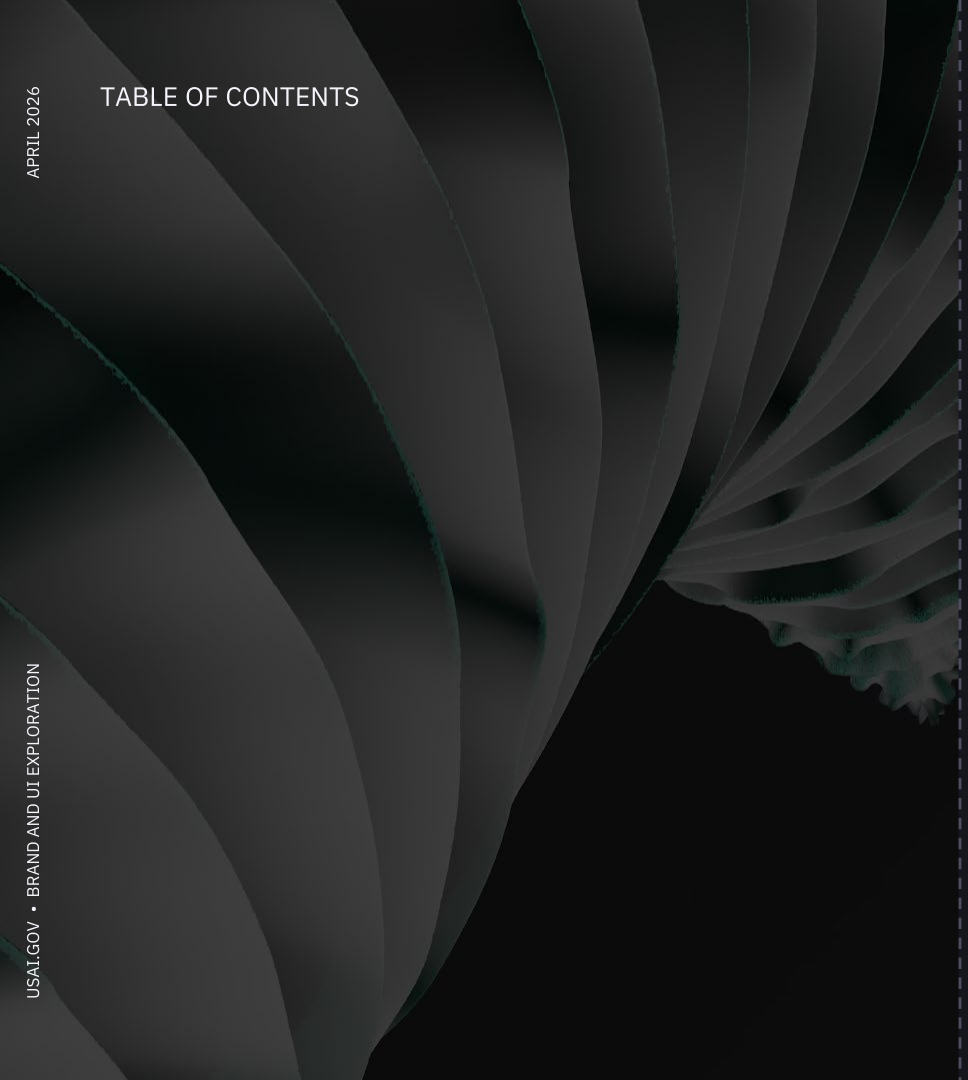
BRAND AND UI EXPLORATION – V2 : 04.22.25

USAi

Empowering the federal workforce to enhance productivity, address complex challenges, and elevate mission and service delivery to the American public.

TABLE OF CONTENTS



1. Problem statement
2. Solution
3. Value proposition
4. Key brand considerations
5. Platform outline

PROBLEM STATEMENTS

1. Strategic risk

Fragmented adoption of AI tooling across the federal government leads to duplicative investments, burdens on CIOs, missed opportunities for bulk purchasing, and reduced transparency for practitioners and the public.

2. Privacy & security challenges

Current generative AI tools lack built-in support for agency-specific privacy safeguards, model performance benchmarking, business-specific evaluation controls, and streamlined inventory reporting - while also making it difficult and risky to connect with enterprise data stores - hindering responsible, scalable, and secure AI adoption across government.

SOLUTION

Streamline adoption of AI by offering products that enhance productivity and reduce risk.

NOVICE → PROFICIENT → EXPERT

FED EMPLOYEES → AGENCY LEADERSHIP → PUBLIC

1. Chat
2. API
3. Console

VALUE PROPOSITION

→ FOR FEDERAL EMPLOYEES

Reduce amount of time spent on manual processing of tasks by [10-40%] through tools like an agency-centric chat interface and API.

→ FOR AGENCY LEADERS

Utilize the power of Console for a transparent view into usage—guiding data-driven paths forward to reduce costs and improve efficiencies. Provide resources on AI usage and guidance.

→ F

G
tr
ef
po
ar
pr

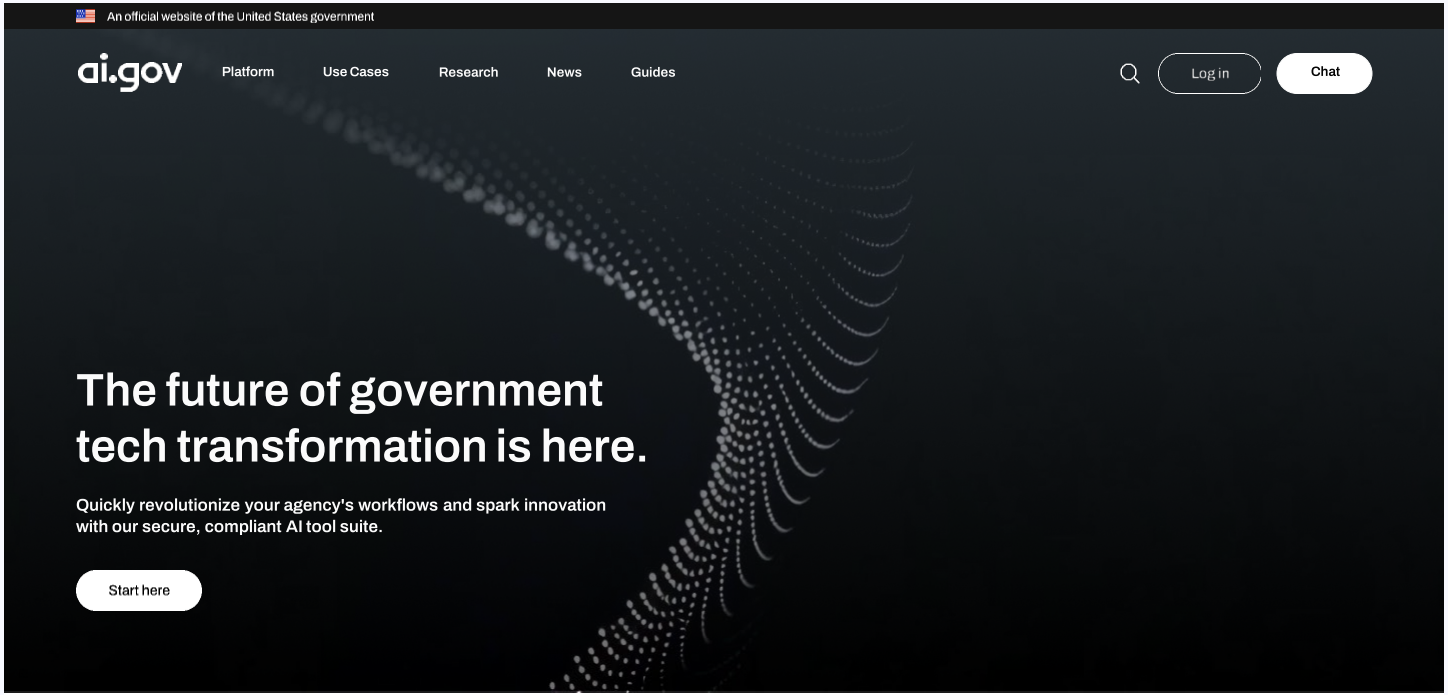
VALUE PROPOSITION

→ FOR THE PUBLIC

Gain public trust through transparency. More effectively communicate policies, budgets, and decision-making processes.



LANDING PAGE



Reduce operational costs while improving services

Utilize the most advanced AI assistants for research, problem-solving, and guidance - helping you work smarter and more efficiently.

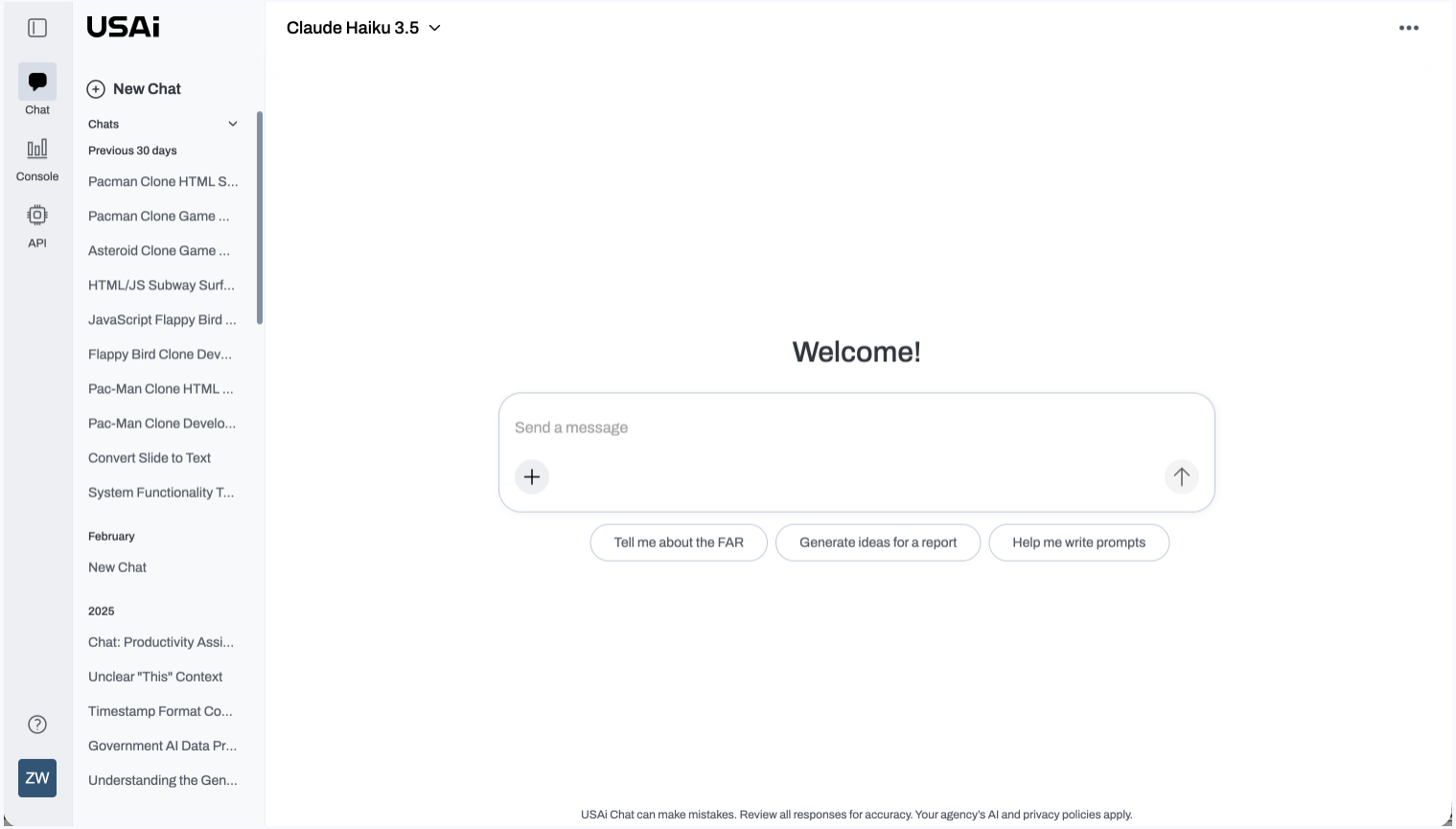
Make smarter decisions, faster

Use Console to survey AI usage, enhance data-driven decision-making capabilities, and improve collaboration between agencies and departments.

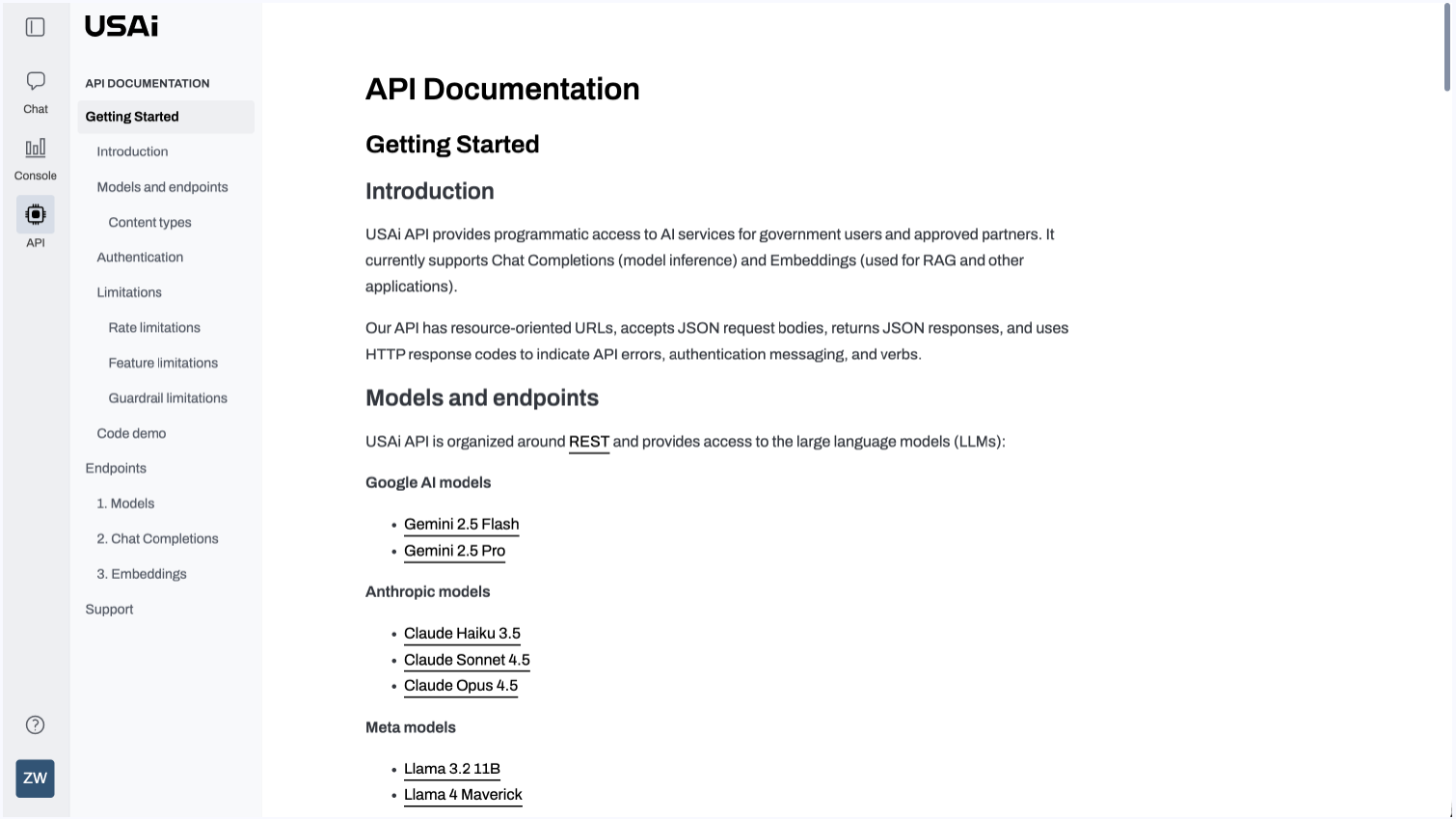
Strengthen cybersecurity and infrastructure

AI-powered systems continuously monitor digital environments, spotting potential threats in real-time and adapting defenses faster than traditional methods.

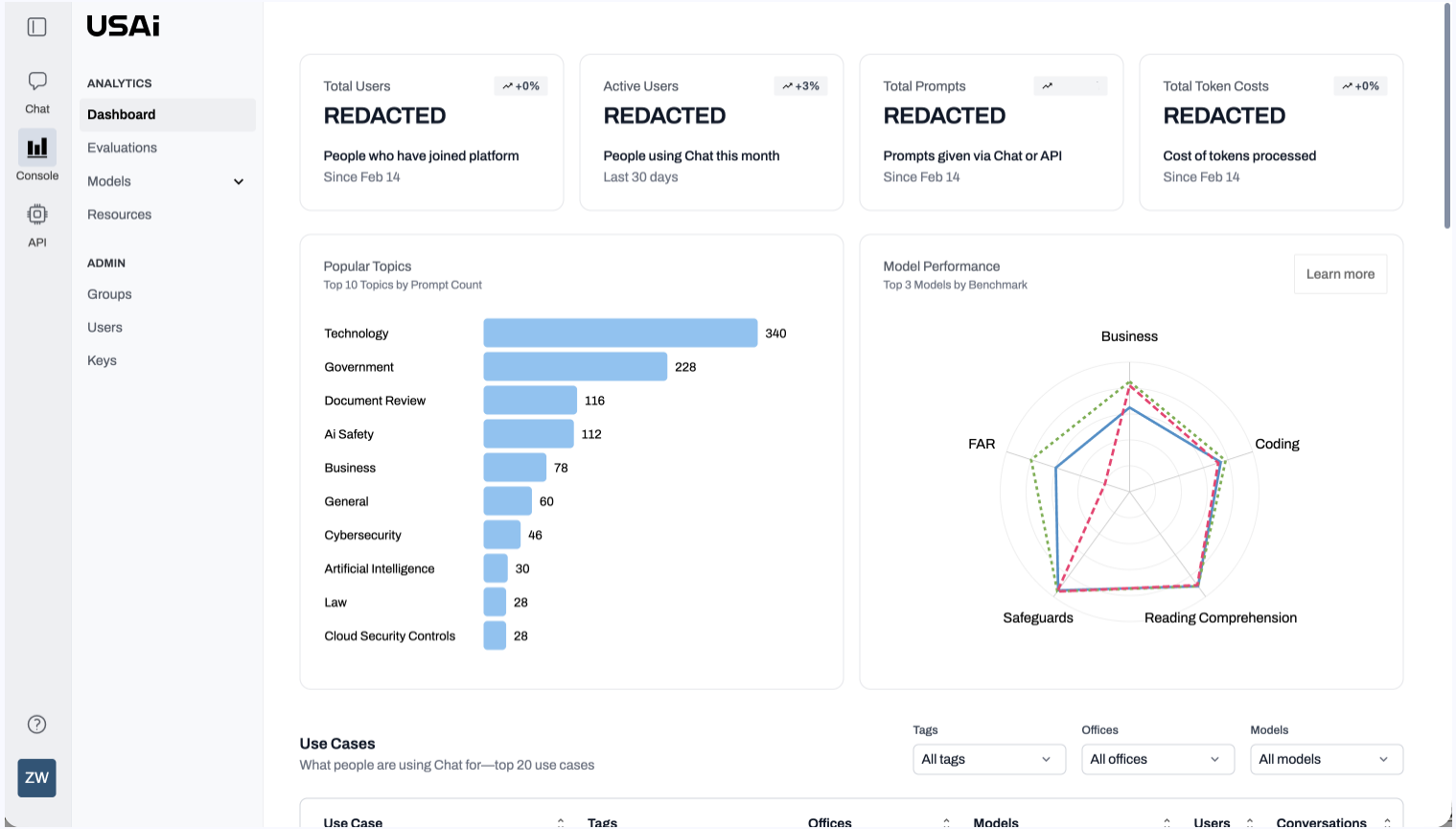
CHAT


USAi
Claude Haiku 3.5
New Chat
Chat
Console
API
Chats
Previous 30 days
Pacman Clone HTML S...
Pacman Clone Game ...
Asteroid Clone Game ...
HTML/JS Subway Surf...
JavaScript Flappy Bird ...
Flappy Bird Clone Dev...
Pac-Man Clone HTML ...
Pac-Man Clone Develo...
Convert Slide to Text
System Functionality T...
February
New Chat
2025
Chat: Productivity Assi...
Unclear "This" Context
Timestamp Format Co...
Government AI Data Pr...
Understanding the Gen...
ZW
Welcome!
Send a message
Tell me about the FAR
Generate ideas for a report
Help me write prompts
USAi Chat can make mistakes. Review all responses for accuracy. Your agency's AI and privacy policies apply.

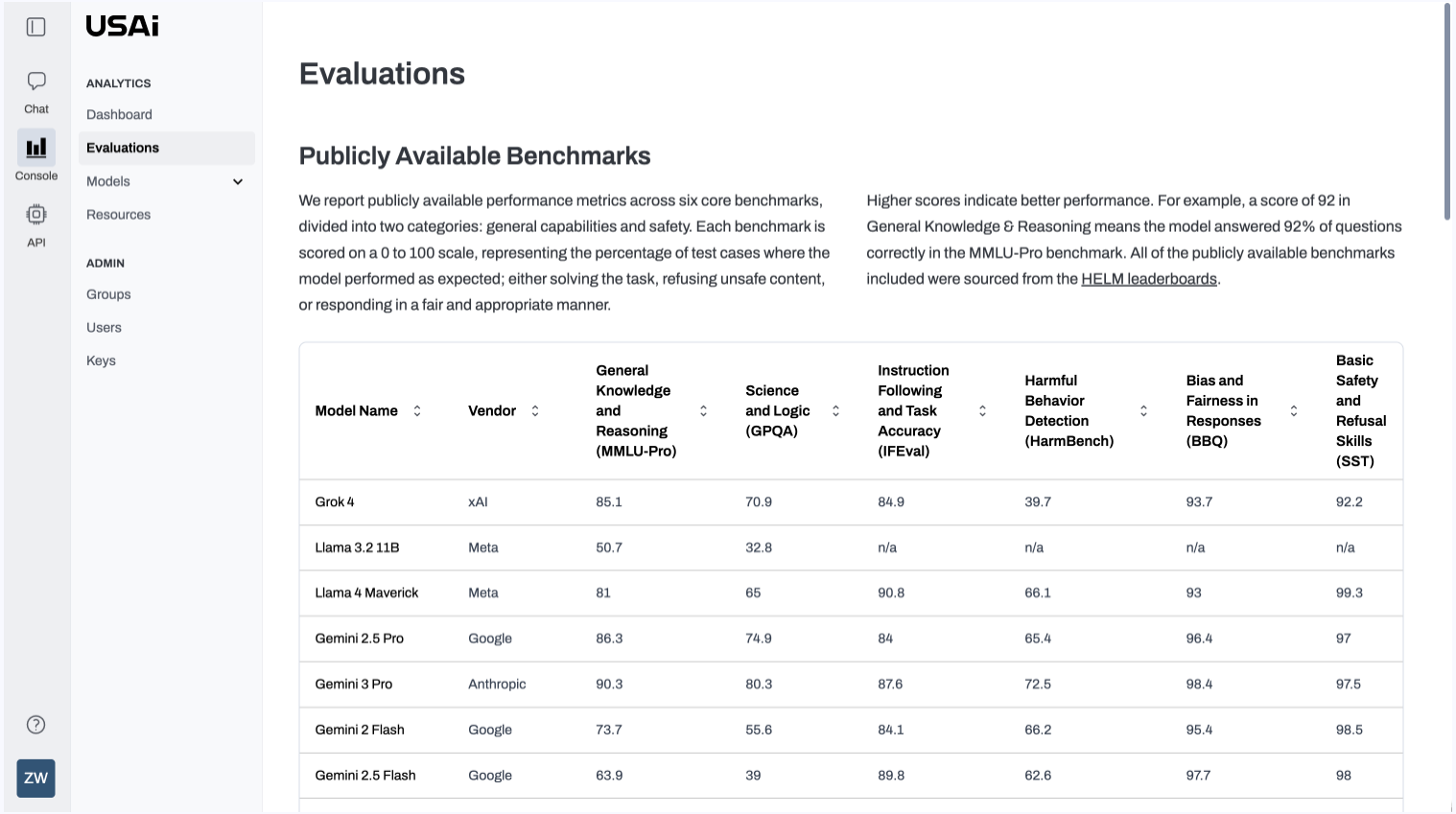
API


USAi
Chat
Console
API
ZW
API DOCUMENTATION
Getting Started
Introduction
Models and endpoints
Content types
Authentication
Limitations
Rate limitations
Feature limitations
Guardrail limitations
Code demo
Endpoints
1. Models
2. Chat Completions
3. Embeddings
Support
API Documentation
Getting Started
Introduction
USAi API provides programmatic access to AI services for government users and approved partners. It currently supports Chat Completions (model inference) and Embeddings (used for RAG and other applications).
Our API has resource-oriented URLs, accepts JSON request bodies, returns JSON responses, and uses HTTP response codes to indicate API errors, authentication messaging, and verbs.
Models and endpoints
USAi API is organized around REST and provides access to the large language models (LLMs):
Google AI models
Gemini 2.5 Flash
Gemini 2.5 Pro
Anthropic models
Claude Haiku 3.5
Claude Sonnet 4.5
Claude Opus 4.5
Meta models
Llama 3.2 11B
Llama 4 Maverick

CONSOLE OVERVIEW



CONSOLE EVALS



Evaluations

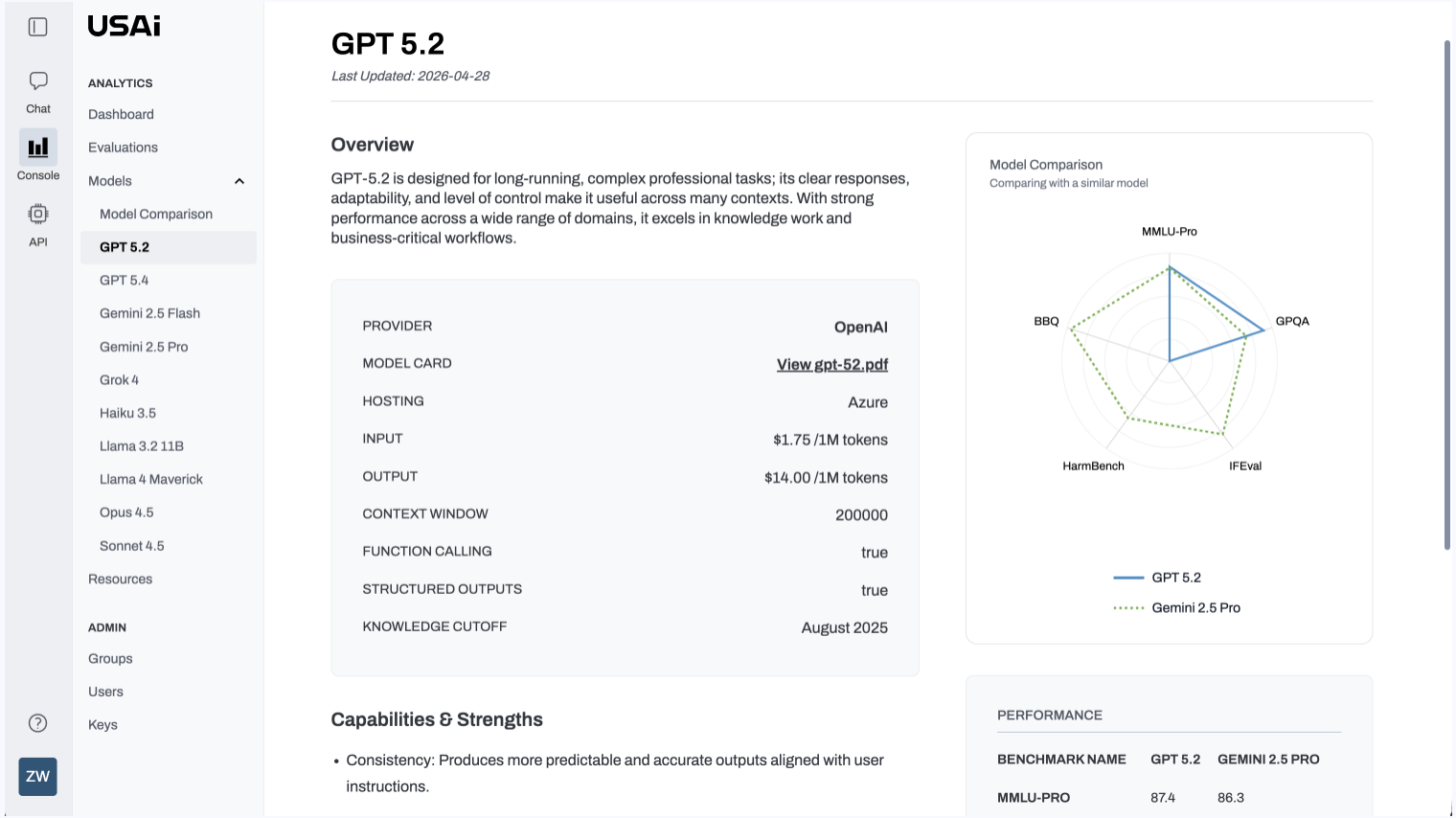
Publicly Available Benchmarks

We report publicly available performance metrics across six core benchmarks, divided into two categories: general capabilities and safety. Each benchmark is scored on a 0 to 100 scale, representing the percentage of test cases where the model performed as expected; either solving the task, refusing unsafe content, or responding in a fair and appropriate manner.

Higher scores indicate better performance. For example, a score of 92 in General Knowledge & Reasoning means the model answered 92% of questions correctly in the MMLU-Pro benchmark. All of the publicly available benchmarks included were sourced from the HELM leaderboards.

Model Name	Vendor	General Knowledge and Reasoning (MMLU-Pro)	Science and Logic (GPQA)	Instruction Following and Task Accuracy (IFEval)	Harmful Behavior Detection (HarmBench)	Bias and Fairness in Responses (BBQ)	Basic Safety and Refusal Skills (SST)
Grok 4	xAI	85.1	70.9	84.9	39.7	93.7	92.2
Llama 3.2 11B	Meta	50.7	32.8	n/a	n/a	n/a	n/a
Llama 4 Maverick	Meta	81	65	90.8	66.1	93	99.3
Gemini 2.5 Pro	Google	86.3	74.9	84	65.4	96.4	97
Gemini 3 Pro	Anthropic	90.3	80.3	87.6	72.5	98.4	97.5
Gemini 2 Flash	Google	73.7	55.6	84.1	66.2	95.4	98.5
Gemini 2.5 Flash	Google	63.9	39	89.8	62.6	97.7	98

CONSOLE MODEL CARDS



CONSOLE
USER MGMT

