# Case 1\_ Professor Periwinkle

* Your research team is made up of staff and students- do you have a dedicated staff person performing data analysis- or are staff capable of using data analysis systems?
  + "you are my dedicated data person. And I expect you to train my staff, or leave instructions for them on how to use data analysis systems."
* What format do you use for your NetCDF files?
  + "I do not understand the question. NetCDF is a data format. You should know that."
* What size is all of your CSV files in uncompressed and raw data so far?
  + "200GB"
* What size is all of your TSV files in uncompressed and raw data so far?
  + "500GB"
* what kind of linked data would you like to see? Linked data mean data linked to other data that have connections. Example: a excel file on average food prices are linked to another excel file on names of restaurants in nova scotia.
  + "Obviously I would like any of the relevant data in the data sets to be linked. Otherwise how would I able to query it?"
* what sort of categories do you want to see for the organization of your data?
  + "I want to be able to search by organism, geographical feature, depth, environmental conditions. I would be interested to see if you have any other ideas."
* What data do you consider to be the most relevant in retaining?
  + "All of my data is relevant and has value!"
* Where is the data currently being stored?
  + "As I have already mentioned in my initial request it stored in various ways, locally, hard drives, DVDs, CDs, etc. And a few files are in a Dropbox account."
* I am emailing you in response, to your request for a data management plan. We noticed that you are interested in data that has been collected outside of Canada. Do you expect that some of that data collected may be in a language other than english?
  + "yes, some data might be in a different language. I plan on hiring translator for those cases."
* wanted to ask about ownership of all of the different types of data you collect: Are you the owner (ultimately accountable for the quality of different datasets), or is the University? During the lifecycle of a given project, who is responsible for data quality on a day-to-day basis? You mentioned that students in your research teams share data with in-coming students when they graduate or they can take the data with them. It would be helpful to know what sort of ownership or other rights current and past members of your teams have to the data, if these have been defined.
  + "I have never bothered to get specific licenses. After your request I took a look for the citizen scientists and my own contract. Their data is open data and I own the data I produce.  
    There is not procedure or contract in place that the students (incoming or leaving) have to sign."
  + We wanted to know if you have copies of all of the data that students take away with them? If not, do you have a way to contact former students to get access to their data?
    - "yes, I do have copies. I would never let them leave with the only copy."
* Can you let us know if you have any funding allocated for your Data Management Plan and if so, how much is available to us for budgeting purposes?
  + "there is funding in my grant for the Data Management Plan. The funding is more than enough to cover all your expenses. That does not mean though that I would not like for you to explain to me your budgeting choices."
* I wanted to know whether the wildlife sightings reported by citizen scientists via your website are data that you would *eventually* like to use and, if so, what format the data are in.
  + "yes, I definitely want to use that data. I download them as tsv files."
* can you tell me what format the field notes from the various field experiments you conduct are in, including those that document the animals captured and tagged, from mark-recapture population estimation experiments, and from observational studies conducted during your classes?
  + "The field notes follow the Darwin Core: <http://www.iobis.org/manual/darwincore/>All the required filed must be filled, otherwise the data are dismissed. "
* Finally, are you able to give me a basic idea of how the data you share with your team have been structured up to now?
  + "Everyone has their own system in sharing the data, and their own structure. I believe in letting everyone work how they think is best!"

# Case 2 - Professor Green

* you know we have a mix document files (PDF, Word and plain text). Could you tell us the average size of those documents?
  + "the file size ranges from a few hundred kilobytes to 25-30MB."
* Could you please provide us the name or the code for each hospital that you are planing to examining
  + "I work with a few hospitals in Nova Scotia, one hospital in France, one in Germany and a few more in Calgary."
* Should we consider that he has already got the license to gather the data?
  + "of course I have a license."
* Does that mean the data is not shared with public but shared with a project team?
  + "The data is not shared with the public. As I mentioned, it can identify people. What I share with my team are the notes I make (recordings, spreadsheets, text files) after I have gone through the data. And that is only shared with my project members. If they want more information, they have to ask me."
* This case did not mentioned the size of some data collection , how we can know the data size?
  + "You are the expert! There are 15 interviews, each interview is an hour long, and I am encoding them in mp3 format, 128kbps."
* Is the data that was collected in 2002 related to the data that was mentioned in this case?
  + "There might something relevant in the 2002 binder with the printed spreadsheets. I would like for you to look through it. It does not contain anything that can identify participants. "
* "which data will retain value after the life of the project?" should I be specific about the data type and how I can know if it has value or not?
  + "All my data is important and of value!"
* Is there a specific size for the zotero database ? Also,how much space do you wish to have for your future research (How much do you think youll use in 5 years from now?
  + "the current size of all my files right now fits in my USB drive which is 64GB and it is not full. There are about 40GB left over. I predict that the data influx will increase in the near future. My research is very important! You say five years? It will be at least ten. I predict it will triple in size."
* There are 383 documents with several formats. Would you like to have them available on the original format or would you be willing to have one standard format for your database ?
  + "Why would I want them in one format? Do you plan on changing my original data??"
* How long will the data be collected and how often will it change?Who is responsible for managing the data? Who will ensure that the data management plan is carried out?
  + "It is your responsibility to manage the data after I am satisfied with what you have come up with. So you have to make sure that it is carried out. After this round of funding is over and I get my new grant (yes, I am sure that my grant will be renewed!), whoever takes your place will be responsible for carrying out the data management plan."
* 'll make use of a specific data storage service. As a Dal faculty member you can make use of the university's space to store your data. Is Dal's repository ok, or do you have any specific preference ?
  + "the Dal repository is fine for now, but I would like to know what other options are available. Not too many, two or three would be good."
* If you are required to share interview instruments or anonymized data to conduct your studies, how is this done? What is the language of communication?
  + "I am not required to share the original files or interviews. Only information that does not allow for subject identification. That data needs to be translated to the official language of each collaborator's location."

Case 4 - Professor Chartreuse

* Is the "science of science" considered part of social sciences? ( I need to know as to select the proper metadata standard for your research)
  + "The science of science (SciSci) is based on a transdisciplinary approach that uses large data sets to study the mechanisms underlying the doing of science—from the choice of a research problem to career trajectories and progress within a field"
* What's the size of the data that you have so far? and how long do you think your research is going to last?
  + "My data collection is about 20GB is size, and I plan on continuing until the day I am no longer here. "
* What type of metadata would you like to see happen with the files? Metadata is how you label your files. Example: title of the file, the date of the creation of the file. the subject of the file and so on.
  + "I would like the metadata to include relevant information for me to be able to query it later on. So, definitely date, name, keywords, publications that used the data, the authors. Those a mandatory, if you think that you can add more, so that I can make more complex queries, let me know."
* What type of openness would you prefer to see for the data? For example, the database could be close to only your view or it can be open to everyone.
  + "Since I collect my data from public sources (mainly Pubmed), I would like to keep their licensing and definitely make it available for non commercial use only."
* What data do you consider to be the most relevant in retaining?
  + "Any type of data that will lead to a publication."
* Where is the data being stored?
  + "Right now the data lives in Excel files in my hard drive."
* How do you organize your data in excel?
  + "Simple excel files with a maximum of 3 worksheets."
* Is there a specific format structure in the way you store your data in your external hard drive, that you would like to keep (i.e Different Files structure)?
  + "I store it by date. I have found that it makes it hard to go back and search for things. So I am not adamant in keeping it."