POTENTIAL FOR MIS IN GOOGLE CLOUD AND GOOGLE BIG QUERY

Prof.dr. Dalia Kriksciuniene ; PV250 Marketing Information Systems

Masaryk University, Faculty of Informatics, Lasaris Lab, BRNO, 2019

Google Cloud Platform

- Google Cloud Platform, offered by Google, is a suite of cloud computing services that runs on the same infrastructure that Google uses internally for its end-user products, such as Google Search and YouTube.[1] Alongside a set of management tools, it provides a series of modular cloud services including computing, data storage, data analytics and machine learning.[2] Registration requires a credit card or bank account details.[3]

- Google Cloud Platform provides Infrastructure as a service, Platform as a service, and Serverless computing environments.

- In April 2008, Google announced App Engine, a platform for developing and hosting web applications in Google-managed data centers, which was the first cloud computing service from the company. The service became generally available in November 2011. Since the announcement of App Engine, Google added multiple cloud services to the platform.

Categories of Action

- [Cloud computing -](#) Compute, storage, databases
- Analytics and machine learning - BigQuery, machine learning APIs
- [Identity and security -](#) Cloud Identity, Cloud Security Command Center
- [Collaboration and productivity -](#) G Suite with Gmail, Docs, Drive, Hangouts
- [Google Maps Platform -](#) Maps, Routes, Places
- [Browser, hardware, and OS -](#) Chrome, Android, Jamboard
- [Services -](#) Consulting, training, certification

Big Data products - BigQuery

- [BigQuery](#) is Google's fully managed, low cost, serverless data warehouse that scales with your storage and computing power needs.

- With BigQuery, you get a columnar and ANSI SQL database that can analyze petabytes of data in seconds.

- REST-based APIs integrate with existing applications and developers can even create new apps using familiar programming languages like Java and Ruby. Analyze Geographic Information Systems (GIS) with SQL and BigQuery GIS Beta data in BigQuery. And quickly build and operationalize ML models on planet-scale structured or semi-structured data using simple SQL with BigQuery ML.

- Plus, BigQuery offers data transferring services, flexible data ingestion, and pay-for-what-you-use pricing.

Big Data products - Cloud Dataflow

- Cloud Dataflow is a fully-managed service for transforming and enriching data in stream (real time) and batch (historical) modes with equal reliability and expressiveness.

- Cloud Dataflow's serverless approach frees you from operational tasks like capacity planning, resource management, and performance optimization while only paying for what you use.

- Plus, Cloud Dataflow not only works with Google's ingestion, data warehousing, and machine learning products, but also third-party tools like Apache Spark and Apache Beam.

Big Data products - Cloud Dataproc

- Cloud Dataproc is a fast, easy-to-use, fully-managed cloud service for running Apache Spark and Apache Hadoop clusters in a simpler, more cost-efficient way.

- Operations that used to take hours or days take seconds or minutes instead — and you pay only for the resources you use with per-second billing.

- Cloud Dataproc integrates with storage, compute, and monitoring services across Google Cloud products, giving you a powerful and complete data processing platform.

Big Data products - Cloud Dataprep

- [Cloud Dataprep](#) by Trifacta is an intelligent data service for visually exploring, cleaning, and preparing structured and unstructured data for analysis.

- Because Cloud Dataprep is serverless and works at any scale, there is no infrastructure to deploy or manage.

- Your next ideal data transformation is suggested and predicted with each UI input, so you don't have to write code.

- And with automatic schema, datatype, possible joins, and anomaly detection, you can skip time-consuming data profiling and focus on data analysis.

Big Data products - Cloud Pub/Sub

- Cloud Pub/Sub is a simple, reliable, scalable foundation for stream analytics and event-driven computing systems. You can send and receive messages between independent applications and syndicate data across projects and applications running on cloud, on-premises, or hybrid environments.

- You can leverage Cloud Pub/Sub's flexibility to decouple systems and components hosted on GCP or elsewhere on the internet.

- And Cloud Pub/Sub is designed to provide "at least once" delivery at low latency with on-demand scaling to tens of millions of messages per second.

Big Data products - Cloud Composer

- [Cloud Composer](#) is a fully managed workflow orchestration service that empowers you to author, schedule, and monitor pipelines that span across clouds and on-premises data centers.

- Built on the popular Apache Airflow open source project and operated using the Python programming language, Cloud Composer is free from lock-in and easy to use. Plus, with end-to-end integration for GCP workloads, you can orchestrate a full pipeline with all of Google Cloud's big data products.

Big Data products - Google Data Studio

- Google Data Studio allows you to unlock the power of your data with interactive dashboards and engaging reports that inspire smarter business decisions.

- It makes insights easy to read, share, and customize. Integrations between BigQuery, Data Studio, and Sheets and popular BI and ETL tools translate to flexibility for how data is ingested and presented.

- Data Studio also uses Google Drive's sharing model, enabling teams to collaborate in real time.

- And with pre-built report templates, teams can focus on telling powerful stories, not handling data.

Big Data products - Cloud Data Transfer

- Cloud Data Transfer offers solutions that meet your unique data transfer needs and move data to Google Cloud Storage, BigQuery or Dataproc quickly and securely.

- Whether you have 50 gigabytes or 50 petabytes of data, one-time or recurring transfers, or a T1 line or a 10 Gbps network connection, our data transfer services are here to meet your unique needs.

Big Data products - Cloud Bigtable

- [Cloud Bigtable](#) provides a massively scalable NoSQL database suitable for low-latency and high-throughput workloads. It integrates easily with popular big data tools like Hadoop and Spark and it supports the open-source, industry-standard HBase API.

- Cloud Bigtable is a great choice for both operational and analytical applications, including IoT, user analytics, and financial data analysis.

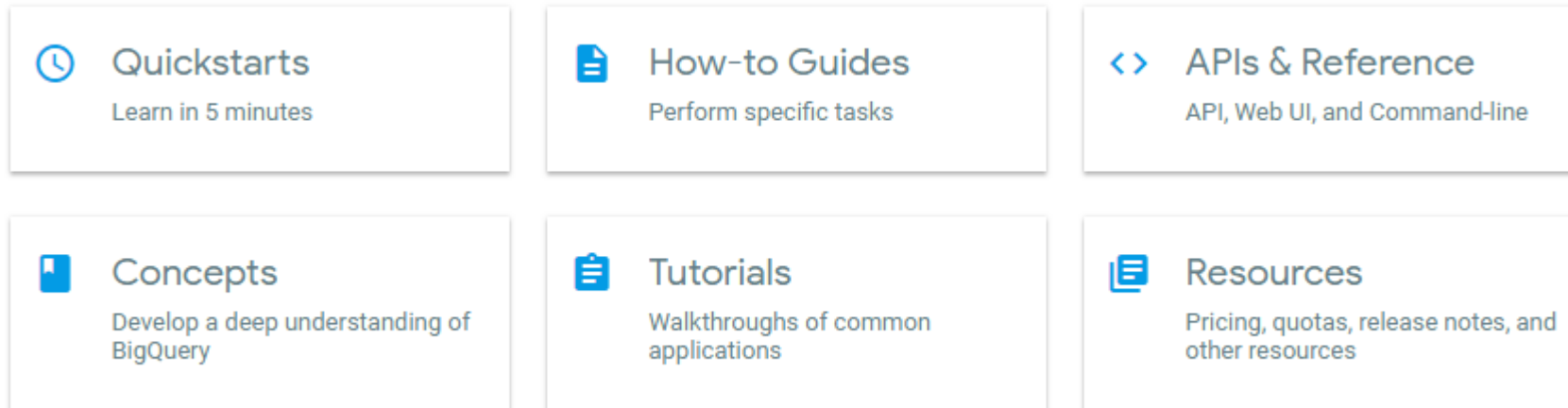# Big Data products - Google Cloud Storage

- [Google Cloud Storage](#) is a unified object storage solution that allows world-wide storage and retrieval of any amount of data at any time.

- You can use Cloud Storage for a range of scenarios, including serving website content, storing data for archival and disaster recovery, or distributing large data objects to users via direct download.

Big Data products - Cloud Datalab

- Cloud Datalab is a powerful interactive tool created to explore, analyze, transform, visualize data, and build machine learning models on Google Cloud Platform.

- It is an interactive notebook based on Jupyter and it's integrated with BigQuery and Cloud Machine Learning Engine to provide easy access to key data processing services.

- And with TensorFlow or Cloud Machine Learning Engine, you can easily turn data into deployed machine-learning models ready for prediction.

Google BigQuery Documentation

- BigQuery is Google's fully managed, petabyte scale, low cost analytics data warehouse. BigQuery is NoOps—there is no infrastructure to manage and you don't need a database administrator—so you can focus on analyzing data to find meaningful insights, use familiar SQL, and take advantage of our pay-as-you-go model.

Quickstarts Learn in 5 minutes	**How-to Guides** Perform specific tasks	**APIs & Reference** API, Web UI, and Command-line
Concepts Develop a deep understanding of BigQuery	**Tutorials** Walkthroughs of common applications	**Resources** Pricing, quotas, release notes, and other resources

https://cloud.google.com/bigquery/docs/

START from
https://cloud.google.com/bigquery/docs/

BigQuery

Product Overview

Documentation

Quickstarts

All Quickstarts

Using the Web UI 🧪

Using the Command-Line Tool

Using the Client Libraries

Using the Classic Web UI

How-to Guides

All How-to Guides

▸ Interacting with BigQuery

▸ Running and Managing Jobs

▸ Working with Datasets

▸ Working with Table Schemas

▸ Working with Tables

▸ Working with Partitioned Tables

▸ Working with Clustered Tables 🧪

▸ Working with Views

Creating and Managing Labels

▸ Loading Data Into BigQuery

▸ Querying BigQuery Data

▸ Controlling Costs

▸ Using External Data Sources

Protecting Data with Cloud KMS Keys

▸ Monitoring and Logging

▸ BigQuery API Basics

Tutorials

All Tutorials

Creating an Authorized View in BigQuery

Visualizing BigQuery Data Using Google Data Studio

Visualizing BigQuery Data in a Jupyter Notebook

Importing Firebase Analytics Data into BigQuery

Real-time logs analysis using Fluentd and BigQuery

Analyzing Financial Time Series using BigQuery