Anthony Camacci

INFO 3401-001

Problem Set One

1. This quote was a bit confusing to me, but to my understanding it means that the people that are collecting data shouldn’t be the ones finding a way to represent the data. I’m not sure why he would want to frame data science in this way, but my guess would be that the person collecting the data and finding a way to represent it shouldn’t be the same person. I think this statement is problematic because I think that if you are collecting data you should also know how to present it. You may have insights on the data that you collected that someone else trying to make sense of the data may not have.
2. The biggest thing that stood out to me between these definitions is the specificity that has been added throughout the definitions. The first definition is very vague in comparison to the second and third definitions. The very first definition is very broad and basically says data, the internet and applications. Compared to then, we know have many specified fields for many different types of technology, data and applications. There are many more related fields mentioned in the last quote including engineering, annotators and many more.
3. Three technological factors that contribute to the growth of data include the creation of computers, cellphones and the internet. Three major sources of data that have helped contribute to the growth of data are data from Google, Amazon and the NSA.
4. If you change your directory to the root, you can see all the parents of all the files, lets you know what exactly is in your system and is known as a common reference point. Allows you to say you know how to find things such as users. You can use cd.. to move up to a parent. ./name allows you to access a certain file within a directory. cd~ takes you to your home directory.
5. cd ~: Stands for change directory, changes it to whatever directory name comes first. Can change to a specific file if you put the name after cd. After changing cds, you can use pwd. The tilde takes you to your home directory.

mkdir ./problem\_set\_1: Shows the list of files that are in the current directory. There are other special commands that can be done with dir such as listing hidden files.

cd ..: Changes the current directory to the parent directory, takes you back to previous directory.

pwd: Stands for print working directory, tells you what directory you are currently working from. Helpful because you can’t always be sure what directory you are working from and gives a sense of what files are available for immediate use.