



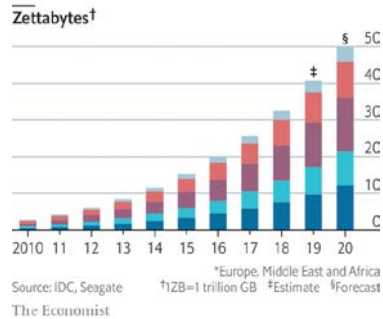
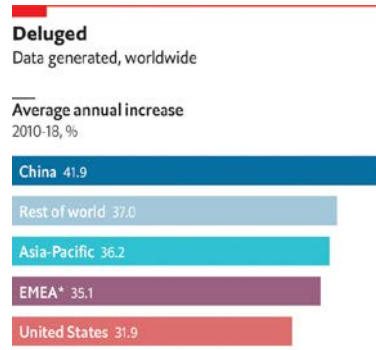
# NEXT-GEN DATA MANAGEMENT, ARCHIVING AND PRESERVATION

EGI CONFERENCE 02 NOV 2020

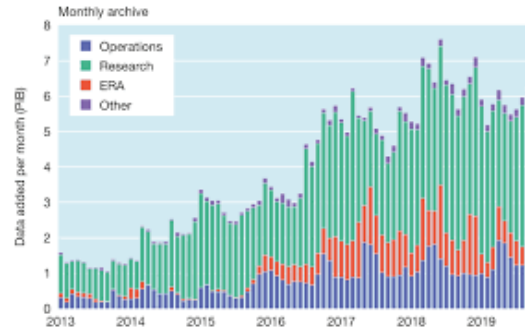
Jurry de la Mar

**T** · · Systems ·

# Data Management, Archiving and Preservation Next-Generation Challenges



**Global Data Increase  
2010 - 2020**



**Ex.1: ECMWF Data Archive Increase**

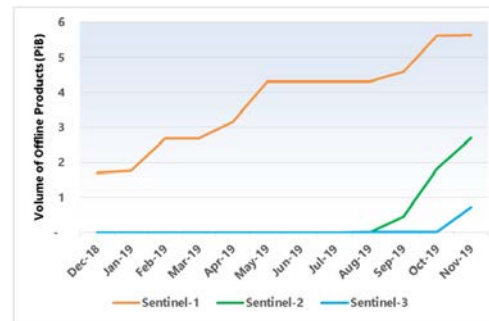


Figure 61: Cumulative growth of Offline product volume (PiB) per mission

**Ex.2: Sentinel Data Archive Increase**

**HOW TO EVOLVE FROM  
DATA SILOS  
AND  
„BLACK HOLE“ ARCHIVES  
TOWARDS  
DYNAMIC AND COST-EFFICIENT  
SOLUTIONS  
FOR  
DATA LIFECYCLE  
MANAGEMENT**

# Problem



***Communities growing; Data volumes growing; Needs of advanced functionality increase***

***Objective: perform R&D to demonstrate long-term preservation and archiving services for scientific data in the PB range under F.A.I.R. principles, using commons solutions, while research groups keep total stewardship of their data sets***

- ❄ Standardize resulting services under the OAIS reference model
- ❄ Hybrid Deployments
- ❄ PB data volume; High sustained ingest rates (1-10 Gbps)
- ❄ Access through GÉANT network; Federated AAI
- ❄ Cost-effective and transparent business models
- ❄ Reduce fragmentation expanding existing solutions to several scientific domains

# T-Systems Starting Point.

## EXPERIENCE



Providing Commercial Archiving services since 1992 (ImageMaster). Various FP7 and H2020 projects.

**STATE-OF-THE-ART.**

## COMMUNITY



Established partner in leading and thriving communities to collaborate with users and developers.

**STRONGEST TEAM.**

## PUBLIC CLOUD



Open Telekom Cloud is a leading public cloud service from Germany, scalable, secure and cost-effective.

**SIMPLE, SECURE, AFFORDABLE.**

## NETWORKS

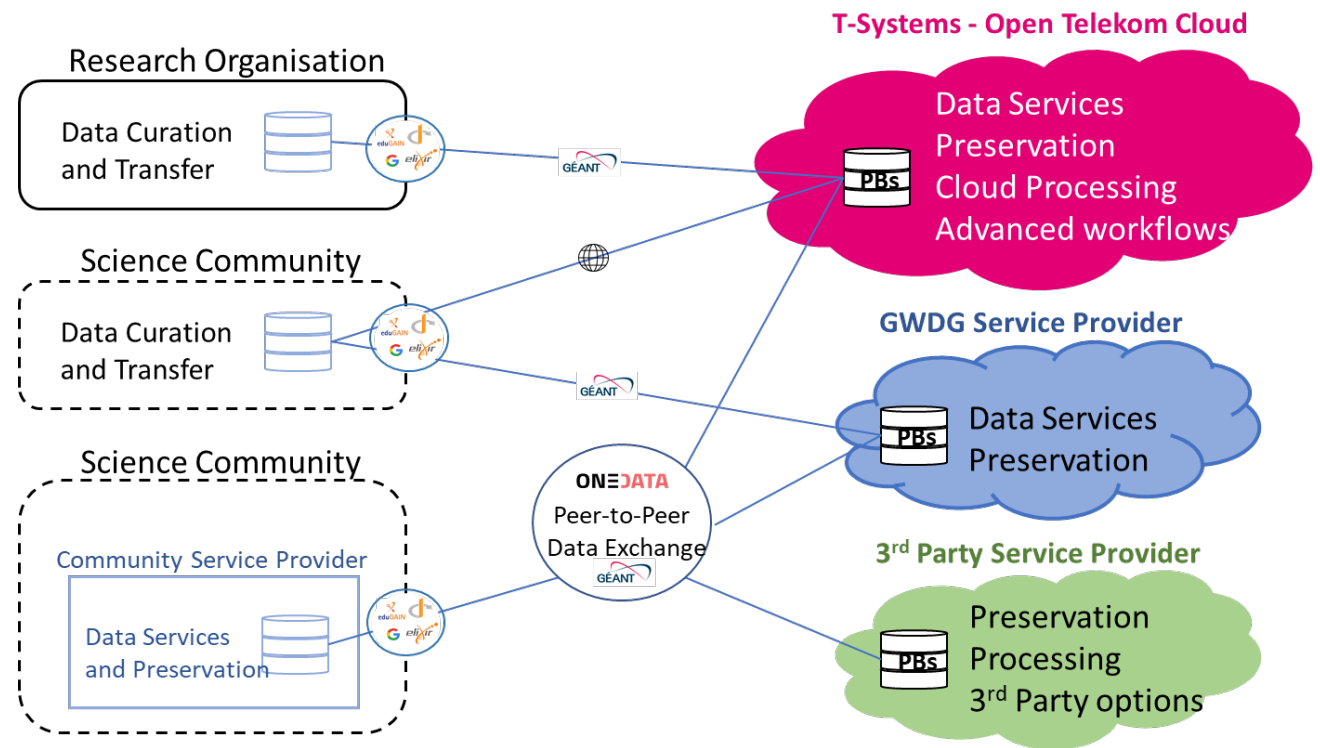


Registered GÉANT IaaS provider and established nx10G network peering.

**BEST ACCESS.**

# Design Phase results compared vs. State-of-the-Art

- Modern – full open-source - data access and management platform, supporting hybrid deployments and automated push/pull ingestion from legacy data sources
- Creation of arbitrary size (OAIS-)standards compliant archive-packages in cost-effective manner
- Highly scalable and flexible metadata extraction and data discovery framework based on OpenFaaS and ElasticSearch
- Support for flexible setup and control of archiving pipelines based on BPMN/Flowable
- Fully deployable on Kubernetes



## All Services

## My Favorites



## Computing



Elastic Cloud Server (10)  
Elastic, scalable computing servers



Bare Metal Server (0)  
Provides dedicated physical servers for tenants



Image Management Service (2)  
Self-service image management



Cloud Container Engine (0)  
container service that features high availability and elastic scalability



Auto Scaling (0)  
Dynamically adjusts computing resources



Dedicated Host (0)  
Dedicated physical servers

## Storage



Elastic Volume Service (12)  
Elastic, scalable block storage



Cloud Server Backup Service (1)  
Secure, reliable cloud server backup



Storage Disaster Recovery Service (0)  
Storage disaster recovery service



Volume Backup Service (1)  
Secure, reliable block storage backup



Object Storage Service  
Scalable cloud storage



Scalable File Service (0)  
Elastic, scalable file storage

## Network



Virtual Private Cloud (3)  
Provides securely isolated virtual networks



Elastic Load Balancing (0)  
Distributes traffic across multiple ECSs



Direct Connect (0)  
Provides high-speed, stable network access services



Private Link Access Service (0)  
High-quality, secure and dedicated network access service



Domain Name Service  
Stable, secure, fast domain name resolution



NAT Gateway (0)  
provides source NAT service



Virtual Private Network (0)  
Enables remote secure access to VPC networks



CDN (Akamai)  
Easy-to-use, reliable, quick content distribution



Elastic IP (10)  
Flexible public network access

## Security



Anti-DDoS  
Provides Anti-DDoS protection



Web Application Firewall  
Filters malicious web traffic



Key Management Service (0)  
Easily manage the keys used to encrypt your data

## Management &amp; Deployment



Cloud Eye  
Resource monitoring and alarm notification



Identity and Access Management  
Manages user access and encryption keys



Resource Template Service  
Provides orchestration for resources



Cloud Trace Service  
Records operations performed on cloud resources



Log Tank Service  
Log collection, query, and storage



Tag Management Service  
Efficient resource management with tags

## Application



**NEW Archiving and Preservation**



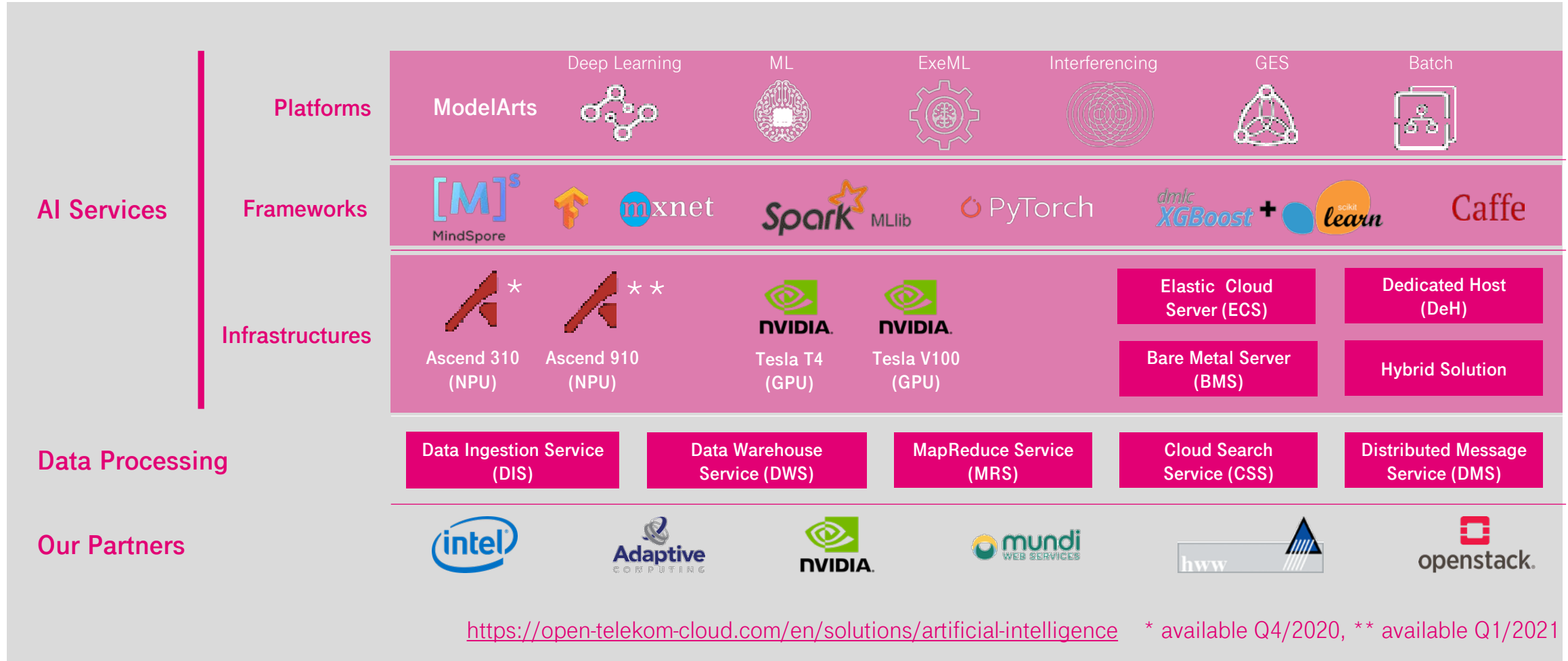
Simple Message Notification  
Provides simple and reliable message notification service



Software Repository for Container (0)  
Secure and reliable container image management

# Open Telekom Cloud

## Complementary AI-Portfolio for Data Analytics



**THANK YOU**



**LIFE IS FOR SHARING.**