**Module 2 Activity: Identifying Types and Stages of Data**

**Please read this case study and answer the following questions.**

Researchers were interested in impacts of potential future climate change and coastal flooding on two Massachusetts communities. These communities would both be affected by coastal flooding, but in one community flooding would mostly affect residential areas, while the other community’s commercial areas would be most vulnerable. Researchers held a series of workshops with residents of both communities to explore possible responses to increased coastal flooding resulting from sea level rise. Researchers designed the earlier workshops in the series to gain understanding of the residents’ existing knowledge of climate change, with the goal of linking this knowledge to the communities’ potential to adapt to climate change in later workshops. The workshops were organized by local community and non-governmental organizations, and were held over the course of several evenings. Audio recordings were made of the workshops with the consent of the participants.

The goal of the first two community workshops was to better understand the residents’ existing understanding of climate change. In the first workshop, residents were asked to list all words they could think of when they considered the term ‘climate change’. The words were called out by residents and recorded on flip charts by organizers. These words included terms related to environmental impacts (e.g. droughts, deforestation, global warming) as well as public health impacts (e.g. pollution, famine, epidemics). Residents were then asked to vote on the words they thought represented the potential climate change impacts that were most important. This voting process reduced the word list to 47.

In the second workshop, participants were given a pile sorting exercise. They were provided with the 47 words generated from the previous workshop, each on a separate piece of paper, and asked to sort the words into piles of related terms. After the workshop, the pile sorting results were entered into a spreadsheet. Multidimensional scaling analysis was performed on this data set, and the residents’ pile sorting results were used to generate a model of perceived proximities between terms. The ‘relatedness’ of these terms was then visualized.

In the third and final workshops, the researchers presented potential 100-year coastal flood maps of the communities that had been generated based on various emissions scenarios for the years 2030 and 2070. Researchers then presented various options that would address sea level rise and help residents adapt to coastal flooding in their communities. Option choices were informed by the residents’ knowledge of climate change as elucidated in the first two workshops. These options included a modular sea wall, construction of natural buffers such as a beach/dune system, evacuation plans and routes to move residents away from the flood-prone areas of their communities, and various types of structural flood-proofing. After the presentations, residents were divided into four groups. Each group was assigned a moderator/facilitator who transcribed responses to these discussion questions:

1. Which of the options seems most feasible/attractive?

2. Which options would you object to and why?

3. What obstacles are in the way to getting the options in place?

4. What needs to happen in the community to adapt to sea level rise?

Following the workshop, researchers consolidated and summarized the discussion responses, and looked for similarities and differences in responses between the communities. They also evaluated residents’ levels of preparedness and engagement, based in part on outcomes of the word generation and pile sorting exercises.

***Questions:***

*1. Briefly list the data generated and used in the course of this study.*

*2. What file types/formats might be involved in the study?*

*3. Taking your list of data can you develop a list of potential file names?*

*4. Can you structure the file names into a file plan with folder names?*

*5. For each of the items on your data list, match it to one of the following data stages*: *Raw data, Processed data, Analyzed data, Finalized/published data.*

*Raw data: What is being measured or observed? This is the data that is being generated during the research project. An example of raw data might be daily measurements of temperature in Lake Superior. Raw data coincides with the ‘creating data’ section of the research data lifecycle.*

*Processed data: How can the raw data be made useful/manipulable? To continue with the example above, lake temperature data may become processed once researchers remove clearly erroneous temperature measurements from the data set, and enter the remaining temperatures into a spreadsheet for manipulation and analysis. Processed data coincides with the ‘processing data’ section of the research data lifecycle.*

*Analyzed data: What does the data tell us? Is it significant? How so? For example, daily lake temperature data could be analyzed by finding average temperatures, looking at seasonal fluctuations, and generating graphs that demonstrate these changes. Analyzed data coincides with the ‘analyzing data’ section of the research data lifecycle.*

*Finalized/published data: How does the data support your research question? For example, a plot of our average lake temperatures for 2013 may show statistically significant differences when compared to the same data from 1913 and 1963. Finalized/published data coincides with a few sections of the research data lifecycle including preserving, giving access to, and re-using data.*

\* Case study adapted from Douglas, E.M., P.H. Kirshen, M.J. Paolisso, C. Watson, J.Wiggin and M. Ruth. 2012. Coastal Flooding, Climate Change, and Environmental Justice: Identifying Obstacles and Incentives for Adaptation in Two Metropolitan Boston Massachusetts Communities, Mitigation and Adaptation Strategies for Global Change, Vol. 17, No. 5, pp. 537-562