Cross-Domain Information Integration in Government: Hierarchies and Responsibilities

Zogheib, Ciara University of Toronto, Canada | ciara.zogheib@mail.utoronto.ca
Mahetaji, Kaushar University of Toronto, Canada | kaushar.mahetaji@mail.utoronto.ca

ABSTRACT
Cross-domain integration of information is increasingly identified as a priority across public sector contexts because (in theory) it enables the use of *more* information, including information from groups and communities historically excluded from public sector decision making. In this paper, we reject the tendency to take 'integration' for granted, arguing the need to position cross-domain integration as an information practice, and conducting mixed methods thematic analysis of government strategic documents to validate the utility of this approach. We find that depending on the type of information proposed to be integrated — digital data versus the knowledge of peoples and communities — our sample of Canadian government institutions treats cross-domain integration with differing levels of procedural rigour and detail. Reflecting ASIS&T 2024 themes of prioritizing responsibility and reflexivity in information practice, and of cultivating community partnerships through practice, not merely in name, we discuss the information hierarchies that emerge in the cross-domain information integration in government and the associated impacts on stakeholder communities.

KEYWORDS
cross-domain integration, public sector information use, knowledge integration, data integration

INTRODUCTION
A cursory look at published data and information strategies of governments and public sector organizations reveals that the integration of data and information across domains or contexts is seen as a desirable outcome (Government of Canada, 2023; United States Government, 2021; Central Digital and Data Office (UK), 2021). On the face of it, information integration as a goal for public sector institutions is intuitive: in our time of both big and broad data, there are vast amounts of information being generated from vastly different sources and domains of work (Swan, 2015; Hendler, 2014). It makes sense that organizations are seeking to integrate this information so that they can make effective use of more of it, but the goals of integration extend beyond merely expanding the *quantity* of usable information. Cross-domain integration is also advocated as a means of incorporating lived experience, Traditional Knowledges, and insights from groups and peoples often excluded from data and/or information work by government organizations.

Less intuitive, and certainly less interrogated, is exactly what this cross-domain integration looks like, and how it is to be implemented. This is potentially problematic: when we uncritically adopt 'integration' as a non-specific goal for our information use, we run the risk of being integrative in name only, limiting any attempt at centring "nuanced understanding and peripheral perspectives" (Freshwater & Fisher, 2014). If we value inclusive, people- and community-oriented usage of information in public institutions, it is our responsibility to take a reflexive, critical stance, and look more closely at the intent and actions behind cross-domain information integration.

Such is the aim of this paper. We briefly explore the ways in which cross-domain integration has been discussed, broadly and in public sector contexts, then make the case that adopting an information-focused lens is an effective way to explore the topic. To demonstrate the applicability of information approaches to cross-domain integration in the public sector, we report the results of a mixed methods thematic analysis of public Government of Canada documents and discuss the ways in which the findings support our contention that cross-domain information integration *must* be conducted responsibly and reflexively to produce intended outcomes.

RELATED WORK
Clarifying Terminology
It is helpful before exploring the substantive content of this paper to clarify terminology. Most studies of cross-domain information integration happen in fields outside of information studies (see the following subsection). Because of this, words like 'data', 'information', and 'knowledge' tend to be used colloquially, and occasionally interchangeably, rather than with the specific meanings implied by information scholars or in government documents. As information scholars, we engage discipline-specific understandings of 'data', 'knowledge', and 'information.' 'Data', synonymous here with 'information objects', refers to granular, often digital objects such as datasets, documents, or forms; data often lacks context, requiring interpretation to be rendered useful. 'Concepts' or 'knowledge and perspectives' on the other hand refer to more abstract ideas and encompass individual and collective knowledge. Knowledge and data are interlinked so that when 'data' is processed, it yields 'knowledge'. We use

'information' in a broad sense, as it is used in 'information studies', to refer to the study of information phenomena in general, including data and knowledge.

What is Integration?

Any definition of 'integration' will necessarily be accompanied by caveats and conditions, given the multiplicity of contexts in which the term is used. For our purposes, we combine multiple definitions to use 'integration' to refer to the synthesis (or merging, incorporating, blending…) of knowledge, information, and data (or methods, perspectives, ideas, paradigms, tools…) across domains (Aboelela et al., 2007; Hitchcock & Onwuegbuzie, 2022; Klein, 1990). Hjørland (2017) describes domains in detail; here we adopt Hjørland's conclusion that a domain is "a body of knowledge, defined socially and theoretically as the knowledge of a group of people sharing ontological and epistemological commitments" (2.4), for example, sectors of industry, academic disciplines, or stakeholder groups (Klein, 2022).

Social and communicative elements of cross-domain information integration at a high level — that is, integration of concepts, knowledge, or knowledge frameworks — have been conceptualized by scholars of interdisciplinarity (Canonico et al., 2017; Laursen et al., 2022), mixed methods research (Onwuegbuzie et al., 2009; Schoonenboom, 2022), and philosophy of science and knowledge (Boon & Van Baalen, 2019; Tobi & Kampen, 2018). We encounter similar language in works discussing the integration of Traditional Knowledges or community knowledge into institutional workflows in attempts to address historic exclusion and dismissal (Montenegro, 2019; Strunk, 2023).

There is a largely separate literature based in computer sciences, data sciences, and information systems that explores integration in terms of the combination of domain-specific datasets and other information objects. For example, Finkel et al. (2020) and Xu and Yu (2017) discuss computational means of 'integrating' heterogeneous datasets through shared data storage or merging strategies, while Wagner et al. (2023) explore how domain norms governing digital and physical file formats impact 'integrative data re-use'.

We thus see two general 'shapes' of cross-domain information integration. The first is integration on a conceptual level, where the entities being integrated are knowledge, ideas, or shared understandings of some phenomena — 'information' in the broad, colloquial sense. The second is integration on a technical level, where the entities being integrated are individual information objects (e.g. datasets, files) and their contents via linking, merging, or shared infrastructure. It is often implicitly assumed in the interdisciplinary literature that the second kind of integration facilitates the first, but the specific mechanisms by which this occurs are underexplored. In other terms, it is unclear how the integration of information objects from multiple domains enables or influences the practices of integrating ideas or bodies of knowledge to produce some understanding or insight that could not have been found by either domain in isolation.

Information Integration in the Public Sector

We can observe both 'shapes' of cross-domain information integration in the public sector. With regards to the first, conceptual integration: to Torfing (2019), collaboration and innovation are seen as desirable in the public sector to the end that they can enable "the integration of ideas into proper solutions" (p. 3). To Crosby and Bryson (2010), governments should aim to integrate public leadership structures and practices by "bringing diverse groups and organizations together in semi-permanent ways" and integrating their communications-related information (p. 211).

It is also common to discuss public sector integration technically, in terms of data and information objects. Gil-Garcia et al. (2009) suggest that information integration in government comprises four distinct but related elements: social networks, shared information, integrated data, and interoperable technical infrastructure. The idea of process integration, information sharing, and systems interoperability ('INT-IS-IOP') as fundamentally intertwined in government is thoroughly explored by Scholl et al. (2012). Comparable concepts to these understandings of information-centric integration also appear in policy documents by different names—for example, a 'comprehensive approach' to national security (Essens et al., 2016) or 'whole of government' decision making (Livingstone, 2007).

There has been some evaluation of integrative activity in the public sector, such as Scholl et al.'s (2012) analysis of success and failure conditions for INT-IS-IOP project reports from local and regional governments across Europe, Wang's (2018) case studies of two Chinese cities exploring information exchange and interoperability of government information systems, and Ren and Yang's ASIS&T 2023 exploration of cross-domain diffusion from science to policy. In this area of scholarship, however, mirroring what we observe in the literature on information integration more generally, there is no particular engagement with the assumption that technical, object-level integration automatically leads to the broader knowledge integration that seems to be the end goal.

An Information Studies Approach to Integration

Summarizing the preceding subsections: we know that cross-domain integration is seen as a desirable outcome in public sector contexts. We know that integration tends to involve data, knowledge (individual and collective), and information to varying degrees (from specific information objects to shared and represented, conceptual bodies of

knowledge). We know very little about how cross-domain information integration *actually happens* in public sector contexts — does integrating information objects facilitate the integration of concepts and bodies of knowledge? If so, how? What are the implications and impacts of these integration activities?

It is an intuitive next step, faced with this gap in our understanding, to position cross-domain integration as an information practice (or a collection thereof), since it is a way of using or interacting with information (Bates, 2010). Information scholars have studied cross-domain integration as a feature of academic research settings by investigating the information practices that comprise interdisciplinarity (Palmer, 2013; Zogheib, 2023); here we attempt to apply that same framing, along with terminology from information studies more broadly, to cross-domain integration in public sector settings.

The value of this framing of cross-domain integration is that we can build on existing, ongoing work from information scholars about domain-analytic approaches to information work (Hjørland, 2002; Luo et al., 2022), and that we can engage with our two categories of integration ('technical' and 'conceptual') by drawing on critical information scholarship to question the ways in which institutions value different types of information (Thomer et al., 2022; Sadowski, 2019) and the implications that those value hierarchies have for stakeholder communities (Poole, 2023).

RESEARCH OBJECTIVES

The research objectives of this paper are to

1. identify if there is an empirical basis for positioning cross-domain integration in government as an information practice, validating information studies as a productive approach for studying this topic, and
2. demonstrate the application of concepts and framing from information studies to the topic of cross-domain integration in the government.

We focus on the sample case of integration in a Canadian government context to restrict analysis to a manageable scope and make a novel contribution to existing research, where to our knowledge no studies of cross-domain integration in the Canadian federal public sector currently exist.

METHODS

To explore information integration in our chosen government context, this paper draws on a mixed methods document analysis approach. For a subject such as integration, where definitions are fluid and often inconsistent, published documents can demonstrate the different ways that topics are conceived of in official contexts and illuminate government priorities or challenges (Cardno, 2019).

Data Collection

Documents for analysis are collected from publicly accessible Government of Canada webpages in June and July of 2023. Here, we focus specifically on ministerial mandate letters and departmental plans. Per the Office of the Prime Minister (2021), mandate letters "outline the objectives that each minister will work to accomplish, as well as the pressing challenges they will address in their role". Published mandate letters and their utility as a means of understanding government priorities have been the subject of some debate in the public sphere (Lang, 2022; Novovic, 2022), but they continue to be used as research material (Lucyk, 2020; McKelvey & Macdonald, 2019; Waubert de Puiseau, 2016).

Ministers of the Crown are responsible for carrying out a portfolio assigned to them by the Prime Minister. This portfolio can include the administration of federal government departments or ministries. These departments and ministries publish annual departmental plans which describe "priorities, strategic outcomes, programs, expected results and associated resource requirements" (Treasury Board Secretariat, 2017). Departmental plans have been discussed in research contexts as signifiers of policy priorities (Diabo, 2020; Graff, 2021).

In total, a corpus of 38 mandate letters and 30 departmental plans (68 documents in total) is collected here. Official Government of Canada publications are made available in both French and English; only English versions are analyzed.

Data Analysis

Exploratory quantitative analysis was used to guide our initial inquiry. The Requests library in Python (Reitz, 2023) was used to scrape the text contents of the 2021 Ministerial Mandate Letters and 2022-2023 Departmental Plans. Texts were searched to locate the following keywords of interest:

Integrate, Integrated, Integrating, Integration

The frequency of these keywords in each type of document was then tabulated. We collected mandate letters and departmental plans that were found to contain at least one instance of the aforementioned keywords of interest. The

documents were then explored further through qualitative thematic analysis. From these documents, we extracted text sections — i.e. document subsections or paragraphs, where applicable, or lists for text in point form — which discussed integration or similar ideas. In total, 153 text sections were reviewed and thematically coded.

Themes were identified by one of the authors following the iterative, qualitative process described by Braun and Clarke (2012). Potential thematic groups were identified during a preliminary reading of the selected documents, with this 'first pass' sensitized by concepts from the information studies and interdisciplinary literature. As these initial thematic groups resurfaced during subsequent re-readings of the documents, they were codified as themes. Text selections were read through the lens of these preliminary themes; where themes appeared in particular text selections, their presence was recorded. This process was neither wholly deductive (i.e. themes identified based on theory or existing ideas from the literature) nor wholly inductive (i.e. themes emerging entirely based on a 'bottom-up' exploration of the data). Rather, in the joint approach discussed by Byrne (2022), the author explored the texts with a particular subject in mind (integration) but with openness to finding the as-yet-unknown themes that occurred 'around' the subject. Theme frequency is visualized using the ggplot2 library in R (Wickham, 2009) and palettes from Tol (2021).

Additional quantitative analysis was used to support further exploration of the themes identified qualitatively. We investigated what Armborst (2017) calls 'thematic proximity', a term that describes the tendency of certain thematic codes to occur together. Thematic proximity can be used to develop meta-themes and to explore potential relationships between themes (Oleinik, 2011). In this paper, thematic proximity was evaluated using the cooccur package in R (Griffith et al., 2016), which has been used across various other research contexts (Miles et al., 2020; Zwetsloot et al., 2021), including for the analysis of thematic data from Tweets and interviews relating to cultural services (Johnson et al., 2019). The use of simple descriptive statistics is intended here to complement qualitatively-derived themes, rather than to imply generalizability of findings or representativeness of our sample documents of some hypothetical larger corpus. Likewise, p values are used to support interpretation rather than to assign undue weight to an arbitrary significance threshold.

Reproducibility, Reliability, and Reflexivity Considerations

The value of attempts to quantify reproducibility and reliability in qualitative thematic research is the subject of some controversy (O'Connor & Joffe, 2020). In this paper, we undertake thematic analysis as an interpretative research approach rather than a positivist one but nevertheless include as an exercise in completeness a quantitative evaluation of intercoder reliability. This evaluation, the Python and R scripts used for quantitative analyses and text extraction, and the Microsoft Excel spreadsheets containing the results of the thematic analysis, are available in a GitHub repository: https://github.com/zogheibc/IntegrationHierarchies.

In conducting our thematic analysis, we also foreground reflexivity, acknowledging that both our data materials and our experiences processing said materials are 'contextually-situated' (Braun & Clarke, 2022). We recognize that the codes and themes are produced, not discovered; they are interpreted, not simply identified; and they are open to further investigation (Braun & Clarke, 2022).

FINDINGS

Integration and related keywords are mentioned in 12 out of 38 mandate letters (approximately 32%) and in 26 out of 30 departmental plans (approximately 87%). In general, as shown in Table 1, all keywords occur more frequently in the text of departmental plans than in mandate letters, and cross-domain integration is generally positioned as an activity to be undertaken in support of a goal or mandate, rather than a mandated outcome in itself.

Keyword	Frequency in Departmental Plans	Frequency in Mandate Letters
Integrate	42	5
Integrated	52	9
Integrating	13	0
Integration	77	3

Table 1. The frequency with which keywords of interest appear in departmental plan documents and mandate letter documents.

Digging deeper into the relevant text sections, where our keywords of interest were mentioned, 153 text selections were extracted and analyzed. Through the iterative process of re-reading and coding, 9 themes were identified. These themes are described in Table 2, and further detailed in the codebook (in the GitHub repository).

Theme 5, ‘Communities and People’, groups discussion of cross-domain integration as it pertains to equity-seeking groups and communities into a single theme. The intention here is not to generalize the individual needs and relationships of these communities but to highlight the common ways they tend to be positioned in relation to integration in mandate letters and departmental plans.

Theme 9, ‘Other’, is intentionally left broad to capture instances where keywords are mentioned, but in a sense not related to the subject of this paper. Themes are emphatically not mutually exclusive — for example, an excerpt may discuss data and information integration while also touching on themes of research and innovation. The relative frequencies of the different themes are visualized in Figure 1.

Theme	Occurrence (Out of 153 text sections)	Usage
Data and Information Integration	32	Cases where integration is discussed in terms of processes – aggregation, exchange, use – that happen to data and information. In exploring this theme, we can understand information as a discrete ‘piece’ of knowledge (e.g. a string of text submitted as an entry to a field of an online form).
Knowledge and Perspective Integration	49	Cases where, rather than a particular piece of data or information, the entities being integrated are the more broad knowledge or perspectives (a frequently-used term) of a given group.
Program and Organization Integration	99	Cases where integration takes the form of coordination of activities, policies, or initiatives across multiple agencies.
Technology	33	Integration as related to technological developments, e.g. ‘integrated platforms’ or digital tools for facilitating integration across government.
Communities and People	47	This theme is intended to capture the relationship that emerges between integration and the government’s consideration of equity-seeking groups and communities.
Crossdisciplinarity	11	Integration occurring in the context of inter-, multi-, or -transdisciplinary approaches.
Research	18	Integration discussed alongside explicit uses of the word ‘research’ but also related ideas such as science, experiments, and evidence.
Innovation	16	Positioning of integration as an innovative activity or of ‘innovation’ as the entity being integrated.
Other	29	For uses of ‘integration’ distinct from the usage of interest in this paper (e.g. social integration of new immigrants, integration of job seekers into workplaces)

Table 2. Overview of themes identified during thematic analysis of discussions of ‘integration’ in departmental plan and mandate letter documents.

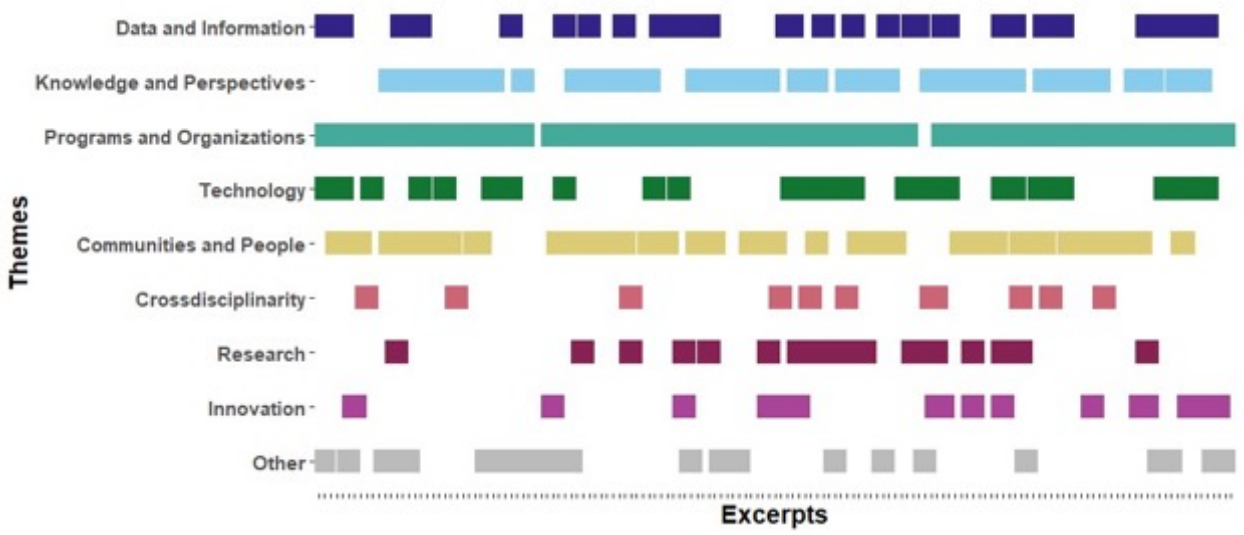

Figure 1. A visualization of the presence of themes in text sections; more colour in a horizontal row shows that the relevant theme appears in more sections. Each 'tick' on the x-axis represents a text section.

Theme A	Theme B	Cooccurrence less than expected	Cooccurrence more than expected
Data and Information Integration	Technology		***
Data and Information Integration	Other	*	
Knowledge and Perspective Integration	Technology	**	
Knowledge and Perspective Integration	Community and Diversity		***
Knowledge and Perspective Integration	Other	***	
Program and Organization Integration	Community and Diversity		**
Program and Organization Integration	Crossdisciplinarity	*	
Program and Organization Integration	Other	***	
Technology	Community and Diversity	***	
Community and Diversity	Other	**	
Crossdisciplinarity	Research		**
Research	Innovation		***

Table 3. Significant results of cooccurrence analysis of themes using the cooccur package in R. $p < 0.05$ is indicated by *, $p < 0.01$ is indicated by **, and $p < 0.001$ is indicated by *. Smaller p values indicate that cooccurrence is less likely to have occurred by chance. p thresholds are arbitrary and support our exploration of relationships rather than marking 'truth'.**

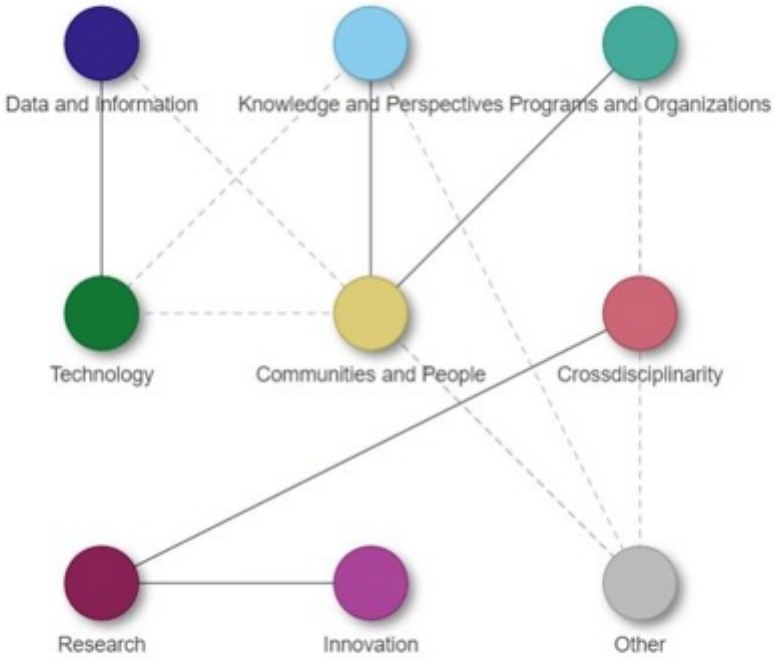

Figure 2. A network diagram demonstrating relationships between themes identified during thematic analysis. Solid lines between nodes indicate more than expected cooccurrence, whereas dotted lines indicate lower than expected cooccurrence.

We turn to the cooccur package in R to explore thematic proximity, or whether two themes tend to occur together more frequently than expected, assuming as our null hypothesis that every theme is equally likely to cooccur with every other theme. Cooccurrence between themes is visualized as a network in Figure 2, with those relationships that meet or pass a conventional threshold for statistical significance ($p<0.05$) detailed in Table 3. We emphasize again that quantitative exploration of relationships between themes is conducted to complement and challenge qualitative observations, rather than to suggest 'truth' or objectivity.

Some results worthy of particular attention: 'Data and Information Integration' has a strongly positive association with 'Technology.' This relationship is intuitive. If a department's aim is to integrate data and information objects, which are largely stored in digital formats, it makes sense that that the department would also consider how to use technologies as tools for integration. Conversely, 'Knowledge and Perspective Integration' has a lower-than-expected cooccurrence with 'Technology,' but a significantly higher-than-expected cooccurrence with 'Communities and People'. This relationship, too, supports something we might have observed qualitatively from our thematic analysis: language discussing the integration of "Indigenous knowledge" or "Gender-Based Analysis+ (GBA+) perspectives" is common throughout our corpus, but given the broader nature of knowledge as belonging to or produced by an entire group or community (versus 'concrete' data as digital artefacts), this language tends not to include specific technical solutions. Instead, the integrative actions or solutions relating to community knowledge tend to relate to 'Program and Organization Integration', though — importantly — they still tend to lack specific mechanisms. That is to say: cross-domain information integration is discussed with differing amounts of procedural specificity and rigour depending on the kinds of information being integrated.

DISCUSSION

Cross-Domain Integration as Information Practice

The government documents explored here demonstrate a strong relationship between government understandings of integration and — to use a phrase from Bates (1999) and later work by Hartel (2020) — the 'red thread of information'. This thread encompasses topics including data, information, knowledge, research, information systems (technical and otherwise), and the institutions and organizations that steward them (Hartel 2020). The thematic analysis conducted here shows that where integration is mentioned in mandate letters and departmental plans, it is frequently in association with one or more of these information topics. Thus, addressing our first research objective, our proposed framing of cross-domain integration in the government as an information practice is validated by the language used by the organizations involved.

Information Integration Hierarchies in Government

Our second research objective was to apply our information practice lens to articulate and engage with particular subjects of concern relative to cross-domain information integration in the government. We do so by exploring a particularly fascinating finding that emerged during thematic analysis and was reinforced by the results of the cooccurrence analysis of themes.

We can recall (from Table 3 and Figure 2 in the 'Findings' section) that in excerpts of mandate letters and departmental plans where integration was mentioned, cases where data and information objects were the things being integrated had higher than expected associations with themes of technology, while cases where knowledge and perspectives were the things being integrated had strong associations with diversity, equity, and inclusion considerations relating to people or communities. These distinctions, even solely on the level of word choice, are not neutral. 'Data' is not the same thing as 'knowledge' (Liew, 2007). This distinction matters in the current moment, characterized by large volumes of open data that continues to increase in velocity, variety, scope, and resolution (Kitchin 2014). Governments are leveraging this explosion in open data for decision making, impacting all spheres of life and making data ethics an important consideration for policymakers (Roy, 2016). 'Data', in a modern context, is valued as political and economic capital; to possess and invoke data is to take advantage of the authority, power, and influence afforded to data-as-concept (Mejias & Couldry, 2019; Sadowski, 2019). Data, accordingly, is frequently discussed in relation to 'data sovereignty', 'technological sovereignty', and 'digital sovereignty', or the control and management of data (Couture & Toupin, 2019) – we care about what data can tell us, so it matters what we do with it.

Here is where our findings, when we look deeper, become concerning. When Government of Canada documents discuss the integration of data, specific systems or practices with which that integration will take place are described. In contrast, when Government of Canada documents discuss the integration of knowledge and perspectives, the communities or peoples whose knowledge is being integrated are mentioned, but *few to no specific practices or systems of integration are described*. These practices and systems need not be technological – indeed, we previously critiqued exclusive focus on technical integration at the expense of conceptual integration – but we might expect some comparable level of detail in terms of describing *how* community knowledge will be integrated into existing work or policy, rather than merely stating that they will be. To draw on specific examples: almost every departmental plan mentions 'integrating' GBA+ lenses into their policy and operations. The practices through which GBA+ lenses and results will be integrated are not described. Several departmental plans mention the integration of Indigenous Traditional Knowledges. Likewise, specific practices or mechanisms for this integration are not described. We can see, in effect, a hierarchy in which the integration of certain types of information (generally digital data) is described in greater detail than others (generally knowledges associated with specific communities).

It is unsurprising that data integration is treated more rigorously than conceptual integration of knowledge or perspectives. The failure of the Canadian government to take seriously the integration of Indigenous knowledge and perspectives in particular is reflected elsewhere in the recent literature — even leaving aside as out of scope the topic of government use and abuse of data from and about Indigenous peoples—see Kukutai & Taylor (2016) for a thorough discussion. Westwood et al. (2023), in a review of government-affiliated forest science reports, find that, "although nearly a quarter of [reviewed publications] mentioned Indigenous, local or community knowledge, none clearly attributed members or organizations from those communities at the level of item authorship" (p. 13). Alexander et al. (2019), doing comparable work in the area of freshwater research, found that of publications that mention integrating Indigenous knowledge, fewer than half included authorship attribution for the Indigenous stewards and sharers of that knowledge. It is clear from the frequency with which integration is mentioned alongside specific (often marginalized) communities in mandate letters and departmental plans that there is downward pressure supporting greater integration of knowledge and perspectives from diverse communities in the Canadian federal government. However, when these kinds of integration are simply noted with little to no explanation of the specific mechanisms and practices to carry out integration, there is again a risk of integration in name only (Freshwater & Fisher, 2014). Future discussions of integration in the public sector could benefit from treating the integration of knowledge and perspectives as something requiring as much planning and infrastructure as the integration of more easily tangible information objects.

Limitations

The findings presented in this empirical study should be considered in light of limitations of the research approach. Paterson (2021) raises the topic of 'fugitive documents', government publications that may, through workflow or labeling issues, remain absent from public or searchable archives. These archival gaps could have affected the initial selection of documents. Likewise, our extraction of text sections for further thematic analysis was limited to those containing keywords related to 'integration'. We chose to focus our work on integration in particular to meet our first research objective and supplement existing work on integration as a popular 'buzzword' and phenomenon of interest. It is possible though that in doing so we overlooked discussion of cross-domain 'synthesis' or other related

concepts. Also worth considering is the fact that Government of Canada administrative documents published online may be more likely to contain information about strategic- or policy-level goals than about the internal procedures by which those goals will be pursued — that is, it is very possible that additional information about specific information practices for integration in government may be in internal documents unavailable to the public (Cardno, 2019). This in itself would inhibit responsible and reflexive engagement with cross-domain integration practices by preventing public scrutiny.

Looking Forward

This study is intended as a starting point. We use document analysis to validate an information-centred approach to exploring cross-domain information integration in government; future scholarship could expand on this approach through interviews or observational studies of the public servants and information specialists doing integrative work to question whether *actual* integration practice corresponds to integration practice *as articulated* in the strategic documents explored here. Incorporating perspectives from critical data studies can further explain the discrepancies in the integration of data and knowledge in government and communities through methodologies that locate moments of 'data friction' and track data flows across software infrastructure (Bates et al., 2016).

CONCLUSION

This empirical work used the case of a particular set of publications from a particular government context to demonstrate:

1. that cross-domain integration can be meaningfully conceived of as an information practice; and
2. that further investigation of this practice suggests that cross-domain integration in its current form may be perpetuating information hierarchies, undercutting the intended function of integration.

We take no issue with 'integration' as a desirable outcome for government data and information strategies. In fact, we see it as worthwhile and important work. Because of this conviction, we argue that in order for cross-domain information integration to have a meaningful impact, government information stewards and policymakers must actively avoid the tendency to discuss and formalize integration solely in terms of datasets and technological information objects, and bring to the fore the people, communities, and other knowledges involved.

GENERATIVE AI USE

We confirm that we did not use generative AI tools/services to author this submission.

AUTHOR ATTRIBUTION

First Author: conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, visualization, writing – original draft, writing – review and editing; Second Author: conceptualization, formal analysis, validation, writing – original draft, writing – review and editing

REFERENCES

Aboelela, S. W., Larson, E., Bakken, S., Carrasquillo, O., Formicola, A., Glied, S. A., Haas, J., & Gebbie, K. M. (2007). Defining interdisciplinary research: Conclusions from a critical review of the literature. *Health Services Research*, 42(1 Pt 1), 329–346. https://doi.org/10.1111/j.1475-6773.2006.00621.x

Alexander, S. M., Provencher, J. F., Henri, D. A., Taylor, J. J., Lloren, J. I., Nanayakkara, L., Johnson, J. T., & Cooke, S. J. (2019). Bridging Indigenous and science-based knowledge in coastal and marine research, monitoring, and management in Canada. *Environmental Evidence*, 8(1), 1–24. https://doi.org/10.1186/s13750-019-0181-3

Armborst, A. (2017). Thematic Proximity in Content Analysis. *SAGE Open*, 7(2), 1–11. https://doi.org/10.1177/2158244017707797

Bates, M. J. (1999). The invisible substrate of information science. *Journal of the American Society for Information Science,* 50(12), 1043–1050. https://doi.org/10.1002/(SICI)1097-4571(1999)50:12<1043::AID-ASI1>3.0.CO;2-X

Bates, J., Lin, Y.W., & Goodale, P. (2016). Data journeys: Capturing the socio-material constitution of data objects and flows. *Big Data & Society, 3*(2), 1–12. https://doi.org/10.1177/2053951716654502

Boon, M., & Van Baalen, S. (2019). Epistemology for interdisciplinary research – shifting philosophical paradigms of science. *European Journal for Philosophy of Science*, 9(1), 1–28. https://doi.org/10.1007/s13194-018-0242-4

Braun, V., & Clarke, V. (2022). Conceptual and design thinking for thematic analysis. *Qualitative Psychology*, 9(1), 3–26. https://doi.org/10.1037/qup0000196

Braun, V., & Clarke, V. (2012). Thematic analysis. In *APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 57–71). American Psychological Association. https://doi.org/10.1037/13620-004

Byrne, D. (2022). A worked example of Braun and Clarke's approach to reflexive thematic analysis. *Quality & Quantity*, 56(3), 1391–1412. https://doi.org/10.1007/s11135-021-01182-y

Canonico, P., De Nito, E., Esposito, V., Martinez, M., & Iacono, M. P. (2017). The adoption of knowledge integration mechanisms in an interdisciplinary research project. *Management Research Review*, 40(5), 604–622. https://doi.org/10.1108/MRR-04-2016-0099

Cardno, C. (2019). Policy Document Analysis: A Practical Educational Leadership Tool and a Qualitative Research Method. *Educational Administration: Theory and Practice*, 24(4), 623–640. https://doi.org/10.14527/kuey.2018.016

Central Digital and Data Office (UK). (2021). Integrated data service: Service Standard Report. GOV.UK. https://www.gov.uk/service-standard-reports/integrated-data-service

Couture, S., & Toupin, S. (2019). What does the notion of "sovereignty" mean when referring to the digital? *New Media & Society*, 21(10), 2305–2322. https://doi.org/10.1177/1461444819865984

Crosby, B. C., & Bryson, J. M. (2010). Integrative leadership and the creation and maintenance of cross-sector collaborations. *The Leadership Quarterly*, 21(2), 211–230. https://doi.org/10.1016/j.leaqua.2010.01.003

Diabo, R. (2020). Trudeau Government's "Recognition & Self-Determination" Tables. *Indigenous Policy Journal*, 31(1). http://indigenouspolicy.org/index.php/ipj/article/view/696

Essens, P. M., Thompson, M. T., Karrasch, A., & Jermalavičius, J. (2016). Building Effective Collaboration in a Comprehensive Approach. NATO. http://resolver.tudelft.nl/uuid:6897c228-618c-47d6-9732-2fe9b64239cb

Finkel, M., Baur, A., Weber, T. K. D., Osenbrück, K., Rügner, H., Leven, C., Schwientek, M., Schlögl, J., Hahn, U., Streck, T., Cirpka, O. A., Walter, T., & Grathwohl, P. (2020). Managing collaborative research data for integrated, interdisciplinary environmental research. *Earth Science Informatics*, 13(3), 641–654. https://doi.org/10.1007/s12145-020-00441-0

Freshwater, D., & Fisher, P. (2014). (Con) Fusing Commerce and Science: Mixed Methods Research and the Production of Contextualized Knowledge. *Journal of Mixed Methods Research*, 8(2), 111–114. https://doi.org/10.1177/1558689814526804

Gil-Garcia, J. R., Chun, S. A., & Janssen, M. (2009). Government information sharing and integration: Combining the social and the technical. *Information Polity*, 14(1,2), 1–10. https://doi.org/10.3233/IP-2009-0176

Government of Canada. (2023). 2023–2026 Data Strategy for the Federal Public Service. https://www.canada.ca/en/treasury-board-secretariat/corporate/reports/2023-2026-data-strategy.html

Graff, D. (2021). Ethics and Arms Sales: A Discourse Analysis of Canadian Foreign Policy [Thesis, Université d'Ottawa / University of Ottawa]. https://doi.org/10.20381/ruor-26422

Griffith, D. M., Veech, J. A., & Marsh, C. J. (2016). cooccur: Probabilistic Species Co-Occurrence Analysis in R. *Journal of Statistical Software*, 69(2016), 1–17. https://doi.org/10.18637/jss.v069.c02

Hartel, J. (2020). The red thread of information. *Journal of Documentation*, 76(3), 647–656. https://doi.org/10.1108/JD-04-2019-0067

Hendler, J. (2014). Data Integration for Heterogenous Datasets. *Big Data*, 2(4), 205–215. https://doi.org/10.1089/big.2014.0068

Hitchcock, J. H., & Onwuegbuzie, A. J. (Eds.). (2022). *The Routledge handbook for advancing integration in mixed methods research*. Routledge.

Hjørland, B. (2002). Domain analysis in information science: Eleven approaches – traditional as well as innovative. *Journal of Documentation*, 58(4), 422–462. https://doi.org/10.1108/00220410210431136

Hjørland, B. (2017). Domain analysis. In *Encyclopedia of Knowledge Organization*. ISKO. http://www.isko.org/cyclo/domain_analysis

Johnson, M. L., Campbell, L. K., Svendsen, E. S., & McMillen, H. L. (2019). Mapping Urban Park Cultural Ecosystem Services: A Comparison of Twitter and Semi-Structured Interview Methods. *Sustainability*, 11(6137), 1–21. https://doi.org/10.3390/su11216137

Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1), 1–12. https://doi.org/10.1177/2053951714528481

Klein, J. T. (1990). *Interdisciplinarity: History, Theory, and Practice*. Wayne State University Press.

Klein, J. T. (2022). Interdisciplinary Collaboration: Insights from Art-Science, Digital Humanities, and SHAPE-ID. *Scientia Poetica*, 26(1), 201–214. https://doi.org/10.1515/scipo-2022-008

Kukutai, T., & Taylor, J. (Eds.). (2016). *Indigenous Data Sovereignty: Toward an agenda*. ANU Press. https://doi.org/10.22459/CAEPR38.11.2016

Lang, E. (2022). Ministerial mandate letters: Another nail in the coffin of cabinet government. *Policy Options*. https://policyoptions.irpp.org/magazines/february-2022/ministerial-mandate-letters-another-nail-in-the-coffin-of-cabinet-government/

Laursen, B. K., Motzer, N., & Anderson, K. J. (2022). Pathways for assessing interdisciplinarity: A systematic review. *Research Evaluation*, 31(3), 326–343. https://doi.org/10.1093/reseval/rvac013

Liew, A. (2007). Understanding Data, Information, Knowledge and Their Inter-Relationships. *Journal of Knowledge Management Practice*, 8(2), 1–16.

Livingstone, A. (2007). Interdisciplinarity and the Whole of Government: Theory and Praxis. 9, 221–229. http://ismor.com/cornwallis/cornwallis_2007/CXII_2007_15_Livingstone.pdf

Lucyk, K. (2020). Commentary - Intersectoral action on the social determinants of health and health equity in Canada: December 2019 federal government mandate letter review. *Health Promotion and Chronic Disease Prevention in Canada : Research, Policy and Practice*, 40(10), 314–323. https://doi.org/10.24095/hpcdp.40.10.03

Luo, P., Hong, L., Wang, J., Wang, S., Guo, X., Gao, Z., & Cho, S. W. (2022). Learning Domain-specific Semantic Representation from Weakly Supervised Data to Improve Research Dataset Retrieval. *Proceedings of the Association for Information Science and Technology*, 59(1), 205–214. https://doi.org/10.1002/pra2.616

McKelvey, F., & Macdonald, M. (2019). Artificial Intelligence Policy Innovations at the Canadian Federal Government. *Canadian Journal of Communication*, 44(2), 43–50. https://doi.org/10.22230/cjc.2019v44n2a3509

Mejias, U. A., & Couldry, N. (2019). Datafication. *Internet Policy Review*, 8(4), 1–10. https://policyreview.info/concepts/datafication

Miles, L. A., Bowman, R. L., Merlinsky, T. R., Csete, I. S., Ooi, A. T., Durruthy-Durruthy, R., Bowman, M., Famulare, C., Patel, M. A., Mendez, P., Ainali, C., Demaree, B., Delley, C. L., Abate, A. R., Manivannan, M., Sahu, S., Goldberg, A. D., Bolton, K. L., Zehir, A., … Levine, R. L. (2020). Single-cell mutation analysis of clonal evolution in myeloid malignancies. *Nature*, 587(7834), 477–482. https://doi.org/10.1038/s41586-020-2864-x

Montenegro, M. (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. https://doi.org/10.1108/JD-08-2018-0124

Novovic, G. (2022). In defence of ministerial mandate letters. *Policy Options*. https://policyoptions.irpp.org/magazines/ministerial-mandate-letters/

O'Connor, C., & Joffe, H. (2020). Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. *International Journal of Qualitative Methods*, 19, 1–13. https://doi.org/10.1177/1609406919899220

Office of the Prime Minister. (2021, December 16). Mandate Letters. Prime Minister of Canada. https://pm.gc.ca/en/mandate-letters

Oleinik, A. (2011). Mixing quantitative and qualitative content analysis: Triangulation at work. *Quality & Quantity*, 45(4), 859–873. https://doi.org/10.1007/s11135-010-9399-4

Onwuegbuzie, A. J., Slate, J. R., Leech, N. L., & Collins, K. M. (2009). Mixed data analysis: Advanced integration techniques. *International Journal of Multiple Research Approaches*, 3(1), 13–33. https://doi.org/10.5172/mra.455.3.1.13

Palmer, C. L. (2013). *Work at the Boundaries of Science: Information and the Interdisciplinary Research Process*. Springer Science & Business Media.

Paterson, S. (2021). The Government information landscape in Canada (IFLA Professional Reports, pp. 27–47).

Poole, A. H. (2023). Data Flourishing: Developing Human-Centered Data Science through Communities of Ethical Practice. *Proceedings of the Association for Information Science and Technology*, 60(1), 338–352. https://doi.org/10.1002/pra2.793

Reitz, K. (2023). Requests: HTTP for Humans (2.28.2) [Python]. https://requests.readthedocs.io/en/latest/

Ren, C., & Yang, M. (2023). Study on the Characteristics of Cross-Domain Knowledge Diffusion from Science to Policy: Evidence from Overton Data. *Proceedings of the Association for Information Science and Technology*, 60(1), 368–378. https://doi.org/10.1002/pra2.795

Roy, J. (2016). Data, Dialogue, and Innovation: Opportunities and Challenges for "Open Government" in Canada. *Journal of Innovation Management*, 4(1), 22–38. https://doi.org/10.24840/2183-0606_004.001_0004

Sadowski, J. (2019). When data is capital: Datafication, accumulation, and extraction. *Big Data & Society*, 6(1), 1–12. https://doi.org/10.1177/2053951718820549

Scholl, H. J., Kubicek, H., Cimander, R., & Klischewski, R. (2012). Process integration, information sharing, and system interoperation in government: A comparative case analysis. *Government Information Quarterly*, 29(3), 313–323. https://doi.org/10.1016/j.giq.2012.02.009

Schoonenboom, J. (2022). Developing the meta-inference in mixed methods research through successive integration of claims. In J. H. Hitchcock & A. J. Onwuegbuzie (Eds.), *The Routledge handbook for advancing integration in mixed methods research* (pp. 55–70). Routledge.

Strunk, K. K. (2023). QuantQueer: Renovating Quantitative Methods through Queer and Critical Theoretical Traditions. *Educational Studies*, 559(4), 1–14. https://doi.org/10.1080/00131946.2023.2276229

Swan, M. (2015). Philosophy of Big Data: Expanding the Human-Data Relation with Big Data Science Services. *2015 IEEE First International Conference on Big Data Computing Service and Applications*, 468–477. https://doi.org/10.1109/BigDataService.2015.29

Thomer, A. K., Starks, J. R., Rayburn, A., & Lenard, M. C. (2022). Maintaining Repositories, Databases, and Digital Collections in Memory Institutions: An Integrative Review. *Proceedings of the Association for Information Science and Technology*, 59(1), 310–323. https://doi.org/10.1002/pra2.755

Tobi, H., & Kampen, J. K. (2018). Research design: The methodology for interdisciplinary research framework. *Quality & Quantity*, 52(3), 1209–1225. https://doi.org/10.1007/s11135-017-0513-8

Tol, P. (2021). Qualitative Colour Palette. https://personal.sron.nl/~pault/#sec:qualitative

Torfing, J. (2019). Collaborative innovation in the public sector: The argument. *Public Management Review*, 21(1), 1–11. https://doi.org/10.1080/14719037.2018.1430248

Treasury Board of Canada Secretariat. (2017). Government of Canada Strategic Plan for Information Management and Information Technology 2017 to 2021. https://www.canada.ca/en/treasury-board-secretariat/services/information-technology/strategic-plan-2017-2021.html

United States Government. (2021). 2021 Action Plan—Federal Data Strategy. https://strategy.data.gov/

Wagner, T. L., Fenlon, K., & Sorensen, A. (2023). "Garbage Bags Full of Files": Exploring Sociotechnical Perceptions of Formats within the Recovery and Reuse of Scientific Data. *Proceedings of the Association for Information Science and Technology*, 60(1), 396–407. https://doi.org/10.1002/pra2.798

Wang, F. (2018). Understanding the dynamic mechanism of interagency government data sharing. *Government Information Quarterly*, 35(4), 536–546. https://doi.org/10.1016/j.giq.2018.08.003

Waubert de Puiseau, F. (2016). On the Drama of Canadian Public Arts Funding: An Analysis of Justin Trudeau's Mandate Letter to Mélanie Joly. *Canadian Theatre Review*, 167, 7–11. https://doi.org/10.3138/ctr.167.002

Westwood, A. R., Hutchen, J., Kapoor, T., Klenk, K., Saturno, J., Antwi, E. K., Egunyu, F., Cortini, F., Robertson, M., Le Noble, S., Wang, J., Falconer, M., & Nguyen, V. M. (2023). A systematic map of knowledge exchange across the science-policy interface for forest science: How can we improve consistency and effectiveness? *Ecological Solutions and Evidence*, 4(1), 1–17. https://doi.org/10.1002/2688-8319.12214

Wickham, H. (2009). *Ggplot2: Elegant graphics for data analysis*. Springer.

Xu, W., & Yu, J. (2017). A novel approach to information fusion in multi-source datasets: A granular computing viewpoint. *Information Sciences*, 378, 410–423. https://doi.org/10.1016/j.ins.2016.04.009

Zogheib, C. (2023). Interdisciplinary Integration as Information Work. *Issues In Interdisciplinary Studies*, 41(1), 57–78.

Zwetsloot, M. J., Leeuwen, J., Hemerik, L., Martens, H., Simó Josa, I., Broek, M., Debeljak, M., Rutgers, M., Sandén, T., Wall, D. P., Jones, A., & Creamer, R. E. (2021). Soil multifunctionality: Synergies and trade-offs across European climatic zones and land uses. *European Journal of Soil Science*, 72(4), 1640–1654. https://doi.org/10.1111/ejss.13051