Better data management, one nudge at a time

Daria Orlowska[1], Colleen Fallaw[2], Yali Feng[3], Livia Garza[4], Ashley Hetrick[5], Heidi Imker[6], Hoa Luong[7]

Abstract

How do you help people improve their data management skills? For our team at the University of Illinois at Urbana-Champaign, we decided the answer was "one nudge at a time".

A study conducted by Wiley and Mischo (2016) found that Illinois researchers are aware of data services available but under-utilize them. Many researchers do not consider data management as a concern distinct from researching and producing scholarly work products. In 2017, the RDS piloted the Data Nudge – a monthly, opt-in email service to "nudge" Illinois researchers toward good data management practices, and towards utilizing data services on campus. The aim of the Data Nudge was to address the gap between knowing about a service and using it by highlighting best practices and campus resources.

The topics covered in the Data Nudge center around data. Some topics are applicable to everyone, such as data back-up, documentation, and file naming conventions. Other topics are specific to Illinois, like storage options, events, and conferences.

After four years, the Data Nudge has accumulated over 400 subscribers through word-of-mouth, marketing channels on campus and inclusion in subject liaisons' instructional workshops. It receives stable open rates averaging at 52% (compared to 19.44% average industry rate for Higher Education*) and many compliments from subscribers. We expect the Data Nudge to continue supplementing workshops and training as an effective means of communication to reach researchers on our campus. In the spirit of re-use, we are in the process of archiving the Data Nudge topics in a reusable format, readily adaptable by other institutions.

Data Nudge link: https://go.illinois.edu/past_nudges

Keywords

Data Nudge, data management, research data service, subject librarians, community engagement, service promotion, research data

Introduction

As an R1 research university, the University of Illinois at Urbana-Champaign boasts over 150 centers, labs, and institutes, in addition to active research programs of more than 2,700 faculty (Office of the Vice Chancellor for Research & Innovation, 2019). In FY18, the university's research and development expenditures amounted to $652 million, with $353 million in federal research expenditures (Office of the Vice Chancellor for Research & Innovation, n.d.). The release of the 2013 Office of Science and Technology Policy memorandum calling for "increasing access to the results of federally funded scientific research" bolstered the recommendation to establish specific research data services and infrastructure in Illinois' 2013-2016 strategic plan, under the goal of "foster[ing] scholarship, discovery and innovation." The strategic plan called for the creation of a "research data service and accompanying research education initiative in the curation, use and dissemination of large amounts

of data" (Strategic Working Group, 2013). Since its inception in 2014 with four staff members, the Illinois Research Data Service (RDS) has aided researchers by offering a digital repository (the Illinois Data Bank), curating deposited datasets to ensure completeness and accessibility, and reviewing data management plans for federal grant applications.

Around the same time, Fearon et al. (2013) surveyed 73 ARL member libraries to gain a sense of how research libraries were meeting data service demands driven by an increased emphasis on open data and DMP requirements. The survey indicated that libraries were still in the early stages of development. Although data education was not covered as its own category, most education occurred through website resources and consultations (N = 48 for DMPs and other research data management) (Fearon et al., 2013, pp. 13-14; Fearon et al., 2013, p. 42). Less than half of respondents reported running workshops for DMPs (N=33) (Fearon et al., 2013, p. 13) and other research data management education (N=35) (Fearon et al., 2013, p. 42). Service challenges touched on themes such as campus-wide collaboration and engagement and marketing (Fearon et al., 2013, pp. 82-90). Similarly, a 2014 survey of eighty-one UK libraries by Cox and Pinefield found that UK academic libraries offered a limited research data services with only twenty-four libraries hosting web resources and twenty-eight offering consultations (Cox and Pinefield, 2014, p. 18), but many saw data services as a priority, with medium-term goals of consultations and training (Cox and Pinefield, 2014, p. 22).

However, offering these services alone did not completely meet the need to help researchers take advantage of resources to support better data management practices. A study conducted by Wiley and Mischo (2016) found that Illinois researchers were aware of campus data preservation and management services, but rarely used these resources. They further found that researchers did not consider data management as a separate element of scholarly communication workflow, highlighting an area that could benefit from improved outreach.

In 2017, the RDS piloted the Data Nudge, a monthly opt-in email service, to prompt and offer support to Illinois researchers about data services on campus. The aim of the Data Nudge was to address the gap between awareness and practice by promoting best practices and campus resources in a bite-sized monthly subscription service.

While many recent surveys have identified the areas in which libraries are providing data support services, to our knowledge, they have not explicitly gathered information on how data education is being implemented. In a recent article, Ohaji, Chawner and Young (2019) argue that the role of a data librarian encompasses skill building, developing policy, and supporting data infrastructure, with a specific focus on fostering new connections across campus. This suggests that hiring a data librarian as a response to data needs moves past a traditional liaison role, and encourages a new approach to engaging with the research community. A scan of websites from other universities in the Data Curation Network (DCN), a group of data curation collaborators who offer data services similar to the University of Illinois, revealed that all universities within the DCN offer a website and consultation services, and most offer library guides and workshops as their primary source of data education (Table 1). However, the scan also revealed a plethora of new forms of data education, suggesting that universities are spearheading a more experimental approach to building awareness and competency within their research community.

Universities in the DCN run community workshops on demand, upload online tutorials, and engage members of their research community through social media. In addition to continuing to offer more traditional data education like workshops, they have also ventured into listservs and newsletters. Of the five universities offering this service, most use it as a communication tool: Washington University offers "regular updates about events and news", Cornell University offers "low traffic mailing list for updates, upcoming events, and news", John Hopkins University offers "the latest data management and sharing information, upcoming Data Services events and training sessions", and Duke University offers "a biweekly newsletter for upcoming events, exhibits, resources, services, and other library news". In contrast, the University of Illinois has developed a dedicated data listserv offering resources and highlighting best practices: the Data Nudge.

The purpose of this paper is to provide insight into the success of turning a listserv into a tool that not only nudges the research community into action but also connects them with resources across campus. We hope that through describing our processes and publishing our materials openly online, we can offer other university libraries a framework and ready-made content to connect with their own communities.

	Workshop Series	Workshop on Demand	Online Tutorial	Consultation	Library Guide	Landing page	Listserv/ Newsletter	Blog posting	Twitter handle
University of Illinois	x	x		x	x	x	x		
University of Michigan	x	x		x	x	x		x	
University of Minnesota	x	x	x	x		x			
Washington University	x	x		x	x	x	x	x	x
Cornell University	x	x		x	x	x	x		x
John Hopkins	x	x	x	x	x	x	x		
Penn State	x	x		x	x	x			
New York University	x	x		x	x	x		x	x
Duke University	x	x	x	x	x	x	x	x	x

Structure of the Data Nudge

1. Topic selection and development

Inspiration for Data Nudge topics comes from a variety of sources, including campus events and consultations with researchers in which it is clear that being aware of information made a positive difference to the research process. Increasingly, we receive requests from campus partners supporting research with specific messages they hope to share with the Illinois community through the Data Nudge. While these can be opportunities to share new information about emerging and changing services available to researchers, they can also be challenging as the content must be reworked to match the format, tone, and purpose of the Data Nudge.

However, the seed for a Data Nudge most often sprouts from personal experiences, when someone on the team learns something in the course of their work and thinks, "I wish I had known that sooner!" For example, while evaluating file types in the Illinois Data Bank and encountering compressed and archive file formats, various team members were surprised by different aspects of the distinctions, characteristics, and options available. Thus, the August 2018 Data Nudge focused on four common compressed file formats widely used in research and offered readers facts about files that do not compress well (Figure 1).
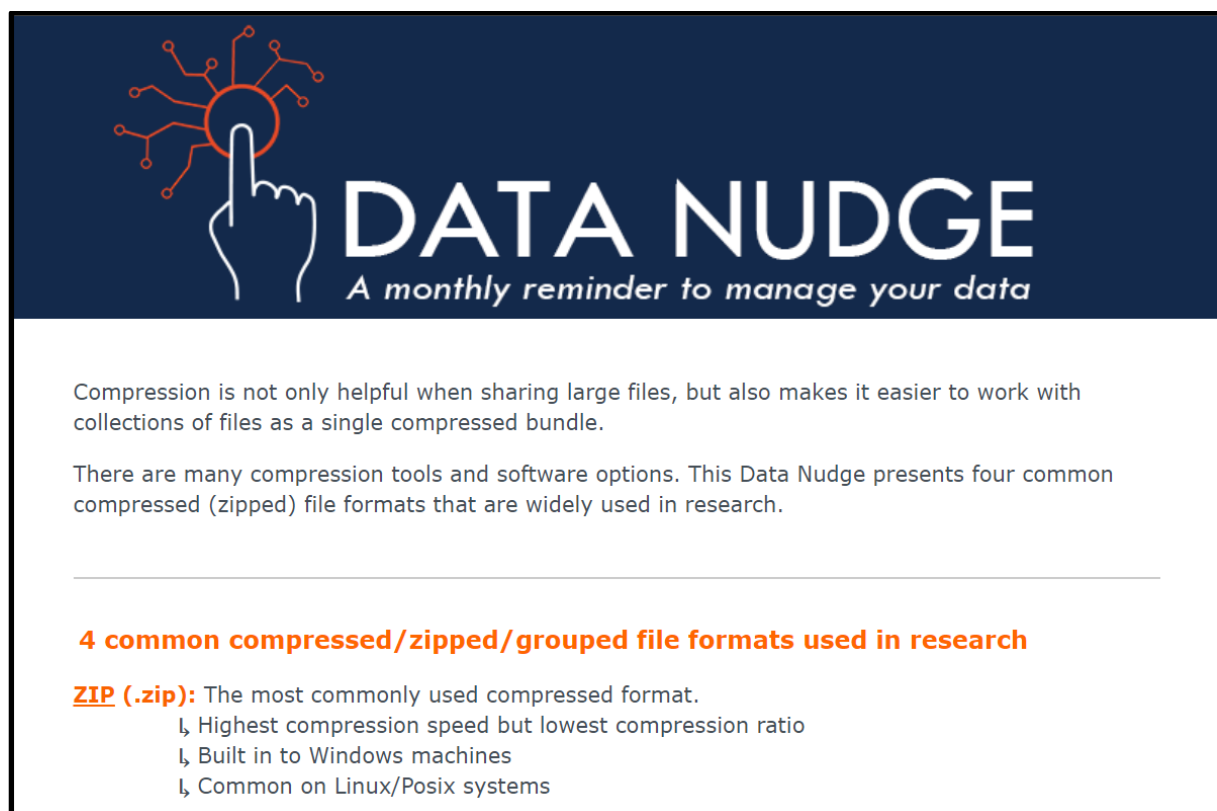

Figure 1. A section of the August 2018 Data Nudge, "Common zipped file formats used in research"

Regardless of the source of inspiration, the idea for a Nudge begins when a proposer pitches a topic to the team. Initial acceptance of a topic requires that it can be represented in a concise and visual manner and has dedicated support on campus (only for topics that specific to Illinois), demonstrated

through the inclusion of who to contact with questions. Once the pitch convinces the team, accepted topics are scheduled depending on campus events or in an order that allows topics to build on each other. After the brainstorming process, a member of the team commences work on an initial draft of the Data Nudge, aiming to work ahead 2-4 months before the release date. This work consists of research, collaborating on how to phrase and structure the content, and creating visual elements when appropriate. All draft materials are transferred into a template, where they are checked for size and adjusted using HTML. Any images are uploaded and linked from U of I Box, and alt-text is included for accessibility.
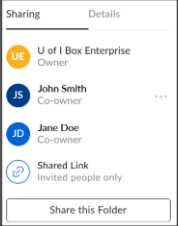
At least one month in advance of release, the draft Data Nudge is presented at a weekly RDS meeting to receive feedback from other team members. Team members consist of partners within and outside of the University of Illinois Library, including subject liaisons, functional specialists, and IT professionals. This mix of diverse perspectives is crucial to crafting messages that are free from jargon and speak across disciplinary or professional boundaries. Drafts are displayed on a large monitor and modified live based on group feedback for clarity, brevity and appeal to the target audience. In the case where more research is needed or larger scale rework is required, the group provides suggestions during the meeting, and the team makes adjustments afterwards. If the complexity of the topic still remains high, then it is postponed, and another is chosen that better fits the guiding principles. After several reviews, the content is finalized and scheduled to be sent out on the last Tuesday of the month.

While we did not start out with a fully-articulated set of guiding principles, a de facto sense of a Data Nudge style has formed from a consensus of I-know-it-when-I-see-it responses as the team shapes a draft into published form. We are guided substantially by empathy for our busy researchers, and thus we shape content around anticipations of how it will be experienced by our subscribers. We aim to offer pragmatic, specific information that a researcher could use immediately, or describe a practice that could improve the body of research data available in the world. We evaluate drafts in terms of careful accuracy for scan-and-skim and every-word readers, striving to offer a valuable take-away from a glance. We close by highlighting available resources, linking to relevant materials, and providing contact information for those who could answer further questions.

The January 2020 Data Nudge on "Illinois Box Features" is a particularly good example of reworking content based upon empathetic reading. Originally, the Data Nudge content was adapted from text provided by a campus partner. The text provided a high-level overview of useful features, security best practices for data stored in Box, and handy training videos to help users get started. Because our primary audience is researchers and we want to follow the aforementioned guiding principles, the RDS meeting members decided to emphasize and demonstrate features we know to be of high interest to researchers. These includes team folders for collaboration, favorites for quickly locating frequently used folders and files, and security concerns (Figure 2).
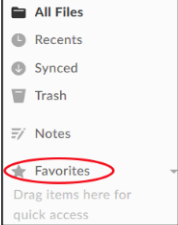
Figure 2. Final version of the January 2020 Data Nudge, "Illinois Box Features".

2. Data Nudge platform

During its first year (between 2017-01 and 2018-01), the Data Nudge was delivered through Mailchimp (mailchimp.com), a commercial marketing automation platform and email marketing service. Subscribers could view the email in their choice of browser, in HTML format. Beginning in 2018, the University's Technology Services implemented DMARC (Domain-based Message Authentication, Reporting and Conformance) to authenticate emails from third party email services to reduce phishing and increase email security, and Mailchimp was among several email services Technology Services identified as not being configured to work with the new DMARC. To continue supporting our researchers, we switched to the Email+ platform in Webtools, a resource provided by University Public Affairs (Figure 3). In either case, what was most important was a user-friendly platform for both our subscribers and for us as the content creators and providers. In particular, it was essential that we had access to analytics so we could evaluate usage.
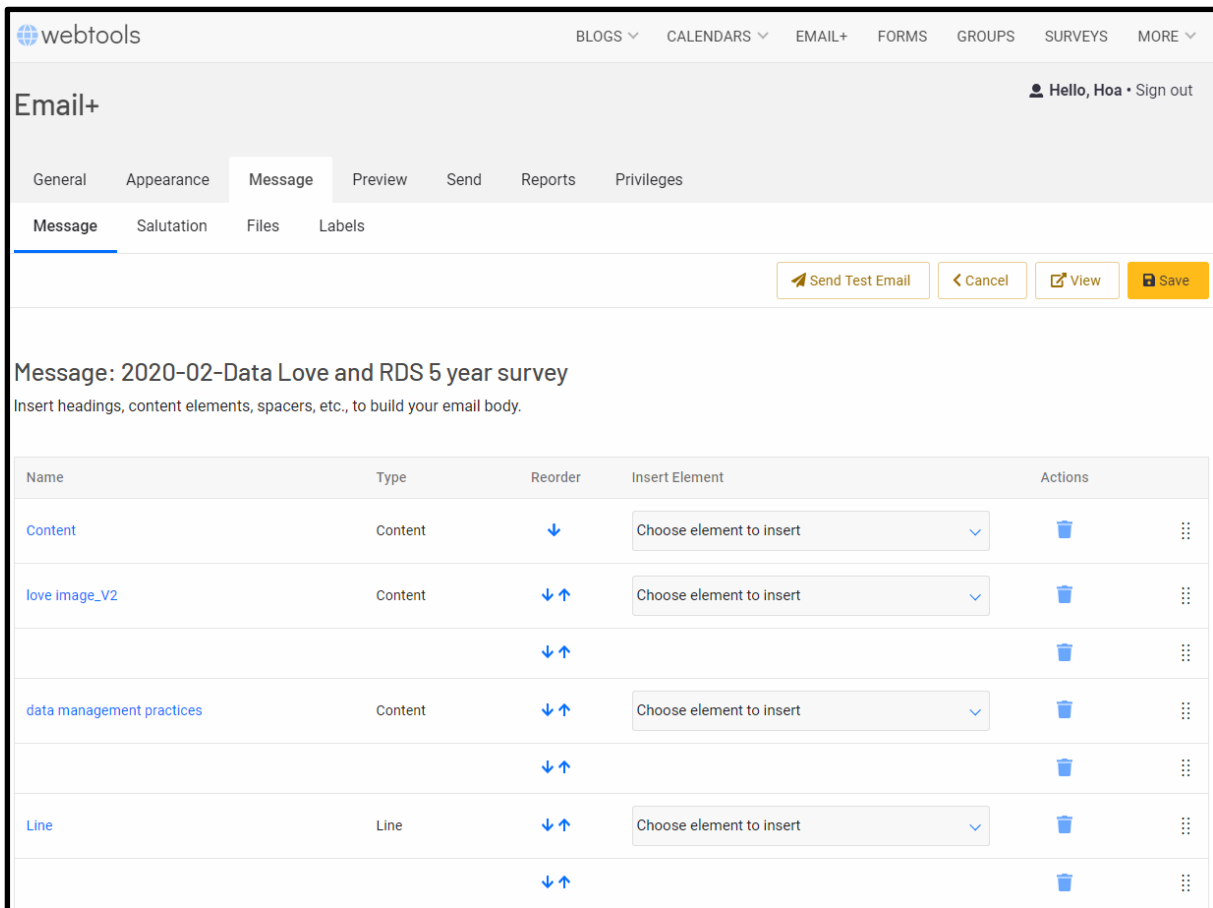
Figure 3. To conform to campus security requirements, content is developed in Email+ platform, powered by an Illinois service called Webtools.

3. Marketing channels

To bring the Data Nudge to the campus community, we send announcements to two different e-bulletins containing general interest campus information and opportunities for faculty/staff or graduate students at the beginning of each semester. We have created Data Nudge cards not only to pass out at university events for IT professionals, graduate students, and faculty, but to make them available as take-away promotional literature in the library's digital scholarship center. This card template has been transformed into a digital sign that is displayed regularly at the Main Library and library branches with heavy RDS users such as our campus engineering and life sciences libraries. We also advertise the Data Nudge on the RDS website, with a link to past Nudges.

Another avenue is subject librarian promotion, such as an announcement in a monthly departmental newsletter or through information sheets created for undergraduate instructional sessions. Other advertisements include electronic slides displayed on departmental monitors or by posting physical Data Nudge promotion flyers with a QR code, making it convenient for users to subscribe through their smartphones (Figure 4).

Figure 4. Data Nudge promotional flyer

As the Data Nudge has grown, we have been pleasantly surprised by strong word-of-mouth advertising, with users unexpectedly championing the Nudge without prompt at community events. In many cases, internal subscribers moving on from the University of Illinois continue to subscribe through a personal email or by resubscribing with a new address. Other external subscribers include those who have never been affiliated with the University of Illinois, and we can only speculate whether they stumbled on Data Nudge organically or through shout-outs from external sources such as webinars hosted by current subscribers.

4. Usage figures

Thanks to all of our marketing channels, the number of Data Nudge subscribers has been steadily increasing. From humble "library friend" opt-ins in 2017, the Data Nudge has accumulated over 400 subscribers as of December 2020, with 82% University of Illinois subscribers and 18% external subscribers, including individuals from government organizations and other academic institutions (Figure 5, 6).

Reports are recorded a week after the release of the emails, with an average of 52% for email open rate and 11.4% for link click rates between January 2017 through December 2020. In comparison, the industry open and click rate (Constant Contact, 2021) for higher education is around 19.55% and 7.94%, respectively (Figure 8).
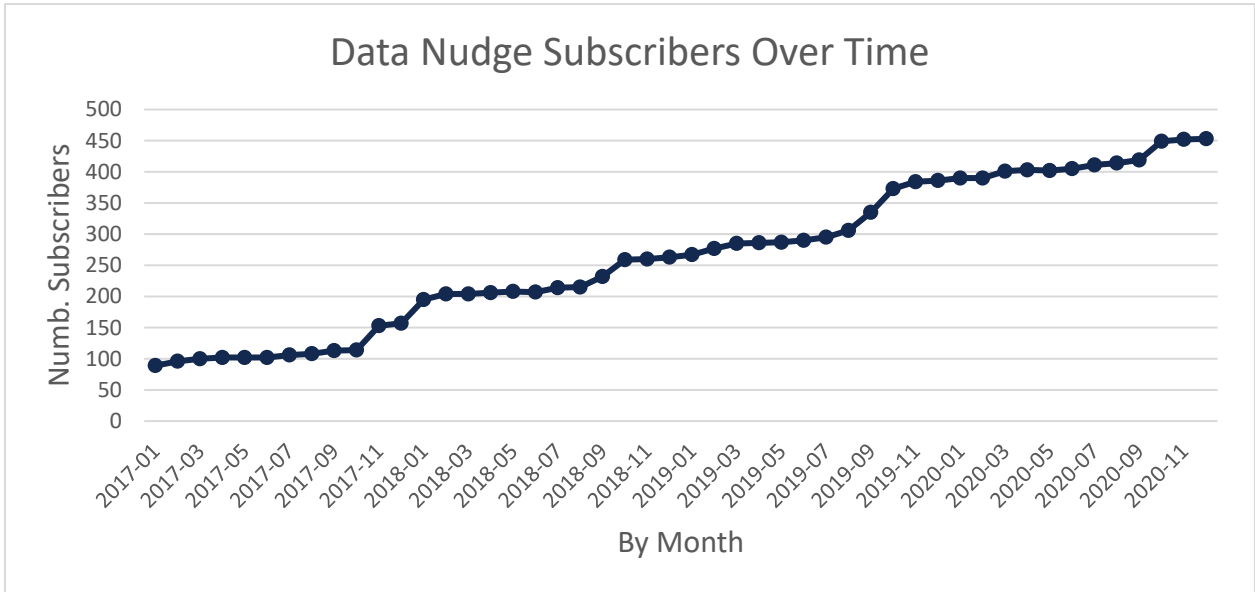
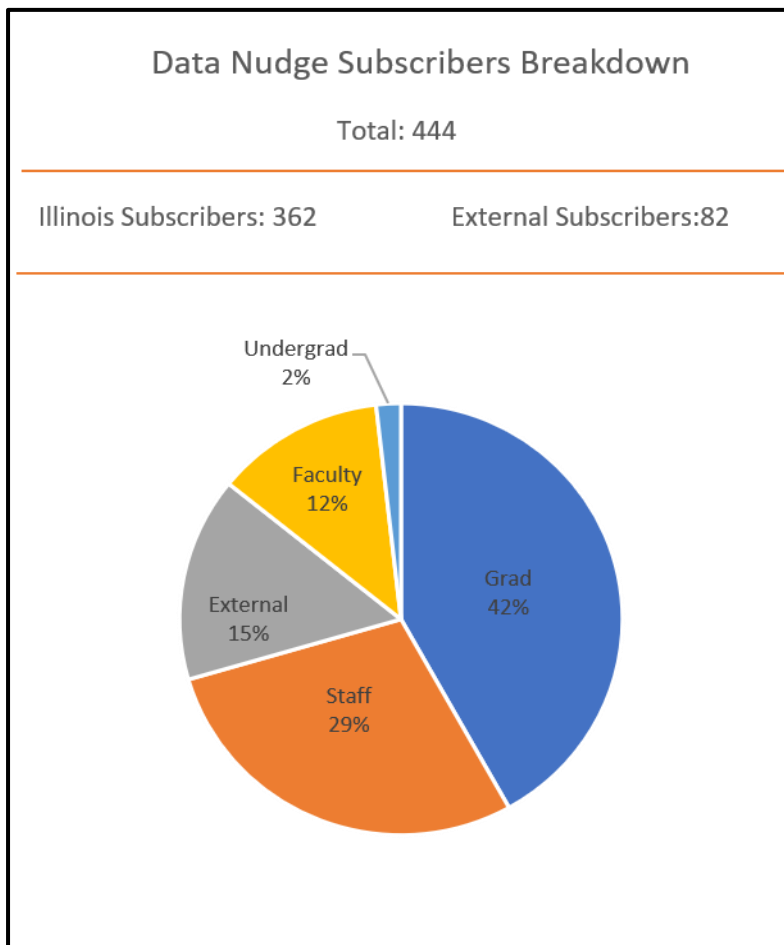Figure 5. Number of subscribers between inception in January 2017 through December 2020.

Figure 6. Breakdown of Data Nudge subscribers.
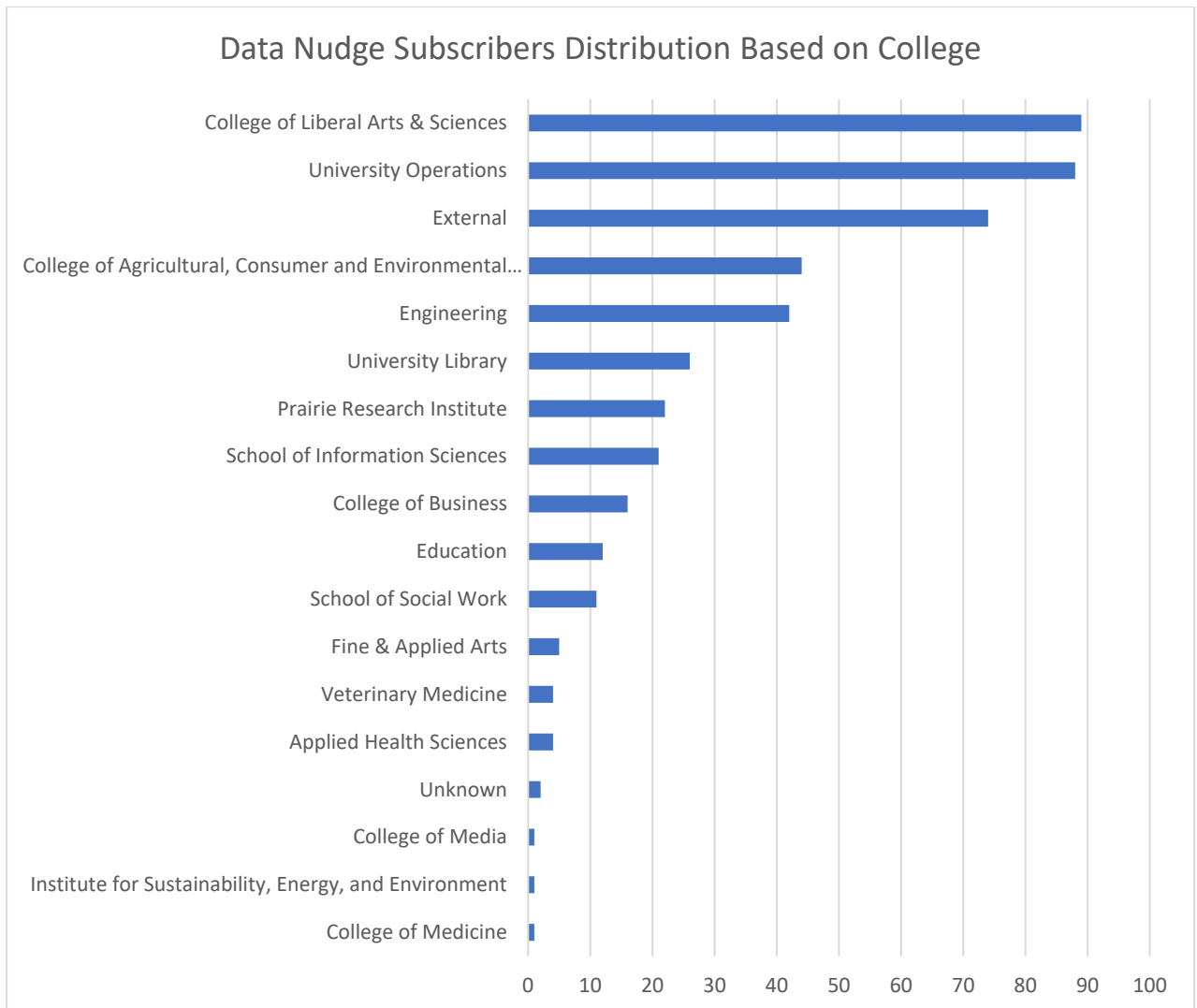
Figure 7. Number of subscribers break down by college. Note: "University Operations" refers to subscribers who administer or support services, like Technology Services / Research IT, NCSA, IGB, Sibel Design Center, OVCRI, Graduate College, etc. "External" refers to subscribers who are located outside of Illinois. "Unknown" refers to subscribers who we could not locate their department.
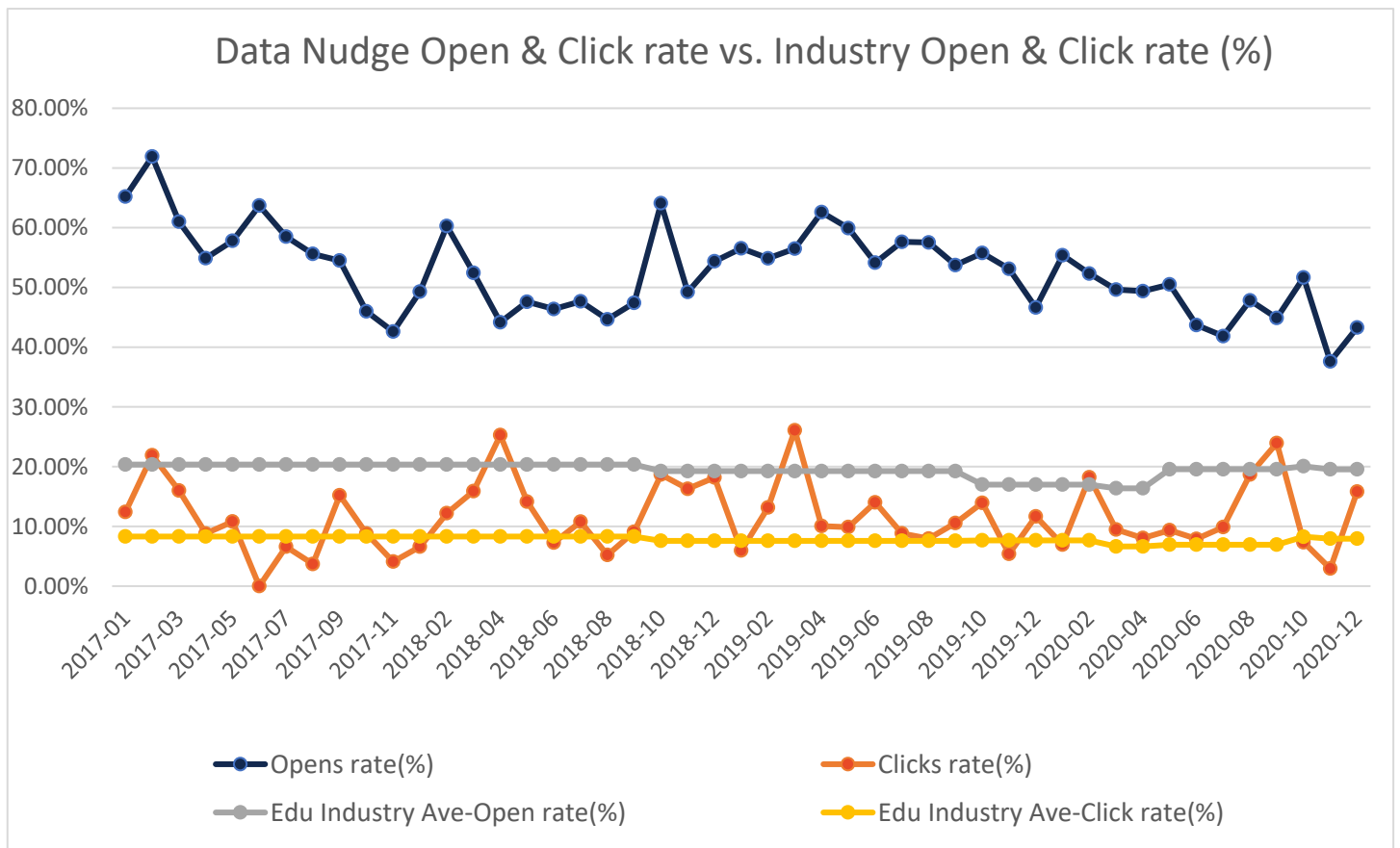
Figure 8. Open and click rate of each Data Nudge topic compared with the educational industry average rate, according to Constant Contact (2021).

Data Nudge Topic	Open Rate
2017-02 File organization	71.90%
2017-01 Backup	65.20%
2018-10 Scary data stories from IL Community	64.09%
2017-06 Box tips	63.70%
2019-04 Data Cleaning	62.59%

Table 2. Top five most popular Nudges by open rate.

 Orlowska, Daria; Fallaw, Colleen; Feng, Yali; Garza, Livia; Hetrick, Ashley; Imker, Heidi & Luong, Hoa (2021) Better data management, one nudge at a time, IASSIST Quarterly 45(2), pp. 1-16. DOI: https://doi.org/10.29173/iq1010

5. Quotes from users

We embed a one-question survey in the footer of each Data Nudge to learn if subscribers have been nudged into action by that month's topic. After explaining how they have been nudged, subscribers are provided an optional field for their name and campus address to receive a swag bag. Examples of feedback are provided below:

"I just tagged my Box data!! Thanks for the fabulous tip!"

"I have gone through your Data Nudges and I love them. I have been tasked as the lead of a data initiative [...] to improve the current data infrastructure. I would love to hear more about your thought processes."

"[...] I did want to share how much I enjoy the Nudge and compliment you on your content. The Halloween cartoon yesterday was especially well done, sometimes audiences equate simplistic with simple and here's a great example of how wrong that view is. Kudos on your consistently well-crafted messaging."

"I've subscribed to these data nudge newsletters for a while, and I've found them pretty useful. This one [ORCID] in particular is super important, especially because the culture of academia makes it so that certain people's work is acknowledged and recognized more than others'. Little things like this can go a long way in changing that culture. Thanks for sending these out!"

"I just wanted to say again how much I enjoy the Data Nudge. I learn something valuable every time."

Adaptability and reusability topics for Data Nudge

Data Nudge content is under a CC-BY license, and all topics are available online: https://go.illinois.edu/past_nudges. We are currently in the process of converting content into a reusable format (HTML and PDF), readily adaptable by other institutions and archived at IDEALS, known as Illinois institutional repository: https://www.ideals.illinois.edu/handle/2142/109236.

Some Data Nudge topics can be easily customized by other institutions, while others are specific to Illinois and require modification. Nudge topics listed as "Generally Applicable" may require some changes to linked resources, but otherwise provide universally applicable content. On the other hand, Nudge topics listed as "Illinois Specific" mostly focus on Illinois resources and would require significant modification to customize for another institution. "Mixed" Nudge topics contain Nudges that fall into both the "Generally Applicable" and "Illinois Specific" categories and require further consideration before adapting (Table 2). The following table is a quick guide to determine the amount of institutional customization a Data Nudge topic will need:

Topic	Data Nudges	Generally Applicable	Illinois Specific	Mixed
Back Up Data	2017-01, 2017-11, 2019-03			x

Campus Events / Resources	2017-09, 2018-04, 2018-11, 2019-09, 2020-08, 2020-09		x	
Cloud Storage	2017-06, 2017-07, 2018-03, 2020-01, 2020-03		x	
Documentation	2017-05, 2019-10	x		
Data Analysis	2020-04	x		
Data Curation	2020-11		x	
Data Cleaning	2019-04	x		
Data Destruction	2019-11	x		
Data Loss	2020-05	x		
Data Management Plan and practices	2017-03, 2017-10, 2020-02	x		
Data Repositories	2017-08, 2018-07, 2019-02			x
Data Sharing	2018-05, 2018-06, 2018-09, 2019-01, 2019-07, 2019-08, 2020-06	x		
Data Visualization	2019-05, 2019-06	x		
File Formats	2018-01, 2019-08	x		
File Naming	2017-02, 2018-02	x		
Risk Assessment	2017-04, 2020-07	x		
Scary Data Stories (Halloween comic)	2018-10, 2020-10			x

Table 3. Categorization of Data Nudge topics by reusability.

Discussion/Conclusion

What started out as a small pilot program intending to periodically nudge researchers about their interest in data has grown into an effective tool for the promotion of effective data management. Along the way, we have learned that it takes a long time to craft a short message, it takes a networked team to craft a useful message, and that people really enjoy comics. The biggest lesson learned might be that the time and effort are worth it because our community values and uses our messages.

The Data Nudge has become an outlet when we are bursting with a potential solution that we wish everybody had in their toolboxes. When librarians talk about RDS to their units, subscribing to the Data Nudge has become a concrete step furthering the relationship between our service and crucial collaborators. The archive of past Data Nudges has become a resource for subscribers to go back and check on information about services or for RDS team members to share right-sized guides on topics that come up in consultations. With our curated topics freely available for adaptation, we hope our work can be useful to other universities who want to re-use this content to jump-start a similar effort to engage with their own communities. We hope to someday see the Nudge come full circle and adapt topics offered by other institutions in a mutually supportive network.

Acknowledgements

References

Center for Data and Visualization Sciences (no date). Center for Data and Visualization Science. *Duke University* [online]. Available at: https://library.duke.edu/data/ (Accessed: 22 February 2020)

Constant Contact. (2020). Average Industry Rates for Email as of December 2019. Available at: https://knowledgebase.constantcontact.com/articles/KnowledgeBase/5409-average-industry-rates?lang=en_US (Accessed: 25 January 2021)

Cox, A.M. & Pinfield, S. (2014). Research data management and libraries: Current activities and future priorities. *Journal of Librarianship and Information Science*; 46(4), pp.299–316.

Data Services. (no date). Data Services, John Hopkins University [online]. Available at: https://dataservices.library.jhu.edu/ (Accessed: 22 February 2020)

Data Services. (no date). Data Services, New York University [online]. Available at: https://library.nyu.edu/departments/data-services/ (Accessed: 22 February 2020)

Data Services. (no date). Data Services, Washington University in St Louis [online]. Available at: https://library.wustl.edu/services/data/ (Accessed: 22 February 2020)

Fearon, D. Jr., Gunia, B., Lake, S., Pralle, B.E., & Sallans, A.L. (2013). SPEC Kit 334: Research Data Management. *Association of Research Libraries*. Available at: https://doi.org/10.29242/spec.334 (Accessed: 21 February 2020)

Holdren, J.P. (2013). Increasing Access to the Results of Federally Funded Scientific Research. *Office of Science and Technology Policy* [online]. Available at: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf (Accessed: 20 February 2020)

Information Technology (2018). Email Fraud Defense. *University of Illinois at Urbana-Champaign* [online]. Available at: https://techservices.illinois.edu/content/email-fraud-defense (Accessed: 20 February 2020)

Mailchimp. (no date). [online]. Available at: https://mailchimp.com (Accessed: 20 February 2020)

Office of the Vice Chancellor for Research & Innovation (no date). 2019 Research Report. *The University of Illinois at Urbana-Champaign* [online]. Available at: https://research.illinois.edu/sites/research.illinois.edu/files/upload/research_report_digital.pdf (Accessed: 15 March 2020)

Office of the Vice Chancellor for Research & Innovation (no date). By the Number. *The University of Illinois at Urbana-Champaign* [online]. Available at: http://research.illinois.edu/research-illinois/numbers (Accessed: 20 February 2020)

Ohaji, I. K., Chawner, B. and Yoong, P. (2019). The role of a data librarian in academic and research libraries. *Information Research*, 24(4), p. N.PAG. Available at: http://search.ebscohost.com/login.aspx?direct=true&db=lls&AN=140844395&site=ehost-live (Accessed: 14 February 2020).

Research Data Management Service Group (no date). Research Data Management Service Group, Cornell University [online]. Available at: https://data.research.cornell.edu/ (Accessed: 22 February 2020)

Research Data Service (no date). Research Data Service*, The University of Illinois at Urbana-Champaign* [online]. Available at: https://www.library.illinois.edu/rds/ (Accessed: 22 February 2020)

Research Data Services (no date). Research Data Services, Penn State University [online]. Available at: https://libraries.psu.edu/research/research-data-services (Accessed: 22 February 2020)

Research Data Services (no date). Research Data Services, The University of Michigan [online]. Available at: https://www.lib.umich.edu/research-data-services (Accessed: 22 February 2020)

Research Data Services (no date). Research Data Services, The University of Minnesota [online]. Available at: https://www.lib.umn.edu/datamanagement (Accessed: 22 February 2020)

Strategic Plan Working Group (2013). The Illinois Strategic Plan, The University of Illinois at Urbana-Champaign [online]. Available at: https://strategicplan.illinois.edu/2013-2016/goals.html (Accessed: 20 February 2020)

Wiley, C., & Mischo, W.H. (2016). Data Management Practices and Perspectives of Atmospheric Scientists and Engineering Faculty. *Issues in Science & Technology Librarianship*; 85(1) [online]. Available at: https://doi.org/10.5062/F43X84NJ (Accessed: 20 February 2020)

Endnotes

[1] Daria Orlowska is Data Librarian at Western Michigan University, email: daria.orlowska@wmich.edu

[2] Colleen Fallaw is Research Programmer for the Research Data Service, University of Illinois at Urbana-Champaign, email: mfall3@illinois.edu

[3] Yali Feng is Behavioral Sciences Research and Data Services Librarian at University of Illinois at Urbana-Champaign, email: yalifeng@illinois.edu

[4] Livia Garza is former Pre-professional Graduate Assistant at Research Data Service, University of Illinois at Urbana-Champaign, email: liviag2@illinois.edu

[5] Ashley Hetrick is former Assistant Director of Research Data Engagement Education at Research Data Service, University of Illinois at Urbana-Champaign, email: ahetrick@illinois.edu

[6] Heidi Imker is Director of Research Data Service, University of Illinois at Urbana-Champaign, email: imker@illinois.edu

[7] Hoa Luong is Associate Director of Research Data Service, University of Illinois at Urbana-Champaign, email: hluong2@illinois.edu