BIG DATA TRENDS IN 2019 AND BEYOND

Data Curators:

Data curation means ensuring that people can easily find and use the data now and in the future.

- Identify the most relevant data sources
- Sourcing (getting) the data from the data owners
- Fixing any issues with the data such as missing values, unexpected values, and mysterious variables.
- Using annotations, tags, appropriate and crisp documentation, etc. to help make the data easy to find, use, reuse and present.
- Preserving the data such that it is available for use, reuse, etc.

I.e.

https://coronavirus.jhu.edu/
https://www.worldometers.info/coronavirus/
https://datasetsearch.research.google.com/

- Implicitly, job of data custodian is to safe place data.

Data Sources |Data Curation| Data Analysis |Data Translation| Business Decision

-

Data Engineer |Data Curator| Data Analyst | | Decision Maker

DARK DATA

The information assets organizations's collect, process and store during regular business activities, but generally fail to use for other purposes.

Dark Data is subset of Big Data. Dark Data constitutes biggest portion of the large volume of Big Data.

EDGE COMPUTING

Edge computing allows data from IoT devices to be analyzed at the edge of the network before being sent to a data center or cloud.

- Cost Effective
- Faster Response time
- Security and compliance
- Interoperability between legacy and new devices
- Dependable operations with sporadic internet connectivity

CHATBOTS WILL GET SMARTER

Act as personal assistant.

STREAMING THE IoT MACHINE LEARNING

- Healthcare
- Smart Cities
- Transportation
- Automobile
- Retail

CONTAINER(ed)

COMMODITIZATION OF VISUALIZATION

APIs

- Tableau
- PowerBI
- QlikView
- JasperSoft

HADOOP IS FUNDAMENTAL AND WILL REMAIN SO!

Courtesy: Text Book.