст для того, щоб інтегрувати ці дані або використовувати, або перевіряти їх в рамках якихось інших досліджень. Це коротко. Я сподіваюся, говорила не дуже швидко.

І ви встигли уловити основні позиції. Якщо є запитання, я їх відповім.

[Мовець 2]

Пані Людмилі. Можна? Так, дякую.

Ви багато говорили і багато питань. Одне з таких питань щодо DWI. Зараз у нас інститут, як і багато інших, заповнює цей метадата менеджер.

І там пише, що от в кінці 25-го року він припине своє існування, а що далі? І я не можу зрозуміти, що далі з тим буде. Бо він досить зручний функціонал, можливо, ви знаєте.

У нас просто в інституті багато видальної, і зараз є оце вимога, щоб кожна стаття молодої. І, власне, з тим питанням, куди далі, як далі. Можливо, у вас є якась інсайдерська інформація.

Власне, дякую вам дуже.

[Мовець 1]

Дякую, на жаль, у мене немає інсайдерської інформації, я не знаю. Але я думаю, що у вас буде практична частина, де вам будуть розповідати про можливості Dataverse, і яким чином функціонують там репозиторії, і ви зможете отримати відповіді на питання. Але якщо у вас є DWI, то ці дані мають бути збережені.

Вони нікуди не зникнуть.

[Мовець 2]

Я маю на увазі системи. Ця система має на увазі цей кросреф метадата менеджер. Пише, що він до 2025 року, а далі, що він закінчує, що його не буде цього.

І питання, де далі, бо в кінець року і в кінці року зазвичай вже всі статті виходять, які треба... Тобто ми викладаємо на сайт оперативно і реєструємо DWI. Питання, а що насправді далі виходить.

Я не знаю, чи може якась інформація, яка надалі буде у цей менеджер, бо досить зручний функціонал, і там легко в ньому працювати.

[Мовець 1]

От Сергій Якович зараз вам прямо відповідь. Він знає відповідь на ваше запитання.

[Мовець 3]

Доброго дня. Це питання не до нас, а це питання до тих, хто, будемо казати, надає вам доступ. Це екедемперіодик.

Ми використовуємо іншу систему. І отримуємо DWI тільки для наборів блянків. Але DWI залишиться, менеджер зміниться, а доступ кросрефт вам надасть.

Тобто це питання, яке ви задаєте безпосередньо, будемо тим, хто супроводжує цю систему в Україні.

[Мовець 1]

Так. Значить, мені важливо, для мене важливо дуже, щоб у вас склалось загальне розуміння про те, що таке відкрита наука і навіщо воно потрібно. Для того, щоб далі всі інші позиції вписувалися в контекст цієї відкритої науки.

Чому так робиться, чому так будуються процеси, навіщо це потрібно. Сподіваюся, що це було зрозуміло. Якщо так, ми підемо тоді далі.

Про те, яким чином описувати дані, як це робиться, у вас буде практичне заняття. Я думаю, що вам буде надзвичайно цікаво, і ви з цим дуже добре розберетеся після практики.