

**Національний університет  
«Києво-Могилянська академія»  
Докторська школа імені родини Юхименків**



**Управління даними досліджень:  
*навчальна програма курсу***

**Research Data Management:  
*Syllabus***

**Управління даними досліджень: навчальна програма курсу для студентів PhD програми НаУКМА /Упорядники Борисова Т.О., Ярошенко Т.О. – К.: НаУКМА, 2017. – 13 с.**

Схвалено на засіданні виконкому Докторської школи імені Родини Юхименків НаУКМА, \_\_ жовтня 2017 р.

Цей твір ліцензовано за [ліцензією Creative Commons Із Зазначенням Авторства 4.0 Міжнародна](#).

## ІНФОРМАЦІЯ ПРО КУРС

Цей курс із управління даними досліджень розроблений у 2017 р. для студентів PhD програм НаУКМА із усіх спеціальностей. Має на меті ознайомлення та вивчення усіх складових укладання плану із управління даними досліджень, його практичного застосування та впровадження у загальний дослідницький процес умінь із планування, документування, зберігання та поширення даних.

**3 кредити ECTS** / Заняття в аудиторії: 14 год (7 занять по 2 академічні години).  
Самостійна робота слухачів курсу: 46 год

Необхідні умови: студенти повинні мати чітке уявлення про тему свого дослідницького проекту

## COURSE CONTENT

The purpose of the course is to enable PhD students to properly manage research data as a means to foster efficiency, ensure proper stewardship of data, and exceed funder requirements for planning, preservation and sharing. Major course content includes the following topics:

- Overview of research data management, definitions and best practices
- Types, formats and stages of research data
- Metadata (data documentation)
- Data storage, backup and security
- Legal and ethical considerations of research data
- Data sharing and reuse
- Archiving and preservation

**3 ECTS credits**

## Опис курсу

Цей курс із управління даними досліджень розроблений для аспірантів та докторантів НаУКМА із усіх спеціальностей. Має на меті ознайомлення та вивчення усіх складових укладання плану із управління даними досліджень, його практичного застосування та впровадження у загальний дослідницький процес умінь із планування, документування, зберігання та поширення даних.

## Зміст курсу

Курс побудований як комбінація лекційного викладу та активного включення студентів через дискусію та практичні заняття як під час занять в аудиторії, так і в рамках самостійної роботи. Курс складається із семи модулів, кожен із яких передбачає короткий лекційний огляд та обговорення і виконання практичних завдань слухачами курсу. Загальний зміст курсу включає наступні 7 тем:

- Вступ та загальний огляд управління даними досліджень;
- Типи та формати дослідницьких даних;
- Документування та організація даних;
- Зберігання та захист даних;
- Правові та етичні застереження в управлінні даними;
- Поширення даних, обмін та повторне використання;
- Архівування та довгострокове збереження даних.

## Очікувані результати курсу

Після завершення курсу, його слухачі повинні:

- Уміти пояснювати необхідність обміну та поширення даних, розуміти життєвий цикл даних та його значення для загального дослідницького процесу;
- Ідентифікувати потенційних повторних користувачів їхніми даними та формулювати стратегію їх поширення;
- Укладати план управління даними та визначити обов'язки дослідницького персоналу щодо різних його етапів;
- Знати типи, формати та стадії дослідницьких даних та конкретного програмного забезпечення, що потрібне для роботи із ними на різних етапах циклу;
- Уміти документувати дані через метадані для подальшого зберігання та поширення;
- Планувати поширення та заходи безпеки щодо власних даних;
- Знати та вміти пояснити правові та етичні умови поширення даних;
- Планувати довгострокове збереження та можливість повторного використання даних.

## **План оцінювання**

Студенти отримають залік з курсу за умови набору 61 балу за 100 бальною шкалою.

Участь у аудиторних заняттях - 15 балів.

Виконання практичних завдань - 35 балів.

Проміжні оцінювання - 20 балів.

Фінальний тест - 30 балів.

Участь у аудиторних заняттях передбачає активне включення в дискусії із викладачем та іншими слухачами курсу, задавання питань, поширення власних думок, вражень та досвіду. Також оцінюється активна участь у груповій діяльності.

Виконання практичних завдань після кожного модулю передбачає як читання рекомендованої літератури та наступне її обговорення під час аудиторних занять, так і безпосереднє виконання практичних завдань на матеріалах власного дослідження.

Проміжні оцінювання проводиться шляхом тестування, двічі за період курсу за інформацією, засвоєною під час вивчення модулів 1-3 та 4-7.

Фінальний тест (залік, екзамен) відбувається шляхом публічного представлення плану із управління даними для власного дослідницького проекту слухача курсу.

# План занять курсу

## Модуль 1.

### Вступ. Загальний огляд управління даними досліджень.

- 1.1. Вступ до курсу, інструктаж слухачів.
- 1.2. Що таке дані досліджень?
- 1.3. Значення ефективного управління даними для успіху дослідницького проекту.
- 1.4. Цикл управління та зберігання дослідницьких даних.
- 1.5. Контекст (світовий та європейський) для управління та поширення даних.
- 1.6. Визначення вимог до плану управління даними, які використовуються для характеристики та планування життєвого циклу даних.

### Рекомендовані матеріали до Модулю 1.

1. DataONE Best Practices. <http://www.dataone.org/best-practices>
2. H2020 Programme Guidelines on FAIR Data Management in Horizon 2020 - [http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf)
3. Joint Data Archiving Policy. <http://datadryad.org/pages/jdap>
4. Planning for the future: developing and preserving information resources in the Arts and Humanities <http://www.dcc.ac.uk/resources/developing-rdm-services/dmps-arts-and-humanities>
5. Promoting the Stewardship of Research Data, Ensuring the Integrity, Accessibility, and Stewardship o
6. Research Data in the Digital Age (2009): pages 95-99 [http://books.nap.edu/openbook.php?record\\_id=12615&page=95](http://books.nap.edu/openbook.php?record_id=12615&page=95)
7. Simon Hodson and Sarah Jones (2013). Seven Rules of Successful Research Data Management in Universities <http://www.theguardian.com/higher-education-network/blog/2013/jul/16/research-data-management-top-tips>
8. Whitmire, Amanda. 2013. "Data Management Throughout the Research Lifecycle" (August 16). <http://dx.doi.org/10.6084/m9.figshare.774628>
9. Why Share Data? UK Data Archive <https://www.ukdataservice.ac.uk/manage-data/plan/why-share>

### Діяльність під час аудиторного заняття:

Подумайте над проблемами із управління та поширення даними, що є у Вашій галузі. Сформулюйте два ключові питання та обговоріть їх із колегами в своїй групі.

### Практичне завдання до модулю 1:

Укладіть попередній план управління даними, використовуючи як приклад плани [https://dmptool.org/public\\_dmps](https://dmptool.org/public_dmps)

## Модуль 2

### Типи, формати та організація даних.

2.1. Що таке дослідницькі набори даних, які існують типи і формати даних?

2.2. Які із дослідницьких даних краще використовувати для вирішення різних типів дослідницьких питань?

2.3. Загальні формати для зберігання даних.

2.4. Способи запису даних, специфічні для різних дисциплін.

Рекомендовані матеріали до Модулю 2.

1. Ball, Alex (2012). Review of Data Management Lifecycle Models.  
<http://opus.bath.ac.uk/28587/1/redm1rep120110ab10.pdf>

2. Data Types and File Formats: (Oregon State University Libraries):  
<http://guides.library.oregonstate.edu/data-management-types-formats>

3. Defining Research Data: Data Management <https://www.lib.ncsu.edu/data-management/define>

4. File formats and software <https://www.ukdataservice.ac.uk/manage-data/format/file-formats>

5. Library of Congress. Sustainability of Digital Formats.  
[http://www.digitalpreservation.gov/formats/fdd/browse\\_list.shtml](http://www.digitalpreservation.gov/formats/fdd/browse_list.shtml)

6. Supported Formats. Bio-Formats 5.7.1 documentation. OME  
<https://docs.openmicroscopy.org/bio-formats/5.7.1/supported-formats.html>

#### Практичне завдання до модулю 2

Наведіть приклад для кожного із етапів формування даних у Вашому дослідженні (необроблені дані, оброблені дані, проаналізовані дані, опубліковані дані). Вкажіть, які із форматів даних будуть прийнятними для Ваших даних.



## **Модуль 3.**

### **Документування та організація даних**

3.1. Конфігурації назв файлів та папок, логічна організація даних.

3.2. Стандартні операції для управління та документації даними в процесі дослідження.

3.3. Що таке метадані та чому вони важливі?

3.4. Стандарти збору та обробки метаданих у різних дисциплінах.

Рекомендовані матеріали до Модулю 3.

1. File Naming Conventions from the University of Minnesota

<http://researchdata.wisc.edu/manage-your-data/file-naming-and-versioning/>

2. ICPSR. Best Practices in Creating Metadata.

<http://www.icpsr.umich.edu/icpsrweb/content/deposit/guide/chapter3docs.html>

3. Introduction to Metadata: Setting the Stage (Getty Research Institute)

[http://www.getty.edu/research/publications/electronic\\_publications/intrometadata/setting.html](http://www.getty.edu/research/publications/electronic_publications/intrometadata/setting.html)

4. UK Data Service Catalogue metadata [https://www.ukdataservice.ac.uk/manage-](https://www.ukdataservice.ac.uk/manage-data/document/metadata)

[data/document/metadata](https://www.ukdataservice.ac.uk/manage-data/document/metadata)

5. Version Control and Authenticity [http://www.data-archive.ac.uk/create-](http://www.data-archive.ac.uk/create-manage/format/versions)

[manage/format/versions](http://www.data-archive.ac.uk/create-manage/format/versions)

#### **Діяльність під час аудиторного заняття**

Організуйте та правильно конфігуруйте дані у наведеному прикладі (роботу у групах, кожна із яких отримує своє завдання). Обговоріть результати спільно, надайте рекомендації для покращення організації даних для кожної із груп.

#### **Практичне завдання до модулю 3:**

Використовуючи стандарт метаданих, що застосовується у Вашому тематичному репозитарії (<https://www.re3data.org/>) опишіть два різних за типами набори даних у Вашому дослідженні. Якщо Ви ще не маєте власних згенерованих даних, візьміть за приклад існуючий датасет, близький за типом до Ваших даних у дослідженні.

#### **Проміжне оцінювання 1**

Тестування за інформацією, що засвоєна за результатами Модулів 1-3

## **Модуль 4.**

### **Зберігання та захист даних**

4.1. Чому потрібне зберігання, резервне копіювання та захист даних?

4.2. Методи та найкращі практики зберігання, резервного копіювання та захисту даних.

Рекомендовані матеріали до Модулю 4.

1. Backing Up Data (UK Data Archive): <http://www.data-archive.ac.uk/create-manage/storage/back-up>
2. How to start and keep a laboratory notebook, Jennifer Thomson, *Professor, Department of Molecular and Cell Biology, University of Cape Town, South Africa*  
<http://www.iphandbook.org/handbook/ch08/p02/>
3. Inadequate security practices expose key nasa network to cyber attack  
<https://oig.nasa.gov/audits/reports/FY11/IG-11-017.pdf>
4. Store your data (UK Data Archive) <https://www.ukdataservice.ac.uk/manage-data/store>
5. Waard de, A., Cousijn H. 10 aspects of highly effective research data. Good research data management makes data, December 2015 - <https://www.elsevier.com/connect/10-aspects-of-highly-effective-research-data>.

#### **Діяльність під час аудиторного заняття**

Обговорити в групах по одному прикладу втрати даних та можливі способи цього запобігання. Представити усім короткий план із зберігання та захисту даних для запропонованих кейсів.

#### **Практичне завдання до модулю 4**

Підготувати план зберігання та резервного копіювання даних, що буде містити інформацію про: місце для зберігання даних, місце зберігання резервних копій, спосіб синхронізації основного сховища та резервних копій, як часто та хто буде робити резервне копіювання, яким захистом будуть послуговуватися сховища із даними та хто за нього відповідатиме.

## **Модуль 5.**

### **Правові та етичні застереження в управлінні даними**

5.1. Хто володіє правами на дані у дослідженні?

5.2. Які потенційні правові застереження існують у випадку поширення даних для різних галузей знань та в різних країнах?

5.3. Приватність та конфіденційність даних.

5.4. Ліцензії та особливості їх використання.

5.5. Атрибуція та цитування даних.

Рекомендовані матеріали до Модулю 5.

1. Altman, M. & King, G. (2007). A proposed standard for the scholarly citation of quantitative data. D-Lib Magazine, 13(3/4),

<http://www.dlib.org/dlib/march07/altman/03altman.html>

2. Data Citation Initiatives and Issues by Matthew S. Mayernik, 2012 -

[http://www.asis.org/Bulletin/Jun-12/JunJul12\\_Mayernik\\_DataCitation.html](http://www.asis.org/Bulletin/Jun-12/JunJul12_Mayernik_DataCitation.html)

3. How to Cite Datasets and Link to Publications <http://www.dcc.ac.uk/resources/how-guides/cite-datasets>

4. Information Security and Privacy Laws and Regulations, from UW's Chief Information Security Officer, <http://ciso.washington.edu/laws/> (Links to an external site.) [Links to an external site.](#)

5. Ownership and Intellectual Property -

<http://guides.library.uwa.edu.au/c.php?g=325196&p=2177355>

6. Protecting Personal Health Information in Research: Understanding the HIPAA Privacy Rule [https://privacyruleandresearch.nih.gov/pr\\_02.asp](https://privacyruleandresearch.nih.gov/pr_02.asp)

#### **Діяльність під час аудиторного заняття**

Знайти 5 наборів даних із близької тематики та оформити цитування на них. Обговорити питання комерційного використання дослідницьких даних, використовуючи матеріал - <http://guides.library.uwa.edu.au/c.php?g=325196&p=2178559>.

#### **Практичне завдання до модулю 5**

Описати усі правові застереження, що можуть виникнути із даними у Вашому дослідженні. Обрати ліцензії, що підходять для даних Вашого дослідження.

## Модуль 6.

### Поширення даних, обмін та повторне використання

- 6.1. Причини та переваги поширення дослідницьких даних.
- 6.2. Політики видавців, грантодавців та інституцій для поширення даних досліджень.
- 6.3. Відкриті дані та відкрита наука

Рекомендовані матеріали до Модулю 6.

1. Dan L. Longo, M.D., and Jeffrey M. Drazen, M.D. Data Sharing. N Engl J Med 2016; 374:276-277 January 21, 2016. 10.1056/NEJMe1516564
2. DataCite consortium: "Why Cite Data?" <http://datacite.org/whycitedata>
3. Editorial: Alan E. Guttmacher, Elizabeth G. Nabel, and Francis S. Collins. Why data-sharing policies matter. PNAS 2009 106 (40) 16894; published ahead of print September 18, 2009, doi:10.1073/pnas.0910378106
4. H2020 Programme. Guidelines on FAIR Data Management in Horizon 2020 - [http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf)
5. Open data: The researcher perspective - <https://www.elsevier.com/about/open-science/research-data/open-data-report>
6. Open Research Data <http://ec.europa.eu/research/openscience/index.cfm?section=monitor&pg=researchdata#1>
7. Overview of Scientific Data Sharing and Reuse Policies of the Federal Government at The Value of Shared Access and Reuse of Publicly Funded Scientific Data, 2010 [http://sites.nationalacademies.org/PGA/brdi/PGA\\_059258](http://sites.nationalacademies.org/PGA/brdi/PGA_059258)
8. Piwowar HA, Vision TJ. (2013) Data reuse and the open data citation advantage. PeerJ 1:e175 <https://doi.org/10.7717/peerj.175>
9. Борисова Тетяна. Що треба знати, перш ніж поширювати дані? : [презентація до семінару "Цифрова підтримка наукових досліджень", 9 жовтня 2017, НаУКМА, Київ] - <http://ekmair.ukma.edu.ua/handle/123456789/11976>
10. Ярошенко Тетяна. Від відкритого доступу до відкритої науки: виклики часу: [презентація до семінару "Цифрова підтримка наукових досліджень", 9 жовтня 2017, НаУКМА, Київ] - <http://ekmair.ukma.edu.ua/handle/123456789/11980>

#### Діяльність під час аудиторного заняття:

У вигляді дебатів із двома командами обговорити переваги та загрози поширення даних.

#### Практичне завдання до модулю 6

Підготувати короткий огляд (2-3 сторінки) публікацій за останні два роки, що стосуються поширення даних у Вашій галузі.

## Модуль 7

### Архівування та довгострокове збереження даних

- 7.1. Різниця між поточним зберіганням та довгостроковим збереженням даних.
- 7.2. Які стандарти та формати даних використовуються для архівування даних?
- 7.3. Які дані варто/можна поширювати?
- 7.4. Типи та можливості репозитаріїв для довгострокового збереження даних.

Рекомендовані матеріали до Модулю 7.

1. 4TU.ResearchData. Long-term storage - <http://researchdata.4tu.nl/en/publishing-research/long-term-storage/>
2. Digital Repository Audit Method Based on Risk Assessment (DRAMBORA) : <http://www.repositoryaudit.eu/about/>
3. Preservation Pressure Points: Evaluating Diverse Evidence for Risk Management <http://www.repositoryaudit.eu/images/PreservationPressurePoints.pdf>
4. Re3data - <https://www.re3data.org/>
5. Trusted Repository Audit Checklist (TRAC) <https://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying/trac>
6. Whyte, A. & Wilson, A. (2010). "How to Appraise and Select Research Data for Curation". DCC How-to Guides. Edinburgh: Digital Curation Centre. <http://www.dcc.ac.uk/resources/how-guides>
7. Кучма Ірина. Практики управління науковими даними: Європейський досвід: [презентація до семінару "Цифрова підтримка наукових досліджень", 9 жовтня 2017, НаУКМА, Київ] - <http://ekmair.ukma.edu.ua/handle/123456789/11979>
8. Підготувати план із архівування та довгострокового зберігання даних, використовуючи як шаблон [http://researchdata.4tu.nl/fileadmin/editor\\_upload/pdf/Preservation\\_Plan/4TU.Preservation\\_Plan.pdf](http://researchdata.4tu.nl/fileadmin/editor_upload/pdf/Preservation_Plan/4TU.Preservation_Plan.pdf)

Практичне завдання до модулю 7 (5 балів)

Скласти план довготривалого зберігання для даних дослідницького проекту слухача.

#### Проміжне оцінювання 2

Тестування за інформацією, що засвоєна за результатами модулів 4-7.

#### Фінальний екзамен

Представлення детального плану управління даними власного дослідження.