



ПОЛИТЕХ

Информационно-библиотечный
комплекс

Идентификация объектов в глобальном цифровом пространстве

Национальный центр идентификации
научных данных

24.11.2016

Содержание

- Цели и задачи идентификации цифровых объектов в Интернет
- Инфраструктура DOI
- Регистрирующие агентства
- DataCite и научные данные
- Кто регистрирует научные данные
- Что дает присваивание DOI



Анализ текущей ситуации

- «Цифровизация или переход (поворот) внимания общества от материальных носителей культурного содержания в пользу их цифровых аналогов и репрезентаций является прогрессирующей тенденцией современной культуры»
- Open Access, Open Science, Open Education
- Лицензии Creative Commons действуют в РФ с 01.01.2015
- Digital Humanities (гуманитарная информатика)
- Big Data, Data Mining



Цифровые объекты и цифровизация

Свойства цифровых объектов отличаются от материальных (аналоговых)

- Гранулированность произвольного уровня
- Принципиально полноценная воспроизводимость
- Нелокальность во времени и пространстве
- Изменение структуры стоимости владения – отсутствие затрат на тиражирование копий, но затраты на сохранность и доступность



Положительные стороны

- Облегчение доступа к культурному наследию
- Рост связности объектов в сфере образования, науки и культуры
- Объективный и оперативный анализ использования (чтение, цитирование)
- Цифровой объект дает принципиально новые возможности для получения новых знаний

**НО! Все работает неэффективно
без глобальной идентификации
цифровых объектов**



Отрицательные стороны цифровизации

- Рост количества нарушений прав на РИД
- Снижение мотивации к заучиванию и запоминанию информации
- Смена способа работы со знаковыми системами в массовом масштабе
- Издательская деятельность, научный и образовательный процессы требуют реорганизации

**Идентификация цифровых объектов
содействует преодолению отрицательных
сторон**



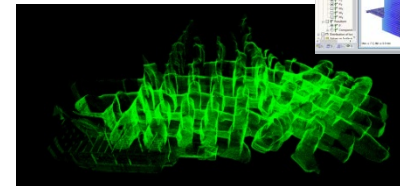
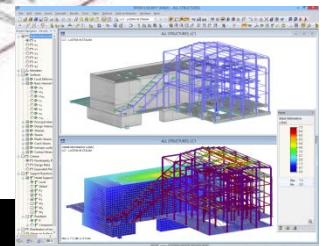
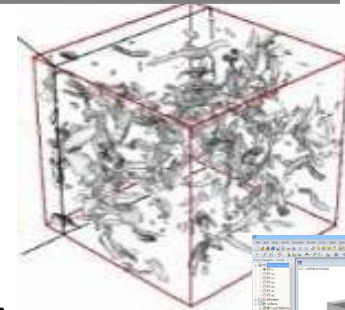
Смена парадигм науки

- **Тысячи лет назад:**
 - наука эмпирическая: описание наблюдаемых явлений
- **Сотни лет назад:**
 - науки теоретическая: модели, обобщения
- **Десятки лет назад:**
 - вычислительные модели
- **Сегодня:**
 - eScience – объединение теории, экспериментов и симуляции

Jim Gray, eScience Group, Microsoft Research



$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G\rho}{3} - K\frac{c^2}{a^2}$$



Пирамида научных данных



Пример связанных данных

The dataset: - Набор данных

Storz, D et al. (2009):

Planktic foraminiferal flux and faunal composition of sediment trap L1_K276 in the northeastern Atlantic.

<http://dx.doi.org/10.1594/PANGAEA.724325>

Is supplement to the article: - Приложение к статье

Storz, David; Schulz, Hartmut; Waniek, Joanna J; Schulz-Bull, Detlef; Kucera, Michal (2009): *Seasonal and interannual variability of the planktic foraminiferal flux in the vicinity of the Azores Current.*

Deep-Sea Research Part I-Oceanographic Research Papers, **56(1)**, 107-124,

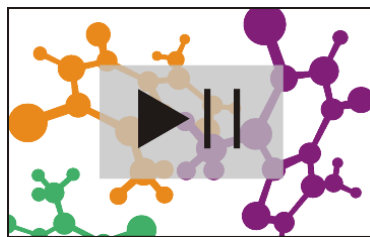
<http://dx.doi.org/10.1016/j.dsr.2008.08.009>



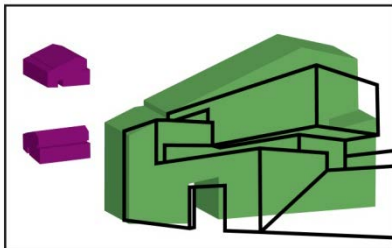
Новые виды ресурсов (не только «классические» научные публикации)

```
-<state>  
<language>de<  
<position/>  
<length/>4</  
<name>peter</  
<favourites>  
  -<favourite>
```

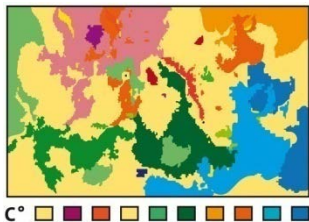
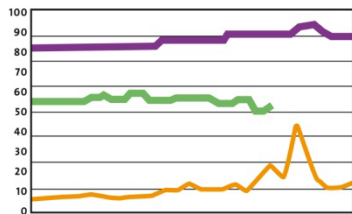
Программы



Научные фильмы



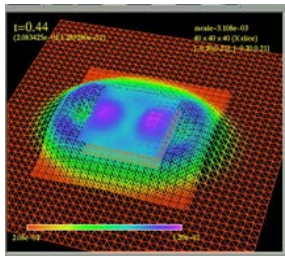
3D объекты



«Сырые» данные

TexteTexte	
TexteTexte Texte	
TexteTexte	
TexteTexte Texte	
TexteTexte	
TexteTexte	
TexteTexte	TexteTexte
TexteTexte Texte	TexteTexte Texte
TexteTexte	TexteTexte

«Серая» литература



Симуляция



Применяемые системы идентификации

- ISBN (International Standard Book Number)
- ASIN (Amazon Standard Identification Number)
- ETTN (Electronic Textbook Track Number)
- ISAN (International Standard Audiovisual Number)
- ISMN (International Standard Music Number)
- ISWC (International Standard Musical Work Code)
- ISSN (International Standard Serial Number)
- SICI (Serial Item and Contribution Identifier)
- **DOI (Digital Object Identifier)**



DOI – стандарт ИСО с 2012 г.



- Стандарт ISO 26324:2012 «Информация и документирование. Система цифровых идентификаторов объектов»
- Наименование DOI является идентификатором физических, цифровых или абстрактных объектов в цифровых сетях. Содержит информацию об объекте, включая местоположение в интернете, где может быть найден объект или информация о нем
- Используется, в первую очередь, для обмена с заинтересованными сообществами и для управления интеллектуальной собственностью
- Описывает синтаксис, определения и интерпретацию функциональных компонентов системы цифровых идентификаторов объектов, а также базовые принципы создания, регистрации и администрирования имен DOI





www.doi.org

HOME | HANDBOOK | FACTSHEETS | FAQs | RESOURCES | REGISTRATION AGENCIES | NEWS | MEMBERS AREA

The DOI® System

ISO 26324



This is the web site of the International DOI Foundation (IDF), a not-for-profit [membership organization](#) that is the governance and management body for the [federation of Registration Agencies](#) providing Digital Object Identifier (DOI) services and registration, and is the registration authority for the ISO standard (ISO 26324) for the DOI system. The DOI system provides a technical and social infrastructure for the registration and use of persistent interoperable identifiers, called DOIs, for use on digital networks.

Система DOI была создана в 1998 году Международной организацией DOI, некоммерческой организацией членом, организованной по инициативе нескольких издательских компаний. Международная организация DOI является регистрирующим органом для ISO 26324.



Цифры и факты

- Более 133 млн. DOI зарегистрировано
- Около 15% - ежегодный прирост
- Более 5000 организаций, присваивающих DOI ресурсам (издательства, научные дата-центры, киностудии, пр.)
- Более 5 млрд. обращений ежегодно (поиск через DOI)
- 90% самый рейтинговых журналов в WoS имеют DOI (по анализу 2014 г.)
- 9 регистрирующих агентств

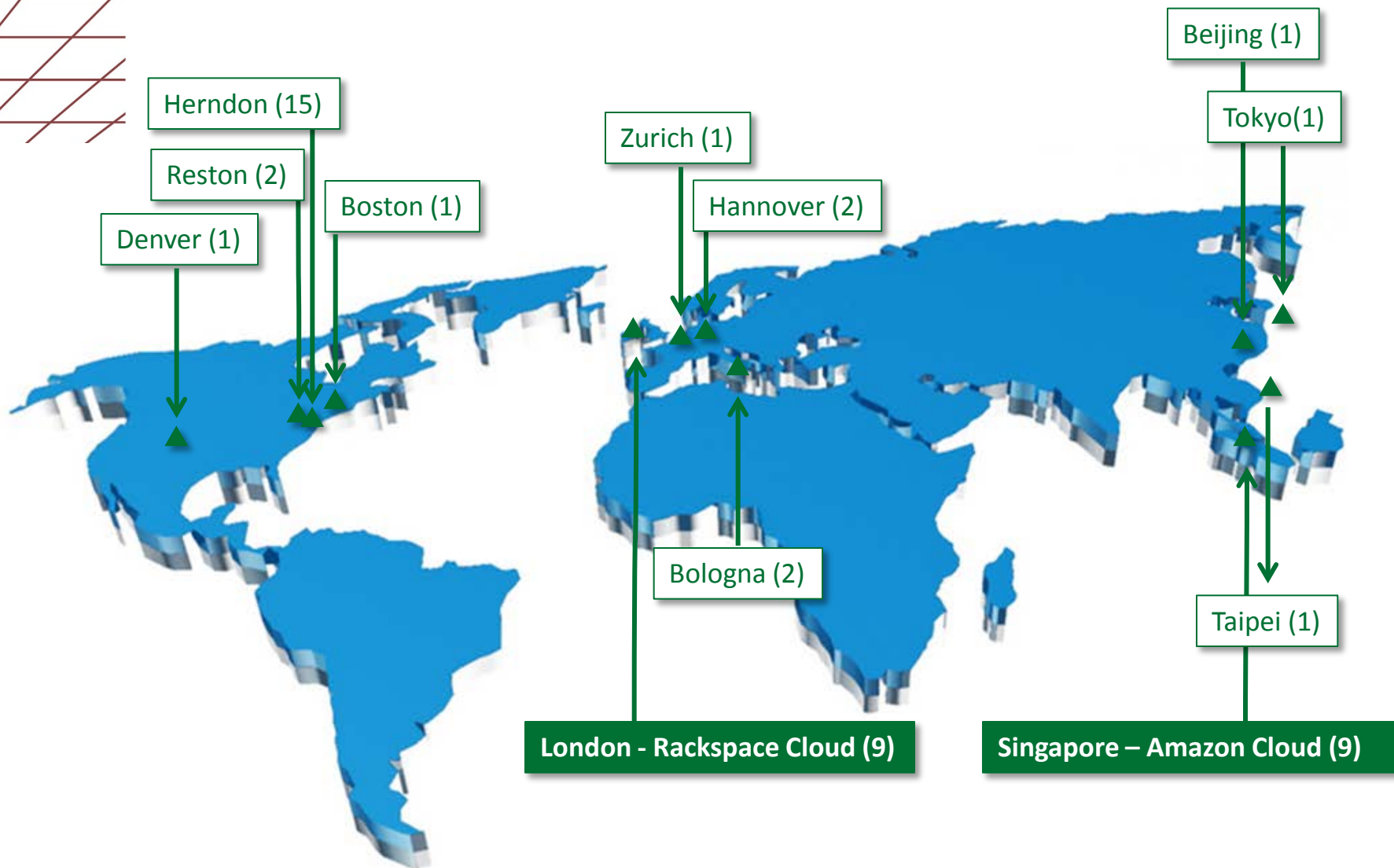


Модель управления

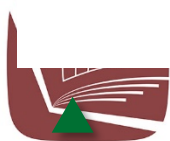
- **Международный фонд DOI (International DOI Foundation, IDF)**
 - Обеспечивает техническую и социальную инфраструктуру
 - Решение общих задач (стандарты, лицензии, пр.)
- **Регистрирующие агентства (РА)**
 - Автономны и независимы друг от друга
 - Самостоятельно определяют бизнес модель взаимодействия со своими пользователями
 - Все РА разделяют затраты на поддержание инфраструктуры DOI (IDF)
 - Все РА поддерживают стандарты системы DOI
 - РА могут иметь собственные доп. сервисы
- **Пользователи** сервисов регистрирующих агентств



Техническая инфраструктура DOI



DOI Handle Servers (46 сервера в 12 узлах распределенной системы)



Регистрирующие агентства

Цифровые библиотеки Китая



[airiti, Inc.](#)



[Crossref](#)

Базы знаний Китая



[China National Knowledge Infrastructure \(CNKI\)](#)



[DataCite](#)

Система идентификации для фильмов и телепередач



[EIDR \(Entertainment Identifier Registry\)](#)



[ISTIC \(The Institute of Scientific and Technical Information of China\)](#)

DOI для японских научных статей



[JaLC \(Japan Link Center\)](#)



[Korea Institute of Science and Technology Information \(KISTI\)](#)

Европейское регистрационное агентство для отдельных стран



[mEDRA \(Multilingual European DOI Registration Agency\)](#)



[OP \(Publications Office of the European Union\)](#)

Издательства

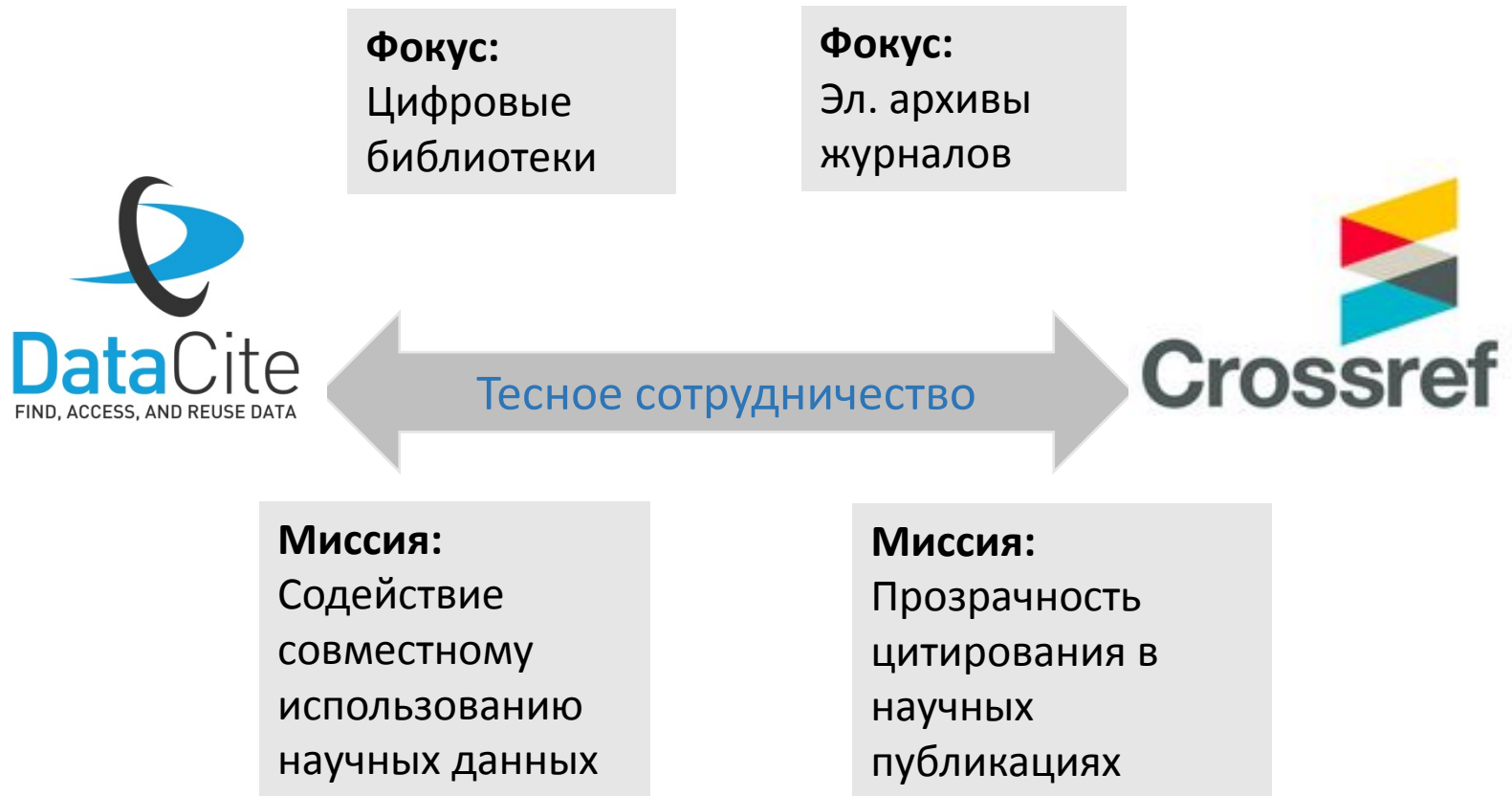
Библиотеки, университеты, научные организации

DOI для китайских журналов и диссертаций

KISTI - институт научно-технич. информации Кореи

Издательство для организаций ЕС, публикующих офиц. информацию

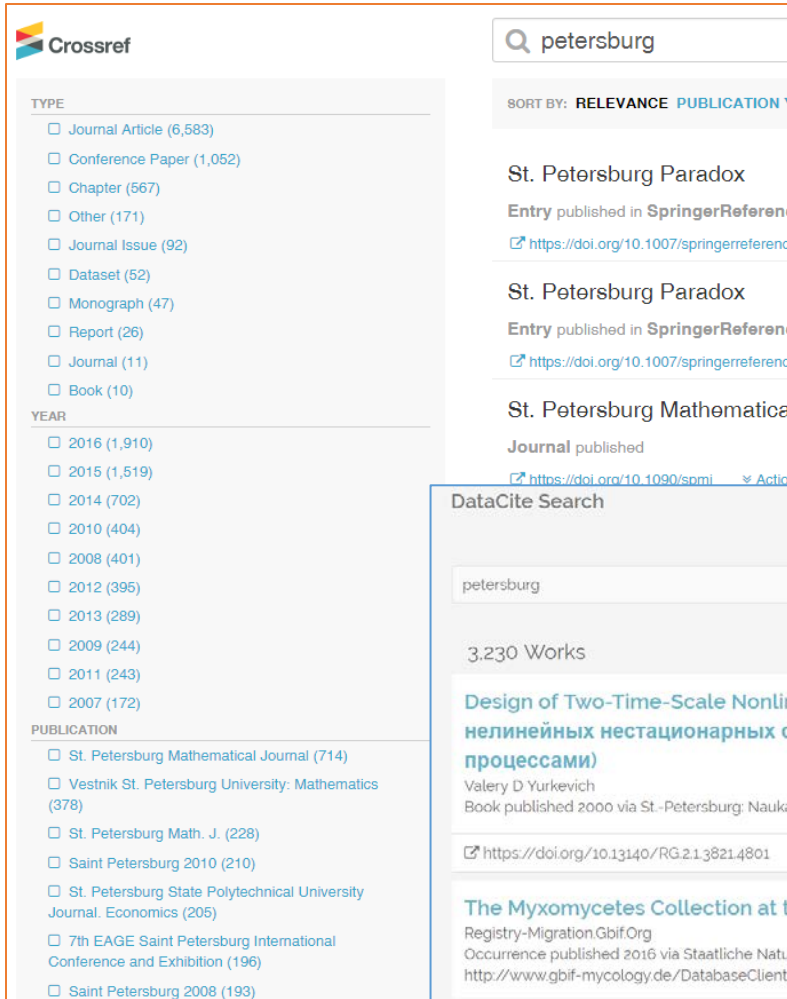
Глобальные поставщики DOI



Сервис поиска по метаданным поддерживается в каждом РА собственной инф. системой, это не связанные поисковые системы.



CrossRef и DataCite



Crossref

TYPE

- Journal Article (6,583)
- Conference Paper (1,052)
- Chapter (567)
- Other (171)
- Journal Issue (92)
- Dataset (52)
- Monograph (47)
- Report (26)
- Journal (11)
- Book (10)

YEAR

- 2016 (1,910)
- 2015 (1,519)
- 2014 (702)
- 2010 (404)
- 2008 (401)
- 2012 (395)
- 2013 (289)
- 2009 (244)
- 2011 (243)
- 2007 (172)

PUBLICATION

- St. Petersburg Mathematical Journal (714)
- Vestnik St. Petersburg University: Mathematics (378)
- St. Petersburg Math. J. (228)
- Saint Petersburg 2010 (210)
- St. Petersburg State Polytechnical University Journal. Economics (205)
- 7th EAGE Saint Petersburg International Conference and Exhibition (196)
- Saint Petersburg 2008 (193)

Q petersburg

SORT BY: RELEVANCE PUBLICATION YEAR

St. Petersburg Paradox
Entry published in SpringerReference
<https://doi.org/10.1007/springerreference>

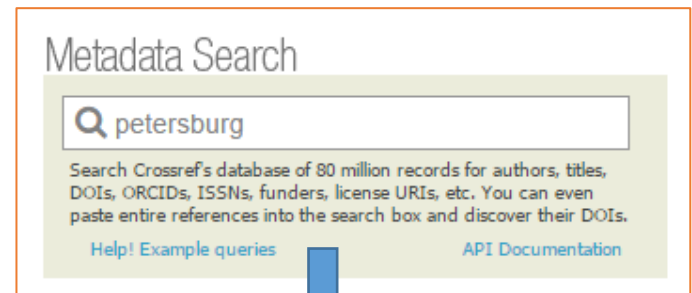
St. Petersburg Paradox
Entry published in SpringerReference
<https://doi.org/10.1007/springerreference>

St. Petersburg Mathematical
Journal published
<https://doi.org/10.1090/spmi>



DataCite
FIND, ACCESS, AND REUSE DATA

petersburg Search

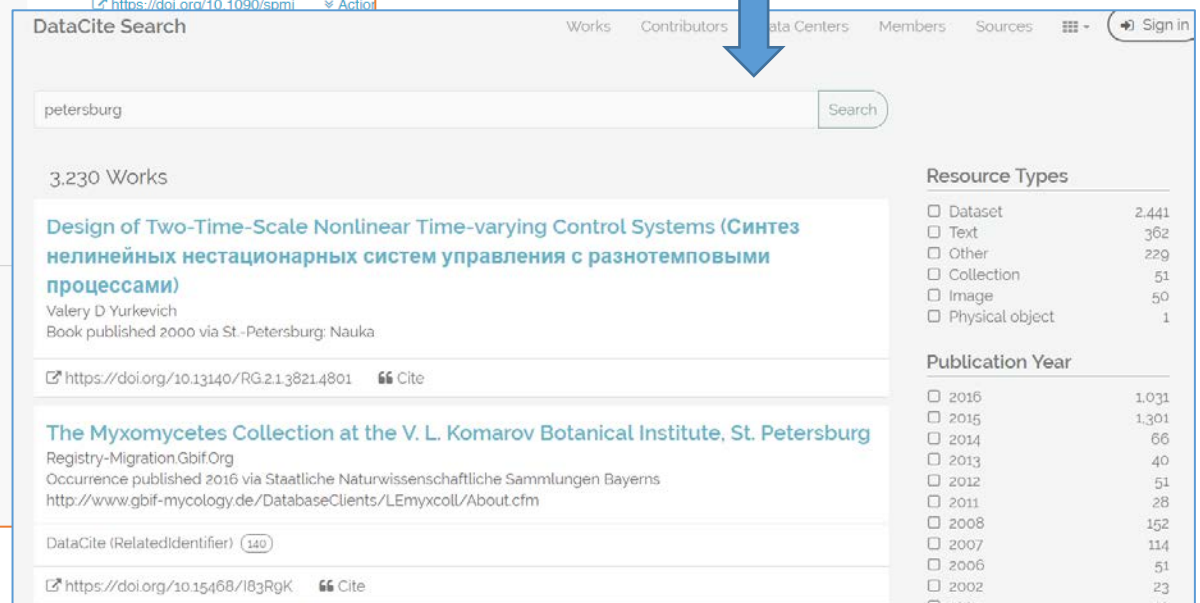


Metadata Search

Q petersburg

Search Crossref's database of 80 million records for authors, titles, DOIs, ORCID, ISSNs, funders, license URIs, etc. You can even paste entire references into the search box and discover their DOIs.

[Help! Example queries](#) [API Documentation](#)



DataCite Search

Works Contributors Data Centers Members Sources Sign in

petersburg Search

3,230 Works

Design of Two-Time-Scale Nonlinear Time-varying Control Systems (Синтез нелинейных нестационарных систем управления с разнотемповыми процессами)
Valery D Yurkevich
Book published 2000 via St.-Petersburg: Nauka
<https://doi.org/10.13140/RG.2.1.3821.4801> Cite

The Myxomycetes Collection at the V. L. Komarov Botanical Institute, St. Petersburg
Registry-Migration.Gbif.Org
Occurrence published 2016 via Staatliche Naturwissenschaftliche Sammlungen Bayerns
<http://www.gbif-mycology.de/DatabaseClients/LEmyxcol/About.cfm>
DataCite (RelatedIdentifier) 140
<https://doi.org/10.15468/183RgK> Cite

Resource Types

- Dataset 2,441
- Text 362
- Other 229
- Collection 51
- Image 50
- Physical object 1

Publication Year

- 2016 1,031
- 2015 1,301
- 2014 66
- 2013 40
- 2012 51
- 2011 28
- 2008 152
- 2007 114
- 2006 51
- 2002 23



Члены DataCite

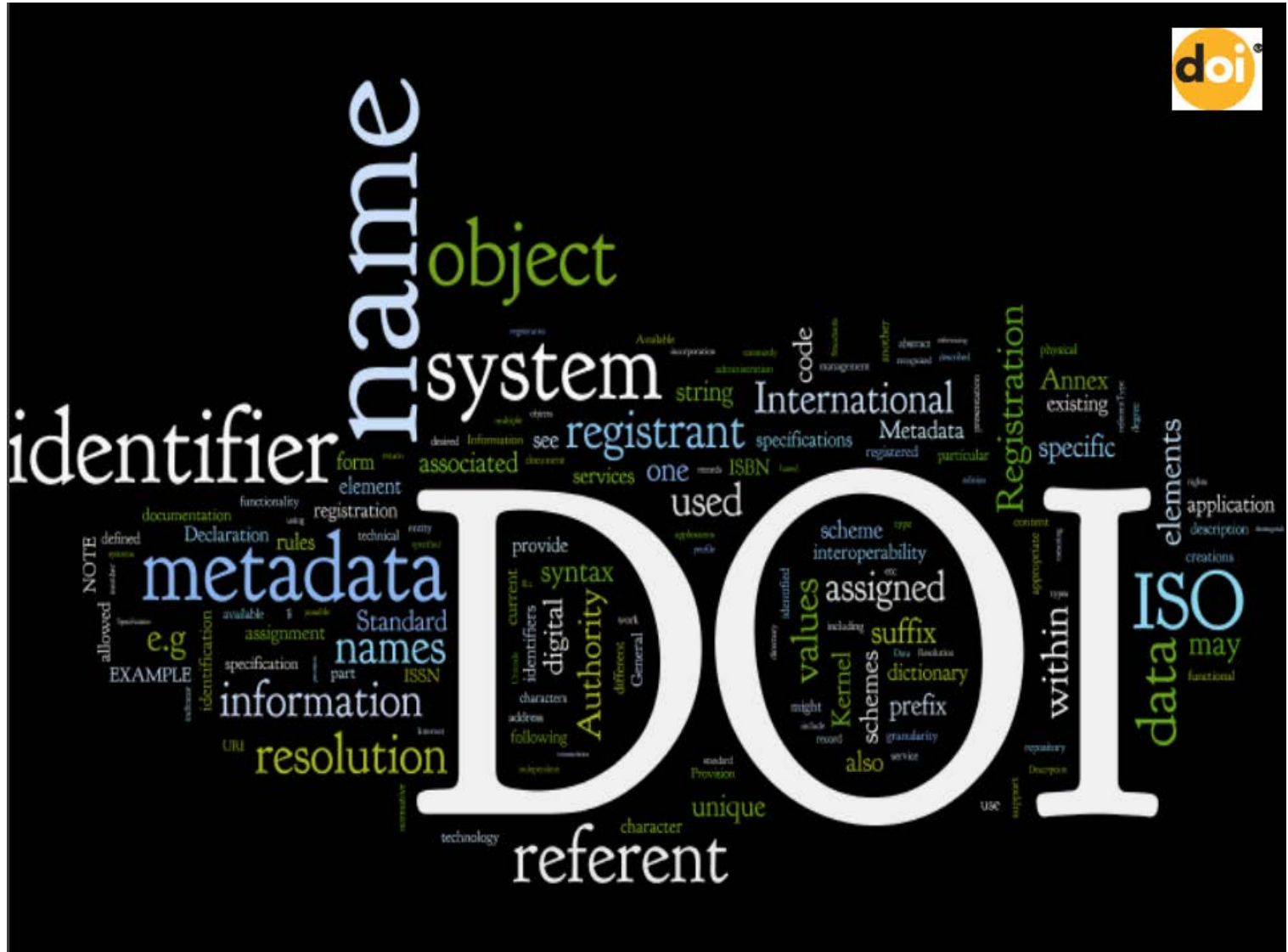


История DataCite

- 2004 - небольшой проект в Германии:
 - первый DOI **18.03.2004:**
doi:10.1594/WDCC/EH4_OPYC_SRES_A2 (DOI #1)
- 2009 - создание международной Ассоциации DataCite, Национальная научно-техническая библиотека Германии (библиотека Лейбниц университета Ганновера) обеспечивает техническую поддержку
- Цель: доступность научных данных различных видов и развитие практики их цитирования, которое:
 - дает возможность повторного использования и верификации данных
 - позволяет проверить корректность сделанных выводов на основании коллекций исходных данных
 - создает инфраструктуру, где приветствуется и признается публикация научных данных разных видов



Цифровой идентификатор объекта



Цифровой идентификатор объекта

- *Не идентификатор «цифрового объекта»*
- Объект - любая сущность (физическая, цифровая, абстрактная)
 - документы, цифровые данные, лицензии, пр.
- Цифровой идентификатор – система «активных связей» ("click on it and do something")
- Общая универсальная схема
- Изначально фокус на документах/мультимедиа, т.е. статьях, массивах данных
 - сейчас начинает использоваться в других секторах
- Не является решением, применяемым исключительно издательствами (цифровая конвергенция)
- Международный глобальный охват



Цифровой идентификатор объекта

- Постоянное имя и гиперссылка
- Обычно выражается через URL
- Определяет ресурс – не место его размещения
- Имеет связанные с ним метаданные
- Уникальная алфавитно-цифровая строка, назначенная регистрирующим агентством



Назначение DOI

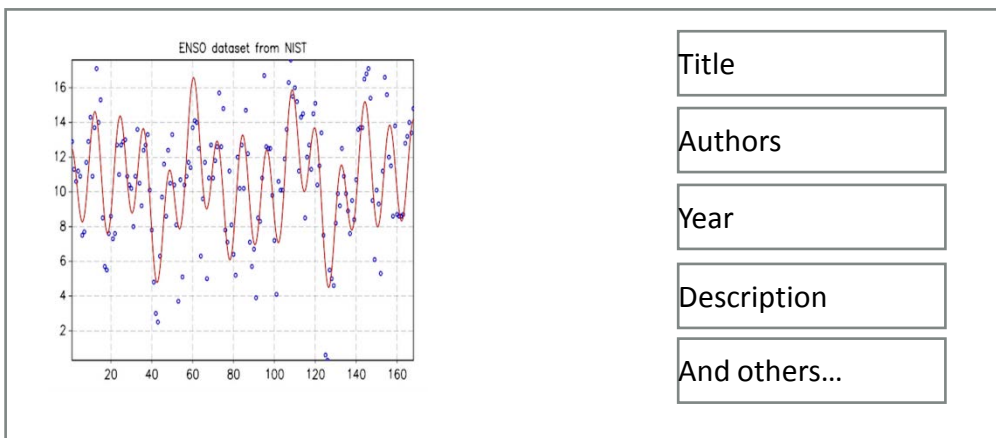
DOI обеспечивает разрешимую, постоянную и интероперабельную ссылку:

- **разрешимая** – стандартный синтаксис идентификатора и механизм отображения ссылок в сети (резольвер, Handle System)
- **постоянная** – на основе:
 - инфокоммуникационной инфраструктуры (реестр БД, поддержка проху, пр.)
 - социальной инфраструктуры (обязательства регистрирующих агентств)
- **интероперабельная** – благодаря модели данных (семантическая интероперабельность)



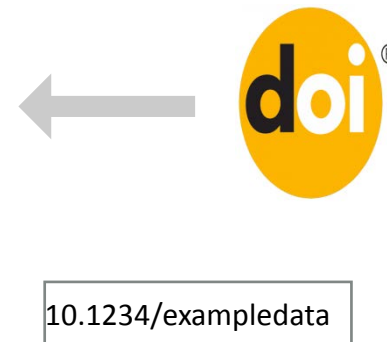
Использование DOI

1. Получаем набор данных



2. Описываем

3. Назначаем DOI



4. Используем и цитируем!

ATLAS Collaboration, "Data from Figure 7 from: Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC: $H \rightarrow \gamma\gamma$,"
<http://doi.org/10.7484/INSPIREHEP.DATA.A78C.HK44>

✓ Unique

✓ Persistent

5. Получаем преимущества

Легкость поиска	Выявление цитирований
Повторное использование	Измерение импакт показателей



Data from Figure 7 from: Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC

ATLAS Collaboration (for the collaboration)

Cite as: ATLAS Collaboration (2013) HepData, <http://doi.org/10.7484/INSPIREHEP.DATA.A78C.HK44>

Description: -2 log Likelihood for the $H \rightarrow \gamma\gamma$ channel in the $(\mu_{ggF+tH} * B/BSM, \mu_{VBF+VH} * B/BSM)$ plane for a Higgs boson mass $m_H = 125.5$ GeV.

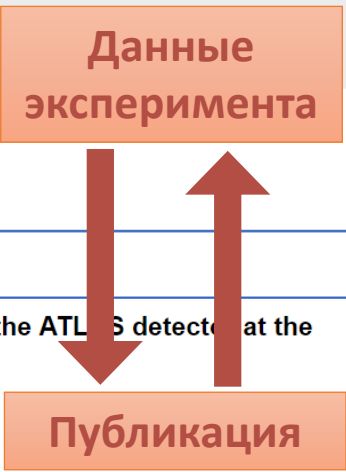
Preview not available

Note: * Temporary entry *

This dataset complements the following publication:

[Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC](#)

Record added 2013-09-11, last modified 2013-12-16



Информация | Ссылки (121) | Цитирования: (538) | Файлы | Графики | Data

Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC

ATLAS Collaboration (Georges Aad (Freiburg U.) *et al.*) [Показать все 2923 авторов](#)

Jul 4, 2013 - 32 pages

Phys.Lett. B726 (2013) 88-119

Erratum: Phys.Lett. B734 (2014) 406-406 (2013-10-07)

DOI: [10.1016/j.physletb.2013.08.010](http://dx.doi.org/10.1016/j.physletb.2013.08.010), [10.1016/j.physletb.2014.05.011](http://dx.doi.org/10.1016/j.physletb.2014.05.011)

CERN-PH-EP-2013-103

e-Print: [arXiv:1307.1427](https://arxiv.org/abs/1307.1427) [hep-ex] | PDF

Experiment: [CERN-LHC-ATLAS](#)

Abstract (Elsevier)

Measurements are presented of production properties and couplings of the recently discovered Higgs boson using the decays into boson pairs, $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ \rightarrow 4\ell$ and $H \rightarrow WW \rightarrow \ell\nu\ell$. The results are based on the complete pp collision data sample recorded by the ATLAS experiment at the CERN Large Hadron Collider at centre-of-mass energies of $\sqrt{s}=7$ TeV and $\sqrt{s}=8$ TeV, corresponding to an integrated luminosity of about 25 fb^{-1} . Evidence for Higgs boson production through vector-boson fusion is reported. Results of combined fits probing Higgs boson couplings to fermions and bosons, as well as anomalous contributions to loop-induced production and decay modes, are presented. All measurements are consistent with expectations for the Standard Model Higgs boson.

[Abstract \(arXiv\)](#)

Note: *Temporary entry*

Note: 23 pages plus author list (38 pages total), 13 figures, 10 tables, published in Phys. Lett. B 726 (2013) 88. All figures including auxiliary figures are available at <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/HIGG-2013-02/>

Keyword(s): INSPIRE: [Higgs particle: hadroproduction](#) | [Higgs particle: coupling](#) | [vector boson: fusion](#) | [p.p. scattering](#) | [CERN LHC Coll](#) | [ATLAS](#) | [Higgs particle: decay modes](#) | [vector boson: pair production](#) | [vector boson: leptonic decay](#) | [mass spectrum: two-photon](#) | [mass spectrum: \(4lepton\)](#) | [dilepton: mass spectrum](#) | [transverse energy: missing-energy](#) | [Higgs particle: mass](#) | [experimental results](#) | [7000..8000 GeV:cms](#)



OpenScience и DOI

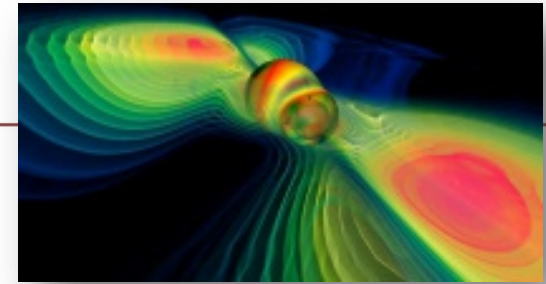
<http://doi.org/10.7935/K5MW2F23>

Уникальные данные о коллизии двух «черных дыр»

- собраны на оборудовании LIGO
- цитируются через DataCite DOI
- **и доступны всем**

LIGO Open Science Center
California Institute of Technology и
Massachusetts Institute of
Technology

предоставил в открытом доступе все данные, включая технические отчеты, графики, параметры калибровки и аудио файлы



LIGO LIGO Open Science Center
LIGO is operated by California Institute of Technology and Massachusetts Institute of Technology and supported by the U.S. National Science Foundation.

Getting Started

Tutorials
Data & Catalogs
Timelines
My Sources
Software
GPS – UTC
About LIGO
Student Projects
Acknowledgement

Data release for event GW150914

This page has been prepared by the LIGO Scientific Collaboration (LSC) and the Virgo Collaboration to inform the broader community about a confirmed astrophysical event observed by the gravitational-wave detectors, and to make the data around that time available for others to analyze. There is also a technical details page about the data linked below, and feel free to contact us. This dataset has the Digital Object Identifier (doi) <http://dx.doi.org/10.7935/K5MW2F23>

Summary of Observation

The event occurred at GPS time 1126259462.39 s = September 14 2015, 09:50:45.39 UTC. The false alarm rate is estimated to be less than 1 event per 203,000 years, equivalent to a significance of 5.1 sigma. The event was detected in data from the LIGO Hanford and LIGO Livingston observatories.

- There are Science Summaries, covering the information below in ordinary language.
- There is a one page factsheet about GW150914, summarizing the event.

How to Use this Page

- Click on the section headings below to show available data files.
 - (click to Open/Close all sections)
- There are lots of data files available in the sections below, look for the word **DATA**.
- Click on each thumbnail image for larger image.
- See the papers linked below for full information, references, and meaning.
- Many of the data files linked below have heterogeneous formatting; if you have any questions, please contact us.

The GW150914 detection paper:
Observation of Gravitational Waves from a Binary Black Hole Merger
For full details see LIGO DCC, arXiv, or Phys. Rev. Letters
This paper and all the companion papers can also be found at papers.ligo.org

Estimated source parameters

Quantity	Value	Upper/Lower error estimate	Unit
Primary black hole mass	36	+5 -4	M sun
Secondary black hole mass	29	+4 -4	M sun
Final black hole mass	62	+4 -4	M sun
Final black hole spin	0.67	+0.05 -0.07	
Luminosity distance	410	+160 -180	Mpc
Source redshift, z	0.09	+0.03 -0.04	
Energy radiated	3	+0.5 -0.5	M sun

TABLE I. Estimated source parameters for GW150914. We report the median value as well as the range of the 90% credible interval. Masses are measured in the source frame; to convert masses to detector frame, multiply by $(1+z)$. The source redshift assumes standard cosmology.

Hanford, Washington (H1) Livingston, Louisiana (L1)

click for DATA click for DATA (L1 only)



Структура наименования (имени) DOI

- Уникальная строка из букв и цифр, состоящая из двух частей: префикс и суффикс, разделенные “/”.

10.1234/data567

- **10.1000** — префикс, или *идентификатор издателя*, составленный из признака идентификатора (10) и строки, указывающей на издателя (1234)
- **data567** — суффикс, *идентификатор объекта*, указывающий на конкретный объект
- Префиксы издателей распределяются регистрационным агентством (DataCite)
- Суффикс формируется владельцем хранилища данных или издателем
- Идентификатор цифрового объекта может объединить существующие идентификаторы, такие как ISBN или SICR



Префикс

DOI неразрывно связаны с пространством имен своего префикса (хранилищем объектов).

Рекомендация - использовать отдельный префикс, когда:

- Разные создатели (издатели) объектов
- Объекты размещена на разных платформах/серверах
- Объекты ориентированы на различных пользователей



Суффикс

- Определяется организацией, регистрирующей объект
- **Рекомендация** - «слепая» строка (имена могут меняться)

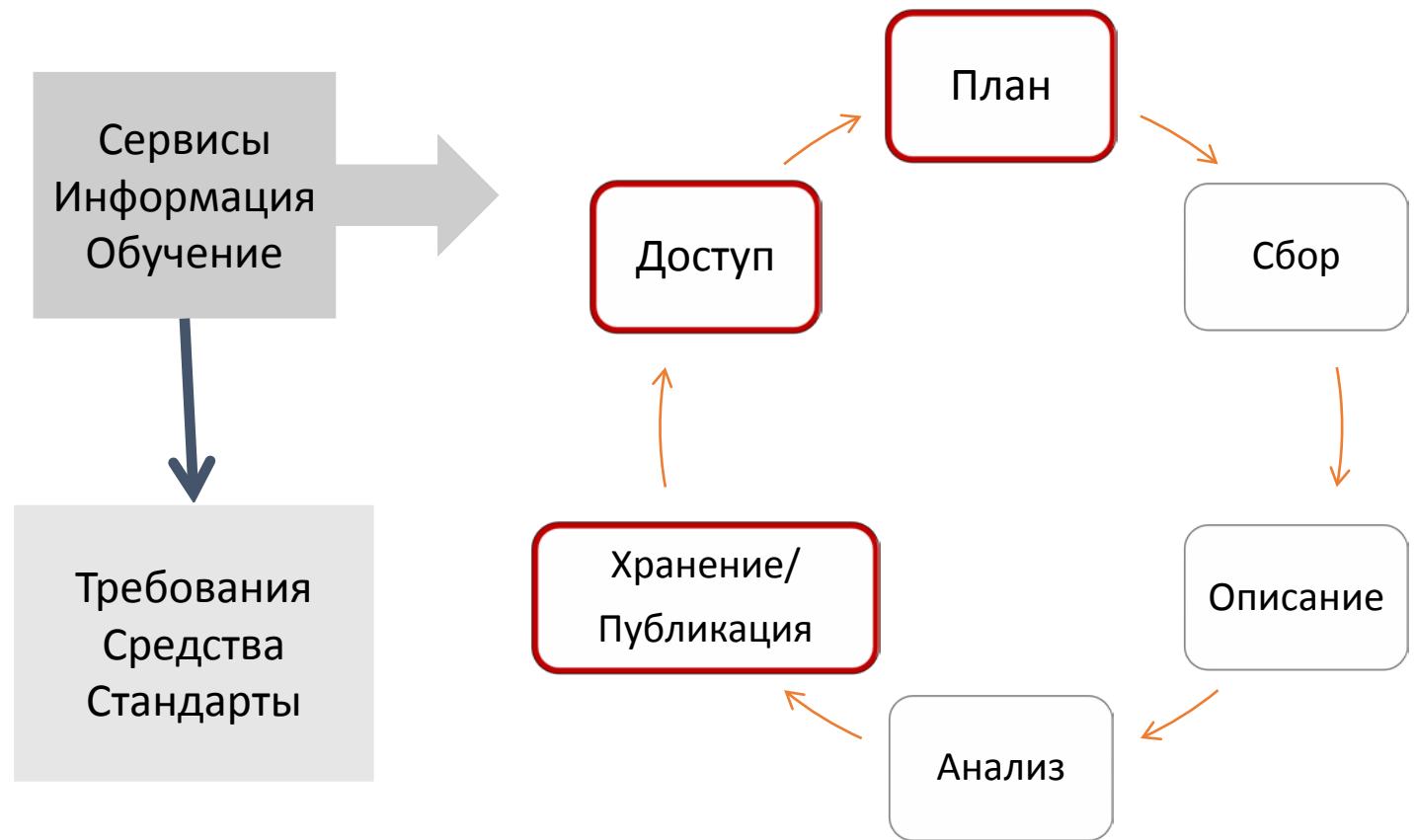
Prefix Suffix

doi:10.2314/529827735

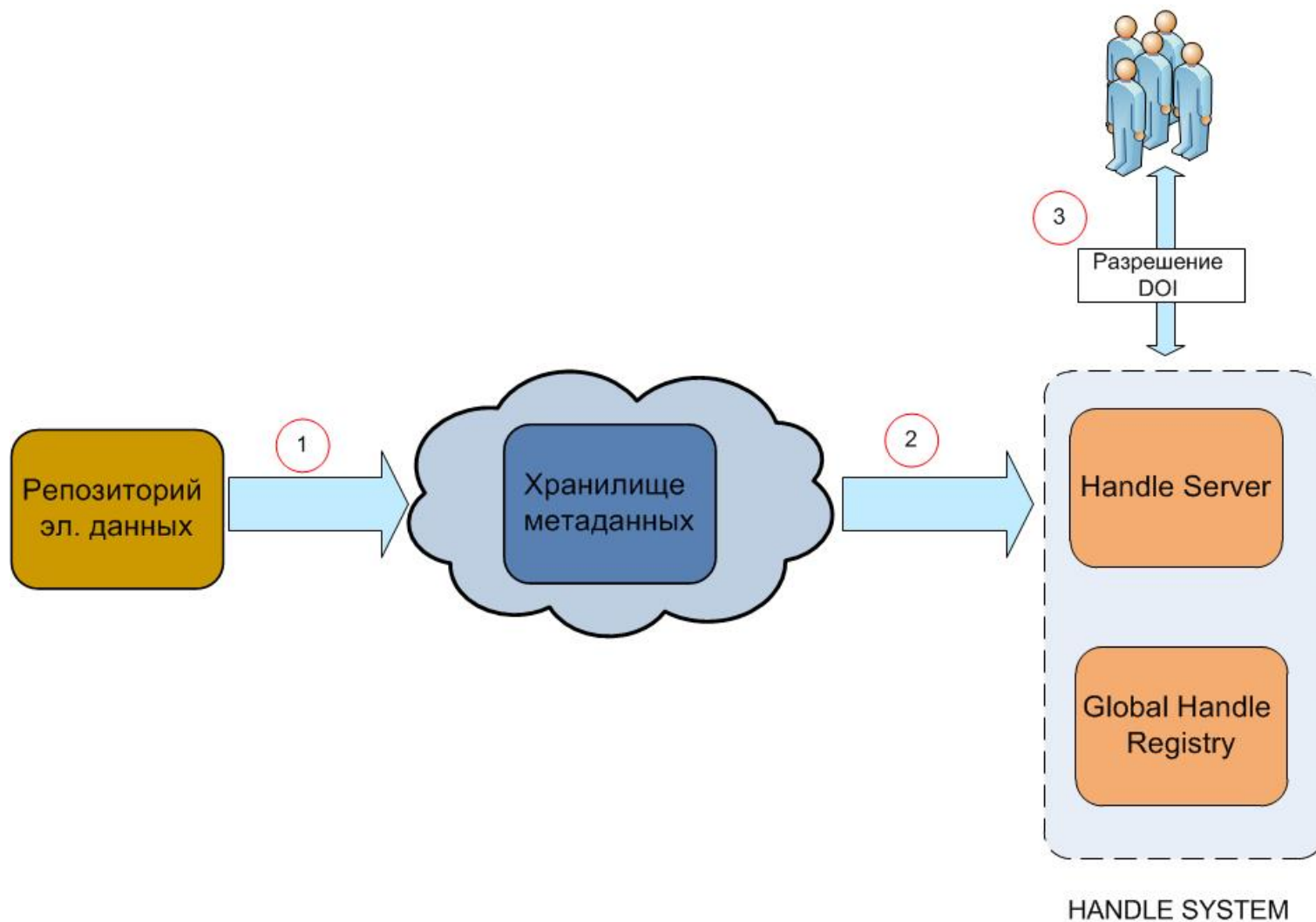


Управление научными данными

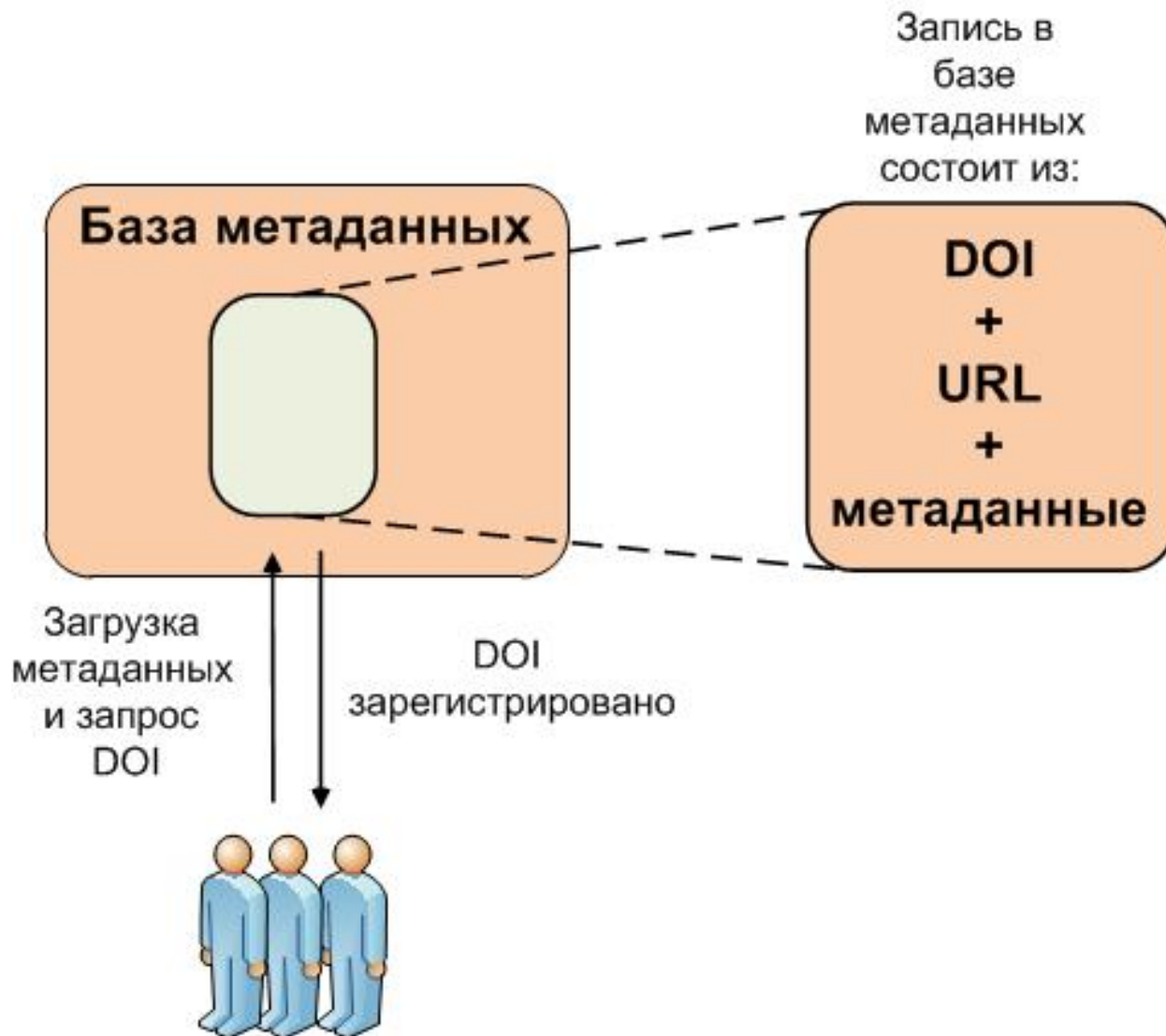
– новая роль библиотек



Инфраструктура DataCite



Процесс регистрации



Открытый доступ и лицензирование

- Открытый доступ – главная цель OpenScience
- Но если необходимо, можно ограничивать доступ. Landing page («посадочная» страница) должна информировать об условиях доступа
- Метаданные и landing page должны быть доступны всегда



Что это дает?

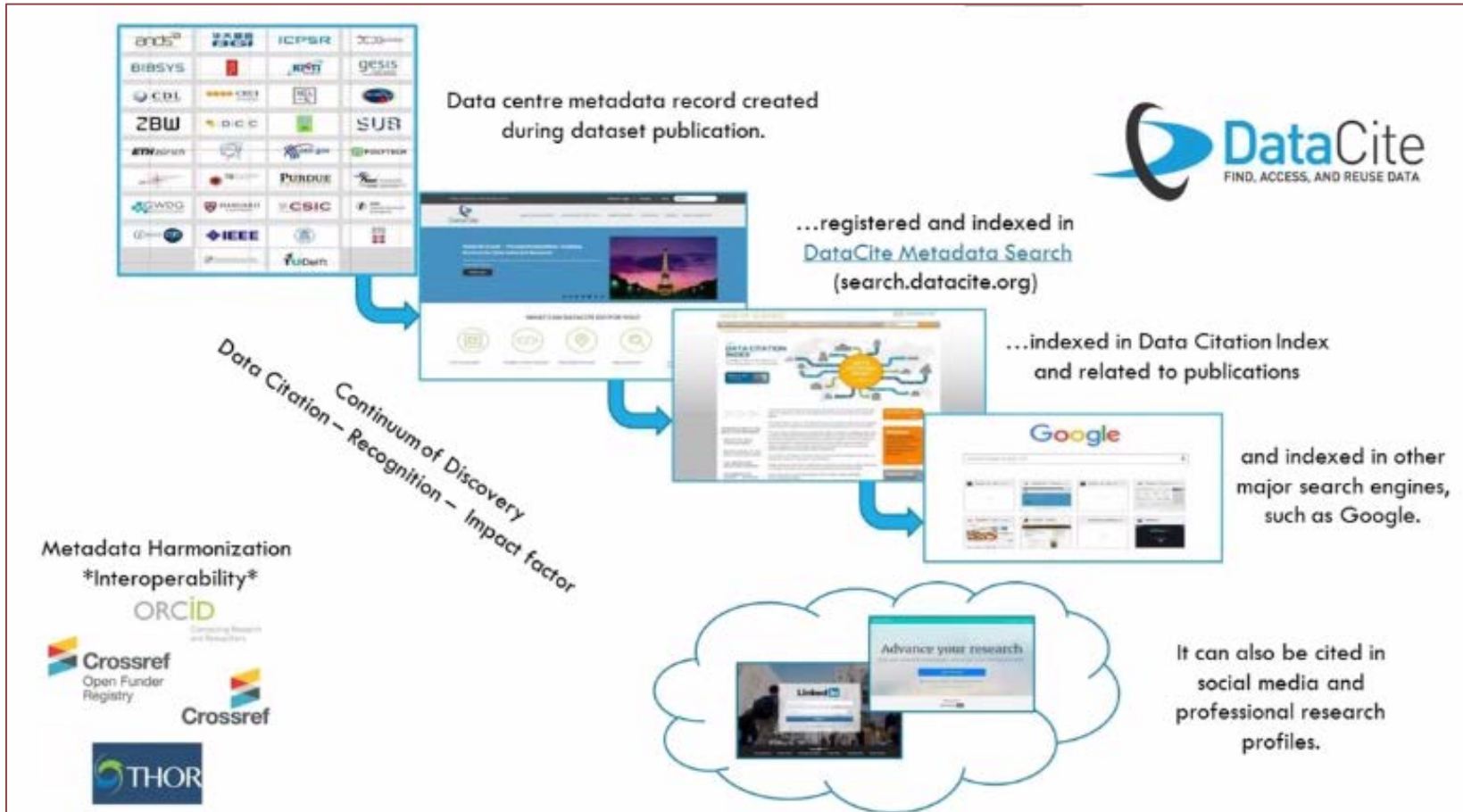
- Повышение статуса ресурса
- Подтверждение достоверности исследования
- Сохранение и распространение неопубликованных материалов, артефактов и пр.
- Облегчение процесса цитирования
- DOI - основа современных научных коммуникаций
- Содействие поиску и цитированию ресурса, в том числе в зарубежных работах

Новая роль и новая услуга библиотеки в цифровой среде



Как это происходит

Публикация +DOI -> Размещение в базе метаданных DataCite -> Передача в другие индексы -> Цитирование, импакт-фактор



Национальный центр идентификации научных данных



Санкт-Петербургский политехнический университет Петра Великого
Информационно-библиотечный комплекс

ПОЛУЧИТЬ ПАРОЛЬ

ВОЙТИ

ОБ ИБК

ЧИТАТЕЛЯМ

ПРЕПОДАВАТЕЛЯМ

АВТОРАМ

ПОИСК ИНФОРМАЦИИ

УСЛУГИ

ОБУЧЕНИЕ

🏠 > Об ИБК > Структура > Национальный центр идентификации цифровых научных данных

Национальный центр идентификации цифровых научных данных

Администрация:

Директор: Домрачева Юлия Александровна

Тел/факс: +7 (812) 552-76-54

e-mail: julia@unilib.spbstu.ru

Местоположение:

ул. Политехническая, д. 29, главный учебный корпус, ауд.138

Режим работы:

понедельник - пятница: с 9.00 до 18.00

суббота, воскресенье: выходные дни

Сведения о подразделении:

Национальный центр идентификации научных данных (НЦИ) был организован для повышения позиции Университета и формирования имиджа крупного научного и образовательного центра за счет выявления и идентификации цифровых научных данных СПбПУ, применимых для их повторного использования и цитирования в России и мире.

НЦИ участвует в деятельности Ассоциации Datacite. СПбПУ является членом Datacite с 2016 года.

Национальный центр идентификации научных данных осуществляет следующие виды деятельности:

- разработка научных основ и технологий идентификации научных данных, представленных в цифровом виде, внедрение их в Университете;
- регистрация хранилищ научных данных Университета. Выделение префикса для идентификаторов DOI для научных данных, размещаемых в хранилище;
- ведение реестра хранилищ научных данных Университета на основе учета заявок, поступающих от подразделений Университета;
- предоставление услуг по описанию, хранению и идентификации научных данных для организаций, расположенных на территории РФ, и для физических лиц;
- организация и проведение подготовки специалистов по направлениям деятельности Центра, включая создание программ дополнительного и дистанционного

<http://library.spbstu.ru>



