What data will be created or collected (type, size, format, etc.)

**(DML** is abbreviation of Data Manipulation Language. It is used to retrieve, store, modify, delete, insert and update data in database. **DDL** is abbreviation of Data Definition Language. It is used to create and modify the structure of database objects in database.)

|  |  |  |
| --- | --- | --- |
| Professor Periwinkle | Professor Green | Professor Pinkerton |
| - 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.(Discussion)  - All of my data is relevant and has value!(Discussion)  - Current data is stored in various ways, locally, hard drives, DVDs, CDs, etc. And a few files are in a Dropbox account. (Discussion)  - CSV files in uncompressed and raw data (200GB) (Discussion)  - size is all of your TSV files in uncompressed and raw data (500 GB) (Discussion)  - wildlife sightings reported by citizen scientists are downloaded as tsv files (Discussion) | - Currently he is using two primary methods of collection, Textual and content analysis. (Case)  - Textual is 383 individual documents describing - Team, Team outcomes, Team Practices. (Case)  - Data is obtained from healthcare organization.(Case)  - Currently it is in the form of PDF, Word, plain text, and other documents all in digital formats (Case)  - Currently it is stored in Zerotero database (Size unknown) (Case)  - Qualitative data is stored in excel spreadsheet. (Case)  -Interviews recorded in mp3, then transcribed in Word. (Case) # Challenge to keep the anonymity of the interviewee |  |

What licenses apply to the data

<http://pitt.libguides.com/copyright/licenses>

|  |  |  |
| --- | --- | --- |
| Professor Periwinkle | Professor Green | Professor Pinkerton |
| -Never bothered to get specific licenses (Discussion)  -Their data is open data and I own the data I produce. (Discussion)  -There is not procedure or contract in place that the students (incoming or leaving) have to sign. Do you think that it is necessary to have that written down somewhere? (Discussion) |  |  |

What facilities and equipment will be required (hard disk space, backup server, central repository, off site repository, etc.)

|  |  |  |
| --- | --- | --- |
| Professor Periwinkle | Professor Green | Professor Pinkerton |
|  |  |  |

What data management practices (backups, storage, access control, archiving etc.) will be used

|  |  |  |
| --- | --- | --- |
| Professor Periwinkle | Professor Green | Professor Pinkerton |
|  |  |  |

Who will own and have access to the data

|  |  |  |
| --- | --- | --- |
| Professor Periwinkle | Professor Green | Professor Pinkerton |
| -Their data is open data and I own the data I produce. (Discussion)  - There is not currently a data sharing Agreement in place, merely a collection of ad hoc sharing  opportunities.  In contrast, NOAA in the US has The IOOS system which spans the country, and the European Commission has EMODnet. (Case)  - Everyone who starts with Dr. Periwinkle’s Team Is Told How the data is formatted and structured, and data is primarily shared within the team using external hard drives and USB keys. (Case)  - There is a help document from 2003 describing how to work with the data sets, and how to contribute to the datasets; this help document is on a shared drive somewhere. As students graduate they Share their data with current students in the lab, or take it with them on their personal devices (Case) |  |  |

Which data will retain value after the life of the project

|  |  |  |
| --- | --- | --- |
| Professor Periwinkle | Professor Green | Professor Pinkerton |
| - Past data is stored in various old formats on her office shelf - floppy disks, ZIP Disks, CDs, DVDs, BluRay, and external hard drives (Case) |  |  |

What metadata and linked open data strategies will be employed

|  |  |  |
| --- | --- | --- |
| Professor Periwinkle | Professor Green | Professor Pinkerton |
|  |  |  |

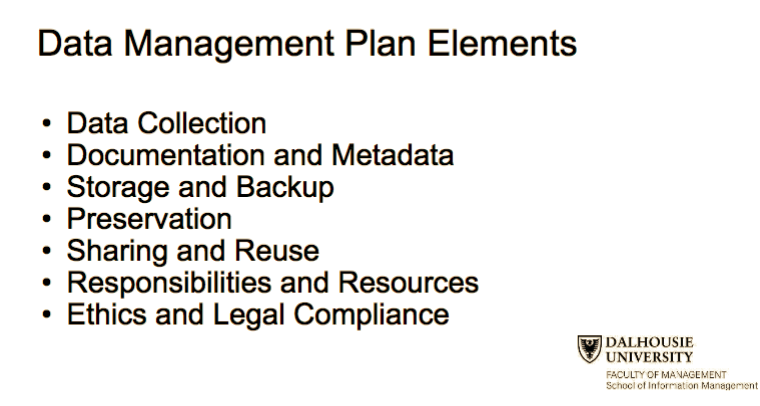
How will its reuse be enabled and long term preservation ensured after the original research is completed

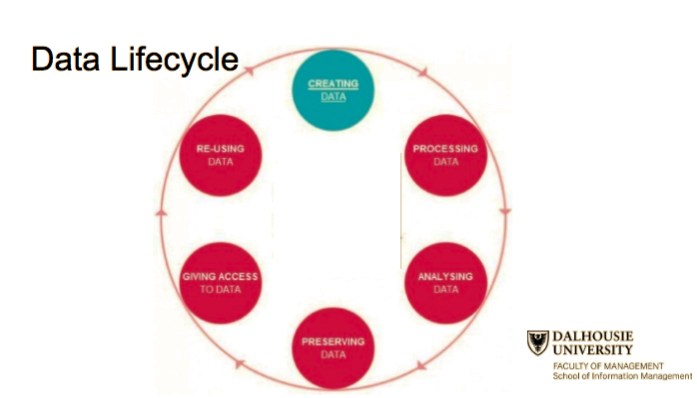
|  |  |  |
| --- | --- | --- |
| Professor Periwinkle | Professor Green | Professor Pinkerton |
|  |  |  |

How much will the storage of this data cost (cloud and/or hard drives)

|  |  |  |
| --- | --- | --- |
| Professor Periwinkle | Professor Green | Professor Pinkerton |
|  |  |  |

<http://pitt.libguides.com/copyright/licenses>





What data will be created or collected (type, size, format, etc.)

What licenses apply to the data

What facilities and equipment will be required (hard disk space, backup

server, central repository, off site repository, etc.)

What data management practices (backups, storage, access control,

archiving etc.) will be used

Who will own and have access to the data

Which data will retain value after the life of the project

What metadata and linked open data strategies will be employed

How will its reuse be enabled and long term preservation ensured after the original research is completed

How much will the storage of this data cost (cloud and/or hard drives)