NSF Expedition Workshop Panel

Nikunj C. Oza, Ph.D.
Leader, Data Sciencees Group
NASA Ames Research Center
nikunj.c.oza@nasa.gov

Two Facets

- How Earth Science should influence Data Mining
- How Data Mining should influence Earth Science

Earth Science -> Data Mining

Big Data is a misleading term

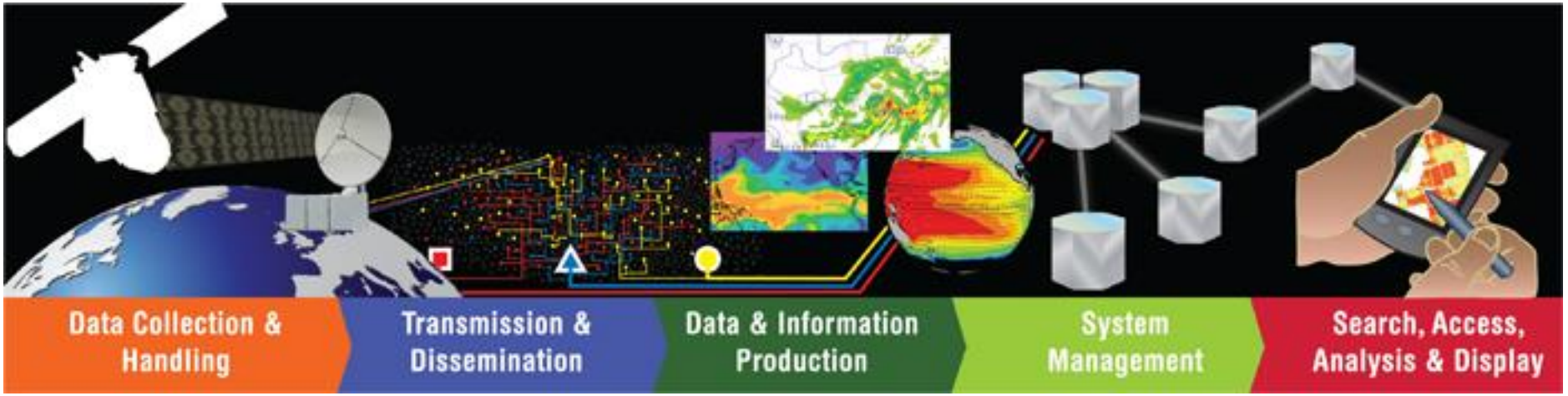
- “Big Data”--- volume, variety, velocity, veracity
- Most people think of volume only.
- Volume is least important part
- Variety: not just numeric. Text, process models, numeric has various resolutions, accuracies, etc.
- Velocity: Online learning data volumes enormous
- Veracity: Varying levels of noise, uncertainty, bias, processing errors

Data Mining -> Earth Science

Earth science still not seen as data science

- Missions need to account for all the data we have.
- We are collecting new data but…
- We have touched < 1% of collected data.

Next 5-10 Years



- Focus on end-to-end process: from raw data to Earth science insight
- And back!
- Integrate data mining with automated planning, reinforcement learning, etc., to aim for practical applications (e.g., farming, aviation)