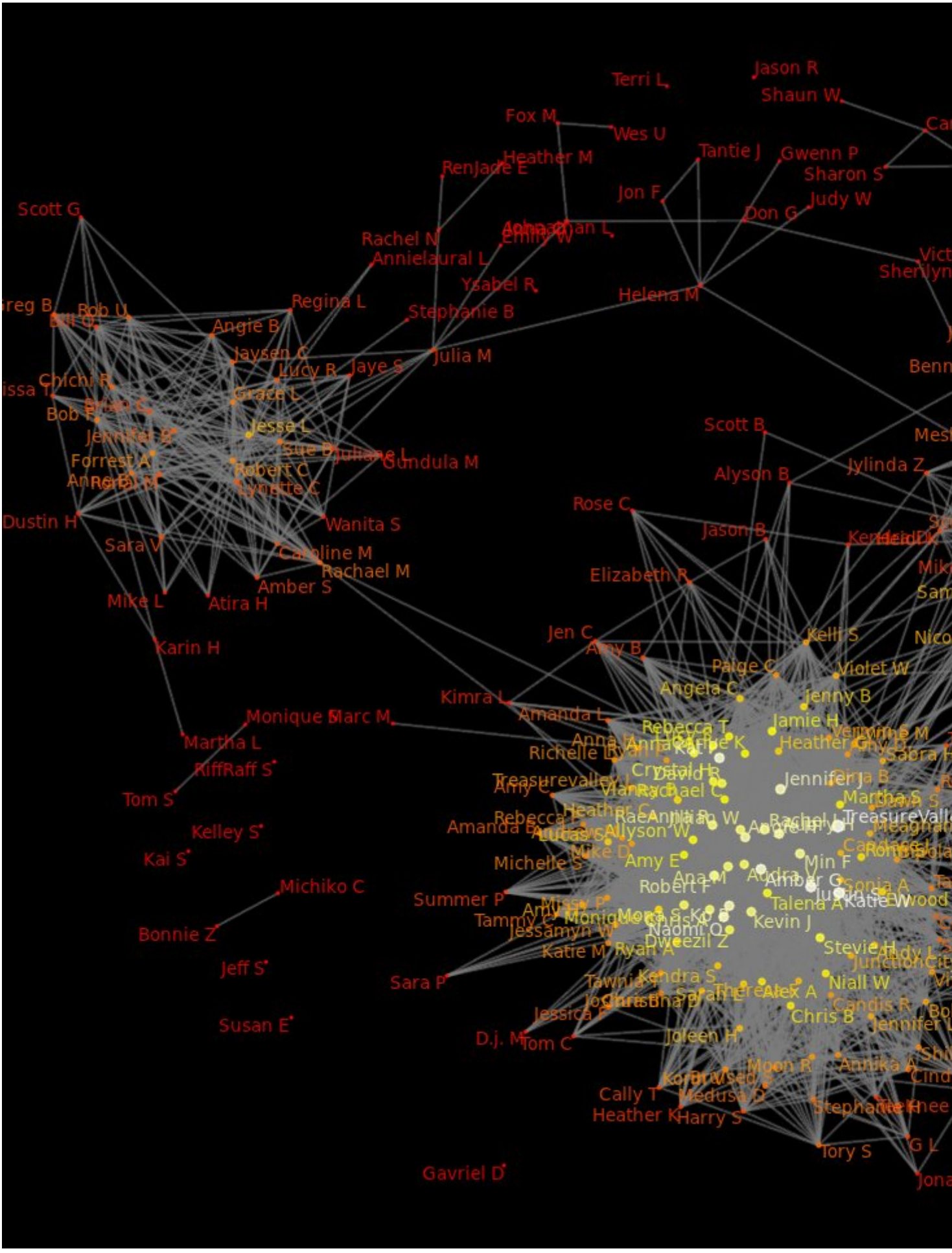
Thumbnail Image:


Terri L
Jason R
Shaun W
Fox M
Wes U
Heather M
RenJade E
Tantie J
Gwenn P
Sharon S
Jon F
Judy W
Don G
Scott G
Rachel N
Emily W
Annielaural L
Ysabel R
Helena M
Regina L
Stephanie B
Angie B
Julia M
Jaysen C
Lucy R
Jaye S
Chichi R
Grace L
Jesse L
Sue B
Juliane L
Gundula M
Robert C
Lynette C
Forrest A
Dustin H
Wanita S
Sara V
Caroline M
Rachael M
Amber S
Mike L
Atira H
Karin H
Monique S
Marc M
Martha L
RiffRaff S
Tom S
Kelley S
Kai S
Michiko C
Bonnie Z
Jeff S
Susan E
Scott B
Alyson B
Jylinda Z
Rose C
Jason B
Elizabeth R
Jen C
Amy B
Kelli S
Violet W
Paige C
Angela C
jenny B
Kimra L
Amanda L
Jamie H
Rebecca T
Richelle B
Heather K
Crystal H
David
Jennifer J
Amy C
Rachael C
Rebecca L
Amanda B
Allyson W
Michelle S
Amy E
Min F
Ana M
Robert F
Amber G
Sonia A
Summer P
Talena A
Tammy C
Kevin J
Naomi O
Dweezil Z
Katie M
Ryan A
Stevie H
Sara P
Kendra S
Niall W
Alex A
Chris B
Jessica P
D.j. M
Tom C
Joleen H
Moon R
Annika A
Cally T
Medusa D
Heather K
Harry S
Tory S
Gavriel D

Aug 29, 2024 by Ali Fisher

Data-Driven Diplomacy: A Practical Guide [1]

Diplomats and scholars now have the opportunity to gather and analyze vast quantities of data. Data has allowed us to explore how social media users responded differently to a variety of international broadcasters in the wake of the use of chemical weapons in Syria and the actions of conservative cyber-activists in Iran. Data science can empower diplomats to evaluate a season of programs from an international broadcaster, the coverage of a global summit about ending sexual violence, or develop more ways to leverage an international sport event such as the London 2012 Olympic games.

One way this data can be put to use is by creating three-dimensional network visualizations which can be viewed on a tablet. Another is to offer insight into the complex systems which make up the Jihadist Swarmcast (a dispersed network of social media accounts which provides resilience and persistence for jihadist content as part of a fully integrated multiplatform social media zeitgeist). This insight can empower public diplomats to develop strategies to challenge the media mujahedeen.

As highlighted by "Data-Driven Public Diplomacy," published by the U.S. Advisory Commission on Public Diplomacy, it is becoming increasingly important for both the scholarship and practice of diplomacy to be able to integrate the insights available from data science. This will allow data analysis to be incorporated into the workflow, to be used in planning and program design, and be applied to the systematic understanding of the extent to which programs engage and influence foreign audiences.

It would be absurd to suggest that diplomacy should be conducted only on the basis of big data. However, it would be equally absurd to conduct public diplomacy without using big data when it is available. The greatest opportunity for influence comes from the visualization and synthesis of big data insights with the nuance, experience, and understanding developed by generations of diplomats.

Moving forward with data-driven public diplomacy will be a collaborative process between those with data science expertise and those with detailed "domain" knowledge about the practice of public diplomacy. An understanding of the nature of data science and the roles diplomats and data scientists would play at each stage of research and analysis is critical to success.

What is Data Science?

"Data science requires skills ranging from traditional computer science to mathematics to art," according to Mike Loukides. In his valuable *An Introduction to Data Science*, Jeffrey Stanton writes:

> Data Science refers to an emerging area of work concerned with the collection,

preparation, analysis, visualization, management, and preservation of large collections of information. Although the name Data Science seems to connect most strongly with areas such as databases and computer science, many different kinds of skills - including non-mathematical skills - are needed.

The potential of data science is to draw large data sets into the study and practice of diplomacy, and allow diplomats and scholars to become comfortable engaging with and analyzing increasingly large and often unstructured data. As Google Chief Economist Hal Varian claimed, "the ability to take data—to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it—that's going to be a hugely important skill in the next decades."

The challenge of data science is often characterised by three Vs: volume, velocity, and variety – with the veracity of data being a common fourth 'V'. This data can come from social media or other huge data sets, such as the quarter billion record database GDELT (Global Data on Events, Location, and Tone).

However, the use of data science must maintain a focus on issues meaningful for diplomacy – and provide insights relevant to diplomacy. The focus must be on the diplomatic rather than the data challenges. To this end, the use of data science in diplomacy will often require teamwork between computer or data specialists and subject matter experts.

Defining roles

As team members are often significantly more familiar with one part of the project than the other, defining clear roles can help ensure that important aspects of research do not fall down the cracks between data specialists and subject matter experts.

Jeffrey Stanton's "four A's of data" are useful categories through which team members can identify the roles and responsibilities at each stage of a research project:

- *Architecture* refers to the form and format of the data and the design of the research.
- *Acquisition* includes the identification, collection, and storage of data, including how that data will be presented for analysis. This consideration is particularly important when the subject area specialist is not also technically capable of accessing the data from the format in which it is being stored.
- *Analysis* combines data science elements such as the summarizing the data, using samples to make inferences about the larger context, and data visualization. It also includes the process through which data will be made available to the subject matter expert, in this case diplomacy, to be analyzed within the context of that field.
- *Archiving* refers to preservation, long-term storage and, where appropriate, making the data reusable.

Laying out the responsibility at each stage of the project can help ensure that a data science approach can best contribute to greater understanding of public diplomacy based on questions developed in collaboration with subject matter experts. The table below is intended to be illustrative rather than exhaustive.

Area of expertise	Project phase			
	Architecture	Acquisition	Analysis	Archive
Diplomacy	What questions are relevant to diplomacy?	Provide rationale for the data being collected	Interpret data in the context of diplomacy.	Will ongoing access to data be required?
	Where is data relating to diplomacy?	Nuance between easily found data and data highly relevent to core elements of the field.	How do results contribute to contemporary theories of diplomacy?	Will publications require raw data to be available?
	What do you want to be able to do with the data?	Domain specific data (such as appropriate search terms)	What conclusions can be drawn and what further research questions do these results suggest?	Is any aggregated form of the data required?
Data science	In what form is the data available? How robust is the data? Is the data of appropriate size?	Provide rationale for the type, size of data and collection method	Machine driven analysis and visualisation	Practical long-term storage in a form that data can be reused. Where physically will data be stored? In what format is the data stored?
	What privacy concerns need to be considered?	Conduct practical data collection and storage	Machine learning, such as text analysis and natural language processing	Can the raw data be shared with others, are there limitations imposed by the terms of service?
	What bias might the focus on this data introduce?	Deliver data for analysis & guide domain specialists to sections of the data.	Statistical models and tests - e.g. regression, cluster, and classification, or social network analysis.	Does the data have to be stored securely? Can it be accessed by others? Will it require specialist knowledge to access?

The availability of data is pushing the boundaries of what is imagined possible in public diplomacy. However, while tools to analyze data have expanded rapidly and users can now search large amounts of data quickly and efficiently, it is still up to subject matter experts to interpret the results and the public diplomats who bridge the last three feet to deliver results.
