

# Applying a Critical Lens across the Research Data Life Cycle to Foster Greater Data Inclusivity: An LGBTQ+ Case Study

Berenica Vejvoda, Research Data Librarian, University of Windsor  
Academic Director, University of Windsor Research Data Centre  
PhD student (Social Work)

March 16, 2022



# LGBTQ+ Population (in Canada)

- Focus on sexual orientation
  - Sexual orientation largely used as the primary way to describe a person's sexuality
  - Complex construct
- Extension of this brilliant piece!
  - Gofman, A., Leif, S., Gunderman, H., & Exner, N. (2021). Do I Have To Be An “Other” To Be Myself? Exploring Gender Diversity In Taxonomy, Data Collection, And Through The Research Data Lifecycle. *Journal of EScience Librarianship*, 10. <https://doi.org/10.7191/jeslib.2021.1219>



# Starting Assumptions

- RDM situated in a social, political and cultural context
- Data not neutral
- Structural oppressions shape data → increases marginalization
  - Harm in reduced value of research data (inaccuracies and dropping from study)
  - Harm to respondents
- Data inclusivity reduces invisibility and discrimination against marginalized groups AND empowers
  - anti-LGBTQ+ laws, social stigma, violence



# Starting Assumptions

- Inclusivity affects all stages of the research data life cycle
  - Inaccurate, incomplete, irrelevant or misleading data
- Changes over time and status (shifts in sexual orientation, “sexual fluidity”)
- Critical lens needed to evaluate how LGBTQ+ data is used (critical data studies, critical theory and post-structuralism)



# Ethical Considerations

Principles (based on LGBT Foundation adapting ESCR, UKRI principles)

## 1. Integrity and Quality

**Are research questions worth asking / do they add to existing knowledge?**

- LGBTQ+ communities not always strongly represented in existing research
- Good opportunity when research is identified where LGBTQ+ people are barely represented
- Ensure that LGBTQ+ participants are central to research
- How new knowledge benefits the community
- Amplify voices of LGBTQ+ participants rather than speak for them



# Ethical Considerations

## 1. Integrity and Quality

**Do I have a good knowledge of the community I intend to work with?**

- Trans community research: how well you understand trans identities and issues?
- Familiarize yourself with full spectrum of gender identity, social and political issues affecting the lives of trans people
- Inclusive language



# Ethical Considerations

## 2. Fully Informing Researchers and Participants

**How informed and aware are the research staff?**

- Make sure that researchers are not ignorant/prejudiced towards LGBTQ+ people
  - Participant discomfort, compromises research, and harm
- Ensure researchers are knowledgeable about LGBTQ+ issues and the topic you are researching to adequately inform participants



# Ethical Considerations

## 3. Confidentiality and Anonymity

**How might my research data affect confidentiality and anonymity?**

- All topics involving LGBTQ+ participants should require **extra care to** preserve confidentiality and anonymity
- Regardless of topic, not all LGBTQ+ people will be comfortable with their identities being known, so responsibility to make certain information is not shared outside of the research team





# Ethical Considerations

## 3. Confidentiality and Anonymity

**How might my research methods affect confidentiality and anonymity?**

- Group-based methods —important to stress to participants not to disclose identities of anyone in attendance—this could ‘out’ participants without their permission
- Consider setting carefully to avoid ‘outing’ participants without permission (e.g., school or work)



# Ethics Approval

## 4. Avoiding Harm

**What sensitivities will my research include?**

- Discrimination can manifest in many ways and not be immediately obvious.
- What ways might discrimination be involved in the topic you are investigating?
- Is the language you are using LGBTQ+-inclusive?
- Try to make sure that you do not ask anything of your participants that could be considered offensive, disrespectful or invasive.



# Collecting Data

- Intentional – theory and research design
- Only collect IF required for research question
  - Minimally, what is necessary?
- Identifying lacuna (consequences of not collecting for other future research)



# Collecting Data

- Accurately represent sexual orientation in all its complexity
  - Gender is the de facto foundation for categorizing sexuality
  - Include gender?
  - Is it sexuality or gender?
  - How is sexual orientation defined if one is attracted to masculinity regardless of the sex of the person presenting or embodying it?
  - Two-spirited doesn't differentiate
  - Focus (social, health, economic)



# Collecting Data

- Multiple options (choose as many as apply)
  - Allows for more than one orientation
  - Generational differences
  - Some categories may be more stigmatized



# Collecting Data

- Open ended
- Avoid “other categories”
- Optional responding
- To enhance accurate responding and willingness, consider:
  - Mode (e.g., survey vs interview)
  - Interviewer training
  - Enhance privacy wherever possible



# Collecting Data

UCLA School of Law Williams Institute

Best Practices for Asking Questions about Sexual Orientation on  
Surveys

Sexual Minority Assessment Research Team (SMART)

November 2009



# WHAT TO ASK

**Self-identification:** How one identifies one's sexual orientation

- a) Heterosexual or straight
- b) Gay or lesbian; or,
- c) Bisexual

**Sexual behaviour:** the sex of sex partners (i.e. individuals of the same sex, different sex, or both sexes)

In the past (e.g. year) who have you had sex with?

- a) Men only
- b) Women only
- c) Both men and women
- d) I did not have sex

**Sexual attraction:** the sex or gender of individuals that someone feels attracted to

Which best describes your feelings? Are you:

- a) Only attracted to females
- b) Mostly attracted to females
- c) Equally attracted to females and males
- d) Mostly attracted to males
- e) Only attracted to males
- f) Not sure

\*\*\*Ask separately from marital status and cohabitation





# Canadian Community Health Survey

## Canadian Community Health Survey - Annual Component (CCHS)

Canadian Community Health Survey - Annual Component (CCHS)	Variable(s)	Summary of changes	Other reference periods	Related products
<b>Sexual orientation of person, status</b>				
<b>Person</b> refers to an individual and is the unit of analysis for most social statistics programs.				
The data for this variable are reported using the following classifications and/or lists:				
<ul style="list-style-type: none"><li>• <a href="#">Sexual orientation</a></li><li>• <a href="#">Sexual orientation, LGBTQ2+ status</a></li></ul>				



# Canadian Community Health Survey

## Classification of sexual orientation, LGBTQ2+ status

Classification of sexual orientation, LGBTQ2+ status

### Classification structure

Code	Title
1	LGBTQ2+
2	Non-LGBTQ2+



# Census of Canada, 2016

## Classification of opposite/same-sex married spouse or common-law partner

### Classification of opposite/same-sex married spouse or common-law partner

#### 1 - Opposite-sex married spouse or common-law partner

This category includes persons whose married spouse or common-law partner is of the opposite sex.

#### 2 - Same-sex married spouse or common-law partner

This category includes persons whose married spouse or common-law partner is of the same sex.



# Processing and Analyzing Data

## Storage

- Secure

## Processing

- Open-ended text fields (consistency)

## Analysis

- Recoding to a new variable for umbrella category
- Link with another e.g. older dataset (harmonization)



# Processing and Analyzing Data

## Analysis

- Ensure you keep disaggregated data for inferential analysis
  - **1.7%** — The percentage of Canadians aged 18 to 59 who reported in 2014 that they consider themselves to be homosexual (gay or lesbian)
  - **1.3%** — The percentage of Canadians aged 18 to 59 who reported in 2014 that they consider themselves to be bisexual



# Processing and Analyzing Data

## Confidentiality

- Creating de-identified / anonymized public use files
  - Aggregating (e.g. to higher geography to collapsing variables)
  - Document changes and version files
  - Keep masterfile
- If can't sufficiently anonymise (lose analytic value; small sample) can also set up access by request with strong data use agreement



# Preservation

- “only metadata” isn’t terrible if the data are too sensitive and cannot be sufficiently anonymized
- Granular metadata: study-level and variable-level metadata
- Codebook explaining methodology and data structure of datasets and processing work
- Deposit where LGBTQ+ researchers are likely to find it!
- Advocate for repository for inclusive metadata
- Crowdsourcing capability?



# Sharing and Reusing Data

- Don't just “park it” –advocate to key stakeholders (e.g., researchers and policy makers and the LGBTQ+ community!)
- Empowerment!
- Breakdown communication barriers by disseminating in accessible formats (data highlight reports, linking to open access publications)





# Future Directions

- More standardized, richer and up-to-date metadata in repos
- Intersectionality –e.g., sexual orientation and gender identity
- Other data types (images, video, audio, text)
- More LGBTQ+ data (especially population-based studies)

