SQL Day

14 edycja konferencji SQLDay

9-11 maja 2022, WROCŁAW + ONLINE

Data Community

partner złoty

Future Processing

Infinite DATA

KPMG

partner srebrny

Objectivity

summ-it

UBS

dbWatch DATABASE CONTROL

partner brązowy

elitmind think bright

Paweł Potasiński, Microsoft

The Evolution of Data Professional

How NOT to Become a Dinosaur in the Cloud Age

Disclaimer

The content of this session is based solely on the **subjective thoughts** of its author.

Feel free to **disagree** with my theses.
If you want to talk, let's talk **offline**!

I may not change my mind,
but it's all about **networking** ☺

Data Community

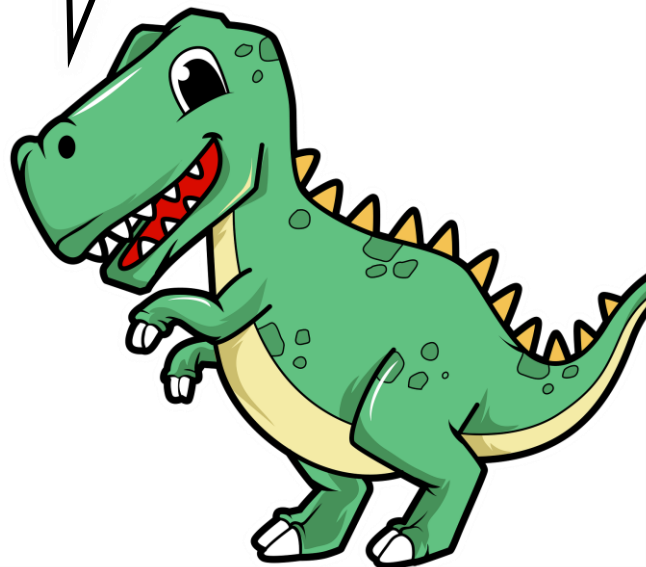
Paweł Potasiński, Microsoft

The Evolution of Data Professional

How NOT to Become a Dinosaur in the Cloud Age

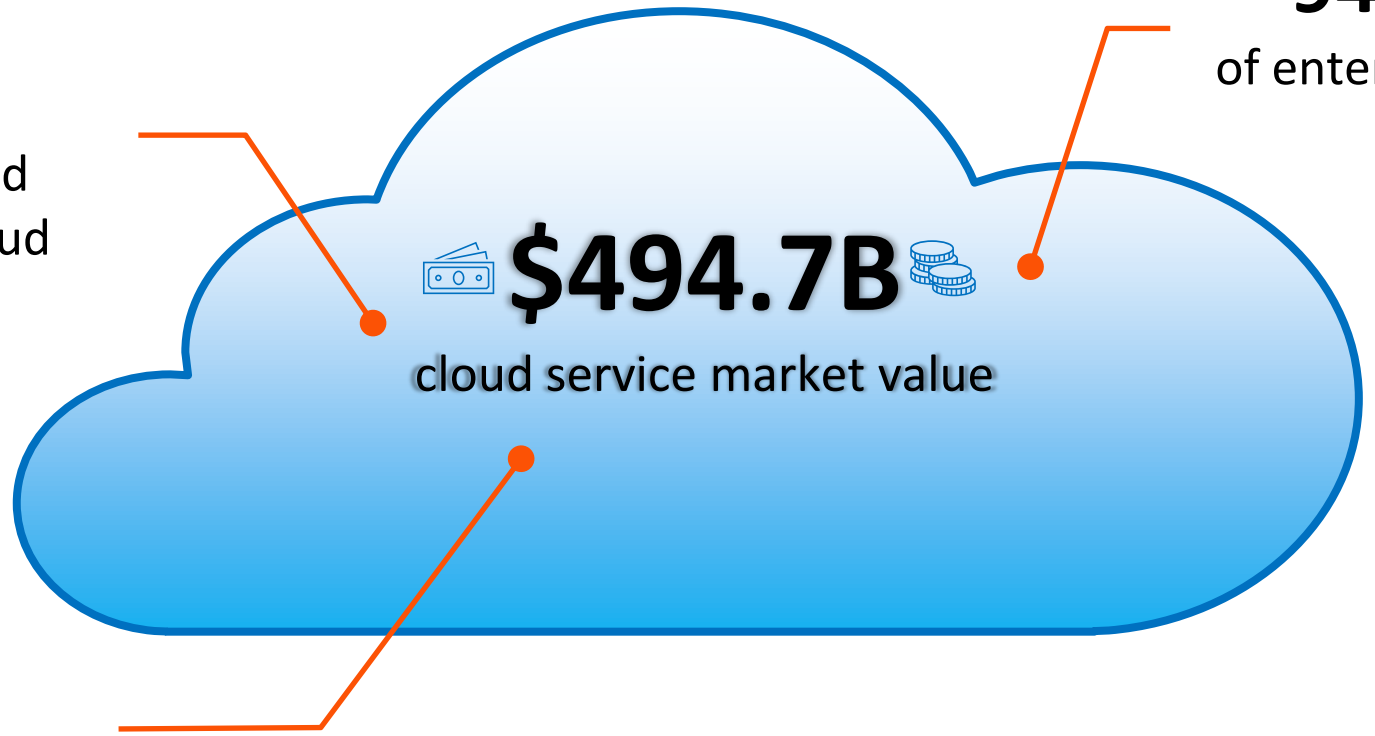
SQL Day

Any dinosaurs in the room?



The Cloud Age

$1.3T
of IT budgets affected by the shift to the cloud



94%
of enterprises

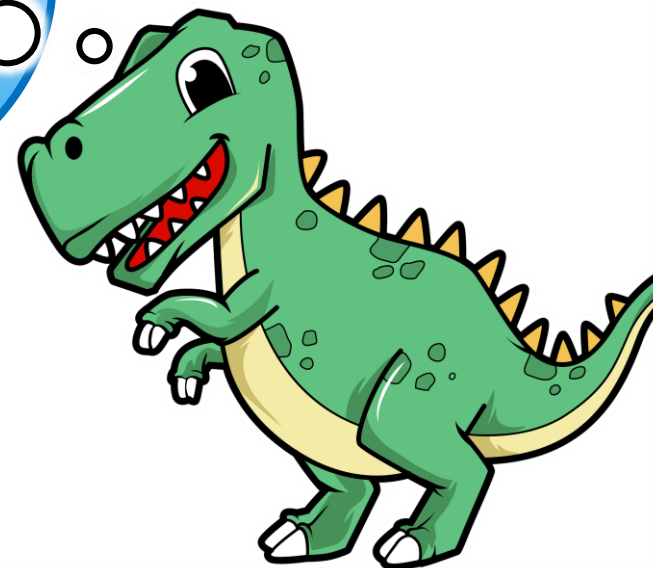
90%
of enterprises depend on hybrid

The Cloud Age

$1.3T
of IT budgets affected by the shift to the cloud

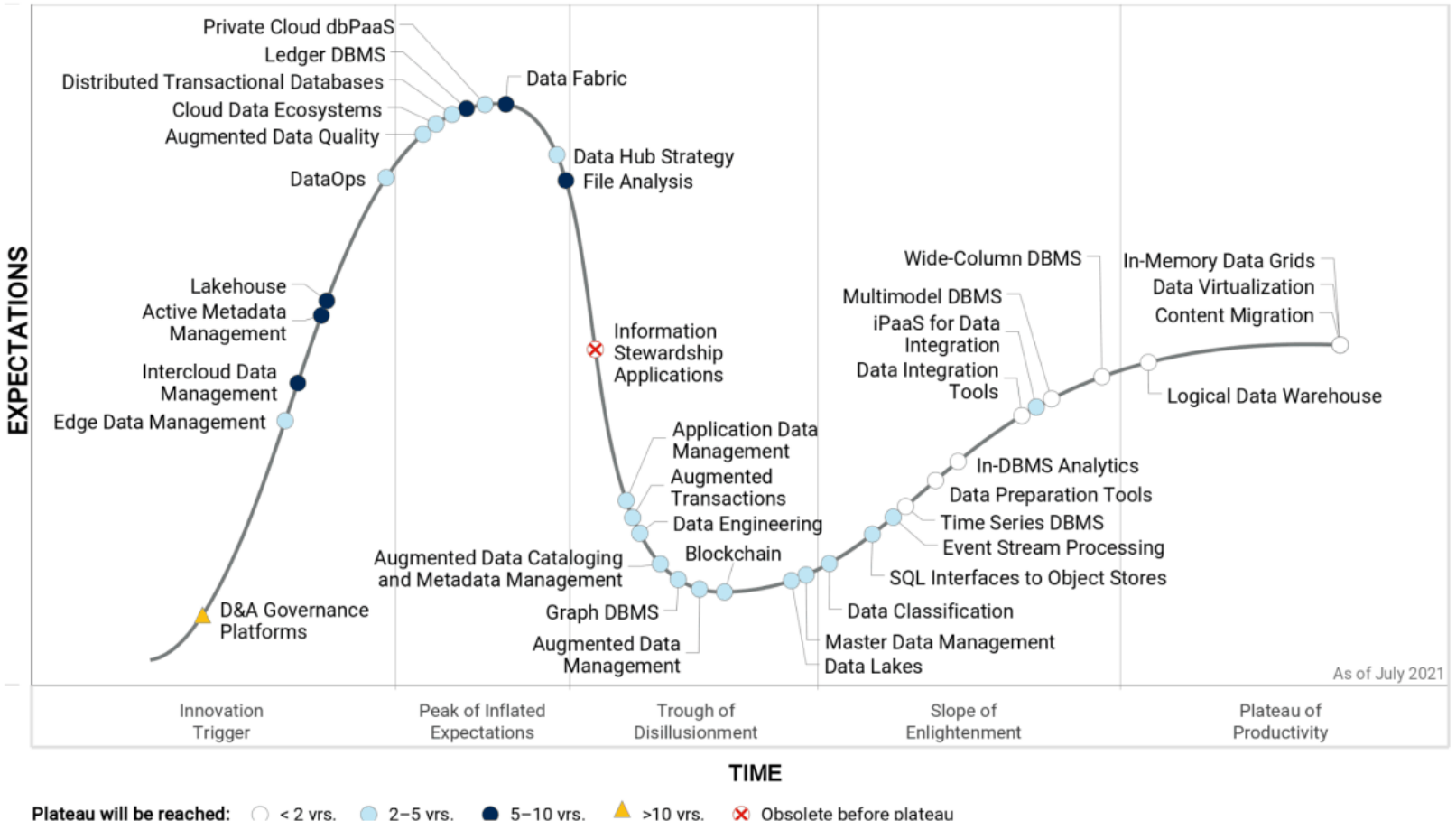
94%
of enterprises

90%
of enterprises depend on hybrid



Let's Talk Hype!

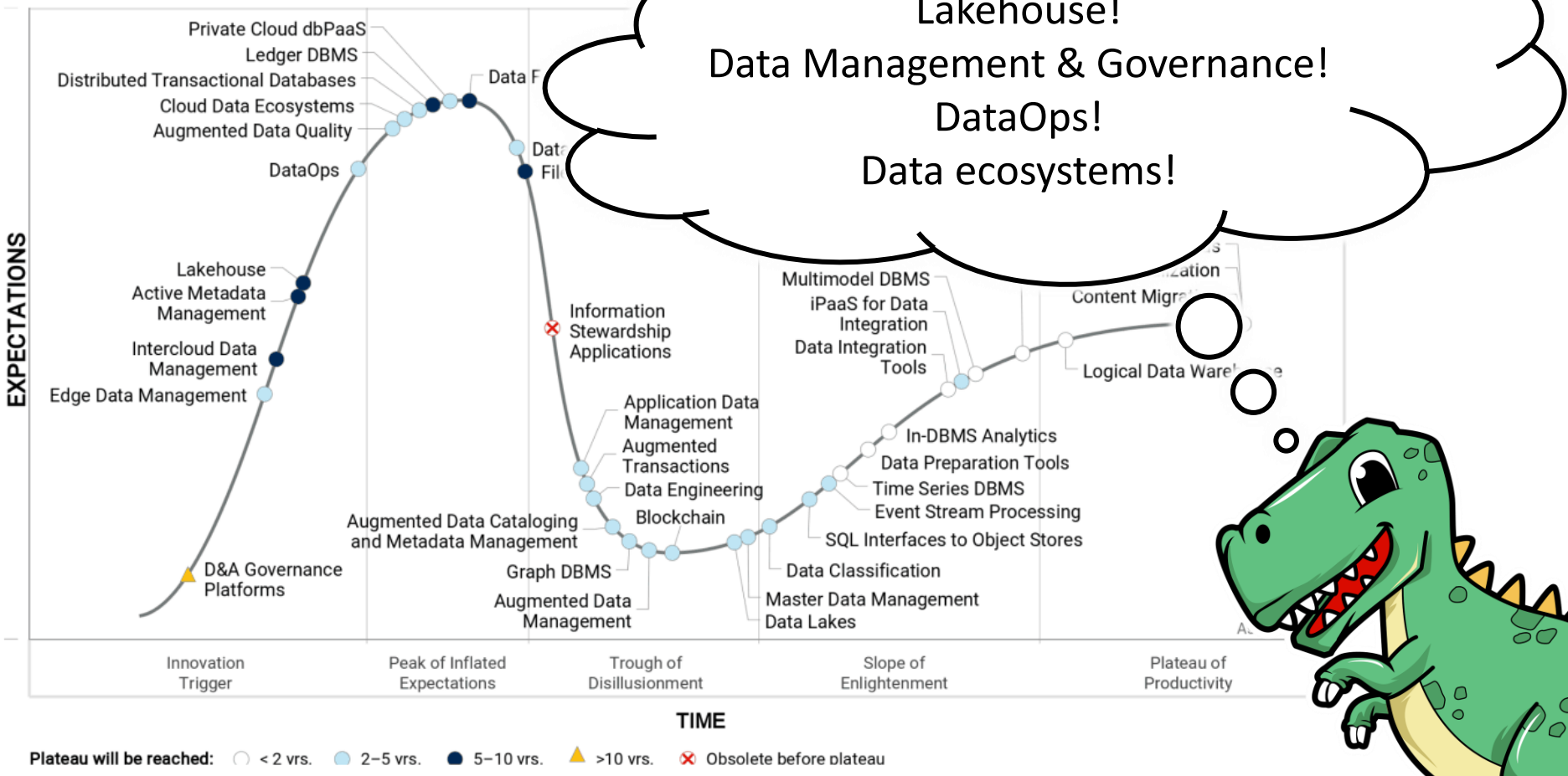
Source: Gartner Hype Cycle for Data Management, 2021



Gartner

Let's Talk Hype!

Source: Gartner Hype Cycle f



Lakehouse!
Data Management & Governance!
DataOps!
Data ecosystems!

Gartner

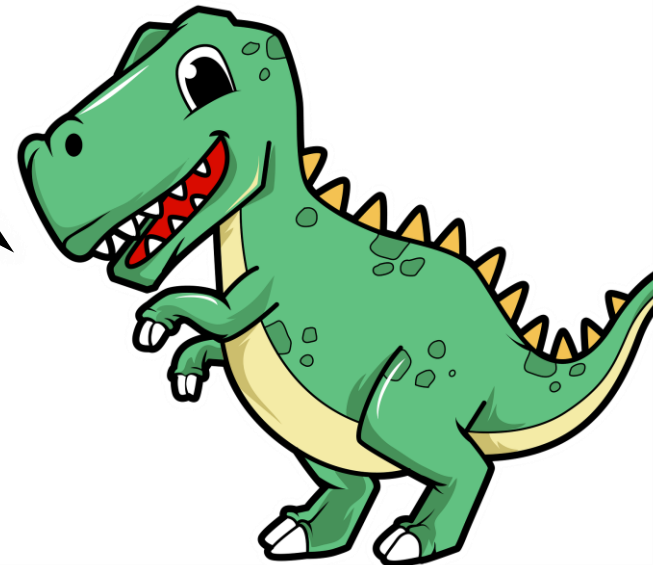
The World Speeds Up

- Quick **Go-to-Market** expected (adapt to „**new normal**”)
- Faster **from data to insights**
- **Measured frequent trials** with **business value** hard to predict
- More **data sources**, both internal and **external**
- **Data first**, schema last
- **Democratization** of data & AI
- Some **budgets** have moved from IT to the business
- Data usage is **regulated** like never before
- **Security** threats are more real than ever

The World Speeds Up

- Quick **Go-to-Market** expe
- Faster **from data t**
- **Measured frequ**
- More **data sour**
- **Data first**, schema
- **Democratization** of data &
- Some **budgets** have moved from IT to the business
- Data usage is **regulated** like never before
- **Security** threats are more real than ever

Change is the only constant...,
but today it's happening
faster and more often.
Learn to adapt!



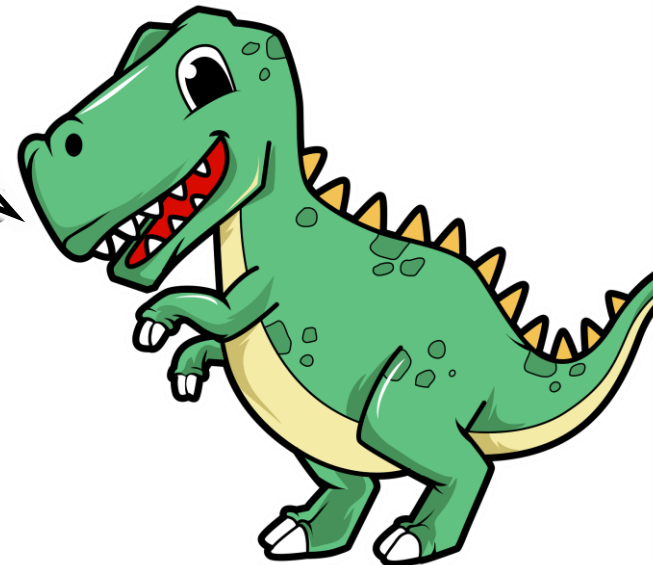
The World of Data Adjusts

- Data is **bigger** and more **distributed** than ever
- Fast pace of **technology evolution** (monthly updates)
- **Near real-time** insights required
- More restrictive **SLAs for data platforms**
- Massive **mix of technologies**, tools and services
- Market leaders consolidating into **uniform platforms**
- **Cost optimization** as an important part of data projects

The World of Data Adjusts

- Data is **bigger**
- Fast
- **Nea**
- More
- Massive **mix of**
- Market leaders consolidating into **uniform**
- **Cost optimization** as an important part of data projects

Track the market for technology **trends**.
Run a lot of **quick and dirty** tests.
Prefer **breadth over depth**.
Cloud is cool, but **costs money**.



Example: Modern Data Warehouse Requirements

- **Requirement #1: Scalability and performance out-of-the-box**
  - **Serverless** (PaaS/SaaS)
  - **Transparent (and independent) scalability** of storage and compute
  - Dealing with **democratization of data** and **hybrid workspace**

- **Requirement #2: Handling different workloads**
  - Analytical queries and data loads **simultaneously**
  - **High-concurrency** supported for analytical queries
  - **Real-time** analytics and data streaming (telemetry, IoT)
  - **Non-relational** data formats (binaries)
  - **Machine Learning**

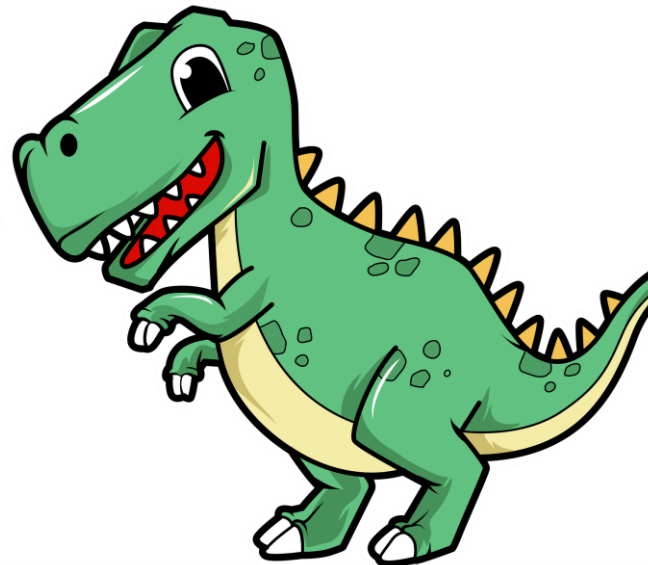
Example: Modern Data Warehouse Requirements

- **Requirement #**
  - **Server**
  - **Tran**
  - Dea
- **Requirement**
  - Analytical queries and data
  - **High-concurrency** supported for analytical queries
  - **Real-time** analytics and data streaming (telemetry, IoT)
  - **Non-relational** data formats (binaries)
  - **Machine Learning**

of-the-box

te

Data MUST drive **business value**.
It's NOT about knobs and tweaks in some hard-to-use systems.



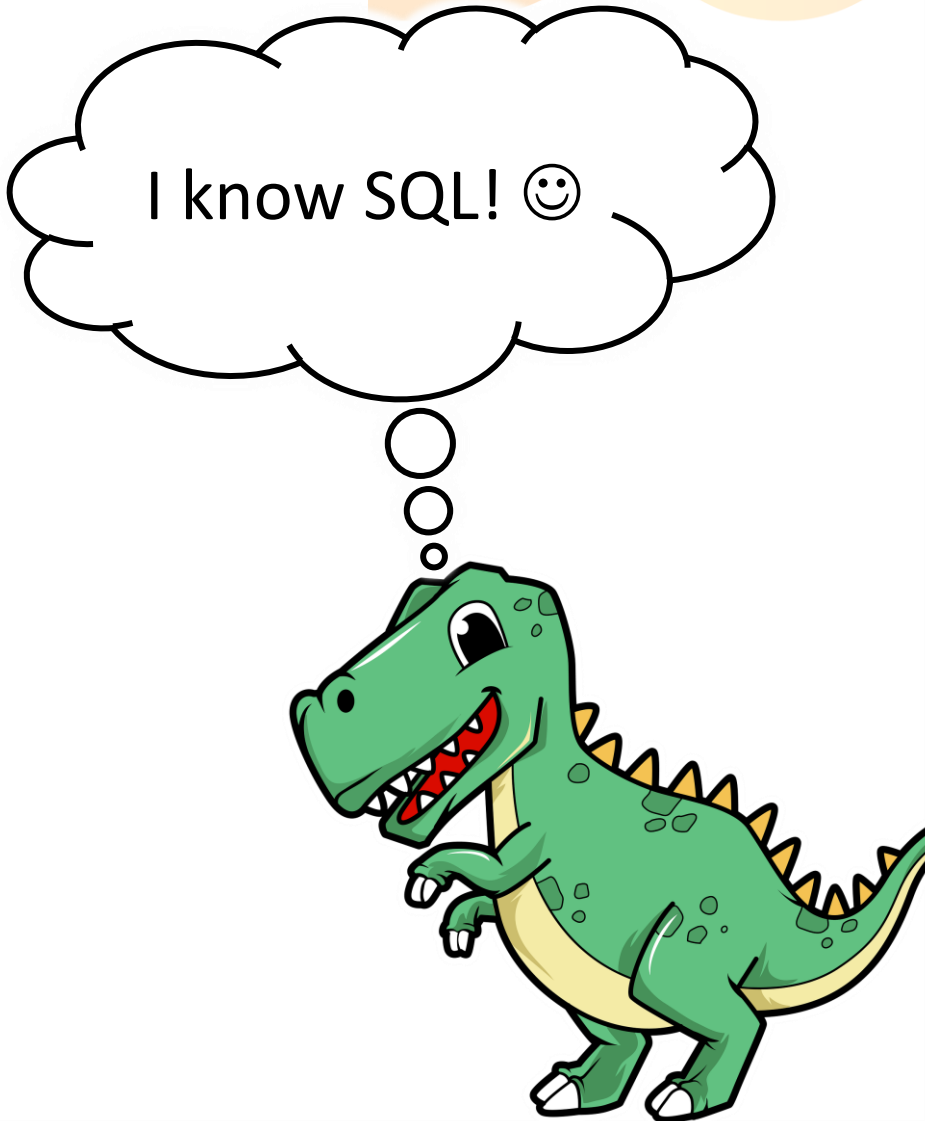
What Did NOT Change?

- You should understand what the **business needs**
- You should know and **understand the data** you manage

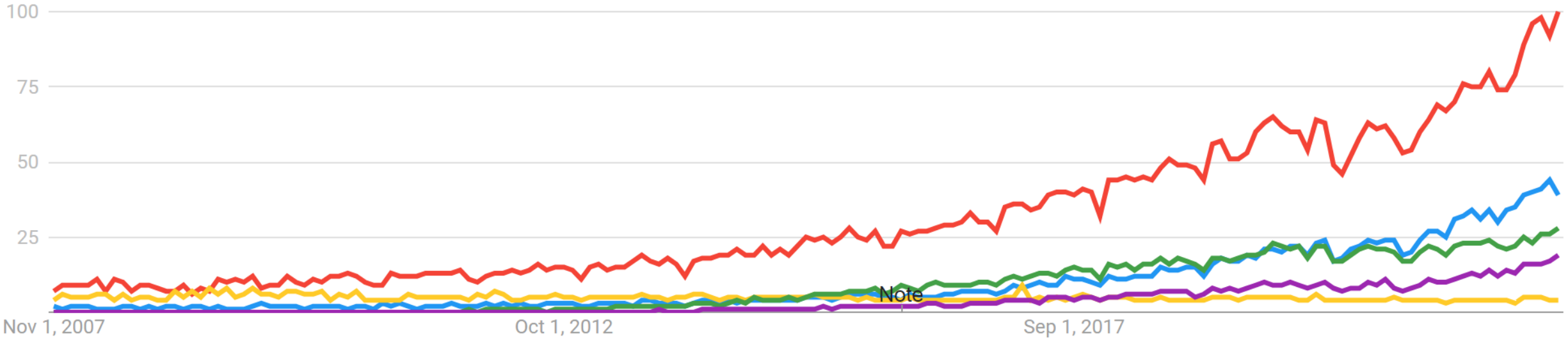
- **Code versioning / source control** is a great thing
- Separate **environments** (DEV/TEST/.../PROD) are a must

- Keep as **few copies** of data as you can
- Track **data lineage** if possible
- Bad things start on the **data(base) schema** level
- **SQL** is still a required language for data professional!

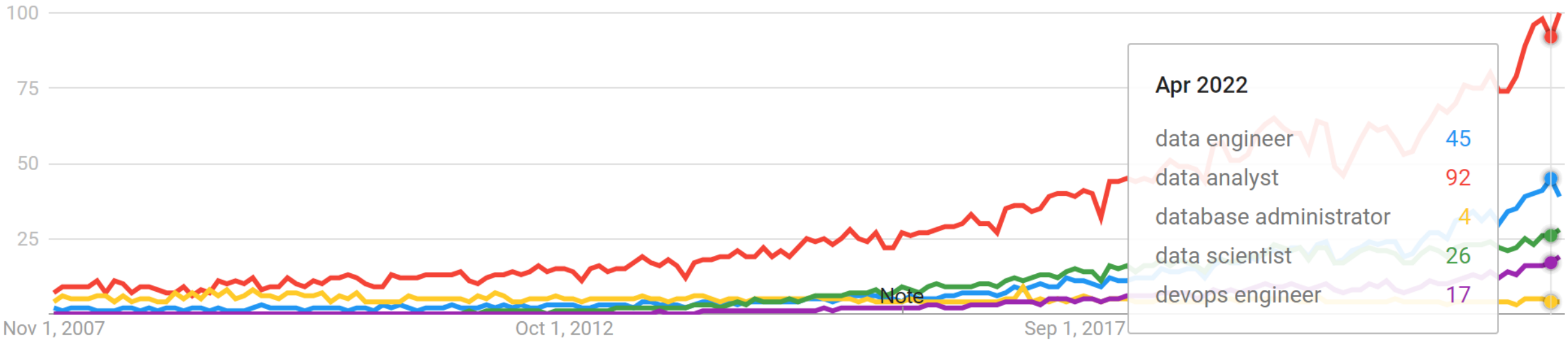
- **You ALWAYS must expect the worst case scenario**



???

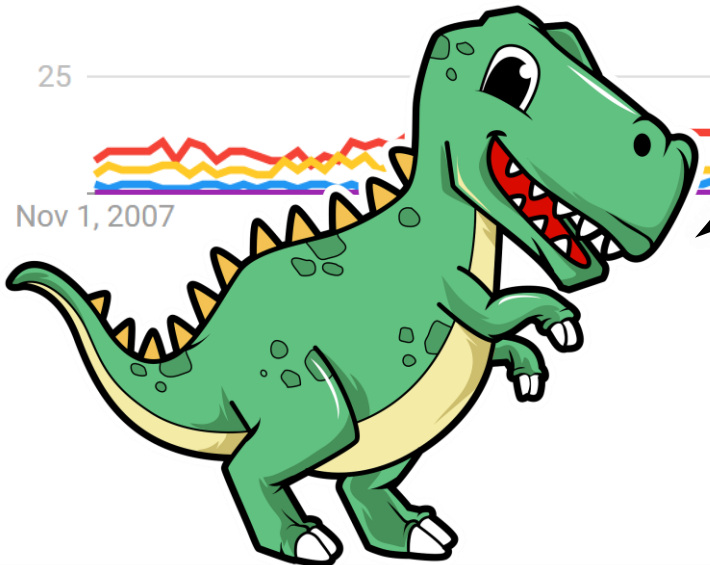

100
75
50
25
Nov 1, 2007
Oct 1, 2012
Sep 1, 2017
Note

Interest in Jobs Has Changed Been Changing



Interest in Jobs Has Changed Been Changing

1. Track the market. Look for **gaps**.
2. Consider **analytics** or **data engineering**.
3. Data scientist may be NOT the sexiest job anymore. If you choose it, learn **feature stores**.



The Evolution of Data Professional Roles

- Database Administrator
  - Automatic backups in PaaS
  - Auto-tuning in PaaS
  - Best practices by default
  - Different toolset (PowerShell, notebooks)
  - Data security & classification
  - HADR, scalability, and performance
  - Hybrid and multi-cloud solutions
- BI Developer
  - Forget single version of truth
  - Modern DWH/lakehouse
  - Pervasive AI in BI
- Data Modeler
  - Big Data (data lake, lakehouse)
  - Data virtualization

PRODUCTS AND SERVICES	*NON-REGIONAL	EAST US	EAST US 2	CENTRAL US	NORTH CENTRAL US	SOUTH CENTRAL US	WEST CENTRAL US	WEST US	WEST US 2	CANADA EAST	CANADA CENTRAL	BRAZIL SOUTH
DATABASES												
Azure Cosmos DB		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Azure SQL Database		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Azure Database for MySQL		✓	✓	✓	✓		✓		✓	✓	✓	✓
Azure Database for MariaDB		✓	✓	✓	✓		✓	✓	✓	✓	✓	✓
Azure Database for PostgreSQL		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Azure Cache for Redis		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
SQL Server Stretch Database		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Azure API for FHIR		✓	✓	✓								
ANALYTICS												
Azure Synapse Analytics		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
HDInsight		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Data Factory		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Data Factory V1		✓					✓	✓				
Azure Integration Runtime		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

https://status.azure.com/en-us/status/

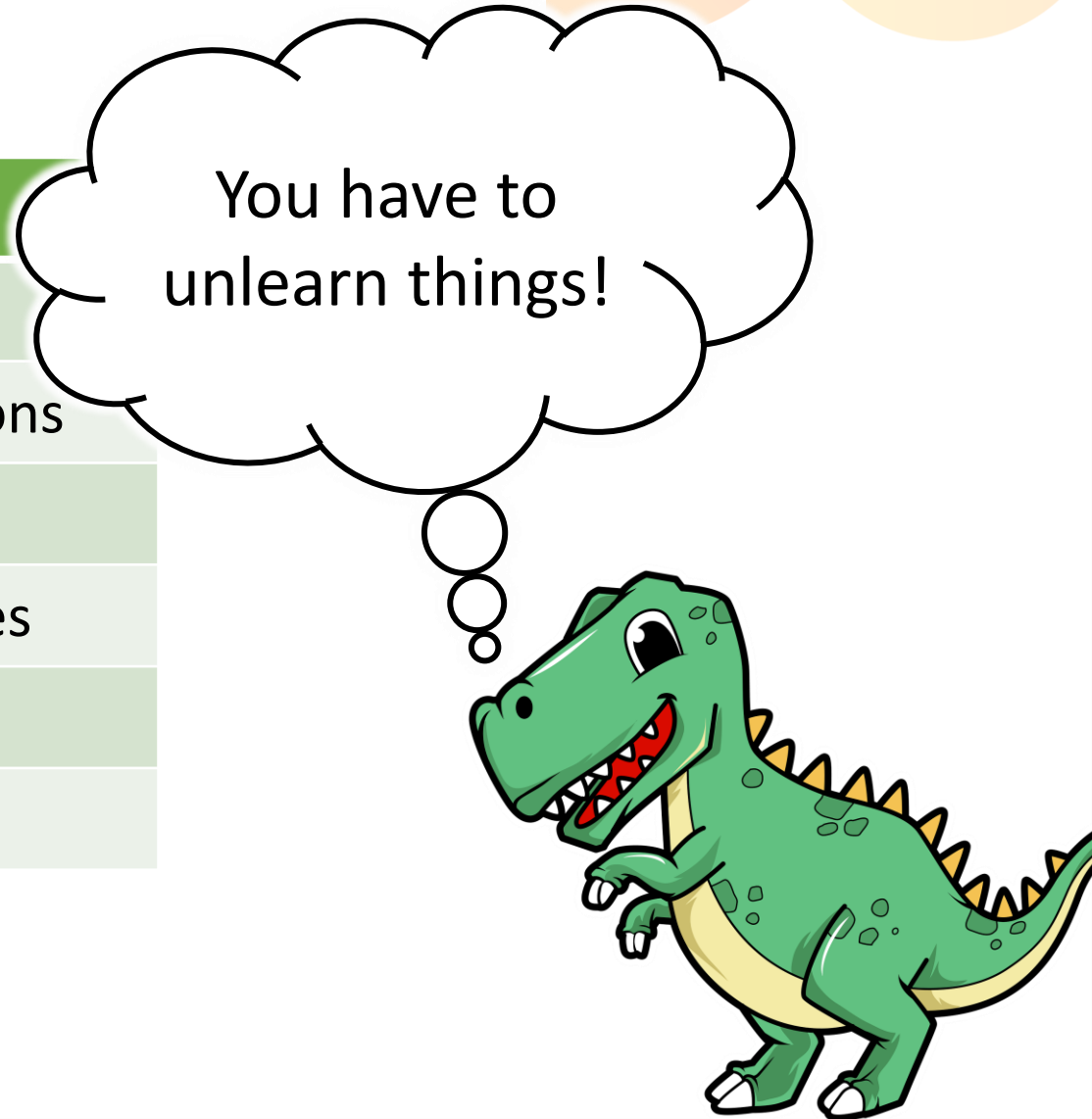
Things CAN Go Wrong in the Cloud

„Cloud is a state of permanent failure.”

by Unknown Cloud Professional

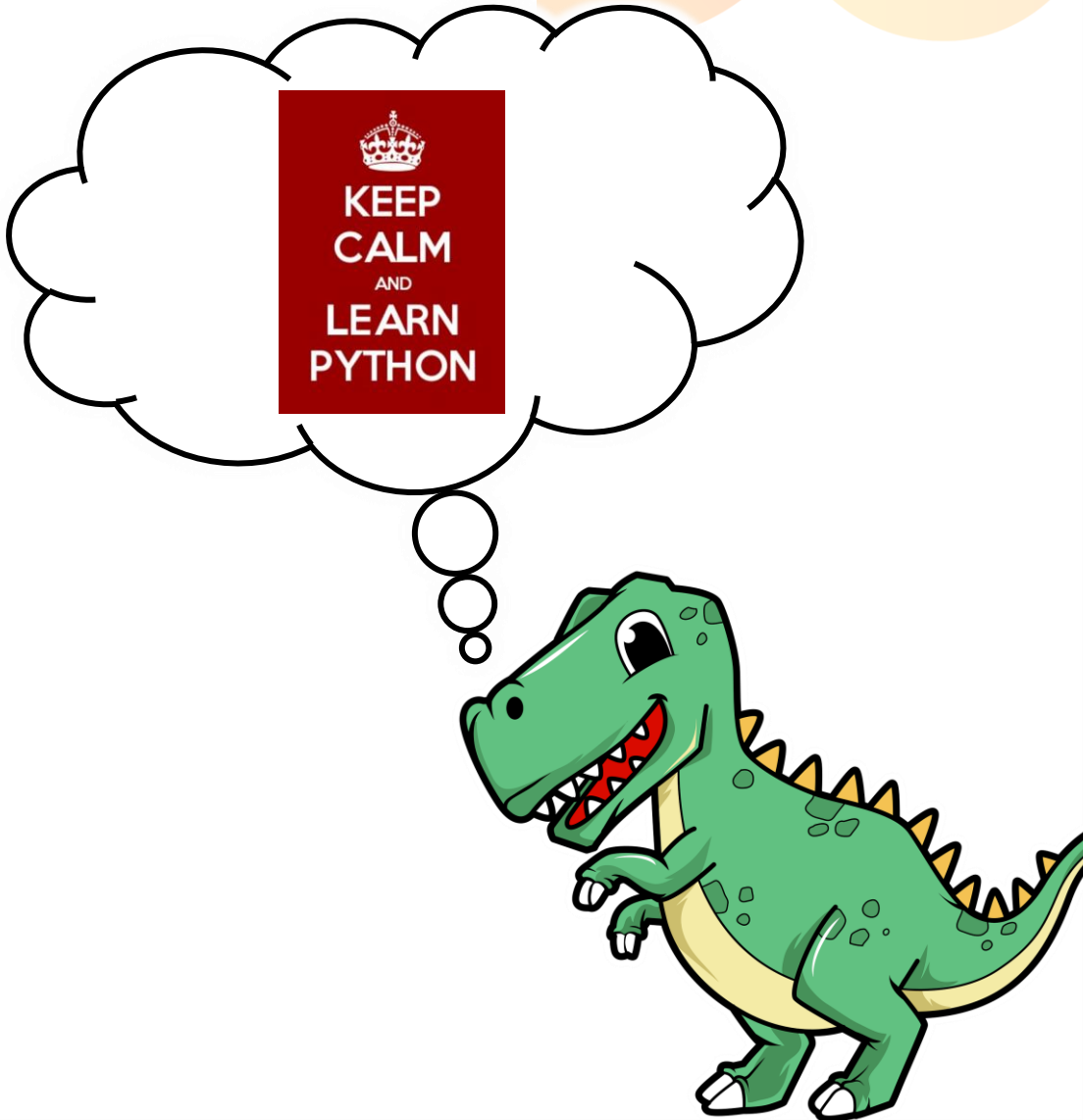
Your Mindset Needs to Change (DataOps)

From	To
Change Fear	Change Velocity
Manual Operations	Automated Operations
Hope for Quality	Integrated Quality
Hero Mentality	Repeatable Processes
Tool Centric	Code Centric
Vendor Lock-in	Diverse Tools



How We Should Act – Technologies to Learn

- SQL, SQL, SQL… – T-SQL, Spark SQL, HQL
- Python – scripting, PySpark, ML
- Open Source – Spark, Delta, Kafka, Airflow
- PowerShell – automation & mgmt, DevOps
- Azure CLI, ARM, Bicep – IaC
- Javascript – front-end, Power BI dev
- KQL – telemetry, monitoring
- DAX – depends on the context… ☺



How We Should Act – It's NOT Only About Technology

- Communication
- Knowledge sharing
- Networking & Community
- Remote teamwork
- Assertiveness and planning
- Embrace learning by trials and failures

How We Should Act - 10 Rules for Cloud Adept

1. Learn systematically. Deep dive knowledge is required occassionally.
2. Know the limitations of services and their integration.
3. The way to implement things in the cloud may differ from on-prem.
4. Automate things in the cloud even more than in on-prem.
5. Practice DevOps, CI/CD, IaC, and PaC.
6. Understand pricing and learn how to optimize billing.
7. Do NOT assume the architecture will be static/fixed.
8. Test and fail often.
9. Be curious of different vendors.
10. Always assume things can go wrong.

Patience!

„Understand this is a long-term process.”

by Jes Borland, Microsoft

Data Community

Paweł Potasiński, Microsoft

The Evolution of Data Professional

How NOT to Become a Dinosaur in the Cloud Age

SQL Day

SQL Day

14 edycja konferencji SQLDay

9-11 maja 2022, WROCŁAW + ONLINE

Data Community

partner złoty

Future Processing

Infinite DATA

KPMG

partner srebrny

Objectivity

summ-it

UBS

dbWatch DATABASE CONTROL

partner brązowy

elitmind think bright