

Data Journalism Resources

List of Data Resources from IJNET

While there are many resources out there for learning data-driven reporting, most focus on the U.S or Europe. But thanks to journalist and digital publishing consultant Phillip Smith, those who want to up their data journalism game can turn to this robust list of more than 50 resources from around the world.

For a recent data journalism workshop Smith led in Venezuela organized by [IPYS](#), he curated this list with help from the NICAR listserve. (You can follow Smith on Twitter at [@phillipadsmith](#) or check out his workshop slides [here](#).)

Below we are posting an edited version with his permission. You can also see the original list in a Google Doc [here](#). If you think of other resources (websites, books, videos, etc.), please add them in the comments, and we'll update the list.

While Smith's list was created with Venezuelan reporters in mind, many of the resources are useful to journalists worldwide. Resources in this version of the list are in English unless otherwise noted.

Guides and articles

- [Data Journalism Handbook](#)
- [Paving the way for data journalism in a divided Venezuela](#)
- [Dataviz catalogue](#)
- [DatosPublicos.org](#) (in Spanish)
- [About "datos públicos"](#) (in Spanish)
- [Qué es Poderopedia](#) (in Spanish)
- [Manual de periodismo de investigación \(Investigative journalism manual\)](#) (in Spanish)
- [Herramientas digitales para periodistas \(Digital tools for journalists\)](#) (in Spanish)

Books

- [Open Data Handbook](#)
- [Verification Handbook](#)
- [A Practical Guide to Designing with Data, Five Simple Steps](#)
- [Naked Statistics: Stripping the Dread from the Data](#)
- [Show Me the Numbers: Designing Tables and Graphs to Enlighten](#)
- [The Functional Art: An introduction to information graphics and visualization](#)
- [Interactive Data Visualization for the Web](#)
- [Visualize This: The FlowingData Guide to Design, Visualization, and Statistics](#)
- [Beautiful Visualization: Looking at Data Through the Eyes of Experts](#)
- [The Visual Display of Quantitative Information](#)
- [Introducción al análisis de datos y mapeo con Google Fusion Tables](#)
- [The Data Visualisation Catalogue](#)
- [Cryptoperiodismo: Manual Ilustrado de Seguridad para Periodistas](#) (in Spanish)

Inspiration and presentations

- [Amanda Cox, Graphics Editor at the New York Times, on data visualization \(a 30-minute video\)](#)
- [Stanford's "Journalism in the Age of Data" \(an hour-long documentary on data journalism\)](#)
- [Any of Hans Rosling's videos about Gapminder](#)
- [Information is beautiful](#)
- [FlowingData](#)
- [The Data Visualisation Catalogue](#)
- Presentations by Alastair Dant of the New York Times and ICFJ Knight International Journalism Fellows [Mariana Santos](#) and [Miguel Paz](#) at Venezuela's [First Data Journalism Boot Camp](#) (in Spanish)

Where to find data and open data

Global and open data catalogs

- [World Health Organization](#)
- [United Nations](#)
- [World Bank](#)
- [DataCatalogs.org](#)
- [The Guardian's world government data portal](#)
- [Google's public data directory](#)
- [The data hub](#)
- [DBPedia Datasets](#)
- [Factual](#)
- [Free GIS data](#)
- [List of open data resources](#)
- [Energy data repositories](#)
- [Data wrangling](#)
- [Quora thread: "Where can I find large datasets open to the public?"](#)
- [Directory of APIs](#)
- [Infochimps](#)
- [Datamarket](#)
- [Offshore Leaks](#)
- [Investigative Dashboard](#)
- [Open Corporates](#)
- [Natural Earth data](#)

Country-specific data for a sampling of countries

- [Afghanistan Election Data](#)
- [American Fact Finder](#)
- [Data.gov - EE.UU.](#)
- [Kenya's open data portal](#)

Latin America-specific data

The resources for Latin America-specific data, and most of the remaining resources on this list, are in Spanish.

- [Open Data Latinoamérica](#)

Understanding Data

To download this report, visit <http://cima.ned.org/publications>

- [Tan conectados como valientes - Uruguay](#)
- [Datos.gub.uy](#)
- [Qué sabés - Uruguay](#)
- [Censo for Humans - Uruguay](#)
- [Datos Abiertos Colombia](#)

Venezuela--National Level

- [Consejo Nacional Electoral \(Registro Electoral, Resultados Electorales\)](#)
- [Tribunal Supremo de Justicia](#) (Sentencias Judiciales)
- [Banco Central de Venezuela](#) (Información estadística, informes económicos)
- [Registro Nacional de Contratistas](#) (Empresas)
- [Instituto Nacional de Estadística](#) (Datos estadísticos sobre Venezuela: población, regiones, indicadores económicos)
- [Ministerio del Poder Popular para la Salud](#) (Datos epidemiológicos)
- [The NYT's linked open data on Venezuela](#)
- [Geonames](#)
- [Gobernación del estado Zulia \(Contrataciones, balances de entes adscritos\)](#)
- [Cámara de la Construcción del estado Carabobo \(Listado de afiliados con información de contacto: dirección y teléfono\)](#)

Venezuela--Municipal Resources

- [Alcaldía de Naguanagua](#) (Ordenanzas, decretos y reglamentos, ejecución presupuestaria)
- [Alcaldía de Maracaibo](#) (Mapa con centros de salud del Municipio, ubicación de los cuerpos de seguridad y de emergencia, rutas de transporte público, contrataciones, mapa proyectos de responsabilidad social)

Other

- [Páginas Amarillas](#) (Teléfonos)
- Universidad Central de Venezuela - [Base de datos de egresados hasta 2004](#)
- [Fondo Nacional para Edificaciones Penitenciarias - Venezuela](#)
- [Transparencia Venezuela](#)
- [Coalición Pro Acceso - Venezuela](#)
- [Base de datos del Banco Interamericano de Desarrollo](#)

Examples of data journalism in Latin America

- [Venezuela's election results mapped as open data](#) - The Guardian datablog
- [Cargografías](#)
- [Latin America's Open Data movement](#)
- [Hacks Hackers Buenos Aires](#)
- [2011 elections - Argentina](#) - HHBA
- [Analice.me](#)
- [D3](#) - HHBA
- [Hackdash Bolivia](#) from a boot camp organized by [Mariano Blejman](#) in Bolivia
- [Hackdash con proyectos de periodismo de datos en Venezuela](#)
- [Hack Cívico](#)
- [Andy Tow's Década Votada](#)

- [La Nación Data](#) (Argentina)
- [La Nación / No voto a ciegas](#) (Costa Rica)
- [La Nación Data — Censo Argentina](#)
- [La Nación — Inundaciones en La Plata](#)
- La Nación — [Mapa Elecciones 2013](#) [Mapa de homicidios - Rosario](#)
- [El Nacional - Tweetómetro](#) (Venezuela)
- Últimas Noticias - [El audio que causó el sacudón](#) (Venezuela)
- El Mundo - [Leche importada por el gobierno se va de contrabando a Colombia](#) - (Venezuela)
- El Universo, El Universal, Reuters, Armando.info - [PDVSA usó a Glencore y Trafigura para proveer derivados a Ecuador](#) - (Ecuador, Venezuela)
- La Nación Data (Costa Rica) / [Especiales](#)
- La Nación (Costa Rica) - [Geografía del Crimen](#)
- La Nación (Costa Rica) - [Escuelas y colegios tienen más de 40.000 alumnos fantasmas](#)
- [Folha de S. Paulo \(Brasil\) - Datafolha](#)
- Gazeta do Povo / RPCTV (Brasil) - [Diários secretos - base de datos](#) - [Diários secretos - Por Dentro](#)
- Consejo de Redacción (Colombia) - [Monitor de Corrupción](#).

List of Data Journalism Resources from GJIN

As our governments and businesses become increasingly flush with information, more and bigger data are becoming available from across the globe. Increasingly, investigative reporters need to know how to obtain, clean, and analyze “structured information” in this digital world.

Here is a list of resources to get you started, but we want to keep updating our community with the best resources available. Do you know of a great data tutorial we haven't listed, perhaps in a language other than English? Help us keep this resource guide comprehensive by sending your favorite resource to: hello@gijn.org. ¿Habla español? For resources in Spanish, [click here](#).

Key Resources

- [The National Institute for Computer-Assisted Reporting](#), a project of Investigative Reporters and Editors, launched in 1989 to train reporters around the world on how to use data as part of broader investigations. In addition to “boot camps” and in-office training, NICAR offers a data library, practice data sets, and hosts the original annual conference on computer-assisted reporting. IRE also publishes the popular book, [Computer-Assisted Reporting: A Practical Guide](#).
- [Poynter](#) offers [Five Tips](#) for getting started with computer-assisted reporting, and [10 Tools](#) to analyze datasets more efficiently.
- [The Center for Investigative Journalism](#) published a [manual on data journalism](#) “for all journalists who want to master the art of interrogating and questioning numbers competently.” [CIJ also provides](#) a slew of additional books, guides and video resources of aspects of data journalism.
- [Data-Driven Journalism](#) offers a [collection of resources](#) for computer-assisted reporters.
- [Periodismo de Base de Datos](#) provides [tutorials and resources](#) on data journalism for Spanish-speaking reporters.
- [Arab Reporters for Investigative Journalism](#) offers [this brief introduction](#) to data journalism (in Arabic).

- [The International Consortium of Investigative Journalists](#) provides a selection of [video tutorials](#) on basic Excel functions, as well as how to background a person or company, or find federal court documents in the U.S.
- [The International Journalists' Network](#) maintains [a blog](#) of the latest trainings, tools, and resources for data journalists.
- [Hacks/Hackers](#) is a global movement bringing together computer programmers and investigative journalists to tell powerful data-driven stories. Trainings offered through regional chapters.
- [The Investigative Dashboard](#) lists tools for data mining, visualization and social network analysis. Google search your tool of choice and you'll surely find tutorials on how to begin.
- [The Data Journalism Handbook](#) is an international, collaborative effort involving dozens of data journalism experts. The free guide is available for download in English, French, Georgian, Russian, and Spanish.

Data Mining

- [Code Academy](#) offers a series of free interactive trainings on the basics of HTML, CSS, JavaScript, Python, Ruby, and PHP.
- [Massachusetts Institute of Technology](#) offers a series [of free online courses](#) in computer programming with Python, Java, and C++.
- [Michael Martl](#) publishes an open-source textbook on how to program with Ruby on Rails.
- [ProPublica](#) ran this ["shopping list"](#) of tools and training guides for scraping data from the web using Ruby.
- [Online Journalism](#) published [an introduction](#) to using [ScraperWiki](#) to obtain data from the web.

Data Analysis

- [Investigative Reporters and Editors](#) provides [a simple tutorial](#) to converting PDFs to Text.
- [Electronic Data Resource Service](#) at McGill provides [a tutorial](#) on how to export a table from PDF to Excel.
- [School of Data](#) offers [a series of tutorials](#) – from finding datasets, to basic Excel skills and using the results to tell a story.
- [Dan Nguyen](#) put together [this tutorial](#) on using Google Refine to clean structured data sets, and also links to other video tutorials on Google Refine.
- [Github](#) offers a ["Gentle Introduction to SQL."](#)

Visualization & Mapping

- Edward Tufte's [books](#) and [courses](#) are industry standards.
- [Flowing Data](#) is run by statistician Nathan Yau, author of [Data Points: Visualization that Means Something](#) and [Visualize This: The FlowingData Guide to Design, Visualization, and Statistics](#).
- [Visualisationofdata.com](#) offers a directory of compelling infographics, how-to info, and more.
- [Esri](#) offers a series of [free online courses](#) for those interested in mapping with ArcGIS.
- Gustavo Faleiros created [JEO, a WordPress theme](#) for launching geodata-based sites. It allows news organizations, bloggers and NGOs to publish news stories as layers of information on digital maps.
- [Peter Aldhous](#) put together [a primer](#) on using Excel's free social network plugin, NodeXL.
- The [Data Visualisation Catalogue](#) is an on-going project to "help you find the right data visualization method for your data".

Statistics

- [OpenIntro](#) hosts this [free textbook](#) on statistics
- [Knight Digital Media Center](#) provides [free, two-day online courses](#).
- [Flowing Data](#) is run by statistician Nathan Yau, author of
- [Coursera](#) offers a number of online statistics courses including:
- [Passion-Driven Statistics](#) through Wesleyan University
- [Statistics, Making sense of data](#), offered through the University of Toronto
- [Statistics One](#), offered through Princeton University
- [Introduction to Statistics](#), offered through the University of California Berkeley
- Recommended Books on Statistics:
- [Damned Lies and Statistics](#), Joel Best
- [Data Analysis for Politics and Policy](#), Edward Tufte
- [Designing Social Inquiry](#), by King, Keohane, and Verba
- [The Drunkard's Walk: How Randomness Rules Our Lives](#), Leonard Mlodinow
- [How To Lie with Statistics](#), Darrel Huff
- [Naked Statistics: Stripping the Dread from the Data](#), Charles Wheelan
- [The Signal and the Noise](#), Nate Silver
- [Thinking, Fast and Slow](#), by Daniel Kahneman

Data & Technology Blogs

- [ProPublica Nerd Blog](#), secrets of data journalists and newsroom developers
- [Data Blog](#), the Guardian's blog on computer-assisted reporting
- [Nacion Data](#), Spanish-language data journalism blog of the Argentinian daily La Nación.
- [Open Knowledge Foundation](#), global movement to open up knowledge around the world and see it used and useful
- [Toledol](#), a Portuguese-language blog about computer-assisted reporting
- [Computational and Data Journalism](#), news and technology articles about data journalism
- [Computational Reporting](#), all about data mining
- [Dajore](#), data journalism research
- [Driven by Data](#), how data journalism is sifting through the facts
- [Vis4.net](#), random thoughts on information visualization and data journalism
- [Reporter's Lab](#), Duke University's blog on tools, techniques and research for public affairs reporting.
- [Tow Center for Digital Journalism](#), Columbia's blog on how technology is changing journalism, its practice and its consumption

Books

- [Computer-Assisted Reporting: A Comprehensive Primer](#), By Fred Vallance-Jones and David McKie
- [Computer-Assisted Reporting: A Practical Guide](#), the E-version by Brant Houston
- [Computer-Assisted Research: Information Strategies and Tools for Journalists](#), By Nora Paul and Kathleen A. Hansen
- [The Data Journalism Handbook](#) is an international, collaborative effort involving dozens of data journalism experts. The free guide is available for download in English, French, Georgian, Russian, and Spanish.
- [Mapping for Stories: A Computer-Assisted Reporting Guide](#), By Jennifer LaFleur and Andy Lehren
- [Precision Journalism: a Reporter's Introduction to Social Science Methods](#), by Philip Meyer

Conferences

- [NICAR](#) hosts the original annual conference on computer-assisted reporting, which is attended by hundreds, and also puts on data-specific boot camps.
- [Data Harvest](#) is a collaboration between the [Journalismfund.eu](#), [Wobbing Europe](#) and [FarmSubsidy.org](#). The next conference is scheduled for May 2014 in Brussels.
- [The International Journalism Festival](#) in Perugia, Italy, includes a School of Data Journalism training.
- [The Global Investigative Journalism Conference](#), held every two years, hosts a broad range of data-specific trainings.
- [Ghana Databootcamp](#) trains participants in Ghana on how to locate, obtain and analyze public data on the extractive industries.

Understanding Data: Can News Media Rise to the Challenge? is a publication of the Center for International Media Assistance (CIMA). The Center is an initiative of the National Endowment for Democracy that works to strengthen the support, raise the visibility, and improve the effectiveness of media assistance programs by providing information, building networks, conducting research, and highlighting the indispensable role independent media play in the creation and development of sustainable democracies around the world. An important aspect of CIMA's work is to research ways to attract additional U.S. private sector interest in and support for international media development.

CIMA convenes working groups, discussions, and panels on a variety of topics in the field of media development and assistance. The center also issues reports and recommendations based on working group discussions and other investigations. These reports aim to provide policymakers, as well as donors and practitioners, with ideas for bolstering the effectiveness of media assistance. For more information on CIMA, please visit <http://cima.ned.org>.