

# What is Data Mining

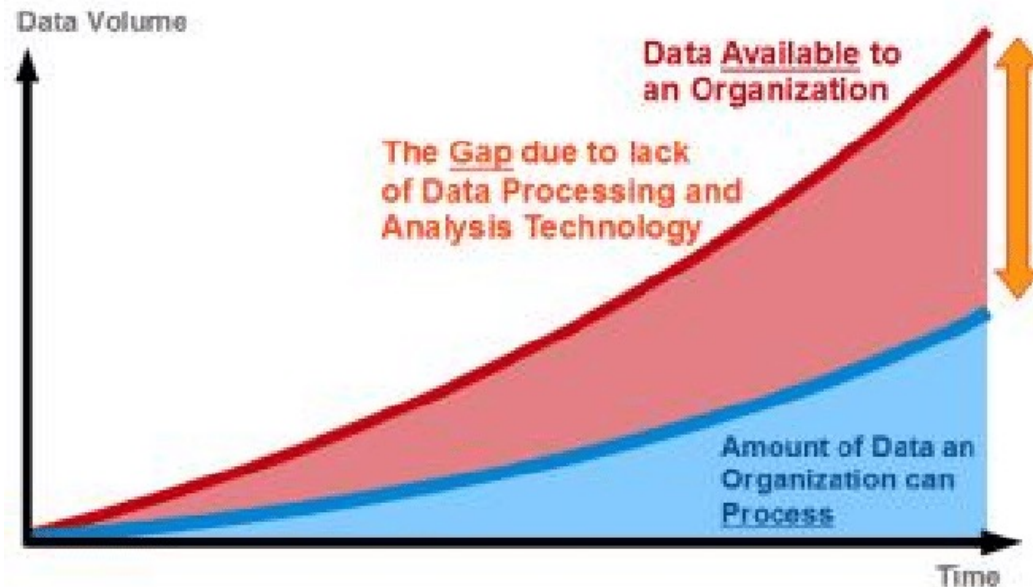
## Data Mining 101 Introduction

**Tim Pengajar:**

- Ratih Ardiati Ningrum, S.Si., M.S., M.Stat.
- Indah Fahmiyah, S.Si., M.Stat.



# What is Data Mining



- ❖ Data mining is the process of automatically discovering useful information in large data repositories.
- ❖ Human analysts may take weeks to discover useful information.
- ❖ Much of the data is never analyzed at all.

Source: Jacob, Samuel. (2017). Seamless Automation and Integration of Machine Learning Capabilities for Big Data Analytics. 8.

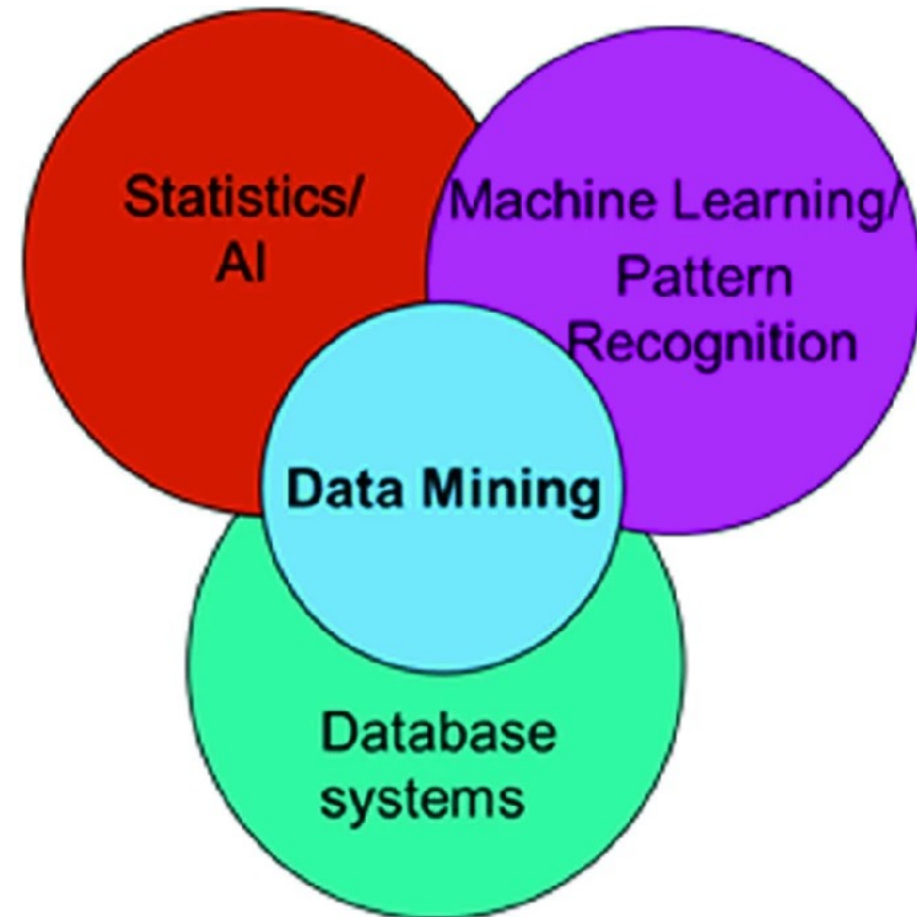


# Origins of Data Mining

Draws ideas from machine learning/AI, pattern recognition, statistics, and database systems

Must address:

- Enormity of data
- High dimensionality of data
- Heterogeneous, distributed nature of data

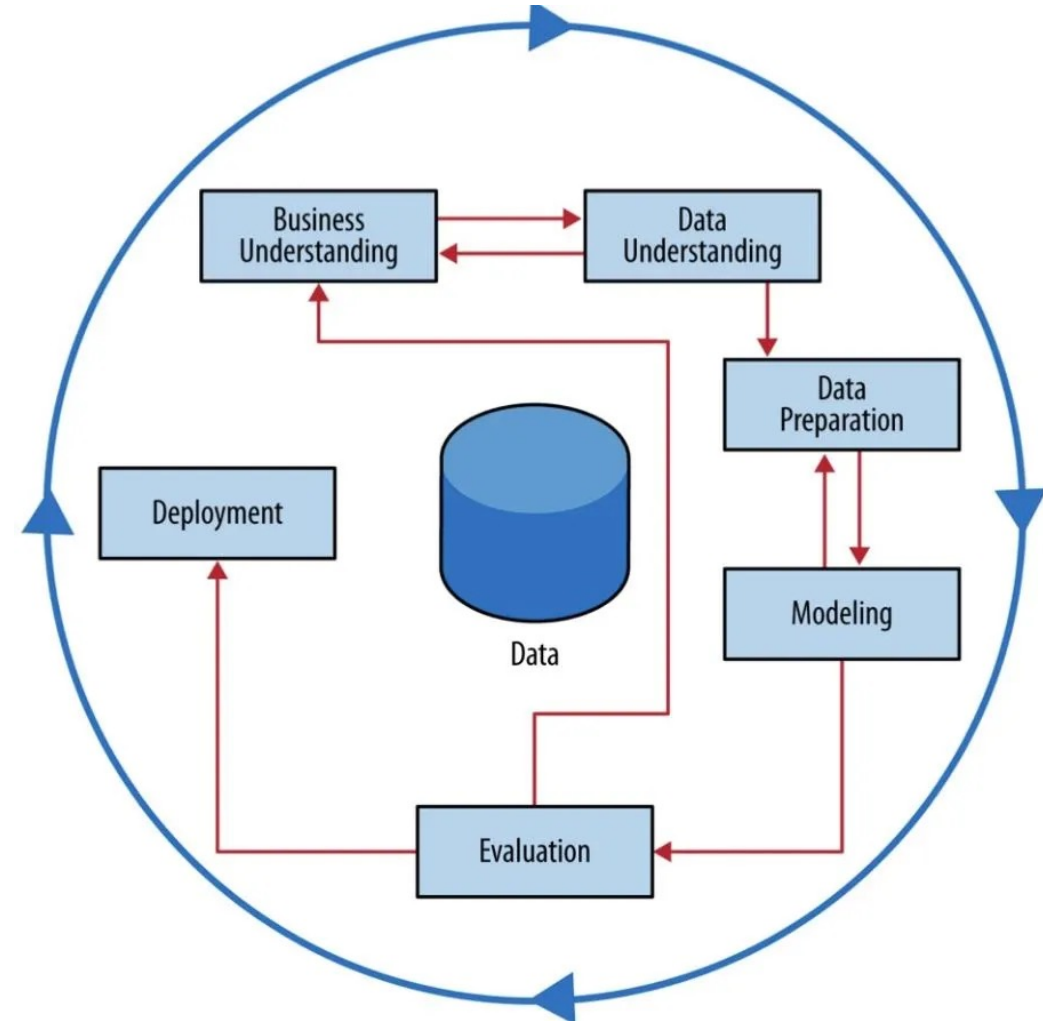


Source: Järvinen, P., Siltanen, P., Kirschenbaum, A. (2021). Data Analytics and Machine Learning. In: Södergård, C., Mildorf, T., Habyarimana, E., Berre, A.J., Fernandes, J.A., Zinke-Wehlmann, C. (eds) Big Data in Bioeconomy. Springer, Cham. [https://doi.org/10.1007/978-3-030-71069-9\\_10](https://doi.org/10.1007/978-3-030-71069-9_10)



# What is Data Mining?

Alternative names: Knowledge discovery (mining) in databases (KDD), knowledge extraction, data/pattern analysis, data archeology, data dredging, information harvesting, business intelligence, etc.



Source: Akyildirim, N.A., 2019, Data Mining Process Basics, <https://medium.com/>



# Terima Kasih!

