

## **E-Portfolio Component – Faceted Data:**

Read Schmitz et al. (2016) article about faceted data.

- Do you think this is an excellent approach to protect systems from data leakage? What are the pros and cons?
- Create a basic outline design of how you would create such a system in Python.

Add your answers to your e-portfolio.

### **1. Do you think this is an excellent approach to protect systems from data leakage? What are the pros and cons?**

Schmitz et al. (2016) proposed a method for dynamic imposition of information flow security in Haskell using faceted computation. This approach does not alter language semantics and can execute secure multi-execution with a single process, improving performance. Implementing faceted values ensures that faulty or malicious programs cannot bypass dynamic protections. Schmitz et al. (2016) library confirms the conventional information flow security trait of termination-insensitive noninterference, regardless of any defects, susceptibilities, or malicious code in the client application.

The approach Schmitz et al. (2016) presented is an excellent way to prevent system data leakage. They proposed a faceted data system that divides the data into multiple facets and then stores each facet separately. This makes it challenging for attackers to access all the data at once. Even if they gain access to one facet, they still cannot see the whole picture.

AlexSpft (2021) mentioned that efficient system integration involves harmonising software and hardware components, resulting in a more productive, precise, and cost-effective ecosystem. There are various forms of system integration, including legacy,

enterprise application, third-party, and business-to-business. Legacy integration updates outdated systems, EAI unifies subsystems, third-party integration enhances functionality, and business-to-business connectivity links disparate systems. To achieve optimal results, businesses must consider their unique needs and select tailor-made solutions like APIs, webhooks, or middleware. System integration can be approached through different architectural models, each with advantages and disadvantages, such as faceted systems. For instance, below are the pros and cons of faceted systems:

**Pros:**

**Security:** Faceted data systems can help improve data security by making it more difficult for attackers to access all the data simultaneously.

**Privacy:** Faceted data systems can also help protect data privacy by making it more difficult for attackers to identify individuals or groups of individuals.

**Cons:**

**Complexity:** Faceted data systems can be more complex to design and implement than traditional ones.

**Performance:** Faceted data systems can hurt performance because data queries may need to access multiple facets.

## **2. Create a basic outline design of how you would create such a system in Python.**

McGrath (2023) defines a faceted system as a highly effective tool that allows users to leverage controlled vocabularies and metadata. It has gained widespread popularity in library discovery interfaces, delivering superior search results to users.

The faceted system has become a widely used technique for browsing and searching through collections of items by applying filters based on their properties. This approach enables users to effortlessly locate the exact thing they are interested in, even in cases where they are unsure of the specific keywords to use in their search (MongoDB, N.D.).

Seaborn is a powerful Python visualisation library that enables the creation of stunning statistical graphics using pre-set styles and colour palettes. It is built on matplotlib and pandas and offers dataset-oriented APIs for better data comprehension. One of its most impactful features is the ability to create facet plots, which would enable data to be broken down based on a categorical variable and displayed in similar plots with the same scale. Seaborn's `Catplot()` method is a popular tool for creating simple facet plots using numerical and categorical variables (GeeksforGeeks, (2020)). Additionally, Seaborn is a reliable Python library designed explicitly for statistical visualisation, with visually appealing default styles and colour palettes and seamless integration with pandas' data structures (Mujtaba, 2021). For instance, below is a basic outline design for creating a Faceted data system in Python based on the ideas of Kaggle (N.D.):

### **1. Install de Seaborn packages:**

```
pip install seaborn
```

### **2. Import libraries:**

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

### 3. Create Seaborn Plots, loading data:

```
data = pd.read_csv('data.csv')

data = pd.DataFrame({
    'x': np.random.randn(100),
    'y': np.random.randn(100),
    'category': np.random.choice(['A', 'B', 'C'], 100)
})
```

### 4. Create a facet plot using seaborn.displot:

```
import seaborn as sns
sns.catplot(
    x='x',
    y='y',
    data=data,
    col=facet_var
)
```

## References:

Schmitz, T., Rhodes, D., Austin, T.H., Knowles, K. & Flanagan, C. (2016). Faceted Dynamic Information Flow via Control and Data Monads. *Lecture Notes in Computer Science*, pp.3–23. doi:[https://doi.org/10.1007/978-3-662-49635-0\\_1](https://doi.org/10.1007/978-3-662-49635-0_1).

AltexSoft. (2021). *System Integration: Types, Approaches, and Implementation Steps*. [online] Available at: <https://www.altexsoft.com/blog/system-integration/>.

McGrath, K. (2023). Musings on Faceted Search, Metadata, and Library Discovery Interfaces. *Cataloging & Classification Quarterly*, pp.1–51. doi:<https://doi.org/10.1080/01639374.2023.2222120>.

MongoDB. (N.D.). *Faceted Search with MongoDB | MongoDB Blog*. [online] Available at: <https://www.mongodb.com/blog/post/faceted-search-with-mongodb>.

Mujtaba, H. (2021). *Seaborn Tutorial in Python for beginners | Data Visualization using Seaborn*. [online] GreatLearning Blog: Free Resources what Matters to shape your Career! Available at: <https://www.mygreatlearning.com/blog/seaborn-tutorial/>.

kaggle.com. (n.d.). *Data Science All in one Note book*. [online] Available at: <https://www.kaggle.com/code/alaatemimy/data-science-all-in-one-note-book>.

GeeksforGeeks. (2020). *How To Make Simple Facet Plots with Seaborn Catplot in Python?* [online] Available at: <https://www.geeksforgeeks.org/how-to-make-simple-facet-plots-with-seaborn-catplot-in-python/>.