Казахский национальный университет имени аль-Фараби
Факультет медицины и здравоохранения
Кафедра фундаментальной медицины

Лекция 11.1

Управление данными научных исследований для раскрытия их полного потенциала

Дисциплина: Дизайн экспериментов в разработке лекарственных средств (DERLS 5208)

Образовательная программа: 7M10117 – Фармация (магистратура)

Лектор: PhD, ассоциированный профессор Калыкова А.С.

Алматы, 2026

Цель

Сформировать компетенции в области управления научными данными: жизненный цикл данных, принципы FAIR, план управления данными (DMP), целостность данных (ALCOA+), репозитории и этические аспекты открытых данных.



Вопросы для обсуждения

- Что такое принципы FAIR и зачем они нужны?
- Из чего состоит план управления данными (DMP)?
- Как обеспечить целостность данных в фармацевтических исследованиях?

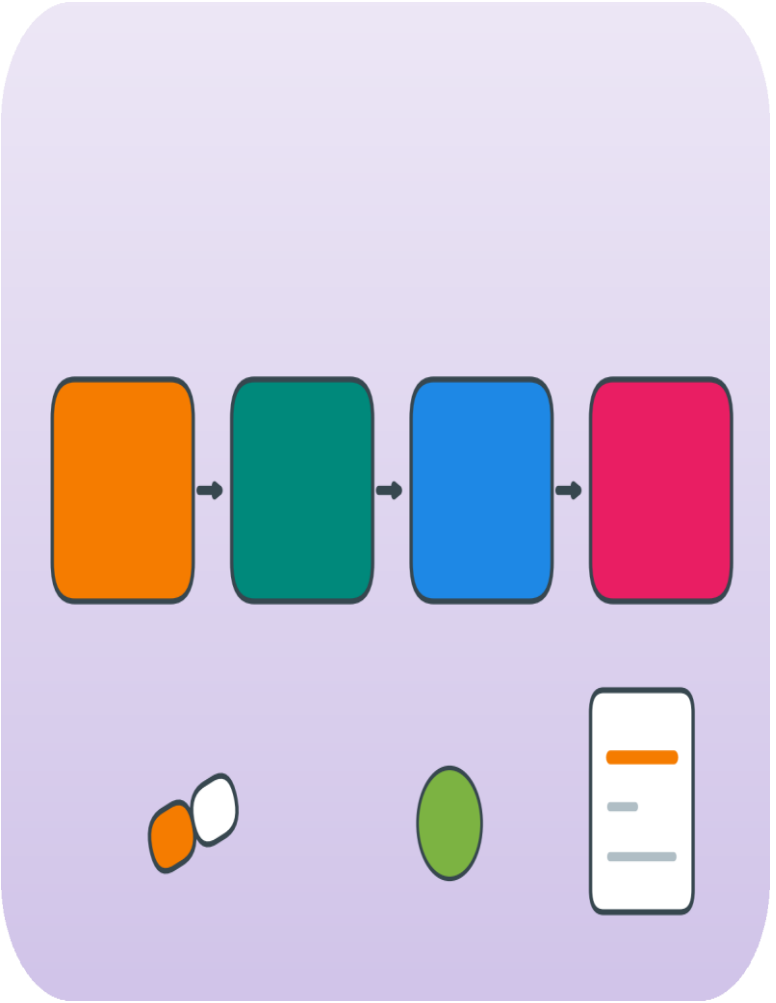
Данные как ресурс исследования

- Современный эксперимент генерирует большие объёмы данных: спектры, хроматограммы, изображения, геномные и протеомные наборы.
- Управление данными (research data management, RDM) – планирование, организация, хранение, защита и публикация данных.
- Хорошо управляемые данные повышают воспроизводимость, позволяют повторное использование и экономят ресурсы.
- Финансирующие организации и журналы всё чаще требуют план управления данными и открытый доступ.



Жизненный цикл исследовательских данных

1. **Планирование** DMP, форматы, ответственность.
2. **Сбор** протокол, метаданные, контроль качества.
3. **Обработка и анализ** очистка, версии, скрипты.
4. **Хранение и защита** резервирование, права доступа, шифрование.
5. **Публикация и обмен** репозиторий, лицензия, DOI.
6. **Долгосрочное хранение** архивирование, повторное использование.



Принципы FAIR (Wilkinson et al., 2016)

01

Findable – находимые

Постоянный идентификатор (DOI), богатые метаданные, индексирование.

02

Accessible – доступные

Стандартный протокол доступа, метаданные сохраняются, даже если данные закрыты.

03

Interoperable – совместимые

Открытые форматы, словари и онтологии, ссылки на другие данные.

04

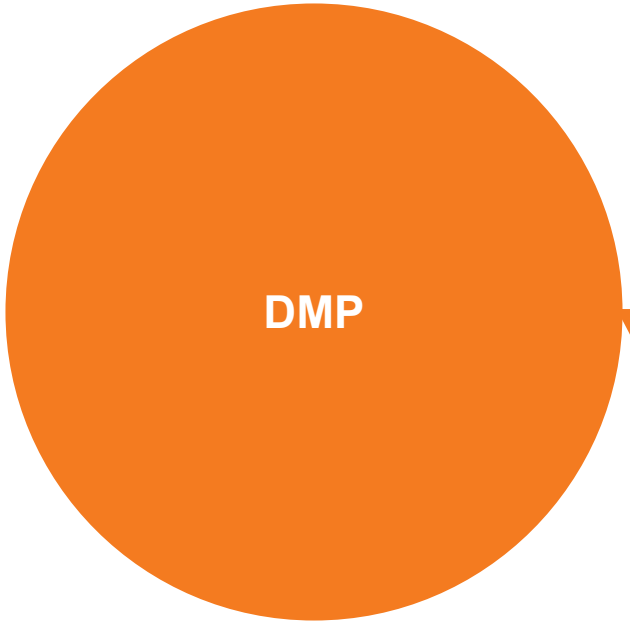
Reusable – повторно используемые

Чёткая лицензия, происхождение данных, соответствие стандартам области.



FAIR не равно «открытые»: данные могут быть защищены, но должны быть описаны и доступны на понятных условиях.

План управления данными (DMP)



- Описание данных: типы, объёмы, форматы, источники
- Стандарты и метаданные: словари, онтологии, структура файлов
- Хранение и резервирование: правило 3-2-1 (3 копии, 2 носителя, 1 внешняя)
- Безопасность и доступ: роли, шифрование, конфиденциальность
- Публикация и лицензии: репозиторий, CC BY, эмбарго
- Ответственность, сроки хранения и ресурсы

Целостность данных: принципы ALCOA+

Принцип	Расшифровка	Пример в лаборатории
A – Attributable	Атрибутируемые	Подпись и дата исполнителя
L – Legible	Разборчивые	Читаемая запись, постоянный носитель
C – Contemporaneous	Своевременные	Запись в момент выполнения
O – Original	Оригинальные	Исходные файлы, а не копии
A – Accurate	Точные	Без исправлений без пометок
+	Полные, согласованные, длительные, доступные	Аудиторский след, архивирование

Руководство FDA «Data Integrity and Compliance with Drug CGMP» (2018); WHO TRS 1033, Annex 4 (2021).

Электронные системы: ELN, LIMS, версионирование

- **ELN (электронный лабораторный журнал):** записи экспериментов, прикреплённые файлы, подписи, время.
- **LIMS:** учёт образцов и результатов, потоки анализа.
- **Контроль версий (Git) и скрипты:** воспроизводимый анализ в R/Python.
- **Аудиторский след:** фиксирует, кто, когда и что изменил (21 CFR Part 11).



Метаданные и стандарты

- Метаданные описывают данные: кто, когда, чем, в каких условиях, какой версией метода.
- Стандарты: Dublin Core, DataCite, MIAME, MIAPE и специфические для химии и биологии форматы.
- Онтологии и словари: ChEBI, MeSH, Gene Ontology для единообразных терминов.
- Файлы именуют по шаблону: дата_проект_образец_версия.



Репозитории и публикация данных

- **Универсальные:** Zenodo (CERN), Figshare, Dryad, OSF.
- **Специализированные:** PDB (структуры), GenBank, ChEMBL, MetaboLights, PRIDE.
- **Поиск репозиториев:** re3data.org – реестр репозиториев.
- **DOI и цитирование:** данные цитируют как публикации.
- **Лицензии:** CC BY, CC0 – условия повторного использования.



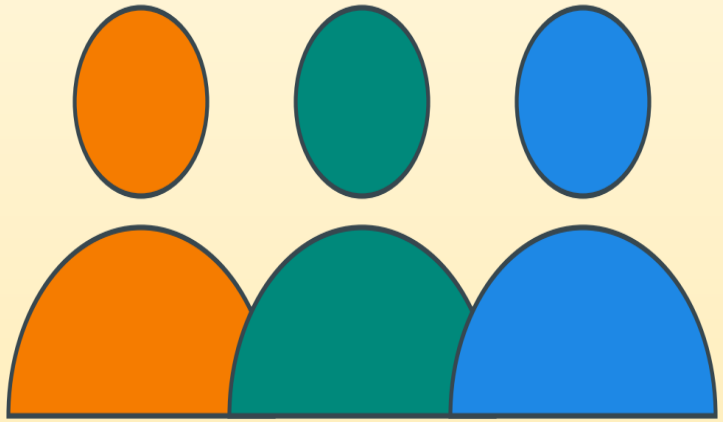
Этика, конфиденциальность и безопасность

- Персональные и медицинские данные: информированное согласие, деидентификация, ограничение доступа.
- Законодательство: Закон Республики Казахстан «О персональных данных и их защите» (2013), GDPR – при работе с данными граждан ЕС.
- Права интеллектуальной собственности и патентная чувствительность: время публикации и эмбарго.
- Принцип: «настолько открыто, насколько возможно, настолько закрыто, насколько необходимо».



Типичные ошибки в управлении данными

- Хранение единственной копии на личном носителе.
- Отсутствие метаданных и описания файлов.
- Правки «исходных» файлов без версий.
- Закрытые форматы, зависящие от конкретной программы.
- Публикация без лицензии и ссылки на исходный метод.



Видеоматериалы и интернет-ресурсы для демонстрации

Видео (YouTube)

Knowledge clip: FAIR data principles
Краткое объяснение принципов FAIR

FAIR Principles in under 60 seconds
FAIR за минуту

Introduction to effective and practical research data management
Введение в управление данными исследований

Сайты и базы данных

GO FAIR – FAIR Principles
Официальное описание принципов

FAIRsharing
Реестр стандартов, баз данных и политик

DMPTool
Конструктор плана управления данными

re3data
Реестр репозиториев исследовательских данных

Zenodo
Универсальный репозиторий с DOI



Заключение

- Управление данными – часть планирования эксперимента и условие воспроизводимости.
- Принципы FAIR делают данные находимыми, доступными, совместимыми и пригодными для повторного использования.
- DMP фиксирует форматы, хранение, доступ и ответственность.
- Принципы ALCOA+ обеспечивают целостность данных в фармации.
- Данные публикуют в репозиториях с метаданными, DOI и лицензией.

Список литературы

1. Wilkinson M.D., Dumontier M., Aalbersberg I.J. et al. The FAIR Guiding Principles for scientific data management and stewardship // Scientific Data. 2016. Vol. 3. P. 160018. https://doi.org/10.1038/sdata.2016.18
2. U.S. Food and Drug Administration. Data Integrity and Compliance with Drug CGMP: Questions and Answers. Guidance for Industry. 2018. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/data-integrity-and-compliance-drug-cgmp-questions-and-answers
3. World Health Organization. Guideline on data integrity. In: WHO Expert Committee on Specifications for Pharmaceutical Preparations: 55th report. Geneva: WHO, 2021. (Technical Report Series, No. 1033, Annex 4).
4. European Commission. Guidelines on FAIR Data Management in Horizon 2020. 2016. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
5. GO FAIR Initiative. FAIR Principles. https://www.go-fair.org/fair-principles/ (дата обращения: 11.11.2026).
6. Закон Республики Казахстан от 21.05.2013 № 94-V «О персональных данных и их защите». https://adilet.zan.kz/
7. Видео: см. слайд «Видеоматериалы» (YouTube; дата обращения: 11.11.2026).