

Data Journalism

HTW Berlin
July 2, 2014



This work is licensed under a Creative Commons
Attribution-ShareAlike 3.0 Germany License

Michael Hörz - hoerz@michael-hoerz.de - [@data_meining](https://twitter.com/data_meining)

Overview

- What is Data Journalism
- How and what data is used?
- How are results presented?
- DDJ vs. Visualization
- Practices in
 - USA
 - UK
 - Germany

What is Data Journalism?

- To use datasets or their combination for showing a journalistic story
- "Interview a dataset" as in interviewing a person (Brant Houston)
- Goes back to 1970s: Computer Assisted Reporting (CAR) and "Precision Journalism" (Philip Meyer)
- Fundamentals by Adrian Holovaty (2006):
holovaty.com/writing/fundamental-change

Basic Process

- Research
- Evaluate data
- Clean up data
- Explore dataset
- Find results
- Present and visualize

blog.ouseful.info/2014/06/28/data-journalism-conversations-with-data-sources/

storify.com/mirkolorenz/workshop-making-data-pretty-school-of-data-journal

What Kind of Data?

- What is the minimum number of entries?
- Structured data
- Machine readable
- Obtained from data repositories
- Own sources: Freedom of Information requests, intelligent collection of editorial data (pandaproject.net), leaked documents, screen scraping from websites

Presentation Possibilities

- Interactive Infographics
- Timelines
- Interactive Maps
- Rich Media ("Snowfall")
- Gamification (bbc.co.uk/news/magazine-22000973)

Infographics Collection:

marijerooze.nl/thesis/graphics

DDJ vs. Visualization

- Data journalism is following journalistic questions with the help of data
- Visualization makes data easily readable
- Graphic patterns (on maps etc.) give hints, are pointers / starting points
- But the story starts with clear connections
- Combining several datasets is usually a criterion

Impact of Data Journalism

- What can Data Journalism specifically discover?
 - Big data + Scalability
 - Geographical patterns
 - Clusters + Networks
- Fact checking in data journalism
 - Transparency - provide the raw data
 - Cross-check with other data sources

github.com/propublica/guides/blob/master/data-bulletproofing.md

Examples in the USA

The New York Times

- Overview: nytimes.com/pages/multimedia
- Election Scenarios by the New York Times: elections.nytimes.com/2012/results/president/scenarios
- Beta Projects: beta620.nytimes.com
- Offspring - Upshot: nytimes.com/upshot
- Analysis: blog.visual.ly/10-things-you-can-learn-from-the-new-york-times-data-visualizations

Examples in the USA II

Washington Post

- Where Americans go to work:
washingtonpost.com/wp-srv/special/nation/census-commuting
- Election Graphics:
washingtonpost.com/wp-srv/special/politics/2012-elections-graphics
- Drone Crash Database:
washingtonpost.com/wp-srv/special/national/drone-crashes/database

Examples in the USA III

ProPublica

- Presidential Pardons:
propublica.org/series/presidential-pardons
- Dollars for Docs: projects.propublica.org/docdollars
- Tampa Bay Times - America's Worst Charities:
tampabay.com/americas-worst-charities
- FiveThirtyEight (Nate Silver + Team)
fivethirtyeight.com

Examples in the UK

The Guardian

- guardian.co.uk/news/datablog/2011/jul/28/data-journalism
- Guardian's Data Blog:
guardian.co.uk/news/datablog
- European Trade in Horsemeat:
guardian.co.uk/uk/datablog/interactive/2013/feb/15/europe-trade-horsemeat-map-interactive

Ampp3d

- mpp3d.mirror.co.uk/2014/05/27/watch-how-many-foreigners-enter-the-uk-every-single-second

Examples in Germany

Key Players: Open Data City: opendatacity.de

(= Marco Maas, Michael Kreil, [Lorenz Matzat])

- Mobile Phone Profile (for Zeit Online/NZZ):
 - zeit.de/datenschutz/malte-spitz-vorratsdaten
 - opendatacity.de/project/vorratsspeicherung-in-der-schweiz/
- Lobbying in Bern (for NZZ):
opendatacity.de/project/lobbying-in-bern
- Lobbyplag/Cloud: lobbyplag.eu/map / lobbycloud.eu
- Geheimer Krieg (for NDR & SZ): geheimerkrieg.de

Examples in Germany II

ZEIT Online

- Theme Site: zeit.de/datenjournalismus
- Members of Parliament: zeit.de/politik/deutschland/abgeordnetenbilanz
- Death victims of far-right violence: zeit.de/gesellschaft/zeitgeschehen/todesopfer-rechter-gewalt

Süddeutsche Zeitung

- Theme Site: sueddeutsche.de/thema/DataGraph
- Zugmonitor: sz.de/1.1651455
- Offshore Leaks: sz.de/1.1639812

Examples in Germany III

Berliner Morgenpost

- Federal Election Results 2013:
berlinwahlkarte2013.morgenpost.de/
- Tempelhof Airfield:
interaktiv.morgenpost.de/tempelhofer-feld
- State Election Results 2011:
morgenpost.de/berlin-aktuell/article1768373/Ergebnisse-der-Berliner-Abgeordnetenhauswahl-2011.html
- Flight Routes and Noise:
flugroutenradar.morgenpost.de/#mein-standort/2013-06-09/52.411234,13.129159

Examples in Germany IV

Spiegel Online

- Theme Site: [spiegel.de/thema/daten](https://www.spiegel.de/thema/daten)
- Population Development since 1855:
[spiegel.de/wissenschaft/mensch/datenlese-175-jahre-im-zeitraffer-bevoelkerung-morphing-a-940443.html](https://www.spiegel.de/wissenschaft/mensch/datenlese-175-jahre-im-zeitraffer-bevoelkerung-morphing-a-940443.html)
- Problems with the Census Data:
[spiegel.de/wissenschaft/mensch/datenlese-zweifel-am-zensus-a-942649.html](https://www.spiegel.de/wissenschaft/mensch/datenlese-zweifel-am-zensus-a-942649.html)
- Federal Election Results if...
[spiegel.de/politik/deutschland/alter-bildung-arbeitslosigkeit-die-alternativen-wahlergebnisse-a-923839.html](https://www.spiegel.de/politik/deutschland/alter-bildung-arbeitslosigkeit-die-alternativen-wahlergebnisse-a-923839.html)

Further Resources

- Data Driven Journalism gathers the essentials: datadrivenjournalism.net
- Data Journalism Handbook: datajournalismhandbook.org
- ProPublica Nerd Blog: propublica.org/nerds
- Simon Rogers (formerly The Guardian): simonrogers.net
- 13 inspirational data journalism projects: journalism.co.uk/news/-editors13-13-data-journalism-projects-big-and-small/s2/a553134
- The rise of Hacker Journalism (PBS): pbs.org/mediashift/2013/05/coding-for-the-future-the-rise-of-hacker-journalism

Further Reading

- Guide to Bad Data Journalism (Andrew Whitby):
prezi.com/pweevqs1hunh/guide-to-bad-data-journalism
- Reading List - Small Data Journalism (Dan Nguyen):
smalldatajournalism.com/readings
- Where to learn about data journalism:
pudo.org/blog/2013/11/13/ddj-resources.html
- Finanzierung von Datenjournalismus (BA Thesis):
geraldgartner.at/wp-content/uploads/2014/06/BA2-2014_Gartner-Kopie.pdf
- Wie wissenschaftlich ist Datenjournalismus? (Weinacht/Spiller):
wpk.org/quarterly/einzelartikel/wie-wissenschaftlich-ist-datenjournalismus.html