

WHAT IS DATA SCIENCE?

- “**Data science**, also known as **data-driven science**, is an interdisciplinary field of scientific methods, processes, algorithms and systems to extract knowledge or insights from data in various forms, either structured or unstructured, similar to data mining.”



WIKIPEDIA
The Free Encyclopedia

WHAT IS DATA SCIENCE?

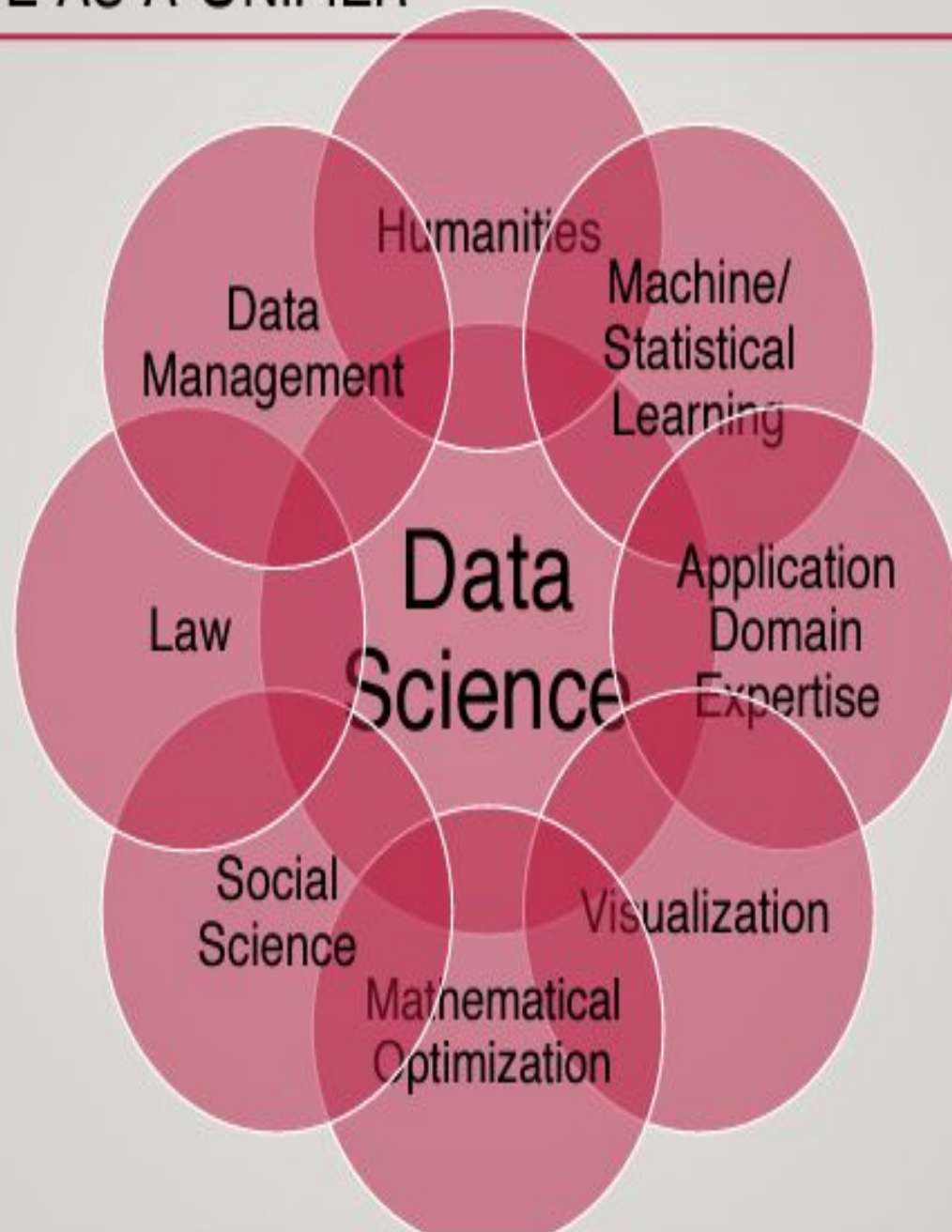
- “**Data science**, also known as **data-driven science**, is an interdisciplinary field of scientific methods, processes, algorithms and systems to extract **knowledge** or insights from **data** in various forms, either structured or unstructured, similar to **data mining**.”
- “Data science intends to analyze and understand actual phenomena with ‘data’. In other words, the aim of data science is to reveal the features or the hidden structure of complicated natural, human, and social phenomena with data from a different point of view from the established or traditional theory and method.”



WIKIPEDIA
The Free Encyclopedia



DATA SCIENCE AS A UNIFIER



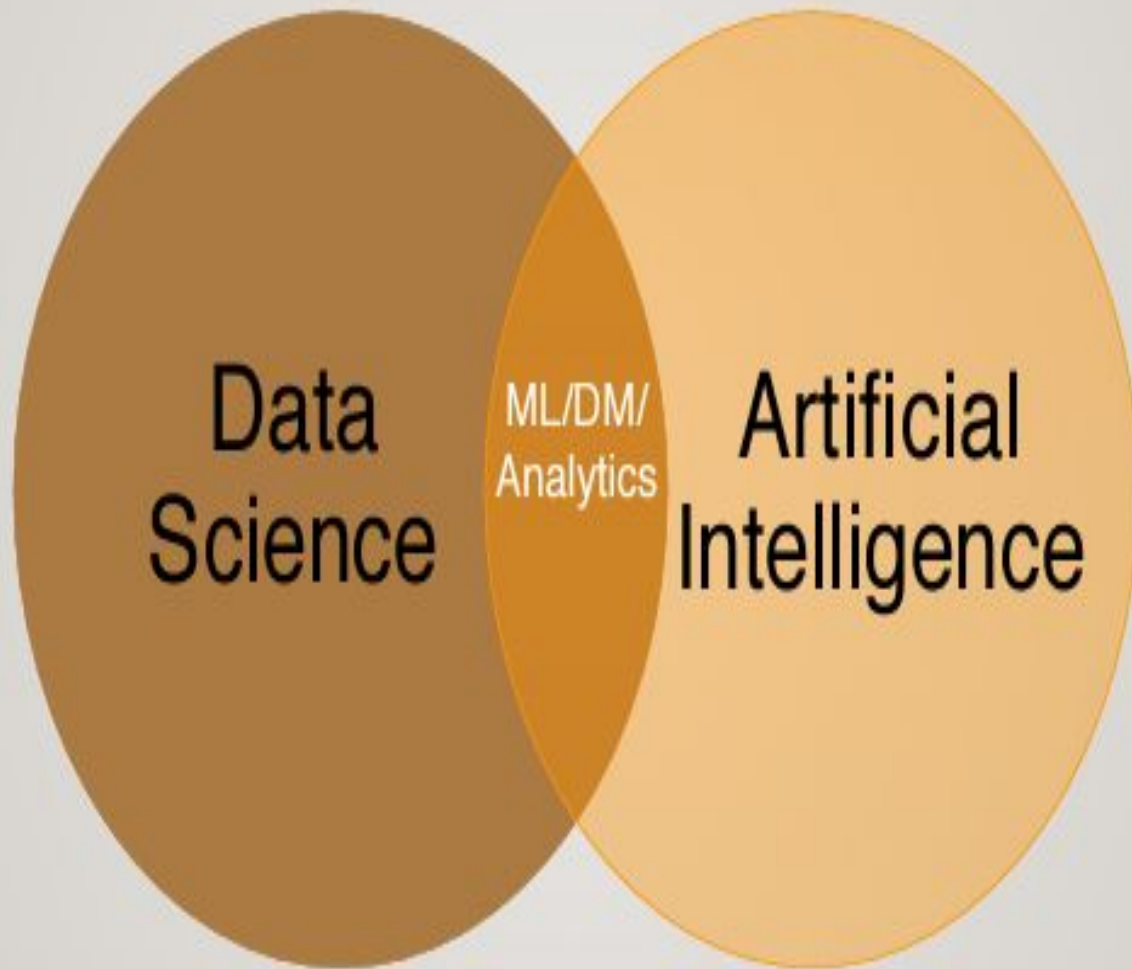
DATA SCIENCE AND BIG DATA

- They are not the “same thing”
- Big data = crude oil
 - Big data is about extracting “crude oil”, transporting it in “mega tankers”, siphoning it through “pipelines”, and storing it in “massive silos”
- Data science is about refining the “crude oil”

Carlos Samohano
Founder, Data Science London



DATA SCIENCE AND ARTIFICIAL INTELLIGENCE



DATA SCIENCE APPLICATION EXAMPLES

- Fraud detection
 - Investigate fraud patterns in past data
 - Early detection is important
 - Before damage propagates
 - Harder than late detection
 - Precision is important
 - False positive and false negative are both bad
 - Real-time analytics



DATA SCIENCE APPLICATION EXAMPLES

- Recommender systems
 - The ability to offer unique personalized service
 - Increase sales, click-through rates, conversions, ...
 - Netflix recommender system valued at \$1B per year
 - Amazon recommender system drives a 20-35% lift in sales annually
 - Collaborative filtering at scale



DATA SCIENCE APPLICATION EXAMPLES

- Predicting why patients are being readmitted
 - Reduce costs
 - Improve population health
 - Find the “why” behind specific populations being readmitted
 - Data lakes of multiple data sources
 - Investigate ties between readmission and socioeconomic data points, patient history, genetics, ...



DATA SCIENCE APPLICATION EXAMPLES

- “Smart cities”
 - Not well-defined



DATA SCIENCE APPLICATION EXAMPLES

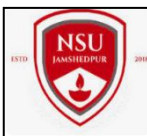
- “Smart cities”
 - Not well-defined



Photograph by Getty / Getty.com / The Atlantic

Stop Saying 'Smart Cities'

Digital stardust won't magically make future cities more affordable or resilient.



DATA SCIENCE APPLICATION EXAMPLES

- “Smart cities”
 - Not well-defined
 - Generally refers to using data ICT to
 - Better plan communities
 - Better manage assets
 - Reduce costs
 - Deploy open data to better en with community



Microtopography / Getty / iStock.com / The Atlantic

Stop Saying 'Smart Cities'

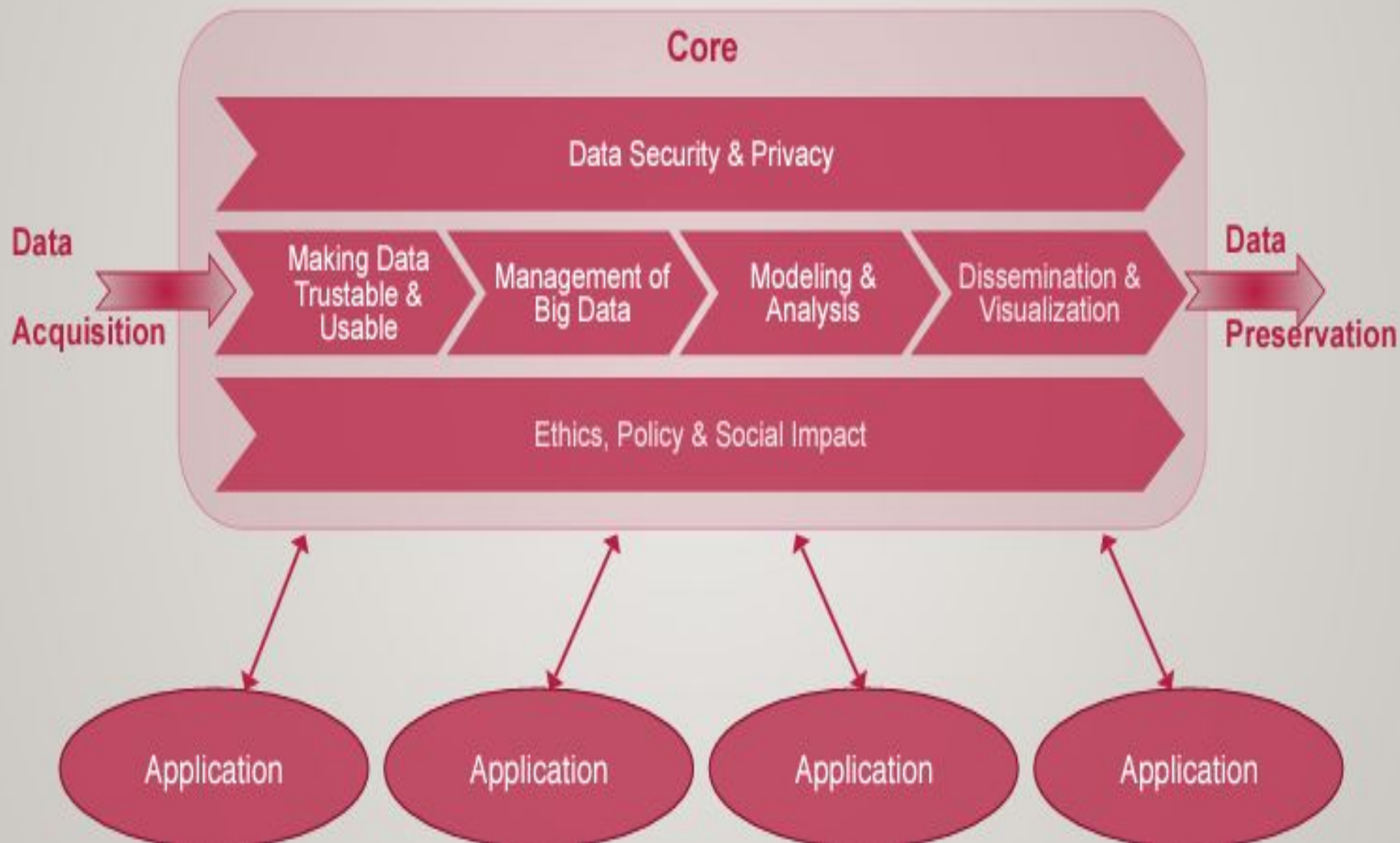
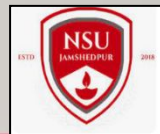
Digital stardust won't magically make future cities more affordable or resilient.

BRUCE STERLING | FEB 12, 2018

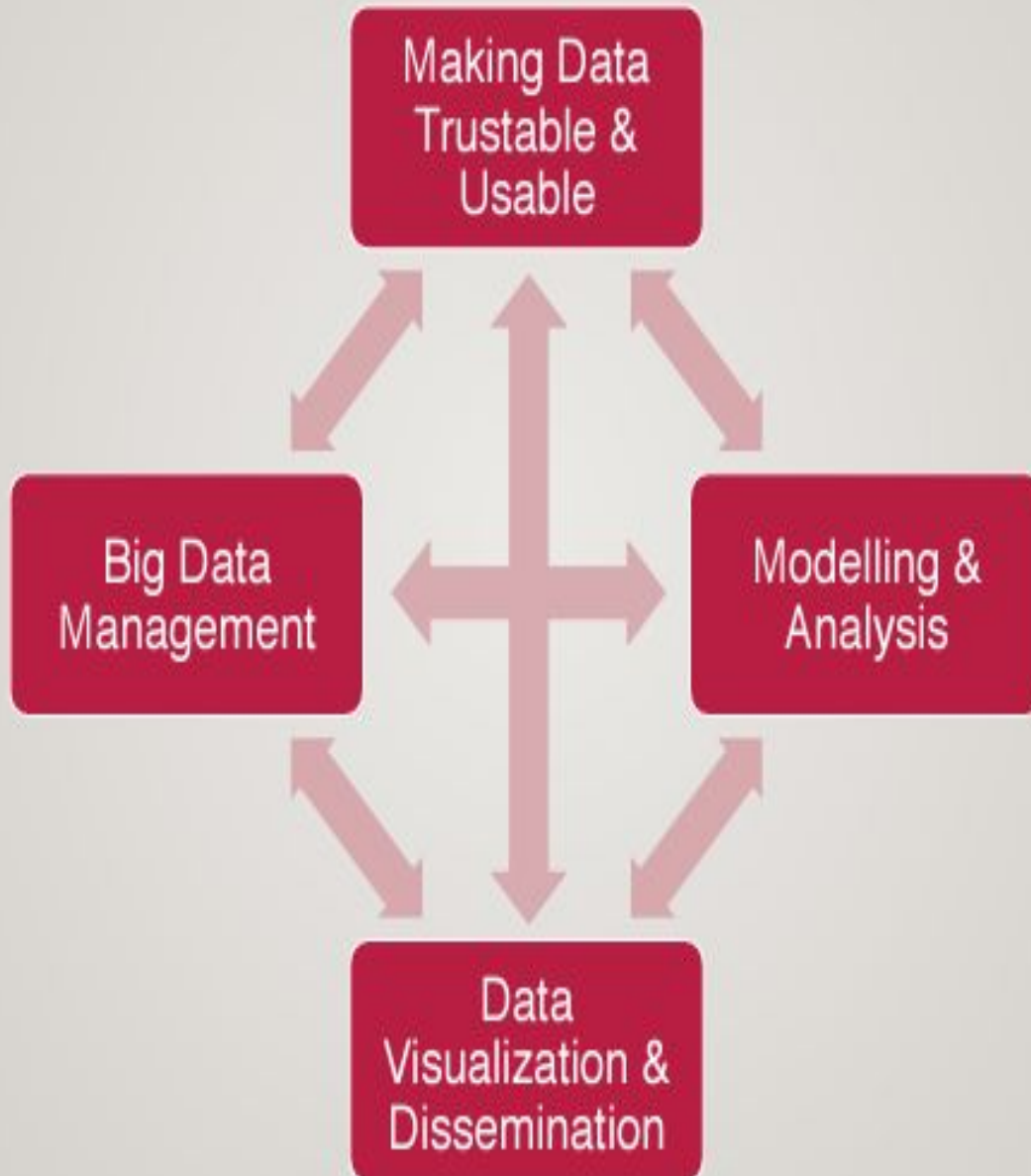
TECHNOLOGY



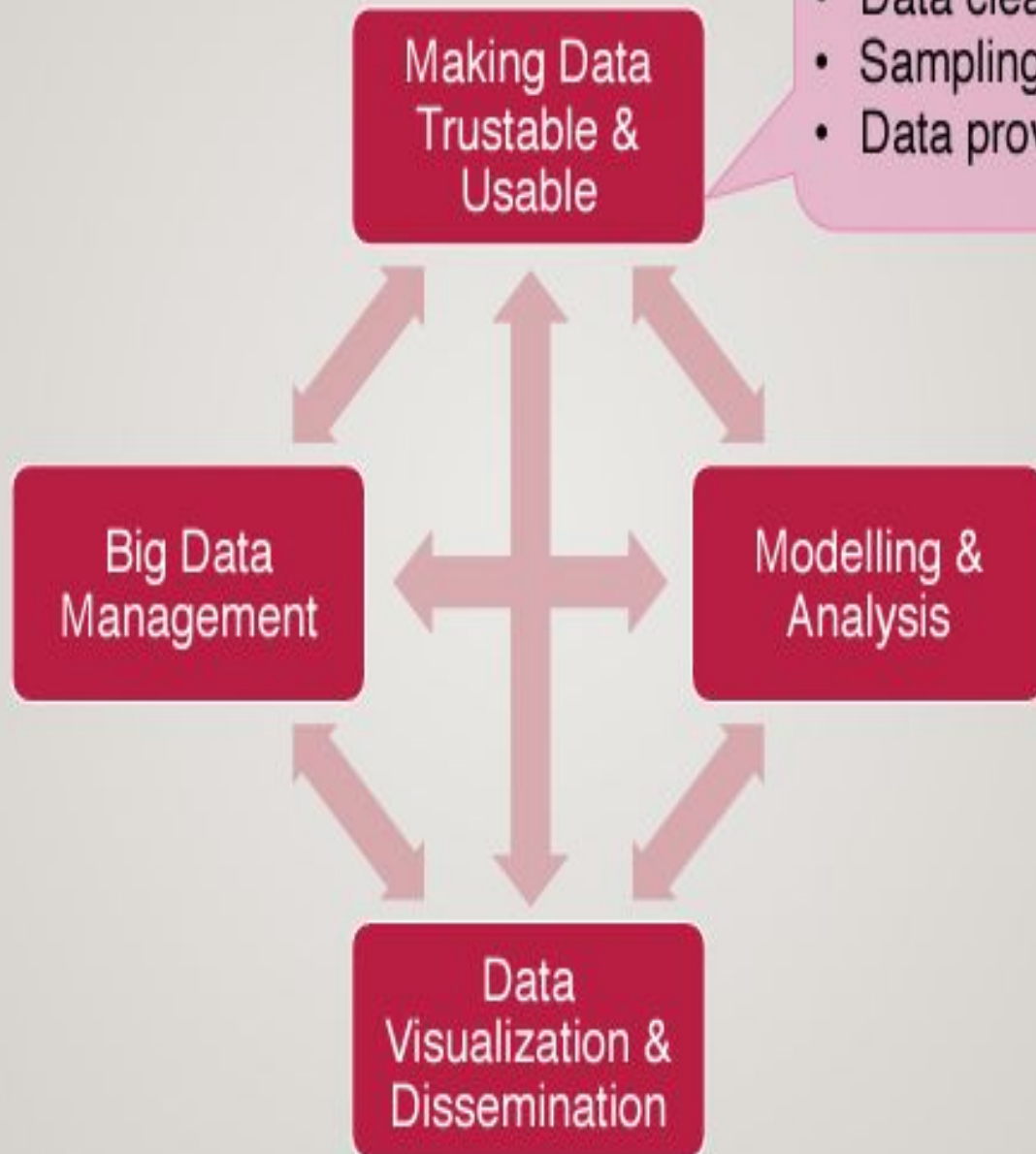
HOLISTIC APPROACH TO DATA SCIENCE



CORE RESEARCH ISSUES & INTERACTIONS



CORE RESEARCH ISSUES & INTERACTIONS



CORE RESEARCH ISSUES & INTERACTIONS

- Data lakes
- Batch & online access
- Platforms

Making Data Trustable & Usable

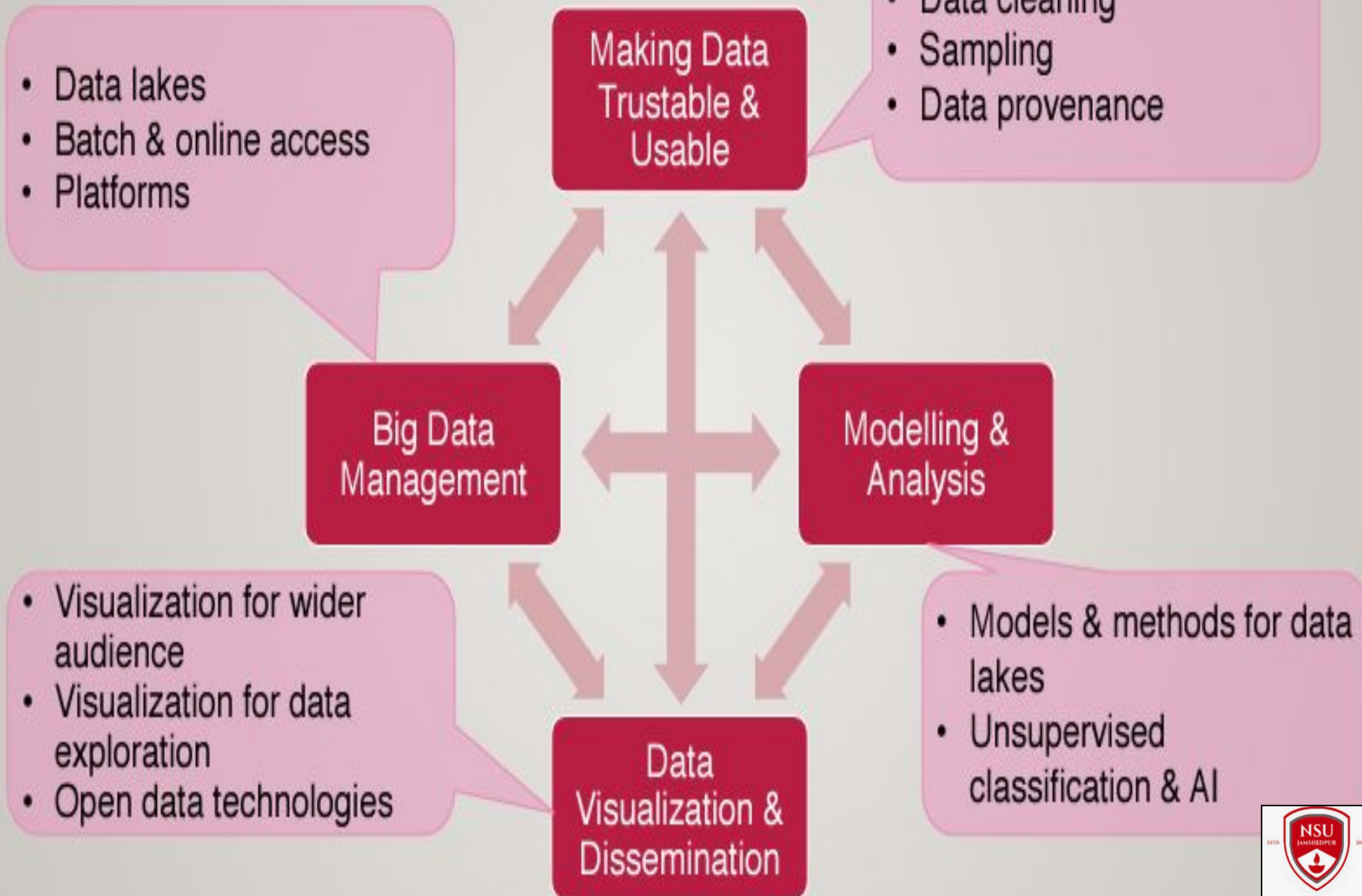
- Data cleaning
- Sampling
- Data provenance

Big Data Management

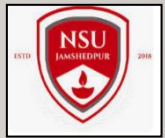
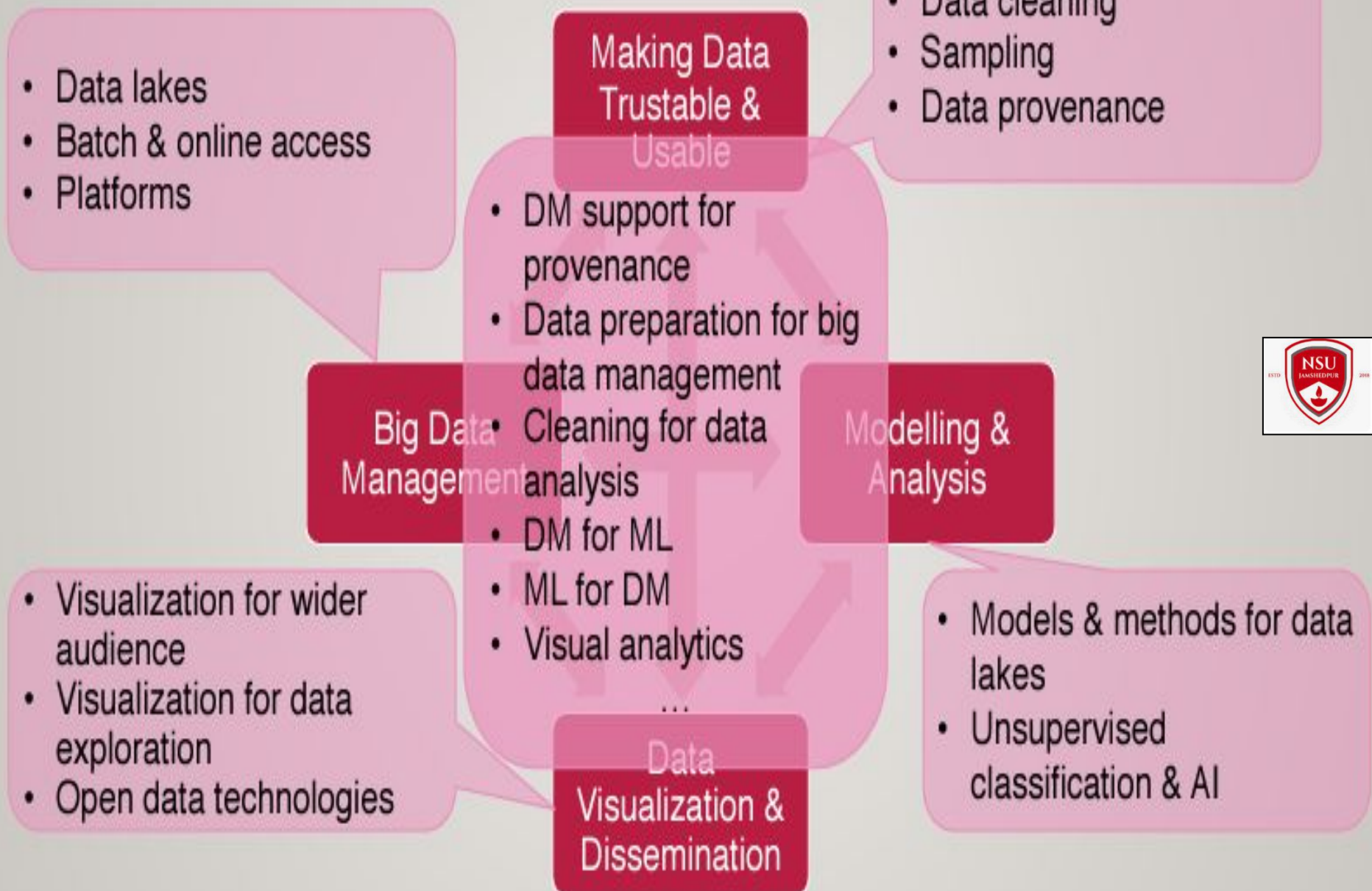
Modelling & Analysis

Data Visualization & Dissemination

CORE RESEARCH ISSUES & INTERACTIONS



CORE RESEARCH ISSUES & INTERACTIONS



THANK YOU .



LIFE AT

INSU

