

6S 2022 Sketch Groups Template
Staccato Project Design

Use this sketch -- either quickly or more comprehensively -- to draw out an ethnographic research design -- for your own project or just for practice. Do this sketch many times for different kinds of possible projects -- thinking of it as calisthenics for ethnography.

For the 6S 2022 workshop, select one of the sites|initiatives|problem-domains below to design a project for -- thinking first about different types of projects that could be done for the site|initiative|problem-domain you have selected, then about a specific project design, filling in the graphic below. Don't describe your own primary project. This is an exercise in rapid research imagining that you should spend about three hours on -- honing your capacity for "analysis as craft." The sketch can be done individually or collaboratively (the latter is more fun). Feel free to reach out to members in your group if you want to do this collaboratively. Responses can be roughly drafted and in bullet points. Fill in as much as you can in the allotted time, purposely working fast. Prioritize work on "Overview / Research Questions," "Methods and Data Resources" and "Theoretical Frames and Data Analysis." All of the sites|initiatives|problem-domains we've listed as options somehow relate to climate change. We fully understand that climate change is not a central focus for many of you (though it is context for all). We chose options within the climate change problem space so that the collection of proposals the group comes up with points to an array of STS research possibilities within any particular problem space.

TITLE

(Ac)Counting for Climate: Data-Based Labor and the Production of Climate Science

RESEARCHER/S

Me :)

ABSTRACT, INTELLECTUAL MERIT, BROAD IMPACT

Climate science depends on massive and massively varied datasets. From atmospheric gas dynamics to on-the-ground recordings of local fluctuations in biodiversity, climate researchers often work to integrate multiple kinds of data in their work. To facilitate this data diversity as well as maximize the use of resources, much contemporary climate science takes place in/on large shared databases. These organizations standardize and configure datasets for convenient reuse, maximizing the value and reach of data within global, interdisciplinary research ecologies. However, databases are not epistemically neutral intermediaries: as work in Critical Data Studies and Cyberinfrastructure Studies has explored, data(bases) shape and constrain how we understand the things they are used to evince. This project follows the World Bank's Climate Change Knowledge Portal (CCKP) in order to explore how these phenomena shape contemporary climate science and scientific discourse. The CCKP is a particularly large database which bridges social and natural sciences, including globally produced data (broken down by both country and watershed) describing phenomena from annual agricultural production to patterns in housing inequity and energy production. In addition to this diversity of data, the CCKP is unique because it circulates data and narrative descriptions thereof to researchers as well as a variety of nonscientific actors, including media and policy professionals. By investigating challenges related to the production, maintenance and circulation of data through the CCKP, this project takes seriously the situated non-neutrality of climate data in producing scientific, governmental, and popular representations of the climate crisis. It considers how data structure which problems are centered and which are silenced, and relates these vocalities to challenges in the collaborative production and maintenance of big data.

OVERVIEW / RESEARCH QUESTIONS

How do data resources structure climate discourses? How do the affordances of databases shape which concerns, peoples, and environments are centered (and who is left out) in climate conversations? How do the quotidian challenges of maintaining large and current internationally collaborative data condition the above?

BACKGROUND AND SIGNIFICANCE

Even among other big data sciences, databasing climate science represents a particular challenge: databases for climate science require international and multilingual collaboration, function within a deeply disunified disciplinary landscape, and are made to speak to a huge variety of concerns spanning from human migration to biodiversity and extinction. Although an abundance of recent work in STS and information science has begun to center how databases structure the production of knowledge, climate databases remain an urgent and underexplored object of interest. In accounting for data, this project centers databases like CCKP as a critical infrastructure of climate discourse, exploring the relationship between data and discursive constructions of what the climate emergency is and who it impacts.

LITERATURE REVIEW

Database studies/eScience/scientific infrastructure studies, particularly in STS, HCI and info science (Bowker and Star 1999 etc)
Specific literature on climate and environmental databasing (maybe start with Edwards 2010 and work out?)
Literatures on climate change discourses, climate and media, climate and scientific production (this would involve some fishing, I'm not particularly familiar with this work)
Methods literature on studying data and its archives
Critical data studies on "data journeys" (maybe start with Leonelli and Tempini 2020 and work out?)

METHODS AND DATA RESOURCES

Ethnographic interviews with CCKP data managers; data labellers; website maintenance staff; etc. Observational ethnography of their daily workflow and common problems, how they engage with data users, how they source datasets and acquire "current" data, etc.
Discourse analysis (incl. some kind of network or bibliometric study) on published research that uses this data--who is making claims? What kinds of claims do they make? About what kinds of places and problems? What's not there? What other data do they rely on?
This would be an easier corpus to produce, probably could just pull it from webofscience
Discourse analysis on where this data appears in other settings--policy briefs, governmental reports, media, etc.--who is making claims? What kinds of claims do they make? About what kinds of places and problems? What's not there? What other data do they rely on?
This would take a lot of time in a couple of databases and some good old fashioned google-hell to assemble

THEORETICAL FRAMES & DATA ANALYSIS

Theoretical Frame (CDS): Data are not neutral or indexical representations of the world. They are a form of power which acts to shape and constrain subsequent knowledge claims, and profoundly condition subsequent inquiry. While climate data are a critical resource for mobilizing scientific and political action, their voices and silences condition the groups of peoples, nonhumans, and environments that are centered in climate discourse. In attending to data, we need to look to who and what it includes; how it is produced; how it is standardized and maintained over time; how it is made mobile; how it is recombined with data from other sources (and their histories); and what it is made to say. Ultimately, the goal of Critical Data Studies projects is not to suggest that any given data is "flawed" (although it often is)--and certainly not that it can be "fixed"--but to show the intricate ecosystems that make data-based knowledge possible and to examine the points of inflection where data shape and are inescapably shaped by situated social practice. This work avoids technosolutionist approaches and attempts to square data-based knowledge production as one way of knowing among many.

PLAN OF WORK

I would probably work backwards. Start with corpus construction and discourse analysis, try to establish patterns of commonality, frame narratives, things that seem like missing pieces. I would use that corpus and preliminary findings to start work with CCKP researchers, if they were interested in this kind of collaboration--in my experience, database workers appreciate having a corpus that shows how far their work has traveled and the kinds of influence it's had, and it's been a good entry to previous database ethnographies for me.

This is a little tough because I don't want to assume collaborative interest or that this would be a valuable project to CCKP workers; on the flip side, trying to assure their interest in advance feels extractive in a different way, asking for time and support on the promise that I might be able to produce something of interest to them. I like doing the discursive work first so that there's something already to show, and if it doesn't pan out the only time and effort lost is my own.

CHALLENGES AND ETHICAL CONSIDERATIONS

Centering data laborers is an important contribution to the moral economy of professional science--where their work has been historically invisibilized, limiting their career trajectories and writing them out of the record--but data laborers on such projects still tend to be privileged knowledge workers in the Global North. Framing this project this way is a choice that merits some further consideration. On one hand it is a project I'm equipped to do, I have cultural competency working with these kinds of people, the power dynamic is relatively symmetric. On the other, there are strong arguments to be made that these are not the problems of most concern and that amplifying these voices and more formal scientific epistemology problems like the relationship between data and knowledge avoids many (most? all?) of the very urgent stakes of the climate crisis and the people most affected.

VALIDITY AND EVALUATION

PREPARATION AND WORK THUS FAR

Nil, though I know the data and infrastructure literatures ok--would need to resurvey thinking more specifically about climate, though. Honestly I would probably need to read for three months minimum before starting a project like this.

REFERENCES

Bowker, G. and Star, S.L. (1999) *Sorting Things Out: Classification and its Consequences*. MIT University Press.
Edwards, P. (2010). *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. MIT University Press.
Leonelli, L. and Tempini, N. (2019). *Data Journeys in the Sciences*. Springer.

DATA MANAGEMENT PLAN

Ethnographic data would probably need to remain private depending on the construction of the IRB (so idk, gdrive? That's where my institution demands that everything live).

The set of scientific and governmental documents that rely on CCKP data could be made public. This would be a useful corpus for other kinds of research questions, as well as a potentially meaningful artifact to justify continued support for programs like CCKP, or to pursue funding for the addition of new categories of data.

FURTHER NOTES

POINTERS

- Make sure to come up with a title (though this is hard and always feels -- and is -- reductive).
- The abstract should describe your project significance, aims, methods, expected findings/contributions (intellectual merit) and expected societal implications (broad impact). Describe each in a sentence.
- In the Overview / Research Questions, try to articulate the scales, systems or objects that will be foregrounded in the project, and its context|location (geographic, ecologic, geopolitical, discursive, etc). Include both theoretical and empirical questions, and a description of the types of data you will generate and mobilize. End with a few statements about what the project will push *against* (methodological nationalism or essentialist constructs of identity or place, for example).
- In the Literature Review section for a literature review, describe two to four topical literatures that you will build on and contribute to through this research. See [Annual Reviews](#) for ideas but reach for [bibliodiversity](#).
- In the methods section, describe *what you will do, where and with whom* -- and the different kinds of data and insight these activities will produce. Consider, for example, how you might include multisited ethnography ([Marcus 1995](#), a tale of implosion ([Dumit 2014](#)), tactile analytics ([Patricia Alvarez Astacio 2021](#)), drawing as analysis ([Rachel Douglas-Jones 2021](#)), or archive ethnography ([Fortun et al. 2021](#)).

- In the section for theoretical frames, describe the basic theoretical insights that you can mobilize in your study design, data collection, analysis, and writing. You could mobilize understanding of “the subaltern,” for example, or Foucaultian ideas about discourse and subject formation. This can be a long list with very cursory descriptions. Note that this section is not usually included in a proposal submitted to funders -- but should be part of your thinking and dialogue with collaborators
- In building your references, reach for bibliodiversity and a transnational field of reference.

sites|initiatives|problem-domains for 6S 2022 April 26 Workshop

- Melting Siberia
 - Troianovski, Anton and Chris Mooney (photo and video by Michael Robinson Chavez). 2019. “Radical Warming in Siberia Leaves Millions on Unstable Ground,” Washington Post. October 3. <https://www.washingtonpost.com/graphics/2019/national/climate-environment/climate-change-siberia/>
 - Struzik, Ed. 2020. “How Thawing Permafrost Is Beginning to Transform the Arctic,” Yale Environment 360. January 21. <https://e360.yale.edu/features/how-melting-permafrost-is-beginning-to-transform-the-arctic>
- Climate Change and Combo Disaster in the United States
- Climate Change, Labor Productivity and Politics
- USAID's Climate Links
- World Bank's Climate Change Knowledge Portal
- WHO et al Environment and Health Compendium
- WHO, UNDP, UNEP and UNICEF have partnered to create a new compendium of 500 actions aimed at reducing death and diseases driven by environmental risk factors, the first such resource to unite this expertise from across the UN system.
- Climate Change and Social Media
- World Economic Forum on Climate Governance