Group Project !

**Professor Periwinkle**

Necessary Data

Licensing Information

Necessary Facilities/Equipment

Recommended Data Management Practices

Ownership and Access to Data

Post-Project Data Value

Metadata Requirements

Relevant Linked Open Data Strategies

Reuse and Long-Term Preservation

Storage Costs

**Professor Pinkerton**

Necessary Data

Licensing Information

Necessary Facilities/Equipment

Recommended Data Management Practices

Ownership and Access to Data

Post-Project Data Value

Metadata Requirements

Relevant Linked Open Data Strategies

Reuse and Long-Term Preservation

Storage Costs

**Professor Chartreuse at JCU (University Professor)**

Research interests = the science of science, ie. how researchers and scientists work together to generate new knowledge

* Primarily uses public data sources (ex. Pubmed)
* Pubmed – available via a web API and served in JSON (J**SON** (JavaScript Object Notation) is a lightweight data-interchange format. It is easy for humans to read and write. It is easy for machines to parse and generate. It is based on a subset of the JavaScript Programming Language) Asp.Net **Web API** is a framework for building HTTP services that can be consumed by a broad range of clients including browsers, mobiles, iphone and tablets.
* Keyword search to find things – and then uses excel files to organize findings
* Wants to combine data from the PubMed searches and Wikidata + other public resources + survey and interview data with scientists (primary data)
* Use = write papers + create exercises for graduate classes + visualizations (Lumira, etc.)
* Students can view at any time, but need permission to download it
* Desires = consolidate data (ex. ER diagram or schema)
* Concerns = consistent and protected
* Helpfuls = a better method for collecting public data, and a better way to share his data with other scientists (outside of just his own university). Share so that they can view but not edit, and they need permission (a license?) to even view it. Available for non-commercial use only.
* Can add to database after being given access to do so
* - 20GB of data over 4 years
* JCU will cover costs of data management if there is a valid reason for the costs
* Will take free if available
* Currently stored in excel files based on keywords. Stored in folders based on dates. Willing to move away from the date option.
* Metadata = definitely date, name, keywords, publications that used the data, the authors
* Wants to retain information relevant to getting published. But public data might be able to go
* Hard drive
* Currently has a lot of data and information storing.
* Chartreuse probably uses the Data Lifecycle (Slide 65)
* Can recommend what to get rid of – because people are data hoarders.

Necessary Data

* Primary data = surveys and interviews with scientists
* Metadata = date, name, keywords, publications that cite the data, and authors
* Nice to keep all the data – but not necessary. 🡪 possible thought, just copy over the primary data and papers, etc. and leave behind the excel queries for now.

Licensing Information

* Creative Common License: BY-NC-SA (slide 57)

Necessary Facilities/Equipment

Recommended Data Management Practices

Ownership and Access to Data

Post-Project Data Value

Metadata Requirements

Relevant Linked Open Data Strategies

Reuse and Long-Term Preservation

Storage Costs