

dr Inez Okulska
AI Innovation Architect

TWARDO NA ZIEMI

CONTEXT GROUNDING I TRUST LAYER

**WŁASNE
MODELE UIPATH**

- CommPath
- Helios (DocPath)

**PRODUKTY I USŁUGI
WYKORZYSTUJĄCE
GENAI**

- Agents
- Autopilot for ...
- IXP/DU
- GenAI Activities

**Integracje zewnętrzne
z usługami genAI****CONTEXT GROUNDING**

(możliwość zbudowania RAGa)

BRAMA LLM**Dostępny API**

- **Wektory**
- **Wykrywanie danych wrażliwych**
- **Struktura chatu**

Możliwość operacyjna

- **Autoryzacja dostępu**
- **Limity użycia**
- **Governance**
- **Telemetria**
- **Monitoring/audyt**

**GUARDRAILS
PAMIĘĆ****KONFIGURACJA LLMów****MODELE (ML/AI)
OD ZEWNĘTRZNYCH
DOSTAWCÓW
ZARZĄDZANE PRZEZ
UIPATH****USŁUGI PLATFORMY UIPATH**

Chmura, Orchestrator, Semantic Search,
Azure Apps, Azure Cache, Automation Ops

LLMy zewnętrznych dostawców zarządzane przez klienta

ChatGPT (Azure OpenAI) Google VertexAI (Gemini) Anthropic (Claude)

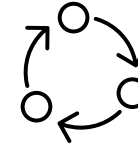
GOVERNANCE



AUTORYZACJA
DOSTĘPU



ZARZĄDZANIE
KLUCZAMI



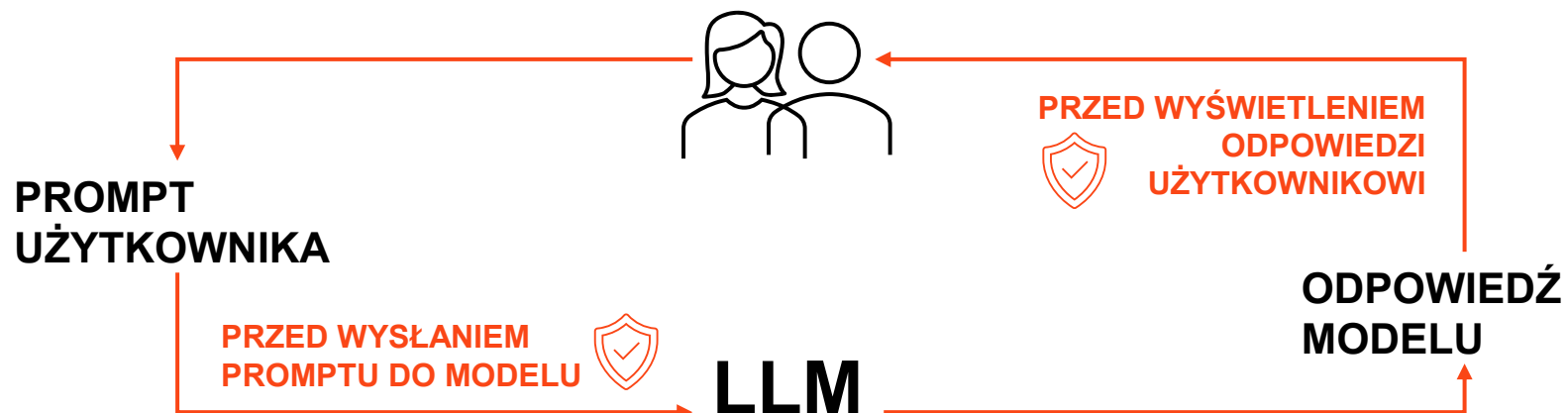
CIĄGŁOŚĆ
POŁĄCZENIA



KONTROLA
KOSZTÓW

Usage summary	Audit	AI governance	Autopilot for Everyone	Context grounding
Manage AI Autopilot usage for your organization, tenants, or user groups with AI policies to guarantee appropriate developer access to Generative AI features. Learn more				
Search		Manage deployments		
Policy name ÷	Deployments (Tenants) ÷	Deployments (Groups) ÷	Deployments (Users) ÷	
AITL	Playground and 1 more	--	--	
Test	--	--	--	
AITrustLayer	--	--	--	

GUARDRAILS – ZAPORY BEZPIECZEŃSTWA

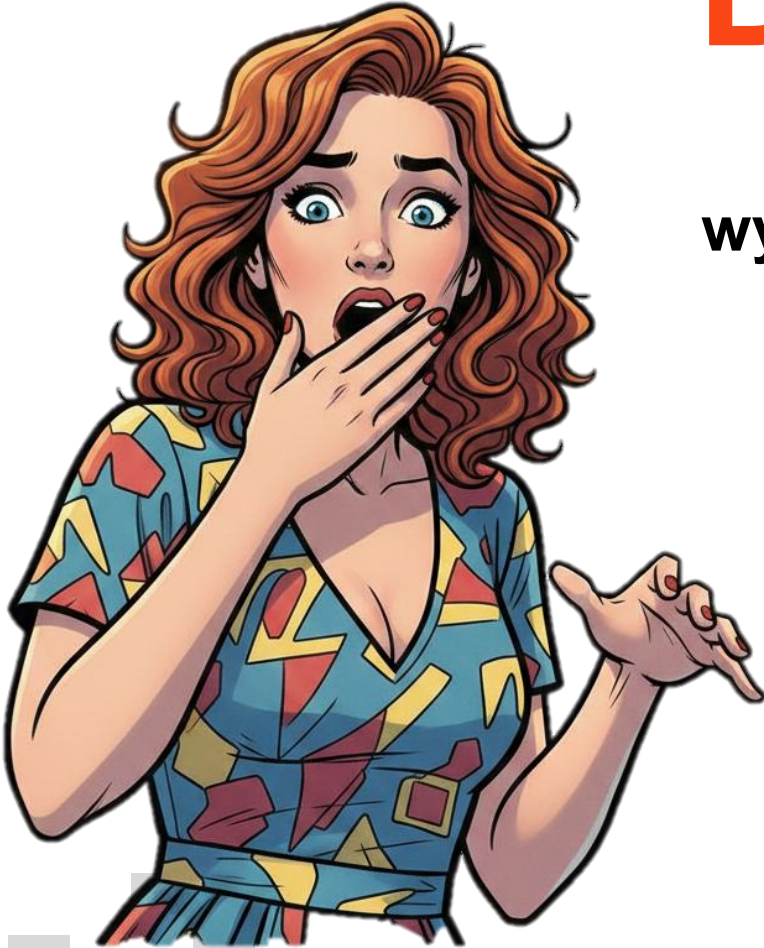


Gwarantowana moderacja treści w oparciu między innymi o kategorie treści szkodliwych wyróżnione w oficjalnych manifestach:

- AWS Bedrock
- Azure OpenAI
- Google Vertex

+ dodatkowo możliwość tworzenia własnych guardrails indywidualnie dla każdego agenta

DANE WRAŻLIWE!



wykrywanie

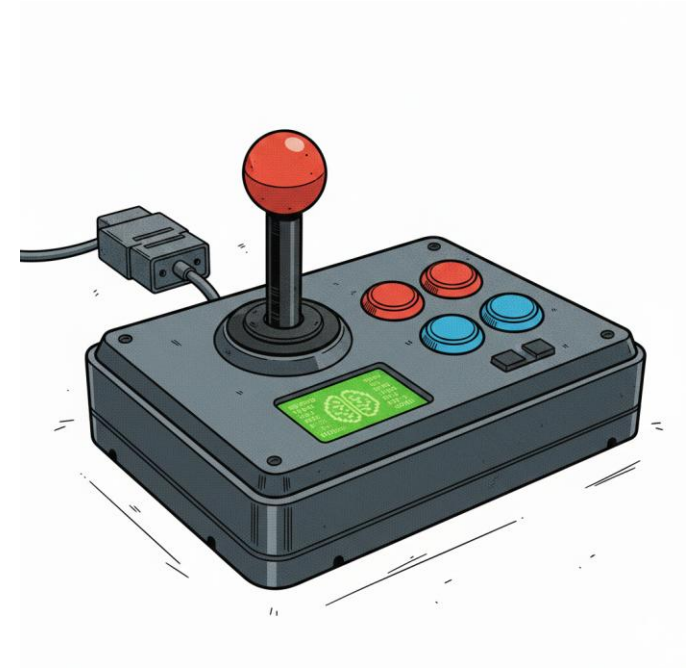
filtrowanie

maskowanie

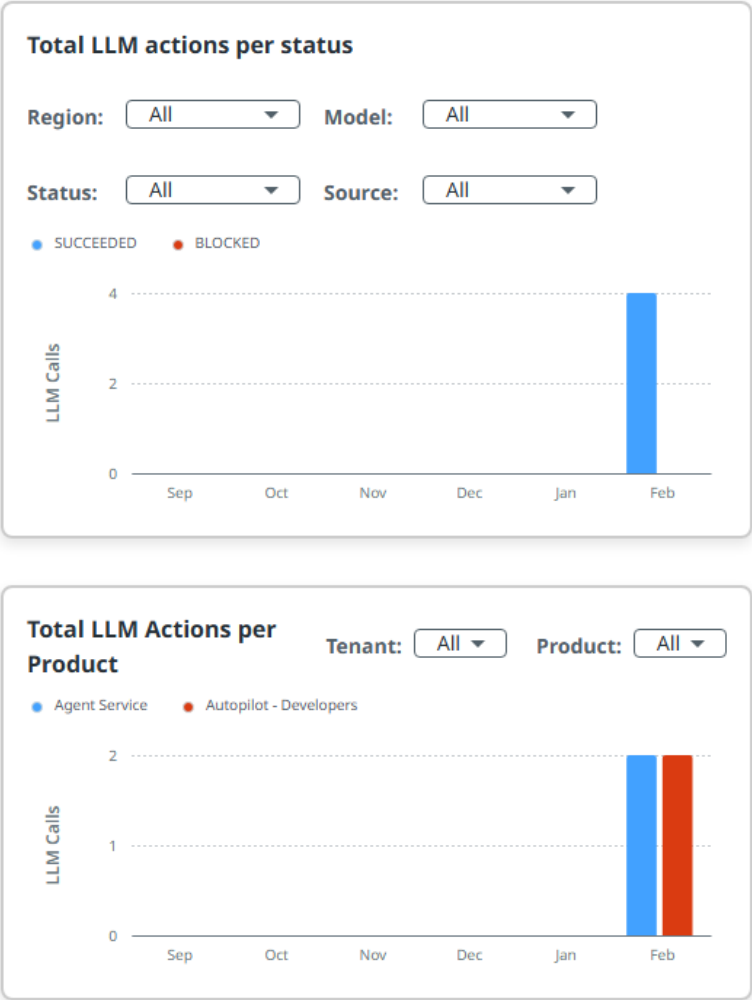
- Model wykrywa, filtruje lub maskuje dane wrażliwe po wpisaniu promptu – **przechwytuje go PRZED** wysłaniem promptu do modelu (np. GPT, Gemini).
- Trzeba tę funkcję samodzielnie włączyć (!), nie działa z automatu – bo może wpływać na jakość odpowiedzi LLM.
- Wykorzystane do tego usługi Azure Cognitive Services – a w nich dedykowane modele ML, które do tej anonimizacji nie wykorzystują LLMów, więc nie wysyłają danych do chmury

AUTOPILOT

- **for Everyone**
- **for Testers**
- **for Developers**
- **for Business Analysts**
- ...



KONTROLA NAJWYŻSZĄ FORMĄ ZAUFAANIA



Monitoring i audyt

Wszystkie modele, które przechodzą przez AI Trust Layer są objęte monitoringiem – można dokładnie obserwować zapisane akcje i statystyki aż do 180 dni wstecz. W zakładce „Audyt” znajdziemy:

- | | |
|------------------|--------------------|
| znaczniki czasu | użycie modeli |
| użycie produktów | lokalizacja modeli |
| użytkownik | konfiguracja |
| tenant | status |

OBSERWOWALNE ŚLADY = KONTROLA I PAMIĘĆ

Dzięki rozbudowanej strukturze monitorowania i zapisywania ścieżek, jakimi podążają modele, sekwencji wywołań oraz interakcji z narzędziami platformy i zewnętrznymi usługami – **agenci mają szansę uczyć się na własnych sukcesach i błędach.**

Nie muszą za każdym razem zaczynać od zera, z tego samego punktu, jakby nie istniało ani wczoraj ani jutro.



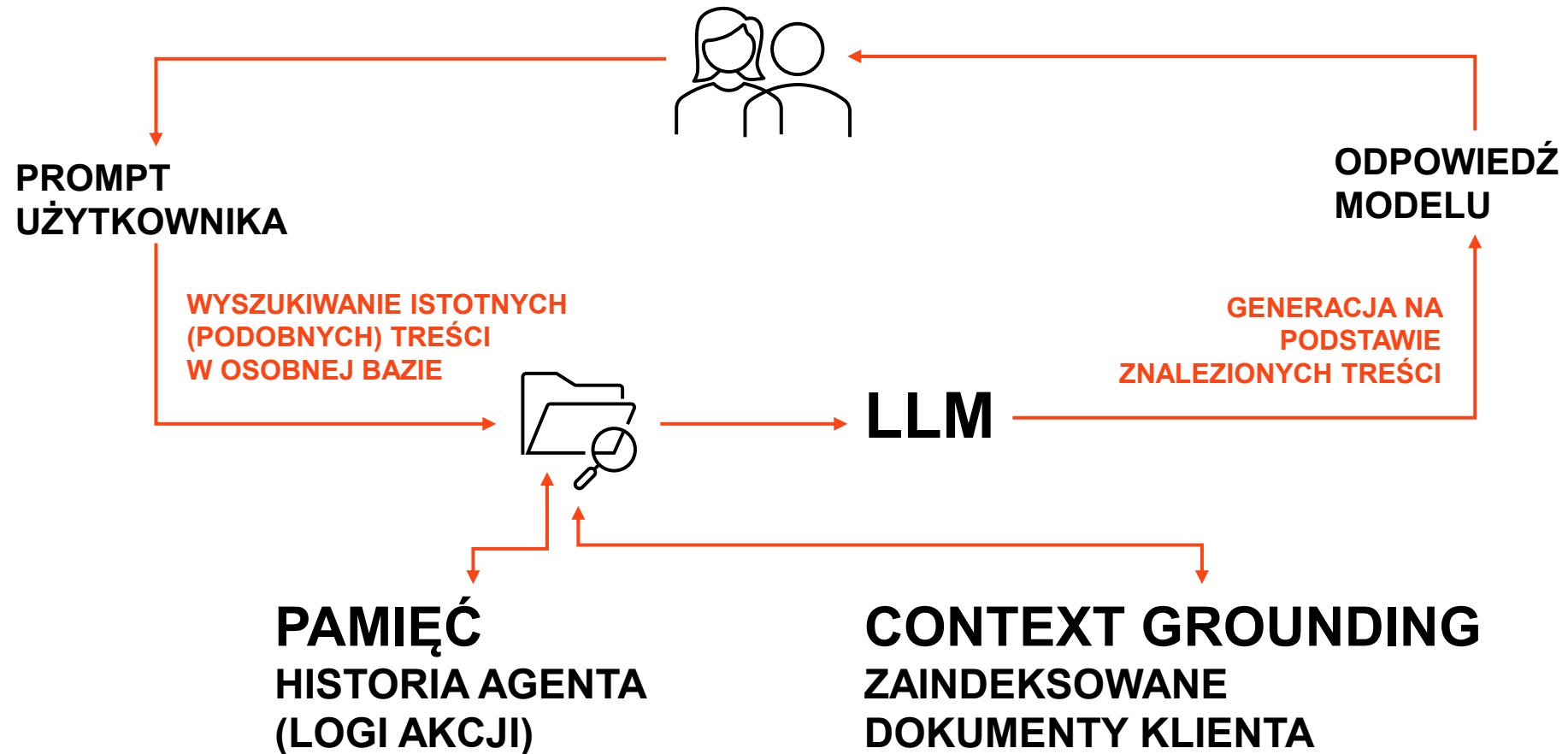
**DZIEŃ
ŚWISTAKA?
NIE DLA
AGENTÓW!**

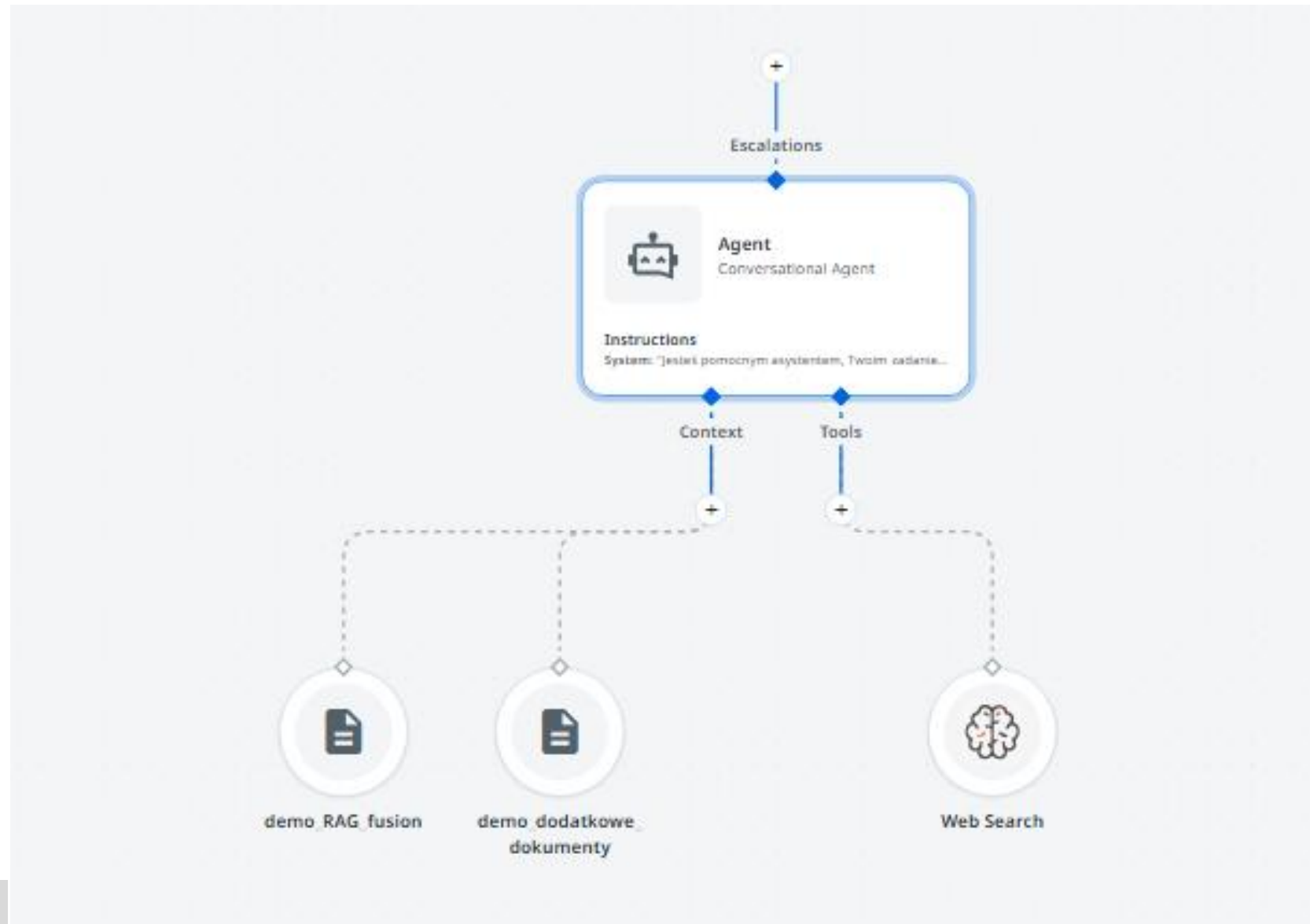
Od dekad znane w teorii sterowania „feedback control”, to sterowanie oparte na minimalizacji **najświeższego** błędu

Szeroki **monitoring i analiza logów** tworzą obraz schematów akcji, z których można wyciągać wnioski i **optymalizować**

KONTEKST MA ZNACZENIE

Mechanizm typu RAG (Retrieval Augmented Generation)





Agent

Ask Autopilot



[Terms](#) | [Privacy Policy](#) | [Chat History Policy](#)

dr Inez Okulska,
AI Innovation Architect

DZIĘKUJĘ!