

## TIF21-31-45

### Data Engineering

#### Rekayasa Data

#### BASIC INFORMATION

<b>Course Credit</b>	3 / 150 minutes per Week3 / 150 minutes per Week
<b>Course Type</b>	RequiredRequired
<b>Course Classification</b>	Engineering TopicsEngineering Topics
<b>Prerequisites</b>	Mathematics, Statistics, Programming, Artificial Intelligence, Computer Science

#### STUDENT AND LEARNING OUTCOMES

##### Covered Student Outcomes

Development of Engineering Solution (b)Development of Engineering Solution (b)	Modern Tools Utilization (e)Modern Tools Utilization (e)
Data and Experiments (d)Data and Experiments (d)	Knowledge Contemporary and Issues (f)Knowledge Contemporary and Issues (f)

##### Learning Outcomes

- LO1** Understand the data engineering paradigm, scheme and architecture
- LO2** Able to store, manage and manipulate data in databases, data lakes or cloud-based storage service using a number of selected data-processing software
- LO3** Able to choose the suitable method to prepare raw data for further analyses based on the data selection criteria
- LO4** Able to perform data analysis using suitable data processing techniques and algorithms
- LO5** Able to provide and discuss initial descriptive insights from a given data set using various graphical and numerical summarisation techniques
- LO6** Able to comply with data privacy protection regulation.

## COURSE DESCRIPTION

This course is intended to equip students with an understanding of the data engineering concept and techniques. Furthermore, this course is expected to guide the students to practically utilise the database systems and data processing tools to store, manage, extract, analyse and present a given dataset based on a number of required criteria. Students will also learn the importance of data and privacy protection in society.

## TOPICS

1. An overview of data engineering
2. Data organisation, modelling, and design techniques
3. Data storage and warehousing
4. Data access and maintenance
5. Data analytics application and algorithms
6. Engineering non-traditional data types
7. Data standards and quality
8. Data presentation and visualisation
9. Effects on society
10. Importance of privacy protection

## REFERENCES

- [1] Witten, I. H., *Data mining : practical machine learning tools and techniques*, Morgan Kauffman Publisher, 2017.
- [2] Han, Jiawei., *Data mining : concepts and techniques*, Elsevier, 2012.
- [3] Rajaraman, A., and Ullman, J. D., *Mining of Massive Datasets*, Cambridge University Press, 2011.