**Case-2:**

**Admin Details**

**Project Name:** Professor Green Data Management Plan

**Principal Investigator / Researcher:** Professor Green

**Consultant Company:** Data Masters

**Institution:** Dalhousie University

1. **Data Collection**

In this case, three data collection tools were used during the experiment:

1- Text documents:

383 individual documents in different format PDF, MS word, and plain text describing the teams, their outcome and practices. It is obtained from healthcare organization and transcription of the interviews in MS word format.

2- Spreadsheet:

Quantitative data about these documents is saved in Excel spreadsheet.

3- Audio file

15 Interviews is saved in mp3 format which are transcribed in word document.

4- TVS file:

The open data set regarding healthcare, including healthcare outcomes, healthcare expenditures, healthcare staffing, enrolment by discipline.

**1.1. File Naming Convention**

This project will have two folders of the datasets. First folder will have the original datasets and the second one will have the datasets after translation. The naming convention for the original main folder of datasets should have the “Original” word, then project name and the date, for example “Origenal\_HealthCare\_20180212”. The subfolders should have the country name to recognize where the data come from. For example, Canada, Germany, and France. The Canada folder will have two folders, named "Nova Scotia, and Calgary", which will have different folders based on the name of institutions. Files' names will have the type of the file, participant’s ID (if applicable), then language codes such as Germany “DE”, Canada “EN” and French” FR”. For example, an audio file should be named “audio\_P1\_EN”.

As we explained above:

1- The naming convention should be descriptive.

2- It should distinguish between the original data and the translated data.

3- It should contain the date of the data collection, in the format (YYYYMMDD).

4- It should show the type of data attached to the participant’s ID if applicable.

We believe this way of naming convention of the folders, subfolders and files, will help other researchers to easily access and use of the data.

1. **Documentation and Metadata**

The metadata will be written in a text file called “README”, that includes descriptive information of the data to ensure that future users will understand the data. The README.txt file will include the information in the table below:

|  |  |  |
| --- | --- | --- |
| General information | Creator | Professor Green |
| Title | Teamwork in Hospital Environments |
| Date | April 10, 2018 |
| Funding agencies/period | 10 Years |
| Keywords | Hospital, Teamwork, Stress. |
| Coverage | A few hospitals in Nova Scotia, one hospital in France, one in Germany and a few more in Calgary. |
| Funding | CIHR (Canadian Institutes of Health Research) |
| Access information | Access restrictions | Project members can access the note taken under the supervision of the main author (professor Green). |
| Copyright | Exist |
| Technical details | File format | Text documents (Docx, plain text, PDF).  Excel Spreadsheets (quantitative data about these documents  Audio MP3 files (interview) |
| Count of files | 383 individual text documents.  15 audio files. |

1. **Ethics and Legal Compliance.**

In terms of accessing the sensitive information, the only selected group members will be given the right to access to the notes that is created by Prof. Green.

For the privacy of the participants, each one will be given ID for identification, to make them anonymous and keep them confidential. The anonymity of the participants in textual data will be preserved by using their IDs in any presentation or publication.

1. **Storage and Backup**

This research should have a storage system that will meet the requirements of the project. It is important to employ multiple methods to backup and copy the project data. We need to look into three aspects of data storage which are storage space, cost and security.

Each document file size ranges from a few hundred kilobytes to 25-30MB and each interview is an hour long, and is encoded in mp3 format at 128kbps. Therefore, we anticipate the maximum size of the whole documents will be around 11.5 GB. Moreover, there are 15 interviews and the length of each audio is 128kbps. So the size of the whole audio files will be 823.5 MB.

In terms of the first aspect, which is storage space, using USB drive with 64 GB will fit the data size for now, because the usage space is 24 GB, which means 40 GB is free. However, the data will expand in the coming 10 years, and there will be more translated data added. Thus, we need more space to save the data.

In terms of cost and security, using USB drives is cheap and easy solution, but it is not secure and can be easily destroyed or lost.

We anticipate that the size of the data will be more than 72 GB if we consider the upcoming years and translated data for these years. Consequently, the data needs to be stored in a large and more secure space. Sensitive data should not be stored on any repository, because no matter how secure they claim the repository is, it could still be hackable.

We recommend the use of multiple forms of storage:

1. The DataVerse: it is open source research data repository software provided by Dalhousie university that allows researchers to deposit and share data openly or privately. The data is hosted on Dalhousie's servers. The service is primarily for those affiliated with Dalhousie University (DalLibGuides, 2018).
2. pCloud: it offers 500 GB with one payment “175$ “for lifetime. With pCloud, researcher can upload any type of files directly to their account, regardless of the size. It is more secure, because information is encrypted using TLS/SSL when it is transferred from the device to the pCloud servers (Pcloud.com, 2018).
3. Sync.com: it offers 1 TB of secure file storage by 5$ per user per month. It uses end-to-end encryption and No third-party tracking for privacy protection (Sync, 2018).

These services are recommend because it is easy to access and control outside the office. In addition, researcher needs to maintain a master copy of the data on an external local such as Dataverse or external remote such as pCloud and Sync.com.

1. **Selection and Preservation**

Interview data cannot easily be recreated or reproduced, so it has a long term value that needs to be preserved. The other data documentations are necessary to validate research findings, so it is required to keep them. Data will be preserved and available for at least 10 years under the control of Professor Green. The data can be stored in the following data storage:

· The DataVerse

· pCloud

· Sync.com

1. **Data Sharing**

The audio files and transcriptions cannot be publicly shared, because they contain potentially identifying information of participants. Other quantitative data cannot be released as well, until this research is published, to protect intellectual property rights of the researcher. Data will only be shared with the research team by Professor Green. If they need more information during data preparation and analysis, the data with anonymity of participants using Dropbox (for audio files), Google Docs (for transcriptions) and Zotero (for documents).

CC BY-NC-ND license will be applied to the project, because it will not allow anyone to make any modification on the original data, or use it for commercial purposes. However, it will allow other researchers to copy and utilize only the original copies of the work, if they have Professor Green's permission.

1. **Responsibilities and Resources**

Our consult group will be responsible to manage the data after Professor Green is satisfied with what we have proposed. After the project is completed, the new consultants will be responsible to manage the project data, if Professor Green has new funding.

**Case-3:**

**Admin Details**

**Project Name:** Professor Pinkerton Data Management Plan

**Principal Investigator / Researcher:** Professor Pinkerton

**Consultant Company:** Data Masters

**Institution:** Dalhousie University

(please use in-text citation in APA format and put it in the reference list in APA as well)

**Data Collection**

The research project involves basically two forms of data

* Excel spreadsheets (around 17,384 Spreadsheets) exists in different versions of Excel.
* Excel CSV files.

These spreadsheets are basically composed from complex text files, quantitative textual analysis, and other data that are either created by the researcher, or collected from different sources based on the researcher field of interest (e.g. government, private corporations, and other researchers). The data is being collected through variety of methods and it is roughly 10 years old.

**Documentation and Metadata**

The below table includes descriptive information about the data included in the plan. This metadata can be used in future for the SQL data option or the cloud storage option. It will help in the naming convention of the tables in the database option, or files and folders in the cloud storage option:

|  |  |  |
| --- | --- | --- |
| General information | Creator | Professor Pinkerton |
| Title | MIS wizard |
| Date | April 10, 2018 |
| Funding agencies/period | Personal & institutional funding |
| Keywords | Managing Information Systems |
| Access information | Access restrictions | Open data accessed publicly.  Closed Data accessed by permission only. |
| Copyright | Exist |
| Technical details | File format | Excel Spreadsheet (different versions).  CSV files. |
| Count of files  Number of Rows  Median Row | 17384 Spreadsheets  1-750,000  1000 |

**Ethics and Legal Compliance**

Reference : <https://mantra.edina.ac.uk/>

**CopyRight and Creative Commons**

For this plan our team recommends that copyright and creative commons should be taking into consideration. As the research data can contain both copyrighted and non-copyrighted data, or creative commons materials. It will be up to Prof. Pinkerton professional eye to keep track of this. However, when in doubt assume it is copyrighted and contact publisher for permissions.

**Consent Form:**

A letter explaining the purpose, approach and distribution strategy (including plans to share data) of the research, and an associated consent form, will be prepared and translated into relevant languages. This letter will be formatted by our team, and that can be used by the researcher team in order to carry any future researches.

**Intellectual Property Rights:**

Principal researchers and their institutions hold the copyright for the research data they generate. By depositing the data with the university, investigators do not transfer copyright, but instead grant permission to distribute the data, and to transform it as necessary. That is to protect respondents’ confidentiality. Online and archival sources will be cited and clearly acknowledged in the database and research outputs.

**Confidentiality:**

This project contains sensitive data (students performance information). Commitments to ensure confidentiality will be maintained, by ensuring details that can be used to identify participants, are removed from transcripts or concealed in write-ups. Our team will avoid transmitting unencrypted personal data electronically.

**Disclosure risk management**:

Once deposited, the data will undergo procedures (screening processes) to protect the confidentiality of individuals, whose personal information may be part of archived data.

**Storage and Backup**

This project’s research data is considered to be very important by Prof. Pinkerton and our team recommends that multiple backups are in order. Storage and backups using both cloud, hard drive copies options should be considered.

First of all, it is recommended to use hard drive as first line of backup, and to keep it up to date as much as possible. One way to do that is by adding and updating the files on the hard drive every month. This would ensure the master copy is safe against any corruptions that could happen on the cloud. High quality products should be used, from reputable manufacturers.

Furthermore, taking into consideration the sensitivity of some of the data in the project, the following cloud services are proposed:

**Plan 1** - **OneDrive:** Dalhousie University offers both students and faculty members the advantage to use their institutional funded OneDrive Cloud. Every Dalhousie faculty member, staff member, student, or different committees (e.g. Computer Science Graduate Society) is allotted their own 1TB of space for storage. Additionally, they have the benefit of using Microsoft Office 365 for collaborative work, and each group is provided with 1TB of storage (separate from individual user storage).

My team recommends that research data could be stored on the University’s server. This is high quality, enterprise-class storage with guaranteed backup and resilience (Dalhousie, n.d.).

**Plan 2 - Amazon Web Services (AWS) :**

For excel the data gets bigger with the passage of time. Big data when shared over the cloud got chances of collapsing. For this its highly recommended to transfer the data from Excel to MySQL. SQL supports structured data in cloud. MySQL for Excel makes the task of getting MySQL data into Excel a very easy one

**Amazon Aurora**

Amazon Aurora is a relational database engine that combines the speed and reliability of high-end commercial databases with the simplicity and cost-effectiveness of open source databases. Amazon Aurora MySQL delivers up to five times the performance of MySQL without requiring any changes to most MySQL applications.Amazon RDS manages your Amazon Aurora databases, handling time-consuming tasks such as provisioning, patching, backup, recovery, failure detection and repair. You pay a simple monthly charge for each Amazon Aurora database instance you use. There are no upfront costs or long-term commitments required.

**Amazon Web Services (AWS) is MySQL compatible**

It means that most of the code, applications, drivers and tools you already use today with your MySQL databases can be used with Aurora with little or no change

For public data repositories, metadata will be generated as above, and accompanying documentation will be published in the appropriate form for the medium (e.g. Readme.md file for Github repositories, online supporting information file for publications associated with data repositories).

AWS is a wholesaler that delivers global compute(what is compute?) from their cloud data centers at a very low price. They offer services that will run your entire IT environment (compute?, networking, storage, database, application services and management) from the cloud and reduce your on-premise hardware footprint.

Following are the reasons to use AWS:

1. Capable of holding extreme structured data like SQL and big data of Excel (thus just a cloud storage service?)

2. 7 years in the market with hundreds of thousands of customers in over 190 countries running every imaginable use case on AWS

3. Groups of data centers *(elaborate more on what does a data centre do why its beneficial)*, which it calls "regions," on the East and West Coasts of the U.S., and in Ireland, Japan, Singapore, Australia and Brazil; it also has one region dedicated to the U.S. federal government

4. Attained most industry standard compliance certifications: HIPAA, SOC 1/SSAE 16/ISAE 3402 (formerly SAS70), SOC 2, SOC 3, PCI DSS Level 1, ISO 27001, FedRAMP, DIACAP and FISMA, ITAR, FIPS 140-2, CSA, MPAA(add a few words why these standard compliance helps with)

Beside all these pros there are some cons for using AWS.(needs better english grammar)

* With AWS one has to setup an account with very low cost
* Billing is extremely confusing;The website recommends going through a reseller for a more detailed monthly bill

Reference <https://npifinancial.com/blog/pros-and-cons-digging-into-amazon-web-services/>

(add plan 3 for datavers by Dalhousie and mention that it is good for preserving and sharing data. This is the most the suitable solution for her because it covers all of her problems (open the site Ychele sent u)

*( “Dalhousie University has a hosted version of Dataverse (v.4.7) for researchers to deposit and share data openly or privately. The data is hosted on Dalhousie's servers. The service is primarily for those affiliated with Dalhousie University.*

*What are some benefits of using Dataverse?*

*Dataverse allows you directly deposit data, share data either openly or privately and discover data. Dataverse can help you receive web visibility, academic credit and increased citation counts. It is also an option to help satisfy research data management plan requirements.” )* All this information should be included in the paragraph, and explain in steps, how the excel data will be inputted, and how professore colleagues will not have to email the professor anymore for permission or requesting certain data via email, because there is data discoverability benefits with dataverse. The proffesor can also preselect which data will be private and which will be public)

**Selection and Preservation**

Prof. Pinkerton has been collecting data for the past 10 years, this dedication proves that preservation is essential part of the plan. Thus, this data has a long term preservation time period and needs to be preserved for at least another 10 years. A good recognized stable long-term repository should be considered such as DalSpace. [DalSpace](http://dalspace.library.dal.ca/) is an institutional repository that collects, preserves and distributes digital content produced by members of Dalhousie University. It is suitable for documentation related to research datasets, such as journal articles. It can also be used to preserve final datasets (those that will not require changes or editing). It is also another appropriate means for sharing these saved data online (DalLibraries, 2018).

Another good stable repository to consider is one that offers curated services such as Dryas Data Repository. This is a good service, because the owner will be able to pass on their data into the hands of a professional data curator, where their main job is to continue to preserve the previous owner’s data (Dryad, n.d.). The data is selected based on Prof. Pinkerton’s interest and all the data should be retained.

Therefore, once Prof. Pinkerton decides it's time to pass on her data she can choose one of the above options, but until then the data will be continue to be preserved by both Prof. Pinkerton and her chosen colleagues.

**Data Sharing**

The data is categorized into two categories:

1)Open Data: This will be shared with all the users.

2)Closed Data: Is shared upon request with colleagues.

Access control can be done using any of the plans offered previously. If DataVerse is selected, the owner will have complete control in deciding which data is private, and which is publically shared. This will help to eliminate answering and keeping track of all data sharing requests received through email (DalLibGuides, 2018).

On the other hand, OneDrive allows data sharing with explicit granted permissions to desired team members. There is a “Share” button for every user, that when it is clicked a share request is sent, and the shared data will be seen in the “Shared with me” section (Dalhousie, n.d.).

As for the AWS plan, it offers a complete access control for the owner. In addition, the owner’s public data sets can be shared via sharing programs such as Amazon EC2, Amazon Athena, AWS Lambda and Amazon EMR. Users can download either one of these programs and use the search interface to locate their desired data they are searching for (Amazon, 2018).

**Responsibilities and Resources**

Our team is responsible for transfering the data. Then, after the completion of this project the responsibility is delegated to the researcher team (postdoctoral).

**Budgeting**

**Hardware**:

This project needs a hard drive and a laptop.The cost for hardware for data management plan is covered by the researcher funding.

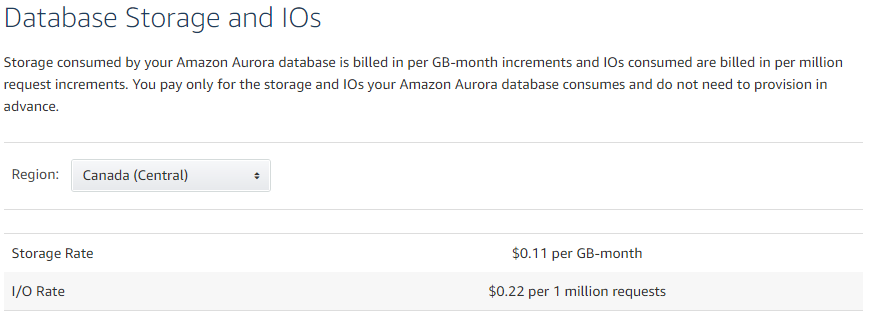
**Human Resources :**

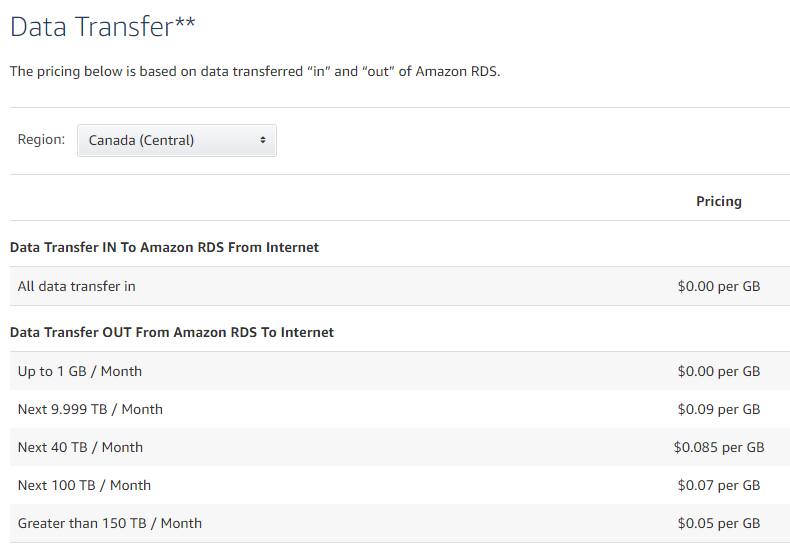
As stated in the case, Neil Gaiman is a colleague of Professor Pinkerton, and he is helping in creating the file structure for this project. This project needs an IT consultant as well, to connect the database to the cloud server, and help in moving the data properly.

**Cloud :**

Dalhousie OneDrive, and DataVerse are institutionally funded, but as it might not support structured data at times, for which Amazon Web Services is recommended. The cost for AWS varies with slight change of the size of data. There is no upfront commitment with Amazon Aurora; you simply pay an hourly charge for each instance that you launch.

(Its better if our team chooses the appropriate prices for her based on her needs (one or two options), and justify your reasoning - put it in a table)





**Case-4:**

**Admin Details**

**Project Name:** Professor Chartreuse Data Management Plan

**Principal Investigator / Researcher:** Professor Chartreuse

**Consultant Company:** Data Masters

**1. Data Collection**

The type of data that is collected in this case is open data, such as data about science of science, which is mostly available from public domain resources. The important collection of data (i.e. medical data) which is collected from database called PubMed, is available via web API and is served and collected in JSON file format. The plan is to keep the JSON format and by pass the Excel sheet format and directly insert (query) the JSON files format into a SQL relational database (Contributors, 2018). By doing this Prof. Chartreuse can keep querying his own inputs as he like and offers less confusion, as to the whereabouts of the data type, and it facilitates data sharing. Furthermore, JSON file format will allow for the data to be reused, and combined with the database, it is easily shared with accessed control in the SQL Server, and is able to be preserved for a long time in the SQL relational database Server Cloud.

For this plan we recommend Microsoft Azure SQL Database, which is an intelligent relational cloud database service. The main advantages this database offers is, it can handle big data, it supports JSON objects and XML objects, cloud server storage and backup for file loss prevention, and version control when data sharing will become necessary (via source control) (Microsoft, n.d.). The fees for this option is displayed in the budget plan proposal.

Another option is using the University cloud OneDrive, it does not offer the same advantage of the structured organization of the SQL database, but it is free, and is an institutional OneDrive cloud that offers storage, collaborative work through Office Documents, access space via the web, personal storage with syncing, data sharing, and reviewing the version history of a file. Furthermore, it offers group storage which is separate of the users own storage, sharing via permission and collaborative editing. This will help with storage and backup, as it offers up to 1TB space for every faculty and student member of the university (Dalhousie, n.d.).

1.1. Public Data Collection Optimization

There are two ways to optimize collection of public data, especially in PubMed, instead of inserting the data manually into Excel. The first method, is to learn the process of Web API public data downloading in JSON format, then to save the JSON files, back them up and input them directly into Azure SQL Database via querying. This is a more technical way and will take our team a while to familiarize with, but it offers a more efficient way of handling the data into a database, and this will lead to a more consistent and structured organization.

The second way is much simpler. In PubMed, after keyword search, a list of results appears. Next to the articles a checkbox is located, and the user can select which articles index information the user wants to download, then the user must go to the ‘Send to’ drop down menu, up in the right corner, and select the appropriate output file format they desire. Choices ranges from TXT files, CSV files, XML files, etc. This is much simpler than doing it by hand; however, this might still lead to a messy excel sheet if CSV file is chosen as an output. If XML format is chosen, it will be easy to query it into the database just like the JSON file format.



***Figure 1:*** *Visual Process of simpler method of public data collecting in PubMed (NCBI, n.d.)*

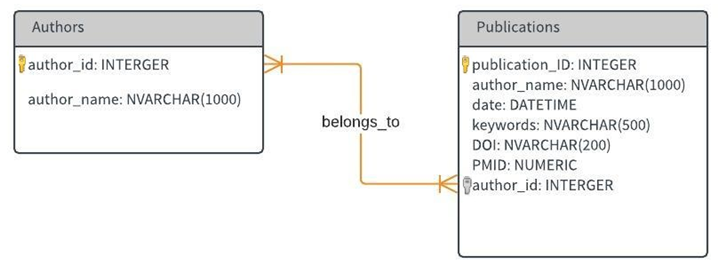
1.2. ER modeling Naming Convention

In order to structure, name the files, and to better understand how the data needs to be organized, an Entity Relationship modeling schema naming convention and procedure is used. This model helps to give a visual layout on how best to organize and label the Metadata in the database. The Metadata labels that is used are date, name, keywords, publication, authors, DOI and PMID. Due to many of the important sources originating from PudMed, we recommend that their unique document identifier label is used (i.e. PMID). DOI (i.e. digital object identifiers) which helps to identify journal articles and scientific articles, is also a Metadata standard that we recommend that is used, as this meta area of research is comprised of science of science.

1.3. Version Control

As this project takes into consideration future collaborative work opportunities with other colleagues, version control is needed. However, for databases this is different. For SQL databases this is called SQL source control. SQL source control is a software, where it is installed and connected with the user’s database (Dave, 2016). After installation the source control keeps track of the database, the transient - which is a local copy of the latest revision of database, and the working base - which is a copy of the database the last time the user ran a commit or pulled from the database (Redgate, 2016). Then a three way comparison is performed and a list of changes is then generated. Here is a list of free recommended SQL source version control that is compatible with Azure; Flyway and Visual Studio database project. Here is a list of commercial priced source controls; Red Gate, ApexSQL source control, and DB Ghost Change Manager (Dbmstools, n.d.). Depending on Prof. Chartreuse’s needs our IP’s will install the one that fits perfectly with his needs.

**2. Documentation and Metadata**



***Figure 2:*** *ER Diagram Schema with Primary and Foreign Key restraints*

**3. Ethics and Legal Compliance**

The most important sources derives from PubMed, and our PI’s familiarized themselves with

PubMed licensing agreement. This data management plan will follow these summarized

Copyrighted, Policy and guidelines, which was gathered from PubMed website:

Copyright and Restrictions (PubMedHealth, n.d.):

1. There are both copyrighted and non-copyrighted, and creative commons materials found on PubMed. It will be up to the user to figure it out. But, when in doubt assume it is copyrighted and contact publisher for permissions. To find out those that are covered by creative commons license can check PMC Open Access Subset.
2. Materials produced by U.S. government employees as part of their official duties in not copyrighted within the US.
3. Works are commonly protected by U.S. and international copyright laws, if they are created by organization other than the federal government.

Restrictions on Systematic Downloading of Articles (NCBI, n.d.):

1. The NCBI`s systems may not be overloaded, by numerous queries.
2. Bulk downloading is prohibited.
3. Crawlers and similar robotic method are not allowed to be used to systematically retrieve content from PubMed web site.
4. There are only two services allowed PMC OAI service and the PMC FTP service are permissible for automated downloading of articles.

Endorsement (NCBI, n.d.):

1. No commercial use.

**4. Storage and Backup**

The Azure SQL Database package comes with cloud services for storage and backup for the database. Thus, the database will be stored and backed up onto Microsoft cloud.

For the research data, which is in JSON format, should be stored and backed up before inputted into the database and multiple backups is needed for this type of project. First the JSON files should continue to be backed up on Prof. Chartruese’s external hard drive. However, this is not enough security against file loss prevention, thus another appropriate cloud storage services is required, and Open Science Framework (OFS) cloud service seems to be a good choice, as it is free, supports JSON objects, can easily archives one's data, offers controlled access and easy collaboration, amplifies one’s workflow, offers version control, and can easily manage one’s project by using his or her ‘one dashboard’ (OFS, n.d.). OFS also offers openness to view what fellow researchers are up to. The user can also choose which data to share and which to maintain private. Another option to consider is the University’s cloud servers, but it might not offer the same openness as OFS, but does offers more storage then OFS.

As this is data collected from public repositories there are no sensitive data to worry about, but Azure Database and Cloud service offers the optimum security with required authenticated access. OneDrive and OFS offers strong security as well.

Both database and research data should be backed up on a weekly basis by Prof. Chartreuse and both OFS/OneDrive and Azure offers sufficient storage, for Azure the first 12 months offers a free package size of 250GB for the database and for OFS its 5Gb on initial registration, however much larger size can be obtained by emailing their support team. If this cannot be done there is always OneDrive, which offers 1TB of space.

**5. Selection and Preservation**

Prof. Chartreuse is very dedicated to his research and plans to continue his work until the day he is no longer here. Thus, this database and research data has a long-term value of at least minimum of ten years and needs to be preserved well. Prof. Chartreuse will continue to control and preserve this data until he sees fit to archive his freely available data on a recognized stable long-term repositories such as Github, OFS, or data repositories that offers data curated services such as institutional repositories or Dryad Data Repository (Dryad, n.d.). The data that must be retained are those that leads to publication and records are selected based on the records of interest to the professor.

**6. Data Sharing**

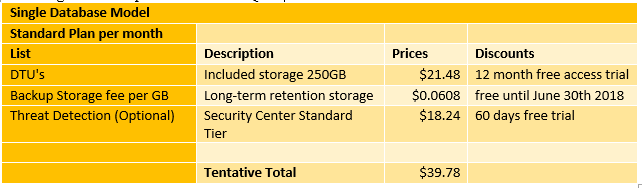
As the data is collected from public sources such as PubMed, their licensing policy must be upheld. Their main policy about data sharing, is that it may not be used for commercial purposes. Their copyrighted and creative commons policy must also be maintained as discussed. This data will be shared with graduate students and fellow colleagues in the same field as Prof. Chartreuse. The professor decides who to give access too. Data sharing will be possible through both Azure SQL Database Cloud Server and OSF cloud server and their search engine. If colleagues wishes to collaborate with Prof. Chartreuse on his Azure SQL database, then they will have to contact the professor first, with their necessary contact information such as name and email, and then he will proceed to make a user credential for the colleague in Azure, with his chosen set of restriction parameters, and then the colleagues will be prompted to choose their username and password.

**7. Responsibilities and Resources**

Our consultant team is responsible with familiarizing themselves with the Web API method of downloading JSON format file in order to input directly these JSON files into to SQL database by querying the input, and then to relate this information to Prof. Chartreuse. In addition, we need to be familiarized with all the copyrights and licensing policy while working on this project.

Resources are provided by open-course providers such as OSF. University funding for Azure SQL Database is conditional on a well-documented reason for the cost.

7.1. Budget Plan Proposal for Azure SQL database



***Figure 3:*** *Monthly Budget Plan (Microsoft, n.d.).*  
Azure SQL Database offers either a standalone database or elastic database pool, and these are priced in 3 tiers; Basic, Standard and Premium.  
DTU (Database Transaction Units) is defined as a melded measure of CPU, memory and data I/O and transaction log I/O in a ratio. This is benched marked by OLTP workload.   
This standard plan is recommended by our team, which is the go-to option for cloud application. It includes support for concurrent queries, collaboration and web application options.  
In the end, Azure SQL database offers advantages such as bulk data migration, migrate SQL-based applications, simple user interface, security, data visualization, controlled access and collaborative work, handles big data, supports JSON and XML objects, Cloud server, storage and backup for file loss prevention. Many of these advantages are not all found in one free database, and with the provided discounts and free trials that are happening right now, it is the best chance to try this plan out now while it's still free.

**References APA format:**

Contributors. (2018, February 19). Work with JSON data in SQL Server. Retrieved April 06, 2018, from<https://docs.microsoft.com/en-us/sql/relational-databases/json/json-data-sql-server>

Dalhousie. (n.d.). Cloud File Storage with Microsoft OneDrive for Business. Retrieved April 06, 2018, from<https://www.dal.ca/faculty/computerscience/current/onedrive.html>

Dave, P. (2016, September 04). Team Database Development and Version Control with SQL Source Control. Retrieved April 06, 2018, from<https://blog.sqlauthority.com/2016/04/20/team-database-development-version-control-sql-source-control/>

Microsoft. (n.d.). Your vision. Your cloud. Retrieved April 06, 2018, from<https://azure.microsoft.com/en-ca/>

Dbmstools. (n.d.). 4 Version Control tools for Azure SQL Database. Retrieved April 06, 2018, from<https://dbmstools.com/version-control-tools/azure-sql-database>

Redgate. (2016, May 18). SQL Source Control 5. Retrieved April 06, 2018, from<https://documentation.red-gate.com/soc5/reference-information/how-sql-source-control-works-behind-the-scenes>

NCBI. (n.d.). PMC Copyright Notice. Retrieved April 06, 2018, from<https://www.ncbi.nlm.nih.gov/pmc/about/copyright/#publisher-links>

Pcloud.com. (2018). *Cloud Storage Pricing & Cost, Cloud Storage Plans | pCloud*. [online] Available at: https://www.pcloud.com/cloud-storage-pricing-plans.html [Accessed 9 Apr. 2018].

PubMedHealth. (n.d.). Copyright - National Library of Medicine - PubMed Health. Retrieved April 06, 2018, from<https://www.ncbi.nlm.nih.gov/pubmedhealth/copyright/>

NCBI. (n.d.). Policies and Disclaimers - NCBI. Retrieved April 06, 2018, from<https://www.ncbi.nlm.nih.gov/home/about/policies/>

OSF. (n.d.). Open Science Framework. Retrieved April 06, 2018, from<https://osf.io/>

Dryad. (n.d.). Dryad Digital Repository. Retrieved April 06, 2018, from<https://datadryad.org/>

Amazon. (2018). Large Datasets Repository | Public Datasets on AWS. Retrieved April 06, 2018, from <https://aws.amazon.com/public-datasets/>

DalLibGuides. (2018). Research Data Management: DataVerse. Retrieved April 06, 2018, from <http://dal.ca.libguides.com/rdm/daldataverse>

Sync. (2018). *Plans & pricing | Sync*. [online] Available at: https://www.sync.com/pricing/ [Accessed 9 Apr. 2018].

DalLibraries. (2018). Research Repository (DalSpace). Retrieved April 06, 2018, from <https://libraries.dal.ca/research/digital-initiatives/digital-publishing/repository.html>