

Structured Data and Human Rights Violations

D. Grant Campbell* and Martin Nord**

* ** Faculty of Information and Media Studies, University of Western Ontario, Canada,

*<gcampbel@uwo.ca>, **<mnord@uwo.ca>

D. Grant Campbell is an Associate Professor in the Faculty of Information and Media Studies at the University of Western Ontario in Canada. His research areas include bibliographic description, classification theory, and the application of Knowledge Organization practices and theories to LGBTQ+ communities and to dementia caregiving.



Martin I. Nord is Librarian at the Shannon Library at Saint Thomas More College, federated with the University of Saskatchewan, and Advising Librarian at the Saskatoon Theological Union Library. He is also a PhD candidate in Library and Information Science in the Faculty of Information & Media Studies at Western University. In his research, Martin is interested in exploring the ways in which documents and information are integrated into our ways of knowing, understanding, and ultimately, living with each other. He has published in Library & Information Science Research and the Canadian Journal of Academic Librarianship.



Campbell, D. Grant and Martin Nord. 2023. "Structured Data and Human Rights Violations". *Knowledge Organization* 50 (6): 427-438. 34 references. DOI:10.5771/0943-7444-2023-6-427.

Abstract: This presentation examines how knowledge organization can help us understand human rights violations in a way that engenders both rational comprehension and emotional empathy. Knowledge organization can foster a useful awareness of injustice through its use of structured data, which serves as an intermediary step between the intimate narrative that appeals to our emotions and the quantified syntheses that engage our reason. This paper lists four possible effects. First, structured data allows for more accurate statistics. Second, structured data provides links from one isolated incident to other similar incidents. Third, structured data lends itself to regularity and consistency of display, thereby appealing to rational comprehension, but on a more direct emotional level. Finally, structured data lends itself to memorial devices that rely on discovery and chance encounters, rather than through memorial installations that people approach purposefully.

Received: 16 August 2023; Revised: 12 October August 2023; Accepted 16 October 2023

Keywords: human rights violations; structured data; impact; ethics in knowledge organization

1.0 Introduction

This presentation examines the relevance of knowledge organization to a problem identified by both cognitive psychologists and humanities scholars: the problem of presenting information about pressing human problems in a way that engenders both rational comprehension and ethical awareness. In particular, we address the presentation of information related to human rights violations such as intimidation, abuse, violence and genocide. As with other urgent issues facing us today—climate change, environmental devastation, polarization of political cultures—human rights violations arise from complex origins and therefore require sophisticated analysis. The causes of rights violations may not be immediately visible. Some may seem innocuous at first sight but acquire greater urgency when amassed on a wide scale and subjected to descriptive and inferential statistical analysis. At

the same time, communities and governments are facing increasingly urgent pressure to respond to the evidence of these violations in a swift and effective manner, a response that demands that we organize the knowledge in the most informative and revealing manner possible, and analyze it at a high level of numeracy as well as empathy. This combination of urgency and complexity subjects these violations to a paradox identified by Scott and Paul Slovic in their 2015 essay collection, *Numbers and Nerves*: while we need “the best available statistic information ... to assess various situations and mitigate problems ... numerical information often fails to register on audiences an effect forceful enough to lead to significant action” (2015, 1). The tension between these two types of reasoning becomes urgent in situations where governments, organizations and citizens must confront evidence of human rights violations on a massive scale and respond justly and effectively.

Knowledge organization is intimately concerned with the ethical implications of providing access to information, a concern which manifests itself in both theory and practice. Sanford Berman's activism through the decades has highlighted numerous problems with the Library of Congress Subject Headings, while Hope Olson's work has greatly enhanced the sophistication of both the critique and the renovation of knowledge organization tools and practices. More recently, theorists and practitioners have built on such foundations to integrate critical theory into standard practices of analyzing classification systems (Martínez-Avila et al. 2016), to raise awareness of newer concerns such as intersectionality (Fox 2016). Leazer and Montoya (2020) identify the task of recognizing "unjust and politically oppressive practice in KO" as an urgent responsibility (370).

Throughout these and other efforts to engage with the ethics of KO, however, we can detect a subtle tension between what our tools say and what they do. Smiraglia (2002), citing Svenonius (2000), argues that knowledge organization "is accomplished through a bibliographic language or, more properly through a complex set of bibliographic languages), with semantics, syntax, pragmatics, and rules to govern their implementation" (330). As with all languages, we can assess how well KO languages convey communicate necessary concepts; we can also assess what principles, assumptions and implications these languages present while communicating. Many critiques of KO tools justify themselves through the pragmatic needs of access and retrieval. Olson's critique of biases such as those of gender, sexuality, race and age rest partly on the prosaic consequences: "Library users seeking material on topics outside of a traditional mainstream will meet with frustration in finding nothing, or they will find something but miss important relevant materials" (2001, 639). And Berman (2016), rejects what he calls "nanny cataloging"—using subject headings as trigger warnings—as "patronizing, arrogant and censorious." Rather, headings should be justified in terms of the work they're intended to do: "to fairly and fully reflect what the material is about, what's in it, and what perspective it conveys" (3).

Implicit in this task of reflection, however, is the expectation that we assign materials appropriate names. In his request to have the LC heading **Armenian massacres, 1915-1923** changed to **Armenian genocide, 1915-1923**, Berman's objection to the term "massacres" rests only partially on the efficiency of retrieval. The choice of "massacres" as the authorized term makes an implied statement that is inaccurate: "the primary heading is wrong, is correction apparently being held hostage to both passivity and political deference" (2016, 3). Similarly, Fox (2016) demonstrates how the mechanics of classification systems, even when they work effectively, result in a form of erasure in cases of intersectionality, in which citation order forces us to assign

primacy to one facet of a subject at the expense of others (378).

The following piece attempts to separate these two aspects of knowledge organization—what our tools do, and what they say by doing it—by exploring practices, not of naming for retrieval, but simply naming for its own sake. Such a practice ironically forces us to invoke some of the most prosaic of all knowledge organization: name authority control. We argue that the simple act of naming, for its own sake, and for the rhetorical visual effects of seeing names listed, play an important role in bridging the two types of reasoning identified by the Slovics.

2.0 Background: Levinas, Gadamer and the "fusion of horizons"

In this paper, we proceed from a core assumption: that a just response, whether individual or collective, requires us to strike a middle ground between indifference on the one hand and despair on the other. Instead, we must engage what tools we have—in this case, knowledge organization—to foster a meaningful connection between ourselves and the others whose suffering we witness. Another term for this connection might be "involvement," as the poet John Donne termed it when he claimed, "Each man's death diminishes me, for I am involved with mankind" (2007). Inspired by the writings of Emmanuel Levinas and Hans-Georg Gadamer, we argue that a just and effective response to evidence of human rights violations requires movement toward a truly ethical relationship between "I" and "Thou." I and Thou are Gadamer's terms for myself and the other as equal selves. Levinas articulates the challenge that knowledge organization faces in working toward this ethical goal, while Gadamer expresses one way for us to respond to that challenge.

Levinas takes an uncompromising and demanding position: that the core of an I-Thou relationship is a mutual physical presence: the face of Thou, which is "inviolable" and Thou's eyes, "which are absolutely without protection, the most naked part of the human body, [yet] none the less offer an absolute resistance to possession in which the temptation to murder is inscribed" (Levinas 1990, 8). If relationships about things are expressed in language, then the face speaks before any word is spoken. In other words, every face-to-face encounter forces a dialogue that is meaningful, even if no words pass either I's or Thou's lips. The interaction is not necessarily verbal: but it is direct and powerful and insistent.

Levinas offers another way to consider the effect of the physical presence that "Thou" has on an encounter: When "I" speak to "Thou" in a face-to-face encounter, "One does not question oneself concerning him; one questions him" (Levinas 2011, 47). "Thou" cannot be the subject of the discussion in which "Thou" participates. Dialogue between two people, as simple as it may seem to us, forces "I" to en-

counter all the ways in which “Thou” is unique. The face-to-face encounter makes “Thou” fully present. As Levinas’ evocation of the eye indicates, the full presence of “Thou” may actually be uncomfortable because “I” cannot control “Thou” as “I” perhaps could as an object of discussion. “I” clearly see things in “Thou”’s eyes that “I” cannot explain.

While face-to-face encounters are frequently impossible, the implications of Levinas’s insistence on physical presence are inescapable: we cannot repair the ruptures of violence and injustice without some direct exposure to the other that transforms the other from a topic of discussion into a real and distinct companion who is entitled to our full ethical consideration. Our challenge in knowledge organization is to find ways to cultivate an ethical relationship with “Thou” even without direct exposure.

Gadamer suggests a different way to consider I’s relationship with Thou. Like Levinas, Gadamer believes that I experience Thou too often as an object of inquiry. Gadamer offers two other options. In the first, I experience Thou as something more than that rationalized object, yet I do not escape the limitations of my point of view as I. Here, Thou is only Thou by virtue of I. However, the second option is characterized by openness and we “experience the Thou truly as a Thou, i.e., not to overlook his claim but to let him really say something to us” (Gadamer 2006, 355). If a point of view is prioritized here, it is that of Thou, who speaks while I listen. However, Thou also listens because “Belonging together always also means being able to listen to one another” (355) Here, Gadamer is referring to his concept of the fusion of horizons, in which I encounter something new that questions I, leading I to ask new questions. Gadamer’s “horizon” refers to a person’s historical-cultural context for a phenomenon, especially as experienced through language. As I encounter a different horizon—something that is new or strange, and that challenges my situation—I engage in a process of dialogue. What is familiar and what is unfamiliar impact each other, fusing together into a new horizon. This process is ongoing so that understanding is never static.

The difference between Levinas and Gadamer is subtle but significant: Levinas focuses on I’s inability to reduce Thou, who can never be fully comprehended, into something familiar. Gadamer shifts our attention from Thou to Thou’s claim, which can be comprehended through the familiarity of a shared language that emerges from dialogue. This shift in thinking emphasizes the limits of knowledge organization while, at the same time, helping us to avoid the “numbing” effect observed by the Slovics. Knowledge organization, on its own, will never foster Levinasian responsibility to Thou. It can, however, harness data in such a way that assists a decision-maker in understanding a cry for justice in a way that fosters both rational understanding and emotional empathy in a productive way.

In the sections that follow, we will first examine how both cognitive psychology and information practice align with the distinction between “numbers” and “nerves.” We will then move to knowledge organization and how our established practices of creating and manipulating structured data could potentially facilitate a sense of relationship and involvement.

3.0 “Numbers and nerves” dichotomy

Since the early 1990s, researchers in psychology have identified two distinct types of cognition: a rapid and intuitive response (Type 1) and a slower and more rational response (Type 2) (Epstein 1994; Kahneman 2011). This phenomenon maps partially to the problem identified by the Slovics. First, in order to grasp the numerical and logical complexity of human rights violations, we need Type 2 reasoning, involving considerable powers of numeracy, a sound grasp of logic and paradox, and facility with large numbers, together with a skeptical scrutiny of information sources. Second, we need to grasp the situation without our Type 1 responses—empathy stimulated by individual narratives—becoming numbed and paralyzed by the bewilderment caused by large numbers (Slovic 2007; Newcomb 2010). Just responses to ethical violations requires, not just a command of the facts, but a capacity for imagination (Kaul and Kim 2015): what Gadamer would call a “fusion of horizons” (2004), and the Slovics would call an “emotional and informational equilibrium” (Slovic and Slovic 2015, 8).

It is important, however, to realize that the Slovics’ distinction does not align precisely with Kahneman’s distinction between Type 1 and Type 2 thinking. For Kahneman, Type 2 thinking does not entail indifference or absence of empathy; rather, it works to ensure that our empathy be grounded in a rational, accurate and comprehensive understanding. Similarly, Type 1 thinking involves reasoning, but reasoning based on rapid heuristic assumptions; the emotional responses such reasoning inspires, therefore, may or may not be accurate or justified. Numbers can appear in both Type 1 and Type 2 reasoning, and to be numerate does not simply mean that one is skilled with numbers or happy working with them. Numeracy also entails placing numbers into a meaningful context: a process that involves making their implications clear and anticipating, to some degree, the Type 1 response that the numbers might inspire. *Harper’s Magazine* illustrates this through its monthly Harper’s Index, which presents numerical data in an artfully-arranged sequence that appeals to a variety of emotions, including whimsy, humor, surprise and anger. Take, for instance, this excerpt from the March 2022 Index:

[Percentage increase] in the number of bird watchers in the United States: 16

In the number of pickleball players: 39
 Percentage by which men are more likely than women to sexually harass a boss of the opposite sex: 10
 By which men are more likely than women to say they use sex appeal to get what they want in the workplace: 10 (Harper's Magazine, March 2022, 9)

The Index arranges its data to take us from the idle—increases in bird watchers and pickleball players—to the serious—harassment in the workplace—to the paradoxical: contrary to prevailing assumptions, men are more likely than women to rely on sex appeal at work. Numbers can be used for rhetorical purposes just as words can.

When presenting data in an ethical fashion, then, we need to consider at least three possible adverse effects:

- Small effects: in which numbers are presented devoid of any content that would trigger empathy or ethical reflection;
- False effects: in which the numbers are framed, intentionally or unintentionally, in a false narrative, thereby inspiring an inappropriate emotional response. This is where we typically find instances of misinformation and fake news.
- Overwhelming effects: in which the size, intensity or scale of the data numbs our responses and causes them to close down, often with a sense of fear or futility.

4.0 Numerical data and information visualization

The field of information studies has already made important contributions to the problem of rendering complex data in meaningful ways. Paul Otlet, in his concept of the Mundaneum, applied architectural metaphors to his concept of knowledge organization as a means “to order human actions and to create a better society” (van den Heuvel 2008, 129). Neurath’s ISOTYPE project sought to develop an iconic form of communication that evaded “roundabout” problems of alphabetic literacy in favour of “immediate visual presentations,” particularly in the expression of statistical phenomena” (Hartmann 2008, 284). Information visualization, particularly through the work of Donald Norman, has shown how the subtle and perceptive presentation of data can appeal to both experiential and reflective cognition (1993). Such research and experimentation anticipated a conclusion reached by Clavier (2019) that the visual representation of data has passed over from being a simple tool to a new, powerful and increasingly prevalent language of its own:

Visual representations are now the predominant modes of expression in our society... [It] is not a question of demonstrating that we live in “a world of im-

ages” . . . but rather that the importance assumed by visual representations—or ‘visuals’ as they are referred to by communication specialists—as forms of semiotization influences the ways in which knowledge is represented and organized. (617-8)

Effective visualization has been used in human rights situations to indicate areas of concentration, such as maps of the United States indicating frequency of racial crimes (Hepworth and Church 2019). They have been used to indicate sharp progressions, such as bar charts indicating the increasing frequency of hate crimes (Rainbow Railroad 2021). Many organizations, such as Rainbow Railroad, employ strategic combinations of quantitative data and stark images that bring the figures to imaginative life. But information visualization rests upon reliable sources of data. And this is where knowledge organization comes in: through an understanding and exploitation of the structure whereby the data is gathered.

5.0 Structured data and knowledge organization

Structured data is data that has been rendered in a standard format, so that all of the data elements are consistent with one another and are accompanied by semantic metadata which identifies the constituent parts of the data and how they relate to each other (Crockett 2023). It therefore presupposes a very different approach to data than that of “reading” unstructured text such as “normal” prose. The literary critic Northrop Frye argued that when reading a text, we experience it two senses: the sequential process of reading the text word by word and page by page, and the subsequent comprehension of the entire structure of the text in all its parts and facets (Frye 1990, 1). Structured data reverses this process, to some degree: the structure comes first, and is subsequently “filled in” with data, to be followed by some kind of narrative interpretation. Structured data, therefore, begins with the overall design: of a bibliographic record, for instance, or a set of tables in a relational database. The data elements are determined from the outset, along with the parameters of those elements: the data types and ranges and relationships. The data elements are surrounded by metadata which identifies them, enabling them to be filtered and combined in queries, sorted in various ways, and displayed in reports.

We can see this reversal at work in the research of Silva et al. (2021) at the Universidade Federal de Paraná, who have extracted expressions of hate speech in Brazilian online communities and established what they call “uma ontologia do discurso de ódio.” The sequential text of messages that display intimidation of marginalized communities has been converted into a structured representation of hateful sentiments (see Figure 1).

Tipo de discurso de ódio	Palavras-chave e expressões de ódio recuperadas
Politico/partidário	"petralha" / "petralha safada" / "coxinha" / "coxinha burro" / "comunista" / "comunista safado" / "fascista" / "coxinha fascista" / "político ladrão" / "compra votos" / "petista vagabundo" / "elite golpista" / "esquerda caviar" / "esquerdopatas" / "esquerdosos" / "vermelhos pervertidos" / "esquerdistas" / "maconheiro libertário" / "Socialista Psicopata" / "vai pra Cuba" / "Volta pra Cuba"
Sexismo	"vadia" / "safada" / "mal comida" / "coisa de mulherzinha" / "falta de rola" / "falta de pica" / "cara de puta" / "odeio mulher" / "feminazi" / "tinha que ser mulher" / "chifrado" / "vaca ordinária" / "quenga" / "biscate" / "pirigüete" / "perigüete" / "piranha"
LGBTQIfobia	"boiola" / "baitola" / "cara de traveco" / "voz de traveco" / "queima rosca" / "meio afeminado" / "coisa de boiola" / "parece uma bixa" / "jeitinho de gay" / "jeito de gay" / "bichona" / "bixona" / "pau no toba" / "homossexual nojento" / "pervertidos" / "viadagem" / "vira homem"
Racismo	"não sou tuas nega" / "tinha que ser preto" / "da cor do pecado" / "preto é foda" / "negrão é foda" / "cara de macaco" / "esse macaco" / "preto safado" / "negro fedido" / "negrão fedido" / "macaco" / "bugre" / "índio preguiçoso" / "índio vagabundo" / "indiarada"
Aporofobia	"Bolsa esmola" / "pobrada" / "bandido favelado" / "favelado" / "filho de papai" / "coisa de pobre" / "mendigo fedido" / "trabalhar vagabundo" / "pobre ignorante" / "fracassado" / "pobretão" / "barriga suja"
Xenofobia	"povo burro" / "nordestino vagabundo" / "muçulmano bomba" / "homem bomba" / "balanice" / "balanada" / "tudo terrorista" / "Volta pra sua terra" / "caíçara folgado" / "caipira burro" / "povo da roça" / "bando de jumento" / "invasão haitiana" / "Volta para África" / "povinho"
Deficiência	"retardado mental" / "tem down" / "aleijado" / "aleijado" / "um demente" / "uma demente" / "leproso" / "aidético" / "coisa de retardado" / "deficiente mental" / "é autista" / "parece cego" / "débil mental" / "abestado" / "mosca morta" / "mongolóide" / "imbecil"
Etariedade	"Velho burro" / "velho babão" / "velho nojento" / "velho safado" / "velho tarado" / "coroa fogosa" / "velho pra isso" / "velha pra isso"

Figure 1: Uma Ontologia do Discurso de Ódio (Silva et al. 2021).

The schematic representation reflects the results of extensive and careful reasoning: establishing and naming the categories, formulating the criteria for category membership, and displaying the data in an effective visual manner. But seeing the hatred laid out in rows is an experience: an emotional one. And Silva and Botelho-Francisco (2018) hope that the emotions inspired by this careful work will lead, not to indifference or despair, but to an acknowledgement on the part of all who see it that such expressions demean us all, and to "assertive strategies to maintain tolerance in online environments (28)."

Structured data is useful in modern information systems, because it makes the data machine-readable and machine-understandable, enabling sophisticated semantic combinations. But the history of structured data goes back before computers: libraries, for example, have historically used of standard structures for bibliographic records, as well as standards for authority records for entities such as names, places and subjects. Structured data in computerized form appeared with the MARC format in the 1960s, as well as through the development of relational database design. It has since been extended into standards for semantic markup such as SGML and XML, and into the standards for Linked Data.

Structured data has also extended its reach beyond mainstream information institutions such as libraries, suggesting a widespread social consensus that structured data is worth

creating. Online communities that maintain resources such as Wikipedia have embraced the crowdsourced conversion of unstructured data, in the form of free text, into structured form. Wikipedia has established a series of what it calls "Infoboxes" that make it easy for those creating a Wikipedia page to create structured data (see Figure 2).

The various boxes established so far are classified into various types and subjects, such as Arts and culture, History and events, Person, Place, Religion and belief, Society and social science, and Technology and applied science. Among the various templates available, we have a template for "Event," which provides a list of fields, along with guidance on how to fill them out. Such fields include: title, duration, date, location, venue, cause, perpetrator. Projects such as Wikidata enable the extraction of structured data elements from free-text resources such as Wikipedia entries; the entry for the Armenian Genocide, for instance, contains structured data providing the date, location, target, attack type, perpetrators, number of deaths and trials (https://en.wikipedia.org/wiki/Armenian_genocide).

The contrast between the structured data and the unstructured prose is complex. Here, for instance, is the opening of the Wikipedia entry on the Armenian Genocide^[1]:

The **Armenian genocide** was the systematic destruction of the Armenian people and identity in the Ot-

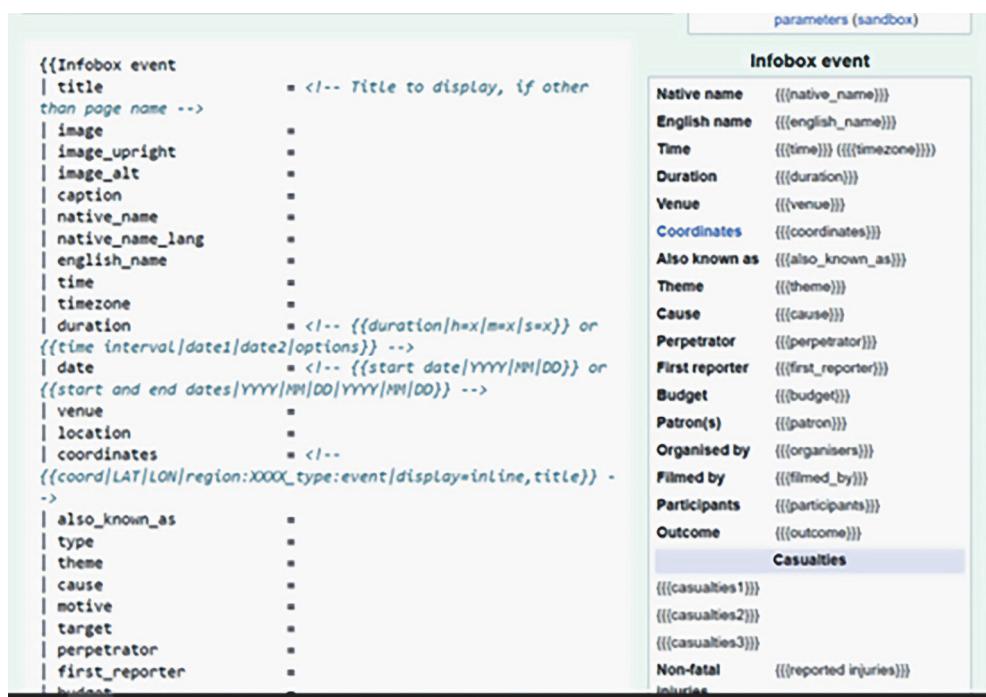


Figure 2: Wikipedia Infobox (https://en.wikipedia.org/wiki/Template:Infobox_event).

toman Empire during World War I. Spearheaded by the ruling Committee of Union and Progress (CUP), it was implemented primarily through the mass murder of around one million Armenians during death marches to the Syrian Desert and the forced Islamization of others, primarily women and children.

The entry itself gives us a narrative account of the genocide, seeded with hyperlinks to further information that provide background context and explanation for names, terms and events that might not be obvious to us a century later. In the infobox to the side of the screen, we find a set of attributes that have been defined by Wikipedia as appropriate for what they call an “Event.” Wikipedia provides documentation for this type of infobox, enumerating the parameters and recommended best practices for the listed attributes in a manner reminiscent of descriptive cataloging rules used in libraries, such as Resource Description and Access (RDA). One such recommendation is that we use this particular type of infobox for “one-off events,” rather than for recurring events, which go in their own infobox type, “recurring events” (https://en.wikipedia.org/wiki/Template:Infobox_event). This interpretation is consistent with the scope note of the concept “Event” as defined by the International Council for Museums and the International Committee for Documentation, which defines event as “an instantaneous change of state” (2018). The Armenian Genocide is thus summed up as follows:

Location:	Ottoman Empire
Date:	1915-1917
Target:	Ottoman Armenians
Attack Type:	Genocide, death march, forced Islamization
Deaths:	600,000-1.5 million
Perpetrators:	Committee of Union and Progress
Trials:	Ottoman Special Military Tribunal

The unstructured narrative tells us, in sequential prose, what happened. The infobox extracts from this narrative a list of facts that have a stark power similar to Silva and Botelho-Francisco's *ontologia do discurso de ódio*. Clearly, we live in a world where we need controlled vocabulary terms for genocide and death marches, and where death counts can land somewhere between 600,000 and 1.5 million.

The case of violinist Tyler Clementi offers a slightly different contrast between unstructured and structured descriptions^[2]:

Tyler Clementi (December 19, 1991– September 22, 2010) was an American student at Rutgers University–New Brunswick who jumped to his death from the George Washington Bridge over the Hudson River on September 22, 2010, at the age of 18. On September 19, 2010, Clementi's roommate, Dharun Ravi, used a webcam on his dorm room computer and his roommate Molly Wei's computer to view, without Clementi's knowledge, Clementi's intimate video chat with another man. Ravi and Wei recorded the video and distributed it to their social network. Clementi committed suicide in response to the harassment and humiliation he suffered as a result of the video.

enti's knowledge, Clementi kissing another man. Clementi found out after Ravi posted about the webcam incident on Twitter.

Out of this, the following attribute-value pairs are derived:

Deaths:	Tyler Clementi, aged 18
Location:	Manhattan, New York City, U.S.
Type:	Teenage suicide, suicide by jumping, drowning
Motive:	Homophobia
Accused:	Dharun Ravi
	Molly Wei
Convicted:	Dharun Ravi

In this instance, the unstructured narrative of Clementi's death differs from the narrative of the Armenian Genocide in its use of hyperlinks. While the hyperlinks for the Genocide provided relevant background information for contextualizing and understanding the background of the event, here the hyperlinks, with one important exception, are irrelevant. They link to the Wikipedia pages which give bland information about the school, the bridge, the river, webcam technology and Twitter. The exception is the link "jumped to his death," which takes the reader to a structured list of "lesbian, gay, bisexual and transgender people whose suicides were deemed sufficiently notable to be reported by the media" (https://en.wikipedia.org/wiki/List_of_LGBT-related_suicides). At some point in this article's life, someone felt that the story of Tyler Clementi needed to be linked to a structured list: that the tragedy of the incident lies not just in what happened to Clementi but the fact that it has happened to many, many people.

By contrast, the structured data of the infobox, while confining itself to the bald facts of the narrative, shows a revealing instance of the cataloguer's judgement. Whoever entered the value for the attribute "deaths" elected to include, not just Clementi's name but his age. Whether Clementi's age makes a difference is a matter for debate: one could argue that a life destroyed is a life destroyed, whenever it happens. But someone clearly decided that the data should insist on the fact that the person who killed himself in the wake of homophobia was not yet 20 years old. As we saw with Harper's Index, the strategic placement and arrangement of structured data can create a stark and brutal eloquence.

6.0 Structured data and meaningful numerical accumulation

In both uses, structured data serves as an intermediary step between the intimate narrative that appeals to our emotions and the quantified syntheses that engage our reason. This

paper suggests that the presence of structured data as this middle ground leads to four possible effects upon our understanding of crimes against humanity.

6.1 Impact through numbers

First, in terms of human rights violations, structured data has one long-standing and obvious use: the capacity to combine and assemble existing collections of data for quantitative analysis. Only through the accurate and effective combination of data from various sources can we have access to the information we need to understand and grapple with human rights violations. Structured data facilitates such amalgamation, allowing for the representation of cumulative effects that might not otherwise be relevant. Furthermore, the use of standard formats and metadata labels renders the principles of accumulation more evident and subject to analysis, reducing the chance of erroneous or fraudulent data displays.

Structured data facilitates not only the effective aggregation but the effective presentation of numerical data, and while the Slovics (2015) are correct to point out the complexities of presentation and dangers of over-simplification, the 2021 Annual Report of Rainbow Railroad reminds us that numbers matter on a visceral level. The organization, which mobilizes support for LGBTQ+ people throughout the world facing persecution, makes effective use of attribute/value pairs (see Figure 3).

In this display, the stark numbers are linked to stark and vivid attributes: "family-based rejection," "sexual violence," "family-based violence," "beat up," "shot at," "chopped," "threat" and "in hiding due to credible threat." Elsewhere in the report, a similar display links numbers to health and welfare concerns: "risk of suicide," "lack of basic needs," "lack of income due to COVID-19," "homeless" and "mental health concerns." The report also shows that 503 individuals were supported in 2020 and 1,812 in 2021.

Numbers are frequently used for rhetorical purposes: in 2015, for instance, Canadian Prime Minister Justin Trudeau proudly presented Canada's first federal Cabinet that was 50% female (Murphy 2015). In the case of human rights violations, however, the numbers signify an important facet of a burgeoning I-Thou relationship. On the one hand, knowing that 39% of the LGBTQ+ individuals facing safety concerns are afraid of being beat up, shot at or chopped provides a forcible reminder to a complacent "I" that there are people living far more precarious lives. On the other hand, data categories such as suicide, lack of basic needs, COVID-19 unemployment and homelessness allude to frightening possibilities for all of us.

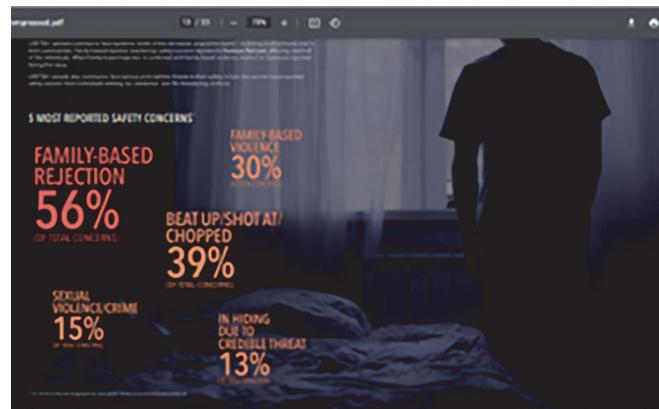


Figure 3: Rainbow Railroad Annual Report 2021.

6.2 Impact through exploration

Second, even without such amalgamation, structured data facilitates exploration and discovery, by enabling an isolated, traumatic incident to lead, through structured links, to documentation of other similar incidents; the Wikipedia entry for the murder of George Floyd, for instance, leads to a structured list of incidents of police brutality ordered by date and another by country. The Wikipedia entry on “Violence against women” provides a subdivided list of violence types: killing, sexual assault and rape, disfigurement and other issues. The topics included under these main headings include bride burning, dowry death, forced prostitution, campus sexual assault, cybersex trafficking, acid attack, gas-lighting and narcissistic abuse (https://en.wikipedia.org/wiki/Violence_against_women). While the sheer number of headings for violence—28 at the time of writing—possesses its own sombre eloquence, the strength of connection also lies not just in the overall number but in the detail and variety of these headings, each of which provide access to those seeking understanding and support. Sometimes, the trauma of one story needs to be balanced, not necessarily by a quantitative summation, but by access to more stories.

6.3 Impact through consistent display

Third, structured data lends itself to regularity and consistency of display, thereby enabling us to enact the process of retrieving and amassing cumulative effects that appeal to rational comprehension, but on a more direct emotional level. In this sense, structured data long predates computer technology; it is a key element of memorials, in which the repetition of a familiar structure heightens the rhetorical effect. The traditional practice of listing names on cenotaphs has been adapted in many imaginative ways, all using some regularity of entry and structure to heighten the emotional impact. The AIDS Memorial in Toronto presents repeated,

regular entries of the names of those who died. In the United States, the National AIDS Memorial contains a digital representation of each panel of the AIDS Quilt, the massive tribute to those who died in the AIDS Epidemic: a resource that was crowd-sourced before we had even invented the word. Each panel represents the individuality of the person or persons who died, resulting in a dizzying variety of colors, designs and motifs. But the names are contained in a structured database, normalized and rendered consistent, enabling us to locate the panel of any individual through the Web interface (see Figure 4).

In a variation on the same method, Brazilian artist Ana Flores has created an exhibit similar in spirit to that of Silva’s and Botelho-Francisco’s ontology: a relentless line of men’s dress shirts, each adorned with a quotation of hate speech from social media (Zavadil 2022). Isolated words and phrases are assembled in regular and consistent order, turning them from passing utterances into a sinister collective expression of hatred and intolerance.

Structured data also extends beyond words. The discovery of unmarked graves of Indigenous children who died in Canadian residential schools has led to multiple displays of children’s shoes, and the many missing Indigenous women and girls in Canada have been commemorated by installations of dresses hanging, unworn, on the boughs of trees. The absence of words reflects the all too frequent absence of names, the inadequate archival records and above all the disruption of families in order “to prevent the transmission of cultural values and identity from one generation to the next” (Truth and Reconciliation Commission 2015, 1). Canadians are slowly and reluctantly learning to accept this namelessness as cultural genocide.

In all of these cases, we can witness the “gathering” effect made possible when data has been cleaned, normalized and structured, thereby enabling us to gain a more accurate sense of the spread, frequency and seriousness of human rights violations. In such cases, however, the final number

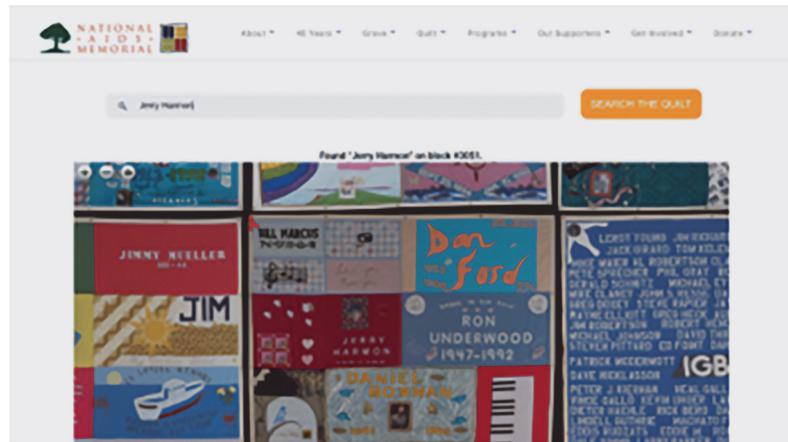


Figure 4: The AIDS Quilt Interface (<https://www.aidsmemorial.org/interactive-aids-quilt>).

is less important than the process of experiencing each data element in sequence. Frequently, such displays provide not numbers but names, words, shapes and artifacts, one following the other with relentless regularity. Such regularity preserves the memory of those who would otherwise be lost, much as the Holocaust survivors in Norman Lebrecht's *The Song of Names* preserve the names of lost comrades through song. The regular pattern of delivery also reminds us that the tragic events commemorated are not exceptions; they are regularities. Not only did they happen; they are happening again, again, and still again. We may not be able to create the mutual physical presence that Levinas advocates; we cannot confront, face-to-face, the "Thou" of those who suffer, or who have suffered violations of their safety, their rights and sometimes their very lives. But structured data, in the form of names, symbols, songs and artifacts, can insist that we recognize their existence.

6.4 Impact through dispersion

Finally, structured data, in its regularity, works in the reverse direction as well. We are all trained to recognize standard and familiar information structures within our environment: red, yellow and green traffic lights, for instance, or signs that direct us to washrooms and elevators. Some memorial projects are facilitating recognition of injustices by seeding memorials in widely dispersed locations, relying on discovery and chance encounters to surprise us into awareness, rather than through memorial installations that people approach purposefully (Helphand 2015). One striking example is the distribution of “sopersteins,” or “stumbling stones” throughout Europe: gold-coloured plates fitted into paving stones in front of houses where victims of the Holocaust once resided. In the words of Barbara Steiner, a

resident of Berlin, these small plates are more meaningful than large memorials: “I think the huge monuments are always about performing memory, when this is really connected to a person” (Smith 2022, 27). Artist Gunter Demnig, who originated the project, argues “that individual stories are more powerful tools than statistics for coming to terms with this history” (Smith 2022, 27) (see Figure 5).

Structured data is fundamental to the *solperstein* project. The plates present their data with the regularity of an old-style bibliographic record, providing the same elements in the same order. One of the *solpersteins* in Figure 5 reads:

HIER WOONDE
KLARA ELZAS-WALVISCH
GEB. 1911
GEDEPORTEERD 1943 UIT
WESTERBORK
VERMOORD 21-5-1943
SOBIBOR

In this plate, the line between structured data and linear narrative begins to fade. We have structured attributes: name, birth date, year of deportation, name of the transit camp, date of murder, and location. But ordered as the elements are, a narrative emerges. The pedestrian who stops to read the plate learns that in the house nearby lived a woman named Klara Elzas-Walvisch, born in 1911. She was deported to the Westerbork transit camp in 1943, and was murdered on the twenty-first of May, 1943, in the Sobibor concentration camp. The adjoining plate indicates that her husband was arrested and jailed a year earlier in 1942 and deported to the Mauthausen camp, where he died on the eighteenth of December in that same year.



Figure 5: Stolpersteine Holocaust Memorials Amsterdam (<https://www.traveldarkly.com/stolpersteine-holocaust-memorials-amsterdam/>).

The stones have a clear and simple structure, and the consistency of the structured data enables the pedestrian to see and recognize the message easily. The clarity and simplicity of the message facilitate rapid comprehension of a very simple narrative. The isolation of the paving stone from others stimulates an initial response to an individual story rather than a collective phenomenon. And the insidious frequency with which the pedestrian encounters these stones throughout Amsterdam, throughout the Netherlands, and throughout all of Europe in a multitude of languages, brings on an incremental awareness of the scope of the Holocaust, while presumably delaying and forestalling the numbing effect brought about by large numbers.

Perhaps this dispersion brings us closest to Gadamer's notion of the fusion of horizons, in circumstances where physical confrontation and dialogic interchange are impossible. The brevity and regularity of these small, repeated records of individual lives might surprise pedestrians past their own points of view into that state of openness that Gadamer argues we need for that fusion of horizons that constitutes a just response to human rights violations.

7.0 The role of knowledge organization: what must we do?

First, we must break down the “numbers” and “nerves” distinction to some degree. Feeling and reason are not diametrically opposed in the cases previously cited. Numbers are full of nerves, and nerves are full of numbers. There is deep and powerful emotion to be derived from slow, careful reasoning of any kind, numerical or nominal, empirical or phenomenological. And we find reason and calculation at work

in the most immediate and impulsive of responses. A more helpful distinction might be, drawing on Kahneman, between ethical responses that have a secure foundation in slow thinking and those derived from rapid thinking. Our ethical responses to injustice of any kind must be founded upon thoughtful, realistic awareness: the kind that comes through patient exercise of the widest possible range of inquiries: empirical, phenomenological, philosophical, aesthetic.

In information studies, we have an opportunity to make as much of this slow and patient work available to any individual or organization seeking to respond to evidence of human rights violations. As we write this, the Manitoba government in Canada must decide whether to instigate a lengthy, costly and possibly hazardous search of a landfill site for the bodies of two murdered Indigenous women (Gowriluk 2023). A pedestrian in Europe may encounter the micro-narratives of person after person murdered in the Holocaust and wonder, as those gold plates keep appearing, how to respond. We must do what we can to help them respond justly and appropriately.

As knowledge organization theorists and practitioners, specifically, two options present themselves. First, as creators and curators of large authority files, both for subjects and for names, we have an existing infrastructure of structured data that can be harvested for projects such as the *solperstein* initiative. Projects that expand the reach of our authority records, as well as the depth and accuracy of their contents and their cross-references, could make a difference. We have procedures and affordances for establishing names in multiple languages, for synonym control, for embedding background information and documenting the sources of

that information. To quote Hope Olson (2001), we who work in knowledge organization possess a certain power to name. Let us use that power carefully, responsibly and courageously.

Second, the wealth of affordances in resources like Wikipedia and Wikidata for creating structured data and links within that data attest to the need for knowledge organization in crowd-sourced projects. We certainly need more projects like that of Silva and Botelho-Francisco, who impose ontological concepts onto the seemingly formless nature of social media, in order to derive a truer and more persuasive picture of destructive online trends and practices. We also need to help non-experts who are reaching in the same direction. Whoever linked Tyler Clementi's story on Wikipedia to a structured list of others who shared his fate was surely trying to do what we've been trained to do: to create a bridge of structured data through a seemingly endless and endlessly confusing information environment. Let us help when and where we can.

8.0 Conclusion

In her novel *Daniel Deronda*, British author George Eliot describes a crucial moment in the life of her heroine, Gwendolyn Harleth^[3]:

There comes a terrible moment to many souls when the great movements of the world, the larger destinies of mankind, which have lain aloof in newspapers and other neglected reading, enter like an earthquake into their own lives. ... [Gwendolyn] was for the first time feeling the pressure of a vast mysterious movement, for the first time being dislodged from her supremacy in her own world, and getting a sense that her horizon was but a dipping onward of an existence with which her own was revolving.

Many of us, as individuals, as organizations, as races, as cultures are finding ourselves face to face, not necessarily with an individual, as Levinas urged, but with the realization that our individual horizons cannot make sense of what we're witnessing. Deciding to search a landfill for missing bodies is not simply a matter of provincial and municipal finance; the scenic pavements of Berlin and Amsterdam have more to tell us than we're prepared to hear on our holiday. As with Gwendolyn, our horizons are dipping: or, as Gadamer would put it, they are fusing with those of others.

The field of knowledge organization possesses theoretical and practical tools to facilitate and mediate those painful passages of realization within individuals, within governments, within populations. Our facility with structured data, whether as a means of producing accurate statistics, or of amassing or dispersing names, can be a key tool at our

disposal. However we choose to curate our own collections of structured data, and however we choose to participate in broader projects beyond our institutional walls, we must strive to enable meaningful fusions that lead us all between the extremes of despair and indifference.

Endnotes

1. Wikipedia, s.v. "Armenian Genocide", last modified November 14, 2023, 01:56, https://en.wikipedia.org/wiki/Armenian_genocide
2. Wikipedia, s.v. "Suicide of Tyler Clementi", last modified November 10, 2023, 08:43. https://en.wikipedia.org/wiki/Suicide_of_Tyler_Clementi
3. Eliot, George. 2005. *Daniel Deronda*. Project Gutenberg. <https://www.gutenberg.org/ebooks/7469>.

References

Berman, Sanford. 2016. "Forum". *Journal of Information Ethics* 25, no.1: 3

Berman, Sanford. 2013. *Not in my Library!: Berman's Bag Columns From The Unabashed Librarian, 2000-2013*. Jefferson, NC: McFarland & Company.

Clavier, Viviane. 2019. Knowledge Organization, Data And Algorithms: "The New Era of Visual Representations." *Knowledge Organization* 46, no.8: 615–21. <https://doi.org/10.5771/0943-7444-2019-8-615>

Crockett, Emma. 2003. "Structured vs Unstructured Data: Key Differences Explained." *Datamation*, February 9, 2003. <https://www.datamation.com/big-data/structured-vs-unstructured-data/>.

Donne, John. 2007. *Devotions upon Emergent Occasions, Together with Death's Duel*. Project Gutenberg. <https://www.gutenberg.org/ebooks/23772>.

Epstein, Seymour. 1994. "Integration of the Cognitive and the Psychodynamic Unconscious." *American Psychologist* 49, no. 8: 709–24. <https://doi.org/10.1037/0003-066X.49.8.709>

Fox, Melodie. 2016. "'Priorities of Arrangement' a 'Hierarchy of Oppression?'": Perspectives on Intersectionality in Knowledge Organization. *Knowledge Organization* 43, no.5, 373–83. <https://doi.org/10.5771/0943-7444-2016-5-373>

Frye, Northrop. 1990. *Words with Power: Being a Second Study of "The Bible and Literature."* New York: Harcourt Brace Jovanovich.

Gadamer, Hans-Georg. 2004. *Truth and Method*. 2nd, rev. ed. Translation revised by Joel Weinsheimer and Donald G. Marshall. London: Continuum.

Gowriluk, Caitlyn. 2023. "Landfill Search for Women's Remains Could Begin in April, Assembly of Manitoba Chiefs Says." *CBC News*, February 10, 2023. <https://www.cbc.ca/news/canada/manitoba/landfill-search-women-remains-1.6480300>

www.cbc.ca/news/canada/manitoba/assembly-manitoba-chiefs-landfill-search-remains-1.6743810.

Harper's Magazine. 2022. "Harper's Index." Harper's Magazine, March 9, 2022.

Hartmann, Frank. 2008. Visualizing Social Facts: Otto Neurath's ISOTYPE Project. In *European modernism and the information society: Informing the present, understanding the past*, edited by Boyd Rayward, 279-93. London: Routledge.

Helphand, Kenneth. 2015. "From One to Too Many." In *Numbers and Nerves: Information, Emotion, and Meaning in a World of Data*, edited by Scott Slovic and Paul Slovic, 89-105. Oregon State University Press.

Hepworth, Katherine, and Christopher Church. 2019. "Racism in the Machine: Visualization Ethics in Digital Humanities Projects." *Digital Humanities Quarterly* 12, no. 4. <http://www.digitalhumanities.org/dhq/vol/12/4/000408/000408.html>.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. Toronto: Doubleday Canada.

Kaul, Susanne, and David Kim. 2015. *Imagining Human Rights*. Berlin, München, Boston: De Gruyter. <https://doi.org/10.1515/9783110376616>.

Leazer, Gregory H., and Robert Montoya. 2020. "The Politics of Knowledge Organization: Introduction to The Special Issue". *Knowledge Organization* 47, no.5:367-71. <https://doi.org/10.5771/0943-7444-2020-5-367>

Levinas, Emmanuel. 1990. *Difficult Freedom: Essay on Judaism*. Translated by Sean Hand. Baltimore: Johns Hopkins University Press.

Levinas, Emmanuel. 2011. *Totality and Infinity*. Tran. A Lingis. Pittsburgh: Duquesne University Press.

Martínez-Ávila, D., Rafael Semidão, and Marcio Ferreira. 2016. "Methodological Aspects of Critical Theory in Knowledge Organization. *Knowledge Organization* 43, no.2: 118-25.<https://doi.org/10.5771/0943-7444-2016-2-118>

Murphy, Jessica. 2015 "Trudeau Gives Canada First Cabinet with Equal Number of Men and Women." The Guardian, November 4, 2015. <https://www.theguardian.com/world/2015/nov/04/canada-cabinet-gender-diversity-justin-trudeau>.

Newcomb, Matthew J. 2010. "Feeling the Vulgarity of Numbers: The Rwandan Genocide and the Classroom as a Site of Response to Suffering." *JAC* 30, no. 1/2: 175-213. <https://www.jstor.org/stable/20866942>

Norman, Donald A. 1993. *Things That Make Us Smart: Defending Human Attributes in the Age of the Machine*. Reading, Mass.: Addison-Wesley.

Olson, Hope. 2001. "The Power to Name: Representation in Library Catalogs". *Signs: Journal of Women in Culture and Society* 26, no. 3:639-68. <https://doi.org/10.1086/495624>

Rainbow Railroad. 2021. *Understanding the State of Global LGBTQI+ Persecution: 2021 Annual Report*. Rainbow Railroad. www.rainbowrailroad.org.

Silva, Luiz Rogério Lopes, Rodrigo Eduardo Botelho-Francisco, and Rafael Cardoso Sampaio. 2021. "Discurso de Ódio nas Redes Sociais Digitais: Tipos e Formas de Intolerância na Página Oficial de Jair Bolsonaro no Facebook." *Galáxia* 46, e-51831. <https://doi.org/10.1590/1982-2553202151831>

Silva, Luiz Rogério Lopes, and Rodrigo Eduardo Botelho-Francisco. 2018. "Da Representação ao Monitoramento: a Criação de uma Ontologia do Discurso de Ódio Online Brasileiro." *AtoZ: Novas Práticas em Informação e Conhecimento* 7, no. 2: 28-33. <https://doi.org/10.5380/atoz.v7i2.67243>

Slovic, Paul. 2007. "If I Look at the Mass I Will Never Act': Psychic Numbing and Genocide." *Judgment and Decision Making* 2, no. 2: 79-95. Cambridge University Press, <https://doi.org/10.1017/S1930297500000061>

Slovic, Scott, and Paul Slovic. 2015. "Introduction: The Psychophysics of Brithness and the Value of a Life." *Numbers and Nerves: Information, Emotion and Meaning in a World of Data*. Corvallis: Oregon State University Press: 1-25.

Smiraglia, Richard. P. 2002. "The Progress of Theory in Knowledge Organization". *Library Trends* 50, no.3: 330-49.

Smith, Clint. 2022, "Monuments to the Unthinkable." *The Atlantic*, November 14, 2022. <https://www.theatlantic.com/magazine/archive/2022/12/holocaust-remembrance-lessons-america/671893/>.

Truth and Reconciliation Commission of Canada. 2015. *Honouring the Truth, Reconciling for the Future: Summary of the Final Report of the Truth and Reconciliation Commission of Canada*. The Commission. https://ircs.ubc.ca/sites/default/files/inline-files/Executive_Summary_English_Web.pdf.

van den Heuvel, Charles. 2008. Building Society, Constructing Knowledge, Weaving the Web: Otlet's Visualizations of a Global Information Society and His Concept of a Universal Civilization. In *European modernism and the information society: Informing the present, understanding the past*, edited by Boyd Rayward, 127-53. London: Routledge.

Zavadil, Ana. 2022. *Fora Das Sombras – Novas Gerações Do Feminino Na Arte Contemporânea Do Rio Grande Do Sul*. Porto Alegre, Brasil. https://issuu.com/jessicajank/docs/cat_logo_fora_das_sombras_-vers_o_digital.