SQL
KONFERENZ
2025



Find your LEGO

an AI-powered image search with Fabric and Power BI



SQL
KONFERENZ
2025

Roger Unholz

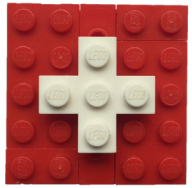




SQL
KONFERENZ
2025



Zürich







SQL KONFERENZ 2025

BSG DATA ANALYTICS

Partner & Solution Architect





Power BI– image search

2000 2024









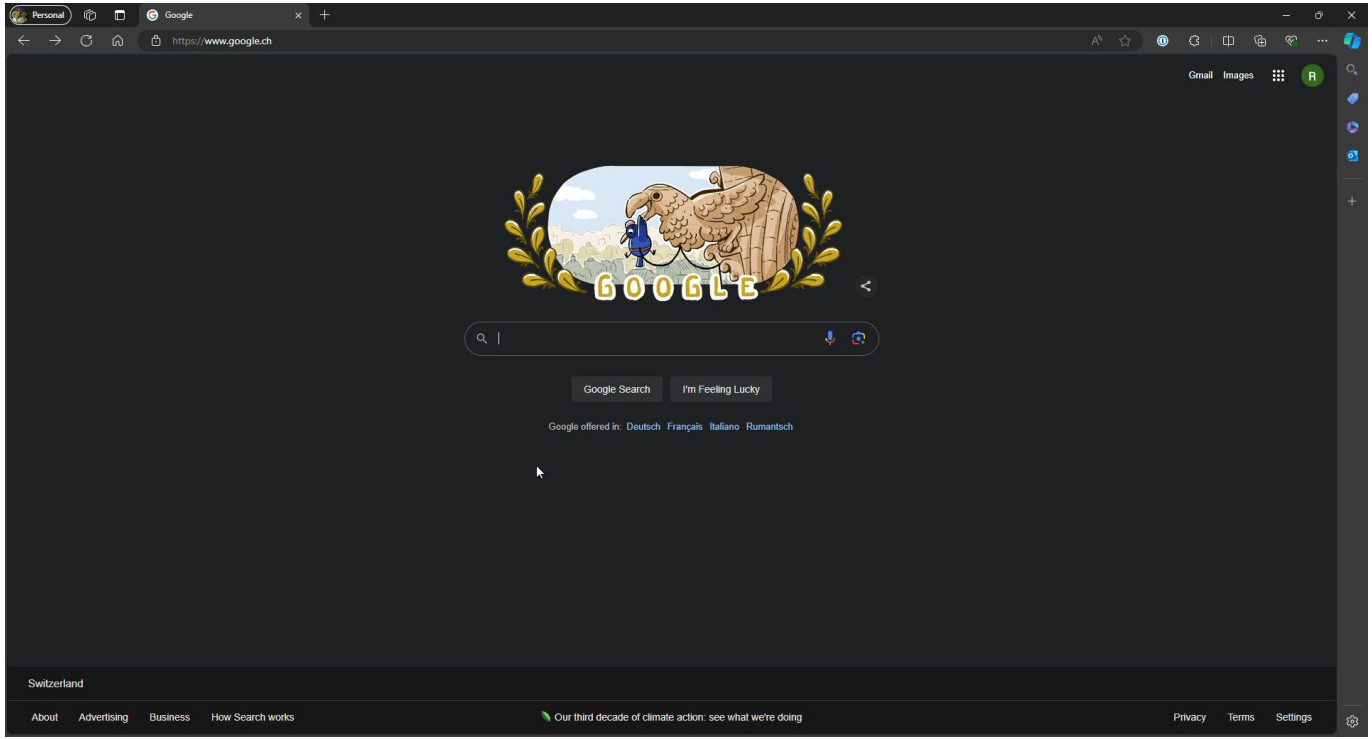

30051-1 - X-wing Fighter - Mini



Lego_Image_Search, Page 1

Live data Data updated on 25.02.25, 08:16

How does reverse image search work?





THE CONCEPT OF VECTOR EMBEDDINGS



The concept of vector embeddings

Imagine you have a massive library of books 📚. Each book can be described by several features:

- Number of pages
- Genre
- Author
- Popularity
- Complexity of language



embeddings automatically turn each book into a **numerical fingerprint** — a list of numbers (vector) that captures the book's "essence" based on its content and context



source



Azure AI Vision

AI

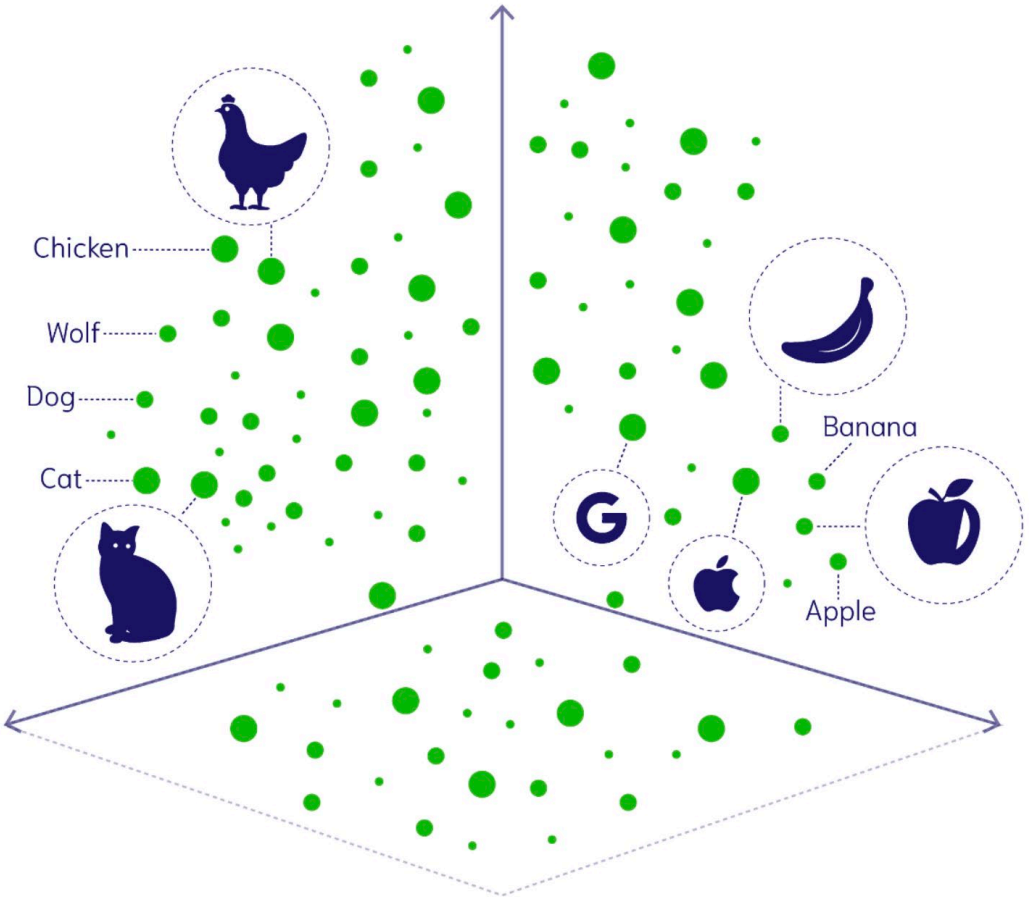
output

0.6	0.3	0.1	0.4	...	0.8

0.2	0.4	0.7	0.3	...	0.5

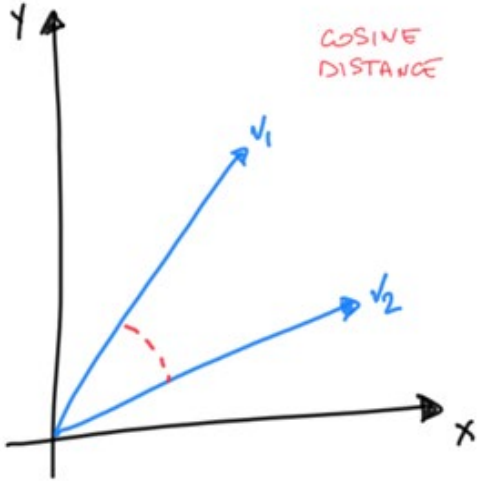


Multimodal embeddings in vector space





Cosine similarity or distance between vectors

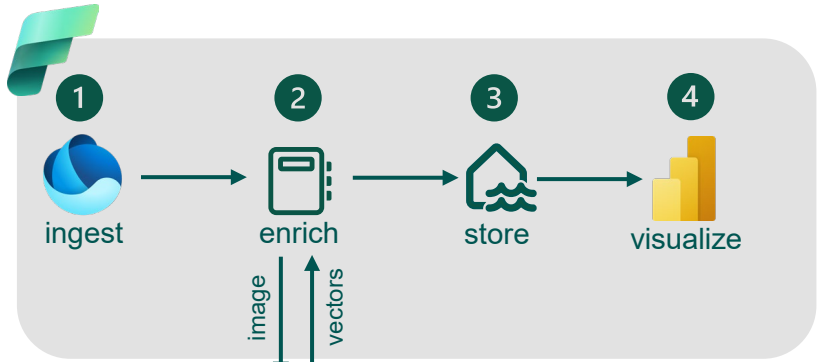


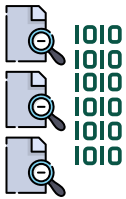
$$\text{cosine similarity} = S_C(A, B) := \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \cdot \sqrt{\sum_{i=1}^{n} B_i^2}},$$



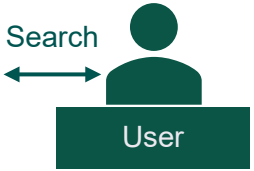
LEGO catalog DB







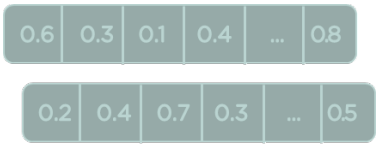
Search





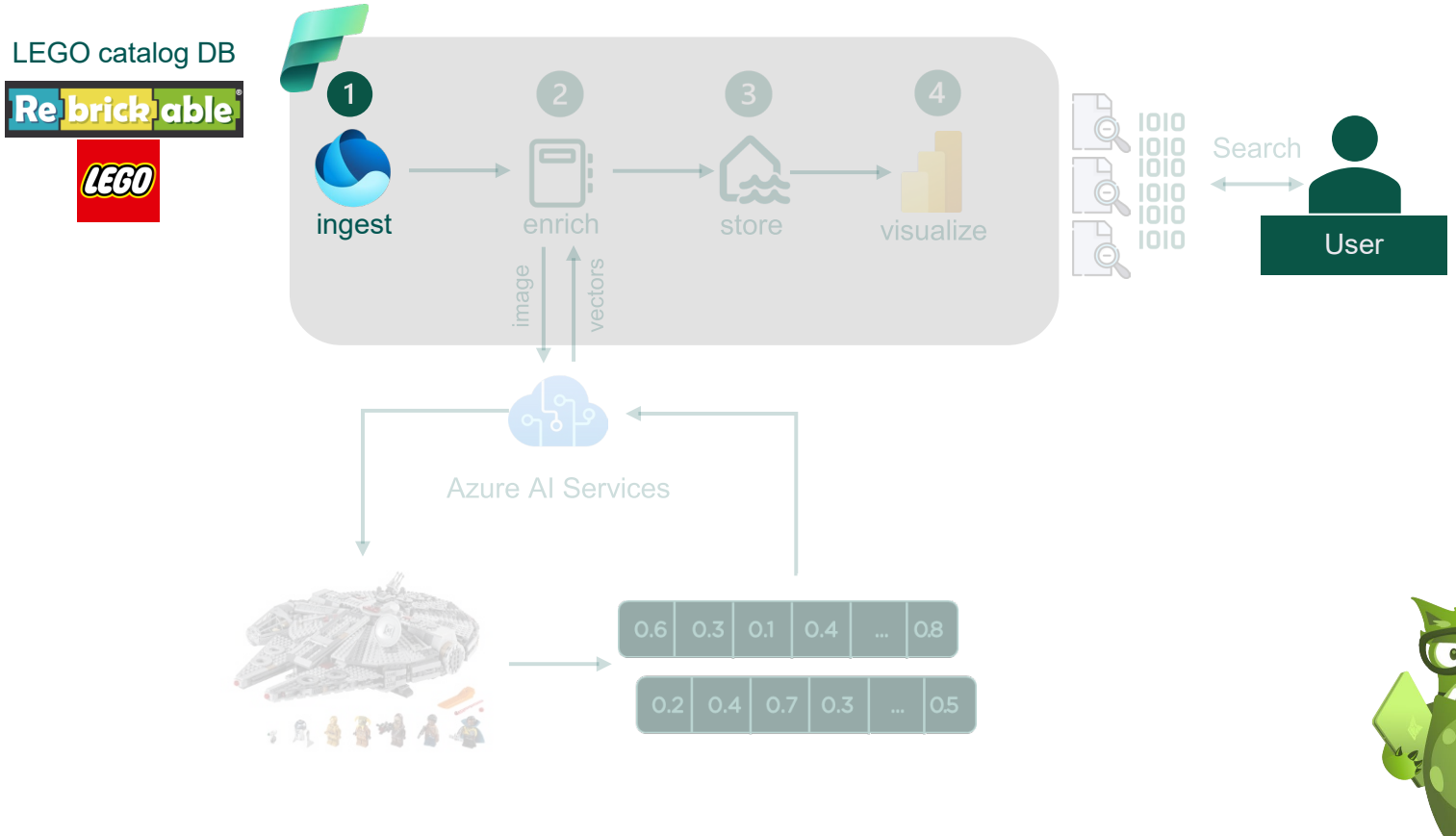
Azure AI Services







1 Ingest LEGO dataset into OneLake. Raw data is a CSV file.

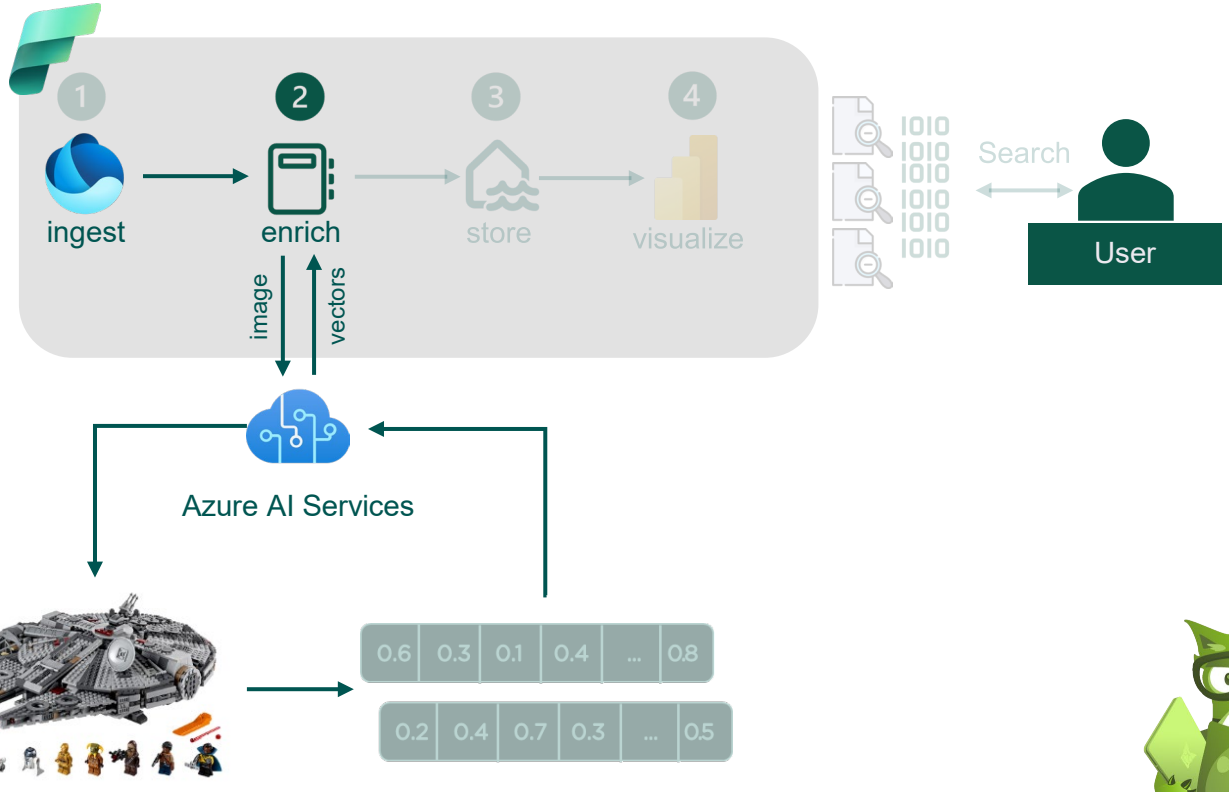




2 Enrich the images of LEGO datasets with their vector representation return by AI service

LEGO catalog DB

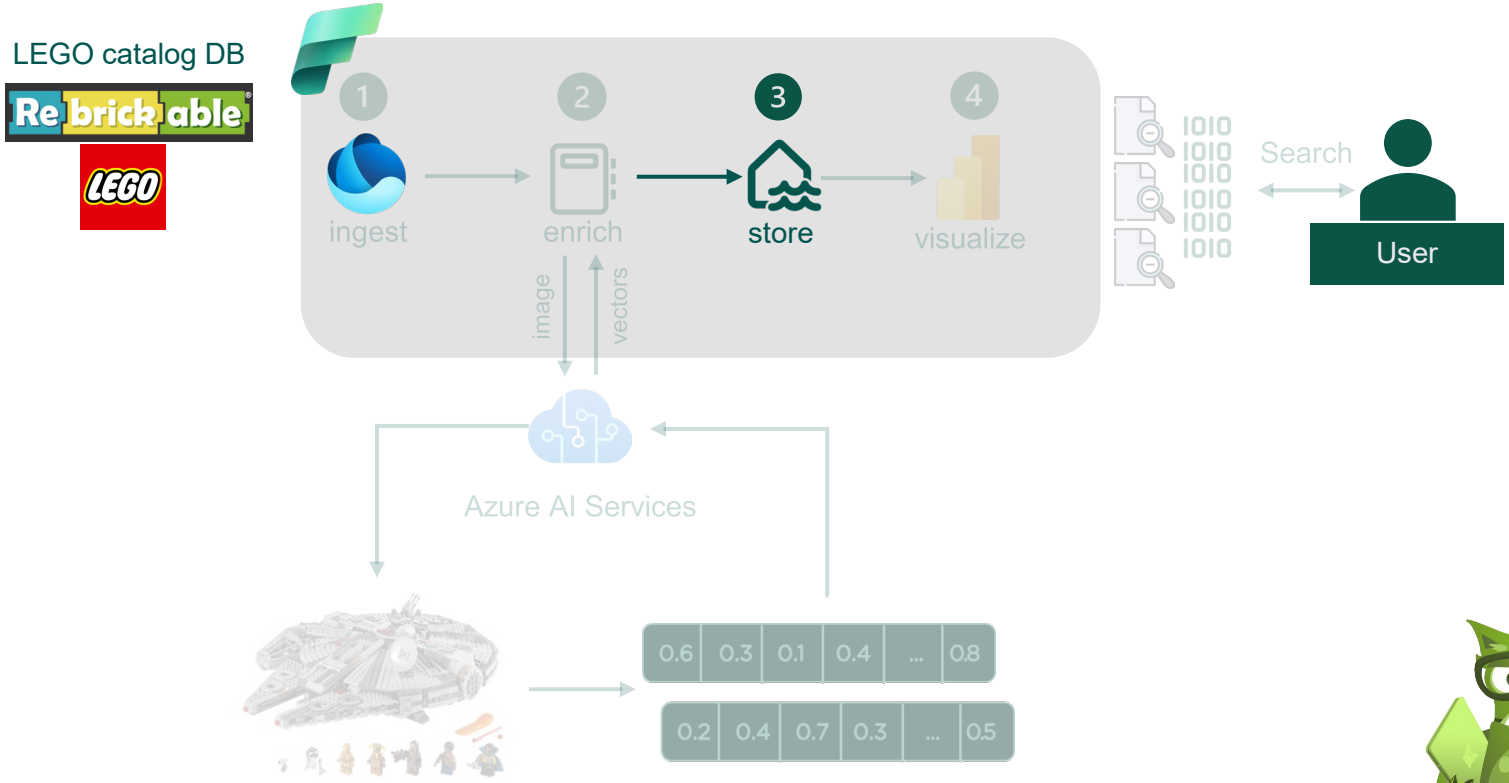






3 Store the LEGO data including the vector embeddings in a Lakehouse for consumption by Power BI

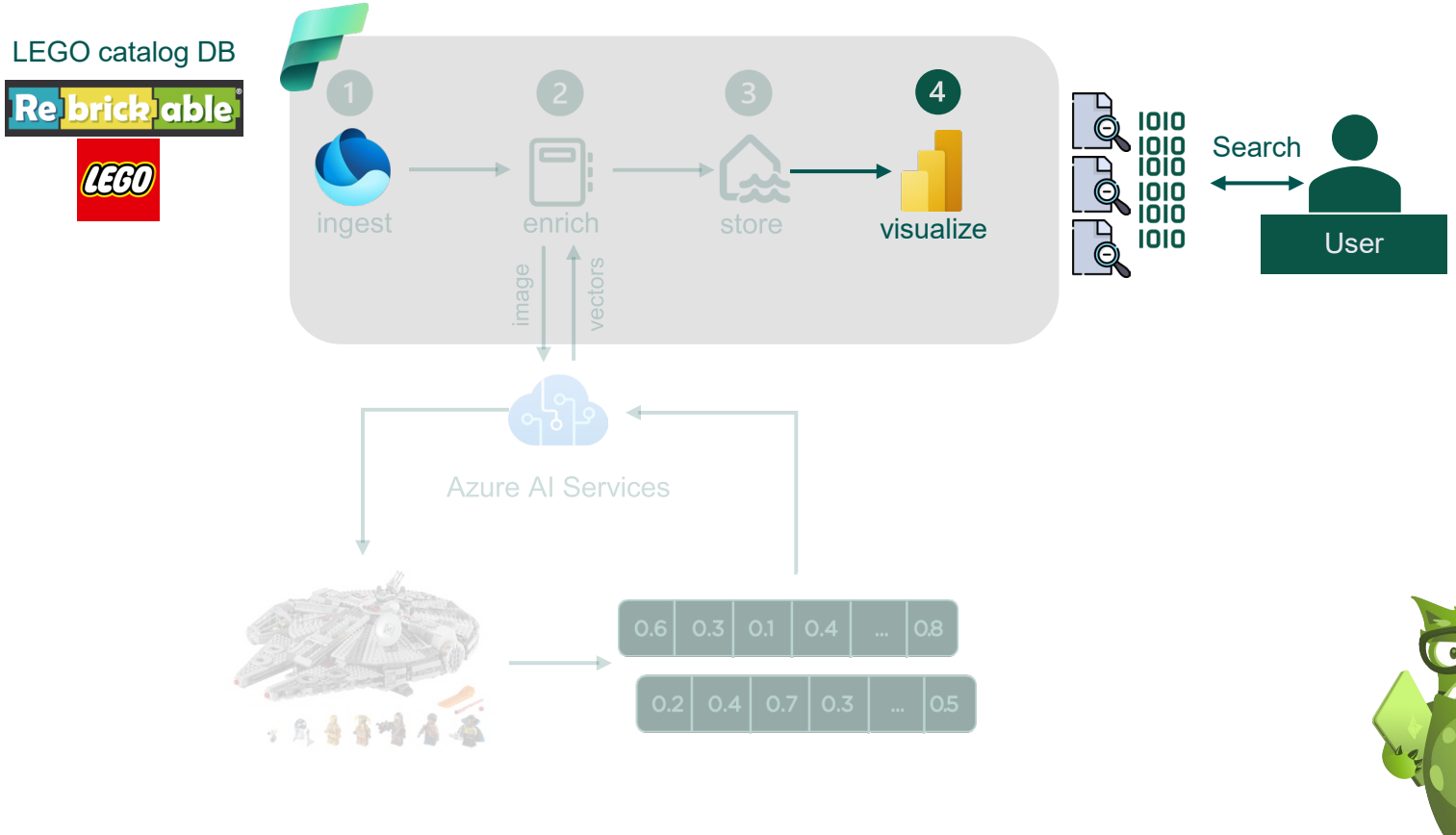
LEGO catalog DB





4 Create a semantic model including a DAX calculation for cosine similarity based on verctors. Build a report to visualize the search results

LEGO catalog DB





It’s time for some







Q & A





roger.unholz@bsgroup.ch

GitHub Repo



SQL KONFERENZ 2025

Session: Mastering Lakehouse Maintenance: From Basics to Automation with Semantic Link Labs

Data Engineering

Tweet Share Share Email

Maintaining your Lakehouse is crucial for ensuring optimal performance and scalability.

In this session, we'll explore the fundamentals of Lakehouse maintenance, why it matters, and how you can streamline these processes. Learn how to leverage the powerful Semantic Link Labs library to automate maintenance tasks with practical examples and actionable insights.

Whether you're a Data Engineer, Data Analyst, or a Fabric Administrator, this session will empower you to optimize your Lakehouse operations with ease.

PRESENTED BY:



Roger Unholz
Partner/Senior Consultant
BSgroup Data Analytics

Type
Session

UNHOLZ200

Give me Feedback

Thank you!







SQL
KONFERENZ
2025