Individual Project Proposal – Megan Ozeran

For my individual project, I plan to investigate and define the ethics of data visualization. These ethics have not been formally defined, and it is crucial that we do so.

Current technologies make data collection, data analysis, and data visualization easier than ever before, even as the amount and variety of data continue to expand. Computer-mediated data analysis is no longer simply in the domain of computer scientists, as more tools and software are made available to researchers with a variety of technical expertise.

This leads to two intersecting issues. The first is that newer technological tools and methods continue to spark ethical quandaries, and the fields of computer ethics, information ethics, and related applied ethics are still in flux. The second issue is that increasingly more people are able to use these new tools and methods, but they are not always taught to consider related ethical questions.

As a visualization librarian at the University of Illinois, I teach people from a variety of disciplines how to visualize their research. I may teach them through a workshop or in a one-on-one consultation session. In either case, I don't have a lot of time, so I have focused primarily on simple, practical principles. It's especially difficult to incorporate a discussion of ethics into a short training session if those ethics haven't even been formally defined.

This is a disservice to my students and to the field of data visualization as a whole. It is crucial that we find a way to teach ethical principles alongside the practicalities of data visualization. To get there, we need to do two things: begin to identify the ethics of data visualization, and design instruction that incorporates those ethics.

Note that I say "begin to identify" because ethical debates are ongoing. A lot of work is involved in creating codes of ethics, like that of the American Library Association, which are subsequently hotly contested and changeable. This doesn't mean that codes of ethics are useless; rather, they help to frame what is otherwise an unmanageable space for ethical debate. Therefore in this research, the important first step is to identify the ethics of data visualization *well enough* to frame future debate and clarification. By identifying the ethics of data visualization well enough, we can also begin designing and assessing ways of teaching those ethics to the populations we serve.

I have already embarked on a first step in this process. A colleague and I are conducting surveys this semester with graduate students and faculty across campus. Our goal is to identify what values people consider when they create data visualizations. By surveying current practitioners

across a variety of disciplines, we hope to identify themes that are common to data visualization generally, rather than values held by specific disciplines.

If I am accepted as a fellow with Visualizing the Future, I would utilize the distributed nature of our cohort to expand these initial surveys to a national sample. The data from these surveys will identify what values practitioners currently consider. Just as crucial will be the values they *should* consider. To identify these values I will continue to review the ongoing conversations about related applied ethics, visual rhetoric, data feminism, and critical studies.

I will make the bold move of proposing ethical principles for data visualization. I fully expect that my initial proposal will be wrong in some way – perhaps misidentifying or overlooking important values. Therefore the proposal will be done in a way that encourages community feedback, such as via a Wiki or Google Doc. Knowing that online debate is not always constructive, I would solicit advice from the cohort and my colleagues on how to design the feedback process to be as equitable and positive as possible.

We can begin to design instruction that incorporates these ethics even as they are debated. Each ethical principle can likely be translated into a learning outcome for which we design modular instruction.

I am excited to embark on this research journey. If I can help identify its ethics, I can fundamentally improve how all of us teach data visualization. This only becomes more crucial as increasing numbers of people collect, analyze, and visualize data.