

DATA SCIENCE

BIG DATA



INSIGHT



ACTION

→ Data Product

ACTIONABLE INFORMATION

HISTORIC DATA + NEAR REAL-TIME DATA



PREDICTIONS



ACTION

STRATEGY

AIM

→ Business objectives

POLICY

→ organizational buy-in

- commitment
- sponsorship
- communication

PLAN

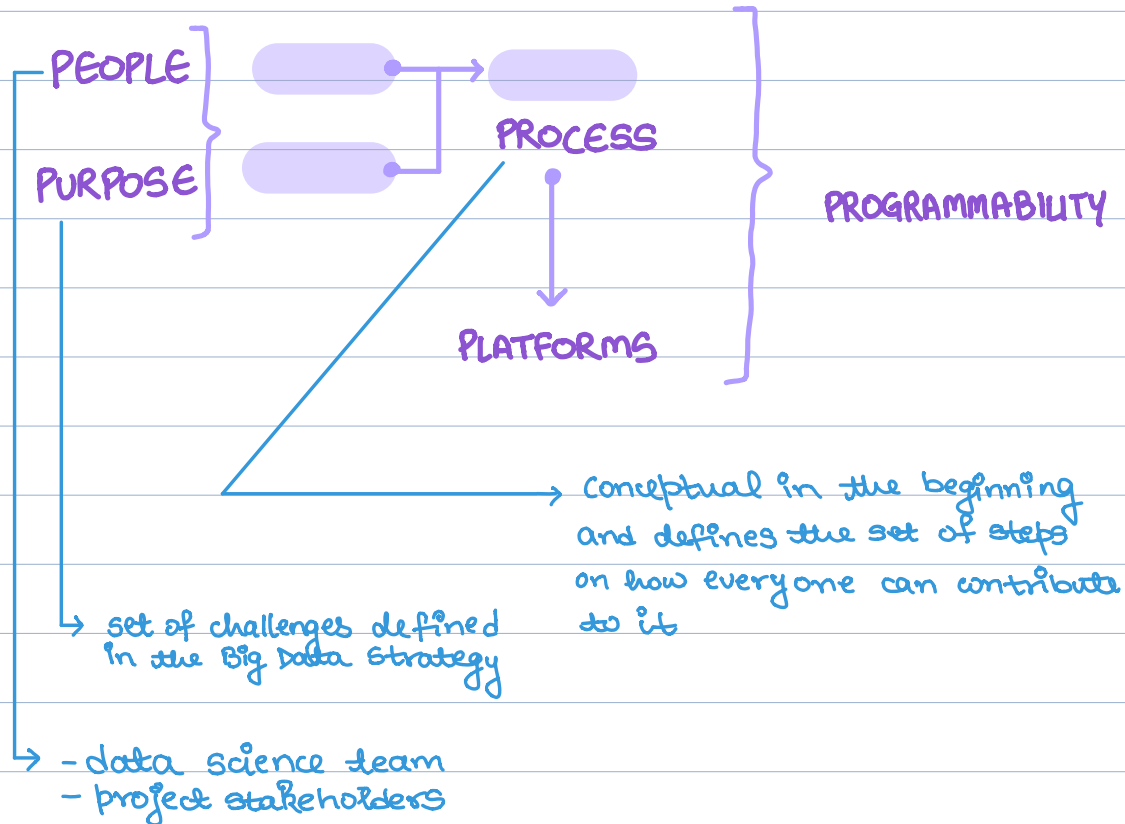
→ Diverse Teams w/ various expertise

→ Training

ACTION

→ ideate & Prototype

FIVE P's OF DATA SCIENCE :

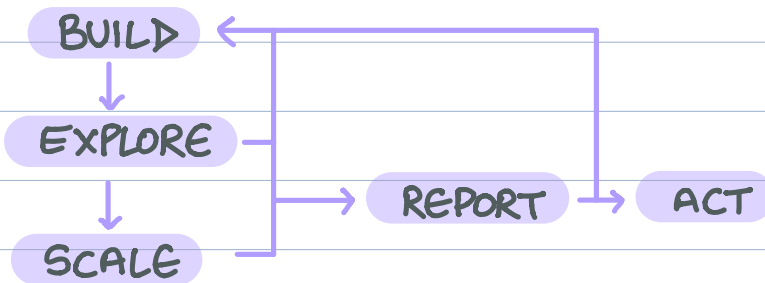


PROCESS

BIG DATA ENGINEERING

COMPUTATIONAL
BIG DATA SCIENCE

ACQUIRE > PREPARE > ANALYZE > REPORT > ACT



PROCESS: Build metrics for Accountability

COST

TIMELINE

PLANNING OF DELIVERABLES

EXPECTATIONS

PURPOSE

STEPS IN THE DATA SCIENCE PROCESS

1. ACQUIRE DATA

- identify datasets
- retrieve data
- query data

2. PREPARE DATA

2.A: EXPLORE

- Preliminary analysis
- understand nature of data

2.B: PRE-PROCESS

- clean
- integrate
- package

3. ANALYZE

- select analytical techniques
- build models

4. COMMUNICATE RESULTS

- Present results in a visual way
- summarize, visualize, post process

5. APPLY RESULTS

- determine action from insights based on the purpose you initially defined.