

Discipline: Media Lab in Online Media
St.Petersburg University, 2016

Introduction to Data Journalism

Irina Radchenko,
Associate Professor @ ITMO University,
Chief Organiser, Open Knowledge International - Russia
@iRadche, <http://www.datadrivenjournalism.ru>

The Data Journalism Handbook

The Data Journalism Handbook

[SIGN UP TO GET UPDATES](#)

The Data Journalism Handbook is a free, open source reference book for anyone interested in the emerging field of data journalism.

It was born at a 48 hour workshop at [MozFest 2011](#) in London. It subsequently spilled over into an international, collaborative effort involving dozens of data journalism's leading advocates and best practitioners - including from the Australian Broadcasting Corporation, the BBC, the Chicago Tribune, Deutsche Welle, the Guardian, the Financial Times, Helsingin Sanomat, La Nacion, the New York Times, ProPublica, the Washington Post, the Texas Tribune, Verdens Gang, Wales Online, Zeit Online and many others.

[O'REILLY BOOK AVAILABLE](#) [FREE WEB VERSION AVAILABLE](#)

Now also available in:
[Russian](#) - [Spanish](#) - [French](#) - [Georgian](#)

<http://datajournalismhandbook.org/>

What is Data Journalism?



Data journalism is an umbrella term that, to my mind, encompasses an ever-growing set of tools, techniques and approaches to storytelling.

It can include everything from traditional computer-assisted reporting (using data as a ‘source’) to the most cutting edge data visualization and news applications.

The unifying goal is a journalistic one: providing information and analysis to help inform us all about important issues of the day.

— Aron Pilhofer, New York Times

What is Data Journalism?

Data-driven journalism is the future. Journalists need to be data-savvy. It used to be that you would get stories by chatting to people in bars, and it still might be that you'll do it that way some times. But now it's also going to be about poring over data and equipping yourself with the tools to analyze it and picking out what's interesting. And keeping it in perspective, helping people out by really seeing where it all fits together, and what's going on in the country.

— Tim Berners-Lee, founder of the World Wide Web



Best practices. The Guardian

The screenshot shows the homepage of The Guardian website. At the top, there is a dark blue header bar with navigation links: 'sign in', 'become a supporter', 'subscribe', 'search', 'jobs', 'dating', 'more ▾', 'International ▾', and the 'theguardian' logo. Below the header is a dark grey navigation bar with categories: 'UK', 'world', 'sport', 'football', 'opinion', 'culture', 'business', 'lifestyle', 'fashion', 'environment', 'tech', and 'travel'. To the right of the navigation bar is a link to 'browse all sections'. The main content area has a light grey background. On the left, there is a sidebar titled 'datablog' with the heading 'Flu tracking in Australia: weekly updates on symptoms around the country'. It includes a small description and a link to the full article. The main content area features a large map of Australia with colored dots representing flu activity across the country. To the right of the map is a sidebar with a thumbnail image of a group of people and the headline 'Bigger than Waterloo: the internet's favourite Eurovision winners'. Below the sidebar are two smaller images and their respective headlines.

sign in become a supporter subscribe search jobs dating more ▾ International ▾

theguardian

UK world sport football opinion culture business lifestyle fashion environment tech travel browse all sections

home

datablog

Flu tracking in Australia: weekly updates on symptoms around the country

This interactive graphic provides weekly updates on the flu situation around Australia, based on a weekly survey of over 20,000 Australians

47

Spill: Turnbull replaces Abbott

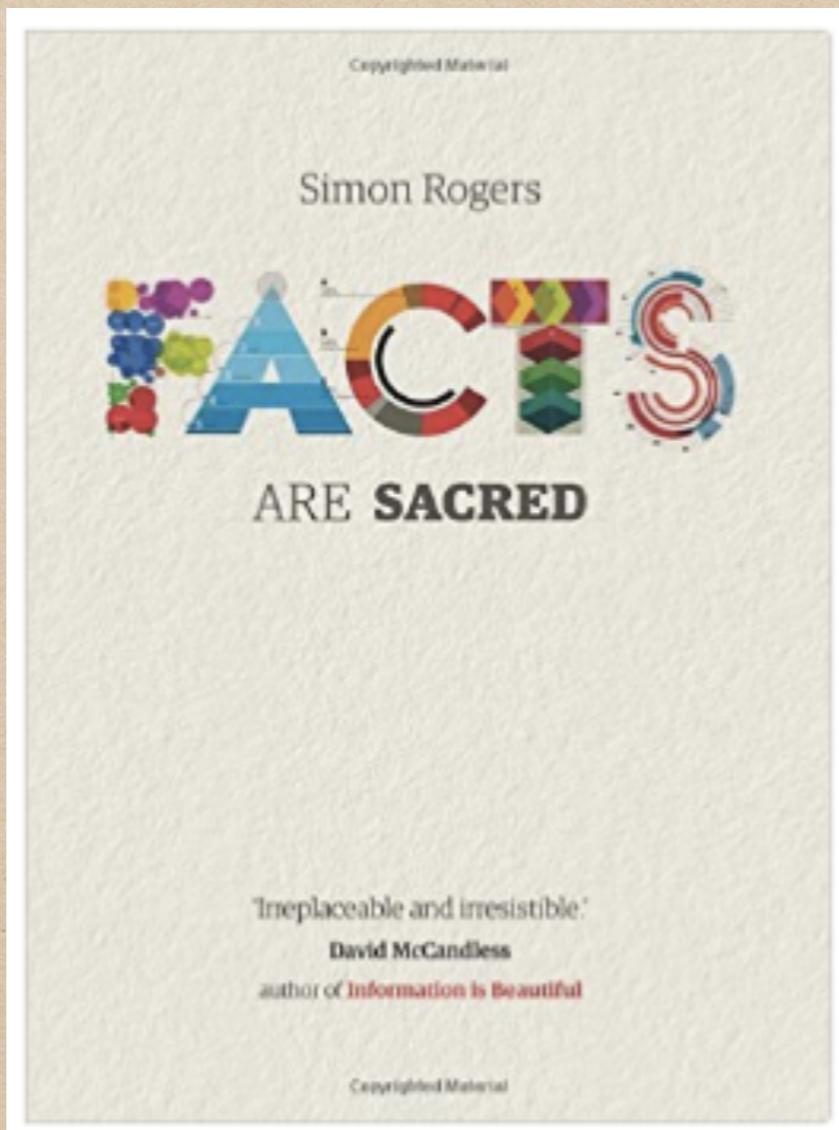
https://www.theguardian.com/tv-and-radio/datablog/2016/may/14/bigger-than-waterloo-the-internets-favourite-eurovision-winners

Bigger than Waterloo: the internet's favourite Eurovision winners

46

<https://www.theguardian.com/data>

Simon Rogers. Datablog



<https://www.amazon.com/Facts-are-Sacred-Power-Data/dp/0571301614>

Data-journalists are the new punks: Simon Rogers at TEDxPantheonSorbonne



<https://youtu.be/h2zbvmXskSE>

New York Times API

The screenshot shows the homepage of the New York Times Developers site. At the top, there's a header with the NYTimes logo and the word "Developers". To the right of the logo is a blue button labeled "Get NYT API Key". Below the header, the main title is "The New York Times Developer Network" followed by the subtitle "All the APIs Fit to POST". A descriptive paragraph explains that NYTimes.com is a premier source of data and encourages hacking it. Under the heading "Getting Started", it says the Times Developer Network is an API clearinghouse and community, followed by a list of three steps: 1. Request an API key, 2. Read the API documentation, FAQ and Terms of Use, and 3. Use the API Tool associated with each API to experiment without writing code. At the bottom, there's a section titled "Using The NYT APIs".

The New York Times Developer Network

All the APIs Fit to POST

You already know that NYTimes.com is an unparalleled source of news and information. But now it's a premier source of data, too — why just read the news when you can hack it?

Getting Started

The Times Developer Network is our API clearinghouse and community. Here's how to get started:

1. Request an [API key](#)
2. Read the API documentation, [FAQ](#) and [Terms of Use](#)
3. Use the API Tool associated with each API to experiment without writing code

Using The NYT APIs

<https://developer.nytimes.com>

New York Times - Linked Open Data

The easy way to get, use and share data

Datasets Organizations About Blog Help Search

Home / Organizations / Linking Open Data Cloud / New York Times - Linked Open Data

New York Times - Linked Open Data

Followers: 1

Organization

Linking Open Data Cloud

See also 2014 augmentation. This group catalogs data sets that are available on the Web as Linked Data and contain

Dataset Groups Activity Stream

New York Times - Linked Open Data

About

From website:

> For the last 150 years, The New York Times has maintained one of the most authoritative news vocabularies ever developed. In 2009, we began to publish this vocabulary as linked open data.

[Download Data Package](#)

Data and Resources

RDF People (SKOS) [More information](#) [Go to resource](#)

RDF Organizations (SKOS) [More information](#) [Go to resource](#)

RDF Locations (SKOS) [More information](#) [Go to resource](#)

RDF Subject descriptors (SKOS) [More information](#) [Go to resource](#)

<https://datahub.io/dataset/nytimes-linked-open-data>

Financial Times Data

ft.com > comment > blogs >

FT Data

Subscribe | Sign in Search for...

Subscribe now - €1 for 4 weeks ►

Home UK World Companies Markets Global Economy Lex Comment Management Personal Finance Life & Arts Columnists The Big Read Opinion FT View Instant Insight EM Squared The Exchange Blogs Letters Corrections Obits Tools

Highlights

Austerity State: Local cuts checker database

Bank chief executives' pay 2009-2014

What Greece owes in the coming months

United States debt to GDP 1970-1990

A world of debt: FT-IMF debt-to-GDP calculator

Could subscriptions fund the BBC?

Datawatch: Afghan opium cultivation

Christopher Campbell Author alerts | Oct 29 07:00 | Comment | Share

Opium poppy cultivation in Afghanistan

Hectares ('000)

Year	Hectares ('000)
2000	~120
2001	~100
2002	~150
2003	~180
2004	~150
2005	~180
2006	~220
2007	~250
2008	~200
2009	~220
2010	~180

FT Data on statistics

About this blog Feedback Commenting

Welcome. FT data is a collaborative effort from journalists across the FT working in data journalism, statistics and data visualisation.

We aim to look at statistical issues in depth and provide original analysis of data. Occasionally we'll also post guest contributions.

<http://blogs.ft.com/ftdata/>

Data Journalist Job

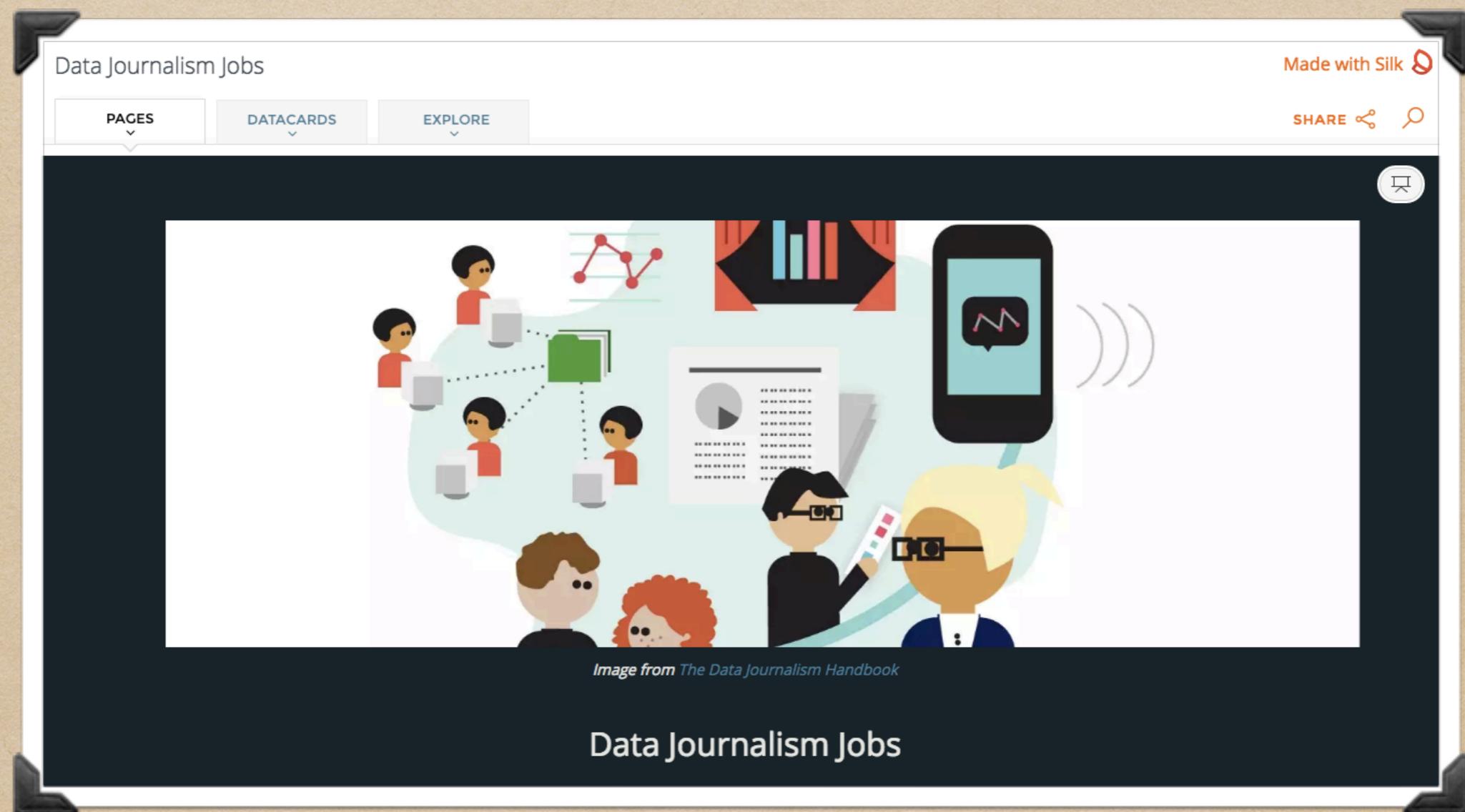
Data Journalism Jobs							
Job	Company	Location	Salary	Skills	Details	Apply	Date submitted
Data reporter	Las Vegas Review-Journal	Las Vegas, NV		Applicants should have experience with programs needed for data manipulation and presentation -- including Excel, Access, Illustrator, Tableau, Google Fusion -- as well as experience with data mapping tools. Intermediate or better coding skills are a definite plus. Applicants should have at least a bachelor's degree in journalism, communications or a related field, reporting experience and a record of delivering results in a deadline environment.	The Las Vegas Review-Journal is seeking a data-driven journalist to work with reporters and editors throughout the newsroom to produce information-rich stories across all platforms. This person must know how to find, mine and analyze data, and ideally, will have the reporting and writing skills to help readers understand the data behind the news. The ideal candidate will have familiarity with existing public data sets and government data sources.	pjohnson@reviewjournal.com	10/26/2016
Editorial Designer	The Marshall Project	New York		The ideal candidate is a well-rounded journalist and designer with 3-5 years' experience in media and/or publishing. The Editorial Designer should have a wide range of visual interests and experience and the ability to juggle a variety of tasks in a small but ambitious newsroom.	The Marshall Project seeks a full-time Editorial Designer to join us in our office in New York and oversee the award-winning visuals of the publication, including digital design, photography and illustration.	Apply here	10/25/2016
Data Visuals	Texas Tribune	Austin, TX		You'll be asked to gather your own data. Our editorial Data Visuals team works in		jobs@texastribune.org	9/29/2016

Chris Persaud made this site. Wanna improve it? [Fork it on Github](#) and you'll get credit!

[TELL THE PROS](#) [TELL EVERYONE](#) [TELL FRIENDS](#) [TELL A FEW PEOPLE](#) [EMAIL THIS](#) [FORK THIS!](#)

<http://datajournalismjobs.com>

Data Journalist Job



<http://data-journalism-jobs.silk.co>

Data Journalist Vacancies @ LinkedIn

The screenshot shows a LinkedIn search interface. At the top, there are search fields for 'Data Journalism' and 'США' (USA), and a 'Поиск вакансий' (Search jobs) button. To the left, a sidebar offers options to receive notifications about new vacancies and includes dropdown menus for 'Region' (listing New York, San Francisco, Chicago, Washington, and Los Angeles) and 'Company'. The main results section displays three job listings:

- Copywriter** at Scout Exchange in Boston, Massachusetts. The listing describes a role where creative spirit turns data into engaging content. It is a new listing (6 days ago).
[Apply]
- Data Editor/Video Producer - CBSNews.com** at CBS Interactive in New York City, NY, US. The listing is a new posting (NOVINKA). It describes a role for its MoneyWatch section.
[CBS News logo]
- Career Opportunities: Data Reporter** at VelvetJobs in Woodland Park, NJ, US. The listing is a new posting (NOVINKA). It describes a role leading data-driven journalism.
[VelvetJobs logo]

<https://www.linkedin.com/jobs/data-journalism-jobs>

Making data mean more through storytelling: Ben Wellington at TEDxBroadway



<https://youtu.be/6xsvGYIxJok>

OpenSpending Project

The screenshot shows the homepage of the OpenSpending Project. At the top, there is a navigation bar with links for OpenSpending, Datasets, Search, Community, Help, About, Contact, Blog, Login/Register, and Language. Below the navigation bar is a large green world map. To the right of the map, the text "It's our money!" is displayed, followed by a subtitle: "By understanding how governments spend money in **our name** can we have a say in how that money will affect our own lives. The journey starts here." A central callout box highlights key statistics: "76 countries", "1091 datasets", and "28323775 entries". Below this, there is a search bar with the placeholder "Search...". On the left side, there is a "Featured Application" section titled "Spending Stories" which says "How much is that really? — See the stories behind the numbers in the news with the Spending Stories application." At the bottom, there are four main sections: "Upload" (with a circular chart showing spending data), "Explore" (with a screenshot of a transaction detail page), "Visualize" (with a screenshot of a data visualization interface), and "Contribute" (with a photo of people working together at computers).

OpenSpending

Datasets Search Community Help About Contact Blog

Login/Register Language

It's our money!

By understanding how governments spend money in **our name** can we have a say in how that money will affect our own lives.
The journey starts here.

76 countries
1091 datasets
28323775 entries

Search...

Featured Application

Spending Stories

How much is that really? — See the stories behind the numbers in the news with the Spending Stories application.

Upload

20.7m United Republic of Tanzania
73.8m More

Explore

Entry: Detailed View of Transaction
This is a single transaction of GBP 25,586 on 2013-10-15 in UK Departmental Spending.

Entity: Defence Science and Technology Laboratory
Supplier: THE SHAWHORSE COMPANY LIMITED

Dimension	Value
Entity	25,586 GBP
Description	

Visualize

Contribute

<https://openspending.org/>

Open Data Institute Courses

The screenshot shows the ODI website's homepage with a focus on the 'Courses' section. The header features the ODI logo and navigation links for About, Get involved, What we offer, Themes, Publications, Events, and News & blog. A prominent blue banner at the top of the main content area is titled 'Courses'. Below this, a dark grey box highlights a specific course: 'Open Data for Smart Cities', described as 'By 2050, around 70% of people on the planet will live in urban areas.' It includes icons for a calendar and people, the date '15th Sep (London)', and a red 'Book' button.

<http://theodi.org/courses>

Data Driven Journalism, European Project

The screenshot shows the 'Resources' section of the Data Driven Journalism website. The header features a logo with three bars and the text 'Data Driven Journalism WHERE JOURNALISM MEETS DATA'. A navigation bar includes links for About, News & Analysis, Events, Featured Projects, Resources (which is underlined), Mailinglist, and Course. A search bar and social media icons for Twitter, Facebook, and RSS are also present. The main content area has a black header bar with 'Home > Resources > Overview'. Below this, a section titled 'Resources' lists two items:

- Tabula 1.0 released: Tool helps to extract data from PDFs**
Many sources release data only in the PDF format - which is causing considerable trouble for data extraction. The free and easy Tabula helps to get the numbers back into readable and usable formats.
[Read more](#)
- Use flowcharts to avoid getting stuck in data journalism projects**
Writing stories is not easy, but well understood. When reporting from data though it is quite common to get stuck. To avoid this create a flowchart for your particular project - as an overview and a means to communicate your goals to others.
[Read more](#)

On the right side, there are two boxes: 'Submit your resource' (with a link to 'Do you want to see your work featured on our website?') and 'Upcoming Events' (listing '3/9/2015 - 3/9/2015 News Impact Summit' with a topic about digital innovation and journalism).

<http://datadrivenjournalism.net/resources>

Doing Journalism with Data

The screenshot shows the homepage of the 'Doing Journalism with Data' website. The page has a white background with black text and features a large, bold title at the top. Below the title, there is a sub-section title and a brief description. A navigation bar with several links is located above a yellow notice bar. The main content area includes three orange call-to-action buttons and a video player.

**Doing Journalism with Data:
First Steps, Skills and Tools**

A free online data journalism course with 5 leading experts.

⌚ Open from May 19th 2014 until March 31st 2015

[HOME](#) [COURSE OVERVIEW](#) [ABOUT](#) [GRADUATES](#) [GIVEAWAYS](#) [FAQ](#) [CONTACT](#) [REGISTER](#)

NOTICE: We are relaunching the course very soon, click here to get notified!

→ LEARN MORE

→ REGISTER NOW

→ LOCAL LEARNING GROUPS

Doing Journalism with Data:
First Steps, Skills and Tools

01:29

vimeo

<http://datajournalismcourse.net/>

School of Data

The screenshot shows the homepage of the School of Data website. At the top, there's a navigation bar with links for About, Learn, Teach, Events, Blog, Support Us, Twitter, and a logo. Below the navigation is a large "Welcome!" section with text about empowering civil society organizations, journalists, and citizens. It includes a call to action to "Get started with School of Data today!". To the right, there's a "Join School of Data" section with a form for email subscription. Further down, there's a "Check out our blog!" section with a "Blog" button, a "Send in your data stories" section, and a "Latest" news feed. The latest news items include "InfoAmazonia and Why We Should Talk About..." and "Data roundup, June 19". At the bottom, there's an "About School of Data" section with a brief description of their mission.

Welcome!

School of Data works to empower **civil society organizations, journalists and citizens** with the skills they need to use data effectively – evidence is power!

Get started with School of Data today!

Join School of Data

Our courses are for everyone: whether you're a data-newbie or a pro looking for inspiration, there's something for you!

Get updates and new courses from the School of Data in your inbox.

Your email

Subscribe

Check out our blog!

Send in your data stories

Have you done something cool with data? Have you recently discovered a new tool, or found a way to work through a tricky problem? Pop us a note to schoolofdata [@] okfn.org. Read our blogging [guidelines](#).

Latest {

InfoAmazonia and Why We Should Talk About...

Data roundup, June 19

We're rounding up data news...

[Go to blog »](#)

About School of Data

School of Data works to empower civil society organizations, journalists and citizens with the skills they need to use data effectively in their efforts to

<http://schoolofdata.org/>

Data Expeditions

Data Expeditions at MozFest

November 14, 2012 in [Data Expeditions](#), [Events](#)

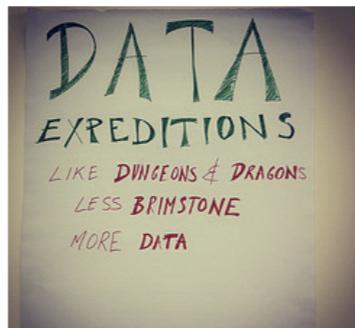
Expeditions into the Data Landscape: the School of Data goes to #MozFest

Find out what happened at MozFest – and see the tools and data sets to recreate it yourself!

Saturday morning at [MozFest](#). A sold out building, full of a thousand hackers, builders, makers, geeks, journalists, thinkers and more. And right at the top on the 9th floor? Three 'data sherpas' in sparkly cloaks...

Data Expeditions

The concept behind the 'Data Expeditions' run by the School of Data at this year's MozFest was simple. Based on the 'Dungeons and Dragons' role-playing game, data explorers would tackle real world problems together, developing their data wrangling skills in the process.



As a first step, explorers were asked to rate their abilities. Can you tell a story? analyse data? code? tweet? draw? The emphasis was on 'doing', but not in any narrow sense – often, it's the data newbie asking a 'stupid question' that sets the team on a fresh track, and becomes the biggest contribution of the day.

Next came the quests. Three Data Sherpas (still sparkling) set out three missions: delving into the data surrounding extractive industries and oil mines; exploring possible causes for a dramatic

Written by
Lucy Chambers

Lucy is the Head of the Knowledge Unit at the Open Knowledge Foundation. She is also the project lead for School of Data, and previously was the Community Coordinator for OpenSpending, Data-Driven Journalism and Spending Stories.

Get updates and new courses from the School of Data in your inbox.

Your name

Your email

Subscribe



Told a data story recently? Ran a successful data driven campaign? Want to share with others how you did it? [Contact us!](#) We're always looking for guest posts.

Have a data question? Got stuck in a data project? [Ask Schoolofdata!](#)

<http://schoolofdata.org/2012/11/14/data-expeditions-at-mozfest/>

Russian Data Journalism blog

The screenshot shows a website with a dark header and a light-colored main content area. The header includes social media links (Facebook, Twitter, Pinterest, Github) and a search bar. Below the header is a navigation bar with tabs: ПРАКТИКУМ, ПОЛЕЗНЫЕ МАТЕРИАЛЫ, О ПРОЕКТЕ, ОБ АВТОРАХ, ПРИСОЕДИНИЯЙТЕСЬ!, ENGLISH, and МЫ НА ГИТХАБЕ. The main content area features two articles. The first article, titled 'Саммит «Открытые данные — 2015»' (Summit 'Open Data - 2015') and dated 24.11.2015, has 0 comments. It includes a thumbnail image of a purple and blue geometric pattern. The second article, titled 'Интерактивный встраиваемый таймлайн' (Interactive embeddable timeline) and dated 06.10.2015, also has 0 comments. It includes a thumbnail image of a timeline creator interface. To the right, there is a sidebar with a 'DATA JOURNALISM' logo and a list of authors: Alina Pinchuk, Alla Rybina, Anastasia Valeeva, Anna Sakoyan, Farishtamoh Gulova, Gulim Amirkhanova, Irina Radchenko, Kristina Kashtanova, Natalia Karbasova, Philipp Kats, and Yulia Mityaeva. A search form is also present in the sidebar.

ЖУРНАЛИСТИКА ДАННЫХ

ПРАКТИКУМ · ПОЛЕЗНЫЕ МАТЕРИАЛЫ · О ПРОЕКТЕ · ОБ АВТОРАХ · ПРИСОЕДИНИЯЙТЕСЬ! · ENGLISH · МЫ НА ГИТХАБЕ

МЕРОПРИЯТИЯ, ОТКРЫТЫЕ ДАННЫЕ

Саммит «Открытые данные — 2015»

24.11.2015 — 0 Comments

HOWTO, ВИЗУАЛИЗАЦИЯ, ИНСТРУМЕНТЫ

Интерактивный встраиваемый таймлайн

06.10.2015 — 0 Comments

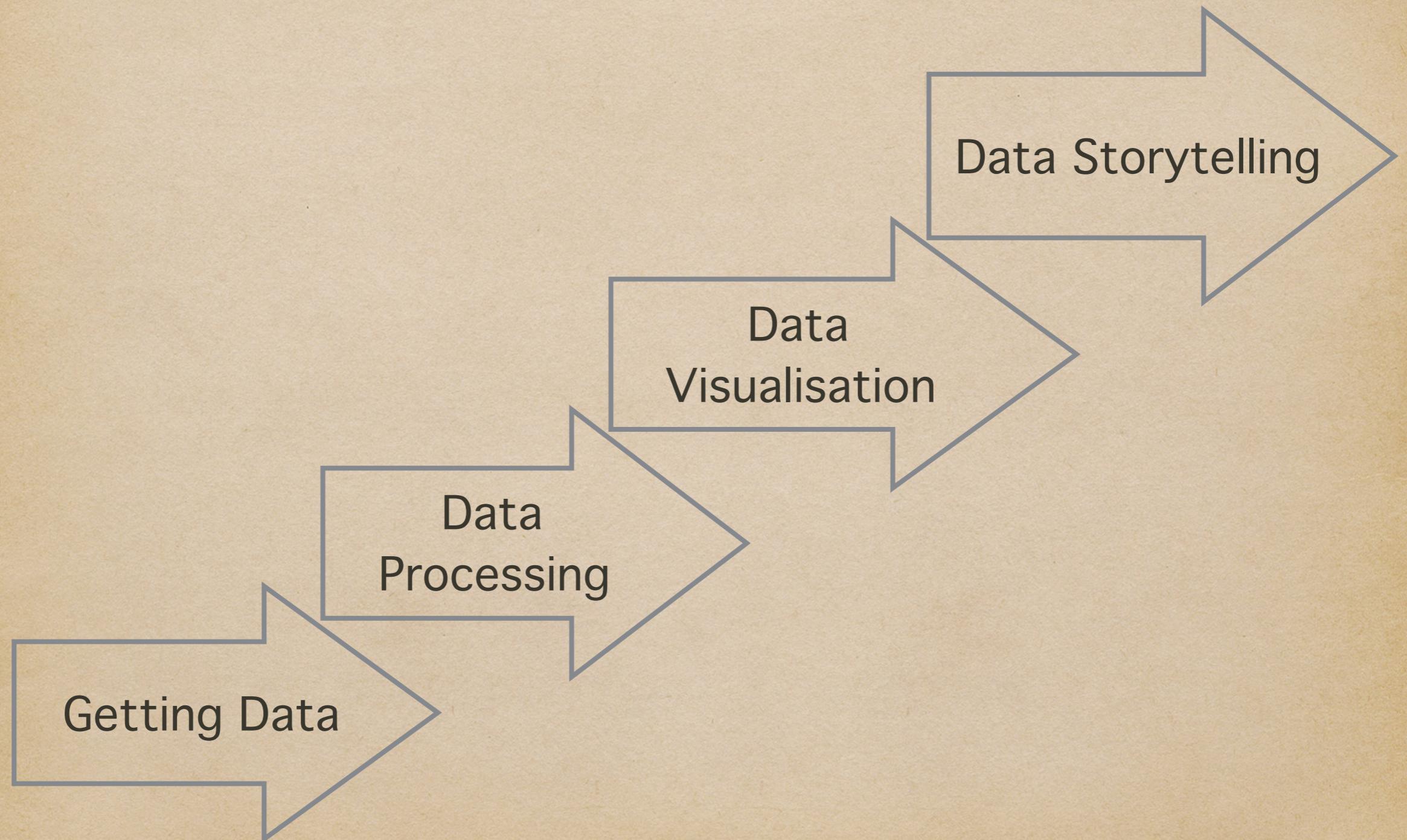
Search form

АВТОРЫ

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

<http://datadrivenjournalism.ru/>

Data Journalism as a Process



Questions?



<http://iRadche.ru>

<http://about.me/Irina.Radchenko>

<http://DataDrivenJournalism.ru>



 @iRadche

 <https://github.com/iRadche>

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

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