COPEBIT

Business Apero
AI in der Praxis: Warum 80% nicht genug sind

Marco Kuendig

21st November 2024

COPEBIT

copebit AG
AWS Advanced Consulting Partner

aws PARTNER Advanced Tier Services

- Public Sector
- Immersion Day
- Solution Provider
- Amazon RDS Delivery
- Amazon EKS Delivery
- Storage Services Competency
- AWS CloudFormation Delivery
- Well-Architected Partner Program
- Amazon EC2 for Windows Server Delivery

Cubotoo AG - The customer

“Thanks to the cooperation with copebit AG, we came across ideas and opportunities during the course of the project that we had not previously focused on. Productivity has already been massively increased in the classic tendering and procurement process, and now we want to improve the experience even further."

Gregor Jann, CEO
Cubotoo AG

AWS
All-in

Serverless
All Serverless with Lambda and Managed Services

AI/ML Tools
Bedrock with Claude Sonnet
RAG with Bedrock KnowledgeBase
Textract

MySQL
as a database, hosted on Aurora, in the future on Autoscaled Aurora Serverless V2

CloudWatch
for Monitoring and Alerting

OpenSearch
For Indexing and Vector Databases

Control Tower
for Foundation and Security

COPEBIT

- Application Layer
- Model Layer
- Infrastructure Layer

Cubotoo wanted to leverage AI to make the ordering and tendering process even faster

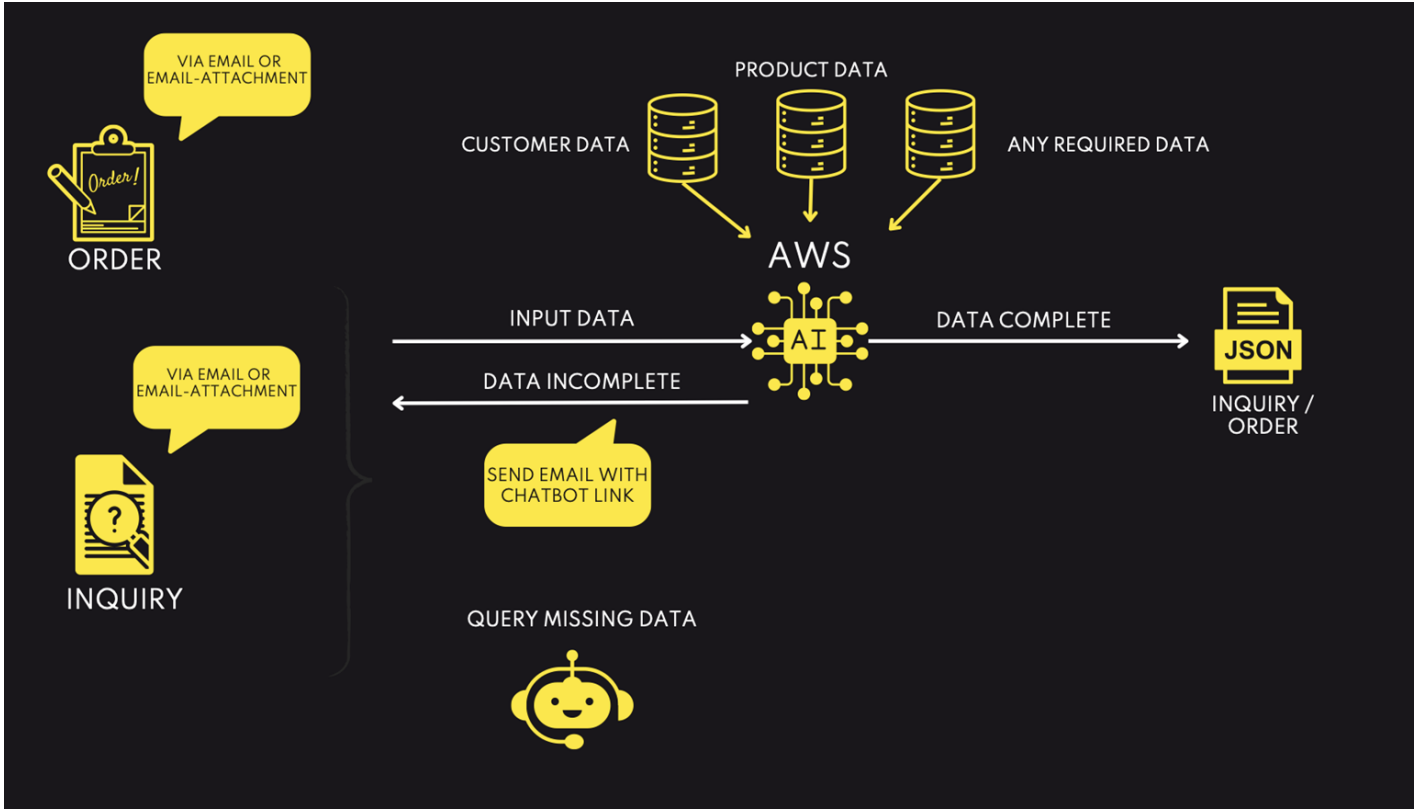
COPEBIT

Cubotoo and AI

- **AI-Powered Marketplace**: Efficient processing of customer inquiries
- **Automated Workflows**: Uses Bedrock-KB and RAG (Retrieval-Augmented Generation) methods
- **Rapid Order Confirmation**: Orders confirmed within minutes
- **Enhanced User Experience**: Streamlined and intuitive processes
- **Award-Winning Solution**: Recipient of the Swiss Digital Commerce Award

COPEBIT

First Whiteboard Brainstorming

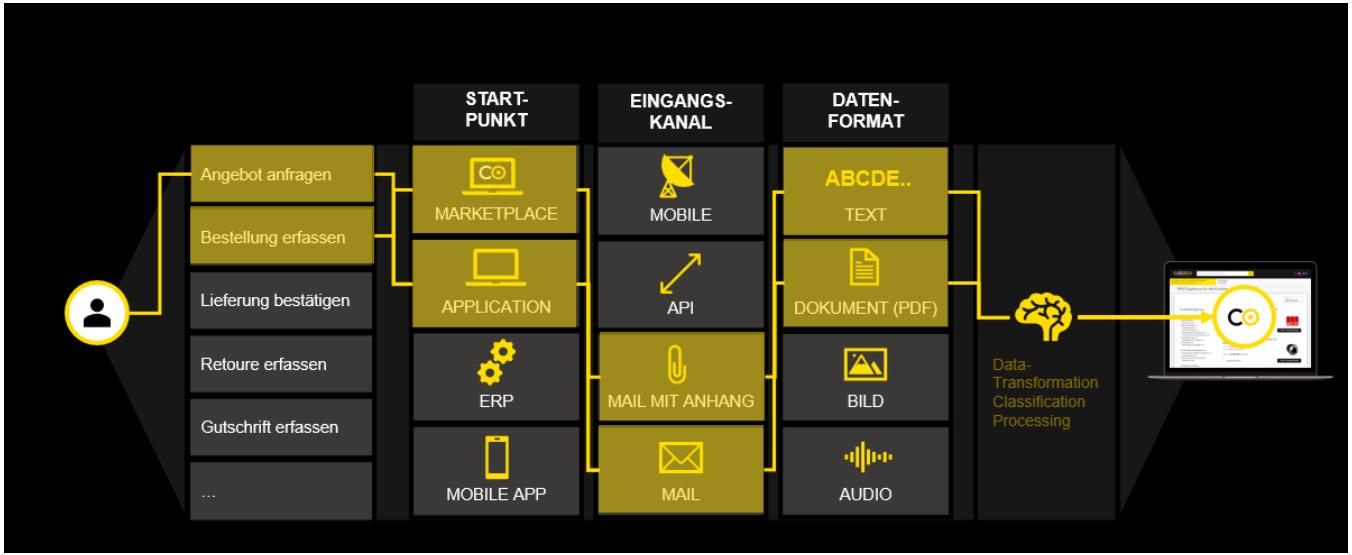


COPEBIT

Technical Overview

What is it?

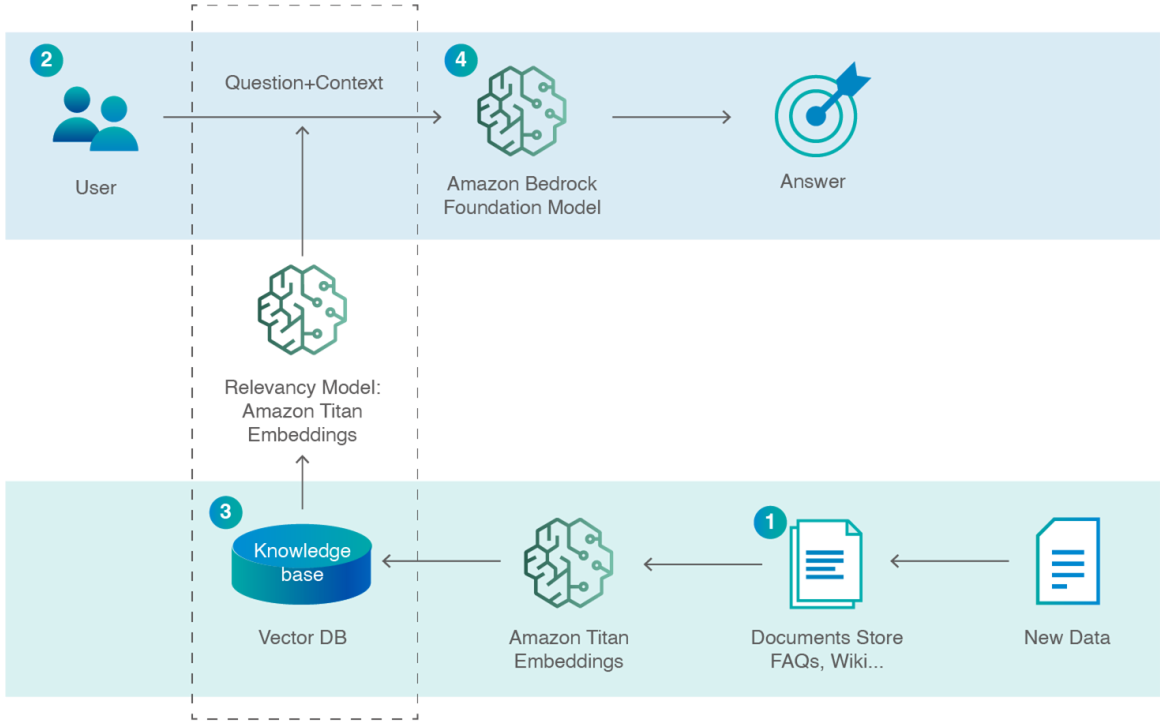
As cubotoo - marketplace for construction material - we would prefer that all users / roles (foreman, procurement manager, wholesales salesperson) that want to order / request offers directly from Manufacturer use our marketplace.
The problem: **Our marketplace loses against Project Specifications and E-Mail**



COPEBIT

How we do it

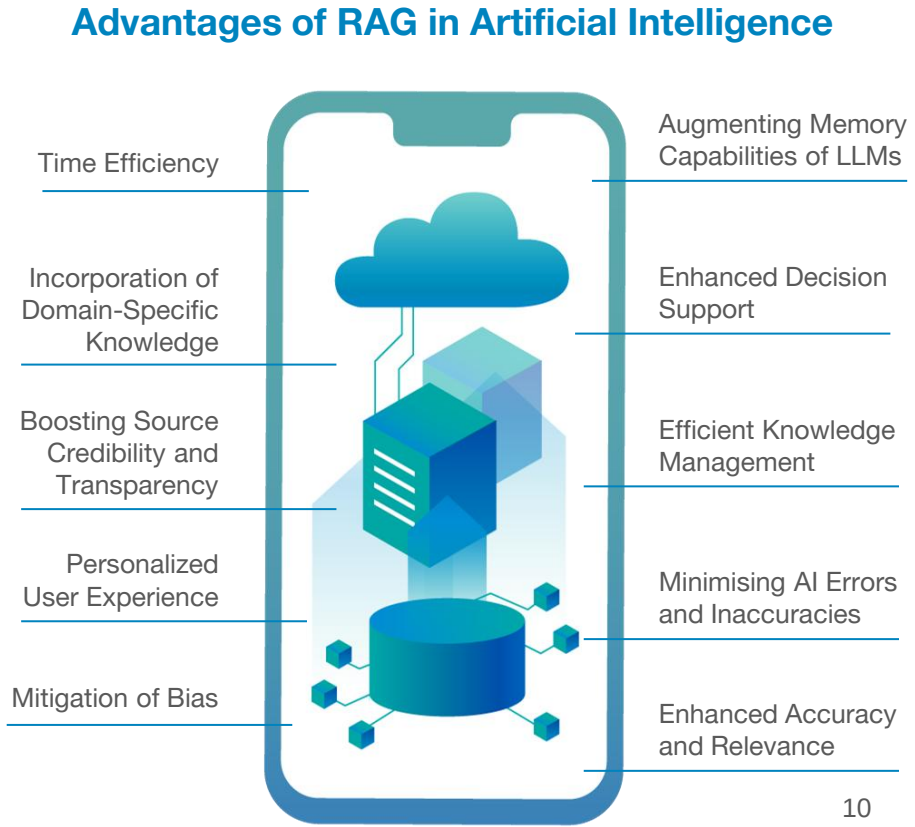
- Large Language Models
- Knowledge Base
- Retrieval Augmented Generation (RAG)
    - Product Catalog
    - Customer Database (CRM)
    - Customer Knowledge Bases
    - Documentation
- Data Enrichment and matching
- Output generation
- API’s
- Iterations



COPEBIT

Benefits

- Access to current information: Vector databases can be regularly updated with new information
- Domain-specific knowledge
- Transparency and trust
- Cost-effectiveness compared to fine tuning a LLM
- Efficient retrieval: Vector databases



COPEBIT

Lessons learned

COPEBIT

Takeaways

- Accuracy is key for acceptance: A first test at cubotoo with 80% accuracy was not good enough
- A human in the loop approach took out the bad AI Apples
- Technical
  - Multiple GenAI Iterations: needed when use case is not specific enough
  - Time consuming: If you have lots of data to prepare the chunking
  - Use tokens instead of RAG where possible

COPEBIT

Thank you!