1. Question was raised, “What is Informatics & Data Science?” I recommend a book from Microsoft Research entitled, *The 4th Paradigm: Data Intensive Scientific Discovery.* I highly recommend this book if you want to understand how big data has changed the process of scientific discovery.
2. So how did we get to where we are today? How did this emerging paradigm shift happen? Well, it started around 2000 when the first draft of the human genome was announced. The torrent of data has simply accelerated since. Sequencers are more sophisticated and powerful. Sensors are more ubiquitous. Just about everything is producing data!
3. IBM estimates that we’re creating 2.5 quintillion bytes of data each day. 90% of the world’s data has been created in just the last 2 years.
4. So data science / informatics has arisen in just the past couple of years to make sense of all this data. It started in the large Silicon Valley companies, in companies like Google and Facebook.
5. What is data science? On the screen, I provide the “infamous” data-science Venn diagram. As you can see, data science lies at the intersect of code, data, and statistics.
6. And on the left, we see the tools scientists use in each of these areas. These tools are my world. This is where I live, consult, and educate. I instruct scientists (both aspiring and experienced) in the art of creating research systems that then enable them to tell compelling data stories.
7. Key point: Data science has changed the way scientists do their work. These tools are now foundational to the scientific enterprise. And that’s why our academic library has an informatics/data science program.