Telling Data Stories Together

Final Report of BU Datalabs

A Fusion CCCP & Civic Media Hub Project

Faculty of Media and Communication
Bournemouth University

2015
Carried out between January 2015 and July 2015, the BU Datalabs project sought to establish a sustainable, iterative model for co-creating effective data stories on sensitive subjects designed for generating societal impact. We worked toward this aim by exchanging knowledge and skills across the fields of journalism, media and communications, design, geography and data science.

Using a hands-on participatory approach to knowledge exchange, we brought together researchers, students, journalists and NGOs to explore the possibilities and challenges associated with the communicative power of visualizations, maps and related interactive digital media for telling data stories on sensitive subjects of societal importance.

Through a series of innovative Datalabs events and workshops that prioritised collaborative exchange, we were able to generate key insights and co-create tangible outputs. In this report we introduce readers to the emergent area of data storytelling and offer a pragmatic overview of our Datalabs process. Covering the opportunities and challenges our workshop model provided, we also share some key insights that arose out of our collaborations. We encourage other universities and organisations to use our model as a template, enabling your organisation to co-create data stories with public stakeholders.
THE RISE OF DATA STORYTELLING

With the growth of open data and big data, it has become more important than ever to understand and access the datasets made available. However, datasets are often large and complex, making it difficult to interpret and analyse. More and more people are turning to visualisations as a way to both interpret and communicate with data.

Adding visual narrative to numbers and bringing storytelling techniques into working with data can help make research and complex information come alive. This increased emphasis on visualising data has brought with it more attention to the importance of understanding the principles of design and storytelling across a range of fields. Storytelling with data involves implementing design techniques and expressive practices of visual narration in order to give data context and meaning in ways that connect to different audiences.

Over the past few years we have seen a rise in the number of software tools, handbooks and online resources available for such data visualisation. Yet, two major challenges currently face those wanting to tell data stories. In terms of audiences, users come with all types and levels of personal skills, education, and tastes. In more traditional areas where data visualisations are used, such as economics or scientific visualization, graphics are intended for highly trained audiences. However, in the growing area of data visualisation, images are used to communicate with everyday audiences and reach the general public (Gershon et al 2001). This poses challenges for designers who are often new to the tools and techniques of data visualisation.

Like any new digital skill, visualising data requires time and resources. This puts NGOs and small organisations at a disadvantage in terms of their ability to tell compelling data stories. This disadvantage manifests itself in the lack of resources to afford certain technologies, as well as to afford know-how in the form of data visualization designers that are being “snatched up” by the corporate sector to work under Non-Disclosure Agreements (Wilson 2015). This basically means that these highly talented in-house data visualisation designers will most likely never be able to share their ideas and work with the public and therefore hinder the development and the innovation for new ways of data storytelling. In addition, while the rise of digital tools and techniques certainly diversifies the kinds of narratives we can tell with numbers, not all data stories are easy to capture. Sensitive subjects often have no straightforward data source, documents are scattered across agencies and organisations. Moreover, data on sensitive topics is often kept hidden, deemed too confidential to be made open.

This ‘uneven transparency’ raises important questions about the duty to document (Larsen and Walby 2012), particularly in regard to issues of security where obtaining health and human rights information on vulnerable populations (prisoners, detainees, those living in conflict zones) becomes difficult, trumped by a greater interest in keeping data protected for national interests.
RESEARCH INTO DATA STORYTELLING

The novelty of data storytelling tools and practices means that most discussions regarding the expanding field of data storytelling takes place in industry forums and blogs, rather than in academia. Thus far, there has been scarce academic research addressing data visualization and storytelling. What has been published is mostly based on critical articles, rather than primary research (Gershon et al 2001, Lancy 2001, Amin 2002, Anderson 2008, Crampton 2009, Elwood 2009, Kosara et al 2013). Most of the research that does exist is either in the field of geography or social media.

In relation to the field of Geography, digital technologies like Google maps have transformed the ways people relate to visual representations of geographic spaces. From zoom functionality on street view to ArcGIS StoryMaps platform, people are engaging their geographical imaginations to produce their own visual representations of the spaces that matter to them. Likewise, utilising the web and social media, people use RSS feeds and APIs to draw information from data sources to create real-time maps (Crampton 2009). Such interactives are often populated with user-generated or crowdsourced content, engaging the activity of citizen scientists and citizen journalists. These practices open up new opportunities for civic authorship and challenge the idea of cartography as fixed, objective and authoritative (Dodge and Kitchen 2013).

DATA LABS PROJECT AIMS

Addressing this emergent field of data storytelling, the aims of our project were to:

• Bring together a multidisciplinary, cross-school team of researchers and students to work in collaboration with external stakeholders

• Establish a sustainable training model for turning data-based research into effective communication resources designed to create societal impact.

• Co-create effective ways of communicating and visualizing complex data sets targeted at policy-makers and journalists to maximize impact.
Our team at Bournemouth University worked with our NGOs, journalists and digital designer stakeholder partners to run hands-on data aggregation, visualisation and digital storytelling workshops designed specifically for addressing civic and humanitarian issues.

Dr Anna Feigenbaum, PI
Senior Lecturer in Digital Storytelling

Dr Einar Thorsen, CI
Principal Lecturer in Journalism and Communication

Dr Shelley Thompson
Senior Lecturer in Corporate and Marketing Communications

Dr Phillipa Gillingham, CI
Senior Lecturer in Biogeography

Dr Dan Jackson
Principal Lecturer in Corporate and Marketing Communications

Dr Nathan Farrell
Lecturer in Communication and Media

Dr Hamid Bouchachia
Associate Professor, School of Design, Engineering and Computing

Dr Brad Gyori
Senior Lecturer in Digital Storytelling

Naciye Demirkol
Research Assistant

Daniel Weissmann
Research Assistant
Coordinated by John Horne, a doctoral candidate at Birmingham University and organiser at Bahrain Watch, our diverse stakeholder partners came from journalism organisations, digital design teams and human rights NGOs.
BAGGAGE
A Brixton based community initiative monitoring neighbourhood gentrification and housing struggles.

Bahrain Watch
An independent organisation that seeks to promote democracy, equality and social justice in Bahrain, through evidence-based research investigations and advocacy.

Bureau of Investigative Journalism
Research, investigations, reporting and analysis which is of public benefit by undertaking in depth research into the governance of public, private and third sector organisations and their influence.

Campaign Against Arms Trade
UK-based organisation working to end the international arms trade.

Corporate Watch
Independent research group, investigating the social and environmental impacts of corporations and corporate power.

Corruption Watch
An investigative organisation that details and exposes instances of corruption and their subjective and objective impact on democracy, human rights and development across the world in order to precipitate strong action against it.

CVG Design
A New York based Graphic Designer specialising in community art as urban intervention.

IRIN News
Independent news agency delivering unique, authoritative and independent reporting from the frontlines of crises to inspire and mobilise a more effective humanitarian response.

Jay Cassano
Independent journalist covering the intersection of technology and politics.

Julio Molina Montenegro
An award-winning filmmaker with over ten years of international experience, Julio Molina combines production work with lecturing in the fields of documentary, sound recording, post-production and media technology.

Minute Works
A multidisciplinary graphic design studio specialising in sustainable communications.

Omega Research Foundation
Non-profit research foundation providing rigorous, objective, evidence-based research on the manufacture, trade in, use of, military, security and police (MSP) technologies.

Reported.ly
A global team of journalists covering stories of international importance through social media and citizen networks with a focus on social movements, civil rights, conflict and human rights.

Small Media
An action lab helping the free flow of information and creative expression in closed societies, with training, technology and research initiatives that focus in Iran.

Tactical Technology Collective
Non-profit organisation, working since 2003 to advance the use of information and digital technologies by advocates and activists worldwide.

Tear Gas Research Connection
Research project bringing together existing knowledge around tear gas and the impacts less lethals have on people and their lived environments. Working with international researchers, NGOs, journalists and tactical technologists, we contribute to news reports, public debate and policy-making on the safety and social impacts of tear gas technologies.
As a team, we hosted three 'Datalab' workshops on-site at Bournemouth University. Each Datalab ran for two days, consisting of an interactive, Masterclass and a discussion-based workshop on Day 1, followed by a guided computer lab session on the afternoon of Day 1 and into Day 2. During these lab sessions, students, staff and stakeholders worked on live datasets using free and open source software for data storytelling.
DATA SCRAPING & CLEANING

Monday March 16 and Tuesday March 17, 2015
This Masterclass and workshop introduced participants to tools and techniques for scraping data from the web. We went over how to format and organise data in spreadsheets. Discussing basic methods for data storytelling, the workshop provided training in how to find data stories within datasets.
Datasets: TheyWorkForYou.com data on MP profiles
Tools: import.io & openRefine

DIGITAL STORYTELLING WITH DATA

Thursday May 14 and Friday May 15, 2015
In this Masterclass and workshop session we introduced participants to the emerging interdisciplinary field of data storytelling. We explored different, emerging forms of digital narratives and discussed the principles behind storytelling with data for different audiences. In the workshop session we reflected on the ethical questions that come with turning numbers in narratives. Then we had a hands-on lab session digging for stories in a data set that could be visualised using the tools we had learned at the first two Datalab sessions.
Datasets: MP travel to MENA region
Tools: import.io, openRefine, CartoDB

DATA VISUALISATION WITH MAPPING

Monday April 27 and Tuesday April 28, 2015
This Masterclass and workshop introduced participants to a variety of mapping techniques and technologies. Offering a beginner lesson in the principles of GIS, participants learned how to work with base maps, data layers and symbolizing to present spatial representations of data. On Day 2, participants were introduced to social network mapping, using twitter APIs to map the use of hashtags.
Datasets: UK Parliamentary Constituencies
Tools: QGIS & CartoDB
DATALABS
PUBLIC EVENTS

In addition to our Bournemouth-based hands-on Datalabs workshops, we also hosted a series of events both at Bournemouth University and in London. These events were designed around the project aims, bringing together a wider network of partners from the local and national community.
INTERDISCIPLINARY RESEARCH WEEK—TURNING NUMBERS INTO NARRATIVES: AN INTRODUCTION TO DIGITAL STORYTELLING WITH DATA

13 July 2015, Talbot Campus, Bournemouth University
In this interactive workshop, we invited charities, local businesses and local councils, introducing them to the opportunities and challenges of data storytelling. We provided insight into visual design, online user experience and digital communication. With 30 participants from local businesses, advertising agencies, charities and the Poole council, the workshop was well attended. We worked with live data from participants’ organisations, doing hands-on activities to enhance how we tell data stories in meaningful and impactful ways.

MAPPING FOR JUSTICE

2 June 2015, with Richmond, American University in London
In this public showcase and discussion we explored how mapmakers engage our geographical imaginations and use the power of maps for social change. From oral history to twitter data-mining, our featured mapmakers use a variety of techniques to make injustice visible. We discussed how, whether made with paper and pen or on open source platforms, maps can give way to new tactics and strategies for intervention. Attended by 80 people, this public engagement event co-hosted with RAUL brought creative, critical and participatory mapping practices to a broad audience.

FESTIVAL OF LEARNING—DATASTORYTELLING: HOW TO GIVE NARRATIVE TO NUMBERS

14 May 2015, Talbot Campus, Bournemouth University
In this Masterclass, we introduced the audience to the emerging interdisciplinary field of data storytelling research. Drawing on a range of contemporary examples, we presented a variety of techniques and technologies for creating data stories. With over 50 participants from across all faculties of the university, as well as outside, our Interdisciplinary Research Week event drew together insights from the fields of communications, journalism, geography, psychology and data science.
CREATING THE DATALABS FORMAT

We imagined our Datalabs to take both the form of hackerspace and of an open-space learning environment. Combining the collaborative, problem-based dimension of the hackerspace with the lab format for learning, we aimed to give participants hands-on experience and insight into the use of data storytelling tools and techniques.

DATALAB AS HACKERSPACE

Traditional hackerspaces were first established in Germany in the late 1980s as "places in the community where local programmers can collectively meet, work, and share infrastructure" (Borland, 2007). Because of the community of programmers skill levels and skill sets were relatively homogenous (Cavalanti, 2013). This level of homogeneity created a high creative output as the programmers could work creatively with a community of similar skill sets, mind-sets and languages. With the hackerspace in mind, we first brought our partners and participants together to present the tools we would be using in the hackerspace format the following day.

However, unlike a hackerspace, where participants can be assumed to have a shared set of skills and experiences with IT, at Datalabs there was a broader range of abilities. This meant that there was a need for scaffolded tool training (Hmelo-Silver et al., 2007). We took participants on a step-by-step guided lesson in how to use data gathering and visualisation tools at the beginning of every workshop.

CHALLENGES & OPPORTUNITIES

McLoughlin and Lee highlight that especially during interdisciplinary workshops “learning focused on knowledge creation and networking, offers the potential for transformational shifts in teaching and learning practices, whereby learners can access peers, experts, the wider community, and digital media in ways that enable reflective, self-directed learning” (2008 p19). This aspect of a transformational shift in learning practices was very important as we had to guide our partners and stakeholders not only on the same skill level, but also to a shared understanding of the language and terms used across fields of digital media, GIS and data science.
Even some basic terms take on different meanings across these approaches to working with data. For our partners and stakeholders to communicate efficiently within our interdisciplinary workshop environment, definitions were key.

Although it comes with challenges, this focus on knowledge creation and networking is facilitated by personalisation, participation and productivity (McLoughlin and Lee 2008). There is personalisation in the form of learner choice, learner agency and customization. The open-space aspect of Datalabs provided these features, particularly during the day two sessions where participants could work autonomously, or in self-selected groups, on their own datasets.

Furthermore, autonomy did not hinder participation. As there was no instructor on the second day, Datalabs participants encouraged communication and collaboration. This autonomous, hackerlab style time and space, encouraged partners and stakeholders to test and apply their knowledge and skills gained during the scaffolding learning sessions. Here they moved beyond using the provided sample datasets to generate and experiment with their own datasets both individually and in collaboration with other Datalabs participants.

McLoughlin and Lee also emphasise that the purpose of Pedagogy 2.0 “is to enable self-direction, knowledge building, and learner control by providing options and choice while still supplying the necessary structure and scaffolding” (McLoughlin and Lee 2008 p17). As the community “spend[s] time together, they typically share information, insight, and advice. They help each other solve problems. They discuss their situation, their aspirations, and their needs. They ponder common issues, explore ideas, and act as sounding boards.” (Wenger et al., 2002 p4).

This description by Wenger was reflected in our experience during the unguided, open-space learning during day two of Datalabs. Overall, and as evidenced in participant feedback, we found that although the mixed format of scaffolded learning on day one plus open-space learning on day two seemed quite challenging at the beginning, once participants became comfortable with the format, the blend of taught and autonomous lab sessions worked well. As few similar workshops have been done in a similar fashion.
Feedback from our survey of Datalabs participants showed that most people’s familiarity with the tools and concepts taught during the workshops increased substantially, with most participants becoming familiar or experienced by the end of their sessions.

Whereas, in terms of concepts and skills participants found it most beneficial how to locate stories within data for storytelling.

In terms of the tools participants found CartoDB and Import.io most beneficial.
Our stakeholders’ outputs varied as different organisations have different needs in terms of tools and techniques they want to use. Here, two organisation stand out as a good example on the different demands and use of tools and techniques provided by the Datalabs workshops.

**Omega**
Omega’s focus is on the manufacture, trade and use of military, security and police technology. They work with news feeds, arms expo data and trade data. This made tools such as Import.io for data extraction and Open Refine for cleaning datasets, most beneficial for them. With these new tools, they are now pulling data from exhibition lists and off of RSS newsfeeds, and comparing datasets from various sources.

**IRIN News**
IRIN is a news organisations focused on crises and humanitarian catastrophes. IRIN News also can make use of the tools for extracting and cleaning datasets, however as they mainly use existing statistics from humanitarian organisations, their primary aim is to narrate and contextualise data, embedding visual and interactive elements into their stories. This helps build audience engagement with what they are reporting. This visualisation and narrative aspect of data storytelling makes tools such as CartoDB most suitable.
Creating visualisations and stories with data comes with a number of challenges. Through our experience with the Datalabs project, we were able to confront, explore and attempt to address these challenges, leading to crucial insights into both how we think about data storytelling, as well as how we put it into practice together.

**DIFFERENT TOOLS FOR DIFFERENT STAGES**

Teaching tools and techniques does not only mean training people in the use of each individual tool, but also, helping people understand how these tools are used in combination with each other along the process of data-driven storytelling. Different stages of the data storytelling process require different tools. Sometimes there is a linear flow—from scraping and cleaning to visualising—but other times you must go back to your dataset and dig for new stories or angles, to look from new perspectives. This requires a flexible approach to the use of different tools at different stages, as well as the ability to identify which tool or technique is most appropriate for the data available.

**CHOOSING THE RIGHT TYPE OF VISUALISATION**

Different data stories require different visual languages and storytelling techniques. Not every dataset can be or should be visualised into a map just because it has geo-locations in it. Nor does a good visualisation necessarily have to be digital. An analogue visualisation can be just as powerful as a digital one if not more so (Cohen 2015; Rivas 2015). The challenge is not only to be able to use the tools provided sufficiently, but also to know when to use certain tools and what its limitations are for the story you are trying to tell.
DISTORTED DATA

Perhaps the biggest data-oriented challenge in data-driven storytelling is data distortion. In every stage of the data storytelling process, from gathering information to circulating a visualisation on social media, distortion can come into play. This might arise from missing data, mis-recorded data, or information displayed out of context. It can happen when designs exaggerate the sizes of bubbles or bars. Or distortion might occur in the analysis, where biases shape the interpretation of findings and this becomes codified in the data story. For this reason we believe that transparency around data storytelling processes and data sources is key. Where it is not a privacy concern to share information, publishing a guide to your process of data storytelling, along with the raw datasets you used is good practice and helps the data storytelling community reflect and grow.

VISUALISING SENSITIVE SUBJECTS

In addition to these data challenges, working with data that deals with human rights abuses, corruption, torture and crisis, provides important insights into the particular challenges that telling data stories about sensitive subjects brings to this emerging area of communications. For these sensitive issues, in which the subjects represented already face difficulty in having their voices heard and experiences shared, storytelling in meaningful ways comes with increased risks.

When working with charts, maps, iconography and pictograms, it is particularly challenging to represent and visualize complex issues like death, torture or suffering in a way that is both accurate and ethical. When giving visual narrative to such sensitive topics, it is crucial to approach visualization with respect for those affected, to protect privacy when necessary, and to avoid turning victims into numbers in ways that lose their humanity.
WORKS REFERENCED


