

Research in regard to data dictionaries provides so much insight to not only the information we are able to obtain for various reasons, but what can go wrong if the data isn't managed correctly. A data dictionary is a descriptive list of the names, definitions, and attributes of data elements to be collected in an information system or database whose purpose is to standardize definitions and ensure consistent use.

To establish a data dictionary for healthcare several things must happen. You would want to work with IT security teams as well as data scientists to make sure the correct metadata is going into the correct fields. Establishing, maintaining, and ensuring the information provided is correct and will not be misused is within the utmost importance within maintaining a data dictionary. Within my research I also want to see if I can find where the data was not accurate and the results of that when the issue was corrected. Information provided in the dictionary should speak for itself and with that understanding we have to know exactly what information should be pulled and left out otherwise the information could cause several issues that cost the healthcare industry money.

The Data dictionary was first established as a concept of providing a means to support clinical information systems pulled from clinical information systems. One of the main issues at that time was terminology. The terminology had to be consistent to make sure the data scientists building the IT infrastructure would match what the medical record was providing. One great aspect of the data dictionary is that it is built by several people, and that keeps more eyes on the information to make sure it is being built ethically. The other great part of the dictionary is that it really is composed of numbers, values - and we pull the information we need to build the research review and information needed to

verify different aspects of healthcare. For example, the hospital that I was working at previously here in Georgia had an Emergency department that was always hitting capacity. Due to one reason or another we couldn't move the patients from the ER onto the units fast enough due to several reasons. I was fortunate enough to be part of a meeting that asked several things of the business analysts. 1. Why can't patients be moved onto the units fast enough? They provided information that for instance the medical unit wasn't discharging patients within a time frame and this would cause a waterfall effect. The business analyst team was able to provide us with time frames of data that gave exact times that the discharge was written, and when the patient was finally discharged from the unit within the computer systems. All this information was available within metadata in the hospital systems that could be utilized to show if we could get patients out of the unit faster when being discharged then it would reduce the amount of time a patient is waiting in the ER to be moved to the unit they are being admitted to. From there, using the information provided within the data dictionary that could be pulled, each unit director was able to plan to overcome the Emergency room hitting capacity as often. The overall question was, does the ER need more beds? Should we establish an overflow unit and have ER nurses watch that unit, or can the issues with transfers be solved? Adding onto a unit like the ER would be awfully expensive, but because we have the meta data going into a data dictionary we were able to isolate some important time frames.

One issue that has come up a few times within issues of a data dictionary is that computers can easily process massive amounts of data, but sometimes the data pulled could have no value. In one of the studies I read it explains that sometimes within healthcare, the clinical data could not make sense. For example, "cold". An elderly patient could come into the Emergency room because he has a cold, that could be documented, and the data pulled. Another person could come into the emergency room and one of their symptoms could be that they are cold. That symptom could also be documented to support a diagnosis, but the meaning of "cold" is much different. It is this type of information that could

be pulled if someone were looking to get information on patients that came into the hospital experiencing symptoms of a virus, versus someone coming into the hospital with hypothermia. Luckily, we can isolate other information, or instance Diagnosis codes that will give us more accurate information than symptoms will. This is information that has evolved within the Data Dictionary.

We will continue to see changes and evolution throughout data provided within healthcare. This is a particularly good thing as laws, regulations, and patient empowerment options continue to change. Healthcare technology continues to grow, and with that we have more information available at our fingertips.

References

Ahima. (2012). Managing a Data Dictionary - Retired. *Journal of AHIMA*, 83(1), 48-52. Retrieved 10 25, 2020, from <http://library.ahima.org/doc?oid=105176>

Ahima. (2013). Assessing and Improving EHR Data Quality (2013 update). *Journal of AHIMA*, 84(2). Retrieved 10 25, 2020, from <http://library.ahima.org/doc?oid=106219>

(Ahima, Managing a Data Dictionary - Retired, 2012) (Ahima, Managing a Data Dictionary - Retired, 2012)