

## Zajęcia 6 - Zadanie 1: Przygotowanie do interview

**Cel zadania – przygotować się do interview,**

Notatka z krótkim opisem najważniejszych narzędzi analitycznych/Big Data/AI. Móc opisać **wybrane** przez siebie zagadnienie/narzędzia analityczne **w wybranej** chmurze Azure/AWS/GCP.

Przykład

1. Robisz PoC na wykrycie anomalii z linii produkcyjnej (jakieś IoT). Jakich narzędzi użyjesz np. w Azure. Jak będzie wyglądało PoC.
2. Zaproponuj architekturę Big Data dla Rockstar Games przetwarzanie danych w celu wykrycia i zapobiegania błędom w grze, satysfakcji użytkowników. Jakie narzędzia, gdzie będzie compute ect.

Podpowiedź - przejrzyj co jest wymagane do certyfikacji np. DP-203, i DP-100, AI-102

<https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE3VEHg>

<https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE4MbYT>

<https://learn.microsoft.com/pl-pl/credentials/certifications/resources/study-guides/ai-102>

<https://learn.microsoft.com/en-us/credentials/certifications/fabric-analytics-engineer-associate/?practice-assessment-type=certification>

---

### **Opis architektury PoC w chmurze Azure:**

#### **1. Zbieranie danych (IoT):**

- **Azure IoT Hub:** Centrala komunikacyjna dla urządzeń IoT przesyłających dane (np. temperatura, drgania, ciśnienie).

#### **2. Przechowywanie danych:**

- **Azure Data Lake Storage Gen2:** Hurtownia danych do długoterminowego przechowywania surowych danych IoT.
- **Azure Blob Storage:** Tymczasowe przechowywanie danych do analiz ad hoc.

#### **3. Przetwarzanie danych strumieniowych:**

- **Azure Stream Analytics:** Przetwarzanie danych w czasie rzeczywistym, agregacja i filtrowanie sygnałów IoT.
- Integracja z detekcją anomalii (reguły, model ML).

#### **4. Modelowanie i wykrywanie anomalii:**

- **Azure Machine Learning:**
  - Budowa modelu (np. Isolation Forest, Autoencoder)
  - Trenowanie i ewaluacja
  - Wdrożenie modelu jako endpoint REST API

#### **5. Wdrażanie i automatyzacja:**

- **Azure Functions:** Lekka funkcja serverless, która wywołuje predykcję modelu dla każdego nowego rekordu.
- **Azure Logic Apps:** Notyfikacje, alerty dla zespołu produkcyjnego.

## 6. Wizualizacja danych:

- **Power BI:** Dashboard prezentujący aktualny stan produkcji, wykryte anomalie, trendy.
- 

## Zalety użytej architektury:

- Skalowalność przetwarzania danych strumieniowych
  - Oddzielenie warstw: zbieranie, przetwarzanie, analiza, wizualizacja
  - Wykorzystanie gotowych komponentów Azure zgodnie z rekomendacjami DP-203 i DP-100
  - Integracja AI i analityki w jednym środowisku
-