A Framework for Data Visualization: Foundations and Threshold Concepts

Delores Carlito and Dorothy Ogdon

University of Alabama at Birmingham Libraries

Background

In 2016, the Association of College & Research Libraries replaced the Information Literacy Competency Standards with the [Framework for Information Literacy for Higher Education](http://www.ala.org/acrl/standards/ilframework). The old standards were prescriptive, but the new framework broadened and recast information literacy for multiple information paradigms. The new framework is conceptually rooted in establishing threshold concepts for information literacy. Threshold concepts are ideas in a discipline that are transformative, irreversible once learned, integrative, and bounded, yet can be troublesome to learn because they might seem counterintuitive or alien (Meyer and Land 4-6).

 Our goal in participating in the Visualizing the Future symposia was to develop a framework of threshold concepts necessary for both creation and evaluation of data visualization. We foresee this as a tool that could be used cross-discipline to inform both instructional practices around data-related topics and to provide guidance to libraries and information-centric disciplines in the development of instruction, workshops, collections, and technology-enabled workspaces.

Literature Review

We began our research with a literature review. We also examined and began mapping known literacies. Literacies we identified as having the most in common with visual representation and evaluation of data are: information literacy, visual literacy, digital literacy, multimodal literacy, metaliteracy, and quantitative literacy. We annotated many works to determine threshold concepts in data organization, representation, and evaluation. We will conduct a survey of faculty and staff who provide instruction and training to students to determine what digital tools are utilized within the UAB community for the purposes of data visualization and the creation of graphic information. We will use this information to develop a detailed concept matrix that maps literacies to data-related activities in student instruction and training.

Preliminary Findings

We have identified four frames that describe data, its behaviors, and its characteristics in the contexts of visualization and communication. We and describe the frames below, in no particular order.

*Data and its visualization can be manipulated to achieve a desired result*.

Data is an information object unto itself. It can be manipulated to achieve a positive result, such as conveying important information, or a negative result, such as promoting disinformation. Users must not only be creators, but critics. With data visualization, the data can take on both representational and generative properties with regard to the dissemination and creation of knowledge (Drucker 65). It can be affected by the transmitter, noise source, or receiver as communication schema are described in Shannon’s “Mathematical Theory of Communication.”

* There are Layers of meaning to any data visualization.
* Visual cues such as font, color, and organization, can affect understanding and perceived validity.
* The form of the data’s input affects what can be done with the data at the creation stage.
* Bias can affect how data is interpreted.
* Data must be viewed in context.

*Data visualization is information, and information has monetary value*

As culture is changing, Frederick Douglass’s idea that freedom comes from reading must be extended to other forms of communication. Information is now currency, and information has a monetary value. According to D’Ignazio, “data has become a currency of power,” and “emancipation [is] achieved through the literacy process” (6). Virginia Eubanks notes “marginalized groups face higher levels of data collection” (6), resulting in a scenario in which the people most impacted by mandatory and automated data collection and interpretation are often those least likely to have a voice in use of that data.

* Financial success is determined by access to and an understanding of information.
* The way information is visualized, and the information, itself, may be presented to monetarily reward one social group over another.
* Information in the form of data is the foundation for automated decision making. Automated decision-making is the next wave of technology innovation, but the ability to build models that are meaningful is limited by an individual’s ability to understand the meaning of both a model and its output. A data visualization can be made to fit any model with enough manipulation.

*Content is culture- and community-based*

Data visualization is a form of communication, and members of a culture or community must have access to the language of delivery when in the role of the receiver in communication models. Social semiotics can affect how data is interpreted by receivers in their social and cultural group. Data visualization can facilitate social interaction or alienate a group.

* Prior knowledge can be necessary to visualization interpretation.
* Visual cues can affect attitudes toward the data representation and interpretation
* Social information is explicitly or implicitly revealed in the way the visualization is presented.
* The meaning of content may drift when the content is experienced outside of the original creator’s frame.

*Data visualization tells a story*

Humans relate to the world through story, as exemplified in many commercials, YouTube videos, and Tweets. Data visualization is a rhetorical argument with visualization and form and data as content. The term “data storytelling” may be used to define communicating through visualizations, and “People hear statistics, but they feel stories” (Dykes). Data storytelling is an essential learning concept and a key qualitative skill necessary for proficiency in data visualization (Ryan 101).

* Visualization can reinforce previous knowledge or be used as a rhetorical device.
* Visualization is a form of multimodal rhetoric, combining visual, data, and information tools to produce a message.

Implications

* A creator or user interacting with a data visualization has several roles: user, designer, creator, presenter, communicator, and evaluator.
* Visualization development is an evolving skill and practice that applies across disciplines and professions; the potential for the application of data visualization and related skills are not limited.
* The role of the library is to facilitate the development of proficiencies in abstract thinking and knowledge synthesis by developing an understanding of design as manipulation and interpretation of data.

Going Forward

We plan on surveying professors about their students’ needs, and we will use these results to design workshops in the Center for Teaching and Learning. These workshops will focus on introducing a framework-based foundation for developing data and data visualization literacies.

Works Cited

D’Ignazio, Catherine. “Creative Data Literacy: Bridging the Gap between the Data-Haves and Data-Have Nots.” Information Design Journal, vol 23, no. 1, 2017, pp. 6–18. John Benjamins, doi: 10.1075/idj.23.1.03dig.

Drucker, Johanna. Graphesis: Visual Forms of Knowledge Production. Harvard UP, 2014.

Dykes, Brent. “Data Storytelling: The Essential Data Science Skill Everyone Needs.” Forbes, 31 Mar. 2016. www.forbes.com/sites/brentdykes/2016/03/31/data-storytelling-the-essential-data-science-skill-everyone-needs/#3aa32e1052ad.

Eubanks, Virginia. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin’s, 2019.

Meyer, Jan, and Ray Land. Threshold Concepts and Troublesome Knowledge: Linkages to Ways of Thinking and Practising within the Disciplines. Occasional Report 4, ETL Project, May 2003. www.etl.tla.ed.ac.uk/docs/ETLreport4.pdf.

Ryan, Lindy, et al. “Teaching Data Visualization as a Skill.” *IEEE Computer Graphics and Applications,* vol. 39, no. 2, 2019, pp. 95-103. *IEEE Xplore*, doi: 10.1109/MCG.2018.2889526.

Shannon, C.E., “A Mathematical Theory of Communication.” Bell System Technical Journal, Vol. 27, Jul., Oct. 1948, pp. 379-423, 626-656. Harvard Mathematics Department, math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf.