



OSDU™ Data Platform on AWS

Wlad Frazao

Energy Principal for Americas

osdu-on-aws@amazon.com

wfrazao@amazon.com

<https://www.linkedin.com/in/wladfrazao/>

There's a lot of data in the Energy industry



SECRETARIA DE
GEOLOGIA, MINERAÇÃO E
TRANSFORMAÇÃO MINERAL

MINISTÉRIO DE
MINAS E ENERGIA



ACESSO GRATUITO AOS DADOS PÚBLICOS TERRESTRES

Para uma melhor experiência de transferência dos arquivos, recomendamos fortemente, que se evite selecionar e transferir vários grupos de arquivos simultaneamente.

Todos os dados estarão permanentemente disponíveis.








Considere navegar entre os diretórios e de modo criterioso, selecionar especificamente os dados de seu interesse.

Transferir simultaneamente vários pacotes, potencialmente poderá afetar a experiência de uso, alguns grupos de arquivos são superiores a 100 GB.

 <p>BACIA DO ACRE - MADRE DE DIOS ATUALIZADO EM : 05/05/2021</p> 	 <p>BACIA DE ALAGOAS ATUALIZADO EM: 05/05/2021</p> 	 <p>BACIA DO AMAZONAS ATUALIZADO EM: 05/05/2021</p> 
 <p>BACIA DO ARARIPE ATUALIZADO EM: 05/05/2021</p> 	 <p>BACIA DE BARREIRINHAS ATUALIZADO EM: 16/05/2022</p> 	 <p>BACIA DE BRAGANÇA - VIZEU ATUALIZADO EM: 05/05/2021</p> 
 <p>BACIA DO ESPÍRITO SANTO ATUALIZADO EM: 05/05/2021</p> 	 <p>BACIA DE JATOBÁ ATUALIZADO EM: 05/05/2021</p> 	 <p>BACIA DO MARAJÓ ATUALIZADO EM: 05/05/2021</p> 

<https://reate.cprm.gov.br/anp/TERRESTRE>

I mean, a LOT of data

 >	<input type="checkbox"/> Name ▲		Size	Modified
	ESTUDOS	⋮	33.1 MB	2 years ago
	GEOQUIMICA	⋮	6 GB	2 years ago
	NAO_SISMICOS	⋮	465.4 GB	2 years ago
	POCO	⋮	153.4 GB	2 years ago
	SISMICA_2D	⋮	28.6 GB	2 years ago
	SISMICA_3D	⋮	578.8 MB	2 years ago
<small>2c7c9338e 2985e71ac f48ef62e6 ff547d53e e673ab8f1</small>	md5_catalogo_Bacia_do_Parnaiba_1618661601.txt	⋮	5.2 MB	2 years ago
6 folders and 1 file			653.9 GB	

A lot of "interesting" formats as well

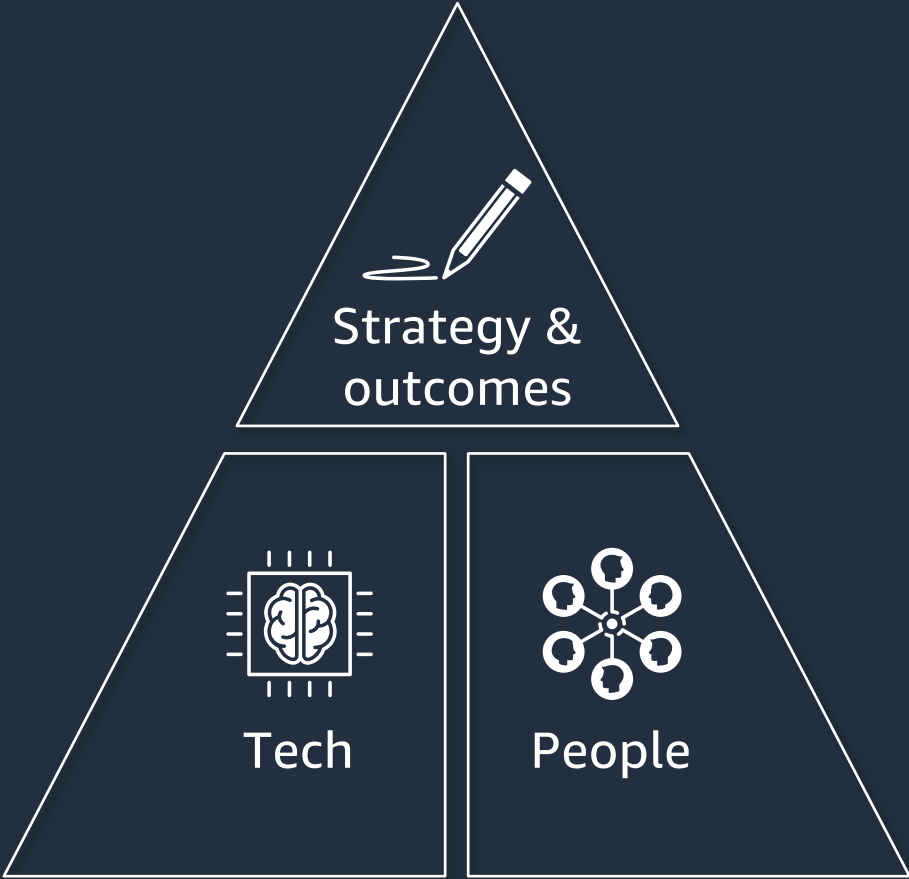
10F3

SIDEWALL CORE DESCRIPTION		COMPANY: Esso Prospecção	WELL: 1-AT-1-
Run No: 1	Type:	DATE: 1 NOV 84	GEOLOGICAL: JCH
DEPTH	REC.	LITHOLOGIC DESCRIPTION	POROSITY
1180m	5mm	MICA SCHIST: svy gy, dkgn-gy, firm to mod hd, w/ dev schistose foliation, abnt musc & biot, occ stabs chlor, com thin Qtzite beds, wht, v hd, Fxln	NUP
1179m	5mm	QTZITE: wht, hd, brit, VF-Fxln, com lam chlor w occ mica, relict thin beds crst: gn-dkdk, mod hd, blk, ncalc	NUP
1178m	NR		
1173m	10mm	SS: gn-gy, fri, F-MG, w scat CG, SR-SA, mod srted, qtz, cly ip	F-GVP
1171.3m	12mm	SH: gn-gy, sft to frm, sub fiss, scat sd grs, ncalc, occ ss lam, wht, fri, VFG, SR, w/ srted, qtz, cly ip, ncalc, EUP	
1168.7m	25mm	SH: gn-gy, sft to frm, fiss, ncalc	
1166.7m	5mm	SS: wht, fri to uncon, FG, SR, v w/ srted, qtz, ncalc	GVP
1163.5m	14mm	SH: fm, frm, fiss, ncalc, thin SCTST lam, wht	
1150.7m	12mm	SH: mg, frm, fiss, ncalc, thin wht SCTST lam	
1145m	25mm		

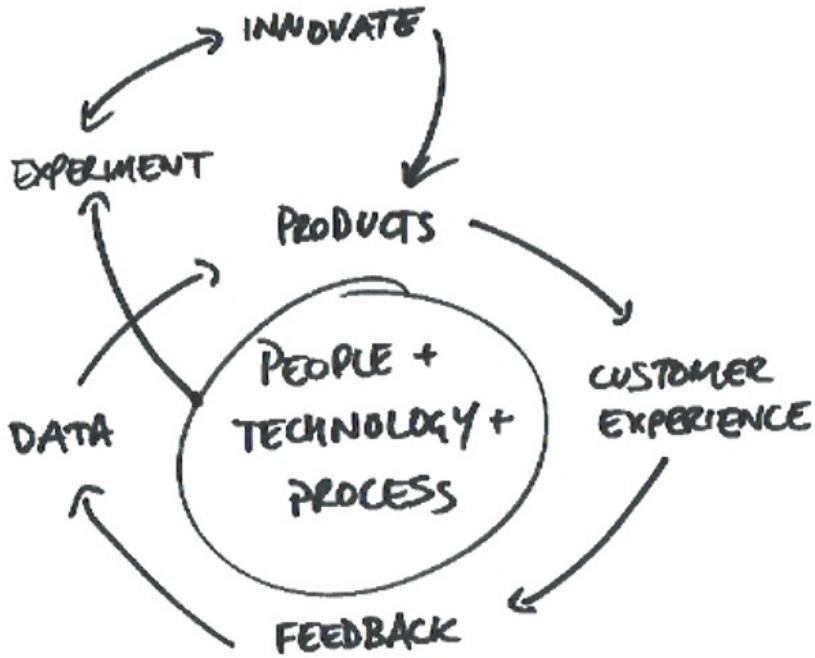
10F3

WELLSITE SAMPLE DESCRIPTION		COMPANY: Esso Prospecção	WELL (ONSHORE/OFFSHORE): 1-AT-1-MA
HOLE SIZE: 8 1/2"	GEOLOGICAL: JCH	DATE: 30 Oct 84	COUNTRY: Brasil
DEPTH FT	LITH %	LITHOLOGIC DESCRIPTION	SHOWS & REMARKS
1040-45	90	SS: wht, uncon, M-VFG, SR-SA, p to mod srted, qtz	N.S
	10	SH: m to dk gy, bn, gn-gy, v. fiss, slty ip, ncalc, frm	
	TR	SCTST: lt to mg gy, mod hd, ncalc, occ micac	
	TR	SH: lt gy, v. hd, VF Fxln	
	TR	SS: wht, v. hd, VFG, SR, v w/ srted, qtz, abnt sil cmt, ncalc, NUP	
1045-50	90	SS: A/A	N.S
	10	SH: A/A	
	TR	SCTST	
	TR	SS: VFG A/A	
1050-55	70	SH: A/A	
	20	SCTST: A/A gradg to SS	
	10	SS: VFG A/A, occ micac occ micac	
1055-60	50	SS: wht to lt gy, hd to brittle, VFG, SR, v w/ srted, qtz, mica, ncalc, no cmt, NUP	
	30	SCTST: crm, v. hd, sdy, ncalc	
	20	SH: A/A	
	TR	PVR	
1060-70	SS: A/A		

Becoming a Data Driven Company



99% of blue-chip companies are investing in data
24% have successfully created a data-driven organization



One day in the life of Anna, the seismic data manager

Anna leads a team of seismic data managers and **has a 5 PB collection of seismic data.**

Seismic workflow requires **multiple tools** from **multiple vendors** with **different data formats**, creating a number of conversions and export/import operations as well as multiple copies of the data.

Each tool has evolved with specialized data formats optimized for performance to best implement their process, and Anna's team uses dozens of different tools. **Her team spends around 60% of their time searching for quality data.**

The typical application runs on **large workstations that cannot leverage HPC.**

Anna's team Technology Landscape



The data perspective from this technology landscape

- Data duplication and siloed
 - No single source of truth
 - Limited storage capacity
- No common nomenclature
- Data ownership and governance
 - Loss of control
 - No documented management practices
- Information locked into proprietary formats
 - no search and query
 - lacking spatial mapping
- Data accessibility
 - Metadata locked inside proprietary files
 - No process to extract attributes
- Few analytics capabilities

Some questions Anna would like answered

- I want to find all the wells and logs that are similar to the one I am designing, so that I can quickly build a well delivery plan. For that, I have to answer questions like these:
- Show me Wells with specific attributes:
 - spud date
 - in a geographic area
 - By trajectory type
- Show me the production rate for my wells
 - What are the average production rates for each well
 - Average production rate by Operator
 - Production rate by operator in a geographical area
 - Heat map with production rate by operator by time
- Show me a detailed wellog location for my wellbore of interest
 - I want to download the file and run ML experiments with it and all the associated data
 - I want to investigate and plot some curves from specific wellogs

What Is OSDU Forum And Its Mission?

Working with The Open Group to create certifiable standards to support interoperability

Using the Open Source model to promote collaboration

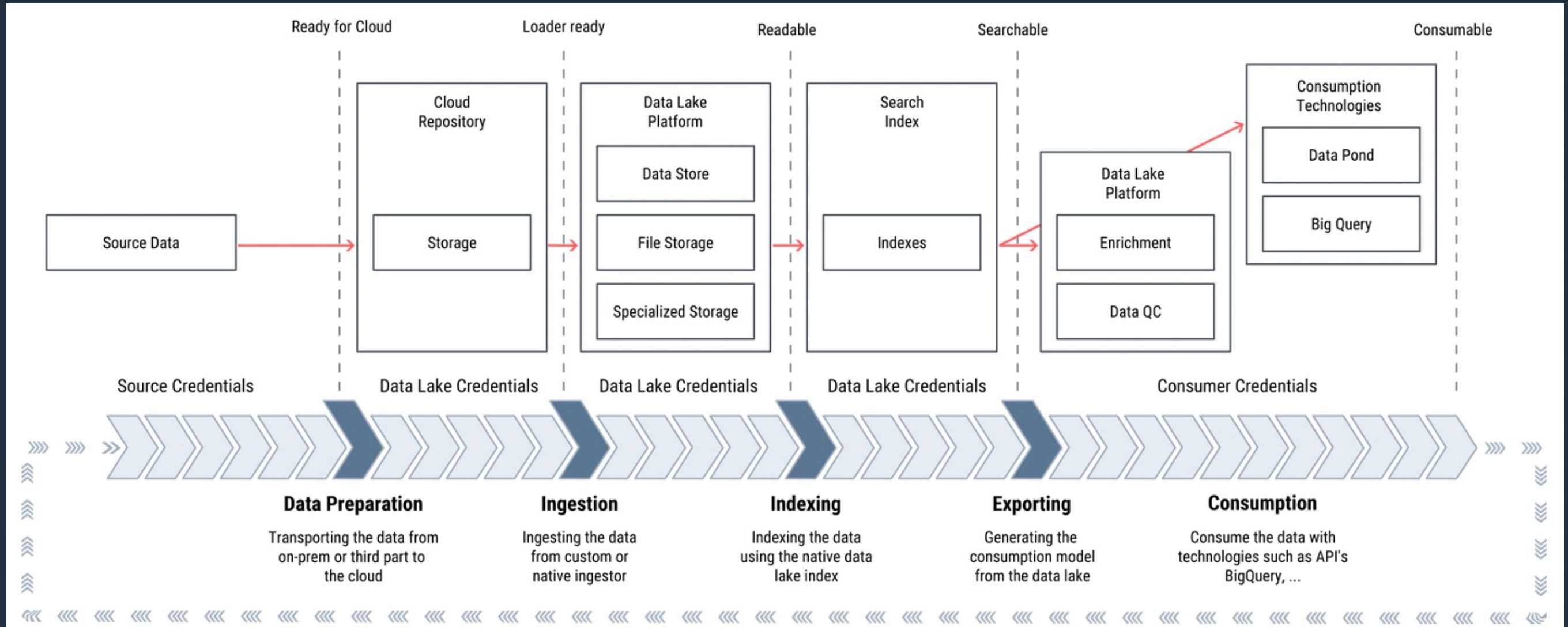
The Open Group OSDU Forum delivers an **Open Source, standards-based, technology-agnostic data platform** for the **energy industry** that stimulates innovation, industrializes data management, and reduces time to market for new solutions.

Its important to note that this solution is built to break down silos and enable cloud services. The Whole value chain, NOT Subsurface.

“Polycloud” – runs on Public Clouds and on premise, taking advantage of the technology available on each

Not a data storage solution or a data management application - but a modern, cloud-native data platform

OSDU implements and end to end data flow



OSDU translates data to well known structures

The screenshot shows the Open Group GitLab interface. The top navigation bar includes the Open Group logo, a search bar for GitLab, and several utility icons. The left sidebar contains a navigation menu with the following items: Schema Definitions (selected), Project information, Repository, Files (highlighted), Commits, Branches, Tags, Contributors, Graph, Compare, Issues (33), Merge requests (0), CI/CD, Security & Compliance, Deployments, Infrastructure, Monitor, and Collapse sidebar. The main content area displays a 'Table of Contents' with the following items:

- [Table of Contents](#)
- [List of Documented Entities](#)
 - ["abstract" Schemas](#)
 - ["data-collection" Schemas](#)
 - ["dataset" Schemas](#)
 - ["manifest" Schemas](#)
 - ["master-data" Schemas](#)
 - ["reference-data" Schemas](#)
 - ["type" Schemas](#)
 - ["work-product-component" Schemas](#)
 - ["work-product" Schemas](#)
- [Other Reports](#)
 - [Supported File Formats](#)
 - [Indexer Hints Summary](#)
 - [Summary of Virtual Properties](#)
 - [Changes compared to Community Mirror Site](#)
- [Superseded Entity Schemas](#)

Below the Table of Contents is a section titled 'List of Documented Entities' with a single item:

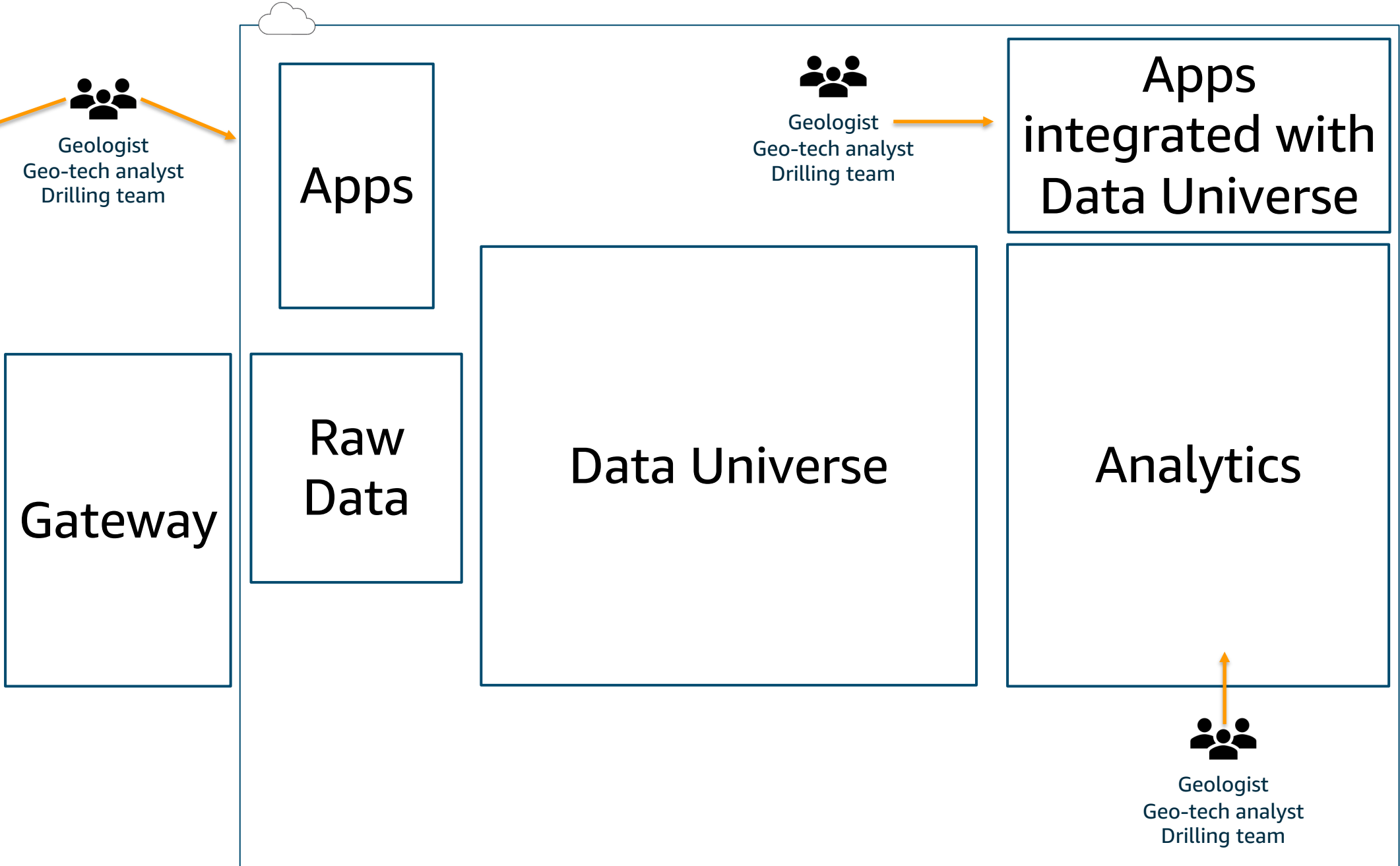
- [Link to entity type documentation — Schema description](#)

A [Back to TOC](#) link is located at the bottom of the main content area.

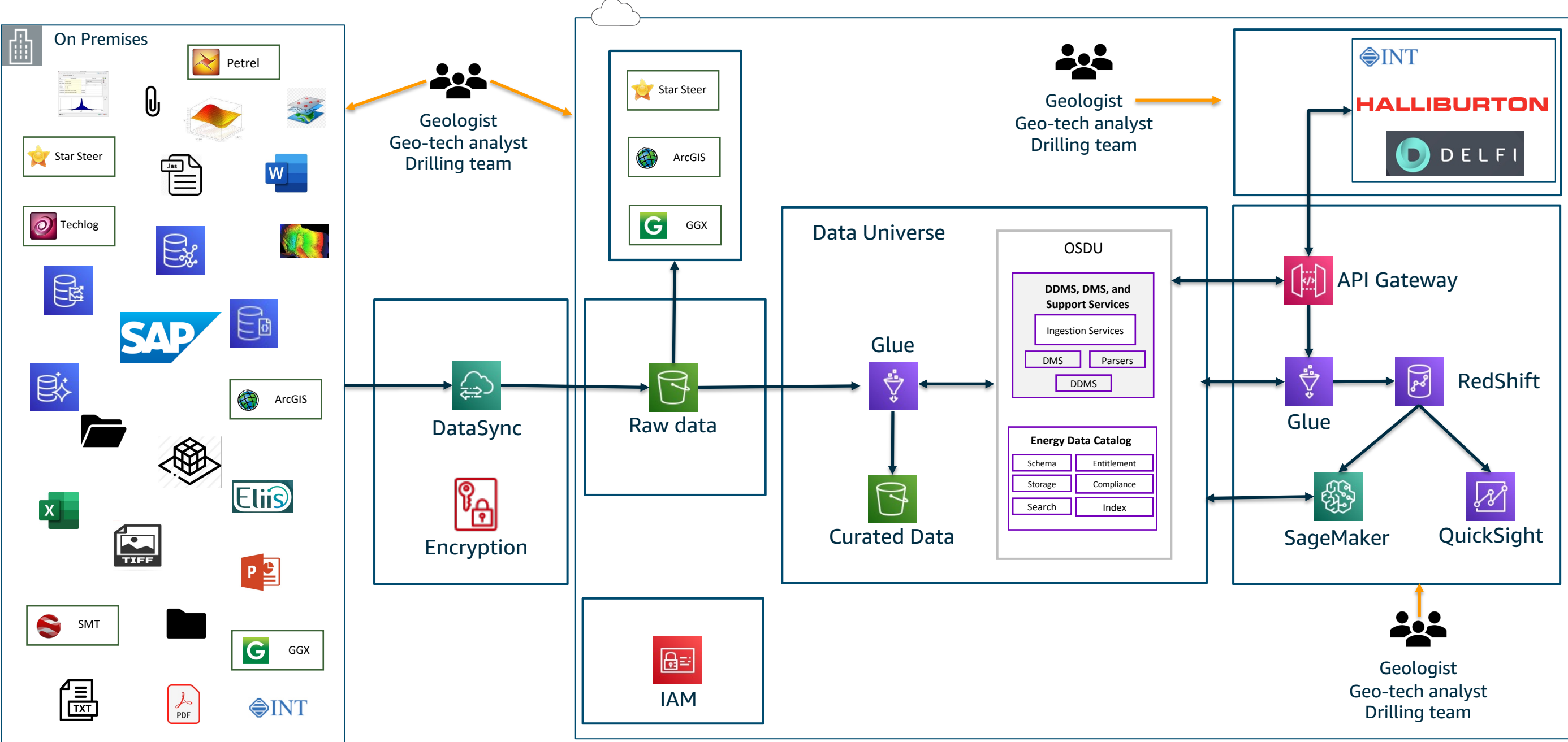
How do I get my data together?

On Premises

On Premises data sources and applications include: Petrel, Star Steer, Techlog, SAP, ArcGIS, Eliis, SMT, GGX, INT, and various file formats like .doc, .xls, .tif, .pdf, .txt.

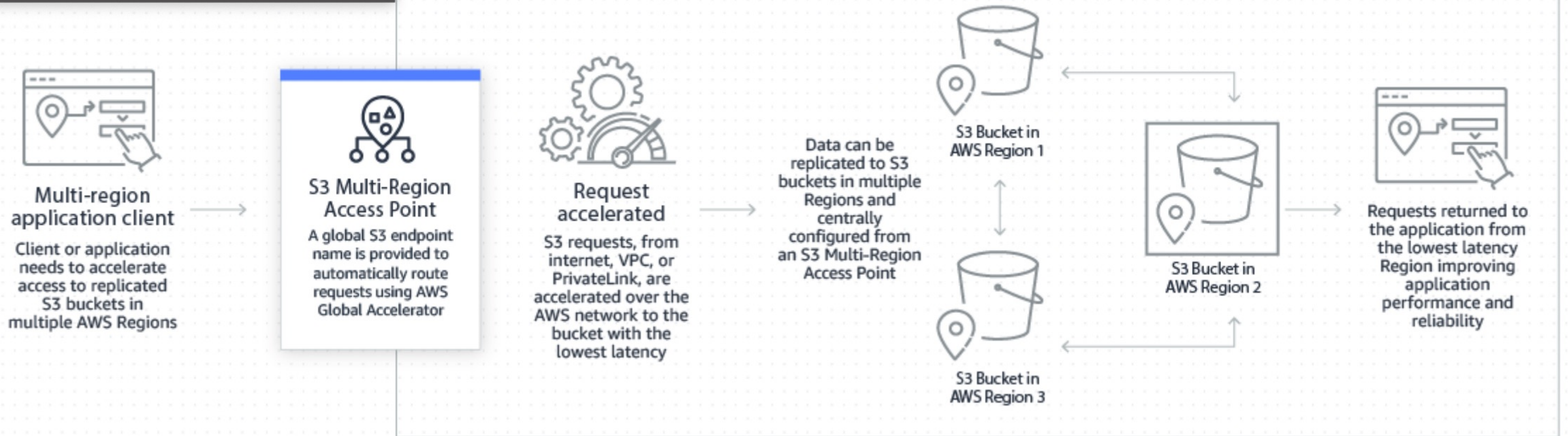


Example Architecture with the Data Universe



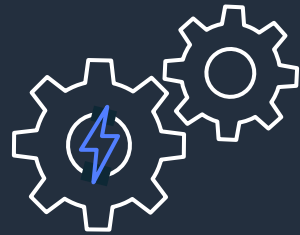
The power of AWS services: Amazon S3

How it works diagram - S3 Multi Region Access Points



Innovation enabled by AWS Nitro System

Modular building blocks for rapid design and delivery of EC2 instances



PERFORMANCE

Better performance across CPU, networking, and storage



SECURITY

Enhanced security that continuously monitors, protects, and verifies the instance hardware and firmware



INNOVATION

Building blocks can be assembled in many different ways, giving us the flexibility to design and rapidly deliver EC2 instances

The power of AWS services: Amazon Alexa





Thank you

Wlad Frazao

Energy Principal for Americas

osdu-on-aws@amazon.com

wfrazao@amazon.com

<https://www.linkedin.com/in/wladfrazao/>