



Research Data Management in the Humanities: Challenges and Opportunities in the Canadian Context

Stefan Higgins, English, University of Victoria, Canada, stefan.higgins@nserc-crsng.gc.ca

Lisa Goddard, Libraries, University of Victoria, Canada, lgoddard@uvic.ca

Shahira Khair, Libraries, University of Victoria, Canada, skhair@uvic.ca

In recent years, research funders across the world have implemented mandates for research data management (RDM) that introduce new obligations for researchers seeking funding. Although data work is not new in the humanities, digital research infrastructures, best practices, and the development of highly qualified personnel to support humanist researchers are all still nascent. Responding to these changes, this article offers four contributions to how humanists can consider the role of “data” in their research and succeed in its management. First, we define RDM and data management plans (DMP) and raise some exigent questions regarding their development and maintenance. Second, acknowledging the unsettled status of “data” in the humanities, we offer some conceptual explanations of what data are, and gesture to some ways in which humanists are already (and have always been) engaged in data work. Third, we argue that data work requires conscious design—attention to how data are produced—and that thinking of data work as involving design (e.g., experimental and interpretive work) can help humanists engage more fruitfully in RDM. Fourth, we argue that RDM (and data work, generally) is labour that requires compensation in the form of funding, support, and tools, as well as accreditation and recognition that incentivizes researchers to make RDM an integral part of their research. Finally, we offer a set of concrete recommendations to support humanist RDM in the Canadian context.

Ces dernières années, les organismes de financement de la recherche du monde entier ont mis en place des mandats de gestion des données de recherche (en anglais, Research data management, RDM) qui introduisent de nouvelles obligations pour les chercheurs en quête de financement. Bien que le travail sur les données ne soit pas nouveau dans les sciences humaines, les infrastructures de recherche numérique, les meilleures pratiques et le développement d’un personnel hautement qualifié pour soutenir les chercheurs humanistes en sont encore à leurs balbutiements. En réponse à ces changements, cet article propose quatre contributions sur la façon dont les humanistes



peuvent considérer le rôle des « données » dans leur recherche et réussir à les gérer. Tout d'abord, nous définissons les techniques de recherche et de gestion de données et les plans de gestion des données (en anglais "data management plan", DMP) et soulevons certaines questions urgentes concernant leur développement et leur maintenance. Deuxièmement, reconnaissant le statut incertain des « données » dans les sciences humaines, nous proposons quelques explications conceptuelles de ce que sont les données, et soulignons certaines façons dont les humanistes sont déjà (et ont toujours été) engagés dans le travail sur les données. Troisièmement, nous soutenons que le travail sur les données nécessite une conception consciente—une attention à la manière dont les données sont produites—et que le fait de considérer le travail sur les données comme impliquant une conception (par exemple, un travail expérimental et interprétatif) peut aider les humanistes à s'engager de manière plus fructueuse dans la RDM. Quatrièmement, nous soutenons que la RDM (et le travail sur les données en général) est un travail qui nécessite une compensation sous forme de financement, de soutien et d'outils, ainsi qu'une accréditation et une reconnaissance qui incitent les chercheurs à faire de la RDM une partie intégrante de leur recherche. Enfin, nous proposons une série de recommandations concrètes pour soutenir la RDM humaniste dans le contexte canadien.

Introduction

Data work is coming to the humanities. In recent years, research funders across the world have implemented mandates for research data management (RDM) that introduce new obligations for researchers seeking funding. In Canada, the three federal research granting agencies officially launched a research data management policy in March 2021, which aims to support the “collection, documentation, storage, sharing and preservation of research data” (Government of Canada 2022). Humanists may not be used to thinking of their research materials as “data,” and may also lack the kind of institutional and technical support typically afforded to the sciences and other traditionally data-heavy fields (Estill 2020; Evalyn et al. 2020; Rockwell et al. 2020). Although data work is not new, digital research infrastructures, best practices, and the development of highly qualified personnel to support researchers in the humanities are all still nascent (O’Donnell 2020; Siemens and Arbuckle 2020a; Siemens and Arbuckle 2020b).

Responding to these changes, this article offers five contributions to how humanists can think about data in humanist research and succeed in its management. First, we define RDM and data management plans (DMP) and raise some exigent questions regarding their development and maintenance. Second, acknowledging the unsettled status of “data” in the humanities, we offer some conceptual explanations of *what data are*, and gesture to some ways in which humanists are already (and have always been) engaged in data work. Third, we argue that data work requires conscious *design*—attention to how data are produced—and that thinking of data work as involving design (e.g., experimental and interpretive work) can help humanists engage more fruitfully in RDM. Fourth, we discuss the need for RDM training for humanist researchers. Finally, we argue that RDM (and data work, generally) is *labour* that requires compensation in the form of funding, support, and tools, as well as accreditation and recognition that incentivizes researchers to make RDM an integral part of their research.

This article constructs its arguments by drawing on proceedings from Research Data Management for Digitally Curious Humanists (RDM4H) a virtual event sponsored by the Social Sciences and Humanities Research Council (SSHRC) on research data management capacity building. The event was held on June 14, 2021, as a Digital Humanities Summer Institute 2021 aligned workshop, and was led by the University of Victoria Libraries and the Electronic Textual Cultures Lab. Following the event, we summarized the proceedings in a report, culminating in our own independent recommendations for researchers, institutions, and funding agencies for the facilitation and support of *humanist* RDM (Higgins, Goddard, and Khair 2023). These recommendations include calls for clearer guidance and the development of pedagogical tools that are relevant to the humanities, as well as additional support in the form of infrastructure and funding needed to advance data work across sub-disciplines.

What is research data management (RDM)? What is a data management plan (DMP)?

Research data management (RDM) is “the processes applied through the lifecycle of a research project to guide the collection, documentation, storage, sharing and preservation of research data” (Government of Canada 2022). In addition to Canada, countries including the USA, UK, Australia, and Europe have implemented mandates that require researchers to store data securely and to document critical related information including ownership, rights, and provenance (NSF 2023; DCC 2023; ARC 2023; OpenAIRE 2023). RDM matters because research data that are not managed properly risk being unusable: whether because they are incomprehensible, corrupted in a technical sense, or cease to exist (Antoniuk and Brown 2020; Boyd 2020; Chun 2008; The Endings Project 2023), imperiling potential access to this valuable resource for other researchers, students, and the general public.

The primary RDM tool is the data management plan (DMP). In Canada, a DMP is defined as “a living document, typically associated with an individual research project or program that consists of the practices, processes and strategies that pertain to a set of specified topics related to data management and curation” (Government of Canada 2021a; see also Portage Network 2020; Lacroix and Rao 2020). DMPs provide researchers, administrators, and funding agencies with information about research projects, including their data formats, metadata schemas, storage choices, security provisions, access and usage rights, and plans related to sharing and publication. A DMP is a document that is unique to its project and is now becoming a standardized part of a grant application. We will discuss some of the risks and rewards of DMP design in the third section. However, as it currently stands, DMPs as standardized documents are at risk of being both too constrained (reducing the meaning and use of a gamut of diverse research materials and procedures) and too open-ended (unable to define and delimit humanities data in ways that prove useful for preservation and reuse).

While we suggest that all researchers, including those in the humanities, are already doing data work, it is nevertheless true that humanists bring a variety of experience with best practices for data storage, description, curation, etc., to the table (Anger et al. 2022; Evalyn et al. 2020; Rockwell et al. 2020). When our event, RDM4H, surveyed attendees, many humanists still expressed a lack of confidence dealing with “data” in their practice. Of 73 respondents (n = 73), only 3% called themselves “experienced” with RDM, 55% called themselves “capable,” and 40% called themselves “inexperienced.” The majority of respondents (66%) expressed a desire to gain further “general knowledge on data management.”

There is therefore, at best, a wide variety of researcher experience with—and confidence in their ability to perform—RDM as a standardized portion of research projects.

What are data?

The *Oxford English Dictionary* defines data in technical terms. Data are “related items of (chiefly numerical) information considered collectively, typically obtained by scientific work and used for reference, analysis, and calculation,” and, in the context of computing, “quantities, characters, or symbols on which operations are performed by a computer, considered collectively.”

It is not our purpose to delve deeper into the technical aspects of data here, but two points stand out. First, data are “considered collectively,” that is, they derive meaning in plural sets. The word “set” derives from the Latin “secta,” which we speak today as “sect,” implying a number, group, or collection of “things” defined by their relation to each other or their use in a particular operation; datasets are not merely a haphazard agglomeration of “stuff,” but created, organized, and delimited data that are thereafter put to the work of analysis and interpretation. How a dataset is organized influences what we can learn from data, and datasets therefore require purposeful *design*. Data and dataset design allow researchers to use data for “reference, analysis, and calculation.” In other words, they allow data, second, to be put to *work*. This work is diverse: historical (and even ancient) practices from weaving to notching to knotting to quilting (insofar as they involve patterning, calculation, and even accounting) are data work, involving reference, analysis, and accounting, and are not reducible in meaning to the numerical (Galloway 2021; see also Holmes 2017). That is, data work is not reducible only to a spreadsheet of numbers or what a researcher does with those numbers thereafter; it is, instead, by definition, a humanities practice, too.

That said, humanists don’t agree on what data are. Some of the first early modern uses of the word “data” in the 17th century were in clerical texts that adopted the Latinate etymology of data as “something given” into things God-given: transcendental gifts to be used for moral reasoning or inferences (Galloway 2011). There is much scholarship in the humanities arguing against the assumption that data (and even statistics) is either given or “raw” (see for example Chun 2016; Chun 2021; Gitelman 2013; Halpern 2014; Joque 2022; Thylstrup et al. 2021). DH scholars like Johanna Drucker (Drucker 2011) have argued instead that data are better thought of as “something taken” in the act of observing, collecting, and analyzing phenomena. Willard McCarty (McCarty 2005) suggests that modelling is a useful exercise for humanists, a means of experimenting, exploring, and testing assumptions. He reminds us that data models,

like data themselves, are neither neutral nor objective, but can best be understood as arguments (McCarty 2005). Beyond theoretical work related to data, many DH projects have produced practical tools for guiding the description and modelling of humanities concepts and research objects. The Text Encoding Initiative, founded and led by humanities researchers, is a widely used set of guidelines for encoding the structure and content of texts (TEI Consortium, 2023). DH projects have published structured ontologies and standardized vocabularies for humanist data modelling, including the CWRC Ontology (CWRC 2023), LINCIS Vocabularies (LINCIS 2023), and the Internet Philosophy Ontology (InPhO 2023). RDM can therefore be seen as an extension of core DH tradition and practice. High quality RDM is a means of improving the maturity of ongoing data work in DH. For computationally intensive DH practices like Computational Literary Studies and Cultural Analytics, RDM offers a path to reproducibility of research, and the reuse of data in new contexts.

We want to suggest that data are best thought of as *produced*: by the situatedness of observation, through the practice of collection, and as a result of the process of analysis. This process, which Geoffrey Bowker and Susan Leigh Star (Bowker and Star 2000) named the act of “sorting things out,” is exemplified in RDM practices like a DMP.

Calling data “produced” best fits a definition of *research data* as “data that are used as primary sources to support technical or scientific enquiry, research, scholarship, or artistic activity, and that are used as evidence in the research process and/or are commonly accepted in the research community as necessary to validate research findings and results” (CODATA 2023). Research data may thus include “experimental data, observational data, operational data, third party data, public sector data, monitoring data, processed data, or repurposed data” (CODATA 2023). Assuming data are produced allows data to be considered both collectively and as objects researchers put to work. Moreover, this assumption stresses that data are not self-same (i.e., inherently interchangeable, transparent, or interoperable), whether in kind or in time (Posner 2021). Before our event we surveyed our attendees (n = 73) on what research materials they work with that they consider to be data. Here are some of their answers:

text; audio files; video files; image files; physical archival records; medieval texts; library catalogue data; policy papers and records; VHS; 8 mm film; camcorder tapes; community relations (impact, contributions, feedback); religion in the 20th century; social movements in the 20th century; organizational memberships, budgets, bureaucratic documents; paper correspondence; interviews; maps; cultural analytics (e.g., Wikipedia links); programming languages; software; geodata; book history; literature (across genres, styles, and historical periods); culture and cultural objects

(e.g., toys, games, clothing, foods); translations; radio tapes; prosopography; historical artefacts (and images of artefacts); state and private financial records (e.g., grant records); text corpora; ephemera like clippings, photos, journals, logs; web usage and analytics; CSV files.

The gamut of these materials suggests, first, that humanist work is also data work. Humanists are engaged in the collective consideration of a plurality of objects that they put to work in tasks like translation, iteration, reconstruction, and interpretation. Humanities data can be defined as data that are produced within the context of disciplines that study “human culture and the cultural record” (Borgman 2015, 161). More broadly, humanities data might be considered as a plurality of objects that researchers put to work in order to know how humans articulate meaning. Second, however, these answers suggest there is no clear (or agreeable) typology of “humanities data.” Humanities data lack a clear “what.” Instead, as our event’s keynote speaker, Miriam Posner (Posner 2021), suggested, where consensus might be found is in the “when” of data: when do we consider our research materials as data versus as something else?

The question of “when” data are is ultimately one of when, exactly, we call data “produced.” When are research materials produced as categories? When are they demarcated and parameterized? How do these categories de-/stabilize their status as research materials? And what does their datafication mean for their connection to the original materials, and their future replicability? Each of these questions may be answered (but also complicated) by considering the process of observation, collection, and analysis. Producing a menagerie of research materials as something else—a dataset, an archive, a critical reflection—cannot be assumed to be standardized for humanists, but it is the feature that renders humanist work as always also data work.

We want to suggest that these considerations—of what and when data are; of their ontological plurality; of the ways in which they are put to work—are a primary justification for attention to RDM in the humanities. RDM has both a practical and a conceptual component, but researchers cannot be expected to attend to RDM ex nihilo. Researchers require funding, training, and support to adequately meet the challenge of framing their work as data work, understanding best practices for data management, including the FAIR Principles (Wilkinson et al. 2016), and determining how to translate their research into formal data management plans. As the guidance for this work is developed (e.g., Harrower et al. 2020) clear examples of the value of high-quality, sustainable, reusable humanities datasets are necessary to convince humanist researchers of the importance of RDM work. These examples should be accessible and should not have to be sought out (see, for example, Tayler et al. 2022; The Endings Project 2023; SpokenWeb 2023).

In addition to practical justifications for accessible RDM examples (e.g., better data work), it is imperative to advocate for exemplars and policies that establish *ethical* examples and guidelines for data work in a diversity of contexts. Many humanists do research that involves, to some degree, human subjects, whether contemporary or historical. People are not objects of study: they are *participants*, what have also been typically referred to by research communities as “stakeholders” in research work (Carroll, Rodriguez-Lonebear, and Martinez 2019). One key consideration of our event was the acknowledgement that ethical concerns must trump data-sharing benefits in all cases (Czaykowska-Higgins 2021).

However, this language of “stakeholders” and “participants” is not sufficient to describe or advocate for Indigenous communities worldwide involved in data work. Though work is ongoing, the issues of Indigenous data sovereignty and governance remain major areas of concern for RDM policies and practices worldwide; there are longstanding questions about how to catalogue and classify Indigenous peoples’ materials (Duarte and Belarde-Lewis 2015; Montenegro 2019). Outside researcher involvement, regardless of discipline or intent, risks being an agent of social and cultural dislocation (Holton, Leonard, and Pulsifer 2022), replicating colonial archives and practices (Thorpe et al. 2021), and ultimately submitting Indigenous peoples’ data to colonial structures of knowledge and information. Indigenous peoples “have always been data creators, data users, and data stewards. Data were and are embedded in Indigenous instructional practices and cultural principles” (NCAI 2018, 1). Data produced in and around Indigenous communities (whether by those communities themselves, or by researchers in partnership with them) do not make those communities stakeholders but rather *rights-holders* (Sarkki, Heikkinen, and Löf 2021). The established concept of the stakeholder tends to prioritize economic interests, risks conflating a diversity of cultural, historical, and situated contexts into the homogenous category of a stakeholder, and also assumes an embeddedness in national and international legal structures of governance and sovereignty that de-contextualize or, worse, imperil Indigenous sovereignties (Sarkki, Heikkinen, and Löf 2021). There is an urgent need for further research into how best to promote and support Indigenous-led data governance (Espinosa de los Monteros 2019; Love et al. 2022).

There is ongoing international collaborative work to develop paradigms for Indigenous data sovereignty—including via paradigms like Ownership, Control, Access, Possession (OCAP) (FNIGC 2023) and Collective Benefit, Authority to Control, Responsibility, Ethics (CARE) (GIDA 2023)—but, as the Canadian Tri-Agencies have themselves noted, these paradigms don’t necessarily “respond to the needs and values of distinct First Nations, Métis, and Inuit communities, collectives and organizations” (Government of Canada 2021b). Likewise, while there are attempts to establish paradigms for project development, there exists a gap in the literature around

project sustainability and the long-term use of digital resources in contexts related to Indigenous data (Strathman 2019). In this context, projects that disappear or die should be regarded as an active threat to Indigenous data sovereignty, governance, and rights-holding.

The question of “what data are” is therefore one that stretches beyond the technical, theoretical, or methodological, and into the ethical. Indigenous data considerations are unique but never irrelevant to broader considerations of ethical data work. Shared examples of ethical data work are not “edge cases” or specific to “marginalized communities” but integral to the development of data work writ large. Seeing examples of good data work and subsequently approaching that work as a process of active researcher production can only strengthen the quality of humanist RDM and research. How, then, might humanists and funders go about creating an environment that is conducive to good data work?

Designing data

To say data are produced is to argue that that they are made or crafted as part of a practice: an argument that may be equally applied to the concept of design. They involve experimentation with and the use of (often complex) tools (Tenen 2016), which beg questions of software and hardware optimization, dependencies, and longevity, not to mention design, programming, and systems operations (Sayers 2016). They involve speculative choices (Hong 2020): about which data to collect, when to collect them, how, and why; and how to classify, represent, and thereafter analyze that data. The craft or design of datasets is labour (an argument we will highlight in the fourth section), and the labour of design—as any craftsperson knows—alters the object produced.

In pragmatic, funder terms, design is both implicit and explicit in a DMP: from researchers picking tools and staking out future terms of classification and analysis, to supplementing those technical and methodological guidelines with conceptual elements like the “what” and “when” of their project’s RDM. Our purpose here is less to discuss the technical ins and outs of project and DMP design, as to establish the argument that *data work requires intentional (rather than ancillary) design*. The gist of the argument is that RDM cannot be “tacked on” to entrenched research practices but must be *systematized*. Part of developing a system of supports for humanist data work involves making space for a variety of meanings and uses of data: in other words, balancing the need for a degree of standardization across the humanities with the flexibility that allows humanists to work with a diverse array of research data.

Much of humanities work in RDM is centred around the management and representation of information—not necessarily numerical datasets to be analyzed, but collections of objects with (often complex) metadata that may be interacted with by

researchers and other users in ways that go beyond a simple reproduction of calculations. In Canada, we might cite the SpokenWeb’s audio recording objects (SpokenWeb 2023); the ongoing dictionary projects for SENĆOŦEN (Elliott et al. 2021), the language of the WSÁNEĆ First Nations on Vancouver Island’s Saanich Peninsula, or for the East Cree language (Junker et al. 2018), spoken in Northern Québec’s James Bay area, which support language revitalization through data work that combines linguistic information with narrative and other resources; or the Map of Early Modern London, whose six databases support a host of interoperable projects ranging from an encyclopaedia, a digital version of a map, an anthology, to a bibliography, and more (Jenstad 2006).

Each of these projects blurs the boundaries between data, information, and narrative. Data and information are intricately related but not self-same. Often, they are colloquially used in a tautology to define each other: data are pieces of information, while information is a collection of data. In reality, the word “information” has Latin roots that differentiate it from “data”: as Galloway (Galloway 2011) notes, the root of information stems from “the act of taking form or being put into form.” In the assumption that data are “raw,” information is, then, “cooked”: it has been shaped, sculpted, or interpreted. Neither is a text, a map, or an audio tape “raw”: they are all best considered as neither given, nor taken, but produced. In the process of production, data render information.

Data also produce narratives. Data tell stories (Chun 2016) because data are used to do things (Thylstrup et al. 2021). For example, the internet database ImageNet—which collected 14 million images, scraped from the internet, and standardized them into categories based on what they were images of, all for the purpose of developing training sets of objects for machine vision—has been widely criticized for producing a classification system that is replete with bias, discriminatory, and therefore actively harmful (Crawford and Paglen 2019). Elsewhere, natural language processors, trained on datasets of social media interactions, have tended to parrot and eventually adopt the same bigoted language that makes up part of the “organic” data they are trained on (Vincent 2016). The point is that data produce narratives or stories, and that all stories have politics (Lampland and Star 2009; see also Apprich et al. 2018). As a result, *data, put to use, have politics too*. This statement does not mean that data are intrinsically a form of domination, or can only be used in that way, but rather that choices about what and how to observe (i.e., RDM) determine data and the insights (i.e., truths) gleaned therein. Data and datasets may go so far as to structure empirical realities, whether in a “positive” sense by producing “facts” and “truth” to add to a system of knowledge, or in a “negative” sense by being denied or misconstrued as “disinformation” or “post-truth” (Harsin 2018). Researchers and funding bodies must be mindful of those

possibilities, and of the concerns that follow, not in order to foreclose data work, but rather to encourage project and data *design* that anticipates how the choices researchers make to extract, organize, and classify data have consequences for knowledge and politics.

These recognitions matter because many humanists may be uncertain about what constitutes “data” in the context of their research projects. Better guidance on defining research data must be developed in consultation with both digital and non-digital humanists from a variety of different disciplines. Humanists should not be guessing whether their research objects, in the act of producing DMPs, change state—from object, to data, metadata, information, or narrative—nor should they be expected to design research projects—including technical tools, practices, or ethical guidelines—in isolation. Funders and institutions must provide a system of support and guidance that not only facilitates grant applications, but can also step in if and when problems arise. Defining research data should be a collaborative activity; otherwise, federal RDM in the humanities risks producing little more than a mess of incongruent and useless data and DMPs.

Training humanities researchers

One solution for developing useful definitions of research data and promoting good research data design *before* projects begin is through the simple axiom of “teaching for all.” Our event survey found that many senior humanities researchers and instructors do not feel that they have enough RDM knowledge to confidently teach the necessary concepts and tools. Likewise, efforts to develop data primers (Tayler et al. 2022) for researchers and librarians evince an ongoing knowledge gap. A great deal of current RDM instruction is aimed at experienced researchers, but it is also necessary to develop instructional resources for undergraduate and graduate audiences.

Teaching humanists how to do RDM—allowing them to ask what kind of data their research objects might be, consider how they will design their projects, what kinds of information and narratives their work will produce, and how it can all be ethical—requires flexibility. If humanist data work is diverse, it is imperative to avoid applying over-standardized solutions (e.g., pedagogy, tools) for diverse research across different disciplines and fields. Although some measure of standardization is necessary for any RDM work at scale, an overemphasis on standardization risks conflating and confusing different types of research and outputs. Researchers applying for funding may shoehorn their research materials to fit overly restrictive DMP templates or simply improperly report what it is they are doing. As RDM policies become more mature, it is imperative to spend time examining edge cases, including, for example, analogue scholarship, and fine arts research processes (Bath 2021). Directed effort must be made to engage

researchers who identify their research as not fitting into current data management policy rather than focussing on successful, masthead digital humanities projects, which generally already have institutional support, funding, and technical capabilities. This recommendation can be summarized as looking at the boundaries and edge cases of policies in addition to the centre.

Should international, federal, and local institutions fail to support researchers in ways that push beyond technical questions and into conceptual ones (like defining data and considering how humanists may produce them differently from, say, scientists), we risk a research landscape that struggles with project sustainability because it continually divorces the technical (e.g., back-end data storage or software use) from the conceptual (e.g., questions about interpretation, information, and narrative).

How, then, can we address this conflict, where it might be more conceptually straightforward to silo project areas, but such divorce will likely also technically inhibit project preservation? One useful step involves more and better training around metadata design. More contextual and interpretive information should be included in project metadata, but also, projects should be designed (with guidance) from inception so that data can stand alone, outside the context of the user interface, without losing comprehensibility. Not only will this produce more reusable data, but it will help a great deal with the problem of project preservation. We recommend that humanist researchers be given much more theoretical and practical training on metadata creation and data literacy, ideally beginning at the undergraduate level. This training could build on work that is already ongoing in the digital humanities to include topics such as data modelling in TEI/XML; introduction to structured vocabularies like Dublin Core, Getty Art & Architecture Thesaurus, or the Homosaurus; exploration and critique of ontologies like CIDOC-CRM or PhilOnto; relational database design and normalization; or topic modelling in the context of text analysis.

The training and tools provided, however, cannot be prescriptive to a degree that they limit or inhibit the breadth of humanities data we have outlined in this article. As we have stressed, humanist data are extremely diverse. Most data are not highly structured or machine-generated, and a significant amount of what might be considered data are not digital. To return to a point made above, for funding bodies, institutions, and humanities researchers, one central task of RDM will be developing infrastructures that achieve a measure of standardization that supports widespread access, while ensuring that researchers do not lose the ability to critically engage with different theories, methods, and practices of categorization in their own work. Research software, publishing platforms, and data repositories need to be flexible enough to support humanist research objects and processes without unduly constraining them.

Data work will not always be perfect from the beginning, and so data as practice involves a willingness to *experiment*, or to be prepared for the changes that projects undergo, and the contingencies they may encounter. It is extremely unlikely that humanist researchers will be able to create accurate and detailed plans at the application stage. It would be unrealistic to expect that they do, and unfair to demand they standardize their initial RDM to a degree that imperils the design and outcomes of their research. Instead, data management planning tools must support the evolving nature of data management plans and, ultimately, humanist RDM.

Labour, support, and recognition

Data work is work: work that cannot be done without funding, support, and tools, and that should be attributed, cited, and recognized. It can be easy to assume that apparently ubiquitous computing capacity and power means all research materials may be seamlessly “made digital,” that computers do that work, and that questions around RDM production, dissemination, and attribution are predominantly questions about tools and tools alone. This assumption can erase critical distinctions about access to computing (Gray, Gerlitz, and Bounegru 2018; Estill 2020), computing power and capabilities, and the myriad hard- and software differences that influence research methods and results. Computational differences are project differences (The Endings Project 2023). These sorts of assumptions can also erase the human labour involved in data work: from researchers developing data management plans, to data librarians setting institutional data policies, to data or technical support officers assisting or leading the technical execution of RDM. Data management policies that do not properly integrate the recognition and attribution of researcher and other participant labour will implicitly punish researchers by requiring extra work that goes unacknowledged at institutional and public levels.

The success of RDM is contingent on the active attention to and maintenance of systems for data collection, sharing, and attribution, and will be improved by practices of recognition and reward that incentivize all researchers (including early career ones) to adopt standards for data work (Alperin et al. 2022; Anger et al. 2022; Moher and Cobey 2021; Schöpfel and Azeroual 2021). This section provides further recommendations to the RDM landscape organized around the theme of labour and the necessity of attribution and recognition of that labour, a topic already under consideration in the sciences (see e.g., Park, You, and Wolfram 2018; Wood-Charson et al. 2022; Zeng et al. 2020) but currently neglected in the humanities.

In the first instance, the support provided to researchers cannot be “one and done,” but must be maintained throughout the lifecycle of a project. Humanist researchers continue to feel that they need support for data management at *all* stages of the research

process: on conceptual and theoretical approaches to data; on guidance for meeting novel funding agency requirements; on making choices about data infrastructure (i.e., platform) use; on defining appropriate metadata frameworks; on capturing and recording metadata according to standards; and on ensuring their research does not change in kind in order to meet data policy. In order to facilitate this support, research computing providers (in Canada, for example, this would include the Digital Research Alliance of Canada) and funders should engage in “meta-science”—training researchers on data work and sharing, establishing benchmarks and auditing projects, evaluating progress, and rewarding researchers for that work in the form of its inclusion as a category in hiring, promotion, etc. (Moher and Cobey 2021). It is not sufficient to offer funding for data work without incentivizing the sharing and proper attribution of data in the research community; early career researchers in particular may be disadvantaged if their work is not properly attributed and rewarded by their institutions, which disincentivizes good data work by inhibiting structures of reward (Schöpfel and Azeroual 2021).

The first step to incentivize data sharing, proper attribution, and credit is funding increases that reflect the additional cost (in time, people, and tools) of research data management. Large research grants may hire project managers, a practice that can and should be extended to, for example, funding increases to hire and train team members who can oversee the design and creation of data and metadata to ensure that practice aligns with the data management plan. While this funding can and should be project or grant specific, as we have suggested, it should also be *systematized* in the form of more support staff at regional and national digital research infrastructure providers and within institutions, who can help researchers before, during, and after research projects, as well as monitor researcher progress and benchmarks. Across different academic disciplines (including the sciences) in different nations, funders currently struggle to monitor and support researcher data work (Anger et al. 2022). An additional benefit of this kind of support from research computing providers and funding agencies is that it would be accessible to all researchers, regardless of the stage of their career or institutional affiliation. As we have discussed, labour increases may be easier to contend with for established, recognized researchers, and systems of incentivization and support at the funder level have the potential to level the playing field for a diversity of researchers.

Humanists require RDM support and training sessions over the course of their whole careers, and not simply when they are ready to apply for funding. Ideally, data management concepts and basic skills will be developed at the undergraduate and graduate level: what Moher and Cobey have called a “core syllabus” (Moher and Cobey 2021, 1535). Asking researchers to try to absorb and apply all of this information at the point of grant application is likely to generate frustration and shallow engagement, as material becomes outdated or forgotten over the award timespans. While the initial

federal and institutional investment in this support and training might be at a higher dollar amount in the short term, it will save money in the long term: first, because it will promote good research practices that avoid previously discussed issues like data corruption, loss, or usability, and second, because researchers who know how to do RDM can teach and help other researchers.

The key, as we have endeavoured to suggest, is people. RDM standards and regulations that limit themselves to technical considerations will miss the forest for the trees. There is no question that platform, software, and tool choices will significantly affect the way in which project data is organized, described, and accessed, and as we have discussed at length, funders and institutions must be prepared to support tool use that is compatible with a diverse range of humanist data. At the same time, they cannot forget that data are *produced*, by *people*, and through *labour*. While there is a critical need for improved access to research tools and infrastructure, technology alone cannot fully address researcher needs. Funded human experts who can provide support and guidance are equally important.

Conclusion

We have attempted to provide an assessment of RDM that acknowledges some of the challenges while highlighting possible areas for researcher success. While it is critical that humanists be provided with the support (in terms of both funding and people) to allow them to do good RDM and produce data that are usable—without imperiling other more traditionally humanist project areas—we have argued that humanist training makes humanists particularly prepared to ask difficult questions like “what are my data?” and “how should this project be designed?”

Nevertheless, it is clear that funders and other affiliated organizations involved in the RDM process need to continue to direct funding and enquiry towards developing sustainable RDM processes, including assessing how they can support humanists in the production of high quality data; engaging with different communities of practice with an eye towards both methodology *and ethics*; adopting RDM and DMP protocols that are open to a diverse array of humanities data; and establishing practices that provide humanists with the time, tools, and people to *design* projects carefully, *experiment* with different practices of production, and fund the requisite *labour* for added RDM components. In sum, this article does not regard RDM as a settled matter, but rather as one that will require ongoing consultation, development, and revision. We provide these recommendations with the hope they will supply humanist researchers (and the institutions and agencies that support them) with an outline of the considerations involved in approaching data work as method, concept, and practice.

Competing interests

The authors have no competing interests to declare.

Contributions***Authorial***

Authorship in the byline is by magnitude of contribution. Author contributions, described using the NISO (National Information Standards Organization) Credit taxonomy, are as follows:

Author name and initials:

Stefan Higgins (SH)

Lisa Goddard (LG)

Shahira Khair (SK)

Authors are listed in descending order by significance of contribution. The corresponding author is LG.

Conceptualization: SH, LG, SK

Formal Analysis: SH

Funding Acquisition: LG, SK

Writing – Original Draft: SH, LG, SK

Writing – Review & Editing: LG, SK

Editorial

Section Editor

Frank Onuh, The Journal Incubator, University of Lethbridge, Canada

Copy and Production Editor

Christa Avram, The Journal Incubator, University of Lethbridge, Canada

Layout Editor

A K M Iftekhhar Khalid, The Journal Incubator, University of Lethbridge, Canada

Translation Editor

Davide Pafumi, The Journal Incubator, University of Lethbridge, Canada

References

- Alperin, Juan Pablo, Lesley A. Schimanski, Michelle La, Meredith T. Niles, and Erin C. McKiernan. 2022. "The Value of Data and Other Non-Traditional Scholarly Outputs in Academic Review, Promotion, and Tenure in Canada and the United States." In *The Open Handbook of Linguistic Data Management*, edited by Andrea L. Berez-Kroeker, Bradley McDonnell, Eve Koller, Lauren B. Collister, 171–182. Cambridge, MA: The MIT Press. Accessed November 29, 2023. <https://doi.org/10.7551/mitpress/12200.003.0017>.
- Anger, Michael, Christian Wendelborn, Eva C. Winkler, and Christoph Schickhardt. 2022. "Neither Carrots nor Sticks? Challenges Surrounding Data Sharing from the Perspective of Research Funding Agencies—A Qualitative Expert Interview Study." *PLOS One* 17(9): e0273259. Accessed November 29, 2023. <https://doi.org/10.1371/journal.pone.0273259>.
- Antoniuk, Jeffrey, and Susan Brown. 2020. "Interface Matters." Digital Research Alliance of Canada. Position Paper Submission #54. Accessed November 29, 2023. https://alliancecan.ca/sites/default/files/2022-03/cwrc_interface_matters_ndrio_white_paper.pdf.
- Apprich, Clemens, Wendy Hui Kyong Chun, Florian Cramer, and Hito Steyerl. 2018. *Pattern Discrimination*. In *Search of Media*. Minneapolis, MN: University of Minnesota Press.
- ARC (Australian Research Council). 2023. "Research Data Management." Accessed November 29. <https://www.arc.gov.au/about-arc/strategies/research-data-management>.
- Bath, Jon. 2021. "Is This Data? Research Process as Data in the Arts." Public talk for *Research Data Management for Digitally Curious Humanists* [online], June 14. Accessed November 29, 2023. <https://osf.io/g947n/wiki/virtual%20panel%20--%20bath/>.
- Borgman, Christine L. 2015. *Big Data, Little Data, No Data: Scholarship in the Networked World*. Cambridge, MA: The MIT Press. Accessed November 29, 2023. <https://doi.org/10.7551/mitpress/9963.001.0001>.
- Bowker, Geoffrey, and Susan Leigh Star. 2000. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: The MIT Press. Accessed November 29, 2023. <https://doi.org/10.7551/mitpress/6352.001.0001>.
- Boyd, Jason. 2020. "DRI, University Libraries and Digital Humanities Research Centres." Digital Research Alliance of Canada. Position Paper Submission #91. Accessed November 29, 2023. https://alliancecan.ca/sites/default/files/2022-03/ryerson_ndriowhitepaper.pdf.
- Carroll, Stephanie Russo, Desi Rodriguez-Lonebear, and Andrew Martinez. 2019. "Indigenous Data Governance: Strategies from United States Native Nations." *Data Science Journal* 18(31): 1–15. Accessed November 29, 2023. <https://doi.org/10.5334/dsj-2019-031>.
- Chun, Wendy Hui Kyong. 2008. "The Enduring Ephemeral, or the Future Is a Memory." *Critical Inquiry* 35(1): 148–171. Accessed November 29, 2023. <https://doi.org/10.1086/595632>.
- . 2016. "Big Data as Drama." *ELH* 83(2): 363–382. Accessed November 29, 2023. <https://doi.org/10.1353/elh.2016.0011>.
- . 2021. *Discriminating Data: Correlation, Neighbourhoods, and the New Politics of Recognition*. Cambridge, MA: The MIT Press. Accessed November 29, 2023. <https://doi.org/10.7551/mitpress/14050.001.0001>.

- CODATA (Committee on Data). 2023. "Research Data." Committee on Data of the International Science Council. Accessed November 29, 2023. <https://codata.org/rdm-terminology/research-data/>.
- Crawford, Kate, and Trevor Paglen. 2019. "Excavating AI: The Politics of Images in Machine Learning Training Sets." The AI Now Institute, NYU, September 19. Accessed November 29, 2023. <https://excavating.ai>.
- CWRC (Canadian Writing Research Collaboratory). 2023. "The CWRC Ontology Specification 0.99.88." CWRC Linked Open Data. Accessed November 29. <https://sparql.cwrc.ca/ontologies/cwrc.html>.
- Czaykowska-Higgins, Ewa. 2021. "'Data' in Indigenous Language Documentation." Public talk for *Research Data Management for Digitally Curious Humanists* [online], June 14. Accessed November 29, 2023. <https://osf.io/g947n/wiki/virtual%20panel%20--%20czaykowska/>.
- DCC (Digital Curation Centre). 2023. "Overview of Funders' Data Policies." Accessed November 29. <https://www.dcc.ac.uk/guidance/policy/overview-funders-data-policies>.
- Drucker, Johanna. 2011. "Humanities Approaches to Graphical Display." *Digital Humanities Quarterly* 5(1). Accessed November 29, 2023. <http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html>.
- Duarte, Marisa Elena, and Miranda Belarde-Lewis. 2015. "Imagining: Creating Spaces for Indigenous Ontologies." *Cataloging & Classification Quarterly* 53(5-6): 677-702. Accessed November 29, 2023. <https://doi.org/10.1080/01639374.2015.1018396>.
- Elliott, John (J,SINTEN), Linda Elliott (ȚOSINIYE), Lou Claxton (SELÁMTEN), Belinda Claxton (SELIYE), Ewa Czaykowska-Higgins, Anter Elliott, Oren Elliott (STOLȚEŁ), Megan Supernault (I,ÍYMEŁWET), Tye Swallow, and David Underwood (PENÁĆ). 2021. "AŁȚEŁ SĆÁ: Intersecting Relationships in Sustainable Language Reclamation - The AŁȚEŁ SĆÁ Team." ICLDC YT, May 3. YouTube video, 30:15. Accessed November 29, 2023. <https://www.youtube.com/watch?v=1fCT2HMfHWg>.
- Espinosa de los Monteros, Pamela. 2019. "Decolonial Information Practices: Repatriating and Stewarding the Popol Vuh Online." *Preservation, Digital Technology & Culture* 48(3-4): 107-119. Accessed November 29, 2023. <https://doi.org/10.1515/pdct-2019-0009>.
- Estill, Laura. 2020. "All Researchers Use Digital Resources: On Campus Support, Grants, Labs, and Equity." Digital Research Alliance of Canada. Position Paper Submission #22. Accessed November 29, 2023. https://alliancecan.ca/sites/default/files/2022-03/ndrio-white-paper-w-centernet-submitted_fixed-1.pdf.
- Evalyn, Lawrence, Elizabeth Parke, Patrick Keilty, and Elspeth Brown. 2020. "Gaps in Digital Research Infrastructure for Canadian Digital Humanities Researchers." Digital Research Alliance of Canada. Position Paper Submission #18. Accessed November 29, 2023. https://alliancecan.ca/sites/default/files/2022-03/ndrio_dh_white_paper_digital_humanities_university_of_toronto.pdf.
- FNIGC (First Nations Information Governance Centre). 2023. "The First Nations Principles of OCAP." Accessed November 29. <https://fnigc.ca/ocap-training/>.
- Galloway, Alexander. 2011. "Are Some Things Unrepresentable?" *Theory, Culture & Society* 28(7-8): 85-102. Accessed November 29, 2023. <https://doi.org/10.1177/0263276411423038>.

- . 2021. *Uncomputable: Play and Politics in the Long Digital Age*. New York: Verso.
- GIDA (Global Indigenous Data Alliance). 2023. "CARE Principles for Indigenous Data Governance." Accessed November 29. <https://www.gida-global.org/care>.
- Gitelman, Lisa, ed. 2013. *Raw Data Is an Oxymoron*. Infrastructures Series. Cambridge, MA: The MIT Press.
- Government of Canada. 2021a. "Tri-Agency Research Data Management Policy – Frequently Asked Questions." Last updated October 28. Accessed November 29, 2023. https://www.ic.gc.ca/eic/site/063.nsf/eng/h_97609.html.
- . 2021b. "Tri-Agency Research Data Management Policy." Last updated March 14. Accessed November 29, 2023. https://www.ic.gc.ca/eic/site/063.nsf/eng/h_97610.html.
- . 2022. "Research Data Management." Last updated May 31. Accessed November 29, 2023. https://science.gc.ca/eic/site/063.nsf/eng/h_547652FB.html.
- Gray, Jonathan, Carolin Gerlitz, and Liliana Bounegru. 2018. "Data Infrastructure Literacy." *Big Data & Society* 5(2). Accessed November 29, 2023. <https://doi.org/10.1177/2053951718786316>.
- Halpern, Orit. 2014. *Beautiful Data: A History of Vision and Reason since 1945*. Durham, NC: Duke University Press.
- Harrower, Natalie, Maciej Maryl, Timea Biro, Beat Immenhauser, and ALLEA Working Group E-Humanities. 2020. "Sustainable and FAIR Data Sharing in the Humanities: Recommendations of the ALLEA Working Group E-Humanities." *Digital Repository of Ireland*. Accessed November 29, 2023. <https://doi.org/10.7486/DRI.tq582c863>
- Harsin, Jayson. 2018. "Post-Truth and Critical Communication Studies." *Oxford Encyclopaedia of Communication*. Accessed November 29, 2023. <https://doi.org/10.1093/acrefore/9780190228613.013.757>.
- Higgins, Stefan, Lisa Goddard, and Shahira Khair. 2023. "Research Data Management Support in the Humanities: Challenges and Recommendations." *UvicSpace*. Accessed November 29. <http://hdl.handle.net/1828/14641>.
- Holmes, Dawn E. 2017. *Big Data: A Very Short Introduction*. Very Short Introductions. Oxford, UK: Oxford University Press.
- Holton, Gary, Wesley Y. Leonard, and Peter L. Pulsifer. 2022. "Indigenous Peoples, Ethics, and Linguistic Data." In *The Open Handbook of Linguistic Data Management*, edited by Andrea L. Berez-Kroeker, Bradley McDonnell, Eve Koller, and Lauren B. Collister, 49–60. Cambridge, MA: The MIT Press. Accessed November 29, 2023. <https://doi.org/10.7551/mitpress/12200.001.0001>.
- Hong, Sun-ha. 2020. *Technologies of Speculation: The Limits of Knowledge in a Data-Driven Society*. New York: NYU Press.
- InPhO Project. 2023. "Welcome to the Internet Philosophy Ontology (InPhO) Project." Accessed November 29. <https://www.inphoproject.org/>.
- Jenstad, Janelle. 2006. The Map of Early Modern London. Victoria, BC: University of Victoria, 2006-present. Accessed March 14, 2024. <http://mapoflondon.uvic.ca>.
- Joque, Justin. 2022. *Revolutionary Mathematics: Artificial Intelligence, Statistics, and the Logic of Capitalism*. New York: Verso.

- Junker, Marie-Odile, Marguerite MacKenzie, Luci Bobbish-Salt, Alice Duff, Linda Visitor, Ruth Salt, Anna Blacksmith, Patricia Diamond, and Pearl Weistche, eds. 2018. *The Eastern James Bay Cree Dictionary on the Web*. Accessed November 29, 2023. <https://dictionary.eastcree.org/>.
- Lacroix, Denis, and Sathya Rao. 2020. "Data Management Plan for Belgians and French in the Prairies (Exemplar)." Zenodo. Accessed November 29, 2023. <https://doi.org/10.5281/zenodo.4062484>.
- Lampland, Martha, and Susan Leigh Star, eds. 2009. *Standards and Their Stories: How Quantifying, Classifying, and Formalizing Practices Shape Everyday Life*. Ithaca, NY: Cornell University Press.
- LINCS Project (Linked Infrastructure for Networked Cultural Scholarship). 2023. "Vocabulary Browser." Accessed November 29. <https://vocab.lincsproject.ca/Skosmos/en/>.
- Love, Robin P., Billie-Jo Hardy, Courtney Heffernan, Amber Heyd, Melissa Cardinal-Grant, Lori Sparling, Bonnie Healy, Janet Smylie, and Richard Long. 2022. "Developing Data Governance Agreements with Indigenous Communities in Canada: Toward Equitable Tuberculosis Programming, Research, and Reconciliation." *Health and Human Rights Journal* 24(1): 21–33. Accessed November 29, 2023. <https://www.hhrjournal.org/2022/06/developing-data-governance-agreements-with-indigenous-communities-in-canada-toward-equitable-tuberculosis-programming-research-and-reconciliation/>.
- McCarty, Willard. 2005. *Humanities Computing*. New York: Palgrave MacMillan.
- Moher, David, and Kelly D. Cobey. 2021. "Ensuring the Success of Data Sharing in Canada." *FACETS* 6(1): 1534–1538. Accessed November 29, 2023. <https://doi.org/10.1139/facets-2021-0031>.
- Montenegro, Maria. 2019. "Subverting the Universality of Metadata Standards: The TK Labels as a Tool to Promote Indigenous Data Sovereignty." *Journal of Documentation* 75(4): 731–749. Accessed November 29, 2023. <https://doi.org/10.1108/JD-08-2018-0124>.
- NCAI (National Congress of American Indians). 2018. *Support of US Indigenous Data Sovereignty and Inclusion of Tribes in the Development of Tribal Data Governance Principles*. Accessed March 30, 2024. <https://www.ncai.org/resources/resolutions/support-of-us-indigenous-data-sovereignty-and-inclusion-of-tribes-in-the-development-of-tribal-data>.
- NSF (National Science Foundation). 2023. "Preparing Your Data Management Plan." Accessed November 29. <https://new.nsf.gov/funding/data-management-plan>.
- O'Donnell, Daniel Paul. 2020. "'Good Things Come in Small Packets': How (Inter)national Digital Research Infrastructure Can Support 'Small Data' Humanities and Cultural Heritage Research." Digital Research Alliance of Canada. Position Paper Submission #1. Accessed November 29, 2023. <https://alliancecan.ca/sites/default/files/2022-03/dri-infrastructure-white-paper-good-things-come-in-small-packets.pdf>.
- OpenAIRE. 2020. "How to Comply with Horizon Europe Mandate for Research Data Management." *Guides for Researchers*. Accessed November 29, 2023. <https://www.openaire.eu/how-to-comply-with-horizon-europe-mandate-for-rdm>.
- Park, Hyungjoo, Sukjin You, and Daniel Wolfram. 2018. "Informal Data Citation for Data Sharing and Reuse Is More Common than Formal Data Citation in Biomedical Fields." *Journal of the Association for Information Science and Technology* 69(11): 1346–1354. Accessed November 29, 2023. <https://doi.org/10.1002/asi.24049>.

- Portage Network. 2020. "Brief Guide: Data Management Plan." Zenodo. Accessed November 29, 2023. <https://doi.org/10.5281/zenodo.4495482>.
- Posner, Miriam. 2021. "What Does 'Data' Mean in the Humanities?" Public talk for *Research Data Management for Digitally Curious Humanists* [online], June 14. Accessed November 29, 2023. <https://osf.io/st28k>.
- Rockwell, Geoffrey, Matt Huculak, Emmanuel Château-Dutier, Barbara Bordalejo, Kyle Dase, Laura Estill, Julia Polych-O'Neill, and Harvey Quamen. 2020. "Canada's Future DRI Ecosystem for Humanities and Social Sciences (HSS)." Digital Research Alliance of Canada. Position Paper Submission #20. Accessed November 29, 2023. https://alliancecan.ca/sites/default/files/2022-03/csdh_ndrio_whitepaper.pdf.
- Sarkki, Simo, Hannu I. Heikkinen, and Annette Löf. 2021. "Reindeer Herders as Stakeholders or Rights-Holders? Introducing a Social Equity-Based Conceptualization Relevant for Indigenous and Local Communities." In *Nordic Perspectives on the Responsible Development of the Arctic: Pathways to Action*, edited by Douglas C. Nord, 271–292. Springer Polar Sciences. Cham: Springer. Accessed November 29, 2023. https://doi.org/10.1007/978-3-030-52324-4_13.
- Sayers, Jentery. 2016. "Minimal Definitions." *Minimal Computing* (blog), October 2. Accessed November 29, 2023. <https://go-dh.github.io/mincomp/thoughts/2016/10/02/minimal-definitions/>.
- Schöpfel, Joachim, and Otmane Azeroual. 2021. "Rewarding Research Data Management." In *WWW '21: Proceedings of the Web Conference 2021*, edited by Jure Leskovec, Marko Grobelnik, Marc Najork, Jie Tang, and Leila Zia, 446–450. New York, NY: Association for Computing Machinery Digital Library. Accessed November 29, 2023. <https://doi.org/10.1145/3442442.3451367>.
- Siemens, Ray, and Alyssa Arbuckle. 2020a. "HQP Pathways: Engaging the Canada's Different Disciplinary Models for HQP Training and Funding to Facilitate DRI Uptake in Canada." Digital Research Alliance of Canada. Position Paper Submission #55. Accessed November 29, 2023. https://alliancecan.ca/sites/default/files/2022-03/dhsi_ndrio-whitepaper_12-14-20.1.pdf.
- . 2020b. "Steps to Success in Ensuring DRI Engages and Mobilizes Humanities and Social Science Research." Digital Research Alliance of Canada. Position Paper Submission #63. Accessed November 29, 2023. https://alliancecan.ca/sites/default/files/2022-03/inke_ndrio-whitepaper_12-14-20.1.pdf.
- SpokenWeb. 2023. "Sounding Literature." Accessed November 29, 2023. <https://spokenweb.ca/>.
- Strathman, Nicole. 2019. "Digitizing the Ancestors: Issues in Indigenous Digital Heritage Projects." *International Journal of Communication* 13: 3271–3278. Accessed November 29, 2023. <https://ijoc.org/index.php/ijoc/article/view/8018>.
- Taylor, Felicity, Marjorie Mitchell, Chantal Ripp, and Pascale Dangoisse. 2022. "Data Primer: Making Digital Humanities Research Data Public / Manuel d'introduction Aux Données : Rendre Publiques Les Données de Recherche En Sciences Humaines Numériques." Borealis V1. Ottawa: University of Ottawa. Accessed November 29, 2023. <https://doi.org/10.5683/SP3/OMLXTZ>.
- TEI Consortium. 2023. "Text Encoding Initiative." Accessed November 29. <https://tei-c.org/>.
- Tenen, Dennis. 2016. "Blunt Instrumentalism: On Tools and Methods." In *Debates in the Digital Humanities 2016*, edited by Matthew K. Gold and Lauren F. Klein, 83–91. Accessed November 29, 2023. <https://doi.org/10.5749/j.ctt1cn6thb.12>.

The Endings Project. 2023. "Building Sustainable Digital Humanities Projects." The Endings Project Team, University of Victoria. Accessed November 29. <https://endings.uvic.ca/>.

Thorpe, Kirsten, Kimberly Christen, Lauren Booker, and Monica Galassi. 2021. "Designing Archival Information Systems through Partnerships with Indigenous Communities: Developing the Mukurtu Hubs and Spokes Model in Australia." *Australasian Journal of Information Systems* 25: 1–22. Accessed November 29, 2023. <https://doi.org/10.3127/ajis.v25i0.2917>.

Thylstrup, Nanna Bonde, Daniela Agostinho, Annie Ring, Catherine D'Ignazio, and Kristin Veel, eds. 2021. *Uncertain Archives: Critical Keywords for Big Data*. Cambridge, MA: The MIT Press. Accessed November 29, 2023. <https://mitpress.mit.edu/books/uncertain-archives>.

Vincent, James. 2016. "Twitter Taught Microsoft's AI Chatbot to Be a Racist Asshole in Less than a Day." *The Verge*, March 24. Accessed November 29, 2023. <https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist>.

Wilkinson, Mark, Michel Dumontier, IJsbrand Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3. Accessed March 14, 2024. <https://doi.org/10.1038/sdata.2016.18>.

Wood-Charlson, Elisha M., Zachary Crockett, Chris Erdmann, Adam P. Arkin, and Carly B. Robinson. 2022. "Ten Simple Rules for Getting and Giving Credit for Data." *PLOS Computational Biology* 18(9): e1010476. Accessed March 14, 2024. <https://doi.org/10.1371/journal.pcbi.1010476>.

Zeng, Tong, Longfeng Wu, Sarah Bratt, and Daniel E. Acuna. 2020. "Assigning Credit to Scientific Datasets Using Article Citation Networks." *Journal of Informetrics* 14(2): 101013. Accessed November 29, 2023. <https://doi.org/10.1016/j.joi.2020.101013>.

