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.