

Власне, те, про що я говорила на початку минулого блоку. Оці феєрдані, вони, налаштування всіх позицій, як вони будуть знаходитись, хто буде присвоювати ідентифікатори, який буде доступ його умови, чи можна їх буде використовувати в інших дослідженнях, чи треба буде запитуватись, можливо, яка термінологія буде використовуватись в описах, які словники і так далі. Це все треба десь спланувати, всі ці позиції треба продумати.

І основою для феєрданих, звісно, виступає процес управління даними досліджень. Слово управління – це, власне, система заходів якихось. Управління передбачає ті заходи, як поводитись з науковими даними протягом усього їх життєвого циклу.

Що є метою цього управління даними дослідженнями? Потрібно, щоб дані були безпечними, надійними, стійкими, доступними для пошуку, зрозумілими, щоб їх можна було використовувати. Тобто, дані мають бути на всіх етапах їх життя такими, щоб вони відповідали принципам феєрда. Що це за життєвий цикл даних? Як вони живуть? Ну, тут зображено у вигляді таблички, але, по суті, це коло, тому що воно ходить по колу, причому на кожній позиції може відбуватися повернення і уточнення до якихось інших етапів, попередніх етапів.

Звісно, дані мають бути заплановані, тобто, які дані потрібні, як їх треба зібрати, якими методами, як, хто їх буде збирати, як їх описувати, на яких умовах їми будуть ділитися, як буде налаштовано доступ. Тобто, це планування закладається в так званому дата менеджмент план, тобто, план управління даними. Це документ такий, про який ми будемо говорити надалі, він тому і виділений тут коліром, тому що план як документ, він складається і визначає всі ці ключові умови, істотні умови.

Наступний етап життя даних – це отримання даних шляхом спостережень, експериментів, опитувань чи там моделювання. Потім відбувається обробка цих даних, структурування, систематизація, переведення даних, наприклад, до машини читаний формат, очищення цих даних. Наступний етап – аналіз, коли вже очищені, підготовлені, систематизовані, структуровані дані за допомогою необхідних методів чи інструментів аналізуються і відбувається інтерпретація цих даних.

Далі треба описати отримані результати, треба задокументувати, що отримали, описати методи застосовані, описати про походження для того, щоб ці дані можна було зрозуміти, можна було повторити той експеримент, можна було застосувати ті методи аналізу, які були застосовані, і підтвердити, ми отримуємо те саме, чи ми не отримуємо те саме, і документування всього контексту того, що відбувалося з цими даними. Коли є опис, треба зберігати дані, організація даних у відповідних системах, наприклад, у репозиторії, де оці наукові результати отримані і набори даних можуть зберігатися. Крім того, наступний етап – це поширення даних.

Потрібно якимось чином, щоб про них дізнались інші, бо якщо про них не дізнались, то цього немає ніякого. Тому треба надати доступ для інших дослідників і для суспільства для того, щоб ці результати могли стати основою для наступних досліджень, для міждисциплінарних досліджень, для інновацій, для використання суспільством, для наукових закладів, для навчання і так далі. Тобто, поширення і публікація.

Далі – питання повторного використання. Власне, це те, що я, власне, проговорила щойно. І архівування.

Дані повинні зберігатися. Зберігатися тривалий час зі збереженням їх цілісності. Власне, це є життєвим циклом даних і на кожному етапі, як я вже говорила, вони можуть повертатись на попередній етап, можуть носитись якісь уточнення, якісь певні коригування.

Наприклад, на рівні збереження, можливо, якщо були виявлені помилки в даних, неправильні результати, вони можуть бути видалені. Залишиться мета даних, лише опис, що таке було, але з

якоїсь причини воно було видалено. Але історично це було, це робилося і якщо є потреба, можна буде тоді в'яснити, звертатись і так далі.

Значить, іще раз, управління даними досліджень, вони відбуваються у процесі ось цього життєвого циклу. Треба відживати заходи, щоб усі ці етапи були пройдені. Звідки взяли взагалі уявлення про потреби того, що це потрібно робити? Це пов'язано з тими дискусіями, які виникли в першу чергу в середовищі бібліотек.

І це були Сполучені Штати кінець 2000-х років, середина 2000-х. І виникли рекомендації з боку, як результат діалогу наукових бібліотек і дослідницьких центрів, фінкентрів, які з'єдналися з проблемою, що їм треба якось зберігати великі обсяги наукових даних. І ось це вивелося в те, що в 2011 році Національний науковий фонд штатів розробив дата менеджмент план, тобто отой документ, в якому прописані всі життєві стадії, він зробив обов'язкову частину всіх заявок на фінансування.

І це вважається такою офіційною відправною точкою в практиці застосування планів управління даними, які відображають отой життєвий цикл, про який ми говорили, відправною точкою для всього світу. Ну і далі ця практика почала розповсюджуватися, обов'язково була введена в практику Європейського Союзу, обов'язково під себе були розроблені відповідні цифрові інструменти, які дозволяють втілити життя в цифровому плані. І це дійшло до України в кінці кінців, про що ми в самому кінці також поговоримо.

План управління даними. План управління даними, він містить весь опис того, що буде з даними протягом життєвого циклу. Тобто, які дані будуть зібрані чи створені, як вони будуть організовані, як треба називати файли, яка буде система папок, хто буде мати доступ до цих даних, хто може тільки переглядати, а хто може вносити зміни, де ці файли будуть зберігатися, або ці об'єкти, який буде доступ, чи можна повторно використовувати в інших дослідженнях, де будуть в довготривалому аспекті зберігатись.

Ну, явно ж не на Google-диску. Тобто, де буде збереження довготривалих цих даних, яким чином будуть дотримані всі принципи FAIR, про які ми поговорили, як захистити дані, як дотриматись принципів етики, особливо в даних, які мають чутливість, там де є персональні дані, наприклад, в медичних дослідженнях, або авторські права, як вони будуть захищені. Тобто, в плані прописується всі позиції абсолютно для того, щоб можна було налагодити ці параметри поводження з даними.

Призначення планів управління даними – це, знову-таки, повторю ще раз, втілення в життя принципів FAIR, про які ми проговорили, тому не буду зупинятись. Але є ще вимоги, вимоги з боку грантодавців і журналів. Тому що надалі ці дані мають бути якимось чином доступними, вони мають бути збереженими.

Це дуже важливо для бібліотек, які збирають також дані, забезпечують доступ до них. Зараз найчастіше на базі бібліотек сформовані інституційні репозиторії, які дозволяють працювати з даними, забезпечувати доступ і до метаданих, і до самих наукових результатів. Плюс план управління даними і продумування того, як з ними поводитися, сприяє тому, що зменшується дублювання знань, зменшується дублювання отих зусиль, а відповідно витрат ресурсів праці і затрат грошей на всю цю діяльність.

Якщо підсумовувати, то ось пан Андрій Швеченко, мені дуже сподобалось в його матеріалах, він запропонував таку ознаку, такий аспект, що для чого потрібен план управління даними. План управління даними це така інструкція, як перетворити файли на комп'ютері на науковий результат, який побачить всі, зрозуміють, який буде цитуватись, який ще буде використовуватись, можна додати. Оце план управління даними.

Хто складає цей план? Взагалі залежить від контексту. Якщо це якийсь дослідник, який подав грантову заявку, і він відповідно реалізує якийсь проект, то це буде робити цей дослідник. Якщо це проектна команда, то керівник проекту, він разом з командою, ну має бути так, що разом з командою, збере і напише, погодить усі позиції, які у них буде план управління даними на всіх життєвих етапах даних.

Або це може бути наукова організація, мало того, це повинна бути організація, або інститут дослідницький, або університет. Це можуть бути вони, тому що на рівні організації теж мають бути встановлені внутрішні стандарти, як поводитися з дослідженнями, які організовуються, які здійснюються на базі цієї організації. Якщо ми говоримо на рівні проектної діяльності, то звісно відповідальність несе керівник проекту.

Але до того, щоб скласти цей план управління даними, треба ще долучити якихось інших фахівців. Наприклад, є спеціальні стюарди даних. Вони забезпечують якість опису, якість метаданих.

Бо метаданих – це ключ для видимості і для знаходжуваності. І треба правильно організувати ці метадані і правильно їх описати. Ось задача стюардів даних.

Ключова задача саме в цьому. Є спеціалісти в репозиторіях, куратори репозиторії. Є IT-відділ, є юристи з інтелектуальної власності персональних даних, які можуть передбачити питання захисту або питання етичні поводження з даними.

Тобто, відповідальний то один, але до роботи залучаються всі інші, якщо це команда, і також інші фахівці, які мають бути, з якими треба проконсультуватись, або які повинні надати відповідні послуги для того, щоб скласти якісний план управління даними. Звісно, це можна зробити на папері, але в нашу цифрову епоху все робиться за допомогою цифрових інструментів. Є інструменти, де по кожному розділу вам будуть поставлені запитання, і ви поступово сформуєте цей план в електронному вигляді.

Його можна опублікувати, зробити доступним для всіх. Два найбільш відомих в Україні і розповсюджених інструменти, вони усі виділені. Один називається ARGOS.

ARGOS – це європейський інструмент, він орієнтований на європейські гранти, і в рамках європейських грантів рекомендують саме зайти на сторіночку ARGOS і там зробити цей план. Для Сполучених Штатів це інструмент DMP Tool. Він значно більш деталізований і, як правило, на Американському континенті, і в усій Америці, і Північної, Південної Америки, найбільше використовується DMP Tool.

Але це не означає, що не можна у нас його використати. Оцей ARGOS, де його можна знайти? Можна зайти на сторіночку OpenAIR, це європейський сайт, знайти сервіси OpenAIR і там знайти для управління даними оцей сервіс ARGOS. До речі, тут всі сервіси, які обслуговують хмару європейської науки, наприклад, ми сьогодні згадували Zenodo, там написано Outreach, і стоїть Zenodo, або Discover, також Zenodo, або Publish, Zenodo, це те, що забезпечує ця платформа.

Але нас зараз цікавить, як створити план управління даними. Для управління даними ось інструмент ARGOS. Це саме зроблений скрін-шот із сторіночки ARGOS.

Там треба зареєструватися, зайти на нього, і ось буде кнопочка на Start New Plan, можна натиснути на неї, і відкриється ось така сторінка, де можна зайти в кожний розділ і описати про збір даних, про документування метаданні, про етику і дотримання умов доступу і нормативних вимог, про те, як будуть зберігатися дані, і так далі. Тобто кожен розділ описується, далі він фіналізується, ну і цей план

можна або залишити закритим, або можна опублікувати, він може бути навіть у відкритому доступі. Складові плану управління даними.

Значить, якоїсь однієї жорсткої структури, жорстко заданої для планів управління даними немає. Кожна організація встановлює свої правила і свою структуру того, що має до нього входити. Або це може бути проєкт, і на рівні проєкту так само відповідальний повинен організувати те, яким буде план, які позиції в нього увійдуть.

Можна скористатись готовим шаблоном, наприклад, Vargos, можна розробити своє і додати необхідні позиції. Приблизно, що може входити? Це загальні відомості про оцю роботу дослідницьку. Яка тема, або яка назва проєкту, хто є учасником, хто координатор проєкту, хто керівник, виконавці, можливо, які робочі пакети, автори, які виконують роботу, анотація цього проєкту, за рахунок чого фінансується, строки, які в виконанні цього проєкту.

Далі, блок отримання даних і яким чином вони будуть зберігатись. Які методи будуть отримання даних, типи даних, формати файлів, назви файлів, обсяг даних, який має бути, яка доступність, хто має доступ, хто не має доступ, це все сюди. Далі, етика отримання та використання даних.

Чи потрібні етичний дозвіл, чи є там якісь чутливі дані, особливо персональні дані. Можливо, треба анонімізувати дані, якщо користуються, залучаються люди до дослідження. Обов'язкове оформлення дозволу на використання персональних даних.

Рішення, можливо, комітету з етики зараз вимагаються в дослідженнях, навіть соціологічних опитуваннях, дослідженнях думки, інтерв'ю, не хочу вже про медичні дослідження. Треба висновок комісії з етики, яка має бути створена в закладі для того, щоб етично використовувати ці дані. Чи потрібна інформована згода, можливо, як її отримати, чи це буде онлайн, чи це поставити галочку в формі опитування, чи це треба якийсь спеціальний документ готувати.

Знову-таки, якщо дослідження пов'язане з людьми, ви будете результати досліджень їм повідомляти? Результати, які будуть отримані, чи не будете? Яким чином будуть етика поведінки з даним, які будуть параметри доступу до наукових результатів? Як вони будуть поширюватись, зберігатись? Тобто, весь блок, всі параметри, в залежності від специфіки вашого дослідження, визначаються в цьому плані. Питання власності, питання того, хто володітиме цими даними, на яких умовах вони будуть використовуватися. Чи потрібно буде запитувати дозвіл тих, хто хоче скористатись, з ким треба буде зв'язатися, до кого треба написати, щоб це вияснити.

Чи потрібні якісь періоди баргу, чи комерційне використання дозволяється чи ні, і так далі. Питання володіння, усі, можливо, консультації з юристами, які потрібно отримати, це все описується в блоці, пов'язаній з адаптальною власністю. Далі, наукові результати треба зберігати і забезпечувати її надійне зберігання, довгострокове, і без спеку.

Де будуть зберігатися дані, які строки, як вони будуть називатися, як вони будуть розпізнаватись, бо треба розуміти, що це за дані, і правила найменувань, файлів і папок, вони дуже важливими. Як буде захищено, як буде проведено резервне копіювання, де воно буде зберігатися, і так далі. Ну, і поширення того обмін даними, це присвоєння ідентифікатора, дої, наприклад, або якогось іншого, це визначення того, в якому репозиторі вони будуть зберігатися, з якою ліцензією, який буде доступ під час поширення цих даних, обсяг доступних даних, ну, період ембарго, тобто, ось така позиція.

Бачите, поширення даних і обмін даними можуть бути частиною якогось іншого блоку. Тобто, оці блоки можна перебудовувати на власний розгляд, виходячи вже з логіки відповідального за цей план і тієї специфіки проєкту чи дослідницького діяльності, яка проводиться. В Україні є Міністерство освіти і науки.

Наше Міністерство освіти і науки видало документ про запровадження, обов'язкове запровадження управління дослідницькими даними в закладах вищої освіти і наукових установах. І в цих рекомендаціях, це рекомендований документ, прописано, який може бути план управління даними. Ось на цьому слайді, я думаю, що ви отримаєте ці презентації, ви зможете подивитись основні позиції, які можуть входити в цей план управління даними.

Вони пропонують загальні відомості про дослідження. Потім, другий блок, це оспливості збирання або генерації даних і забезпечення якості даних. Потім, третій блок, це мета-дані, тобто як він буде описуватись.

Четвертий блок, це етичні і правові питання. П'ятий блок, це зберігання та безпека даних. Шостий блок, це довгострокове зберігання і обмін даними.

Сьоме, це ролі учасників цього проєкту, чи то дослідження, яку відповідальність вони будуть нести протягом всього життєвого циклу, хто за що відповідає, хто відповідає за документацію даних, хто відповідає за правові питання, хто відповідає за довгострокове зберігання і так далі. Ну і ще один блок, це витрати. Які ресурси потрібні і які джерела фінансування дослідження.

Вимоги Міністерства освіти і науки. Ну, вони рекомендаційні поки що. Міністерство освіти і науки вимагає, щоб дослідницькі дані закладу вищої освіти і наукової установи зберігали в електронній формі.

Управління даними на всіх життєвих етапах повинно відповідати принципам ФЕР. Обов'язково МОН рекомендує, щоб наукові установи організували на початку досліджень, щоб дослідники складали плани управління даними. Крім того, доступ до наукових даних має бути організований через репозиторії.

І це забезпечить і поширення цих даних, і можливість підтвердження наукових результатів, і далі показувати їх у звітності цього закладу. Мова йде про всі можливі наукові результати у різних форматах, починаючи від стаття, звіт, доповіді і так далі. Обмеження на доступ даних може бути встановлено відповідно до законодавства України, наприклад, про захист персональних даних, або спеціальними угодами щодо поширення використання цих даних, можливо, умовами грантових проєктів, які виконуються, які можуть також накладати якісь обмеження.

Далі МОН рекомендує визначити структурний підрозділ, який буде займатись управлінням науковими дослідницькими даними. І далі вони прописують, у мене написана деталізація, там детально прописано, що і хто це може робити. Ну і на кого поширюватись повинно оце управління дослідницькими даними, складання планів управління? Воно має поширюватись на всіх абсолютно осіб, які задіяні в наукових дослідженнях, включаючи, якщо це заклад вищої освіти, включаючи студентів.

А хто має відповідати також? МОН деталізує, рекомендує, що от варто призначити таких-то осіб. І деталізація того, що варто робити на кожному етапі життєвого циклу з даними, також МОН деталізує. Я думаю, що ви зможете відкрити ці рекомендації з управління дослідницькими даними нашого міністерства і ознайомитися з ними більш детально.

Остання позиція, про яку я хочу поговорити, це було визначено паном Андрієм Шевченком, який є консультантом Київському академічному університеті. І вони дуже прості, дуже логічні, дуже такі покрокові позиції, що потрібно робити для управління даними. План управління даними, я вже це сьогодні згадувала, це інструкція.

Як перетворити файли на комп'ютері на науковий результат, який бачать, розуміють, ситують. От я би ще добавила, і використовують. Ну, використання може мало залежати від дослідників, але те, що його побачать, всього зрозуміють і його будуть ситувати, це абсолютно може бути під дуже значним впливом або колективу наукового, або окремого дослідника-автора.

Наукова установа, вона повинна визначити роль і відповідальність в тому, як організовано план управління даними. І крім того, науковий заклад, наукова установа повинна забезпечити, встановити мінімальні вимоги до публікації, які мають забезпечити їхню відомість, наприклад, визначитися з репозиторією або вказати, в які строки мають бути розміщені, це щонайменше. Що буде, якщо цього не зробити? Дослідницькі дані без опису і без спланованого плану того, як з ними поводитись, можуть потрапити у видимий світ випадково.

Вони не точно, але не буде організовано того, щоб вони потрапили у видимий світ і щоб їх побачили, зрозуміли і процитували. Як розподілити відповідальність? Керівник, керівник чи то установи, чи то проєкту, він є власником історії і відповідає за те, щоб дані доброчесно використовувалися з належним посиланням. І для цього, щоб це було, треба налагодити режим доступу, як виходячи зі специфіки цього дослідження і усіх умов, про які ми говорили, будуть доступні результати досліджень і набори даних.

Відповідно до цього буде визначено ліцензію. CICI-0 – це найпростіша ліцензія відкритого доступу фактично. CICI-5 – це трошки більш високий рівень, який потребує належного цитування того, який ви запропонуєте.

Ще одна особа – Дата Стюарт. Дата Стюарт, він, скажімо так, редактор видимості, тому що він стежить за тим, щоб дані, оці наукові результати, набори даних, були правильно описані, щоб було достатньо інформації, щоб вона була зрозумілою, щоб були визначені отой мінімальний набір в описі метаданих, про який ми раніше проговорювали в попередній частині. А ІТ-команда, вона дбає про довіру і про безпеку даних, і про їх збереження.

Куратор репозиторії, він перевіряє комплектність того всього, що має бути розміщено в репозиторії. Версію, наявність ліцензії, все, що необхідно. І коли це все є, він проводить дії для того, щоб присвоїти ідентифікатор.

Найбільш нам відомий – це цифрові дої. Далі ще можливо залучати юриста, який може проаналізувати персональні і чутливі дані, права третіх осіб для того, щоб налагодити доступ і дотриматись усіх принцип етики і права інтелектуальної власності при поводженні з цими даними, при їх розповсюдженні. Ось це той набір суб'єктів, який має бути задіяний.

Звісно, юрист за необхідністю.

Порядок роботи. Хто що робить? Власне, це рекомендована позиція, звісно. Автор дослідження повинен підготувати опис даним з тим мінімумом, який забезпечує видимість і довір.

Це мета-дані. Це назва дослідження чи назва проєкту і короткий опис для нас українською і англійською мовою. В тому числі, що це за дані, як вони збиралися, які є зв'язані дані з організацією зв'язок, з проєктом фінансуючим, з якоюсь командою, з попередніми дослідженнями, з іншими наборами даних.

Ключові слова. Ключові слова потрібні для того, щоб був можливий пошук. Пошук не тільки за прізвищем автора, але й за тими ключовими словами.

Варто користуватися словниками для того, щоб пошук вже не тільки людський, а це ще й машинний пошук. І щоб він відбувався ефективно, була знаходжуваність, бо це повинен підбати автор. Також

потрібно вказати ідентифікатори для кожного автора ORCID і організатор, переписую, ідентифікатор для самої установи.

Також визначитися з ліцензією Creative Commons Chooser, як один з варіантів, який допоможе цю ліцензію підібрати. І вона буде регламентувати, показувати, який є режим доступу до даних і визначитися взагалі з цим режимом. Далі описати текстовий файл про це дослідження.

І далі наступний етап Data Steward, який проводить перевірку даних того, що вони відповідали принципам FAA, користуються онлайн інструментами, які дозволяють це перевірити. І після того, як це буде все зроблено, це може бути надіслано до репозиторію для того, щоб присвоїти DOI. На стадії підготовки даних, перш ніж переписуювати DOI, треба проконсультуватися з юристом, якщо потрібно, щоб все, що стосується чутливих даних і питання інтелектуальної власності було враховано.

Ну, а тоді вже можна із куратором в репозиторію, щоб він зробив присвоїв DOI і ці дані були розміщені в репозиторії, стали видимими і доступними для пошуку. І ще раз хочу сказати, що може бути розміщений опис даних, мета дані, це окрема позиція, і самі результати наукові, це окрема позиція. І якщо результати досліджень, наукові дані, вони можуть бути закритими і може бути вкритий доступ, то мета дані доступні завжди.

Дякую за увагу. Це все, що хотілося вам сказати і розповісти з приводу планів управління даними. Я знаю точно, що в СумДУ у них є план управління даними, він обов'язковий до виконання, це не для галочки.

Вони дійсно використовують. У них розроблений, він працює, і всі дослідники повинні слідувати тим принципам, які там визначені. Ваші питання і ваші рефлексії, ваші висновки, що вам запам'яталось найбільше, можливо, проговоріть.

Або що ви собі плануєте зробити? Я думаю, що тут всіх, в принципі, треба просто брати і робити ці плани. Тому залишимо. Бо в нас вже просто перерва.

Я пропоную, щоб люди ще встигли каву попийти. О 12-й Сергій Якович виснов, тоді вже почне наступного лета. А дозвольте запитання.

Це таке напевно болюче запитання для всіх науковців, власне, в Україні. От послуги дата стюарта, вони, так би мовити, безкоштовні для науковців України? Чи це вилітєся в копійку, скажімо так? Це цікаво все. Це робота, це функція, яка має виконуватися.

Найчастіше це працівник, якщо це інституційний репозиторій, то це працівник бібліотеки може бути. Тобто це особа, функції якої входить оце завдання, до посадових обов'язків, в яких входить це завдання. Це може бути призначено організаційно або окрема людина.

Але, скоріш за все, це хтось із фахівців, які вже працюють, для кого це доручено виконувати цю роботу і включено в його посадові інструкції. Але будь-яка робота має оплачуватись. І якщо ця робота, на неї як мінімум час має бути виділений.

Час і кошти відповідні, так. Дякую, дякую. Там є питання стосовно плану управління даними, обов'язкове тільки для установ МОН чи для установ НААД.

Дивіться, це були рекомендації Міністерства освіти і науки. Звучить воно словом рекомендація. Установи Національної академії наук України, вони підпорядковані Національною академією наук.

Немає даних, щоб, звісно, є стратегічні документи, є стратегія, яку ухвалила НАНУ України. І відповідні сфери даними для того, щоб це справдилося і діяло. Щоб ця стратегія втілилася в життя, то треба робити цей план.

Але обов'язку такого немає. Але якщо цього не зробити, якщо не використовувати інструменту цей план управління даними, то воно далі не буде рухатись. Тому що будь-яку діяльність треба організувати, треба її налагодити, треба вписати в процес.

А якщо цього не зробити, то все, що стосується феєр даних, відкритостей і перетворення мертвого грузу на капітал, у нас не відбувається. Тобто формального обов'язку немає. А неформальне є. Ще питання.

Пане Андрій Терехов. У мене таке просте питання. Чи плануються якісь курси, наприклад, для бібліотекарів? Для того, щоб вони розуміли, як робити все те, що Ви сказали.

Оце до рекомендації це дуже добре. Чи будуть якісь курси, чи ще щось? Щоб просто готувати фахівців у всій галузі. Так, я думаю це запит до Володимира Ночвая.