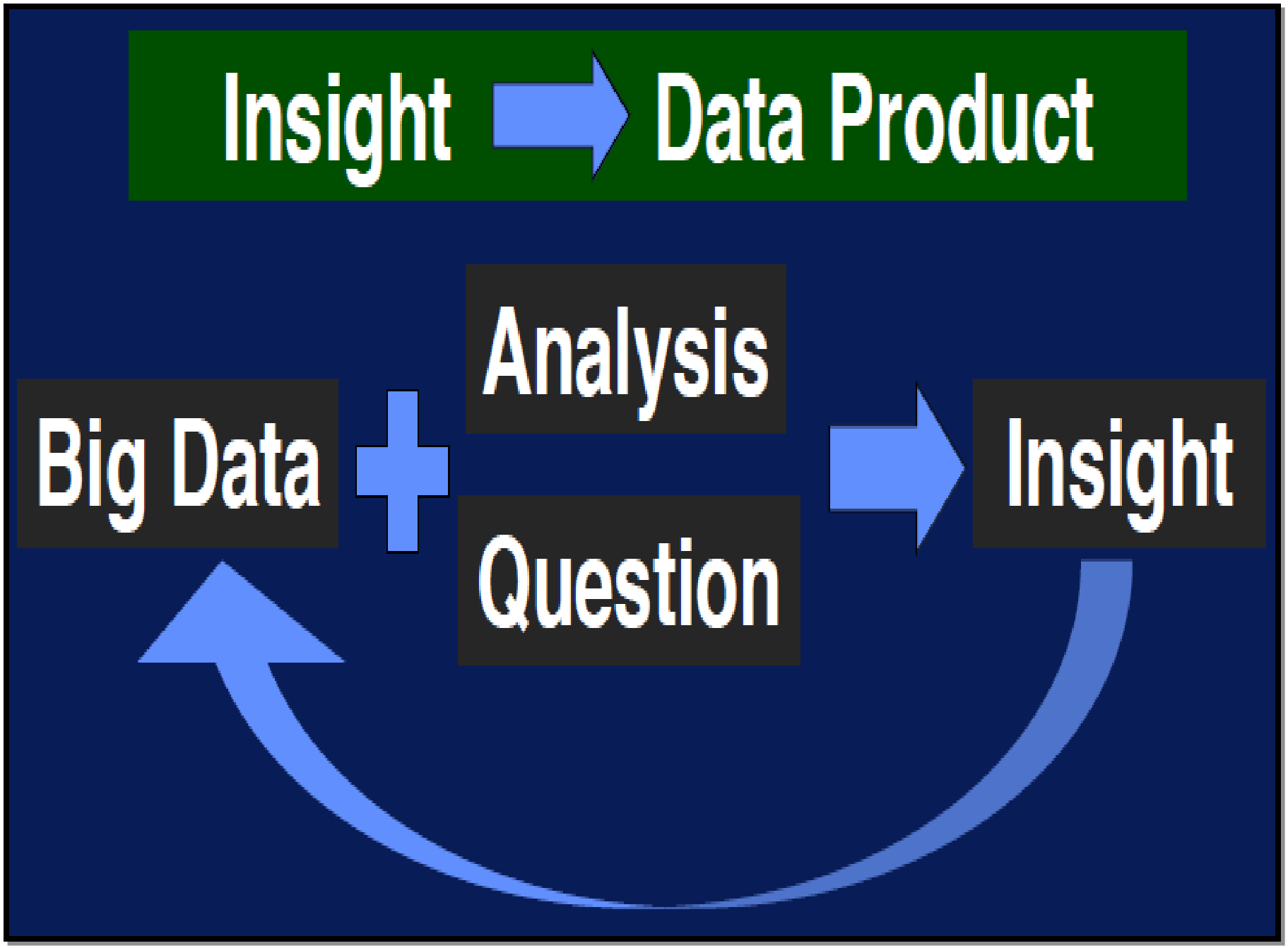
Data Science and Five Ps of Big Data

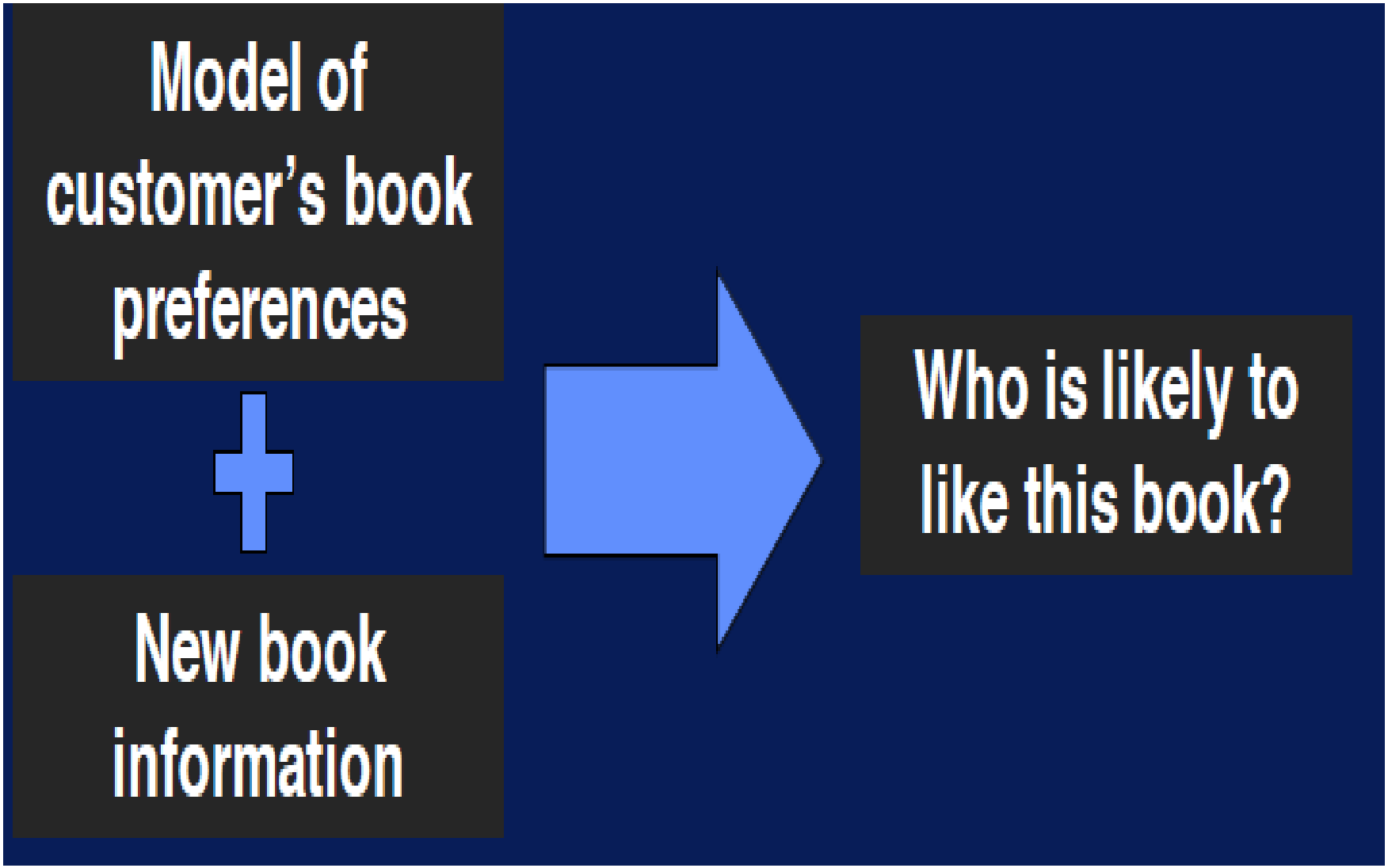
How To get Value out of Big Data


Big Data
Insight
Action
Data Science

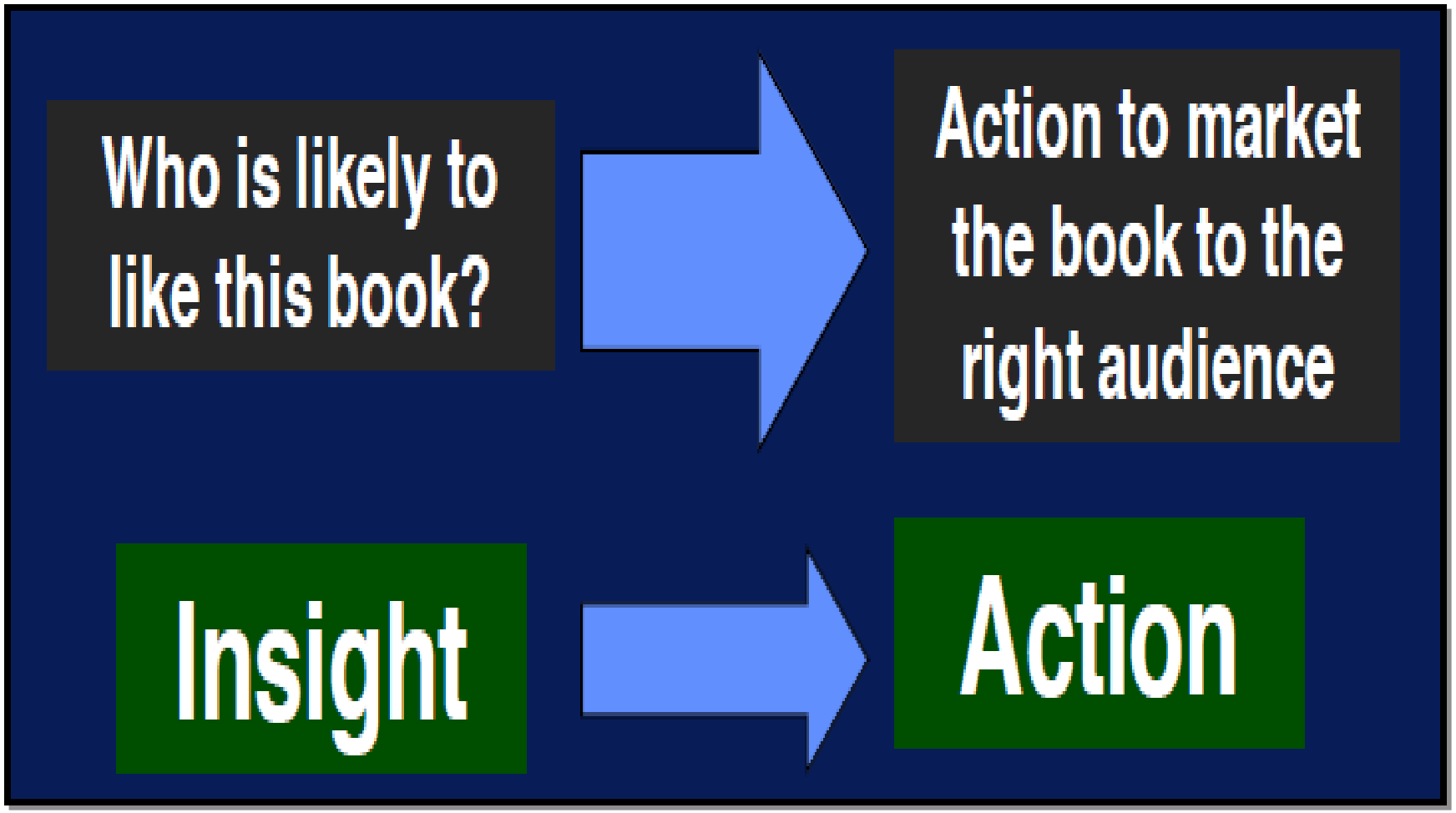

Insight
Data Product
Analysis
Big Data
Question
Insight

Book Recommendation

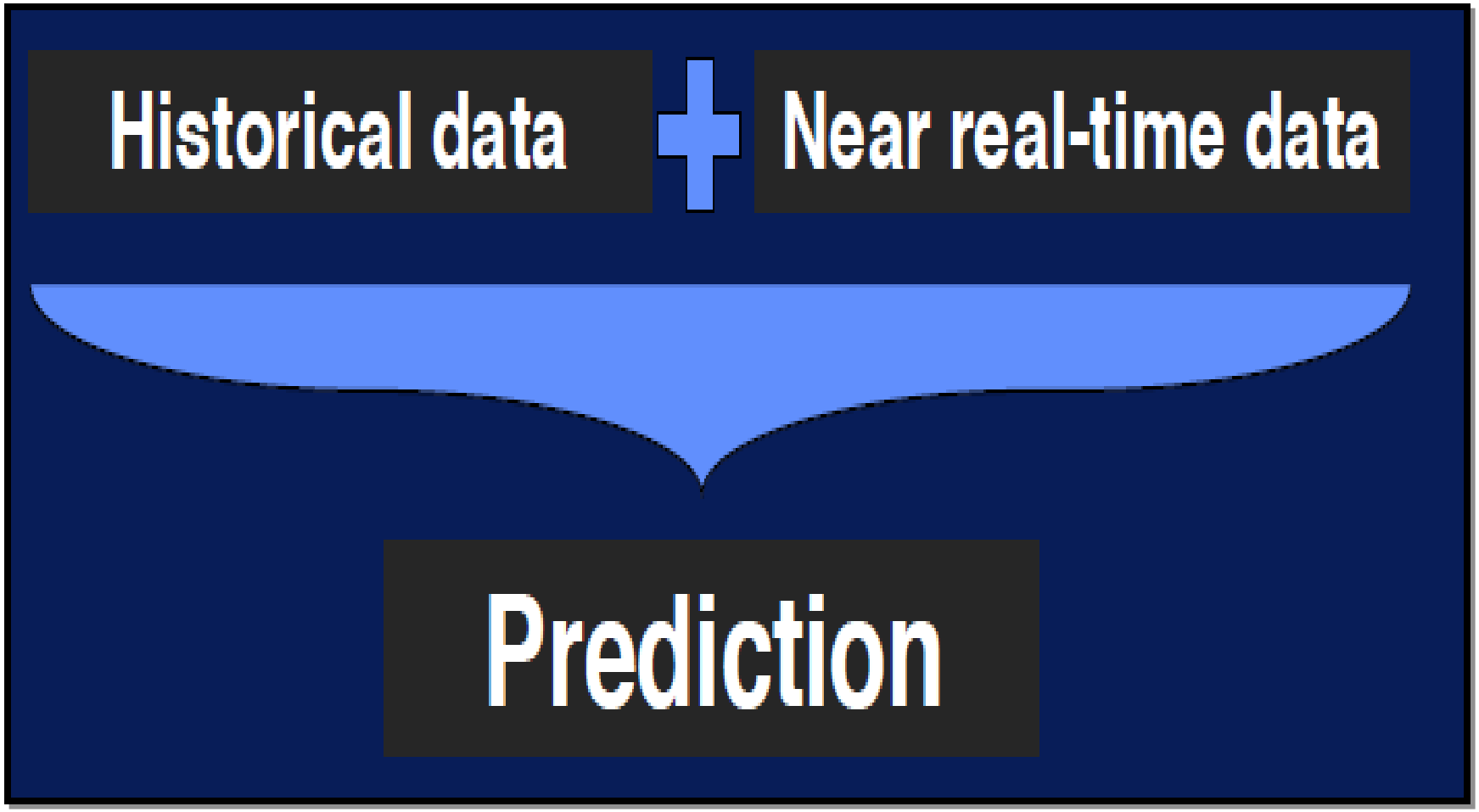
Find Potential Audience for the Book



Market a New Book

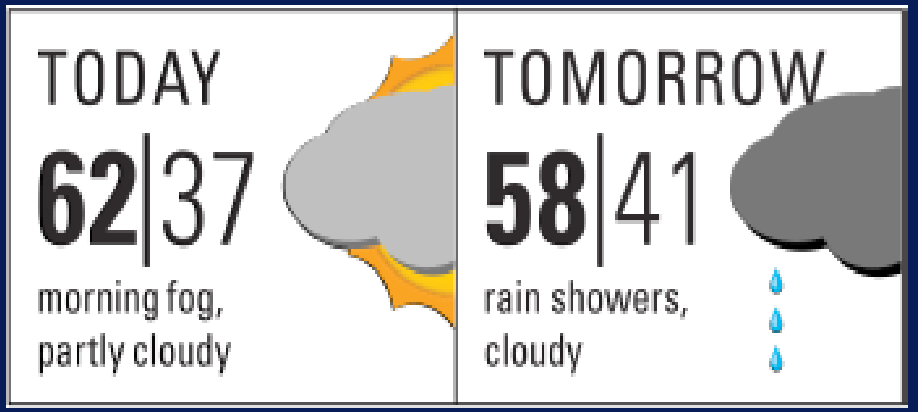


Actionable Information



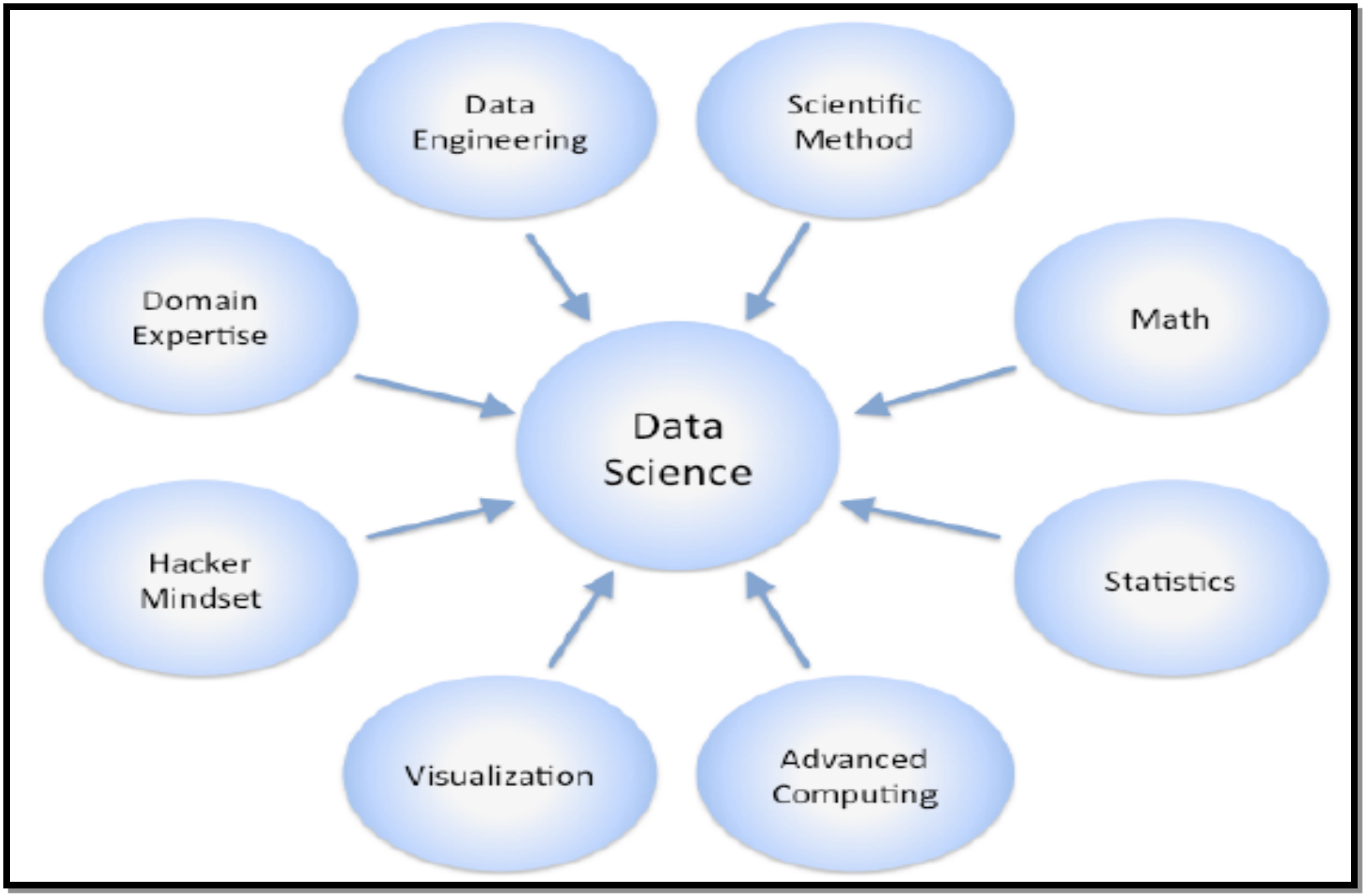
Prediction And Action

Prediction



Action

Data Science is a Team Work



Data Science is a Team Work

Have passion for data

Relate problems to analytics

Care about engineering solutions

Exhibit curiosity

Communicate with teammates

Data Science

The sum is bigger than the parts!

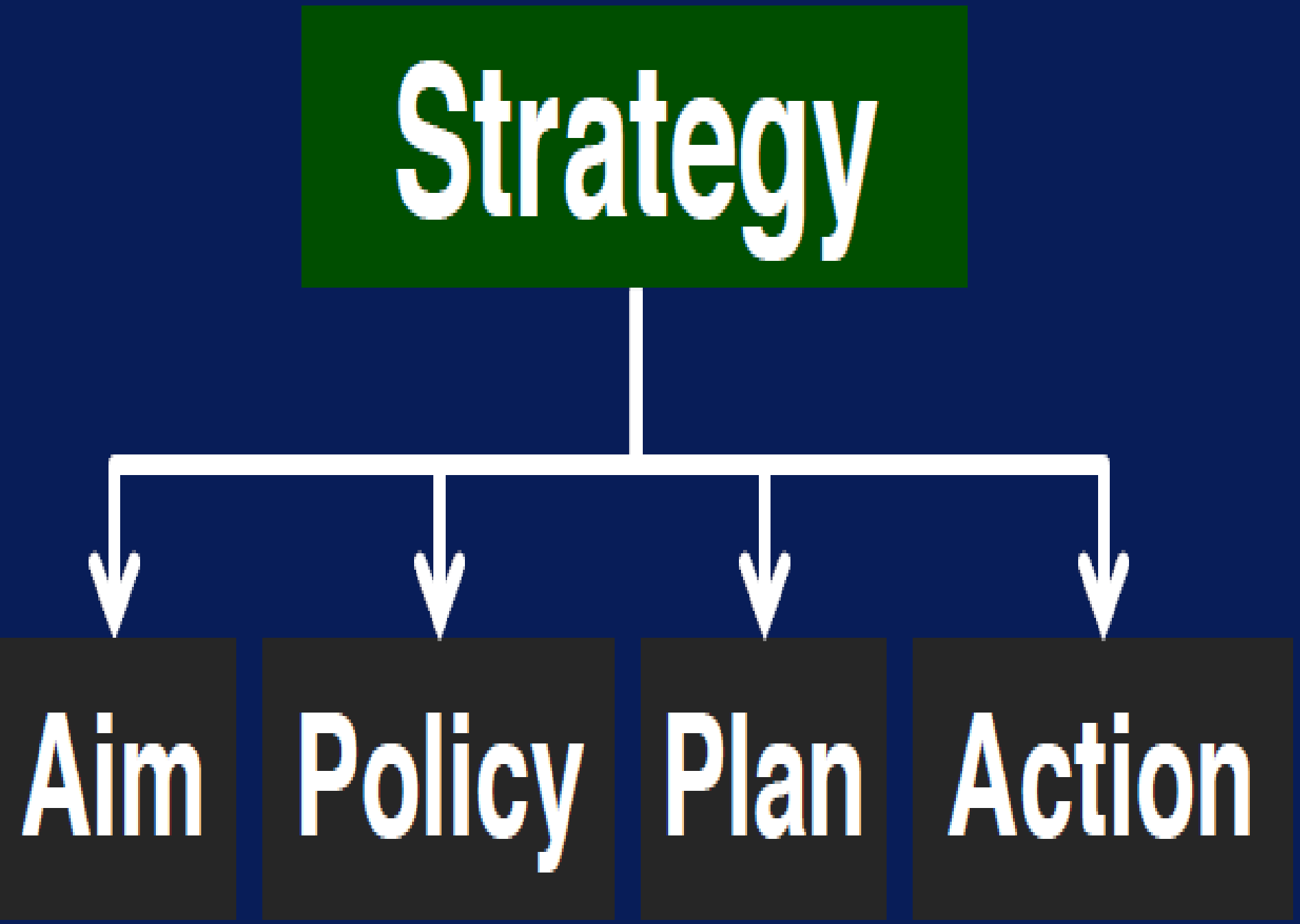
Big Data → Actionable Insight

Data Scientist

- Technical Skills
- Business Skills
- Soft Skills

Building A Big Data Strategy

"a **plan of action** or **policy** designed to achieve an **overall aim**"


Strategy
Aim
Policy
Plan
Action

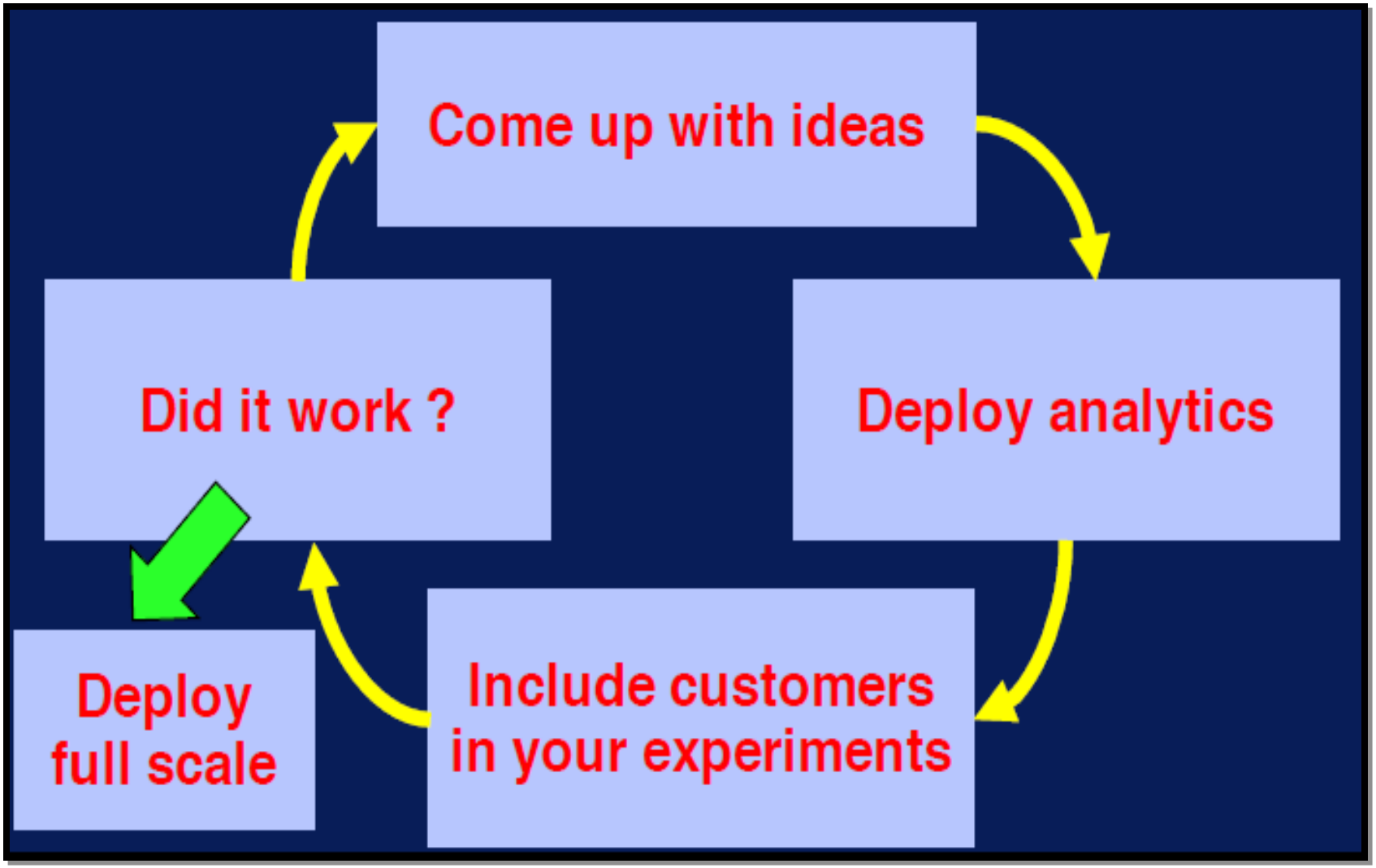
Business Objectives

Provide Organizational Bye In…

Diverse Expertise
Deliver as team

Build In House Expertise

Open Mini Idea Lab



Share Data

Remove barriers to data access

No data silos

Data sharing mindset

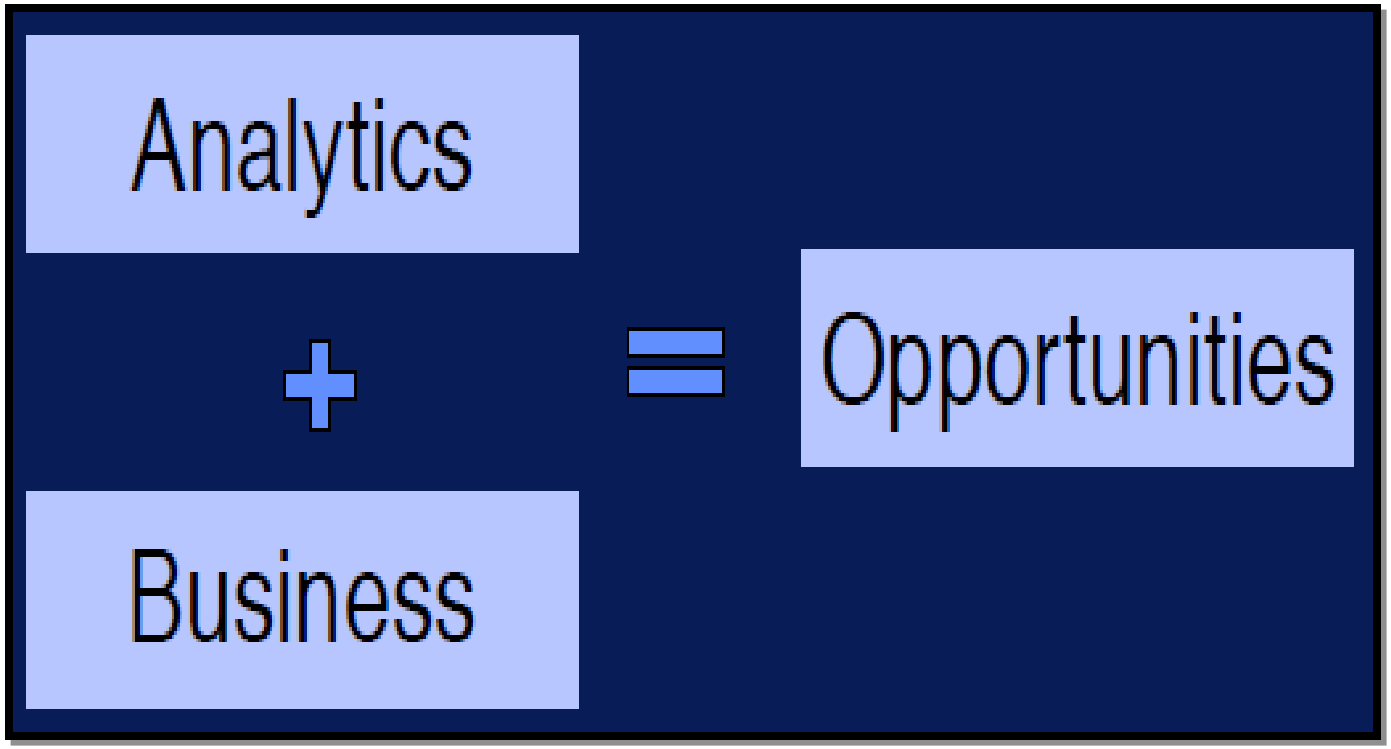
Define Big data Policies

Privacy and lifetime

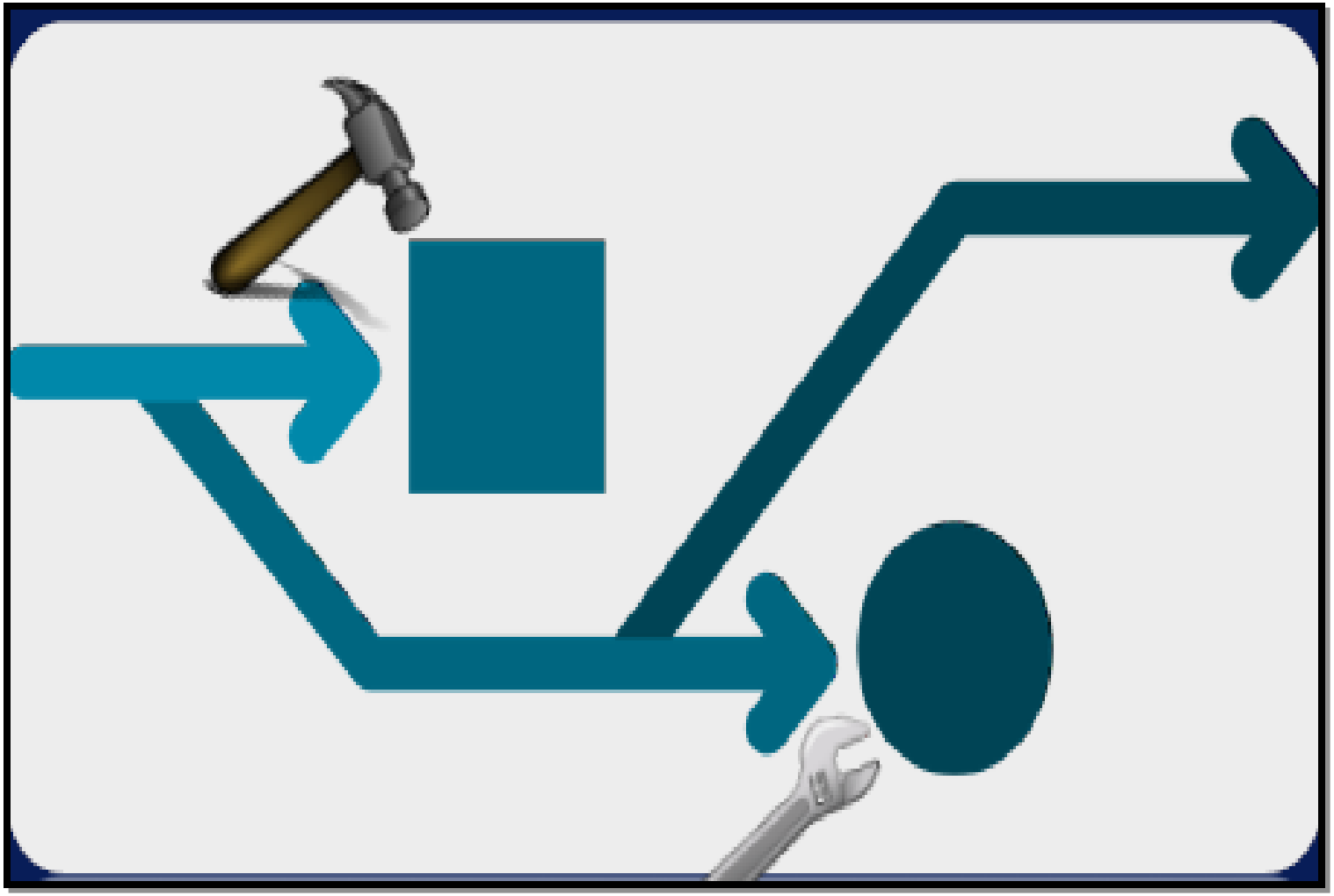
Curation and quality

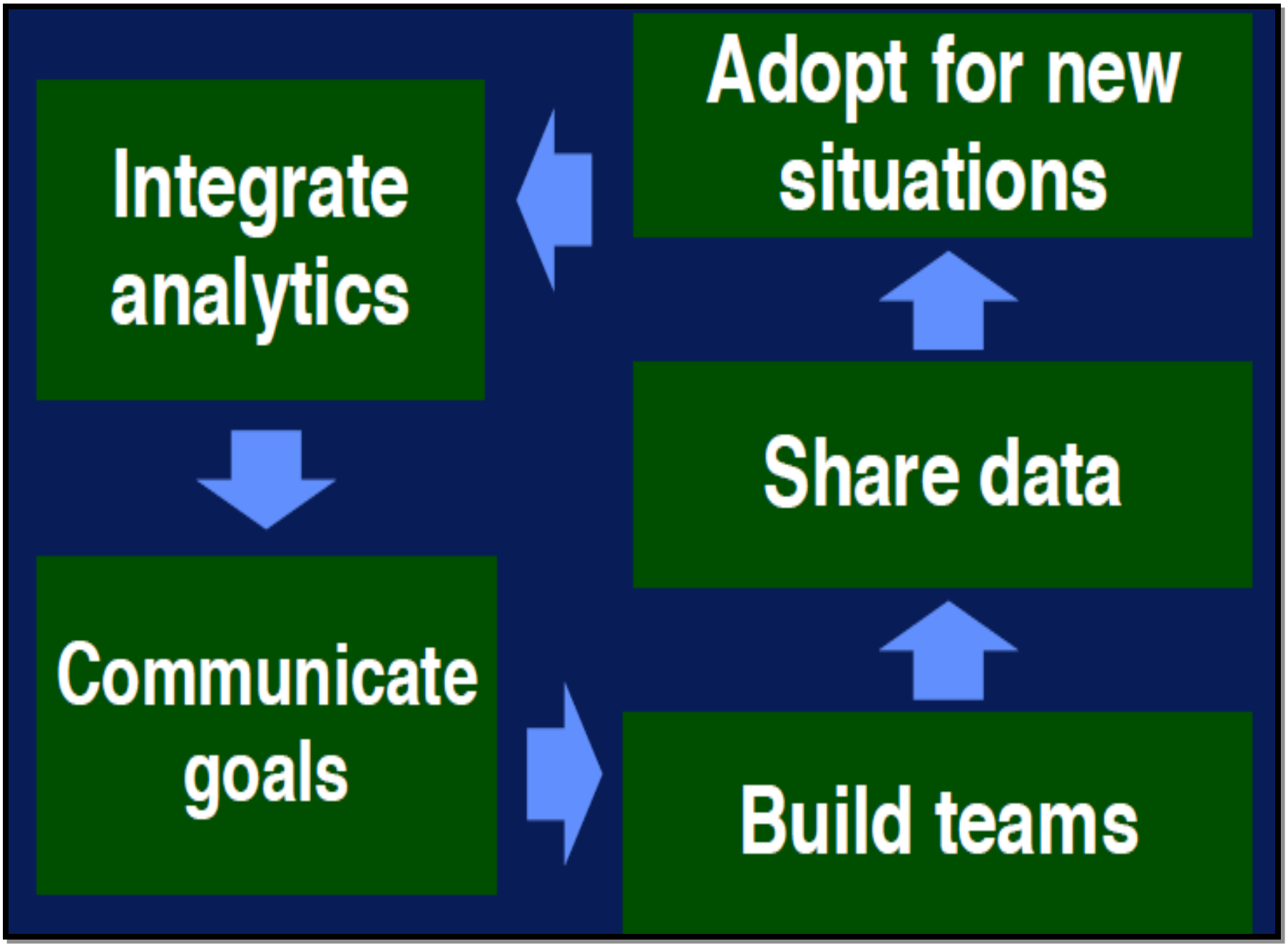
Interoperability and regulation

Cultivate Analytics Driven Culture

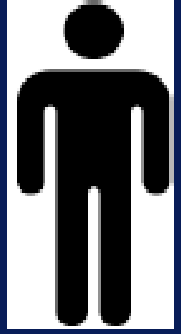


Adapt Strategy to your Use Case



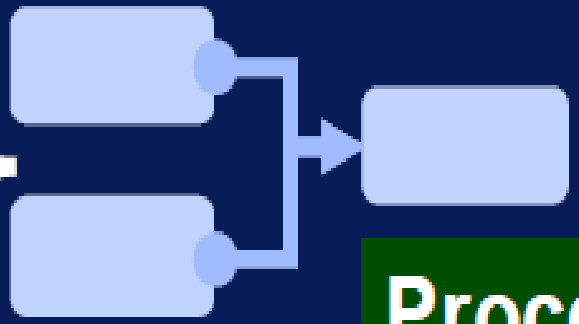

Adopt for new situations
Integrate analytics
Share data
Communicate goals
Build teams

Five Ps
of
Big Data



People

Purpose

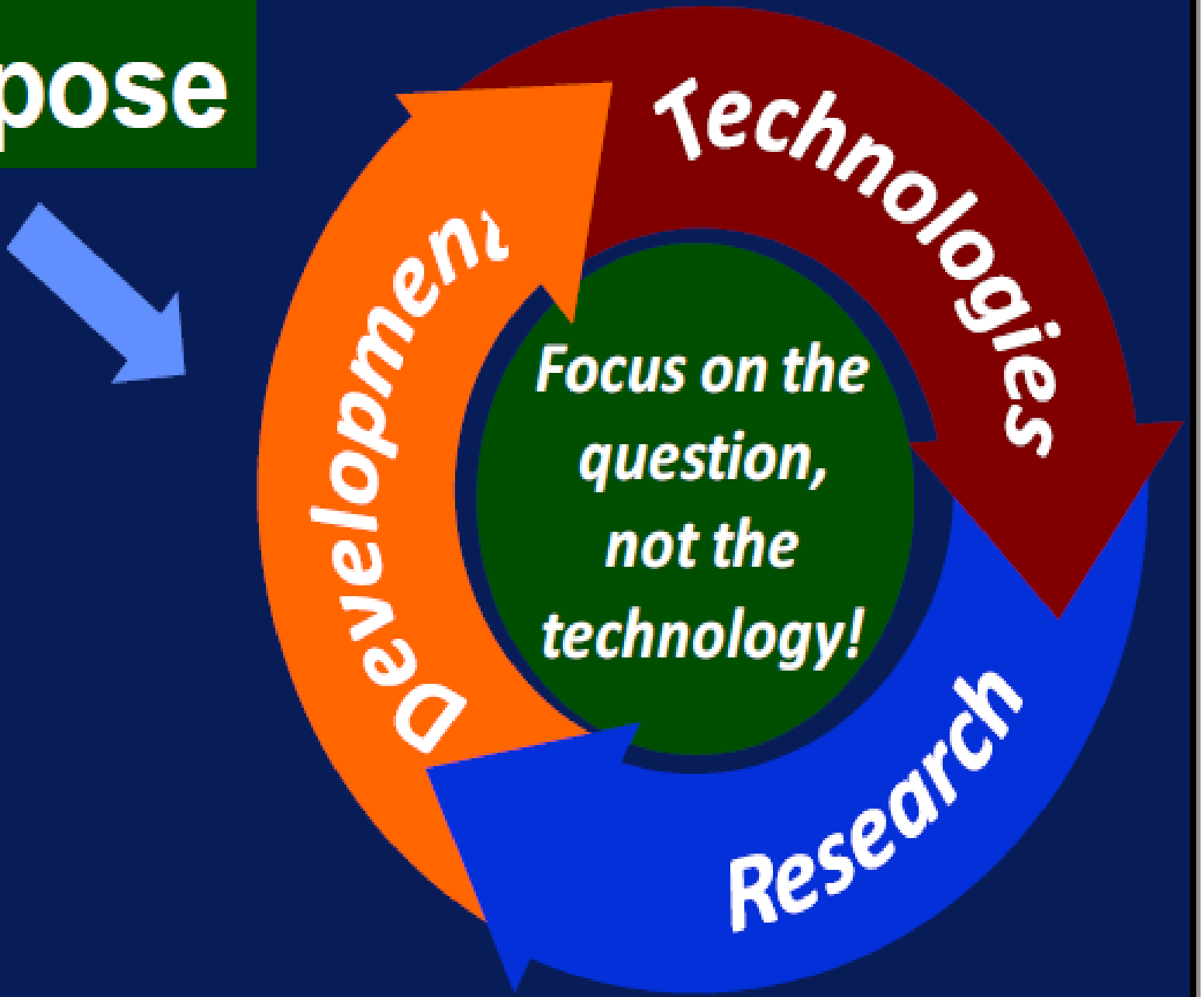


Process

Platforms

Programmability

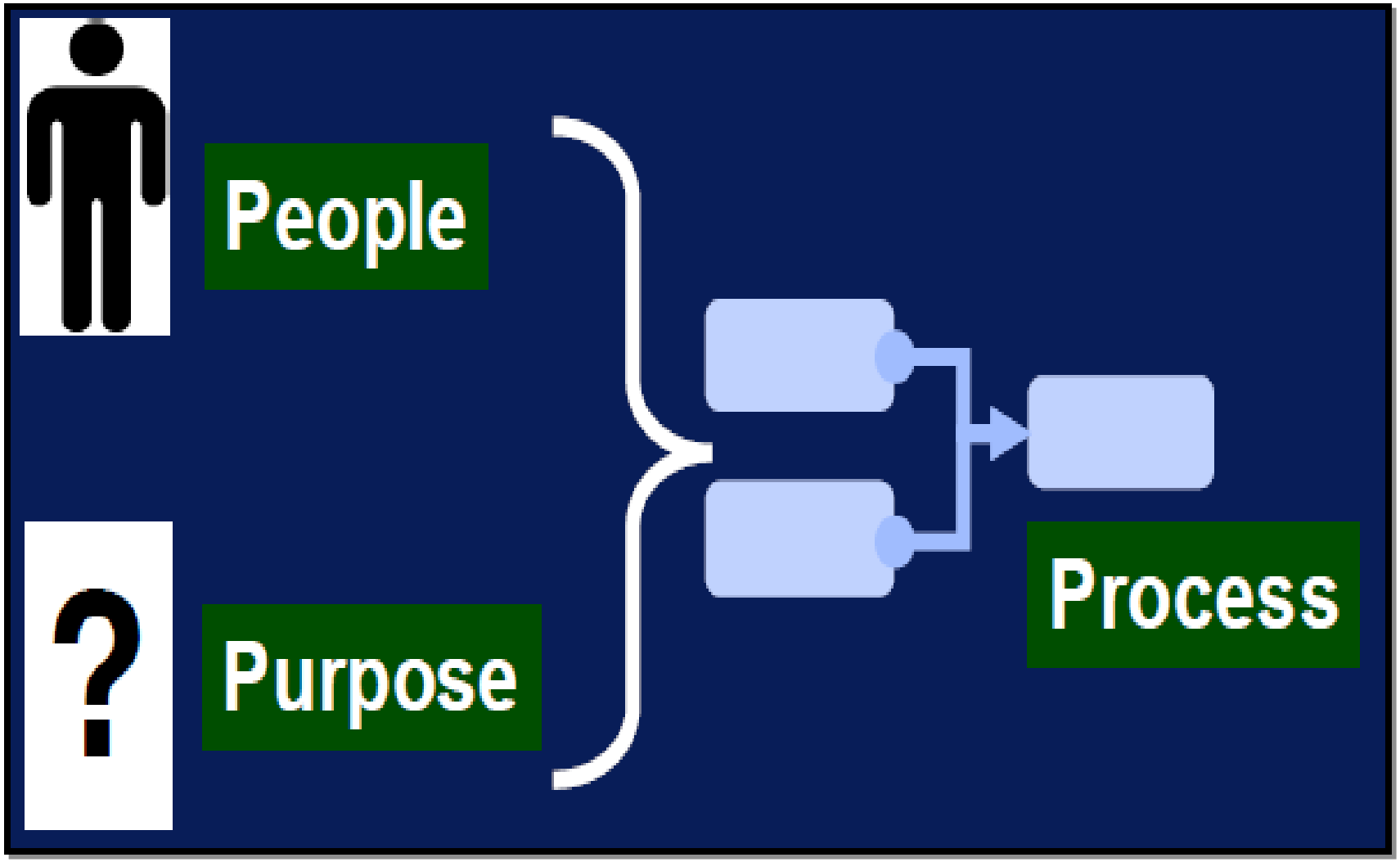
Purpose


Technologies
Research
Development
Focus on the question, not the technology!

Let's not dive into the techniques yet! What is the problem at large? How do you see yourself solving it?

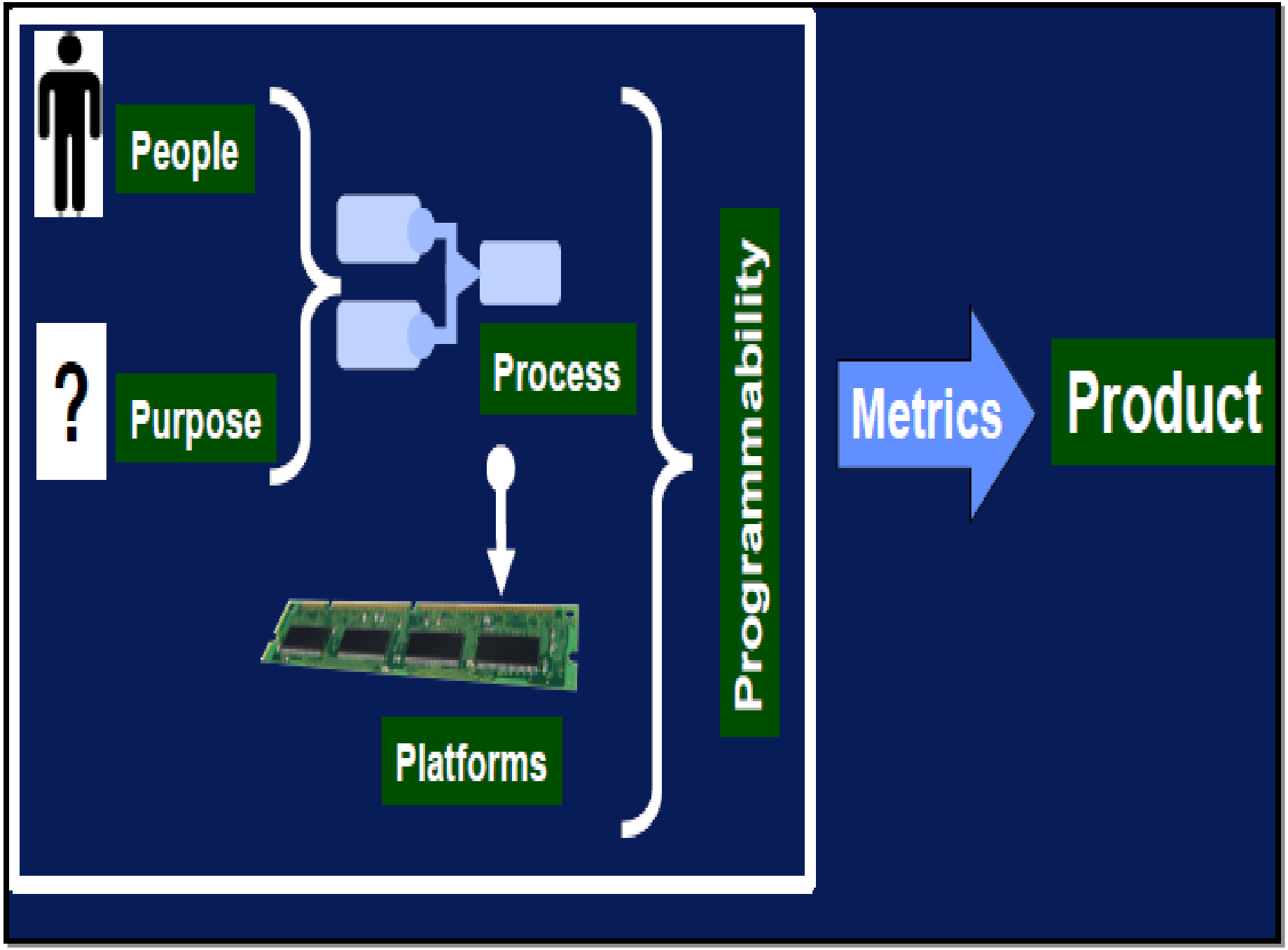
Rate of Spread and Direction

Components of Data Science

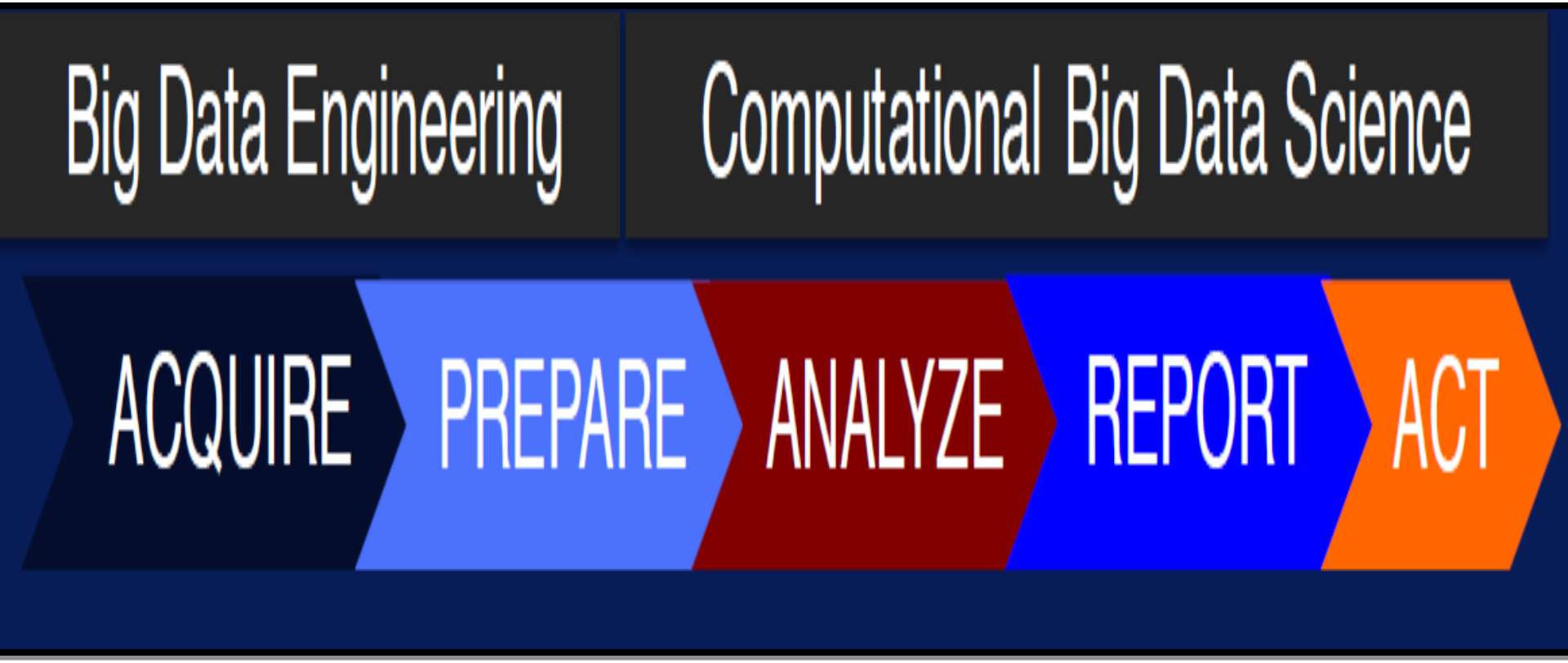


Build Metrics for Accountability

- Cost
- Timeline
- Planning of deliverables
- Expectations
- Purpose

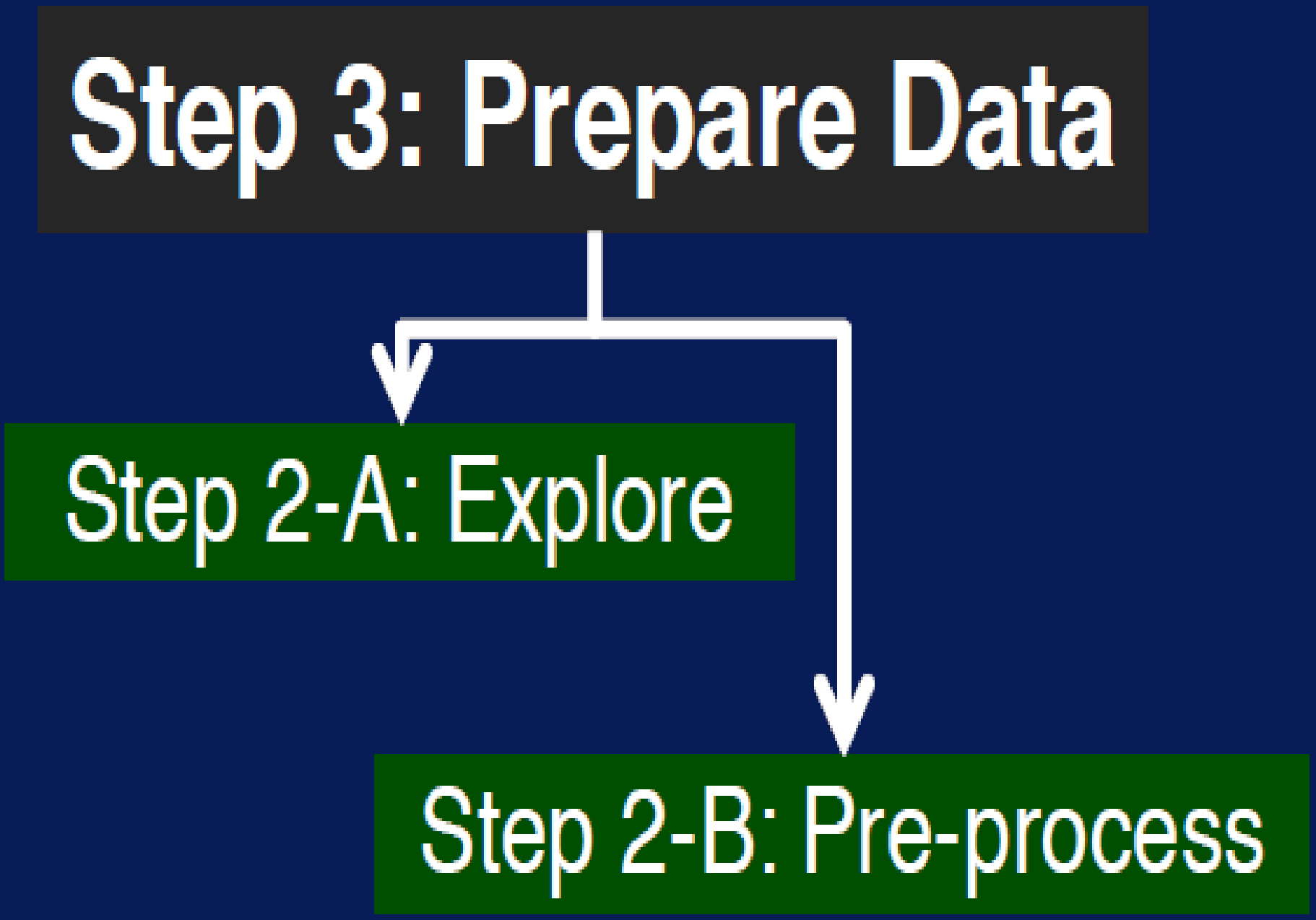

People
?
Purpose
Process
Platforms
Programmability
Metrics
Product

Steps in Data Science Process

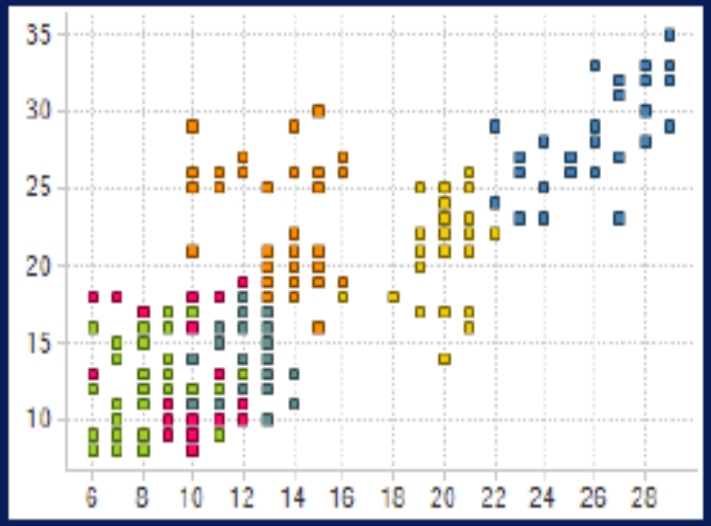

Big Data Engineering
Computational Big Data Science
ACQUIRE
PREPARE
ANALYZE
REPORT
ACT

Step 1: Acquire Data

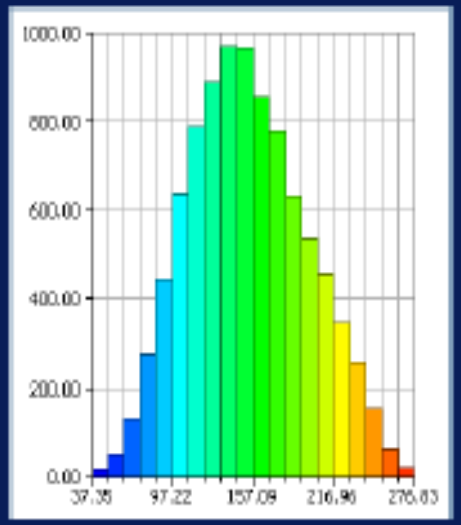
- Identify data sets
- Retrieve data
- Query data


Step 3: Prepare Data
Step 2-A: Explore
Step 2-B: Pre-process

Step 2-A: Explore Data

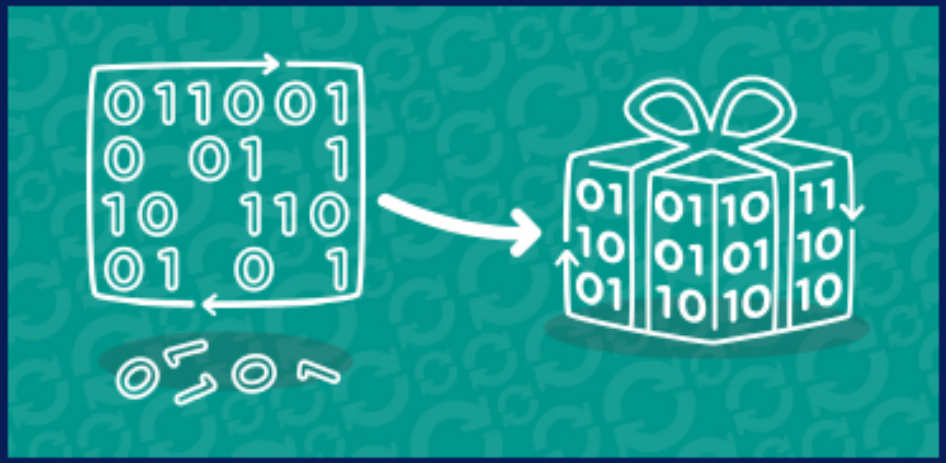


Understand nature of data



Preliminary analysis

Step 2-B: Pre-process Data



Clean

Integrate

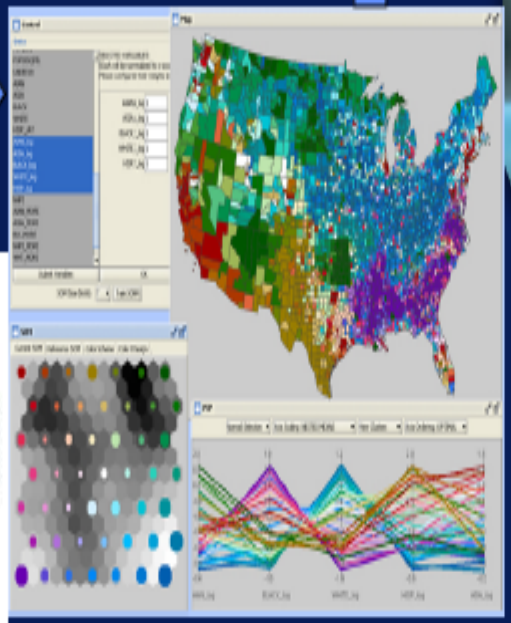
Package

Step 3: Analyze Data

Select analytical techniques

Build models

Step 4: Communicate Results



Step 5: Apply Results