Robert Lee Hotz

Santa Fe Science Writing Workshop 2018

Mapping for Story

No one can deny the power of place in our lives. Data mapping is the creative science of where. The narrative power of maps has been exploited extensively by writers, filmmakers, journalists, and environmental activists to tell non-fictional stories. Scholars use maps to better understand how a narrative is placed in a geography, how a geography has informed or influenced an author or how the narrative is ‘locked’ to a particular geography or landscape.

In the expanding universe of Big Data, maps are taking on a new narrative dimension. They can reveal unexpected relationships by linking social data, demographic trends, pollution reports, wildlife sightings, personal memories or emotions to place. By associating survey responses, municipal noise complaints, or fast food delivery orders with geographic location information — a practice known as [geotagging](https://en.wikipedia.org/wiki/Geotagging) — collections of data can be used to illustrate hidden social patterns, as documented by pizza delivery routes, sexual fantasies, music preferences, or the intensity of tourist activity.

In other words, the potential of maps to reveal and tell stories is virtually unlimited.

**Mapping Tools**

One of the most powerful data mapping and spatial analytics software tools is called ArcGIS, produced by esri Inc. It is especially tailored to harness the power of maps for story-telling. They offer a variety of examples, advice and apps.

https://storymaps.arcgis.com/en/

**Tools for working with data**

* Data acquisition: [DocumentCloud](https://www.documentcloud.org/home), [OpenRefine](https://github.com/OpenRefine/OpenRefine), [EveryBlock](https://www.everyblock.com/)
* Analysis: [Excel](https://products.office.com/en-us/excel), [Google Spreadsheets](https://www.google.com/sheets/about/), [SQL](http://www.w3schools.com/sql/), [CSVKit](https://csvkit.readthedocs.org/en/0.9.1/), [R](https://www.r-project.org/)
* Presentation: [Excel](https://products.office.com/en-us/excel), [Overview](https://blog.overviewdocs.com/), [R](https://www.r-project.org/), [Google Fusion Tables](https://support.google.com/fusiontables/answer/2571232?hl=en), [Tableau Public](https://public.tableau.com/), [QGIS](http://www.qgis.org/)
* [National Institute of Computer Assisted Reporting (NICAR)](http://ire.org/files/static_files/video/toolvideos.html): Tools aiding in the reporting and publishing process
* [ScraperWiki](https://scraperwiki.com/): Useful for acquiring, analyzing, and presenting data
* [“Scraping for Journalism: A Guide for Collecting Data.”](https://www.propublica.org/nerds/item/doc-dollars-guides-collecting-the-data) A list of tools for scraping data (extracting information from websites).

**Tools and databases for visualization**

* [Vizzuality](http://www.vizzuality.com/): Cool maps and graphics on a range of topics
* Flowing Data: [http://flowingdata.com/](https://owa.exchange.mit.edu/owa/redir.aspx?SURL=mISqI8Vp4xH5ylyOP9EY0FZOW6LgR6zgcFEA7E1TlVNL8_z0-sLSCGgAdAB0AHAAOgAvAC8AZgBsAG8AdwBpAG4AZwBkAGEAdABhAC4AYwBvAG0ALwA.&URL=http%3a%2f%2fflowingdata.com%2f)
* Andy Kirk: [http://www.visualisingdata.com/about/](https://owa.exchange.mit.edu/owa/redir.aspx?SURL=97UIJtj6t4NSYJZiCGAWxyAuPo-7v6WSu8kRwzKH2LpL8_z0-sLSCGgAdAB0AHAAOgAvAC8AdwB3AHcALgB2AGkAcwB1AGEAbABpAHMAaQBuAGcAZABhAHQAYQAuAGMAbwBtAC8AYQBiAG8AdQB0AC8A&URL=http%3a%2f%2fwww.visualisingdata.com%2fabout%2f)
* Alberto Cairo: [http://www.thefunctionalart.com/](https://owa.exchange.mit.edu/owa/redir.aspx?SURL=DKqgEuUx3GJaOwbJR-tatOhG144fVPnMsj_lJphoCgZL8_z0-sLSCGgAdAB0AHAAOgAvAC8AdwB3AHcALgB0AGgAZQBmAHUAbgBjAHQAaQBvAG4AYQBsAGEAcgB0AC4AYwBvAG0ALwA.&URL=http%3a%2f%2fwww.thefunctionalart.com%2f)
* [D3](http://d3js.org/): interactive / web-based
* [Inkscape](https://inkscape.org/en/): free version of Illustrator (graphic design software program)

**Programming and Computer Science**

* [Hacks/Hackers](http://hackshackers.com/): A journalism organization that seeks to bridge information seeking and storytelling. Journalists, developers, and designers around the world can meet in local chapters.
* [*Open*](http://open.blogs.nytimes.com/)*:* A *New York Times* blog about code and development
* [Stack Overflow](http://stackoverflow.com/): A question-and-answer site for programmers, from beginning to professional skill levels
* [Code Academy](https://www.codecademy.com/): Learn a variety of programming skills and languages, for free

**General databases**

* [UNData](http://data.un.org/): This tool allows you to search for country level statistics on a wide range of topics from across all UN agencies (e.g., UN Statistics Division, FAO, UNESCO, etc.)
* [NICAR](http://ire.org/nicar/database-library/): This database library from the National Institute for Computer-Assisted Reporting (NICAR) covers a variety of topics, including health and the environment.

**Medicine and Health**

* [HealthData](http://www.healthdata.gov/): Access to over 1000 health datasets.
* The World Health Organization (WHO) has numerous databases, including the [Global Health Observatory](http://www.who.int/gho/en/) and [health statistics and information systems](http://www.who.int/healthinfo/en/).
* [Dollars for Docs](https://projects.propublica.org/docdollars/): Search for general payments pharmaceutical and medical device companies made to U.S. doctors in 2013-2014 (ProPublica database)
* [Centers for Medicare and Medicaid (CMS)](https://dnav.cms.gov/Views/Search.aspx) Data Navigator
* [Agency for Healthcare Research and Quality (AHRQ)](http://www.ahrq.gov/research/index.html): Database covering U.S. medical topics including health care, cost of care, trends in hospital care, health insurance coverage, out-of-pocking spending, and patient satisfaction.
* [Global Trigger Tool](http://www.ihi.org/resources/pages/tools/ihiglobaltriggertoolformeasuringaes.aspx): Measures Adverse Events (AEs) in healthcare organizations
* [Leapfrog Hospital Survey](http://www.leapfroggroup.org/survey-materials): Reports on safety and quality performance of U.S. hospitals. Categories include maternity care, high-risk surgeries, and hospital-acquired conditions.
* [ICD10Data](http://www.icd10data.com/): Index of diagnosis and medical procedure billing codes.
* [Health Statistics and Numerical Data subject guide](https://www.nlm.nih.gov/services/Subject_Guides/healthstatistics/) in the U.S. National Library of Medicine provides a diverse list of references for health data around the world.
* [Human Connectome Project](http://www.humanconnectomeproject.org/): National Institutes of Health (NIH) funded project to map neural pathways.

**Environment and Sustainability**

* [U.S. Geological Survey Science](http://data.usgs.gov/datacatalog/) data catalog
* [National Climatic Data Center](http://www.ncdc.noaa.gov/) Information provided by the National Oceanic and Atmospheric Administration
* [Climate Central](http://www.climatecentral.org/): Find information on the impacts of climate change; independent organization of scientists and journalists focused on telling the story of climate change.
* [Climate Data](http://www.data.gov/climate/): Data related to climate change for America’s communities, businesses, and citizens
* [Sustainability Competitiveness Index](http://reports.weforum.org/global-competitiveness-report-2014-2015/the-measurement-of-sustainable-competitiveness/): World Economic Forum (WEF) index focused on the environmental side of economic competitivenes
* [Environmental Performance Index (EPI)](http://epi.yale.edu/): 2014 EPI country rankings and associated data; data explorer; case studies on indicators in practice
* [Global Forest Watch](http://www.globalforestwatch.org/): Satellite derived maps of deforestation data; create custom maps, analyze forest trends, subscribe to alerts, or download data for their local area or the entire world
* [NASA EarthData](https://earthdata.nasa.gov/): WorldView tool allows visualization of near-realtime imagery from NASA satellites related to fires, dust,ash clouds, air quality, drought, floods, etc.; other mapping and visualization tools such as FIRMS for fires; access to more than a dozen NASA data centers and associated satellite data products
* [UNEP Live](http://uneplive.unep.org/): Environmental indicators on multiple issues, at multiple scales, in multiple formats from the United Nations Environment Program.
* [World Resources Institute (WRI)](http://www.wri.org/): A variety of data and information on the environment. WRI is a non-profit global research organization founded in 1982 with members in more than 50 countries.

**Environment and Sustainability**

* [U.S. Geological Survey Science](http://data.usgs.gov/datacatalog/) data catalog
* [National Climatic Data Center](http://www.ncdc.noaa.gov/) Information provided by the National Oceanic and Atmospheric Administration
* [Climate Central](http://www.climatecentral.org/): Find information on the impacts of climate change; independent organization of scientists and journalists focused on telling the story of climate change.
* [Climate Data](http://www.data.gov/climate/): Data related to climate change for America’s communities, businesses, and citizens
* [Sustainability Competitiveness Index](http://reports.weforum.org/global-competitiveness-report-2014-2015/the-measurement-of-sustainable-competitiveness/): World Economic Forum (WEF) index focused on the environmental side of economic competitivenes
* [Environmental Performance Index (EPI)](http://epi.yale.edu/): 2014 EPI country rankings and associated data; data explorer; case studies on indicators in practice
* [Global Forest Watch](http://www.globalforestwatch.org/): Satellite derived maps of deforestation data; create custom maps, analyze forest trends, subscribe to alerts, or download data for their local area or the entire world
* [NASA EarthData](https://earthdata.nasa.gov/): WorldView tool allows visualization of near-realtime imagery from NASA satellites related to fires, dust,ash clouds, air quality, drought, floods, etc.; other mapping and visualization tools such as FIRMS for fires; access to more than a dozen NASA data centers and associated satellite data products
* [UNEP Live](http://uneplive.unep.org/): Environmental indicators on multiple issues, at multiple scales, in multiple formats from the United Nations Environment Program.
* [World Resources Institute (WRI)](http://www.wri.org/): A variety of data and information on the environment. WRI is a non-profit global research organization founded in 1982 with members in more than 50 countries.

**Population and Socioeconomic Data**

* [Population Estimation Service](http://sedac.ciesin.columbia.edu/tools/population-estimation-mapclient): Hosted by SEDAC (the Socioeconomic Data and Applications Center at Columbia University). Draw a polygon around an area (e.g., a cyclone track or toxic release) and get the population in that polygon in real time; iOS app under development
* [Population Reference Bureau (PRB)](http://www.prb.org/): 2014 World Population Data Sheet including interactive map, population clock, etc.
* [Human Development Report Data](http://hdr.undp.org/en/data): Human Development Index; Public Data Explorer; Multidimensional Poverty Index
* [Socioeconomic Data and Applications Center (SEDAC)](http://sedac.ciesin.columbia.edu/): Gridded population data, poverty maps, infrastructure data (dams, nuclear plants, roads), PM2.5 maps, national estimates of population and land area by climate/biome/elevation (PLACEv3); map gallery available under Creative Commons licenses; country treaty participation data
* [World Bank Data](http://data.worldbank.org/): Includes the World Development Indicators and free and open access to data about development in countries around the globe; thematic portals for agriculture and rural development, climate change, environment and urban development
* [OECD QWIDS](https://stats.oecd.org/qwids/): International development statistics; OECD is the France-based Organization for Economic Cooperation and Development.
* [TIGER data](https://www.census.gov/geo/maps-data/data/tiger.html): From the U.S. Census of population.

**Regional resources**

* [Ekuatorial](http://ekuatorial.com/en/datasets): Indonesia specific datasets on marine, forests, and natural disasters
* [InfoAmazonia](http://infoamazonia.org/datasets/): An aggregation of environmental datasets for the 9 countries of the Amazon basin including forests, watersheds, industries and indigenous lands
* [Open Development Mekong](http://opendevelopmentmekong.net/): Development tracker focused on the Mekong
* [Third Pole Data Network](http://data.thethirdpole.net/): An open source geospatial database: a simple, searchable catalog of water-related datasets sourced from leading organization monitoring water in Asia

*(These resources were compiled during The Kavli Symposium on Science Journalism: Data Mining, with the assistance of the MIT Knight Science Journalism Program)*