

Medium Search

Josef Machytka: Speaker Portfolio — PostgreSQL, DuckDB, Databases in General



Josef Machytka

4 min read · Nov 3, 2024



Listen



Share



More

Expert Talks on PostgreSQL, DuckDB, Databases in general, Data LakeHouse, Data Ingestion and Data Analysis

Introduction

With over 30 years of experience in database technologies, I have cultivated a deep understanding of diverse database systems, including extensive work with PostgreSQL. I worked with multiple heterogeneous data ingestion and data processing pipelines, focusing on delivery of valuable business results. The need to solve often very complicated technical problems equipped me with the expertise necessary for insightful and impactful talks. I am passionate about sharing my knowledge, empowering others in the field of databases and data processing and making complex topics accessible.

My talks are living and evolving, I steadily improve them with new content reflecting progress in new versions of the corresponding software. My speaking style combines practical insights with important technical details, ensuring that audiences walk away with really valuable knowledge they can apply. Slides contain a lot of additional information, making them valuable resources even outside the context of my presentations.

Thank you for considering me as a speaker for your next event. I look forward to the opportunity to share my knowledge and help your audience gain deeper insights into PostgreSQL and database topics.

Short Bio

Name: Josef Machytka

My current job: Professional Service Consultant, PostgreSQL specialist at [credativ GmbH](#)

Education: Engineer's degree, Technical University Ostrava, Czech rep., topic Automation and robotics

Experience

- 30+ years of production experience with different databases:
PostgreSQL 13 years, BigQuery 7y, Oracle 15y, MySQL 12y,
Elasticsearch 5y, MS SQL 5y, Sybase ASE, FoxPro
- 10+ years of experience with high volume and velocity data ingestion pipelines, Data Analysis, Data Warehouse and Data Lakehouse architecture
- 2+ years of practical experience with different LLMs, their architecture and principles

Accounts

- [LinkedIn](#),
- [ResearchGate.net](#),
- [Academia.edu](#),
- [Sessionize.com](#)

Mastodon: [@JosefMachytka@me.dm](#)

Portfolio

(Link leads to detailed article about the talk.)

- [**GIN, BTREE_GIN, GIST, BTREE_GIST, HASH and BTREE indexes on JSONB data**](#)
- [**PostgreSQL and DuckDB: Supercharging ad-hoc Data Analysis and ETL**](#)

- *Building a Data Lakehouse with PostgreSQL: Dive into Formats, Tools, Techniques, and Strategies*
- *PostgreSQL Connections Memory Usage: How Much, Why, and When?*
- *PostgreSQL Partitioning: An Overview of solutions with a deep dive into implementation details and use cases*

My latest speaking engagements

1. 2025.01.29 Prague PostgreSQL Developer Day 2025

“Building a Data Lakehouse with PostgreSQL: Dive into Formats, Tools, Techniques, and Strategies”

2. 2025.01.27 NetApp internal talk

“PostgreSQL Connections Memory Usage: How Much, Why, and When?”

3. 2025.01.23 NetApp internal talk

“Building a Data Lakehouse with PostgreSQL: Dive into Formats, Tools, Techniques, and Strategies”

4. 2025.01.08 PostgreSQL MeetUp for All online event

“PostgreSQL and DuckDB: Supercharging ad-hoc Data Analysis and ETL”

5. 2024.11.26 NetApp internal talk

“PostgreSQL and DuckDB: Supercharging ad-hoc Data Analysis and ETL”

6. 2024.11.15 NetApp Tech Talk

“Many facets of AI hallucinations: factual errors, deep fakes, and creativity”

7. 2024.11.12 Internal Workshop — 3 talks:

“PostgreSQL 17: New Features”

“PostgreSQL Architecture: How PostgreSQL functions under the Hood”

“PostgreSQL Partitioning: Overview of Solutions, Tips & Tricks, Performance”

8. 2024.10.29 Prague PostgreSQL MeetUp October 2024

“PostgreSQL and DuckDB”

9. 2024.10.16 NetApp internal workshop — “AI Workshop: Exploring Artificial Intelligence” — 3 talks:

- “The AI Dilemmas: The Philosophical, Ethical, and Legal Implications of AI”
- “Understanding the Magic of LLMs: A Deep Dive into the Internal Structure of LLMs”
- “Many facets of AI hallucinations: factual errors, deep fakes, and creativity”

10. **2024.10.09 PostgreSQL MeetUp Berlin October 2024**

- “GIN, BTREE_GIN, GIST, BTREE_GIST, HASH and BTREE indexes on JSONB data”

11. **2024.09.27 NetApp internal talk**

- “Beyond the Buzzwords: A pragmatic explanation of LLM terminology”

12. **2024.10.13 NetApp internal talk**

- “The Art and Science of AI Prompt Engineering”

13. **2024.07.16 PostgreSQL MeetUp Berlin July 2024**

- Lightning talk: “Can PostgreSQL have a more prominent role in the AI boom?”

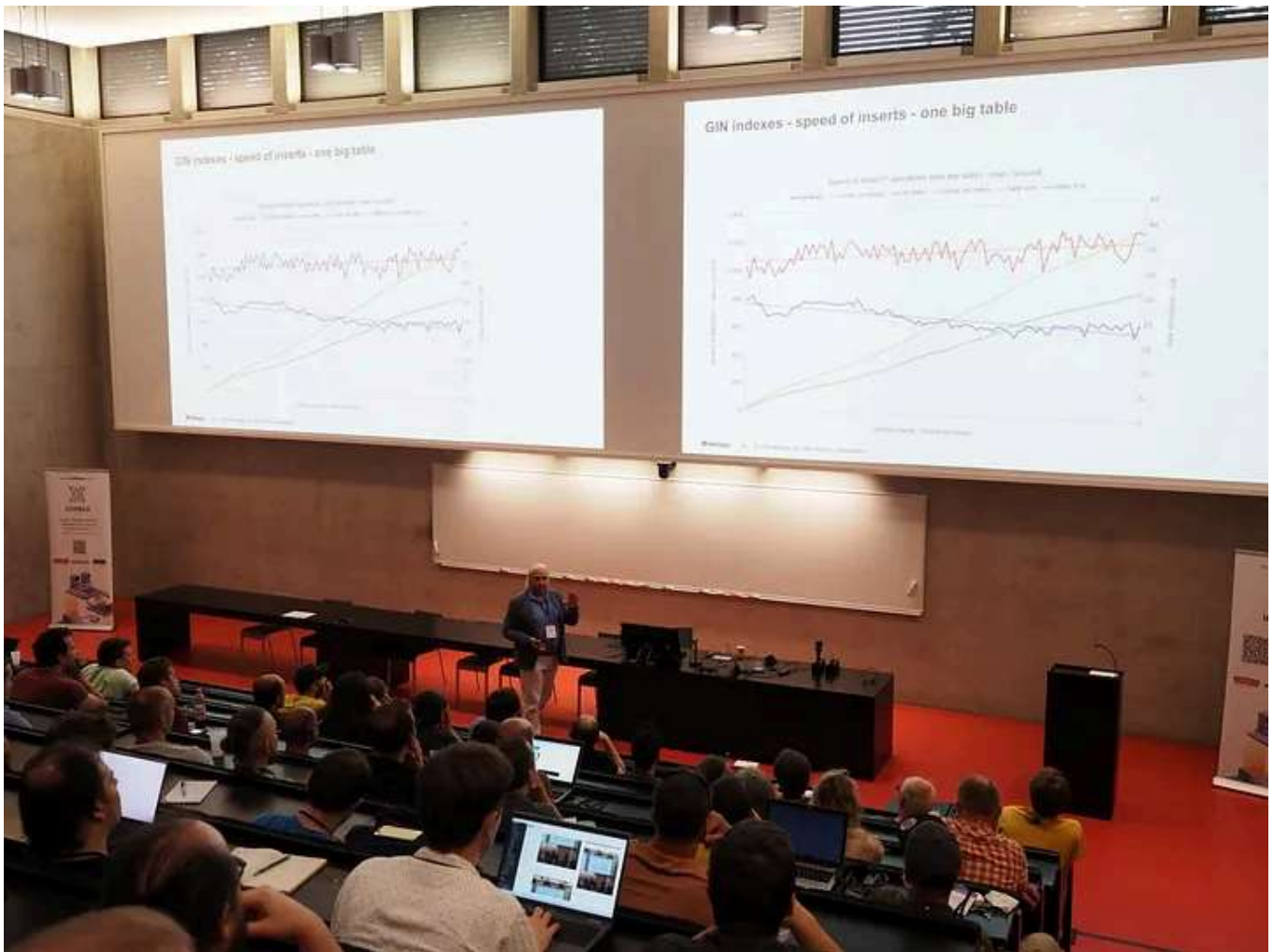
14. **2024.06.27 Swiss PG day 2024**

- “GIN, BTREE_GIN, GIST, BTREE_GIST, HASH and BTREE indexes on JSONB data”

15. **2024.06.05 Prague PostgreSQL Developer Day 2024**

- “GIN, BTREE_GIN, GIST, BTREE_GIST, HASH and BTREE indexes on JSONB data”

Photos from my talks



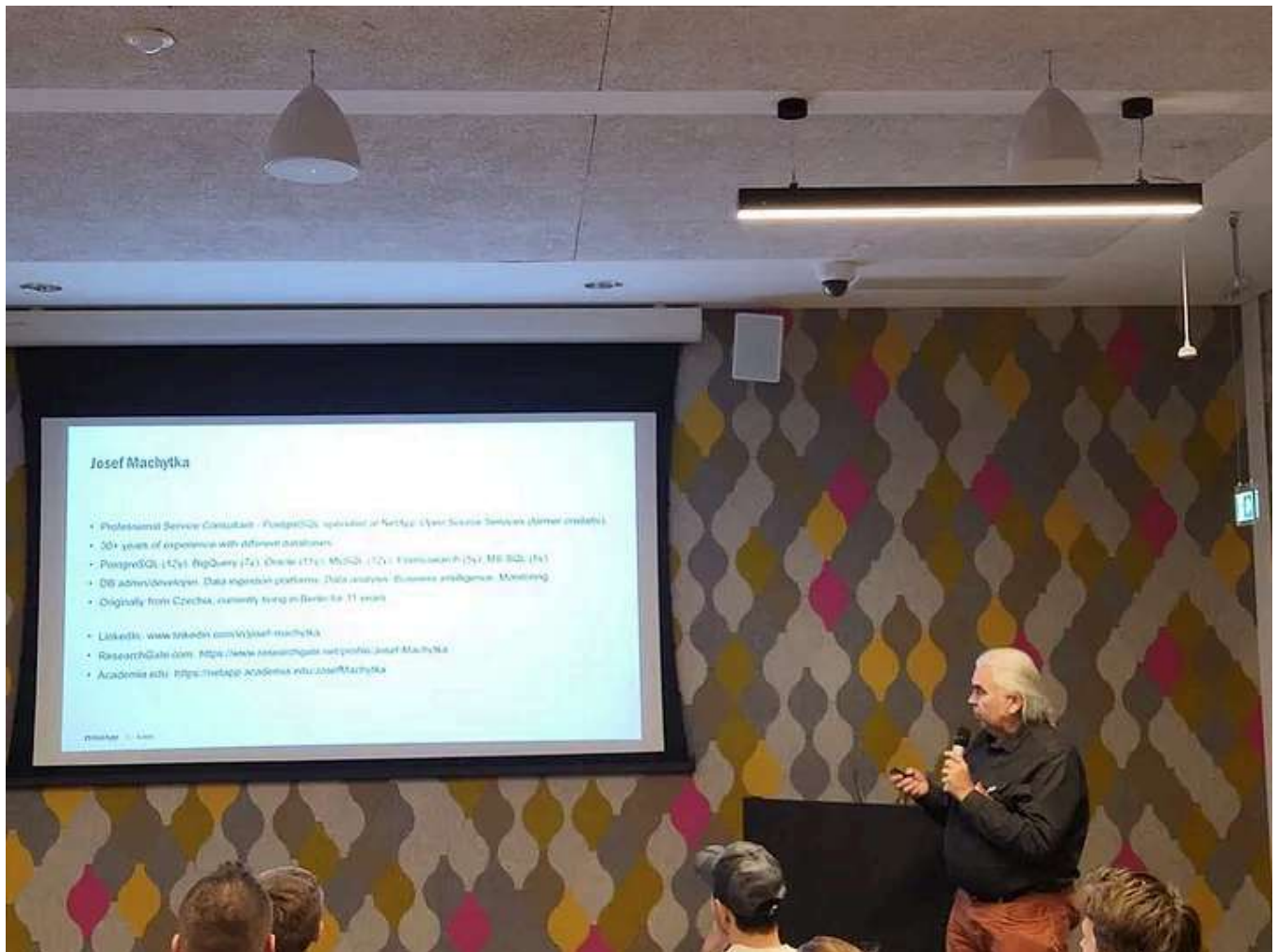
© [Tomas Vondra P2D2](#) — Prague PostgreSQL Developer Day 2024 — talk about indexes on JSONB data



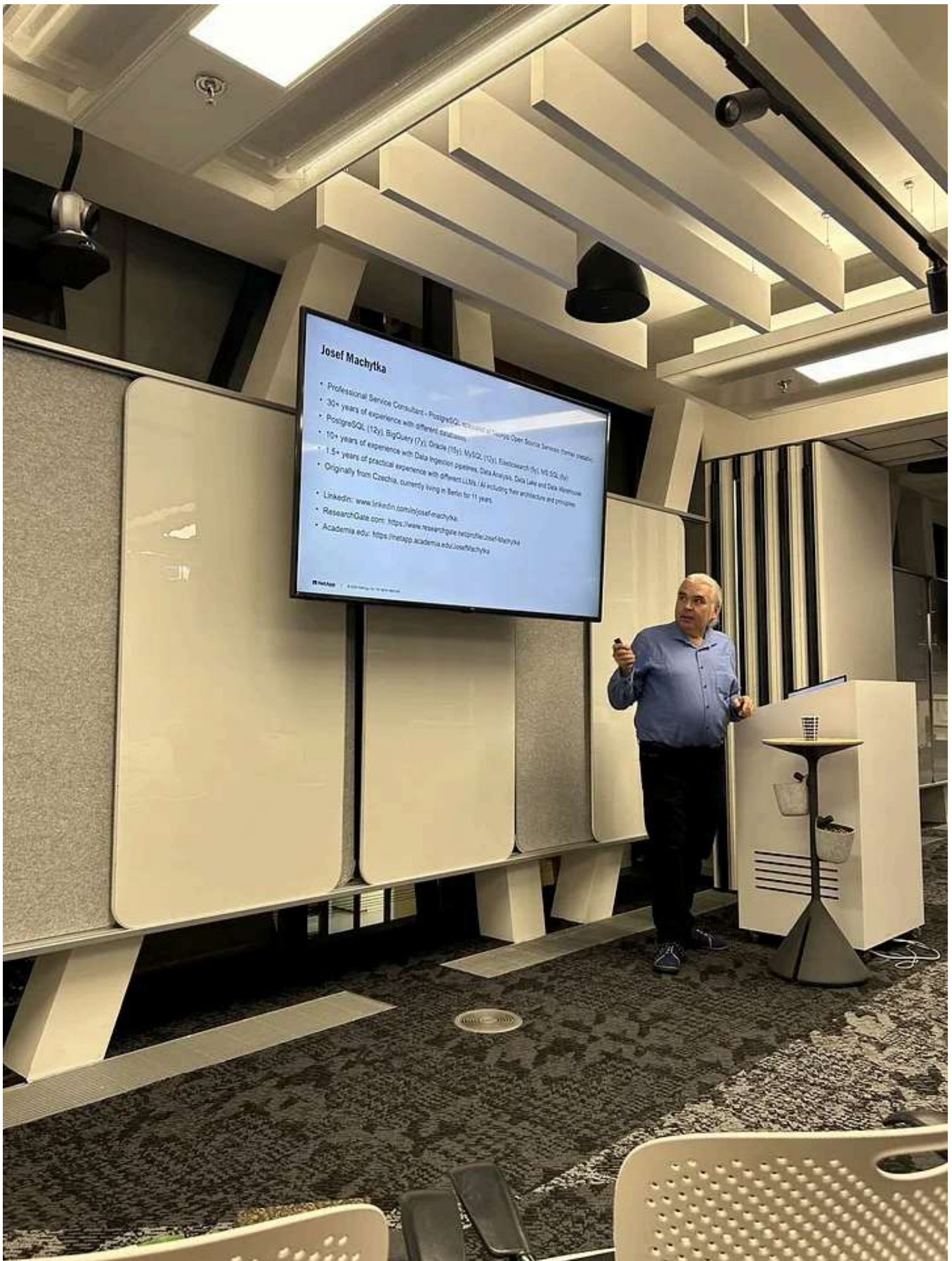
© organizers of Swiss PG day — Swiss PG day 2024 — talk about indexes on JSONB data



© [Andreas Scherbaum](#) — PostgreSQL Meetup Berlin July 2024 — lightning talk about PostgreSQL and AI



© Andreas Scherbaum — PostgreSQL Meetup Berlin October 2024 — talk about indexes on JSONB data



© Igor Gavrilov (LinkedIn post) — Prague PostgreSQL Meetup October 2024 — talk about DuckDB



© organizers of Prague PostgreSQL Developer Day 2025



© organizers of Prague PostgreSQL Developer Day 2025

Postgresql

Data Analysis

Data Ingestion

Duckdb



Edit profile

Written by Josef Machytka

68 Followers · 25 Following

I work as PostgreSQL specialist & database reliability engineer at credativ GmbH.

No responses yet



Josef Machytka he/him

What are your thoughts?

More from Josef Machytka

Dec 16, 2024 26



 Josef Machytka

PostgreSQL JSONB Operator Classes of GIN Indexes and Their Usage

Throughout 2024, I worked on an internal project exploring the use of JSONB data in PostgreSQL and its various indexing options. During...

Jan 8 5



Bob	2100.0	600.0	
Charlie	2300.0	1500.0	1100.0

```
D pivot pg.sales on (product,year) using sum(sales_amount) group by salesperson order by salesperson;
```

salesperson varchar	(Laptop, 2022) double	(Laptop, 2023) double	(Phone, 2022) double	(Phone, 2023) double	(Tablet, 2022) double	(Tablet, 2023) double
Alice	1200.0	1400.0	800.0	900.0	300.0	400.0
Bob	1000.0	1100.0	600.0			
Charlie	1100.0	1200.0	700.0	800.0	500.0	600.0

```
D pivot pg.sales on (year,product) using sum(sales_amount) group by salesperson order by salesperson;
```

salesperson varchar	(2022, Laptop) double	(2022, Phone) double	(2022, Tablet) double	(2023, Laptop) double	(2023, Phone) double	(2023, Tablet) double
Alice	1200.0	800.0	300.0	1400.0	900.0	400.0
Bob	1000.0	600.0		1100.0		
Charlie	1100.0	700.0	500.0	1200.0	800.0	600.0

```
D pivot pg.sales on (year) using sum(sales_amount) group by salesperson order by salesperson;
```

--	--	--

 Josef Machytka

Easy and Intelligent Pivot Tables with DuckDB

After exploring the various capabilities of DuckDB in my earlier articles, I want to focus more on its powerful data analytical...

Dec 4, 2024  7  1



See all from Josef Machytka

Recommended from Medium



In Towards Dev by Merve Gamze Cinar

Optimizing dbt Models with Common Table Expressions (CTEs)

Writing SQL is one thing, but designing a well-structured dbt model is a whole different game. One key aspect that often separates good...

Feb 19  19  1





 Ramesh Ponnusamy

Query Any Data, Anywhere-Simplify Workflows with DuckDB

Engineering is all about simplicity. We often use lightweight technologies to solve problems instead of overly complex ones, wherever we...

★ Jan 6 🖱 48 💬 1



Lists



Practical Guides to Machine Learning

10 stories · 2215 saves



ChatGPT prompts

51 stories · 2605 saves



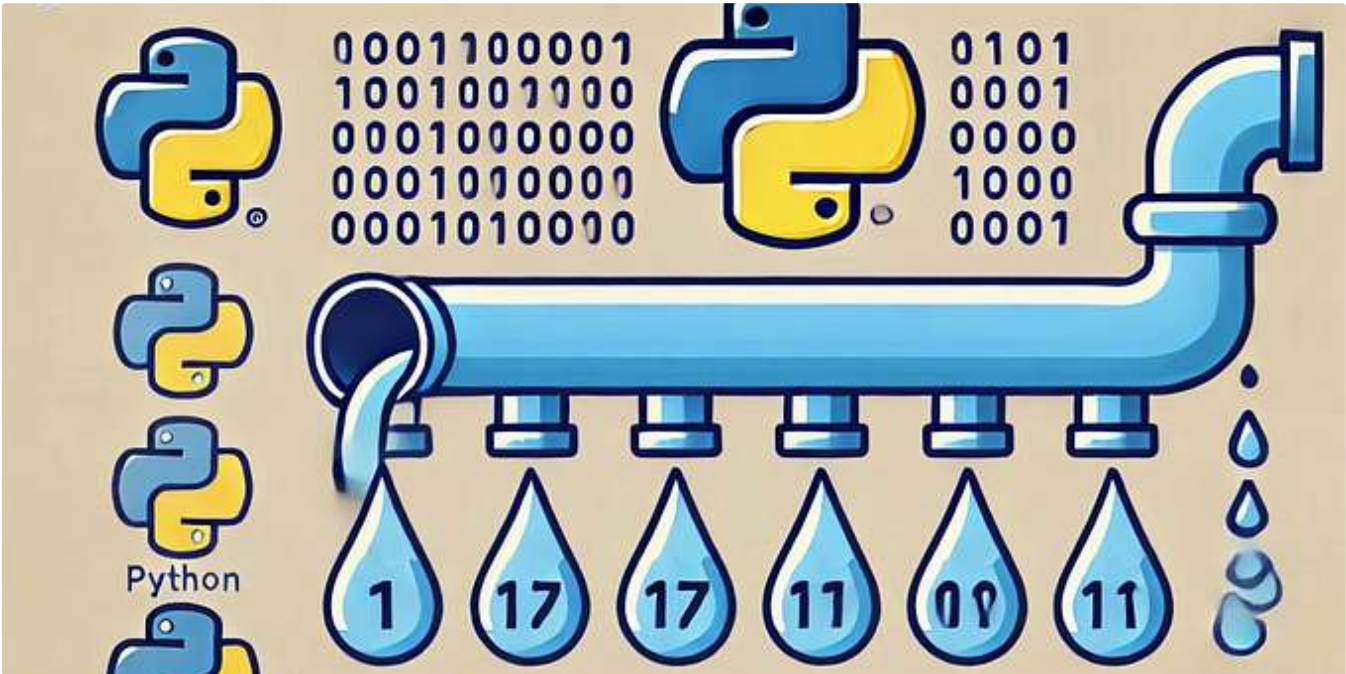
Staff picks


819 stories · 1637 saves



Natural Language Processing


1958 stories · 1605 saves



 In Dev Genius by Ani Palakurthi

10 Essential but Lesser-Known Python Features for ETL Pipelines

Python offers many underrated features and libraries that can supercharge your ETL process.

Feb 7  52  1  

Query Optimization Method	Execution Time	Performance Improvement (vs. No Index)
No Index	42,049 ms (≈42 sec)	Baseline
With B-tree Index	9,684 ms (≈9.7 sec)	77% faster
With Chunk-Skipping Index + Columnstore	304 ms (0.3 sec)	99.28% faster

 In Timescale by Team Timescale

Handling Billions of Rows in PostgreSQL

Here's how to scale PostgreSQL to handle billions of rows using Timescale compression and chunk-skipping indexes.

Jan 17 72 1

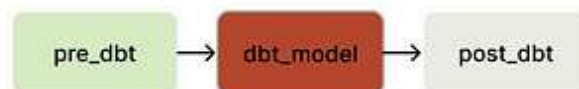


José Cabeda

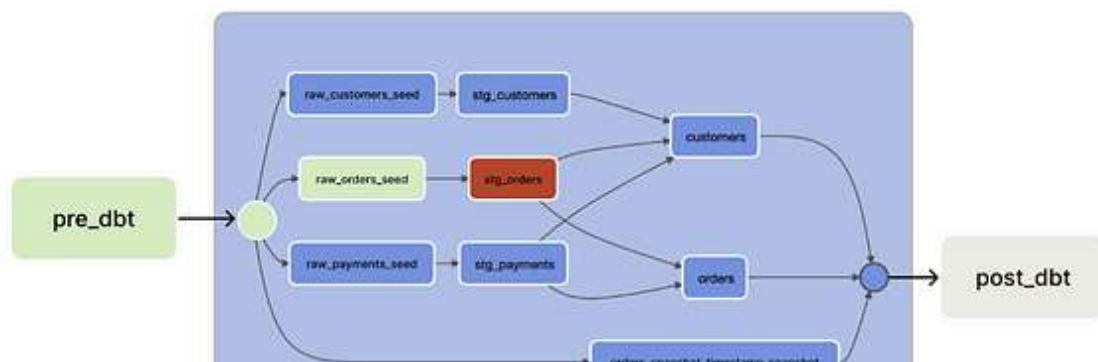
Running DuckDB on AWS Glue

With my increasing use of AWS Glue, I wanted to test some features on the platform, particularly running DuckDB on a constrained instance...

Oct 2, 2024 23 2



With Cosmos



In Apache Airflow by Arjun Anandkumar

Running dbt models on Airflow (MWAA)

Premise

Feb 21  5



See more recommendations