

Kunde: **Medienholding**

Ein deutscher Medienriese nutzt fortschrittliche Analysen mit Hilfe einer Big Data-Plattform in der Google Cloud_

Branche: Fernsehen und Rundfunk

Ort: Deutschland

Mitarbeiter: Über 16.000



Über den Kunden

Ein führendes Unternehmen im Bereich Rundfunk und digitale Inhalte mit mehr als 50 Fernsehkanälen, Streaming-Plattformen und Radiosendern. Die Medienholding bietet ein umfassendes digitales Videonetz und liefert verschiedenste Arten von Inhalten in die ganze Welt.

Anforderung

Ein führendes Unternehmen im Bereich der Erstellung und Ausstrahlung digitaler Inhalte verfügte über ein klassisches lokales Data Warehouse, das für die Erfassung großer Datenmengen kostenintensiv und für die Datenanalyse ressourcen- und zeitaufwendig war. Die vorhandenen Analysetools des Kunden konnten keinen umfassenden Überblick über seine Medieninhalte bieten, um den Bedürfnissen seiner Kunden gerecht zu werden.

Der innovationsorientierte Sender benötigte eine neuartige Datenplattform, die in der Lage ist, riesige Datenmengen (einschließlich Echtzeitinformationen) zu verarbeiten und dabei einfach zu verwalten und zu managen. Zu den wichtigsten Zielen gehörten:

- Optimierung der Speicherung und Analyse großer Mengen von Medieninhalten (insbesondere großer Videodateien) und Ermöglichung ihrer schnellen Verarbeitung in der Cloud.
- Aufbau einer hybriden Infrastruktur, die es ermöglicht, hochauflösende Mediendateien und andere kritische Dateien vor Ort zu speichern, während die Analyse in der Cloud durchgeführt wird - und das alle in voller Übereinstimmung mit den Sicherheitsanforderungen des Kunden und den strengen Urheberrechtsbestimmungen.
- Implementierung einer intuitiven, auf Machine Learning basierenden Empfehlungsmaschine, die den Zuschauern einen einfachen Zugang zu gezielten Inhalten und präzisen Vorschlägen auf der Grundlage der analysierten Präferenzen bietet.

Lösung

Infopulse, als erfahrener Anbieter von Cloud Services mit profunder Expertise im Bereich Google Cloud Platform (GCP) und [fortschrittlicher Datenanalyse](#), unterstützte die deutsche Medienholding bei der Entwicklung ihrer Datenplattform auf GCP, die den Ansprüchen des Kunden an außergewöhnliche Leistung und Flexibilität im Ressourcenmanagement gerecht wird. Google Cloud Platform wurde sowohl wegen seiner Big-Data-Kapazitäten als auch wegen der vollständig verwalteten Infrastruktur mit unabhängiger Zuweisung von Cloud-Ressourcen, der Möglichkeit der Skalierung nach oben oder unten und der Fähigkeit zur Selbstwartung ausgewählt. Darüber hinaus bietet Googles Technologie-Stack umfangreiche Möglichkeiten des Machine Learning zum Aufbau einer Empfehlungsmaschine für die Nutzer.

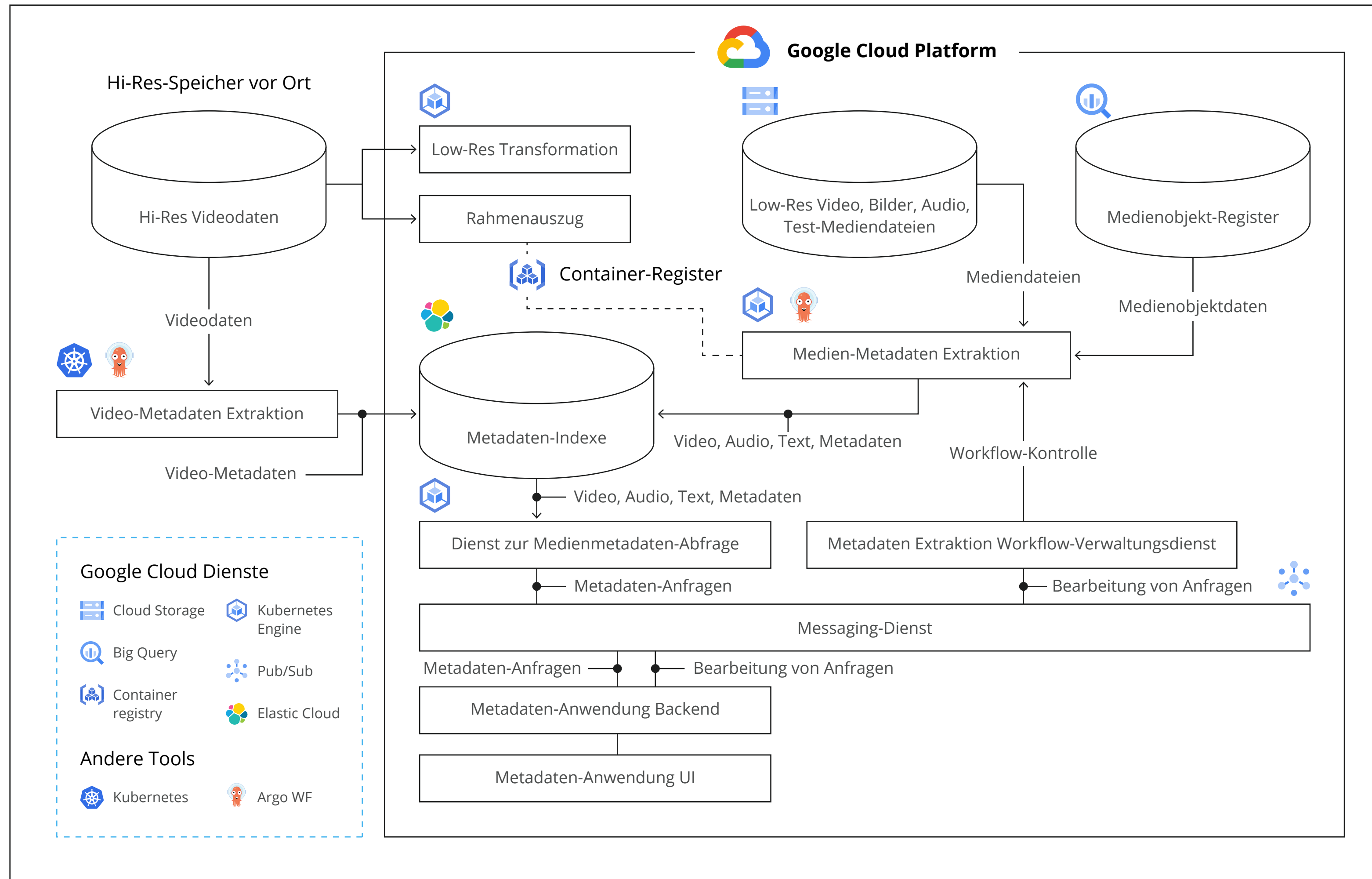
Hier erfahren Sie, auf welche Weise das Data-Engineering-Team von Infopulse dem Medienunternehmen eine robuste und leistungsstarke Datenplattform zur Verfügung stellte:

- Infopulse entwickelte ein System zur Erfassung von Metadaten (z.B. Beschreibung von Filmszenen, Filmbildern, Filminhalten, technischen Parametern usw.), die für eine schnelle Suche nach Informationen in einer Vielzahl von verfügbaren Mediendateien verwendet werden.
- Um solche Metadaten schnell abrufen zu können, entwickelten wir mit Hilfe der Google Cloud ML-Funktionen **Skripte für Machine Learning**.
- Zur Analyse großer Mengen von Videodateien setzte das Team **GPUs vor Ort** ein, da dies im Gegensatz zur Verarbeitung großer Videos in der Cloud kosteneffizienter war.
- Die Speicherung von Metadaten und niedrig aufgelösten Videos und Filmen mit Wasserzeichen in einem **Cloud-Speicher**, der ein Höchstmaß an Flexibilität, Skalierbarkeit und Ressourcenoptimierung gewährleistet, wurde ermöglicht. Zur Aufbewahrung

der strukturierten Daten, vor allem der allgemeinen Metadaten des Films (Erscheinungsdatum, Regisseur, Genre usw.), setzten wir **BigQuery** als relationales SQL-basiertes DWH ein. Zur Speicherung strukturierter Daten im Petabyte-Bereich nutzte das Team **Bigtable** als verteilte NoSQL-Datenbank.

- Angesichts des **hybriden Charakters der Plattform** schufen wir auch die Möglichkeit, die hochauflösenden Mediendateien und die Filme des Senders vor Ort unterzubringen, wobei die Datensicherheitsanforderungen des Kunden berücksichtigt wurden.
- Um eine nahtlose Integration und eine robuste, nachrichtenorientierte Kommunikation zwischen Cloud- und Vor-Ort-Services zu gewährleisten, richtete unser Team eine **Pub/Sub**-Middleware ein, die es ermöglichte, eine Reihe von Integrationsproblemen zu überwinden und gleichzeitig den höchsten Durchsatz zu bieten.

Die Big-Data-Plattform-Architektur auf Google Cloud Platform



- Einrichtung eines Indexierungssystems, das auf der bereits in das GCP-Ökosystem integrierten Google-Partnerlösung **Elastic Cloud Search** basiert, um dem Kunden eine schnelle Metadatenuche auf der Grundlage bestimmter Attribute zu ermöglichen.
- Entwicklung einer Empfehlungsmaschine für Zuschauer, mit der die Metadaten von Mediendateien sowie der Geschmack und die Vorlieben der Nutzer analysiert werden können, um ihnen relevantere Vorschläge zu unterbreiten.
- Um die Entwicklung und Bereitstellung der Lösung zu vereinfachen und die Verarbeitung von Metadaten erfolgreich durchzuführen, stützte sich unser Team auf Microservices, die von **Google Kubernetes Engine (GKE)** orchestriert werden. Wir verwendeten **Docker-Container**, um Skripte für Machine Learning zu verpacken, und **Argo Workflow** mit GKE, um Skripte sequentiell oder parallel auszuführen.

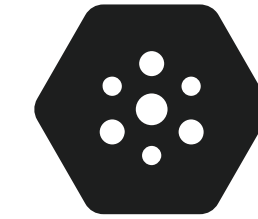
Technologien



Google Cloud Platform



Big Query



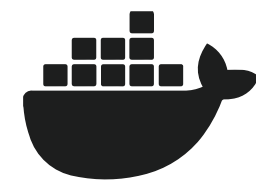
Cloud Pub/Sub



Argo WF



Google Kubernetes Engine



Docker



Elasticsearch



Ergebnis

Als Ergebnis unseres Projekts verfügt die deutsche Medienholding nun über eine **hochflexible, skalierbare und leistungsfähige Datenplattform** zur Speicherung und Analyse ihrer umfangreichen Mediendatenbestände.

- Dank der hybriden Infrastruktur kann der Kunde das erforderliche Maß an Datensicherheit gewährleisten und Daten im Petabyte-Bereich schnell und in Echtzeit analysieren.
- Dank der implementierten Algorithmen für Machine Learning und einer Empfehlungsmaschine können die Zuschauer des Kunden von präzisen Empfehlungen und einer verbesserten Nutzererfahrung profitieren. All dies führt zu einer stärkeren Kundenbindung und Loyalität.
- Dank der Big-Data- und Self-Management-Funktionen der Google Cloud Platform konnte der Kunde außerdem die Kosten für Big Data Warehousing und die Ressourcenzuweisung optimieren sowie die Verwaltung und das Management der Datenplattform-Infrastruktur erleichtern.



Über Infopulse

Infopulse, Teil des führenden nordischen, digitalen Dienstleistungs-Unternehmens Tietoevry, ist ein internationaler Anbieter von Dienstleistungen in den Bereichen Software-F&E, Anwendungsmanagement, Cloud- und IT-Betrieb sowie Cybersicherheit für KMUs und Fortune 100 Unternehmen auf der ganzen Welt. Das in 1991 gegründete Unternehmen verfügt über ein Team von über 2,300 Fachleuten und ist weltweit in 7 Ländern – in Europa sowie in Nord-, Mittel- und Südamerika – vertreten.

Infopulse genießt das Vertrauen vieler etablierter Marken wie Allianz Bank, BICS, Bosch, British American Tobacco, Credit Agricole, Delta Wilmar, ING Bank, Microsoft, Norwegian Oil and Gas Association, OLX Group, OTP Bank, SAP, UkrSibbank BNP Paribas Group, Vodafone, Zeppelin Group und vieler anderer.

Für weitere Informationen besuchen Sie bitte www.infopulse.com/de

Kontaktieren sie uns

DE +49 (69) 505-060-4719

UA +38 (044) 585-25-00

US +1 (888) 339-75-56

UK +44 (8455) 280-080

PL +48 (606) 291-154

BG +359 (876) 92-30-90

BR +55 (21) 99298-3389

 info@infopulse.com

