



Qualitative + Humanities Data

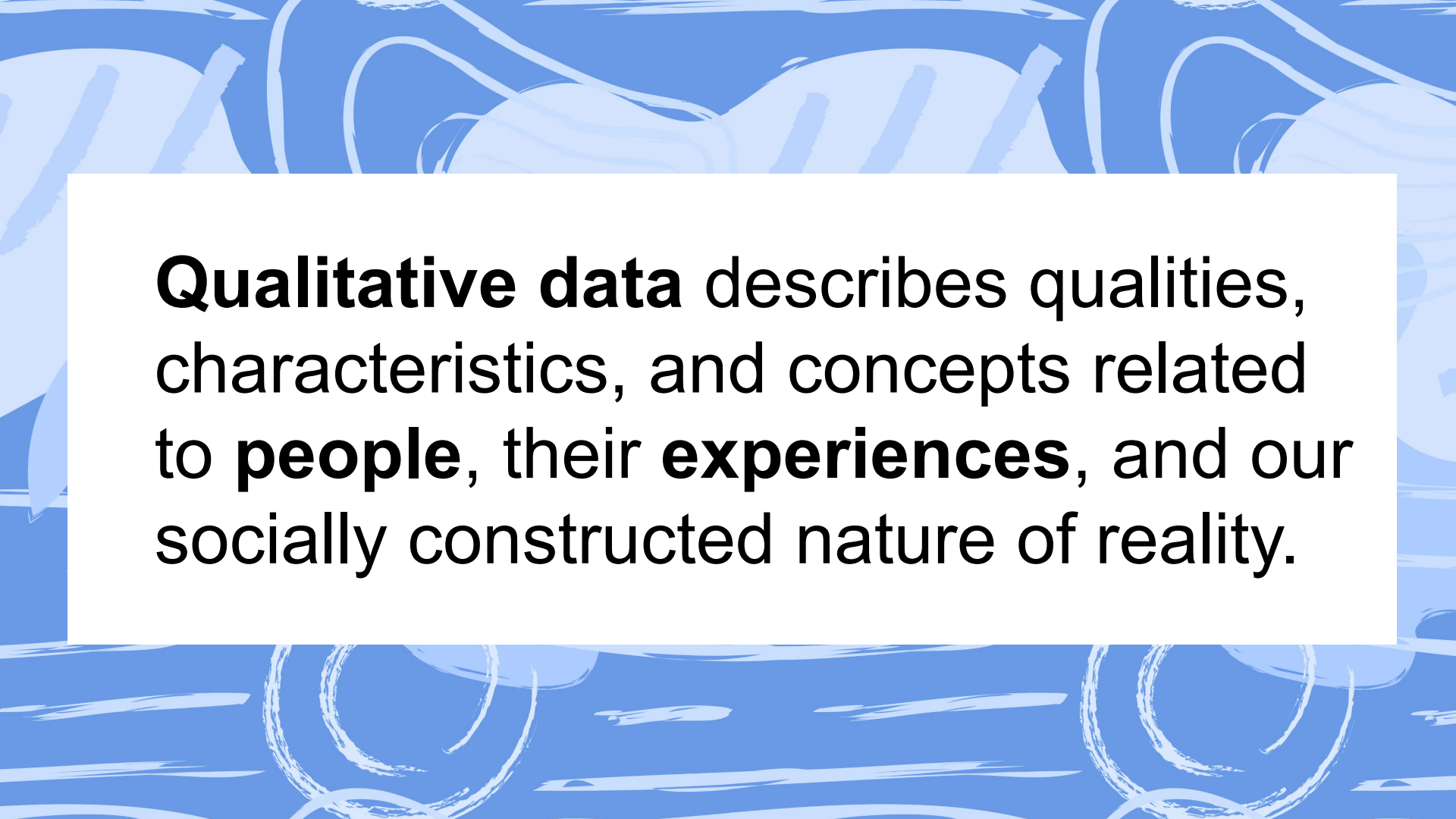
Nickoal Eichmann-Kalwara (she/her)
Data Boot Camp, Winter 2023



Center for Research Data & Digital Scholarship

Objectives

- Build broad familiarity with interdisciplinary data practices and research lifecycles
- Explore some of the types and applications of qualitative and humanities data
- Discuss mixed-methods approaches to data-intensive research and teaching

The background is a solid blue color with white, hand-drawn brushstrokes and circular patterns scattered across it. A white rectangular box is centered on the page, containing the text.

Qualitative data describes qualities, characteristics, and concepts related to **people**, their **experiences**, and our socially constructed nature of reality.

Qualitative Data Analysis

Methodologies

- Case Studies
- Ethnographies
- Grounded Theory
- Phenomenologies
- Narratives

Data Types

- Observations
- Field Notes
- Interviews
- Transcripts
- Demographics

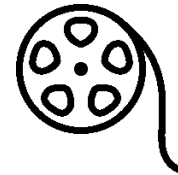
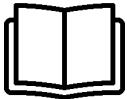
Qualitative Data Analysis

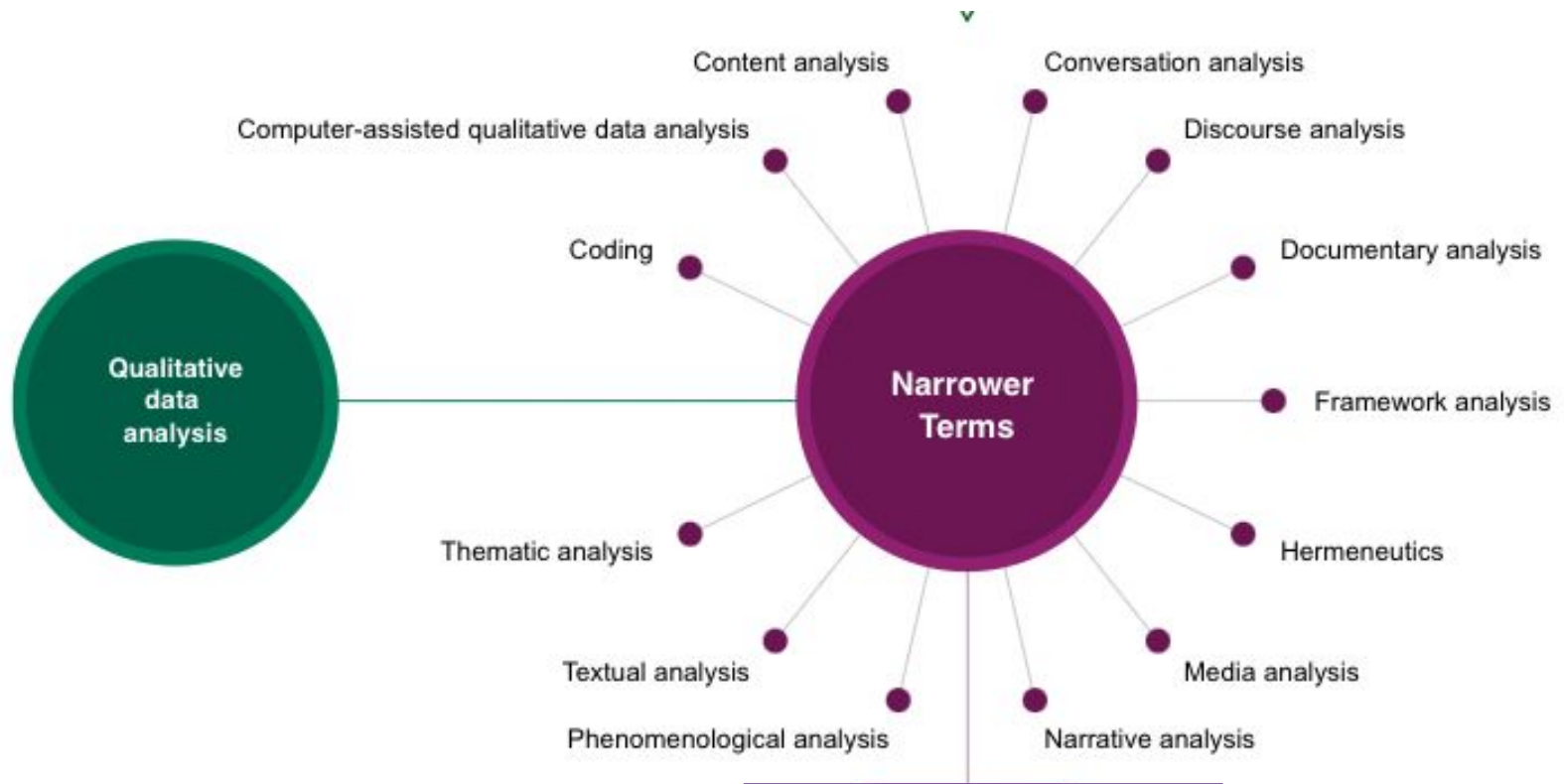
Data Types

- Observations
- Field Notes
- Interviews
- Transcripts
- Demographics


Data formats

- Text
- Video
- Images
- Audio
- Summary statistics





SAGE Research Methods
via CU Boulder Libraries



Humanities data describes qualities, characteristics, and concepts related to **the cultural record**, to explore stories and ideas that help us **make sense of our lives and our world.**

Humanities Data can be defined as:

- data that humanists create and/or use;
- data that represent artifacts traditionally studied in humanities disciplines (e.g., digitized and OCR'd books [corpus], bibliographic metadata, a node-edge list, oral history interviews recordings, web archives)
- critical engagement with data through human-centered / humanistic approaches on data throughout the research lifecycle (e.g., data feminism, data ethics, data sovereignty)

Humanities Data Types

- Books / poems / text
- Images / video
- Music / sound
- Objects
- Social media
- etc..
- Metadata (including TEI)

Artifacts that are
abstracted/represented
for **computability**

Humanities Data Analysis

Methods

- Text Analysis + NLP
- Network Analysis
- Storytelling
- Geo/spatial analysis
- Oral Histories
- Web Archives

Data Types

- Corpora (txt, csv)
- Node/edge lists (csv)
- Multimodal (sound, image)
- Geolocation, maps (lat/long)
- Transcripts (txt)
- 3D data (json, xml)

Humanities Data: A Necessary Contradiction (Posner 2015):

- So it's quantitative evidence that seems to show something, but it's the scholar's knowledge of the surrounding debates and historiography that give this data any meaning. It requires a lot of interpretive work.
- [There is] this tension between the power of computation and the inadequacy of data to truly represent reality.
(normalization)

Context, Abstraction, Process as Method

- Intellectually and culturally informed, indirectly and directly
- Uniformity is rare, especially with crowdsourced data
- To clean or not to clean?
- Data is manufactured, an artifact and text itself subject to interpretation, *whether human or machine created*
- Big data for humanities data suggests heterogeneity in format and time = issues with integration and interpretation
- Computability suggests discrete, clear categories

Collecting / Finding Qual and Hum Data

Qualitative:

- Qualitative Data Repository
- ICPSR Data Archive
- Sage Research Methods

Humanities:

- Internet Archive
- Hathi Trust
- Constellate
- Project Gutenberg
- English Corpora
- Web of Science

**Human Subjects:
Institutional Review Board**

Considerations

- Domain expertise, interdisciplinary flexibility
- Data skills = intellectual contributions
- Collaboration and credit: <https://credit.niso.org/>
- Copyright and Intellectual Property
- Statistical Analyses
- Data Management + Sharing Plans (NEH, NSF)
- “Non-consumptive Research”
- ...etc.

Report Out

- Did any of this resonate with your work?
- Do you see any potential to use qualitative and/or humanities data in your research and teaching?
- What questions do you have about these data?

