

# Framework dan Metode Data Mining

---

Dr. Koko Bachrudin, S.Kom., MMSI.

# Outline

- Introduction
- Framework/model proses – Membutuhkan perulangan dan kehandalan
- Tahapan Utama CRISP-DM

# Analisis Data

 VALUE
 

1

## Descriptive

Based on Live Data,  
Tells what's  
happening in real  
time

Accurate & Handy for  
Operations  
management

Easy to Visualize



2

## Diagnostic

Automated RCA –  
Root Cause Analysis

Explains “why” things  
are happening

Helps trouble shoot  
issues



3

## Predictive

Tells What's likely to  
happen?

Based on historical  
data, and assumes a  
static business  
plans/models

Helps Business  
decisions to be  
automated using  
algorithms.



4

## Prescriptive

Defines future actions  
– i.e., “What to do  
next?”

Based on current data  
analytics, predefined  
future plans, goals,  
and objectives

Advanced algorithms  
to test potential  
outcomes of each  
decision and  
recommends the best  
course of action

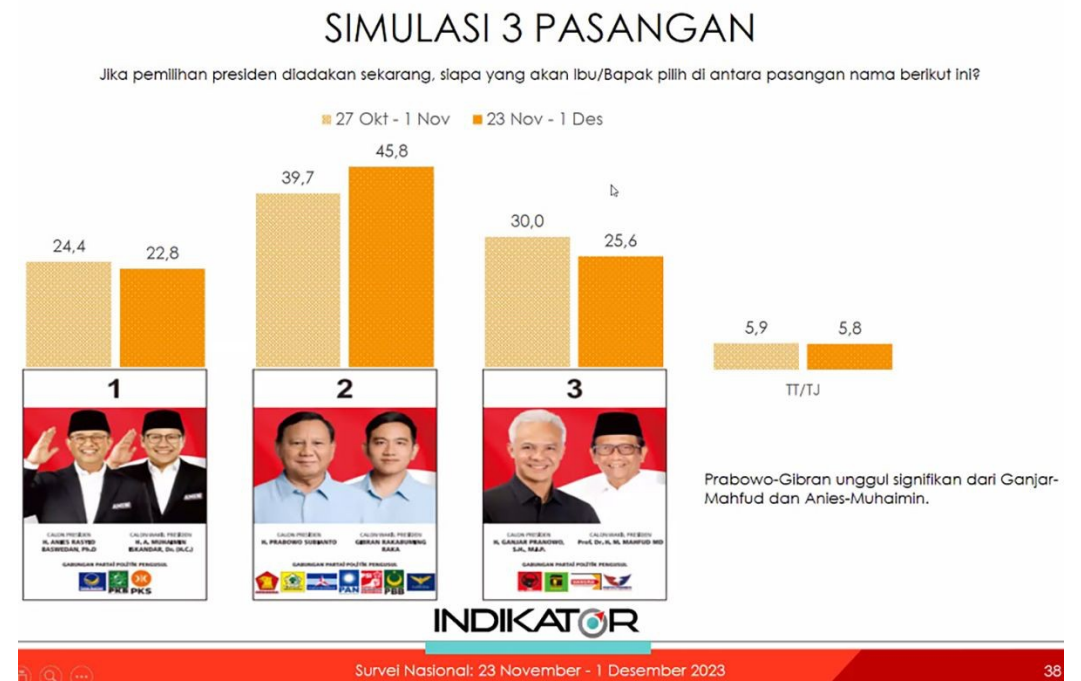

 Complexity
 

# Contoh Analisis

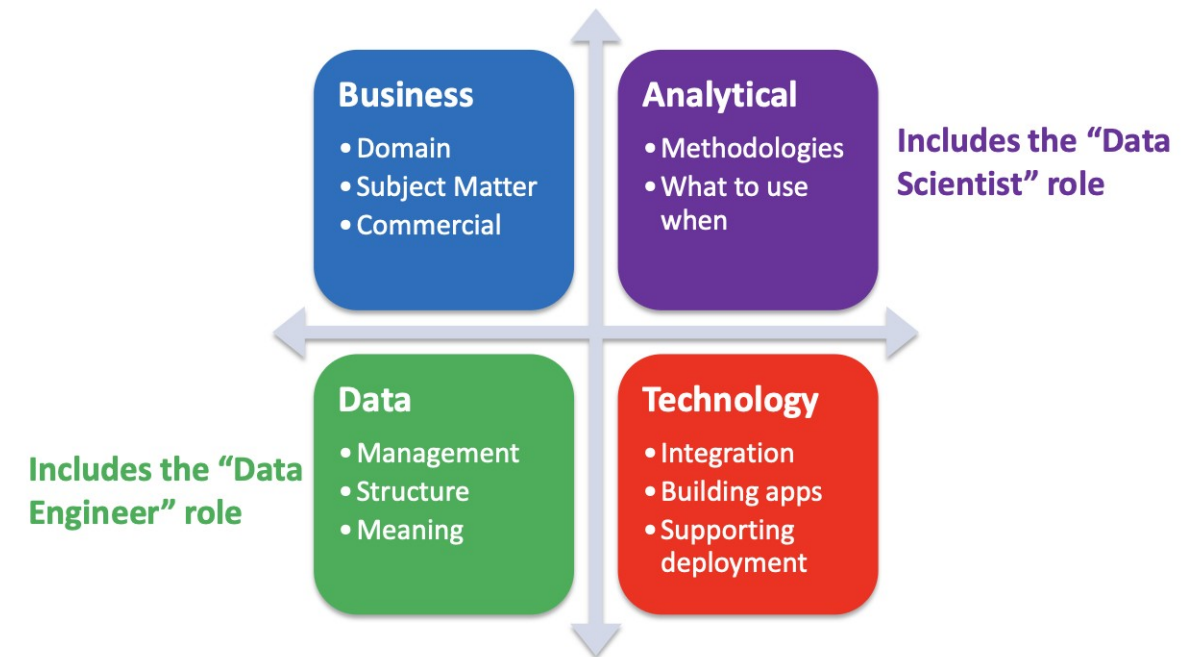
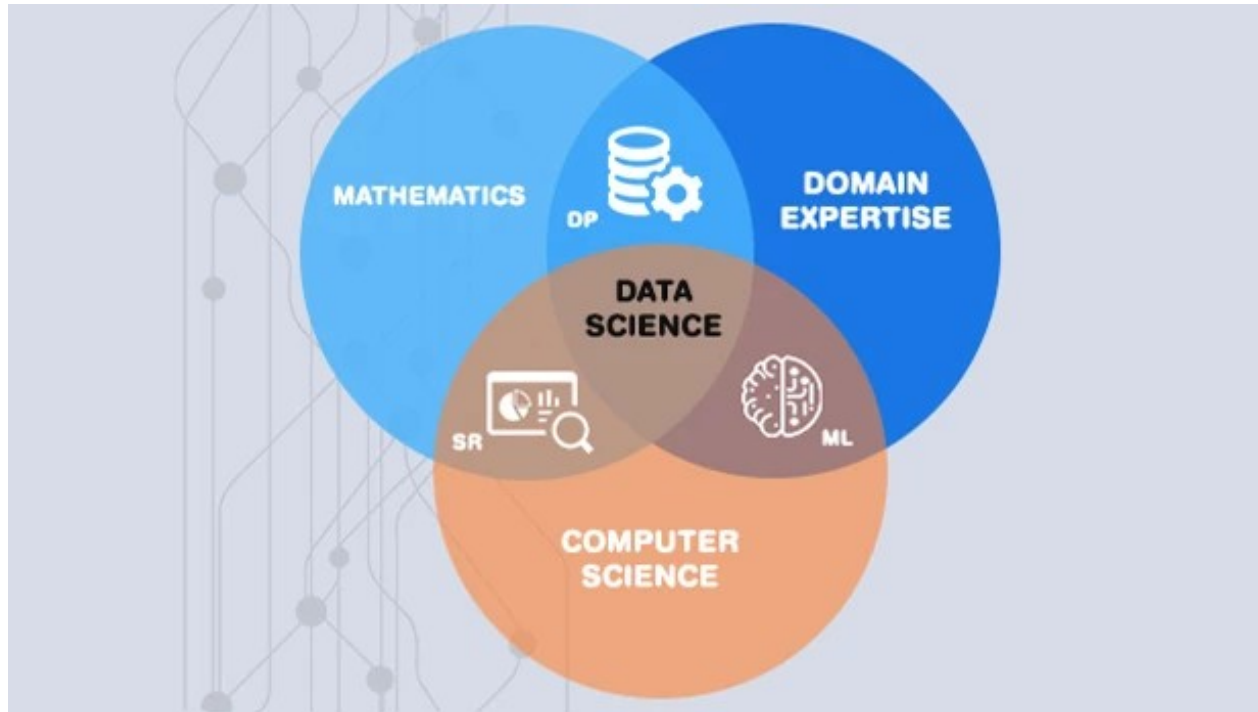
- Mengoptimalisasi Manusia
- Mengoptimalisasi Proses
- Mengurangi Fraud
- Pengambilan Keputusan Berdasarkan data
- **Kualitas Data adalah hal yang penting**

Deputi Politik 5.0 TPN [REDACTED] menyebut, hasil survei internal tak menunjukkan perubahan, 03 tetap teratas. Meski di lembaga survei lain paslon 03 ditempatkan pada posisi ketiga.

"Tentang potensi menang. Saya tadi baru saja lihat rapat triangulasi, terutama lihat big data kami. Posisi saya masih sama dengan kira-kira dua minggu terakhir ya," kata [REDACTED] di Jalan Cemara 19, Jakarta Pusat, Kamis (25/1).



# Data Science dan Kemampuan Analitik

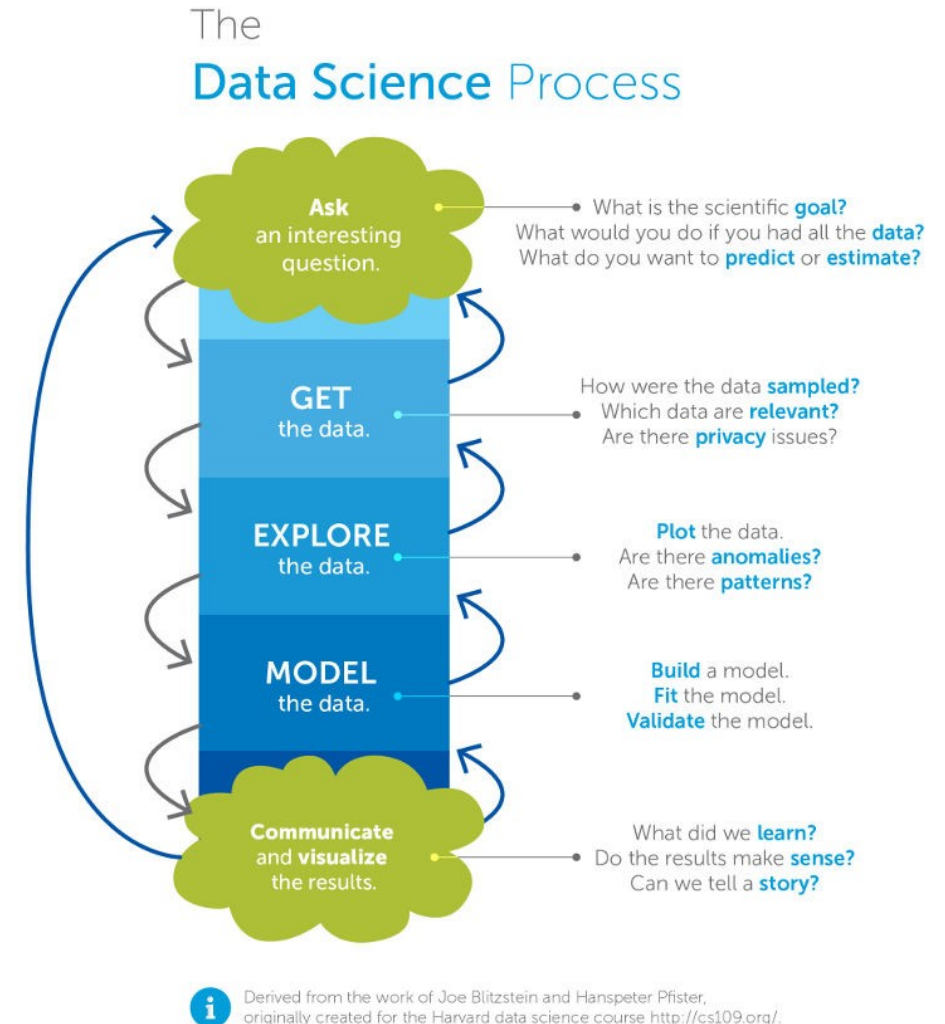


***a real or conceptual structure intended to serve as a support or guide for the building of something that expands the structure into something useful.***

**– Framework–**

# Framework Analitik – Model Proses

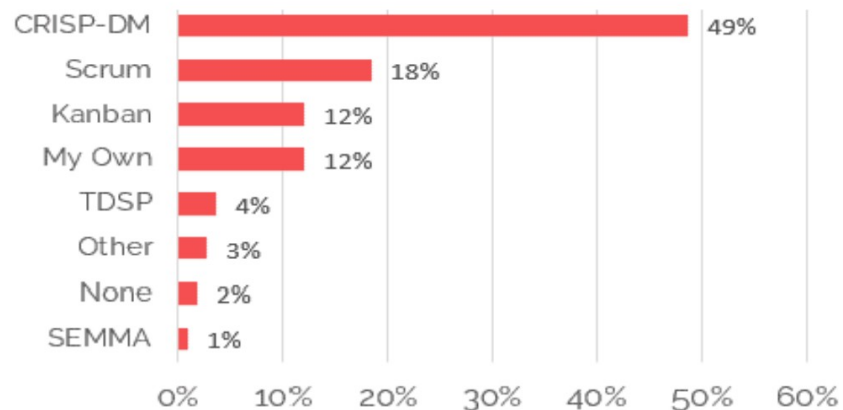
- Untuk mengerjakan proyek data dengan baik, prosesnya harus dapat diandalkan dan dapat diulang
- Kerangka kerja untuk merekam pengalaman
- Memungkinkan proyek direplikasi, ilmiah
- Membantu dan mendorong perencanaan dan manajemen proyek
- Mempermudah dalam pengadopsian
- Menunjukkan kematangan analisis data
- Mengurangi ketergantungan pada ahli analisis data tertentu



# Penggunaan Framework Data Science

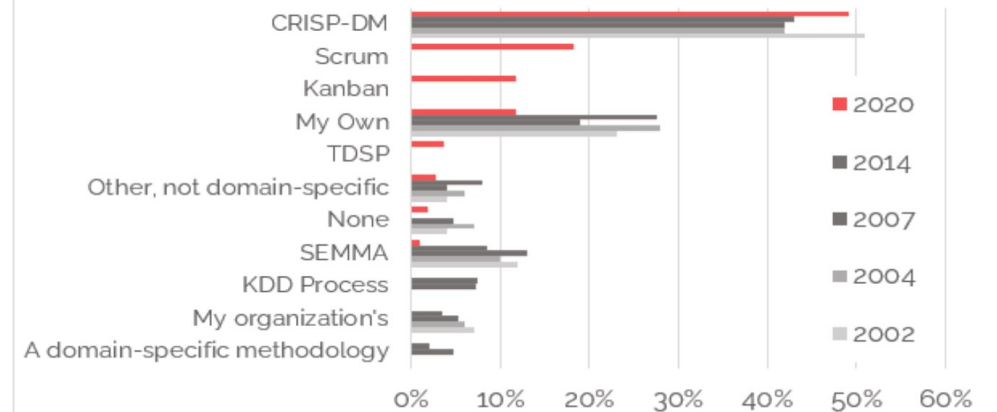
## datascience-pm.com Poll Results

Which process do you most commonly use for data science projects?



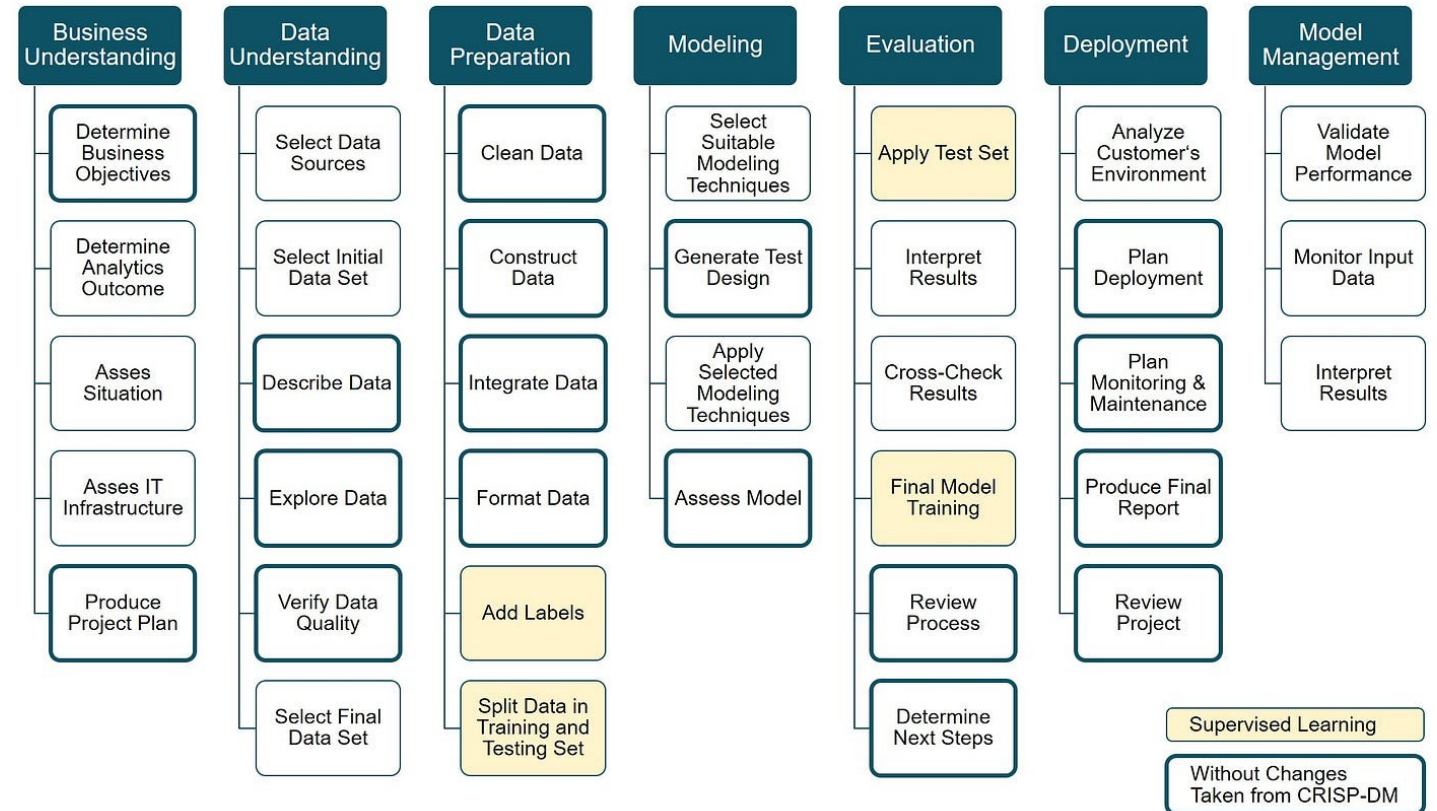
## Poll Result Comparison

Most popular methodologies

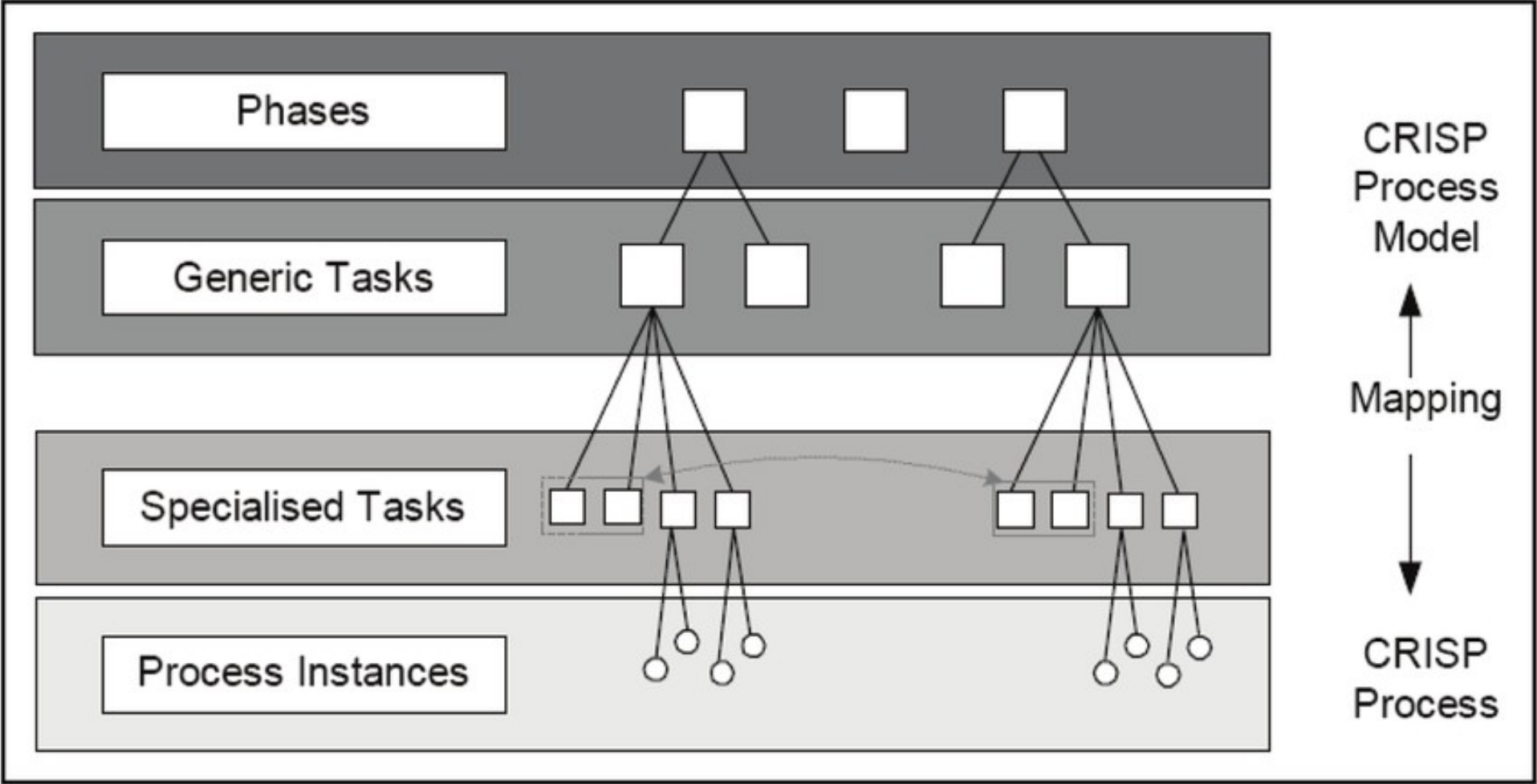




# CRISP-DM Framework



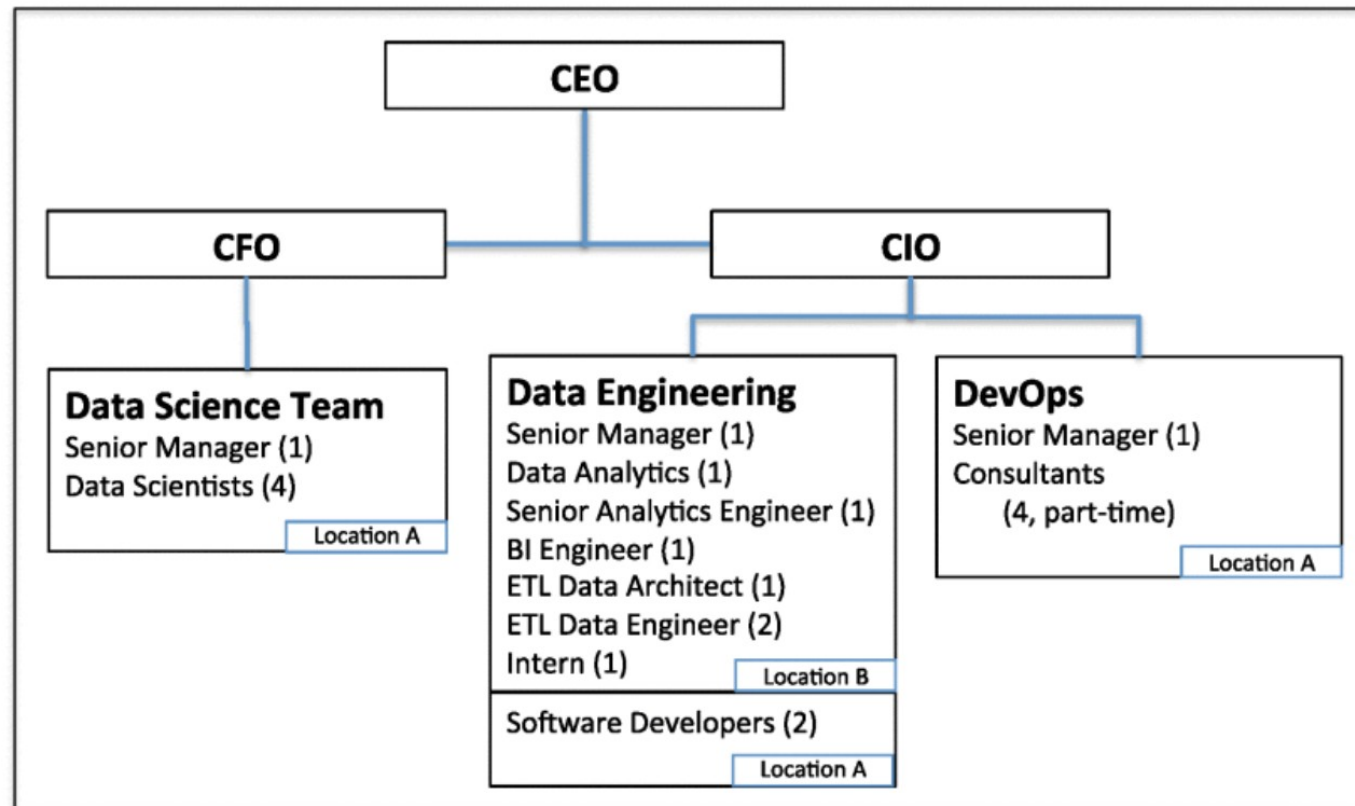
# Model Hirarki dari CRISP DM



# Luaran dari CRISP-DM

Business Understanding	Data Understanding	Data Preparation	Modeling	Evaluation	Deployment
<b>Determine Business Objectives</b> Background Business Objectives Business Success Criteria	<b>Collect Initial Data</b> Initial Data Collection Report  <b>Describe Data</b> Data Description Report  <b>Explore Data</b> Data Exploration Report  <b>Verify Data Quality</b> Data Quality Report	<b>Select Data</b> Rationale for Inclusion/Exclusion  <b>Clean Data</b> Data Cleaning Report  <b>Construct Data</b> Derived Attributes Generated Records  <b>Integrate Data</b> Merged Data  <b>Format Data</b> Reformatted Data  Dataset Dataset Description	<b>Select Modeling Techniques</b> Modeling Technique Modeling Assumptions  <b>Generate Test Design</b> Test Design  <b>Build Model</b> Parameter Settings Models Model Descriptions  <b>Assess Model</b> Model Assessment Revised Parameter Settings	<b>Evaluate Results</b> Assessment of Data Mining Results w.r.t. Business Success Criteria Approved Models  <b>Review Process</b> Review of Process  <b>Determine Next Steps</b> List of Possible Actions Decision	<b>Plan Deployment</b> Deployment Plan  <b>Plan Monitoring and Maintenance</b> Monitoring and Maintenance Plan  <b>Produce Final Report</b> Final Report Final Presentation  <b>Review Project</b> Experience Documentation
<b>Assess Situation</b> Inventory of Resources Requirements, Assumptions, and Constraints Risks and Contingencies Terminology Costs and Benefits					
<b>Determine Data Mining Goals</b> Data Mining Goals Data Mining Success Criteria					
<b>Produce Project Plan</b> Project Plan Initial Assessment of Tools and Techniques					

# Struktur Dari Tim Data Science



Source: J. S. Saltz and I. Shamshurin, "Achieving Agile Big Data Science: The Evolution of a Team's Agile Process Methodology," 2019 IEEE International Conference on Big Data (Big Data), 2019, pp. 3477-3485, doi: 10.1109/BigData47090.2019.9005493.

# Terima Kasih Atas Perhatiannya

---

