Fachhochschule Kiel
University of Applied Sciences

Big Data Technologies – Big Data Applications

Following is **Part 2** of the **Portfolio Examination** of the module **Big Data Technologies** in **Summer Semester 2024**. Please **upload** your **documents** and **program files** to the corresponding course in **Moodle** Learn **latest until 20.06.2024 at 23:59**.

Data Stream Processing

Please prepare a presentation about the message broker Kafka and install, configure, and program it with Python due to the following guidelines:

1. Install and configure Kafka on your laptop. You can use a Docker image or make a native installation for your operating system.

2. Create a topic 'measurements' via the Kafka Bash.

3. Install the Kafka Python Library and create a Python file or Jupyter notebook to test the message broker connection.

4. Develop a Python application that periodically creates random measurements of your choice (e.g., temperature, moisture, wind speed, or other kind of measurement data) with a useful data structure and sends it to the topic 'measurements' to demonstrate writing of messages.

5. Develop a Python application that periodically or event-driven receives messages from the topic 'measurements and displays the data in a structured form to demonstrate reading of messages.

6. Prepare a (PowerPoint) presentation about Kafka that contains at least information about distributed message brokers and log systems, the IT architecture of Kafka including partitions, topics, producers and consumers, the client libraries, typical use cases as well as the installation, configuration, and programming of Kafka.

The **deliveries** are the presentation (as a PDF file) as well as the Python program file(s) or Jupyter notebook(s).

Big Data Application

Please prepare a **presentation of 5 minutes** of a Big Data Application of your **confirmed choice** to be held **in person** on **17.06.2024 at 12:45** in the lecture at university. Consider the following guidelines for the content of your presentation:

1. What is the domain or business sector and the use case of the Big Data Application?

2. What are the data sources being used and which kind of data is processed?

3. How does the IT architecture look like, and which Big Data Technologies are used?

4. What kind of data analytics or artificial intelligence is used and what is the (business) value of the solution?

5. What kind of data visualization or dashboards are used for which user groups?

6. If possible, demonstrate some real-world examples.

The **deliveries** are the presentation (as a PDF file).