

Case for a **Media Holding**

A German Media Giant Enables Advanced Analytics with a Big Data Platform on Google Cloud_

Industry: TV and broadcasting

Location: Germany

Employees: 16,000+



Client Background

A leader in broadcasting and digital content, owning 50+ TV channels, streaming platforms, and radio stations. The media holding provides a comprehensive digital video network and delivers different types of content across the globe.

Business Challenge

A leader in the creation and broadcasting of digital content had a classic on-premises data warehouse that was cost-intensive for big data collection while also resource- and time-intensive for data analysis. The client's existing analytics tools could not provide a comprehensive view of their media content to cover the needs of their customers.

The innovation-driven broadcaster needed to build a novel data platform capable of handling massive amounts of data (including real-time intel) with ease to administer and manage. Among the key goals were:

- To optimize storing and analysis of big volumes of media content (especially large video files) and enable their quick processing in the cloud.
- To build a hybrid infrastructure that would allow for storing high-resolution media files and other critical files on-premises while performing their analysis in the cloud in full correspondence with the client's security requirements and strict copyright regulations.
- To implement an intuitive recommendation engine powered by machine learning and provide viewers with simple access to targeted content and accurate suggestions based on the analyzed preferences.

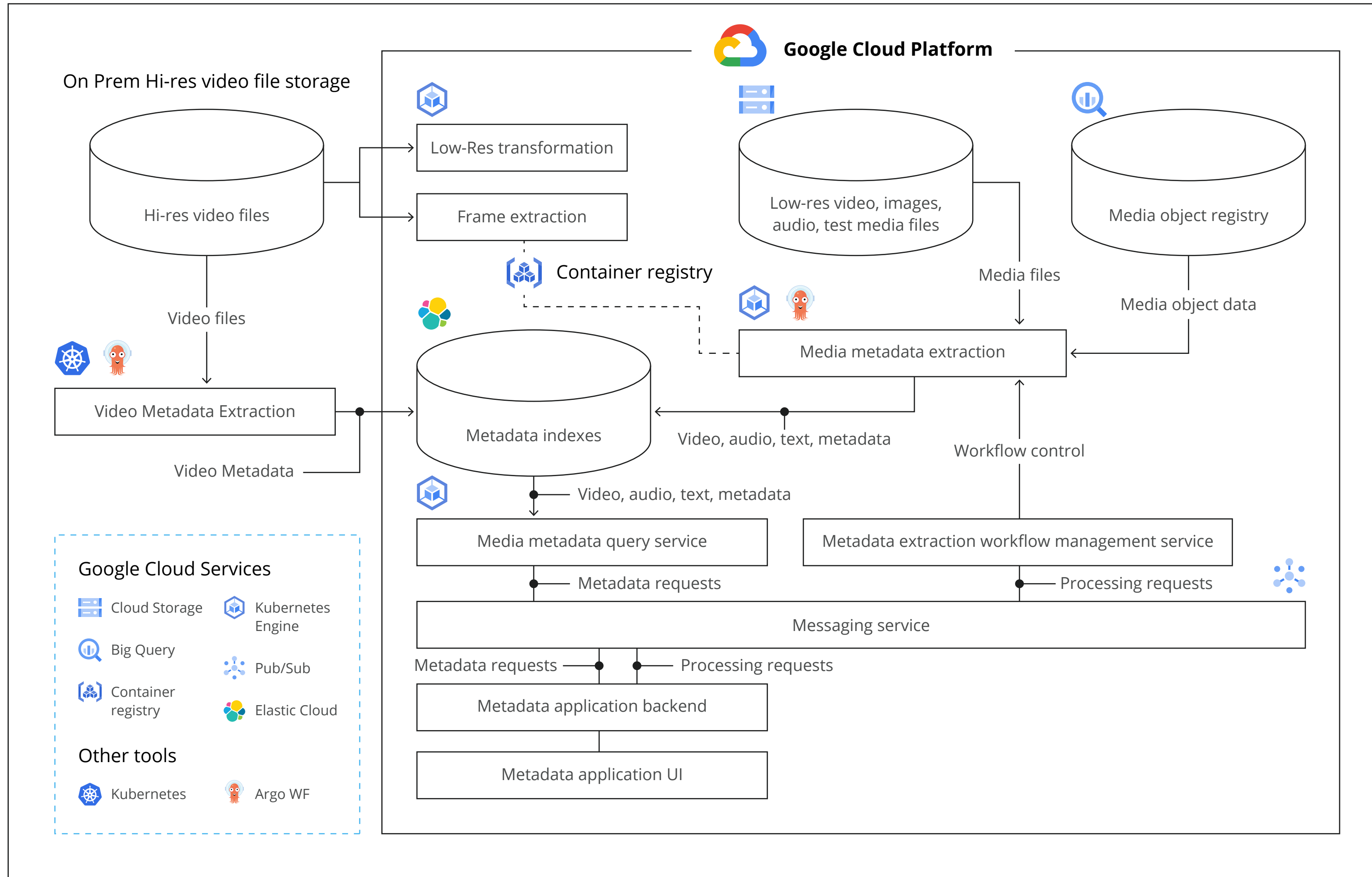
Solution

Infopulse, as an experienced cloud services provider with profound expertise in Google Cloud Platform (GCP) and [advanced data analytics](#), helped the German media holding develop their data platform on GCP that matches the client's requirements in exceptional performance and flexibility in resource management. Google Cloud Platform was selected both for its big data capacity and also for offering a fully-managed infrastructure with an independent allocation of cloud resources, the ability to scale up/down, and self-maintenance capabilities. Besides, Google's tech stack presents extensive machine learning capabilities for building a recommendation engine for the viewers.

Here's how Infopulse's data engineering team delivered a robust and powerful data platform to the media broadcaster:

- Infopulse built a system for metadata collection (e.g., a movie scene description, movie frames, movie content, technical parameters, etc.), used for a quick search of information in a multitude of available media files.
- To swiftly retrieve such metadata, we developed **Machine Learning scripts** with the help of Google Cloud ML capabilities.
- To analyze massive volumes of video files, the team used **on-premises GPU** as it was more cost-efficient in contrast to processing large-scale videos in the cloud.
- Enabled the storing of metadata and low-resolution videos and films with watermarks in **Cloud Storage** that ensures the highest level of flexibility, scalability, and resource optimization. For keeping the structured data, mostly the common metadata of the motion picture (release date, director, genre, etc.), we used **BigQuery** as a relational SQL-based DWH. For holding petabyte-scale structured data, the team utilized **Bigtable** as a distributed NoSQL database.
- In view of the **hybrid nature of the platform**, we also enabled the housing of the high-resolution media files and the broadcaster's movies on-premises, considering the client's data security requirements.
- To ensure the seamless integration and robust messaging-oriented communication between cloud and on-premises services, our team established **Pub/Sub** middleware that allowed overcoming a number of integration issues while offering the highest throughput.

The Big Data Platform Architecture on Google Cloud Platform



- Set up an indexing system resting on Google’s partner solution **Elastic Cloud Search**, already integrated into the GCP ecosystem, to provide the client with a quick metadata search based on specific attributes.
- Built a recommendation engine for viewers that could analyze the metadata of media files, the users’ tastes and preferences, and provide them with more relevant suggestions.
- To simplify the solution development and deployment and successfully perform metadata processing, our team relied on microservices orchestrated by **Google Kubernetes Engine (GKE)**. We used **Docker containers** to package Machine Learning scripts and **Argo Workflow** with GKE to run scripts sequentially or in parallel.

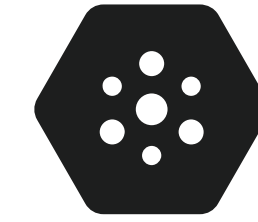
Technologies



Google Cloud Platform



Big Query



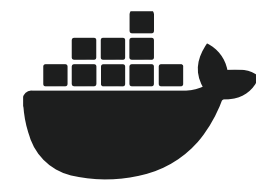
Cloud Pub/Sub



Argo WF



Google Kubernetes Engine



Docker



Elasticsearch



Business Value

As a result of our project, the German media holding now has a **highly flexible, scalable, well-performing data platform** for both storing and analyzing their massive media data sets.

- Due to the hybrid infrastructure, the client can provide the necessary level of data security and quickly analyze petabyte-scale data in real time.
- With the implemented Machine Learning algorithms and a recommendation engine in place, the client's viewers can benefit from accurate recommendations and improved user experience. All this results in enhanced customer engagement and loyalty.
- Moreover, owing to the big data and self-management capabilities of Google Cloud Platform, the client optimized costs for big data warehousing and resource allocation, as well as facilitated the administration and management of the data platform infrastructure.



About Infopulse

Infopulse, part of the leading Nordic digital services company Tietoevry, is an international vendor of services in the areas of Software R&D, Application Management, Cloud & IT Operations, and Cybersecurity to SMEs and Fortune 100 companies across the globe. Founded in 1991, the company has a team of over 2,300 professionals and is represented in 7 countries across Europe and the Americas.

Infopulse is trusted by many established brands, such as BICS, Bosch, British American Tobacco, Credit Agricole, Delta Wilmar, ING Bank, Microsoft, Offshore Norge, OLX Group, OTP Bank, SAP, UkrSibbank BNP Paribas Group, Vodafone, Zeppelin Group, and others.

For more information, please visit www.infopulse.com

Contact us

PL +48 (221) 032-442

DE +49 (69) 505-060-4719

US +1 (888) 339-75-56

UK +44 (8455) 280-080

UA +38 (044) 585-25-00

BG +359 (876) 92-30-90

BR +55 (21) 99298-3389

 info@infopulse.com

