MEMO

1. Write a convincing memo (650 word max) explaining why data curation services are important. Assume that the memo is written for your new director, who is not familiar with data curation, and not convinced whether to keep funding this work. You will want to make sure to introduce data curation within the broader context of data science. You will need to cover the key areas that you think are the most important for data curation at your company. We ask that you incorporate at least two of the following topics into your memo: Provenance, Policy, Metadata, and/or Preservation.

Why is Data Curation Services Important

Data curation is a means of managing data. In current technology era, data is almost everything. Every electronics device or a software solution deals with understand the data as input and producing output in terms of data in one way or the data. It has become very crucial to understand various forms of data. Data representation at a conceptual and implementation doesn't have any concrete form of representation or any guidelines to follows, data management has become one of the critical aspects of technological advancements. Hence having a effective and efficient data curation services has become important to pave way for cutting edge technologies.

Data Curation starts with "data set". These data sets are fundamental building blocks for data curation activities. Broadly the stages of data curation can be listed as Identifying, Cleaning, Transforming. Identifying is the process of identifying different data sets or data sources that are of our interest. Identifying required data-sets is as important as solving a problem. Cleaning is the process in which we remove not so required data and normalize the data to certain standard that we deem important for certain activity or requirement. Transforming is the process in which we convert the normalized data into the data that is required for a solution solution as input data.

In this memo, I would like to express the importance of Metadata and Provenance in the data curation process.

Metadata is a way of referring to the data that supports the discovery, management and understanding of other data and information. Int other words, it is *data about data.* Capturing and maintaining correct metadata is considered key to the reuse and preservation of digital objects. There are a wide range of metadata schema and standards developed based on the requirements such as DDI (Data Documentation Initiative), IEE LOM (Learning Objects Metadata) and many more.

Provenance is defined as the origins, ownership and custody of a data. Since data-sets are reformulated and reworked to create a new data, provenance is important to trace back the origins of the modified data sets. Every data set during the data curation process need to address the requirements of provenance as it guarantees the data creators accountability for their work and provides a chain of information where data cane be back-traced to the creators enabling the researchers and others to adapt the data for their own purposes.

I believe that a thoughtfully designed data curation process and service provided by this process are one of the building blocks for technological breakthrough in future. I strong urge your to continue funding this work to head towards more pivotal changes and process established for enriched data curation services that are going to be the stepping stones in future.