

*Открытая наука и практика управления
научными данными в университетах
Европы*

Соколова Наталия Викторовна,
Санкт-Петербургский политехнический университет Петра Великого

Эволюция «открытости»

Открытый доступ >

открытые данные >

открытая наука



Открытая наука – что это?

- ▶ Зонтичный термин для движения, цель которого — *сделать научные исследования, данные и их распространение доступными для всех уровней заинтересованного общества, будь то любители или профессионалы.* [Википедия]
- ▶ Основными задачами является *популяризация науки и научной деятельности, общественный контроль качества научных публикаций, развитие междисциплинарных исследований, современного института научной рецензии, повышение цитируемости и построение инфраструктуры знаний.* [КиберЛенинка]
- ▶ Издание должно иметь официальный сайт или страницу на сайте в информационно-телекоммуникационной сети "Интернет", где *в открытом доступе ... размещается информация об ... аннотации научных статей и ключевые слова для всех научных статей и обзоров, опубликованных изданием за последние 2 года.* [Приказ Минобрнауки РФ от 25.07.2014 N 793]

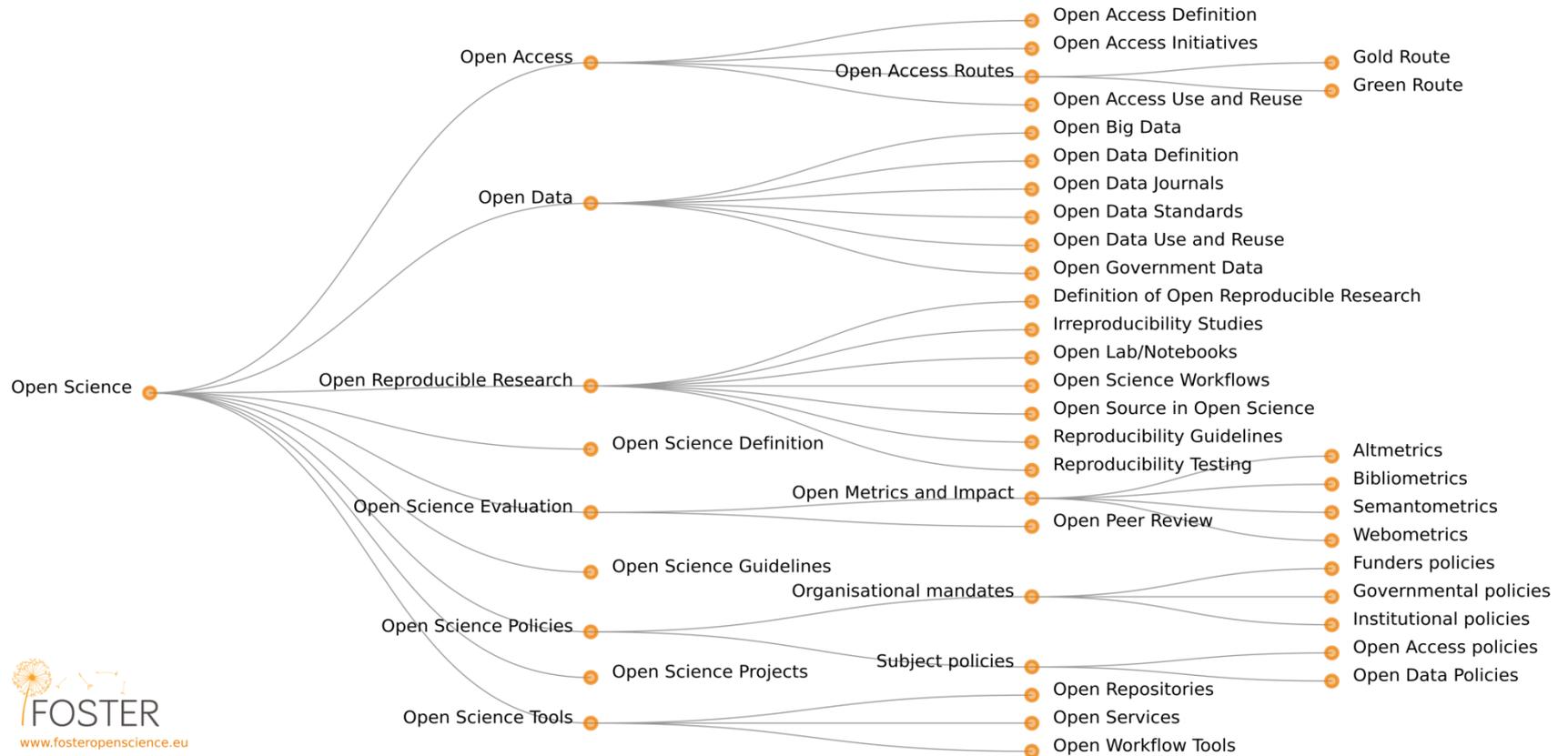
Открытая наука – что это?



«Новый подход к организации научного процесса, базирующийся на совместной работе и новых способах распространения знаний с использованием цифровых технологий и сред для совместной работы»

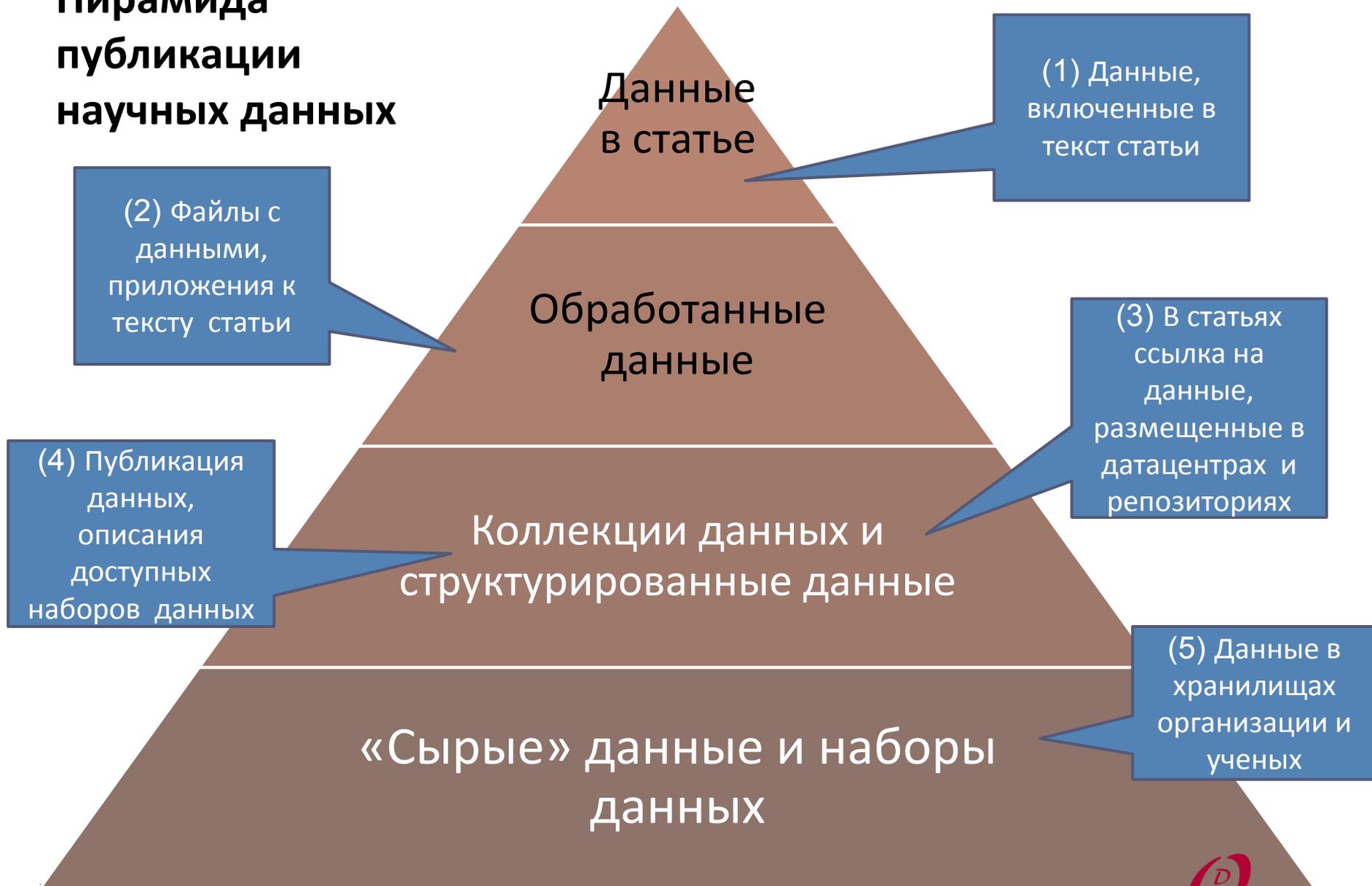
[European Commission, 2016]

Open Science Taxonomy

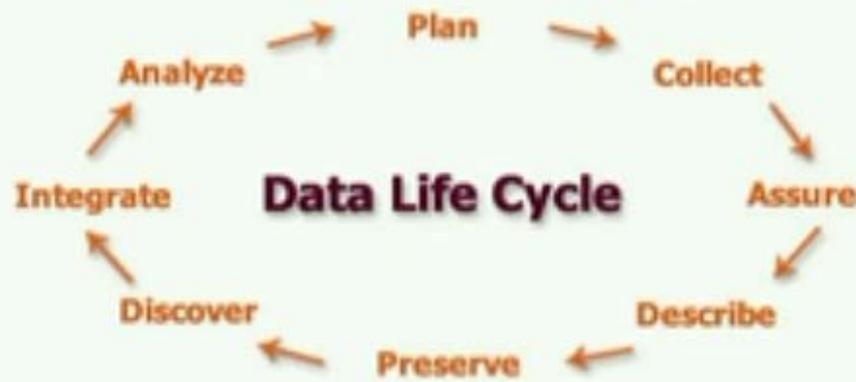


«Open Science это практика организации научной деятельности, при которой каждый может принимать участие и вносить свой вклад , когда научные данные, описания экспериментов и научных процессов доступны на условиях, позволяющих их повторно использовать, распространять и воспроизводить, включая процессы, данные и применяемые методы».

Пирамида публикации научных данных



Жизненный цикл научных данных



Программы и проекты ЕС

- ▶ Horizon 2020, Work Programme 2018-2020, European research infrastructures (including e-Infrastructures)
- ▶

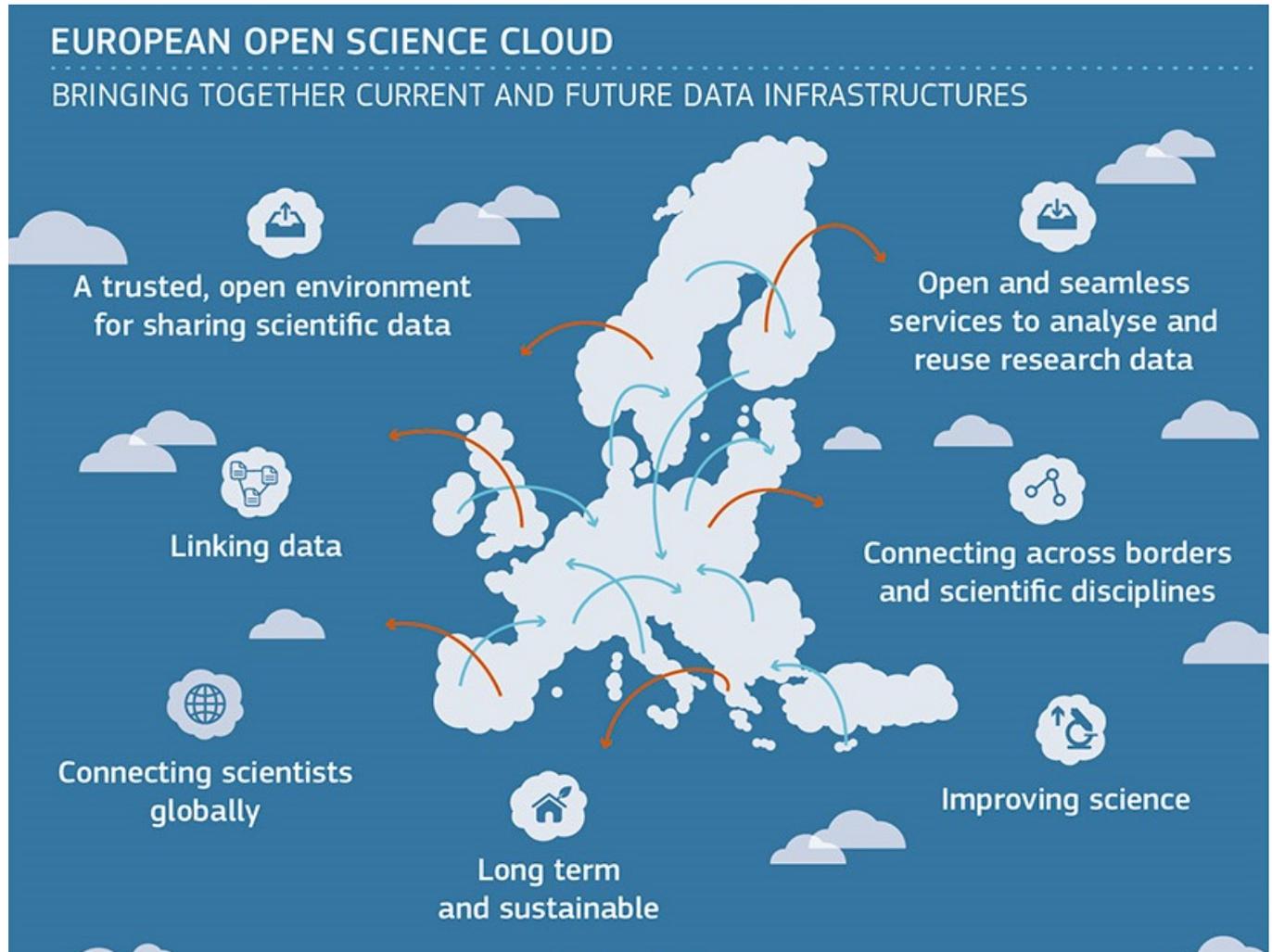
European Open Science Cloud (EOSC)

This is a cloud for research data in Europe. Background, policy information, events and publications related to the EOSC

Из Программы Horizon 2020

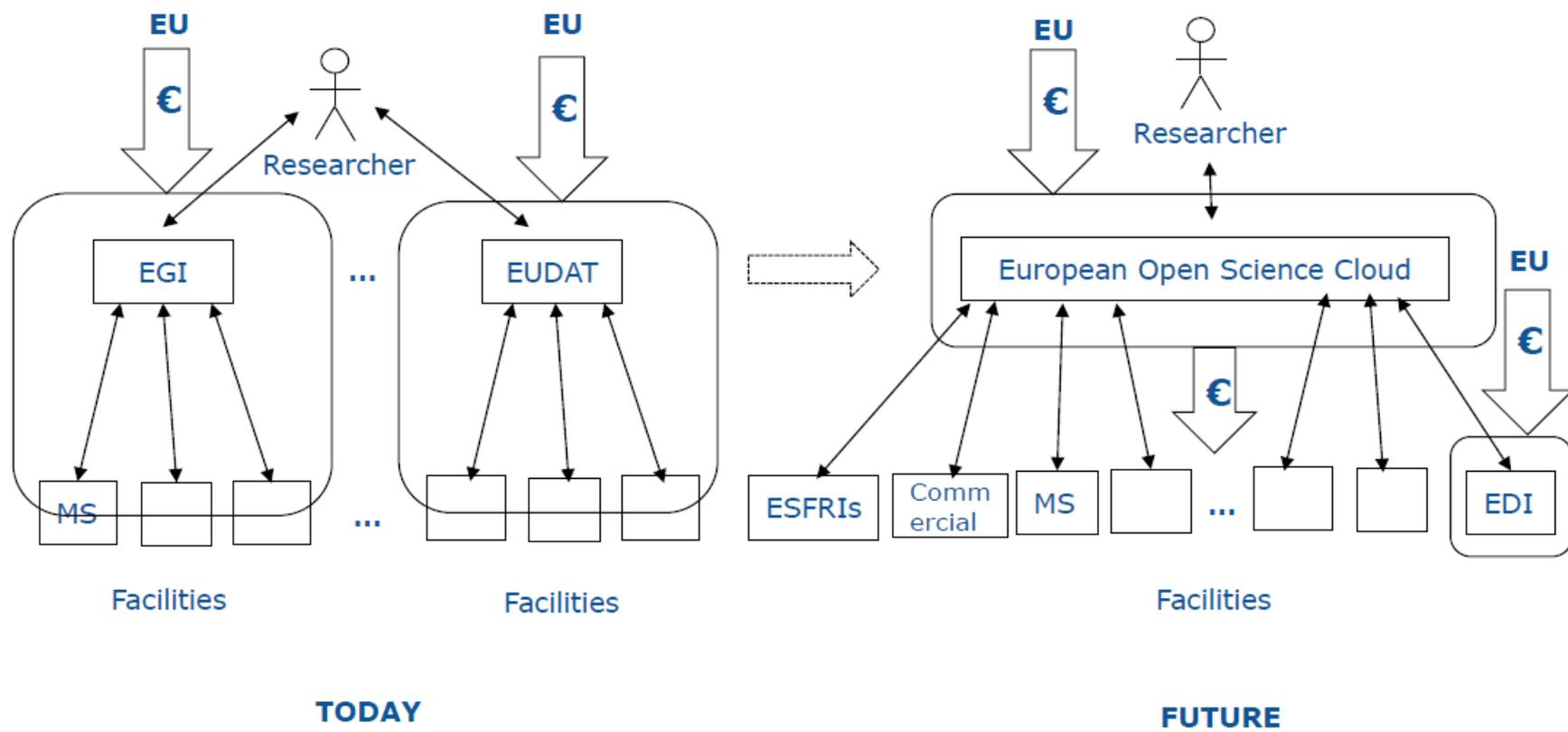
- ▶ Инфраструктура науки – это средства, ресурсы и сервисы, используемые научным сообществом для управления научными процессами и стимулирования инноваций. Инфраструктура науки может быть централизованной, виртуальной или распределенной.
- ▶ Инфраструктура науки включает:
 - ▶ Базовое научное оборудование и наборы инструментов;
 - ▶ Научные знания : коллекции, архивы и научные данные;
 - ▶ Е-инфраструктура: хранилища данных, вычислительные средства и коммуникации;
 - ▶ Иные средства, требуемые для проведения исследований.
- ▶ Инфраструктура науки должна обеспечить к 2020 единое открытое пространство ЕС для онлайн-исследований, включая повсеместно доступные и надежные сетевые и вычислительные сервисы, бесшовный и открытый доступ к оборудованию Е-науки и мировым информационным ресурсам.
- ▶ **Все грантополучатели Программы «по умолчанию» должны обеспечивать открытость научных данных.**

European Open Science Cloud

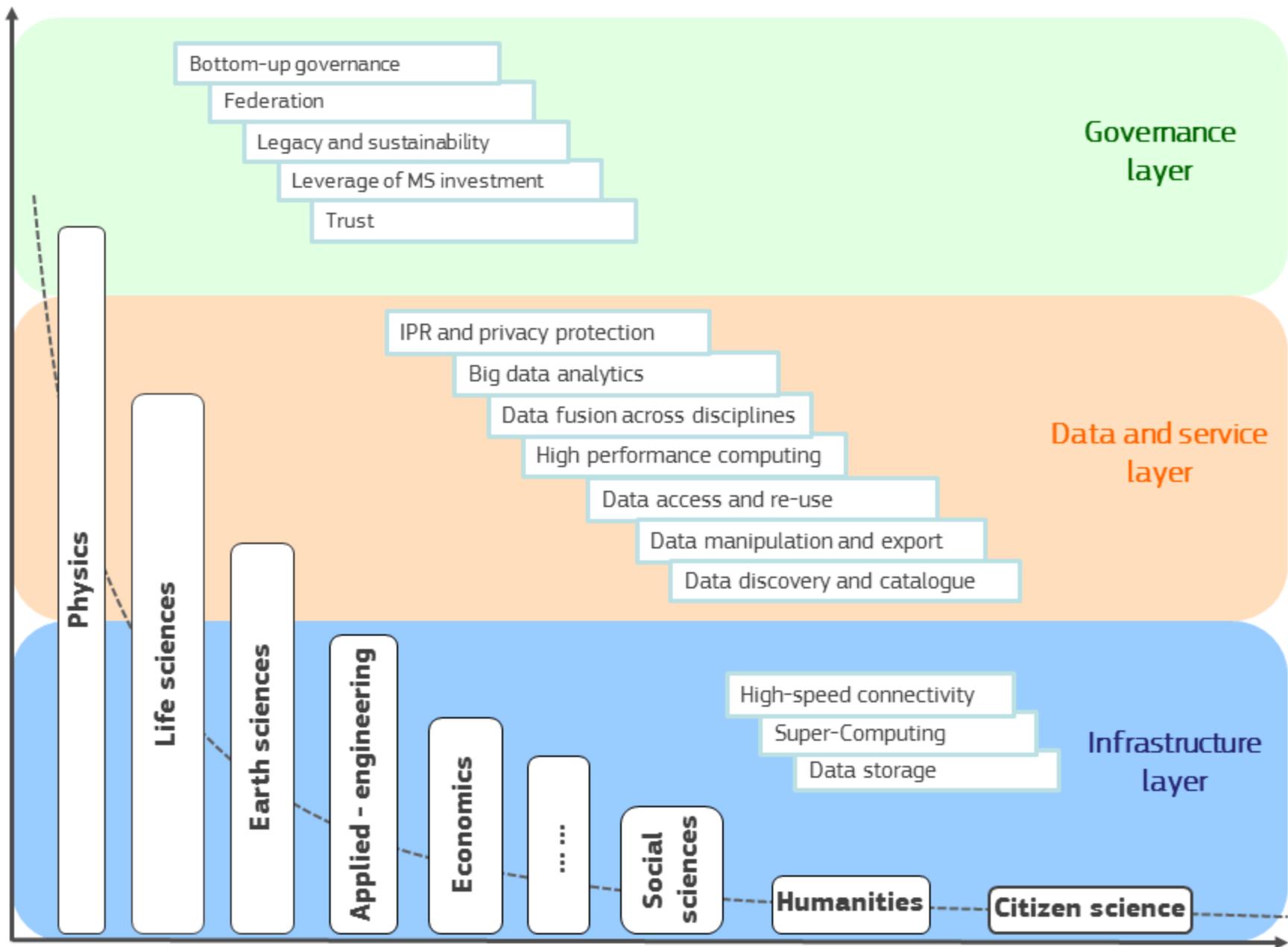


Смена модели взаимодействия

Users, suppliers and access channels



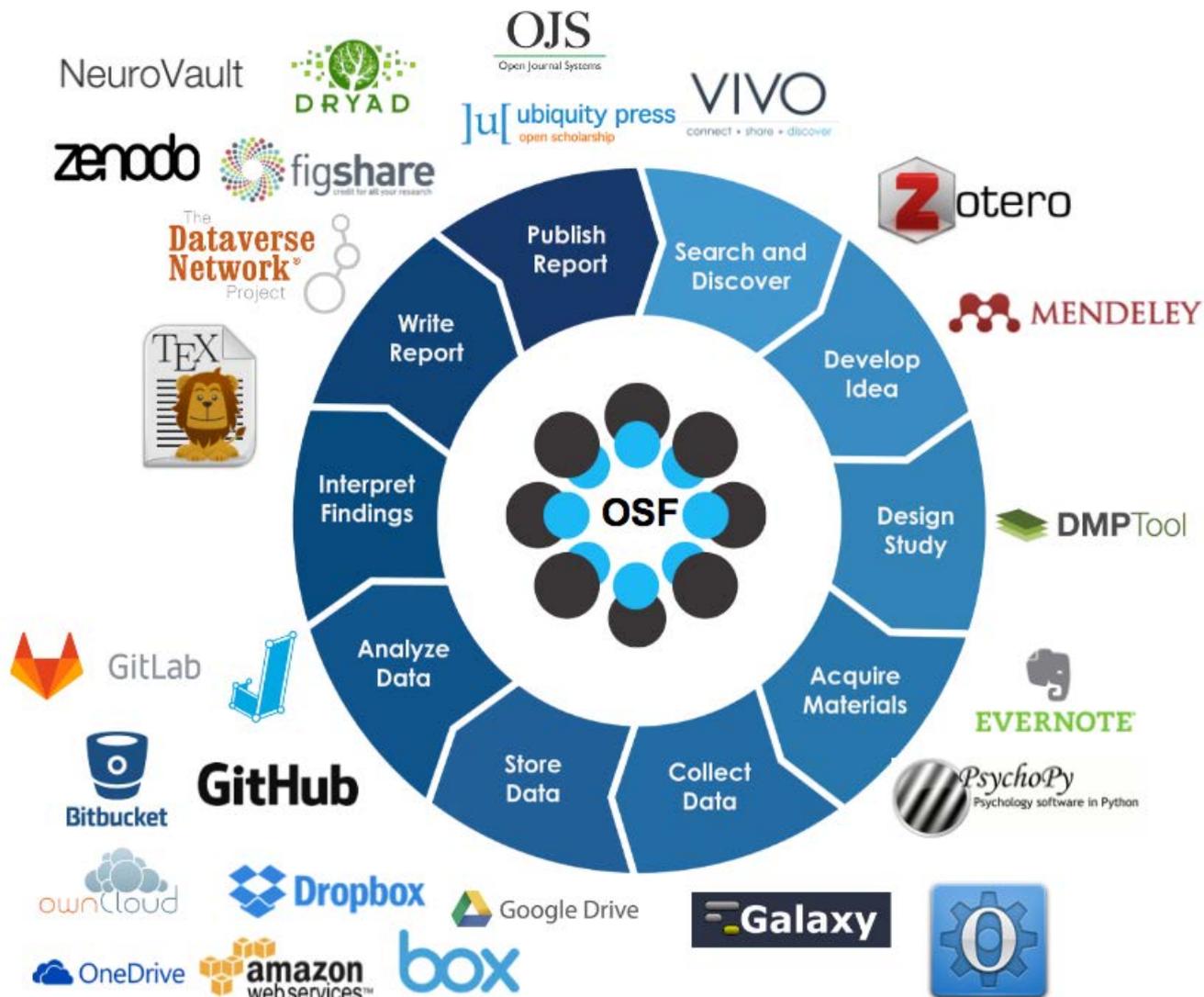
Scale of scientific activity (data-driven science)



Lead scientific users...

...long tail of science

Облачные средства для открытой науки



FAIR научные данные

What is FAIR DATA?



Data and supplementary materials have sufficiently rich metadata and a unique and persistent identifier.

FINDABLE



Metadata and data are understandable to humans and machines. Data is deposited in a trusted repository.

ACCESSIBLE



Metadata use a formal, accessible, shared, and broadly applicable language for knowledge representation.

INTEROPERABLE



Data and collections have a clear usage licenses and provide accurate information on provenance.

REUSABLE

Findable		Accessible		Interoperable		Reusable	
1	(мета)данные имеют уникальный постоянный идентификатор	1	(мета)данные могут быть найдены по идентификатору с использованием стандартного коммуникационного протокола	1	(мета)данные используют формальный, доступный и широко распространенный язык для представления знаний	1	(мета)данные являются «богатými», с множеством точных и релевантных атрибутов
2	данных описаны «богатými» метаданными(см. R1)	1.1	протокол доступа является открытым и широко распространенным	2	(мета)данные используют словари, следующие FAIR принципам	1.1	(мета)данные опубликованы с ясной лицензией, дающей право на использование данных
3	метаданные включают ясный и однозначный идентификатор данных, которые они описывают	1.2	протокол обеспечивает аутентификацию и авторизацию, когда это требуется	3	(мета)данные включают применимые ссылки на другие (мета)данные	1.2	(мета)данные однозначно ассоциированы с источником их происхождения
4	(мета)данные зарегистрированы/ проиндексированы в поисковых системах	2	метаданные доступны даже тогда, когда сами данные становятся недоступны	4	(мета)данные должны быть в стандартном формате обмена	1.3	(мета)данные соответствуют стандартам предметной области

10 рекомендаций библиотекам для работы с научными данными



Ligue des Bibliothèques Européennes de Recherche
Association of European Research Libraries

1. Предлагайте ученым поддержку, включая планы управления данными для грантовых приложений, по вопросам копирайта, информационного обеспечения. Включайте вопросы управления научными данным в учебные планы.
2. Занимайтесь развитием стандартов данных и метаданных, предоставляйте сервис создания метаданных для ученых.
3. Повышайте квалификацию сотрудников в области управления научными данными.
4. Активно участвуйте в разработке политики и организации жизненного цикла научных данных, по возможности, применяя политики открытых данных.
5. Объединяйтесь и сотрудничайте с научными группами. Архивами данных и датацентрами для построения интероперабельной инфраструктуры работы с данными.

10 рекомендаций библиотекам для работы с научными данными



Ligue des Bibliothèques Européennes de Recherche
Association of European Research Libraries

6. Поддерживайте жизненный цикл научных данных, обеспечивая сервисы хранения, поиска и доступа.
7. Пропагандируйте цитирование научных данных с использованием постоянного идентификатора.
8. Создавайте институциональный каталог данных или репозиторий, в зависимости от имеющейся инфраструктуры.
9. Участвуйте в процессе управления данными для предметных областей.
10. Предлагайте свои хранилища или становитесь посредником для обеспечения надежного хранения научных данных в кооперации с ИТ-подразделениями организации, рекомендуйте подходящие облачные хранилища.

Планирование управлением данными

Data Management Planning: Tools

Information on best practices and standards for data management planning.

[Home](#)[Data Creation](#)[Storage and Backup](#)[Metadata](#)[Data Sharing and Publishing](#)[Data Management Plans for Grants](#)[Resources](#)

General Information

NYU
DataServices

NYU LIBRARIES & IT

[Data Services Homepage](#)

[Location](#)

[Staffed Hours: Fall 2018](#)

[Contact Us](#)

[Tell us how we're doing](#)

DMP Tool @ NYU



The [DMP Tool Online](#) is a web-based tool to help you build your data management plan. It provides step-by-step guidance and information specific to many United States granting agencies and their directorates.

NYU affiliates can sign into the DMP Tool using their netID and password from the [institutional log-in page](#). Select "New York University" from the drop down menu and click "Go!"

DMPOnline



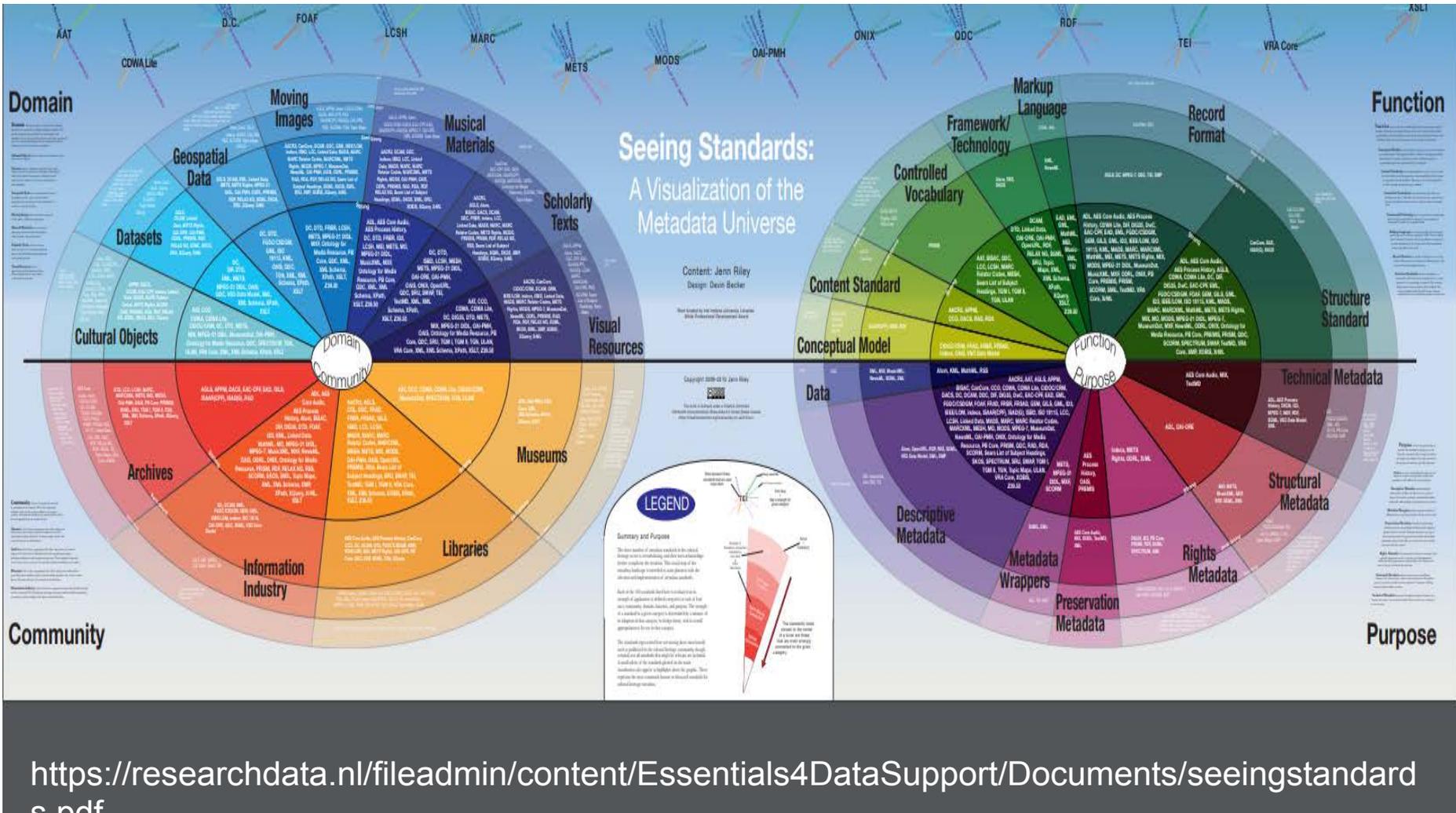
[DMPonline](#) is an UK-based tool developed by the [Digital Curation Centre](#) in close collaboration with research funders and universities. This tool is another online utility that assists researchers in producing an effective DMP to cater to the whole lifecycle of a project. There are many templates in the tool that represent the requirements of a diverse selection of funders and institutions. Users are asked three questions to determine the appropriate template to display; upon selection, guidance is provided to help researchers interpret and answer the questions within the template.

All that is necessary to use this tool is an account, which anyone can sign up for. There is an extensive user community that reports bugs and requests new features directly on [GitHub](#). The past versions are also stored and are accessible for those using the earlier features.

План управления научными данными

1. Общие сведения о проекте и управлении данными – в т.ч. виды данных, которые будут предоставлены для повт. использования: ПО, документация, аудио, видео, фото, версии обработанных данных, сырые данные, т.п.
2. Юридические вопросы – соблюдения норм работы с данными в предм. области, правообладание, пр.
3. Обеспечение «находимости» (findibility) данных
4. Обеспечение доступности (accessibility)
5. Обеспечение интероперабельности
6. Обеспечение повт. использования – вкл. стоимость

Стандарты метаданных



<https://researchdata.nl/fileadmin/content/Essentials4DataSupport/Documents/seeingstandards.pdf>

CDWA Lite

METS

Domain

Metadata

Access

Usage

Preservation

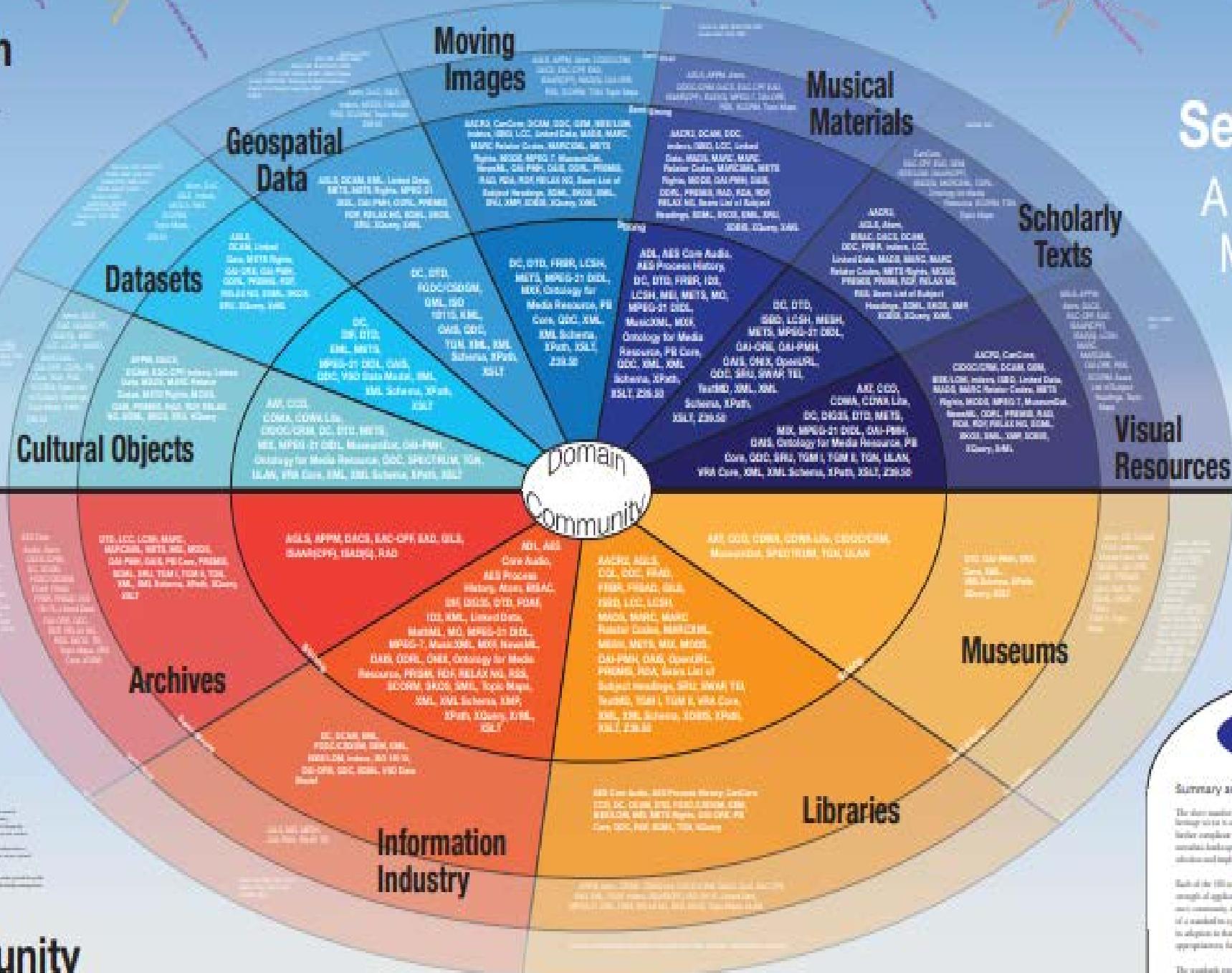
Interoperability

Community

Summary and

Each of the 100 words

The metadata system



See A V M

Summary and

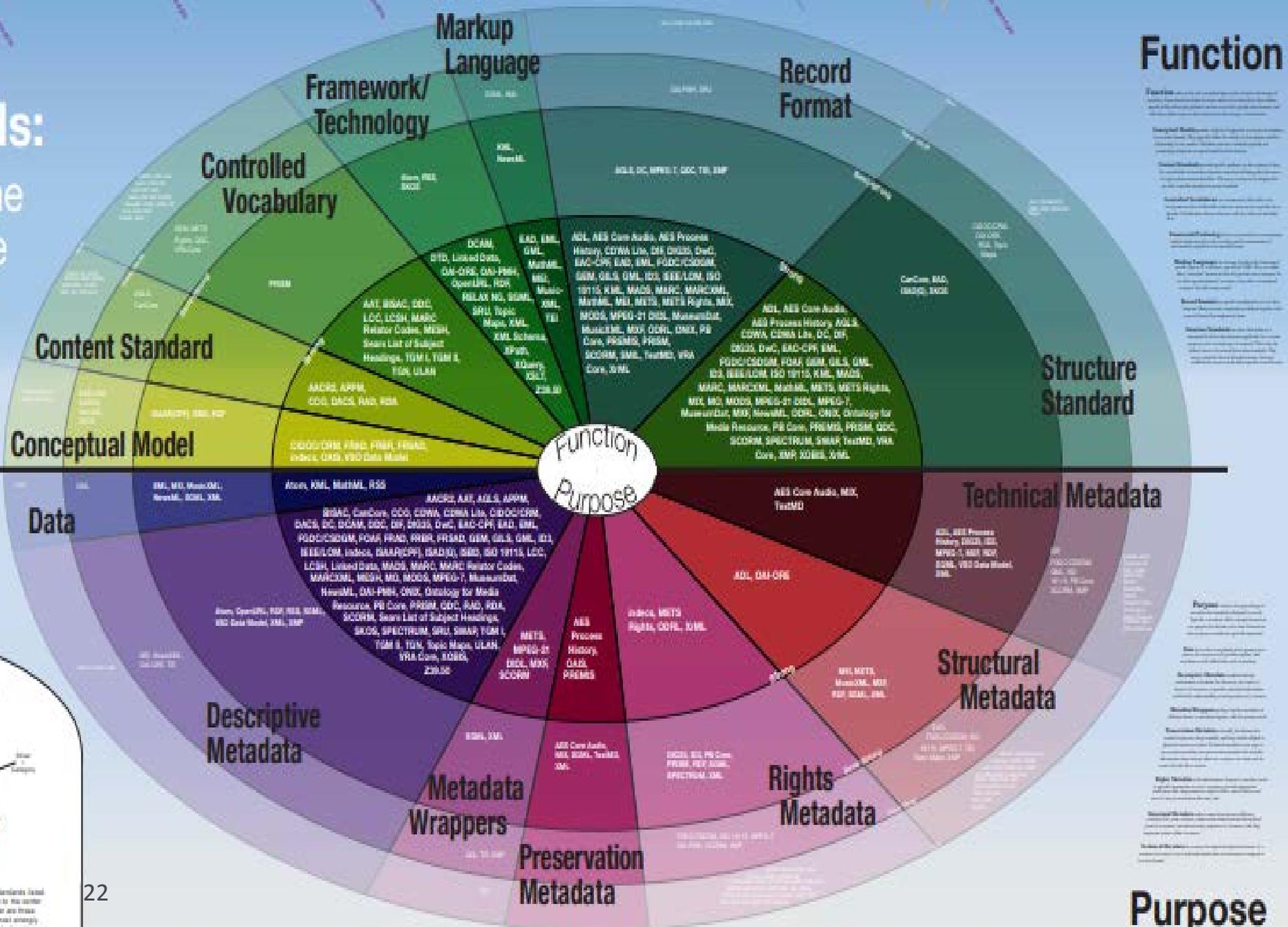
Each of the 100 words

The metadata system

Community

ds:
he
se

Function



Function

Record Format

Structure Standard

Technical Metadata

Structural Metadata

Rights Metadata

Preservation Metadata

Metadata Wrappers

Descriptive Metadata

Data

Conceptual Model

Content Standard

Controlled Vocabulary

Framework/Technology

Markup Language

The standards listed around the center of this diagram are those that are most closely related to the purpose.

Purpose

Что могут библиотекари

Данные :	Возможности для библиотекарей и датацентров:
Находимость (Findability)	<ul style="list-style-type: none"> ✓ Обеспечивайте постоянные идентификаторы. ✓ Участвуйте в разработке единых схем метаданных и единых практик цитирования. ✓ Пропагандируйте использование учеными единых стандартов и средств.
Доступность (Availability)	<ul style="list-style-type: none"> ✓ Помогайте ученым снижать сложность задачи обеспечения доступности данных. ✓ Включайте наборы данных в поисковые индексы и службы.
Интероперабельность (Interpretability)	<ul style="list-style-type: none"> ✓ Обеспечивайте связанность между публикациями и наборами данных. ✓ Помогайте ученым в создании метаданных для наборов данных. ✓ Создайте и поддерживайте базу знаний с наборами данных в контексте предметной области.
Повторное использование (Re-usability)	<ul style="list-style-type: none"> ✓ Обеспечивайте сохранность данных. ✓ Создавайте архивы ПО, требуемого для анализа данных. ✓ Обеспечивайте выполнение условий для повторного использования данных.
Цитируемость (Citability)	<ul style="list-style-type: none"> ✓ Участвуйте в разработке единых стандартов цитирования. ✓ Внедряйте использование постоянных идентификаторов.
Сохранность (Curation/Preservation)	<ul style="list-style-type: none"> ✓ Изучайте и внедряйте лучшие практики управления данными. ✓ Сотрудничайте с создателями данных. ✓ Помогайте ученому в выборе примера управления данными, характерными именно для его предметной области (форматы хранения, документация об экспериментах,...)

Библиотекари будущего?

Tell us how we're doing

@NYUDataServices

Tweets by @nyudataservices

Data Services
@nyudataservices

In the city that never sleeps, data is always at work. Check out NYC's 2018 Open Data report #NYCDataAtWork www1.nyc.gov/site/doitt/abo...#opendata - Finally! A comprehensive list of Open Data that's maintained by the city

59m

Data Services
@nyudataservices

Today's @tableau class at Data Services was packed! Click here to see what they learned and sign up for the next one if you missed it. - guides.nyu.edu/c.php?g=851842...



Embed

View on Twitter

Research Lifecycle



Data management is done at all stages of the research lifecycle.

*each step in the process has its own best practices & standards

Research Data Management Librarians



Nick Wolf

Research Data Management Librarian
Bobst Library, 512



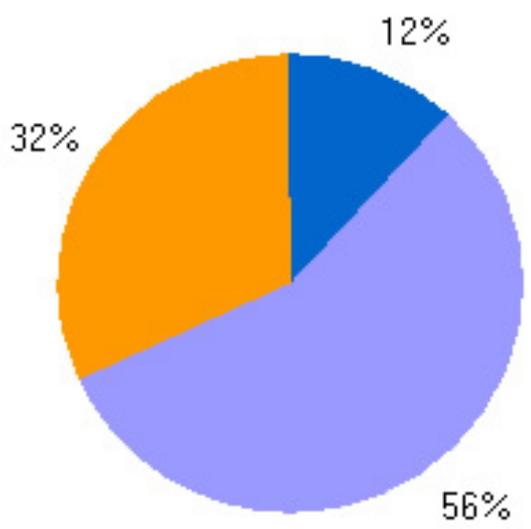
Vicky Steeves

Research Data Management & Reproducibility Librarian
Bobst Library, 517



Developing our own workforce!

Do you feel that your library has the right skills to be prepared for such activities?



- Yes
- Not yet, but the library is investing in developing data managers, data curators, or similar skills
- No, and my library hasn't begun to develop such skills



А мы открыты для открытой науки?

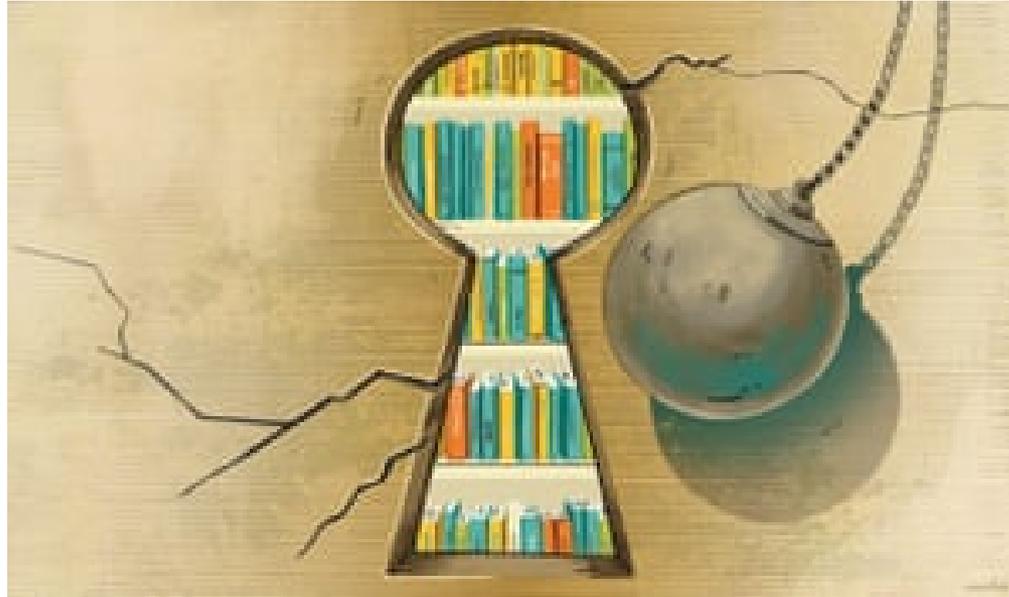


Illustration: Eva Bee
<https://clck.ru/ELftu>

Спасибо за внимание!

