

Дата-экспедиции: продвинутый формат обучения работе с данными

Ирина Радченко (Университет ИТМО)



Aletheia
Business
 РИТ 2019

Профессиональная
конференция
по применению психологии
в управлении и бизнесе





@iradche

<http://www.slideshare.net/iradche>

<https://www.facebook.com/iradche>

iradche@gmail.com

Ирина Радченко

Кандидат технических наук,

Доцент университета ИТМО (Санкт-Петербург)

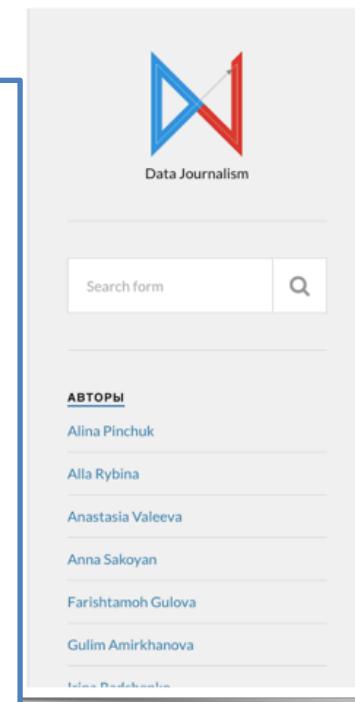
Доцент Высшей школы экономики (Москва)

Сооснователь проекта DataDrivenJournalism.Ru

Что такое data-экспедиции?

Экспедиции данных

Экспедиция данных, или data-экспедиция, – это неформальное общедоступное образовательное онлайн-мероприятие, строящееся на принципах проектного обучения и взаимопомощи. Под “проектным” имеется в виду, что итогом участия в идеале должно стать создание индивидуального или коллективного проекта. Каждая экспедиция сама по себе строится как процесс работы над определенным проектом.



3 Источник: <http://www.datadrivenjournalism.ru/practicum/data-expeditions/>

Дата-экспедиции: история

The screenshot shows the homepage of the School of Data website. At the top, there's a navigation bar with links for Learn, Teach, Fellowship Programme, Events, Blog, About, and a Twitter icon. Below the navigation, a large blue header section features a "Welcome!" message and a "Join School of Data" call-to-action. The "Welcome!" section includes a brief description of the organization's mission to empower civil society organizations, journalists, and citizens with data skills. It also has a button to "Get started with School of Data today!". The "Join School of Data" section encourages users to subscribe by providing fields for "Your name" and "Your email", with a "Subscribe" button. Below this, there's a video thumbnail for the "School of Data Summercamp 2015" featuring a man speaking. To the right of the video, there's a section titled "Meet our 2015 Fellows!" which describes the fellowship program running from now until December, mentioning 7 data trainers from around the world, and links to learn more about them and the 2014 Annual Report. A "Get Involved" link is also present.

Источник: <http://schoolofdata.org/>

Наши дата-экспедиции

The screenshot shows the homepage of the Data Journalism website. At the top, there's a navigation bar with links: ПРАКТИКУМ, ПОЛЕЗНЫЕ МАТЕРИАЛЫ, О ПРОЕКТЕ, ОБ АВТОРАХ, ПРИСОЕДИНЯЙТЕСЬ, ENGLISH, and МЫ НА ГИТХАБЕ. Below the navigation, there are two main news items:

- Мероприятия, экспедиции данных**
Дата-экспедиция №5: введение
19.04.2016 — 0 Comments

- Мероприятия, экспедиции данных**
Дата-экспедиция №5: анонс
15.04.2016 — 0 Comments


On the right side of the page, there's a sidebar with a search form and a list of authors:

Авторы

- Alina Pinchuk
- Alla Rybina
- Anastasia Valeeva
- Anna Sakoyan
- Farishtamoh Gulova
- Gulim Amirkhanova
- Irina Radchenko
- Kristina Kashtanova
- Natalia Karbasova
- Philipp Kats
- Yulia Mityaeva

A large QR code is located in the bottom right corner of the page.

The screenshot shows a list of authors on the right side of the Data Journalism website. The title "Авторы" is at the top, followed by a horizontal line and a list of names:

- Alina Pinchuk
- Alla Rybina
- Anastasia Valeeva
- Anna Sakoyan
- Farishtamoh Gulova
- Gulim Amirkhanova
- Irina Radchenko
- Kristina Kashtanova
- Natalia Karbasova
- Philipp Kats
- Yulia Mityaeva

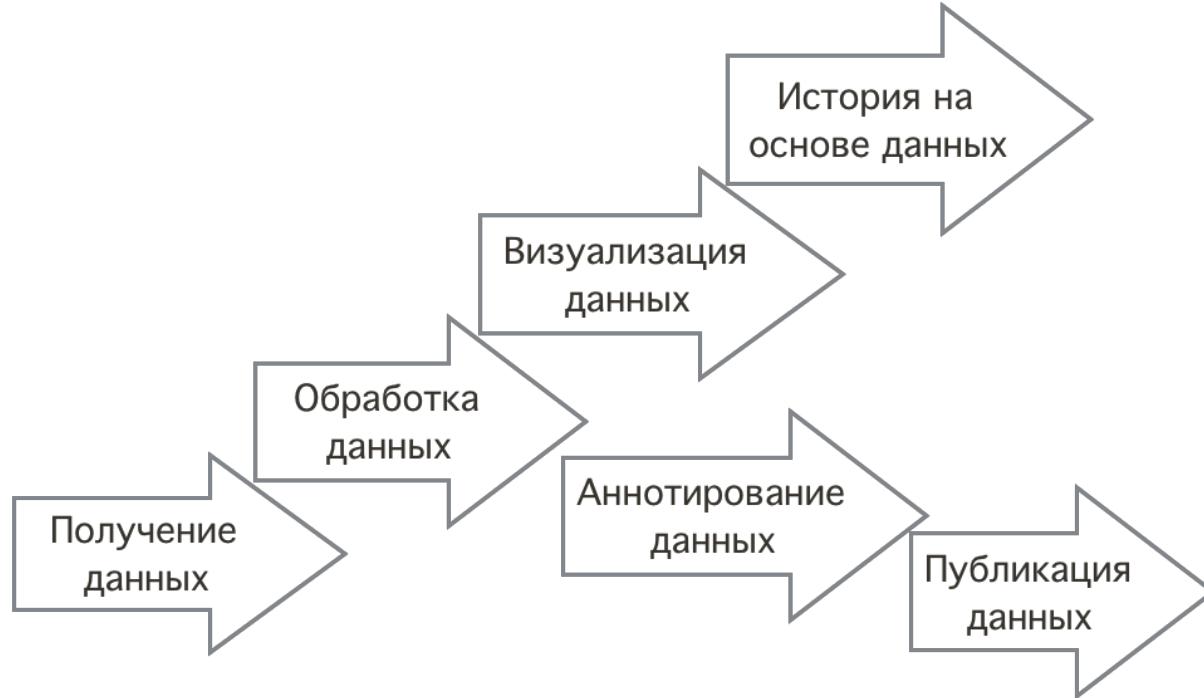


Источник: <http://www.datadrivenjournalism.ru/>

Кто участвует в дата-экспедиции?

- Журналисты
- Аналитики
- Активисты
- Чиновники
- Научные сотрудники
- Студенты
- Все, кто хочет научиться работать с данными

Дата-экспедиции: этапы работы с данными



7 На основе источника: <http://iradche.ru/2013/06/18/журналистика-данных>

Дата-экспедиции: разнообразие задач

Получение данных

- Задачи на поиск открытых данных.
- Постановка задачи.
- Умение работать с платформами данных.

Визуализация данных

- Создание статической визуализации.
- Создание динамической визуализации.

...

Обработка данных

- Очистка данных.
- Обогащение данных.

Распространение данных

- Аннотирование данных.
- Публикация данных.
- Работа с открытыми лицензиями.

Как проводятся data-экспедиции?

- Организация коллабораций
- Работа в малых группах (2-3 человека)
- Работа с онлайн-инструментами
- Использование механизма краудсорсинга
- Прозрачность оценивания результата

Дата-экспедиции: учет и контроль

Файл Изменить Вид Вставка Формат Данные Инструменты Дополнения Справка Последнее изменение: 17 июля 2017 г.

Настройки Доступа

Ф.И.О.	Номер группы	Научные интересы и тема исследования	Ссылка на черновик статьи (в гугл-документе)	Ссылка на журнал дата-экспедиции (
	5	история и социология чтения Русская литература как культурный экспорт	https://docs.google.com/document/d/1ckUQzNfB8cJdkWpQVQMaOl-mZIVONxsUbndS/edit	https://docs.google.com/document/d/1ZLGBMHit77mbz2C0lxNo/edit?usp=sharing
	4	Социология города, социология науки и техники	https://docs.google.com/document/d/18bg-jncndp9VD31mc5FnOYvEKip8HMevw382zqtMU/edit?usp=sharing	https://docs.google.com/document/d/1Ye9UBDY1G7L3XPW_HE/edit?usp=sharing
	2			https://docs.google.com/document/d/1F9PODbc-xoQ7vSHKWMMe/edit?usp=sharing
	2	Социология питания, доверие, солидарность		https://docs.google.com/document/d/1F9PODbc-xoQ7vSHKWMMe/edit?usp=sharing
	5	Научные интересы: социология права и правоприменения	https://docs.google.com/document/d/1ckUQzNfB8cJdkWpQVQMaOl-mZIVONxsUbndS/edit	https://docs.google.com/document/d/1ZLGBMHit77mbz2C0lxNo/edit?usp=sharing
	3			https://docs.google.com/document/d/1A01WR1lbRh0OWhWxp8Y_A/edit
	4		https://docs.google.com/document/d/18bg-jncndp9VD31mc5FnOYvEKip8HMevw382zqtMU/edit?usp=sharing	https://docs.google.com/document/d/1Ye9UBDY1G7L3XPW_HE/edit?usp=sharing
	3			https://docs.google.com/document/d/1A01WR1lbRh0OWhWxp8Y_A/edit
	6	Я интересуюсь анализом некоммерческих проектов развивающих городскую среду. Собираюсь использовать api vk, чтобы собрать данные по подпунктам лайбников этих проектов. Построить сеть проектов по пересечению аудиторий. При этом использовать открытые данные чтобы понять их статус (нко, движение, просто неофициальное сообщество или статус инкента). И посмотреть как связан статус, положение в сети и кластеризация сети. Я буду рад идеям как использовать в проекте другие открытые городские данные. Также буду рад если кому-то будет интересно поучаствовать в проекте.	https://docs.google.com/document/d/1KMWcmYq7ESaxM1Rn1AYQUnk4Gv3h4dAhuwZDQd92u3mhs	https://docs.google.com/document/d/1E9PDRDwv...

Дата-экспедиции: оценивание

Критерий	0	0,25	0,5	0,75	1
Дневник участия в ДЭ (Журнал эксперимента)	<ul style="list-style-type: none">· Дневник ДЭ отсутствует или пуст.	<ul style="list-style-type: none">· По дневнику невозможно воспроизвести последовательность действий исследователя· Не определены некоторые методы очистки, обработки и анализа данных· Последовательность проектных решений определена с пропусками, вносящими путаницу или двусмыслинности· Некоторые проектные решения не обоснованы или обоснования некорректны	<ul style="list-style-type: none">· По дневнику воспроизводится последовательность действий исследователя· Методы очистки, обработки и анализа данных определены частично· Последовательность проектных решений определена с пропусками, вносящими путаницу или двусмыслинности· Некоторые проектные решения не обоснованы или обоснования слабы	<ul style="list-style-type: none">· По дневнику воспроизводится последовательность действий исследователя· Методы очистки, обработки и анализа данных определены полностью· Последовательность проектных решений определена полностью или с незначительными пропусками, вносящими двусмыслинности· Проектные решения полностью обоснованы, но некоторые обоснования слабы	<ul style="list-style-type: none">· По дневнику воспроизводится последовательность действий исследователя· Методы очистки, обработки и анализа данных определены полностью· Последовательность проектных решений полностью задокументирована· Проектные решения полностью обоснованы
Программные средства обработки данных	<ul style="list-style-type: none">· Не использованы средства обработки данных	<ul style="list-style-type: none">· Работа с данными не автоматизирована и проделана вручную	<ul style="list-style-type: none">· Данные частично обработаны с использованием средств обработки данных, не требующими написания программного кода или собственных скриптов	<ul style="list-style-type: none">· В полном объёме и обоснованно использованы средства обработки данных, не требующие написания программного кода	<ul style="list-style-type: none">· В полном объёме использованы средства обработки данных, не требующие написания программного кода.· Дополнительно или вместо указанных программных средств обоснованно использован собственный программный код.
Средства анализа данных	<ul style="list-style-type: none">· Не использованы средства анализа данных	<ul style="list-style-type: none">· Анализ данных произведён с существенными ошибками· Анализ данных не автоматизирован, аналитика проведена вручную или со слабой поддержкой средствами анализа данных	<ul style="list-style-type: none">· Анализ данных произведен без ошибок или с несущественными ошибками· Анализ данных не автоматизирован или частично автоматизирован, аналитика проведена вручную или со слабой поддержкой средствами анализа данных	<ul style="list-style-type: none">· Анализ данных произведен с несущественными ошибками· Анализ данных полностью автоматизирован	<ul style="list-style-type: none">· Анализ данных произведен без ошибок· Анализ данных полностью автоматизирован
Очистка данных		<ul style="list-style-type: none">· Данные очищены с существенными нарушениями, приводящими к искажению результатов· Не все наблюдаемые	<ul style="list-style-type: none">· Данные очищены с существенными нарушениями, приводящими к искажению результатов· Все наблюдаемые единицы разделены по таблицам	<ul style="list-style-type: none">· Данные очищены с несущественными нарушениями, приводящими к искажению результатов· Наблюдаемые единицы	



https://docs.google.com/spreadsheets/d/e/2PACX-1vTLcZV4N8MRtQDf4gwNRFMwgGvHZTIO5UgXg6X_nNAT4qZcFTE0akKKcnY_Dqoxp5p1fFk3_GV3IE8t/pubhtml

Дата-экспедиции: инструменты

- Гугл-группа для рассылки сообщений
 - Гугл-таблица для ведения успеваемости
 - Гугл-документ для организации проекта
 - Телеграм-группа для обмена полезными ссылками
 - Инструменты обработки данных
 - Инструменты визуализации данных
 - GitHub для выкладывания материалов



infogram



+tableau public



Примеры дата-экспедиций.

Казахстан — 2014

- В декабре 2014 года в Караганде, Казахстан
- Количество участников — 20 человек
- Сроки проведения — 2 недели
- Количество очных занятий — 6
- Количество контрольных оцениваний — 5
- Количество завершенных проектов — 16
- Две статьи были опубликованы на ДДЖ

Примеры дата-экспедиций.

Казахстан – 2014

Дата-Экспедиция №4 (02-13.12.2014) в КарГТУ

Протокол дата-экспедиции в КарГТУ

Время: 04-13 декабря 2014 г.

Место: Казахстан, Караганда, КарГТУ

Цель

Поиск, обработка и представление открытых данных по заданной тематике.

Задачи

1. Научиться работать с открытыми данными:
 - а. Научиться искать открытые данные.
 - б. Научиться обрабатывать найденные данные.
 - с. Научиться представлять данные в виде визуализаций.
2. Научиться использовать инструментарий для совместной работы в онлайне.

Планируемые результаты

1. Расширение аудитории ДДЖ.
2. Получение студентами компетенций по работе с открытыми данными.
3. Обучение студентов работе с открытыми данными.

Подготовка к проведению дата-экспедиции

- Организация виртуальной площадки в Etherpad: <https://etherpad.wikimedia.org/p/DataExpedition>
- Разработка и заполнение двух анкет (пре-опросы).
 - Первая анкета ориентирована на международное участие: <http://goo.gl/forms/Bf2XyzVGfW>
 - Вторая анкета предназначена для студентов КарГТУ (г. Караганда, Казахстан): <http://goo.gl/forms/JI923a5ZpR>
- Разработка и заполнение анкеты (пост-опрос): <http://goo.gl/forms/0emYPbnmhA>.
- Заведение общей сводной таблицы результативности: <https://docs.google.com/spreadsheets/d/1w23ulBvAxIMCKs8DqpyYFH1cnElO5vHqUzeMo83YGto/edit#gid=0>
- Организация гуглогруппы (для рассылок и предоставления актуальной информации всем участникам дата-экспедиции): <https://groups.google.com/forum/#forum/dataexpedition4>



Источник: https://docs.google.com/document/d/e/2PACX-1vT4s6itZSnprkQ5nK4so34BJyRdmFIDFUjenoU6xGQ7t5NUW80PswcF0ZQdxV9Wzi_7ayCeqbkbtOpy/pub

Примеры дата-экспедиций.

Казахстан – 2014



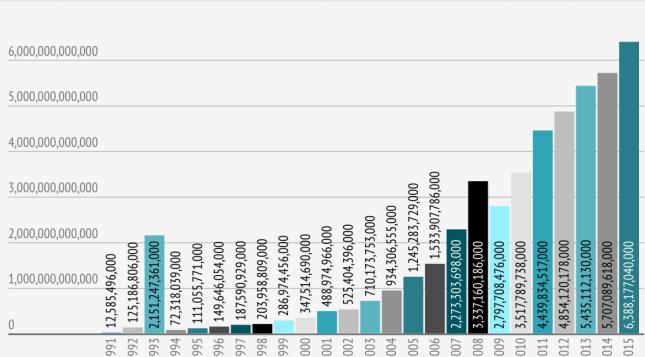
Источник: http://www.datadrivenjournalism.ru/2015/01/14/de4_project_industry_kz/

Примеры дата-экспедиций.

Казахстан – 2014

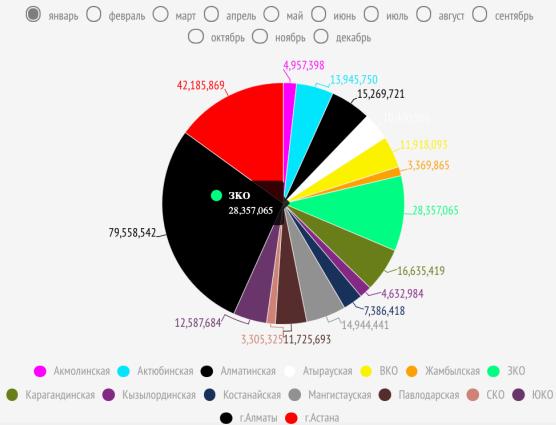
Поступления в бюджет Республики Казахстан с 1991 по 2015 годы

Тема доходной части бюджета меня особенно сильно увлекла. Приятно все-таки считать деньги, пусть даже и свои. Продолжая копаться в данном вопросе и вороче цифры, обозначающие деньги, я выяснил, что кроме понятия "доходная часть бюджета", есть термин "поступления в бюджет", а это более широкое понятие, включающее в себя трансферты, продажу финансовых активов, погашения кредитов, ранее выданных государством. Поэтому далее речь идет именно о поступлениях в бюджет. Подняв все законодательные акты Казахстана с 1991 года, касающиеся утверждения республиканских бюджетов, и был поражен. Я, конечно, понимал, что бюджет сильно вырос, но не представлял себе, насколько сильно. Для удобства привожу ниже следующую таблицу. В этой таблице цифры с 1991 по 1993 год приведены в советских рублях.



По графику мы видим что наибольшая средняя заработная плата в Атырауской и Мангистауской областях и в городах Алматы и Астана, потому что в Атырауской и Мангистауской областях больше всего развита нефтедобывающая промышленность, так как эти 2 области находятся около Каспийского моря. Следующее место занимает Алматы, потому что это крупнейший город Казахстана, известный как "Южная Столица". По данным текущего учета, город насчитывает 1,6 млн. человек населения и является туристическим центром Казахстана, в то время как столица Казахстана - Астана.

Поступление налогов и платежей в государственный бюджет за 2012-2013 год по Республике Казахстан



Источник: http://www.datadrivenjournalism.ru/2015/01/20/de4_project_budget_kz/

Примеры дата-экспедиций.

Европейский университет— 2017

- *В апреле 2017 года в Европейском университете*
- *Количество участников — 14 человек*
- *Сроки проведения — 2 недели*
- *Количество очных занятий — 12*
- *Количество контрольных оцениваний — 5*



Примеры дата-экспедиций.

Европейский университет— 2017

The screenshot shows a GitHub repository page for 'iradche / Data-Expedition-in-EU'. The repository has 15 issues, 0 pull requests, 1 project, and 0 wiki pages. It includes security and insights features. The repository has 3 stars, 0 forks, and 2 contributors. It uses the MIT license. The commit history shows 100 commits, 2 branches, 1 release, 1 environment, and 2 contributors. The latest commit was on March 25, 2018. The commits include:

- Create scraper.py (iradche, 2 years ago)
- Add a text on writing process (academicWriting, 2 years ago)
- Create miscAPI.md (api, 2 years ago)
- Create readme.md (assessment, 2 years ago)
- Add cleaned dataset with recommended VAK journals (datasets, 2 years ago)
- Update README.md (doc, a year ago)



Наши data-экспедиции: результаты опроса

Чем именно была полезна для вас data-экспедиция? Чему вы научились?
(14 ответов)

Узнал много полезных проектов для визуализации данных и не только, ближе познакомился с проектами Google

Данная data экспедиция была очень полезной. Научилась создавать гугл таблицы, гугл документы, работать с открытыми данными, анализировать их...

Получил базовые знания с набирающей популярностью концепцией.

Лучше стал работать с открытыми данными

Теперь я пользуюсь только почтой gmail, умею создавать табл, документы в гугле, работать с инфографикой. Заинтересовалась геоданными.)))

Интересно узнавать новое для себя, а так же узнала много нового о работе в gmail.com, инфографике, в самом google для поиска информации

Получила информацию об открытых данных, научилась искать, обрабатывать данные. Научилась делать инфографики на inform.aim. Работали группами.

- Понять смысл журналистики открытых данных.
- Искать и находить открытые данные.
- Конвертировать открытые данные из формата А в формат Б.
- Пользоваться инструментами визуализации данных.
- Пользоваться электронными таблицами.
- Представлять проект как единое целое.

Для меня data-экспедиция была полезна. Научилась работать на гугл-группе (гугл-таблица, гугл-документ, инфограмма и т.д.), честно говоря, чего я до этого не знала.

Методика проведения дата-экспедиций



Темы занятий для проведения дата-экспедиций

№	Тема
1	Введение
2	Поиск открытых данных
3	Обзор/обсуждение проблем
4	Исследование задач
5	Публикация и цитирование данных
6	Обработка открытых данных
7	Анализ открытых данных
8	Визуализация открытых данных

Дата-экспедиции в вузах

С 2012 года:

1. Высшая школа экономики
2. Университет ИТМО
3. Карагандинский государственный технический университет (Казахстан)
4. Санкт-Петербургский государственный университет
5. Европейский университет

Публикации

2019

- [Collaborative Learning in Data Science Education: A Data Expedition as a Formative Assessment Tool*](#)

2018

- [On Portrait of a Specialist in Open Data*](#)

2017

- [On development of a framework for massive source code analysis using static code analyzers*](#)

2016

- [Principles of Citizen Science in Open Educational Projects Based on Open Data*](#)
- [Open Government Data in Russian Federation*](#)
- [A Case Study of Open Science Concept: Linked Open Data in University*](#)
- [A Survey of Open Government Data in Russian Federation*](#)
- [On some Russian Educational Projects in Open Data and Data Journalism \(full text\)*](#)
- [Transformation of Communication Processes: Data Journalism*](#)

2015

- [Overview of Russian Open Education Projects](#)
- [Experimenting with Data Expeditions](#)

2014

- [The View on Open Data and Data Journalism: Cases, Educational Resources and Current Trends \(full text\)*](#)
- [Data Expeditions and Data Journalism project as OER in Russian](#)



На основе источника: <http://iradche.ru/useful/my-publications/>

Дата-экспедиции: варианты

- Удаленно (в режиме онлайн)
- Сессионно (в режимах онлайн и оффлайн)
- Мастер-классы (в режиме оффлайн)
- Экспресс-практикум (в режиме оффлайн)

Дата-экспедиции: чему могут научить

- Поиск открытых данных
 - Обработка данных
 - Очистка данных
 - Анализ данных
 - Визуализация данных
 - Архивация данных
-
- Публикация данных
 - Распространение данных
 - Управление данными
 - Создание продукта на основе данных
 - Создание аналитического материала
 - Написание научной статьи



Ирина Радченко

@iradche

<http://www.slideshare.net/iradche>

<https://www.facebook.com/iradche>

iradche@gmail.com

