Exploring Visual Representations to Support Data Re-Use for Interdisciplinary Science

Andrea Wiggins
University of Nebraska at Omaha, USA.
wiggins@unomaha.edu

Melissa A. Kenney
University of Maryland, College Park, USA.
kenney@umd.edu

Alyson Young
Indiana University-Purdue University Indianapolis, USA.
youngaly@iupui.edu

ABSTRACT

Data discovery and re-use remain challenging, particularly for interdisciplinary research, but visual representations of data may have potential to better support these tasks than text-only metadata records. Our multi-stage exploratory study used think-aloud and intercept interviews to evaluate how scientists across multiple disciplines interpret the content and evaluate the credibility of climate indicators designed for general audiences as a way to understand the potential of visual representations for supporting interdisciplinary research. Climate indicators provided a convenient alternate representation for describing a data set, featuring a visualization of the data along with vernacular textual description. Our contributions include (i) findings on the preferences of scientists from multiple disciplines for establishing the credibility and assessing the content of data for re-use, (ii) the unexpected observation that a minimal set of core metadata was considered adequate by researchers examining data from outside of their specialties, and (iii) discussion of the implications for the design of data discovery interfaces and future research to support data discovery and re-use by increasingly diverse scientific data consumers.

KEYWORDS

Data representations, metadata, provenance, interdisciplinary science, climate indicators.

INTRODUCTION

Today's scientists are increasingly encouraged to share data, with interdisciplinary research on the rise (Krumholz, 2015; Tenopir et al., 2011; Van Noorden, 2015). The practice of sharing research data has grown substantially in short order (Tenopir et al., 2015; Force & Robinson, 2014), as has emphasis on replicability and open science as demonstrations of rigor and efficient resource use (Farnham et al., 2017). Data discovery and re-use remain difficult in practice (Poole, 2015), although attitudes are increasingly favorable (Curty et al. 2017) and substantial success has been achieved in isolated cases, such as that of clinical trial data (Coady et al., 2017) and bird abundance and distribution data from eBird (Lagoze, 2014). Actual statistics on data re-use are limited, but lags behind sharing; Bishop and Kuula-Luumi (2017) reported 40% of qualitative data sets were never downloaded, and about 25% of data is used just 1-10 times.

Our research centered on the use of climate indicators as visual representations of data with potential to assist researchers in identifying the potential utility of data for interdisciplinary research. We wanted to better understand how researchers evaluate data prior to exploring it in detail and learn what supporting information was required to establish the credibility of the information. Since data exploration is an effortful task, particularly for increasingly large data sets, we argue that the "first blush" initial evaluation of data is an important step in the process of data re-use that is currently overlooked. With current research trends emphasizing the value of data sharing for re-use, there is increasing need for better understanding of and stronger support for data discovery and evaluation in this crucial initial step for data re-use (Sun & Khoo, 2017).

Climate indicators are regularly updated visual representations that summarize key climate phenomenon and impacts relative to a baseline of change (Kenney et al., 2016). To establish a system of indicators as part of a sustained U.S. National Climate Assessment (Buizer et al., 2013), Kenney et al. (2016) describes a process involving over 200 scientific experts to recommend indicators of climate changes, impacts, and responses. This effort used a rigorous process to establish implementation recommendation (Kenney et al., 2014; Kenney et al, in review). The newly released U.S. Global Change Research Program climate indicators (Janetos & Kenney, 2015; www.globalchange.gov/explore/indicators) provided an ideal test case to explore a secondary audience benefit, interdisciplinary data discovery for re-use.

Climate indicators provided a convenient alternate representation to the typical text-only metadata record for describing a data set. Unlike a table of metadata about a data set, climate indicators include a visualization of the data designed to clearly show key trends, along with vernacular textual description, rather than highly technical details of the data (e.g., examples given in Sun & Khoo, 2017). Through our partnership with the U.S. Global Change Research Program, we also sought to identify important design considerations for improving the utility and understandability of data representations intended to communicate long-term, large-scale trends in climate science (Gerst et al., 2017). Although the primary audiences for most climate indicators are members of the public and policy makers, this form of data representation may also have potential value for audiences such as scientists, particularly when they are not specialists in the area that the indicator represents.

Our research therefore focused on data representations as mediators of (i) data discovery, (ii) initial evaluation, and (iii) decisions to further explore data for re-use. These three functions alone require substantial information to support through standard data repository metadata; more complete "research objects" to support replication are still uncommon (Bechhofer et al., 2010). The design of data representations, which could be considered a type of metadata, from a functional perspective, is also likely to influence the processes they mediate (Sun & Khoo, 2017). Through exploratory multi-stage research, we examined the way that scientists across multiple disciplines interpreted the content and evaluated the credibility of climate indicators as a type of visual representations of data. In the initial stages, our primary research question was, "How do scientists assess the content and credibility of climate indicators?" The respondents' strong emphasis on the importance of metadata led to assessment of scientists' general expectations for metadata to accompany data representations, answering the question, "What are scientists' expectations for metadata to support evaluation of data for re-use?"

BACKGROUND

Promoting data re-use is a complex challenge for several reasons, such as (i) research incentive structures that often prioritize novel data collection over re-use and fail to reward replication studies (Farnham et al., 2017), (ii) general lack of awareness of existing applicable data (Tenopir et al., 2015), and (iii) limited or inconsistent metadata (Marcial & Hemminger, 2010; Birnholtz & Bietz, 2003), among others. The specific issue we examine in this work is that of evaluating data as an input to decisions around data re-use. Data re-use has been studied in health (Joo, Kim & Kim, 2017) and biomedical sciences (Federer et al, 2015; Rolland & Lee, 2013), social sciences (Faniel, Kriesberg & Yakel, 2015; Sun & Khoo, 2017; Yoon & Kim, 2017), archeology (Faniel et al., 2013), and engineering (Faniel & Jacobsen, 2010), among others, but rarely for multi- or interdisciplinary research (Parsons et al., 2011). Most of these studies focus on attitudes about re-use, rather than documenting actual data re-use practices or functional criteria for selecting data for re-use.

By extension, we therefore have relatively little understanding of how data are evaluated for re-use, particularly by non-specialist users such as students, since the prior work has focused on single-discipline studies where the norms within each discipline and scarcity of studies limits transferability of results. Among very few exceptions, the work of Gregory et al. (2018) on search processes for data re-use directly evaluates data discovery behavior across multiple disciplinary specialties, Carlson and Anderson (2007) examined data sharing and re-use behaviors across multiple disciplines, and Young and Lutters (2015) examined the re-use of interdisciplinary data sources for meta-study research in the environmental sciences. Cragin and Shankar (2006) and Baker, Duerr and Parsons (2015) noted the value for non-specialist data consumers, while another set of studies considered students' experiences with data re-use (Faniel, Kriesberg, & Yakel, 2013; Kriesberg et al., 2012).

While metadata is not the only information that plays into data re-use decisions (Zimmerman 2007, 2008), it provides initial impressions of credibility and supports development of trust in both the data and the repository (Yakel et al., 2013) that is necessary to support a re-use decision. Similarly, visual design has been seen to influence judgments of credibility in studies of websites (Fogg, 2003). Despite this, the metadata recorded often does not provide sufficient descriptions of the context in which the data were collected, such as choices of metrics and instrumentation for observation to enable prospective re-users to interpret the data for effective secondary analysis (Borgman et al., 2009; Borgman et al., 2007; Wallis et al., 2010). Metadata has also been found to be a source of friction, particularly between scientific collaborators from two or more disciplines, often impeding data sharing (Edwards et al., 2011). While the lack of contextual details is typically less of an issue for proximal disciplines, as similarities in background allow researchers to infer relevant information, scientists from adjacent communities encounter more difficulty due to vastly different scientific practices and use of instrumentation (Borgman et al., 2009; Zimmerman, 2007). Consequently, the process of understanding how variables are constructed is often a highly collaborative and time-consuming task (Rolland & Lee, 2013), requiring engagement with original study staff to obtain the contextual details needed to assess data for relevance (Lee et al., 2009; Edwards et al., 2011).

In light of these challenges, we see indicators and their style of data representation as a potential solution strategy at the data discovery stage of this process. Climate indicators are intended to provide input to decision-making by non-expert audiences who do not have intimate knowledge of the science but need to see credible evidence of causal mechanisms in order to formulate practical and policy interventions. Climate indicators are created by subject matter specialists to communicate complex information with a goal of decision support, which is functionally similar to the needs for data representations to support data re-use. In a data re-use scenario, the representations are created by experts in the data to communicate its potential and content to informed audiences who do not necessarily have intimate knowledge of the data, similar to policy makers in the case of climate indicators. In both scenarios, images accompanying text may be useful for communicating complex information quickly and efficiently, one of the main advantages to data visualization, and professionally designed visual representations of data can support judgments of credibility (Fogg, 2003). Whether visual representations of data are adequate or effective at communicating credibility for would-be data users, however, is not well understood.

METHODS

For this exploratory study, we employed a multi-stage research design, beginning with a small number of extended think-aloud interviews and proceeding to two rounds of intercept interviews at large scientific conferences as the research evolved. Although all stages of the research involved interviews with scientists, think-aloud and intercept interview methods are substantially different in execution and data that they yield, making this a mixed-methods study. Each stage of the research informed the design of the subsequent data collection and analysis.

Think-Aloud Interviews

We conducted 5 in-depth interviews, averaging 90-120 minutes, with scientists in several disciplines (astrophysics, forestry, ecology, marine biology) using a think-aloud protocol that focused primarily on understanding their perceptions of climate indicators from several federal agencies. While interacting with a set of indicators on the U.S. Global Change Research Program (USGCRP) website and extended metadata content on the Global Change Information System (GCIS; data.globalchange.gov), we asked participants to describe their thoughts about the system, the understandability of the climate indicators, and the level of trustworthiness or credibility of the data presented in the indicators. Sampling was constrained due to limited availability during the summer, so interviewees were federal research scientists based at NASA, US Forest Service, and Smithsonian Institution; they included 3 women. The think-aloud method was well suited to answering the research question focused on how scientists evaluated the content and credibility of climate indicators but yielded broad-ranging responses. Based on the initial findings, summarized in the following section, we redesigned our study approach for a more focused scope with an intercept interview design.

Intercept Interviews

After the initial interviews focused on climate indicator designs, we narrowed the scope of inquiry to the use of metadata to establish credibility from the perspective of research scientists. To increase our sample size and diversity, we conducted intercept interviews at the annual meetings of the 2015 Ecological Society of America (ESA) and the 2015 American Geophysical Union (AGU). Both of these meetings draw substantial attendance (approximately 7,000 at ESA and 25,000 at AGU) and include several subdisciplines and application spaces within the broad fields of environmental and Earth sciences. Researchers attending these conferences represent all career stages, employment sectors, and a wide range of nations. The intercept interview format involves "intercepting" random respondents at a public venue and asking a few questions.

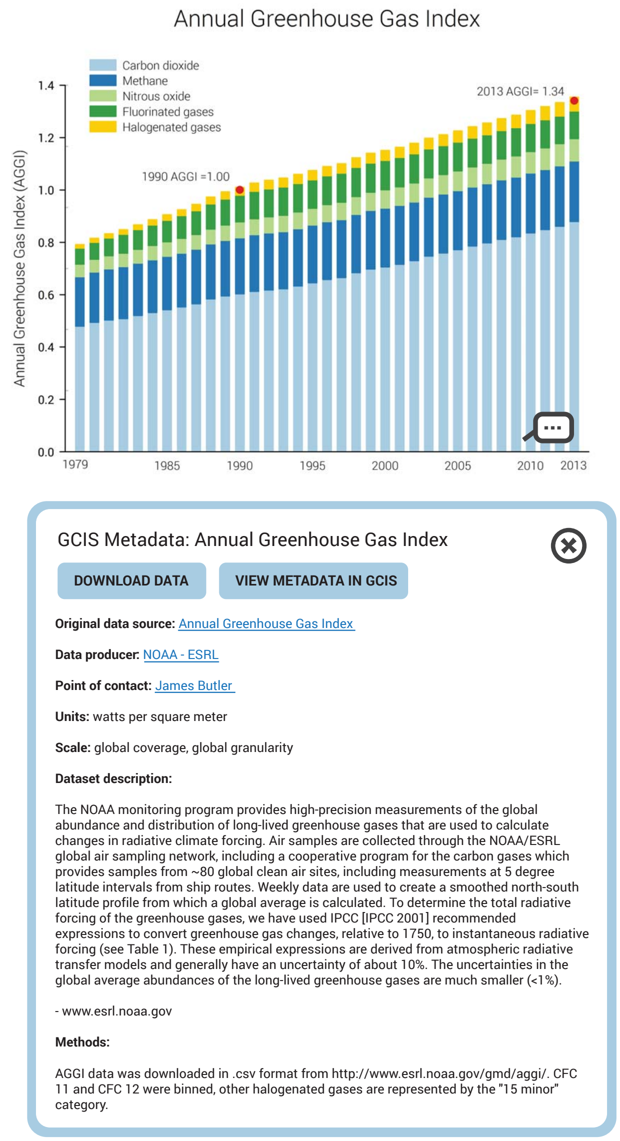

Figure 1. Metadata mockup for the Aggregated Greenhouse Gas Index climate indicator, showing both the indicator's data visualization and the "pop-up" metadata details.

For both meetings, interview team members were positioned near the exhibitor booth for DataONE (which supports data sharing and re-use) in the conference exhibit halls. Interviewers approached attendees at random and asked for five-minute responses to two questions, noting perceived gender and employment sector from affiliations on name badges to verify that a range of perspectives were represented. The ESA intercept interviews included 56 researchers, with 24 females and one of indeterminate gender, and sectors included 43 in academia, six in government, three nonprofit, three private, and one unaffiliated. The AGU intercept interviews included 49 researchers, with 16 females and three of indeterminate gender; sectors included 26 in academia, eight in government, five private, three nonprofit, one K-12 educator, and six undetermined.

After obtaining oral consent, different interview protocols were followed for each meeting. For the ESA interviews, interviewees were shown a randomly selected global climate indicator from the USGCRP website (e.g., vibrio infections, sea surface temperatures, arctic sea ice extent, start of spring, forest cover, atmospheric carbon dioxide) on an iPad and asked to examine the image and text. Interviewees were asked either "How would you interpret the information shown here?" (version 1), or "How would you assess the credibility of this information?" (version 2) Following their initial responses, interviewees were then asked, "What information is missing that would help you interpret it/make that assessment?" 18 interviewees were asked the version 1 questions, 32 were asked the version 2 questions, and six responded to both (3 questions in total).

Based on the results of the ESA intercepts, in which most respondents focused on data provenance in the form of metadata as a primary requirement for information credibility (which was available only through a platform separate from the climate indicators), a new intercept protocol was developed for the AGU meeting. Interviewees were first asked, "In a moment, I'll show you a paper version of a scientific visualization that would be viewed online. In general, what kinds of information or metadata about the visualization would you want to see?" Following their response, alternating participants were shown a high-fidelity mockup of a metadata pop-up interface alongside the climate indicator for either ocean chlorophyll concentrations or aggregated greenhouse gas index (Figure 1), and asked, "How well does this metadata meet your expectations for documenting a scientific visualization like the one here?" The ESA intercept interviews answered the research question focused on scientists' evaluations of credibility and content, while the AGU intercept interviews focused on scientists' expectations for metadata to accompany visual representations of data, as it was difficult to effectively prompt consideration of a data re-use scenario.

Analysis

Deductive and inductive content analysis was applied to all data (Neuendorf, 2016). The deductive analysis used a coding schema based on the web credibility literature (e.g., Barry & Schamber (1998), Fogg (2003), Kubiszewski, Noordewier, & Costanza (2011), Metzger (2007), Rieh & Danielson (2007)). Examples of concepts used to code responses included: accuracy, validity, authority, provenance/metadata, objectivity, informativeness, persuasiveness, quality, trustworthiness, clarity, and consistency. Deductive coding was completed by multiple team members who coded at the level of the entire response to each prompt for the intercept interviews and argued to consensus. Inductive codes were added based on emergent themes, including: intended audience, methodology for raw data collection, and methodology for indicator creation, as well as positive, negative, or neutral assessments of the metadata mockup.

FINDINGS

Several themes were identified at each of the three stages of the study, which focused the evolving exploratory research and refine the inquiry at each successive stage, with increasing emphasis on metadata as a key component of data representations. Table 1 summarizes the research questions and themes identified at each stage of the research.

Research stage	*Research questions addressed*	*Themes*
1: Think-aloud interviews	How do scientists assess the content and credibility of climate indicators?	● Metadata is required to assess data ● Prior knowledge of data impacts assessments ● Machine-readable metadata is not suitable for humans
2: ESA intercept interviews	How do scientists assess the content and credibility of climate indicators?	● Metadata is required to assess data ● Prior knowledge of data impacts assessments ● Disciplinary preferences for metadata ● Data source impacts trust ● Concerns over intended audience
3: AGU intercept interviews	What are scientists' expectations for metadata to support evaluation of data for re-use?	● Metadata is required to assess data ● Disciplinary preferences for metadata ● Data source impacts trust ● Concerns over intended audience ● Basic metadata is adequate for initial data assessments

Table 1: Questions addressed and themes identified across stages of the study

Stage 1: Data representation design and metadata presentation

The initial think-aloud interviews brought three themes to our attention. First, the context of production and processing for data is critical to its evaluation: data presented without metadata is not adequate for scientists' uses. This reflected a strong implicit norm across disciplines of providing basic details about data production and processing. For example, when examining the indicator for Frost Free Season, one participant indicated that he would want to see "details on how [the data] was collected, by whom… there isn't much information from what I saw regarding how they were [collected]... I don't know how that information was processed over time. Was the same methodology employed over time?"

Second, the machine-readable metadata system that was available for interviewees to explore was inadequate for human needs, although well suited to other purposes. Our interviewees consistently reported a strong need for an additional human-friendly interface with simplified navigation and comprehensive views of details relevant to the indicator. As another participant noted, when trying to find additional details about the Sea Ice Extent indicator, "... you shouldn't have to click an image to find the source, to find tabular data … then I have to click on the table, and I get the raw numbers, and then I get what report this appeared in, but I still don't get where this came from." For the GCIS, limited curation of existing machine-readable metadata would be adequate to serve this need.

Third, we observed different types of reactions from our interviewees to data that were and were not familiar to them. This was only partially surprising, as background knowledge is important for evaluating data, but the tone of responses was unexpectedly distinctive. The interviewees were more critical of the representations of data with which they were already familiar. For example, a forest ecologist made extended critical comments about the definitions of a "forest" built into the data that undergirded the Forest Cover indicator; no other interviewee had a similar reaction to this data. As reported below in our analysis of the intercept interviews, these were not isolated instances.

Stage 2: Metadata is essential to establishing scientific credibility

The intercept interviews at the ESA conference yielded several new observations. The most dominant theme, repeated by the majority of respondents, was the necessity of including metadata indicating the source of data and methods of processing. Common critiques of the indicators included statements like, "needs to be more references about who generated the graph and text" and "I want to see primary literature cited, where's data from?" The consistency of these responses was overwhelming, with associated codes for "Source" and "Provenance" occurring most often.

We continued to see differences in scientists' evaluations of the indicators based on whether the data were relevant to their specialty. We had no way of knowing specialties in advance, but reactions to the indicators clearly identified indicators that matched respondents' expertise. One respondent directly connected the credibility of the representation to prior knowledge, saying that "I believe it, but would need more information on the methods if I hadn't been familiar with it already." Responses were more critical when familiar with the data or related field. Those who expressed distaste for the forest cover data set were all clearly disciplinary "insiders" who knew the data set's limitations, while no other respondents made such comments on this indicator. For example, responses to the forest cover indicator ranged from a flat "I don't trust [this data set]" to a completely neutral evaluation accompanied by the caveat, "but I don't know much about [this data set]," clearly demonstrating the importance of background knowledge to evaluating the trustworthiness of data and visual representations of it.

This connected to a subtle theme that emerged during analysis: the source of the data (typically the organization producing it) was taken as a proxy for the methods of data production and interpreted as a signal of trustworthiness. In some cases, even when the specific data set was not familiar, respondents clearly recognized the methods through which these data products were produced due to the organization producing them. These responses were framed in terms of trust in the organization conducting the data collection or the instrument used in the process (e.g., LANDSAT), which bears strong similarity to results reported by Yakel et al. (2013), who reported distinctions in trust in the data versus the repository providing it. Several respondents also described a hierarchy for their trust of data based on who had produced it. Generally speaking, federal research data was considered credible; for example, one respondent said of the indicator source, "since it's government [research], I believe it." Other criteria for evaluation were clearly specific to the respondents' specialties; one respondent who worked primarily on computational modeling, for example, said hardware configuration details were critical for data re-use.

To our surprise, many scientists' interpretations of the factual content of the climate indicators were faulty, particularly when viewing indicators that represented unfamiliar topics or data, which also speaks to the importance of background knowledge in evaluating data. In some cases, they openly admitted to uncertainty in drawing a conclusion from the indicator, and their primary critique was often that the trend or purpose of the indicator was unclear or the implications of the trends and conclusions to be drawn were not obvious enough for confident interpretation. These concerns were raised alongside both correct and incorrect interpretations. One of the more pointed comments was a response to an indicator showing reduction in sea ice extent: "What does it mean to lose a million square miles of sea ice?" While the indicators' textual content satisfied some respondents seeking interpretation and meaning, others remained uncertain when trends were not apparent at a glance.

Finally, respondents commented with unexpected frequency on the intended audience for the representation (14 of 56 respondents) but primarily when struggling to interpret the indicator. These comments most often reflected on specific indicators, such as Heating and Cooling Degree Days, which was considered especially confusing by respondents, or focused on specific terminology, like radiative forcing in the Aggregate Greenhouse Gas Index. The respondents tended to couch their uncertainty about their interpretations by stating that the representation was adequate for their needs but might be less suitable for public and policy audiences. A respondent who was confused about the concept of radiative forcing claimed the indicator had "good credibility to a scientist, but I'm not sure for decision-makers." Another commented on the clarity of the same indicator, saying that "a lot of people don't know what radiative forcing is; I don't know and I studied it!" Overt concern for intended audience to indirectly indicated the respondents' own difficulty with interpretation: if a trained scientist struggled to interpret a representation, they were doubtful that policy makers or members of the general public would understand it. However, one respondent acknowledged that scientists may overthink their interpretations due to specialized knowledge, commenting that "cooling is potentially confusing to a scientist but maybe not so confusing to the average person."

Stage 3: Basic metadata is adequate for initial exploration

Much like the ESA respondents, the AGU attendees also identified specific sub-disciplinary preferences for metadata in their responses to the first prompt about what metadata they would expect to see accompanying a data visualization. Where the ESA attendees repeatedly mentioned error bars and sampling sites, AGU attendees frequently asked for uncertainty estimates, instrumentation details (e.g., for satellite products), and temporal and geographic information. The metadata they envisioned being useful reflected on the norms within their own specialties. One individual specified, "platform and physical devices used for gathering data, and processing to turn raw sensor data into products," while several others mentioned models.

Like the ESA attendees, AGU respondents placed consistent emphasis on metadata that provides information on research methods and the parties responsible for data collection and analysis, with a strong bias toward the identity of the source as a shortcut to establish credibility, often noting that "government versus private, or well-known versus unknown" sources of data would be evaluated differently; some also distinguished between parties collecting the data and those distributing it, similarly to findings from Yakel et al. (2013). A typical summary of expected information was, "when it was collected, who collected it, and where it was collected." A few respondents noted the importance of the intended use, and one person clearly understood the challenges of data re-use, mentioning need for more detailed column headers than typically expected.

Since the indicator metadata mockups that respondents were shown did not usually match their specific disciplinary expertise (again, we had no way to gauge this in advance), a few respondents echoed the concern about the purpose of and audience for indicators, but they were not asked to interpret the indicator content and made less indication of confusion than ESA attendees. To paraphrase a common response, "this data isn't from my field, so the basic metadata looks good enough for an initial evaluation of the data." This contrasted the specificity of their own desired details for imagined representations related to their research focus. The overall responses to the metadata mockups indicated adequacy for entry-level evaluation of the data, with 34 positive, 17 neutral, and 10 negative assessments for the question of whether the mockup met expectations (responses could be coded as both positive and negative if they contained both praise and critique). The generally positive response to the metadata mockups further confirmed the importance of metadata that participants reported in prior stages of the research.

When we examined the deductive coding for the responses to each of the questions, we saw additional evidence of the themes already highlighted. When discussing expectations for metadata, respondents referred to the authority of the content and its source in conjunction with presentation, provenance, informativeness, methods for data collection, and trustworthiness. They were more concerned about the authority of the source and provenance of the data collection methods than information about how the data representation was created. Provenance (metadata) was desirable primarily to demonstrate authority, which appeared to be the preferred source of credibility, as other credibility concepts were rarely mentioned besides sources and methods. Content presentation was also heavily emphasized in combination with clarity, authority, coverage (spatial/temporal range/scale), and provenance information. In the responses to the metadata mockups, strong emphasis was placed on presentation, especially for content clarity and informativeness. The positive assessments focused primarily on its presentation and informativeness, while negative assessments were much more diffuse, suggesting that they reflected personal preferences.

DISCUSSION

Our findings from this study of climate indicators as data representations repeatedly highlighted the utility of basic, standard metadata as an initial point of evaluation and minimum criterion for establishing credibility. As shown in Table 1, several themes were repeated and confirmed in each stage of the research. Contrary to prior work (Edwards et al., 2011), we saw more indications of metadata frictionlessness than metadata friction in our respondents' replies, likely because our respondents had to review the representations as scientific generalists, not disciplinary specialists, and their evaluation focused on a "product" rather than a "process" orientation to the metadata. Our discussion focuses on the tensions between metadata friction and frictionlessness, as well as implications for design and research from this exploratory work.

Metadata friction and frictionlessness

As noted, the responses here primarily reflect a "product" perspective on metadata, as our study only evaluated the initial step of data discovery for a process-oriented perspective and how study participants read and interpreted these visual representation products. Like prior work that reported on data re-use within specific disciplines and metadata friction among specialists (Edwards et al., 2011), we also saw that members of distinct scientific communities have different requirements or "wish lists" for the specific metadata they prefer to use for evaluating data and its potential for re-use, which increase the so-called friction to document and share data. At the same time, we were surprised to see the degree of commonality in core details that all scientists seemed to consider essential for basic description of a data representation. In this respect, we observed metadata frictionlessness in the high-level descriptions that respondents reported would be adequate for certain needs, and strong agreement on many of the basic details, which imply potential for reduction in expected metadata friction to support initial discovery. However, as discussed further below, we cannot determine whether the frictionlessness we observed was impacted by the inclusion of a visual representation of data as well as textual description, since most prior studies have relied on text-only information about data (e.g., Sun & Khoo, 2017), and further work would be needed to clarify whether visual representations provide a faster or easier way to evaluate data for re-use. In addition, our primary focus was on data discovery as opposed to other steps in the data re-use process, and it is not apparent from these results whether visual representations would have benefit for reducing metadata friction at other stages of the data re-use process.

The responses from the ESA and AGU meetings also suggested that researchers didn't make a strong effort to critically evaluate data that was too far removed from their specialty. There are several potential explanations: they may not have felt qualified to make the judgment, the representations we provided may not have given adequate detail at the outset, or the format of the intercept interviews may not have been conducive to a critical evaluation of the data. We were also surprised at the frequency with which trained scientists gave incorrect interpretations of carefully designed visual representations of data. If this issue was not solely based on our methods or indicator design, which seems improbable, it may suggest that there are confounding influences of background knowledge and data fluency that should be further investigated as an input to designing data repositories. In this respect, metadata friction or frictionlessness at the data discovery stage may be far less important than the researcher's background knowledge or their fluency in "reading" data representations, and the design of visual representations could have far more impact on assessments of data for re-use than details embedded in text.

Implications for design

As we refined our inquiry throughout this study, we saw that coupling a visualization of data with a textual summary and human-friendly provenance or metadata was preferred by researchers to support initial assessment of content and credibility, which are important to evaluating data for potential re-use. Our interviewees and ESA respondents insisted that metadata must always be packaged with other summary details, should not be buried within text blocks, and that it should be one click away, not several. The positive responses to the metadata mockups from AGU respondents was premised on a "one-click" model for retrieving additional details, which would make the information easy to obtain without distracting non-technical audiences.

The limitations of using climate indicators for this exploratory study also leaves several points for future work to support effective design of visual representations for data discovery and re-use. In particular, additional study is needed on the use of visual representations versus text-only descriptions; in our ESA interviews, many initial critiques based solely on the image and proximal text were resolved when the respondents reviewed the full text block accompanying the visualization. However, we do not know whether a text-only description would have been as easy to assess within minutes in a busy conference exhibitor hall, or if their interpretations of the content would have been more or less correct without the visualization. We can confirm, however, that respondents put a great deal of emphasis on the specifics of the presentation of information, which suggests the typical data repository record formats, a plain table of text, may be underselling the utility of the data on offer.

We also recommend additional work to explore researchers' responses across a wider range of disciplines, to confirm the utility of visual representations for data discovery, verify agreement on the basic details of data representations, and identify discipline-specific preferences. Replicating this study with researchers from social and other biophysical sciences would likely reveal a core set of key details that are always expected, while further clarifying the additional information of particular interest to specific scientific communities. Identifying the particulars preferred by each discipline would be useful for optimizing interfaces in disciplinary data repositories to support effective evaluation of data by target communities. It could also offer a novel way to assess similarities and differences between scientific disciplines, based on degree of agreement on core elements of research data, to support comparisons of data practices between fields in future studies.

As noted previously, most prior work that focused on supporting data re-use has been constrained to data and specialists within a single field, as opposed to evaluating whether there are general aspects of data representation that transcend disciplinary boundaries. We believe that our results indicate that this is indeed the case, which suggests that data representations may be more useful if they explicitly support multiple levels of detail, with general metadata for general consumption and additional specialist-specific information for disciplinary audiences. This could make browsing for data more efficient for broader audiences without losing the specific decision information that disciplinary specialists rely upon.

Implications for research

This study highlighted the contrasts between seeing data like a generalist as opposed to taking a specialist's view of data. Differences in the expected content, desired provenance and methods details, and even emotive responses to data that were near to versus distant from the respondents' research were striking. Because increasing emphasis is being placed on data sharing and re-use in scientific communities, our findings suggest a need for further research on the utility of alternative metadata and provenance representations. Specifically, research on representations with different levels of granularity and for a wider range of use cases, as discussed by Baker et al. (2015), could also support assessment of utility for collaborative research. An especially important use case for non-specialist data representations is students in training, who are increasingly learning their trade using existing data sources prior to developing the degree of subject specialization that characterizes the participants in most prior studies of metadata and data re-use.

We also believe that these findings point to the potential for data to serve as a neutral touchstone for fostering interdisciplinary collaboration between researchers in proximal disciplines. This bridging function is most likely to emerge when none of the collaborators "owns" the data as a result of specialized data expertise. Shared expectations for high-level information to describe data also suggests that collaborative data exploration exercises could put partners on equal footing in a new investigation, while also revealing their disciplinary differences in productive ways. For example, we can imagine a pair of new research partners simultaneously exploring a data set. The ecologist wonders aloud, "Where do you think we can find information on the sampling sites for this data set?" to which the geoscientist replies, "Oh, I hadn't thought about that yet, because I always look for details about the resolution first. But those are fundamentally related, so let's find out." Although imaginary, such an exercise could help an interdisciplinary team establish a shared vocabulary for research from the outset with greater sensitivity to different uses of similar terminology, and the experience of jointly exploring a data set that is unfamiliar to each party would inevitably highlight the expertise and interests each person brings to the partnership. Both of these potential outcomes are recognized as key factors in building effective interdisciplinary teams (Stokols et al., 2008).

The limitations of the study included its exploratory nature and the disciplines represented, as most participants' backgrounds were in relatively close intellectual proximity. The think-aloud interviews relied on convenience sampling but provided focus for the intercept interviews. The initial intercept interviews focused on the credibility of data as presented in indicators more than the utility of the presentation for data discovery, because the idea of browsing for data seemed foreign to most researchers. Additional research could provide more conclusive results about the utility of visual representations for supporting data discovery and re-use. Finally, the data and indicators presented to interviewees were quantitative and spatial in nature, so the findings reported here may have limited transferability to qualitative and textual research data. The indicators were also all carefully designed by professionals who were more focused on accuracy in representation than we would expect of a typical researcher preparing a data deposit. The data sets typically represented were well known by researchers in associated disciplines; data representations for more obscure data sets may elicit different responses.

CONCLUSION

In this multi-stage exploratory study of climate indicators as visual representations of data for re-use, we found that scientists rely heavily on traditional metadata as a signal of credibility, and often used the affiliated organizations that produced the data, such as federal agencies, as a proxy for methods and quality. Despite their training, many scientists had difficulty interpreting the content of carefully designed visual representations of data from other disciplines than their own, and under these circumstances, tended to report concerns about the suitability of indicators for non-scientific audiences. Given the emphasis that researchers placed on metadata, we further examined scientists' expectations for metadata to support data discovery and evaluating data for re-use. Contrary to expectations, a minimal set of core details were considered adequate by researchers examining data from a generalist perspective, although respondents also highlighted common discipline-specific preferences for metadata to contextualize data sets. We also saw that a greater degree of specificity in data description was expected for data that were closely related to researchers' specialties, while data unfamiliar to them was not met with the same scrutiny. This study complemented related prior work by examining scientists' evaluations of data representations and preferences for metadata in multiple disciplinary audiences, rather than a single disciplinary space, and is unique in evaluating visual representations of data as a support for data discovery.

These results suggested that visual representations of data may be useful for data discovery, which merits further study, and could also have value for mediating initial development of interdisciplinary collaborations. However, visual representations cannot stand alone: traditional metadata on data collection and analysis methods is always expected by scientific audiences. We suggest that in the design of data discovery interfaces, visual representations of data alongside high-level metadata, with one-click access to detailed "specialist-level" metadata, may better serve the interests of both general and disciplinary audiences that have different expectations for information about data. We also suggest that further study of metadata and data representations to support discovery and re-use is needed, particularly with a non-disciplinary and non-expert focus, to support the growing diversity of scientific data consumers, including students, interdisciplinary teams, and the general public.

ACKNOWLEDGEMENTS

This work was supported in part by NSF HRD grant 1008117 and NOAA grant NA14NES4320003. We thank DataONE for sharing space at ESA and AGU, annd our collaborators for contributions to these results, including J. Felix Wolfinger, Michael D. Gerst, Ainsley Lloyd, Amanda Lamoureux, Annibel Rice, Carly Brody, Sarah Ramdeen, and Emily Bondank.

REFERENCES

Baker, K. S., Duerr, R. E., & Parsons, M. A. (2015). Scientific knowledge mobilization: Co-evolution of data products and designated communities. *International Journal of Digital Curation,* 10(2), 110-135.

Barry, C. L., & Schamber, L. (1998). Users' criteria for relevance evaluation: a cross-situational comparison. *Information processing & management*, 34(2), 219-236.

Birnholtz, J.P., & Bietz, M.J. (2003). Data at work: Supporting sharing in science and Engineering. In *Proceedings of the International ACM SIGGROUP Conference on Supporting Group Work*, Sanibel Island, FL. New York: ACM Press.

Bechhofer, S., De Roure, D., Gamble, M., Goble, C., & Buchan, I. (2010). Research Objects: Towards Exchange and Reuse of Digital Knowledge. In *The Future of the Web for Collaborative Science (FWCS 2010)*. Nature Precedings.

Bishop, L., & Kuula-Luumi, A. (2017). Revisiting qualitative data reuse: A decade on. *Sage Open*, 7(1), 2158244016685136.

Buizer J.L., Fleming P., Hays S.L. et al. (2013) *Report on preparing the nation for change: building a sustained national climate assessment process*. National Climate Assessment and Development Advisory Committee.

Carlson, S., & Anderson, B. (2007). What are data? The many kinds of data and their implications for reuse. *Journal of Computer-Mediated Communication* 12, (2). http://jcmc.indiana.edu/vol12/issue2/carlson.html

Coady, S. A., Mensah, G. A., Wagner, E. L., Goldfarb, M. E., Hitchcock, D. M., & Giffen, C. A. (2017). Use of the national heart, lung, and blood institute data repository. *New England Journal of Medicine*, 376(19), 1849-1858.

Cragin, M.H., & Shankar, K. (2006). Scientific data collections and distributed collective practice. *Computer Supported Cooperative Work* 15(2-3). Dordrecht, The Netherlands: Kluwer

Curty, R.G., Crowston, K., Specht, A., Grant, B.W., & Dalton, E.D. (2017). Attitudes and norms affecting scientists' data reuse. *PloS one*, 12(12), e0189288.

Edwards, P. N., Mayernik, M. S., Batcheller, A. L., Bowker, G. C., & Borgman, C. L. (2011). Science friction: Data, metadata, and collaboration. *Social Studies of Science*, 41(5), 667-690.

Faniel, I.M., & Jacobsen, T.E. (2010). Reusing scientific data: How earthquake engineering researchers assess the reusability of colleagues' data. *Computer Supported Cooperative Work*, 19(3-4), 355-375.

Faniel, I.M., Kansa, E., Whitcher Kansa, S., Barrera-Gomez, J., & Yakel, E. (2013, July). The challenges of digging data: a study of context in archaeological data reuse. In *Proceedings of the 13th ACM/IEEE-CS joint conference on Digital Libraries* (pp. 295-304). ACM.

Faniel, I.M., Kriesberg, A., & Yakel, E. (2012). Data reuse and sensemaking among novice social scientists. *Journal of the Association for Information Science and Technology*, 49(1), 1-10.

Faniel, I.M., Kriesberg, A., & Yakel, E. (2016). Social scientists' satisfaction with data reuse. *Journal of the Association for Information Science and Technology*, 67(6), 1404-1416.

Faniel, I.M., & Zimmerman, A. (2011). Beyond the data deluge: a research agenda for large-scale data sharing and reuse. *International Journal of Digital Curation*, 6(1), 58-69.

Farnham, A., Kurz, C., Öztürk, M.A., Solbiati, M., Myllyntaus, O., Meekes, J., ... & Kanninen, L. (2017). Early career researchers want Open Science. *Genome biology*, 18(1), 221.

Federer, L.M., Lu, Y.L., Joubert, D.J., Welsh, J., & Brandys, B. (2015). Biomedical data sharing and reuse: Attitudes and practices of clinical and scientific research staff. *PloS ONE*, 10(6), e0129506.

Fogg, B. J. (2003). Prominence-interpretation theory: Explaining how people assess credibility online. In *CHI'03 extended abstracts on human factors in computing systems* (pp. 722-723). ACM.

Force, M. and N. Robinson, Encouraging data citation and discovery with the Data Citation Index. Journal of Computer-Aided Molecular Design, 2014. 28(10): p. 1043-1048.

Gerst, M.D., Kenney, M.A, Wolfinger, J.F., Baer, A., Wiggins, A., Feygina, I., ... Young, A. (2017). *Effective Visual Communication of Climate Indicators and Scientific Information: Synthesis, Design Considerations, and Examples. A Technical Input Report to the 4th National Climate Assessment Report*. Version 2.0. http://www.umdindicators.org/wp-content/uploads/2018/01/Visual-Design-Guidance-Document-V2.0-1.pdf

Gregory, K., Cousijn, H., Groth, P., Scharnhorst, A., & Wyatt, S. (2018). Understanding Data Retrieval Practices: A Social Informatics Perspective. *arXiv preprint* arXiv:1801.04971.

Janetos, A. C., & Kenney, M. A. (2015). Developing better indicators to track climate impacts. *Frontiers in Ecology and the Environment*, 13(8), 403-403.

Joo, S., Kim, S., & Kim, Y. (2017). An exploratory study of health scientists' data reuse behaviors: Examining attitudinal, social, and resource factors. *Aslib Journal of Information Management*, 69(4), 389-407.

Kenney, M.A., Janetos, A.C., & Gerst, M.D. (in review). Climate indicators for the nation. *Climatic Change*.

Kenney, M.A., Janetos, A.C., & Lough, G.C. (2016). Building an integrated US national climate indicators system. *Climatic Change*, 135(1), 85-96.

Kenney M.A., Janetos A.C. et al. (2014) *National climate indicators system report*. National Climate Assessment Development and Advisory Committee. http://www.umdindicators.org/wp-content/uploads/2016/10/Pilot-Indicator-System-Report_final.pdf

Kriesberg, A., Frank, R.D., Faniel, I.M., & Yakel, E. (2013). The role of data reuse in the apprenticeship process. *Journal of the Association for Information Science and Technology*, 50(1), 1-10.

Krumholz, H.M., (2015). Why data sharing should be the expected norm. *British Medical Journal*, 350: p. H599.

Kubiszewski, I., Noordewier, T., & Costanza, R. (2011). Perceived credibility of Internet encyclopedias. *Computers & Education*, 56(3), 659-667.

Lagoze, C. (2014). eBird: Curating Citizen Science Data for Use by Diverse Communities. *International Journal of Digital Curation*, 9(1), 71-82.

Marcial, L.H., & Hemminger, B.M. (2010). Scientific data repositories on the Web: An initial survey. *Journal of the Association for Information Science and Technology*, 61(10), 2029-2048.

Metzger, M.J. (2007). Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, 58(13), 2078-2091. doi: 10.1002/asi.20672

Neuendorf, K.A. (2016). *The content analysis guidebook*. Sage.

Parsons, M.A., Godøy, Ø., LeDrew, E., De Bruin, T.F., Danis, B., Tomlinson, S., & Carlson, D. (2011). A conceptual framework for managing very diverse data for complex, interdisciplinary science. *Journal of Information Science*, 37(6), 555-569.

Poole, A.H. (2015). How has your science data grown? Digital curation and the human factor: a critical literature review. *Archival Science*, 15(2), 101-139.

Rieh, S.Y., & Danielson, D.R. (2007). Credibility: A multidisciplinary framework. *Annual review of information science and technology*, 41(1), 307-364

Rolland, B. and C.P. Lee, (2013). Beyond trust and reliability: reusing data in collaborative cancer epidemiology research. In *Proceedings of the 2013 ACM International conference on Computer supported cooperative work*. ACM: San Antonio, Texas, USA. p. 435-444.

Stokols, D., Misra, S., Moser, R.P., Hall, K.L., & Taylor, B.K. (2008). The ecology of team science: understanding contextual influences on transdisciplinary collaboration. *American Journal of Preventive Medicine*, 35(2), S96-S115.

Sun, G., & Khoo, C.S.G. (2017). Social science research data curation: issues of reuse. *Libellarium: journal for the research of writing, books, and cultural heritage institutions*, 9(2).

Tenopir, C., et al., (2011). Data Sharing by Scientists: Practices and Perceptions. *PLoS ONE* 6(6): p. E21101.

Tenopir C., Dalton E.D., Allard S., Frame M., Pjesivac I., Birch B., et al. (2015) Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide. *PLoS ONE* 10(8): e0134826.

Van Noorden, R. (2015). Interdisciplinary research by the numbers. *Nature News*, 525(7569), 306.

Yakel, E., Faniel, I. M., Kriesberg, A., & Yoon, A. (2013). Trust in digital repositories. *International Journal of Digital Curation*, 8(1), 143-156.

Yoon, A., & Kim, Y. (2017). Social scientists' data reuse behaviors: Exploring the roles of attitudinal beliefs, attitudes, norms, and data repositories. *Library & Information Science Research*, 39(3), 224-233.

Young, A.L., & Lutters, W.G. (2015). (Re)defining land change science through synthetic research practices. In *Proceedings of the 2015 ACM International conference on Computer supported cooperative work*. ACM: Vancouver, British Columbia, Canada. p 431-442.

Zimmerman, A. (2007). Not by metadata alone: the use of diverse forms of knowledge to locate data for reuse. *International Journal on Digital Libraries*. 7(1-2): p. 5-16.

Zimmerman, A. (2008). New knowledge from old data: The role of standards in the sharing and reuse of ecological data. *Science, Technology, & Human Values* 33, (5). London: Sage.