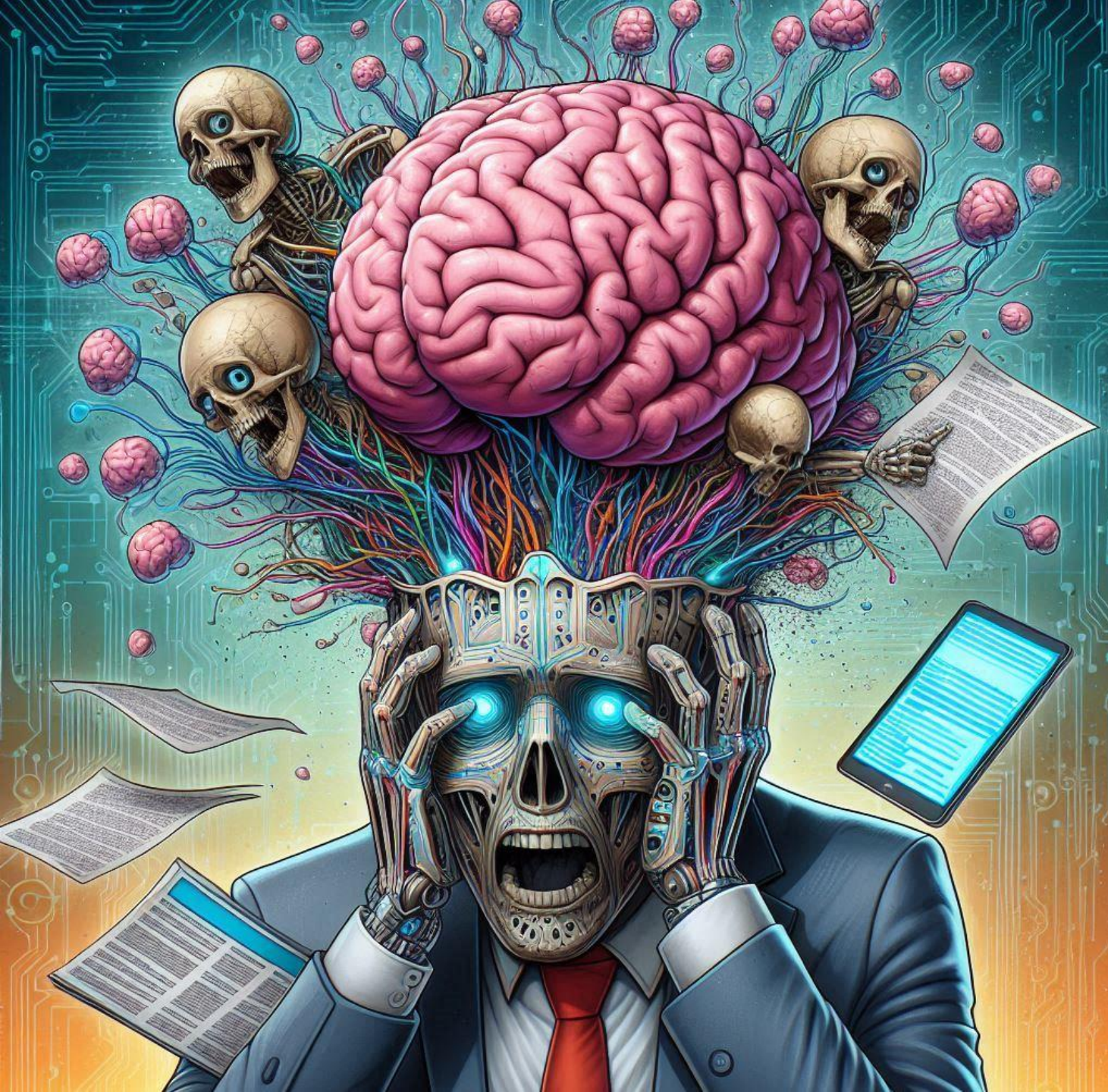




Infrastructure Meets AI: What should Azure infra- pro know about AI

Azure & Friends
22.8.2024



Who has AI
hype
overload?

Markus Lintuala

CLOUD HERO

markus@lintuala.fi | +358 40 585 5531

elisa



- 🧐 Can't say no for anything new
- 🗣️ Public speaker
- 😋 Gastronomic, craft beer and wine geek
- 🌍 Loves travelling and aviation
- 🚗 Waze map editor

@MarkusLintuala

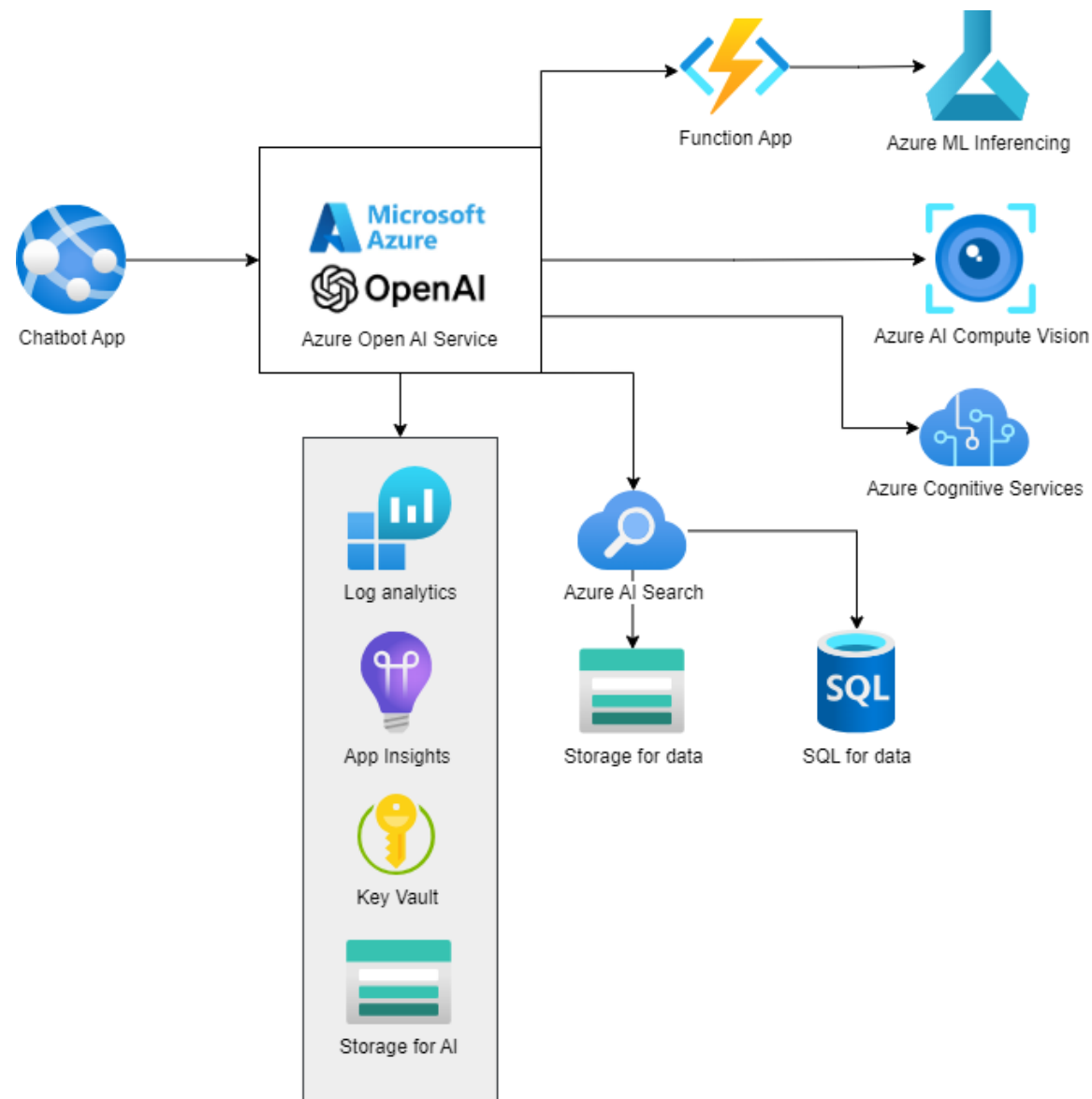
@mmaraa

What should Azure infra-pro know about AI?

- Resources
- Connectivity
- Security
- Published applications
- Published APIs
- Related resources
- Cost structure



Resources



Connectivity

- Internal / External connectivity
- Other networks
- Use Zero Trust concept for networking





Security

- Authentication for Apps, APIs, between resources
- Application attack vectors
- Use protection features, but do not restrict productivity
- Prefer usage of managed identities

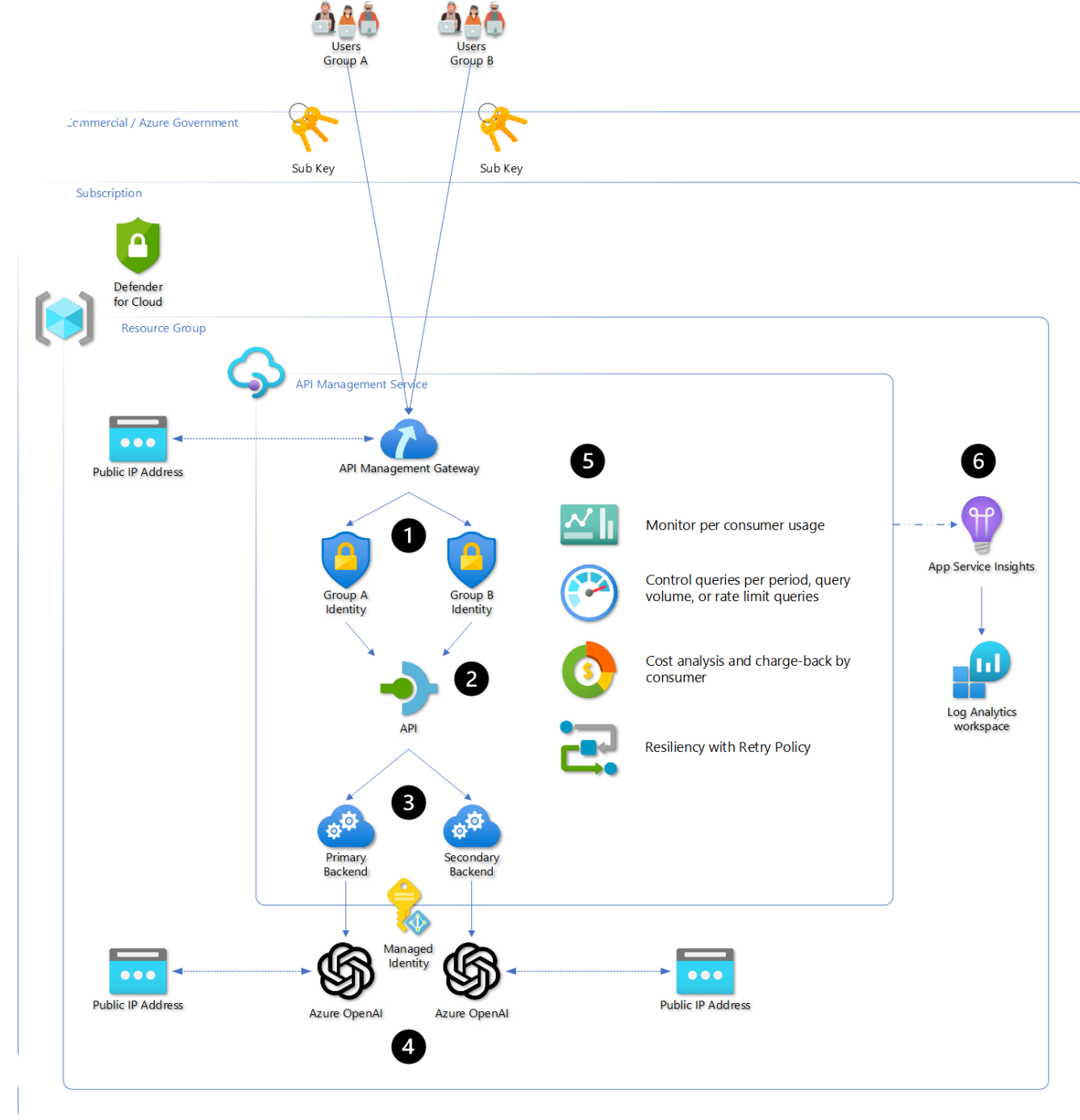


Published applications

- Authenticated or not
- Rate limits
- Minimize lateral movement
- DOS protection and WAF
- Cost protection

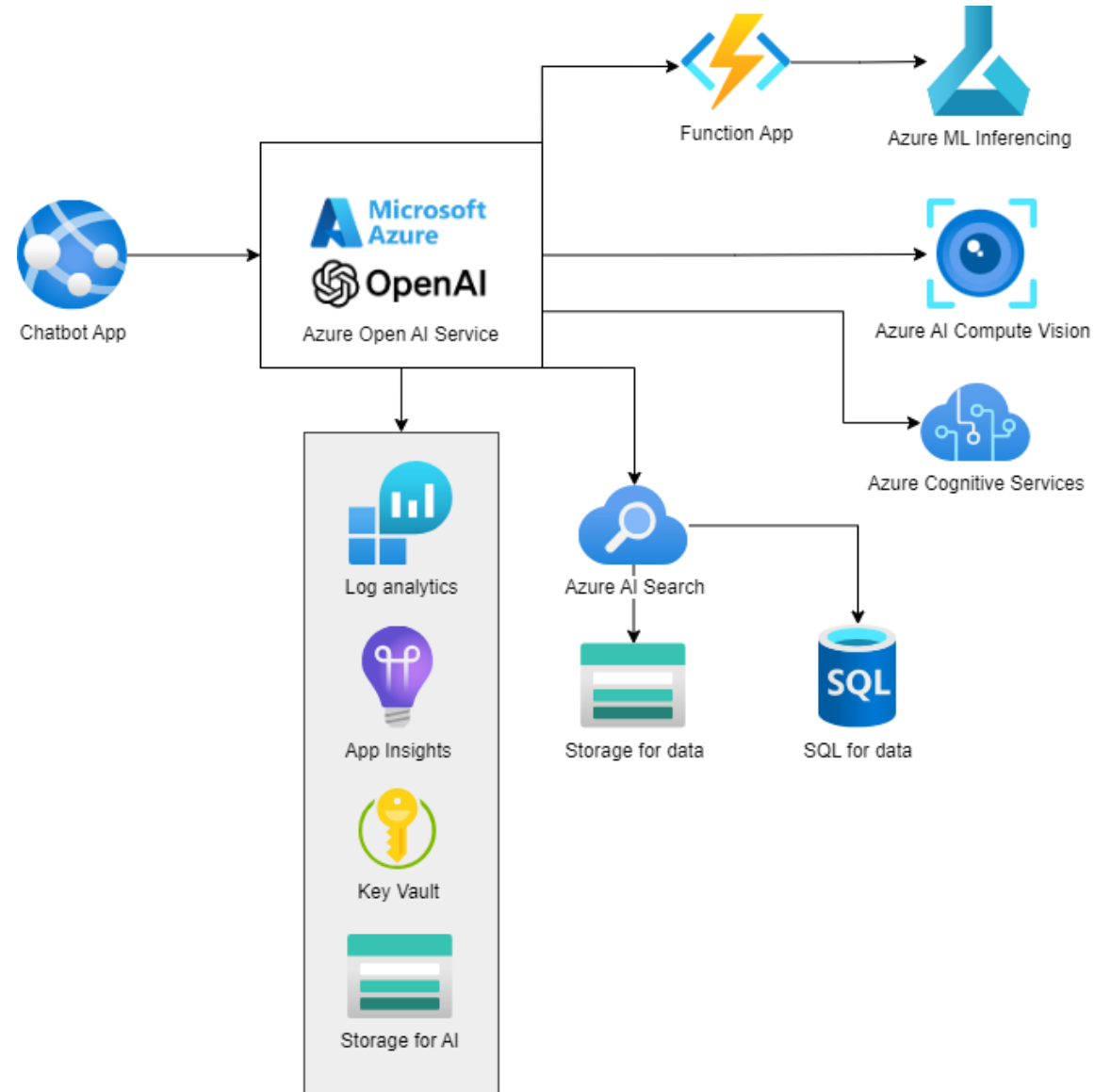
Published APIs

- **No** direct connections
- Use API Management
- Restrict access from network and require authentication



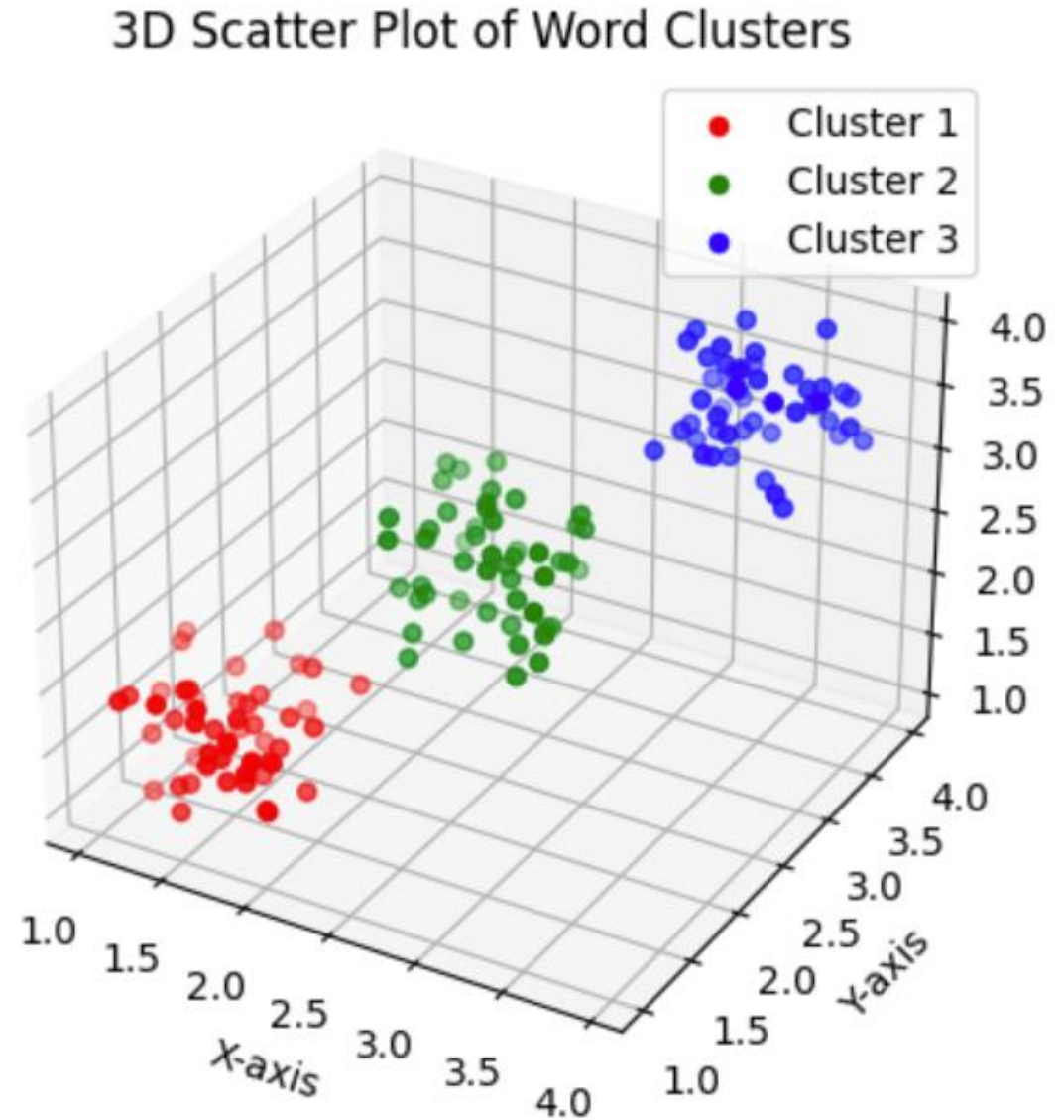
Related resources

- Vector databases
- Storage accounts
- Functions
- Other network connected applications



What is vector database?

- Multi-dimensional database for numeric values (vectors)
- Less DBA work
- DBA work contains much work with data retrievability





AI Cost structure

- OpenAI Service by PAYG
- AI Search by monthly fee (if using)
- API Management has own pricing
- App Service according to App Service Plan



Demo

A composite image featuring a futuristic cityscape with various skyscrapers, including the Petronas Towers, in the background. In the foreground, a brown paper coffee cup with a black lid and a futuristic circuit-like design is placed on a wooden surface. The cup has a glowing blue circular element in the center of its design. The background shows a river with a red boat and a bridge, all under a bright sky.

Takeaways

- **Secure** your applications and APIs
- **Watch out** your cost structure, use budgets!
- **Support** AI developers as much as you can

Q & A

