

How many ways can we teach data literacy?

Yun Dai¹

Abstract

Academic Libraries are ideally positioned to teach data literacy. What is 'data literacy' in the first place? Is it the new information literacy? Will the ways we teach information literacy limit imaginative ways to teach data literacy?

With those questions in mind, the Library of New York University Shanghai has explored multiple ways to teach data literacy to undergraduate students through university events, 'for-class' instruction and workshops, and online casebooks. (1) We initiated the yearlong series of events titled 'Lying with Data', inviting faculty across disciplines to each address one core data literacy question that students of data science may misunderstand (2) We offered workshops and in-class instruction that are up-to-date with the latest technology and that fit with the curriculum. (3) We created online casebooks on various topics in the data lifecycle, tackling user needs at different levels. Essential to our teaching activities are two core values: 'let the quality speak for itself', and 'outreach by teaching'.

Keywords: data literacy, information literacy, instruction, data services, academic library

Introduction

Academic Libraries are ideally positioned to teach data literacy. What is 'data literacy' in the first place? Is it the new information literacy? Will the ways we teach information literacy limit imaginative ways to teach data literacy?

With those questions in mind, the Library of New York University Shanghai² has explored multiple ways to teach data literacy through university events, 'for-class' instruction and workshops, and online casebooks. Our library's data services unit is structured in the larger 'data services ecosystem' of the university, in which we have found our niche as data literacy advocates. Our programming complements course offerings and formal scholarly activities, seeking to bridge the gap between those with rising interest but little knowledge in data and the seasoned data practitioners. We also work with university-wide platforms to promote literacy programs to the more general audience.

While the data services program started to take shape since 2017, we have been broadly experimenting with forms of data literacy instruction. Some of these instructional strategies, such as in-class instruction and workshops, are common practice in academic libraries, while others may be less conventional. Specifically, we initiated the yearlong series of events titled 'Lying with Data', in which we invited faculty across disciplines to each address one core data literacy question that students may misunderstand We also offered for-class instruction and workshops that are up-to-date with the latest technology and that fit with the curriculum. Outside of the classroom, we created online casebooks on various topics in the data lifecycle, tackling user needs at different levels. Essential to our teaching activities are two core values: 'let the quality speak for itself', and 'outreach by teaching'.

In the following sections, I share the ideas that we experimented with, the rationale behind them, and the lessons we learned from these experiences. The bulk of the paper focuses on the event form of data literacy activities, which is believed to be the format less studied.

Revisiting Data Literacy Instruction

In 2004, IASSIST Quarterly dedicated a whole double issue (Vol. 28-2 and 28-3) to the discussion of data literacy and library instruction of data literacy. This was fifteen years ago, long before big data became a big deal; yet the questions asked then are still as relevant, or maybe even more so, today. One question that we are still asking is ‘what is data literacy?’ Another is ‘what is the role of a librarian or a data services support team at an academic library regarding service models and challenges in data literacy education and advocacy’.

To the first question, several approaches to data literacy instruction arise from the body of literature on data literacy instruction in higher education. The first approach aligns closely with information literacy instruction. Across the years, data literacy has been discussed with reference to the different versions of *ACRL Information Literacy Competency Standards for Higher Education*, the most recent one of which is published in 2016 (ACRL, 2016). Data literacy, and its variations (data information literacy, research data literacy and science data literacy in terms of terminology), has been viewed by the library community as a subset, or extension, of information literacy in the time of high performance computing and an increasingly networked environment (Schield, 2004; Stephenson & Caravello, 2007; Carlson, Fosmire, Miller, & Nelson, 2011; Prado & Marzal, 2013; Womack, 2014; Maybee & Zilinski, 2015; Shorish, 2015; Frank & Pharo, 2016).

From this perspective, data literacy is, first of all, critical thinking applied to evaluating data sources and formats, and interpreting and communicating findings. As a companion to, or component of, data literacy, statistical literacy is the ability to evaluate statistical information as evidence; this involves the ability to understand summary statistics and graphs, and gauge the presentation of statistics, including the potential misuse of numbers (Schield, 2004; Gray, 2004; Stephenson & Caravello, 2007; Womack, 2014). Occasionally, it goes beyond discovering and evaluating datasets to include operationalizing research questions into measurable variables (Beauchamp and Murray, 2016).

The second approach builds upon information literacy instruction and centers around data management and curation. Core competencies proposed span knowledge and skills in databases, data acquisition and collection, documentation and standards, data management, curation, reuse, preservation, and sharing (Qin and D’Ignazio, 2010; Schneider, 2013; Carlson et al., 2011; Carlson, Johnston, Westra, & Nichols, 2013; Mooney, Collie, Nicholson, & Sosulski, 2014; Shorish, 2015).

The third approach encompasses the whole data lifecycle, ranging from data discovery and access, data interoperability and manipulation, data analysis, databases, data visualization, metadata, data preservation, data sharing and reuse, and best practices and ethics (Fosmire and Miller, 2008; Prado & Marzal, 2013).

With the first question in mind, it is time that we re-envision the library’s role in data literacy instruction. It is generally agreed that the library should lead efforts in critical thinking, statistical literacy, information literacy, and data visualization literacy (Gray, 2004; Schield, 2004; Womack, 2014). Beyond that, it is the faculty’s role to teach statistical theory within the discipline, while the library’s complementary role can be found in offering practical knowledge to the patrons (Thompson & Edelstein, 2004), and instructing ethics and preservation of data (Shorish, 2015). Data literacy instruction

is believed to be better delivered where it is integrated into the subject curriculum (Hunt, 2004; Stephenson & Caravello, 2007; Carlson et al., 2011; Maybee & Zilinski, 2015) and research projects (Carlson et al., 2013), and where it is combined with the efforts of the strategic partners on campus (Hogenboom, Phillips, & Hensley, 2011).

While information literacy serves as the foundation of the data literacy program, I would argue that data literacy instruction at the library may not be limited within the context of information literacy. For instance, when teaching statistical literacy, data services specialists may introduce students to more sophisticated documents than summary statistics. We may also include the layer of computational skills, which many have, and perhaps some of the methodology component. In the age of technology advancement and big data, data visualization, web-interfacing technology supporting data needs, or even machine learning basics may have already become part of everyday language.

In the section below, I will share a few of our experiments and regular instruction methods that seek to meet new challenges in the curriculum and beyond.

Teaching Data Literacy in Three Ways

1. Campus wide dialogue

We initiated the yearlong series of events titled ‘Lying with Data’, which builds upon our relationship with the Center for Data Science and Artificial Intelligence and Committee on Critical Inquiry by creating a platform to address core data literacy questions. The invited faculty tapped into the courses they have taught previously or their current research projects, which reduced the effort for faculty to prepare their sessions.

Topics included how to view graphs skeptically, getting the truth out of opinion surveys, understanding causal inference beyond correlations, and how to evaluate the appeal of statistics in advertisements. Other proposed topics included blind spots in machine learning and fraudulent reporting on pharmaceutical trials as scientific misconduct. The plan was to cover a wide range of issues arising from various fields and schools of research, certainly not limited to data science or statistics. Together with Trivia, we have hosted four of the six proposed topics.

This event evolved to be a campus wide dialogue once the Vice Chancellor volunteered to give a lecture, with support from two Deans of the academic departments. But more importantly, this event was a collaboration with two partners on campus, the Center for Data Science and Artificial Intelligence and the Committee on Critical Inquiry. The Committee was itself an initiative from the Provost Office. With its Information Literacy Subcommittee, the Committee directs students to participate in interactive discussions and hands-on activities to develop skills to evaluate information. The Center is a research hub with groups of interdisciplinary scholars conducting the most advanced research with quantitative and computational methods. The collaboration also shows how the library’s data services program stands in the middle ground of the ‘data spectrum’ in relation to the scholarly activities carried out by the research institutes and the general information literacy efforts on campus.

In this section, I will summarize these sessions, investigating the contents presented, approaches adopted, and messages delivered. Each session potentially has offered many ways to review, but below I

highlight where it concerns teaching data literacy, situating it in the larger ‘data ecosystem’ within the university. I have given more attention to some cases than others that deserve closer scrutiny.

Advocacy with Graphs. This session analyzed three cases that relied on visual displays to show how advocates sometimes used graphs to persuade their audience to perceive patterns that did not really exist. The essence of this talk rested in where the speaker contrasted and compared alternative ways to array elements of a graph. The cases were examined both in their current forms as well as the counterfactual scenarios that could have left drastically different impressions on the readers.

The first case is the 1987 Gotti case (Raab, 1987), where the speaker presented the graph ‘Criminal Activity of Government Informants’ that John Gotti’s attorneys used to show the jury that the reputation of the witnesses were questionable. The graph arranged the columns with names of seven witnesses and rows with their long list of criminal records, where the most outstanding record was placed in the middle to catch people’s attention. But the graph could have been organized in a different way to reduce the visual prominence of some parts of it, such as aligning some crimes at the bottom. All those adjustments would make the audience think that some witnesses might be actually believable and therefore influence the jury’s decision.

CRIMINAL ACTIVITY OF GOVERNMENT INFORMANTS							
CRIME	CARDINALE	LOFARO	MALONEY	POLISI	SENATORE	FORONJY	CURRO
MURDER	X	X					
ATTEMPTED MURDER		X	X				
HEROIN POSSESSION AND SALE	X	X		X			X
COCAINE POSSESSION AND SALE	X		X	X			
MARIJUANA POSSESSION AND SALE							X
GAMBLING BUSINESS		X		X		X	
ARMED ROBBERIES	X		X	X	X		X
LOANSHARKING		X		X			
KIDNAPPING			X	X			
EXTORTION			X	X			
ASSAULT	X		X	X			X
POSSESSION OF DANGEROUS WEAPONS	X	X	X	X	X		X
PERJURY		X				X	
COUNTERFEITING					X	X	
BANK ROBBERY			X	X			
ARMED HIJACKING				X	X		
STOLEN FINANCIAL DOCUMENTS			X	X	X		
TAX EVASION				X		X	
BURGLARIES	X	X		X	X		
BRIBERY		X		X			
THEFT: AUTO, MONEY, OTHER			X	X	X	X	X
BAIL JUMPING AND ESCAPE			X	X			
INSURANCE FRAUDS					X	X	
FORGERIES				X	X		
PISTOL WHIPPING A PRIEST	X						
SEXUAL ASSAULT ON MINOR							X
RECKLESS ENDANGERMENT							X

Figure 1 Criminal Activity of Government Informants. Reprinted from *Envisioning Information* (6th ed., p. 31), by E. R. Tufte, 1998, Cheshire, CT: Graphics Press. Copyright 1990 by Edward Rolf Tufte.

In the next two cases, the speaker mimicked situations in political campaigns where advocates for Democrats and Republicans use the same sets of data to advance their own agendas. A group of tricks to engage emotions of the readers have been presented and inspected on a graph of the median family income of the U.S. (1947- 2017). These tricks include: censoring time series to flatten the trend and to create an image of stability, or conversely, decreasing the units to increase the fluctuations; adding irrelevant reference data above or below the line being discussed to induce negative or positive moods in the audience; intentionally cutting off parts of the vertical axis to move the line up and down to different levels; selecting data points to shift the slopes in order to strengthen a message; not adjusting

for inflation to distort the real trend. Therefore, advocates always have a compelling story to tell by manipulating the level, trend and volatility of this line graph: ‘when we are in power, things get better; when they are in power, things get worse’ (Lehman, 2018).

Following the analysis, the speaker shared some suggestions for how to make a graph fair but effective, and for how to view a graph with a critical eye. The core message, ‘every chart, every graph, represents a set of decisions about how to represent the information’ (Lehman, 2018), was consistently conveyed in each case.

Getting the Truth Out of Opinion Survey. This session delved into a specific phase of quantitative social sciences research, data collection, through evaluating data sources and sampling schemes of survey design and implementation. Drawing on the experience of collecting data for opinion studies in Hong Kong and Shanghai, the speaker assessed how biases and errors in sampling could affect sample representativeness and data quality. Technical notions, such as sampling frame, population specification, sampling errors, questionnaire and fieldwork design, and social desirability, were integrated into the delivery of the talk with small examples full of details for each concept. The key takeaway was to differentiate between the problems that should be avoided (e.g. measurement error) in survey data collection and the ones that could not be completely resolved (e.g. sampling error), which could therefore be accommodated (Wu, 2018).

How to Make Causal Inference out of Casual Correlation. This session looked into another phase in social sciences research, analyzing and interpreting data. The empirical challenge arises in making causal inference when ‘the fact that two trends seem to fluctuate in tandem does not prove that they are meaningfully related to one another’ (Lu, 2018). The entire session was designed to deal with this challenge.

The session started with an introduction of typical logical fallacies when making causal inference, including omitted variables, survivorship bias, and reverse causality. It then proceeded with concrete examples from the speaker’s own research area in finance, where the speaker illustrated how econometric tools, such as Regression Discontinuity Design and Difference-in-Differences, could be leveraged in causal inference beyond correlation. The delivery style of this session resembled an academic seminar where the speaker presented his/her research, followed by a discussion among the group of audience with questions, comments, and debates.

Stop Being Fooled by Advertisements. This session touched upon a real-world problem outside academic research. It demonstrated how companies’ marketing strategies capitalize on the statistical appeal of scientific evidence featuring numbers to influence consumers’ perception and purchases. In the past, companies have fabricated data, used biased polling and meaningless reference groups, cherry-picked data, and created misleading visualizations to appeal to customers.

This session was constructed in a way that the speaker engaged the audience with a myriad of advertisements across industries and made connections with topics of the preceding lectures on sampling, causality versus correlation, and misleading graphs. A student posed an interesting question to the speaker: if as a marketing expert, could he, himself, effectively avoid the traps? The speaker replied by stating that even when conscious of the tricks of the trade, one could still be misled. However, ‘knowing and recognizing our cognitive biases will help prevent us from being fooled by

misleading ads', since 'individuals tend to embrace information that supports their wishes and reject information that contradicts them.' (Yan, 2018)

Trivia. This session invited students to test their knowledge and critical thinking on topics discussed in the 'Lying with Data' Series with game-based quizzes. Students earned points of one to three during the quiz depending on how difficult the question was. Unlike a lecture, which did not leave much room to exhibit technical details of a graph or a problem, the Trivia session was devised to evaluate the granularities of a problem. Students were encouraged to speak freely for each question, noting what went wrong and how the situation could have been improved or avoided. Following the competition, the library staff reviewed all the questions and offered solutions to the puzzles.

Unfortunately, the series only presented quantitative social science perspectives, though we attempted a session on machine learning and another on critical examination of scientific misconduct. These sessions were later cancelled by the speakers due to conflicts with their schedules.

Materials from the Series will be recycled and reused in two ways. The Social Sciences Capstone Project class will utilize the video recordings as course materials, since each of the topics are relevant to the course objectives. The Series will also grow to be an online mini course consisting of pieces of short videos under the same theme, reproduced for a wider audience outside the university walls.

Moreover, how we marketed this event may also be worth noting. We created a mascot, the 'Lying with Data' Pinocchio, for promoting the event throughout the year (see Figure 1). Pinocchio, a cartoon character whose nose grows when he lies, was selected for its common association with lying to reinforce the title of the series. We built several three-dimensional posters of human size that stood at different locations of the campus building (see Figure 2). A student worker custom designed a limited number of metallic bookmarks in the shape of the mascot (see Figure 3). Students were encouraged to post selfies with the mascot on their social media, in order to generate buzz and publicity about the mascot and the event.



Figure 2. 'Lying with Data' mascot. Illustration by Keyin Wu.



Figure 3 'Lying with Data' poster. Created by Keyin Wu.



Figure 4 'Lying with Data' bookmark. Designed by Keyin Wu.

2. For-class instruction

It is noted that data literacy is best taught when integrated into the subject curriculum (Hunt, 2004; Stephenson & Caravello, 2007; Carlson et al., 2011; Maybee & Zilinski, 2015). We offer such instruction, both up-to-date with the latest technology and that fit with the curriculum. Part of this program consists of stand-alone workshops outside classrooms; part of it are custom sessions designed for a course. We used to refer to the later kind as 'in-class' instruction, but here I am replacing the term with 'for-class' instruction to include those not happening physically in the classrooms but created for a class. One type of such for-class instruction is recorded videos.

We have a case where our staff recorded eight videos on data scraping with Python for the Business Analytics class (Konagai, 2018). These are short video tutorials produced at our studio of the library.

Videos were distributed to the class as recommended resources. The benefit of this practice is that the contents can be recycled for future classes, and that students can view the materials at their own pace. But the risk also exists, especially when the tools are updating themselves quickly and thus the tutorials can be soon outdated.

Conventional in-class instruction sessions are delivered by request, most frequently coming from digital humanities, urban design and occasionally the Business department. For instance, the GIS specialist has been asked to deliver workshops on spatial analysis and geoprocessing. For technical sessions, a typical way is to give out a handout that contains step-by-step instructions of using a tool at the beginning of the class (Luo, 2019). The majority of class time is devoted to the practices of problem solving, where students consult the handouts to complete a task. Before the practices, conceptual questions are first explained, and student readiness are assessed with quizzes. For non-technical topics, a more visual display of the contents is desired to gather students' attention. For instance, a session on the topic of storytelling with data visualization used the platform Story Maps to showcase the narrative revelation of data, rather than using PowerPoint slides. The media, Story Maps, was itself part of the delivery in a session on visualizing GIS data.

3. Online Casebooks

In addition to the events and the for-class instructions, we have created online casebooks that are hosted on the Data Resources and Services website (Dai, 2019). The website has been set up as complementary resources to the pre-scheduled library workshops on various topics in the data lifecycle from data discovery (of Chinese datasets), data preprocessing, data analysis, to data visualization. For each tool, we created tutorials targeted at users of various needs. For instance, topics under 'Coding Smartly' are tricks and tips with examples from real projects to solve particular problems. Topics under 'Cases' are tutorials that demonstrate the full workflow of a project with technical details. The remaining contents are introductory tutorials for a certain step in a project, such as exploratory analysis, transforming variables, handling strings, dates and times etc.

After a few trials, we found that library workshops, even in-class instruction, were better suited for introductory topics. Several attempts at more advanced, or intermediate workshops did not work out as expected in terms of attendance and student reception of the contents. This may be due to the fact that our main student body are undergraduate students. Besides, students more often learn more advanced materials through self-exploration rather than teaching. Therefore, intermediate and more advanced topics have been moved to the online space.

Discussion

1. How literate should we be

There is always the question, how 'literate' we, librarians and technologists, should be in order to teach others data literacy. Part of this question concerns the tools. The question puts more pressure on professionals nowadays as the field and industry of data science are advancing daily; this anxiety and need to upgrade ourselves constantly is real. At a new institution, the stress can double when there is no room for old technology to 'die out'. Yet service providers may have been trained with a small number of tools, while students arrive every year with experience using numerous emerging tools.

The other part of this question concerns support on methodological or conceptual aspects in software assistance. Some libraries refrain from offering such support, while others do not. Perhaps an intermediate step would be asking ourselves to get familiar with the methodological or conceptual knowledge of a domain even though we may not offer direct support in advising or teaching. We should be able to join the conversation alongside data users, if not data scientists, to generate more meaningful communication and collaboration.

2. Marketing

Essential to our teaching activities are two core values: 'let the quality speak for itself', and 'outreach by teaching'. In a way, everything we do is about marketing. In fact, the 'Lying with Data' events probably have said much more about who we are than any brochure could possibly do; the Data Resources and Services website presents to the university community the many facets of our services and capacity. Word of mouth from faculty advocates carry us further than advertising.

Acknowledgement

Huge thanks go to Caitlin MacKenzie Mannion, Xiaojing Zu, Adrian Hodge, Edward Junhao Lim, Jennifer Anne Wood Stubbs, and Fan Luo for your valuable and detailed feedback to the paper.

Reference

ACRL (2016). Framework for Information Literacy for Higher Education. Retrieved from http://www.ala.org/acrl/sites/ala.org.acrl/files/content/issues/infolit/Framework_ILHE.pdf

Beauchamp, A., & Murray, C. (2016). Teaching Foundational Data Skills in the Library. In Kellam, L. & Thompson, K. (Eds.), *Databrarianship: The Academic Data Librarian in Theory and Practice* (pp. 81-92). Chicago, IL: ACRL.

Carlson, J., Fosmire, M., Miller, C. C., & Nelson, M. S. (2011). Determining data information literacy needs: A study of students and research faculty. *Portal: Libraries and the Academy*, 11(2), 629–657. <<http://dx.doi.org/10.1353/pla.2011.0022>>

Carlson, J., Johnston, L., Westra, B., & Nichols, M. (2013). Developing an Approach for Data Management Education: A Report from the Data Information Literacy Project. *International Journal of Digital Curation*, 8(1), 204-217. <<http://dx.doi.org/10.2218/ijdc.v8i1.254>>

Dai, Y. (2019). *Data Resources and Services*. Retrieved from <https://shanghai.hosting.nyu.edu/data/>

Fosmire, M., & Miller, C. (2008). Creating a culture of data integration and interoperability: librarians and Earth Science Faculty collaborate on a geoinformatics course. Retrieved from <https://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1850&context=iatul>

Frank, E. P., & Pharo, N. (2016). Academic librarians in data information literacy instruction: a case study in meteorology. *College & Research Libraries*, 77(4), 536-552. <<http://dx.doi.org/10.5860/crl.77.4.536>>

Gray, A. S. (2004). Data and statistical literacy for librarians. *IASSIST Quarterly*, 28(2/3), 24-29. <<http://dx.doi.org/10.29173/iq793>>

Hogenboom, K., Phillips, C. M. H., & Hensley, M. (2011). Show Me the Data! Partnering with Instructors to Teach Data Literacy. Retrieved from http://www.ala.org/acrl/sites/ala.org.acrl/files/content/conferences/confsandpreconfs/national/2011/papers/show_me_the_data.pdf

Hunt, K. (2004). The challenges of integrating data literacy into the curriculum in an undergraduate institution. *IASSIST Quarterly*, 28(2/3), 12-15. <<http://dx.doi.org/10.29173/iq791>>

Konagai N. *Data Scraping with Python*. Retrieved from NYU Stream NYU community private access.

Lehman, J. (2018). *Advocacy with Graphs*. Lecture presented at the 'Lying with Data' events.

Lu, Y. (2018). *How to Make Causal Inference out of Casual Correlation*. Lecture presented at the 'Lying with Data' events.

Luo F. (2019). *Spatial Analysis and Geoprocessing*. Retrieved from <http://bit.ly/2UHizR8>

Maybee, C., & Zilinski, L. (2015). Data informed learning: A next phase data literacy framework for higher education. In *Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community* (p. 108). American Society for Information Science. <<http://dx.doi.org/10.1002/pra2.2015.1450520100108>>

Mooney, H., Collie, W. A., Nicholson, S., & Sosulski, M. R. (2014). Collaborative Approaches to Undergraduate Research Training: Information Literacy and Data Management. *Advances in Social Work*, 15(2), 368-389. <<http://dx.doi.org/10.18060/15089>>

Prado, J. C., & Marzal, M. Á. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri*, 63(2), 123–134.

Qin, J., & D'Ignazio, J. (2010). Lessons learned from a two-year experience in science data literacy education. Retrieved from <http://docs.lib.purdue.edu/iatul2010/conf/day2/5/>

Raab, S. (1987, March 14). A Weakness in Gotti Case; Major U.S. Witnesses Viewed as Unreliable. *The New York Times*, Retrieved from <https://www.nytimes.com/1987/03/14/nyregion/a-weakness-in-gotti-case-major-us-witnesses-viewed-as-unreliable.html>

Schild, M. (2004). Information literacy, statistical literacy and data literacy. In *IASSIST Quarterly*, 28(2-3), 6-11.

Schneider, R. (2013). Research data literacy. In *European Conference on Information Literacy* (pp. 134-140). Springer, Cham. <http://dx.doi.org/10.1007/978-3-319-03919-0_16>

Shorish, Y. (2015). Data information literacy and undergraduates: A critical competency. *College & Undergraduate Libraries*, 22(1), 97–106. <<http://dx.doi.org/10.1080/10691316.2015.1001246>>

Stephenson, E., & Schifter Caravello, P. (2007). Incorporating data literacy into undergraduate information literacy programs in the social sciences: A pilot project. *Reference Services Review*, 35(4), 525–540. <<http://dx.doi.org/10.1108/00907320710838354>>

Thompson, K. A., & Edelstein, D. M. (2004). A reference model for providing statistical consulting services in an academic library setting. *IASSIST Quarterly*, 28(2), 35-38. <<http://dx.doi.org/10.29173/iq795>>

Tufte, E. R. (1998). *Envisioning Information* (6th ed.). Cheshire, CT: Graphics Press.

Womack, R. (2014). Data visualization and information literacy. *IASSIST Quarterly*, 38(1), 12–17. <<http://dx.doi.org/10.29173/iq619>>

Wu, X. (2018). *Getting Truth out of Opinion Survey*. Lecture presented at the ‘Lying with Data’ events.

Yan, D. (2019). *Stop Being Fooled by Advertisements*. Lecture presented at the ‘Lying with Data’ events.

¹ Contact: Yun Dai, NYU Shanghai Library. Email: yun.dai@nyu.edu

² NYU Shanghai is a portal campus in the NYU global network and a joint venture between NYU and East China Normal University. Founded in 2012, it is a campus with around 1,200 enrolled undergraduate students from more than 70 countries and more than 200 full-time international faculty members.