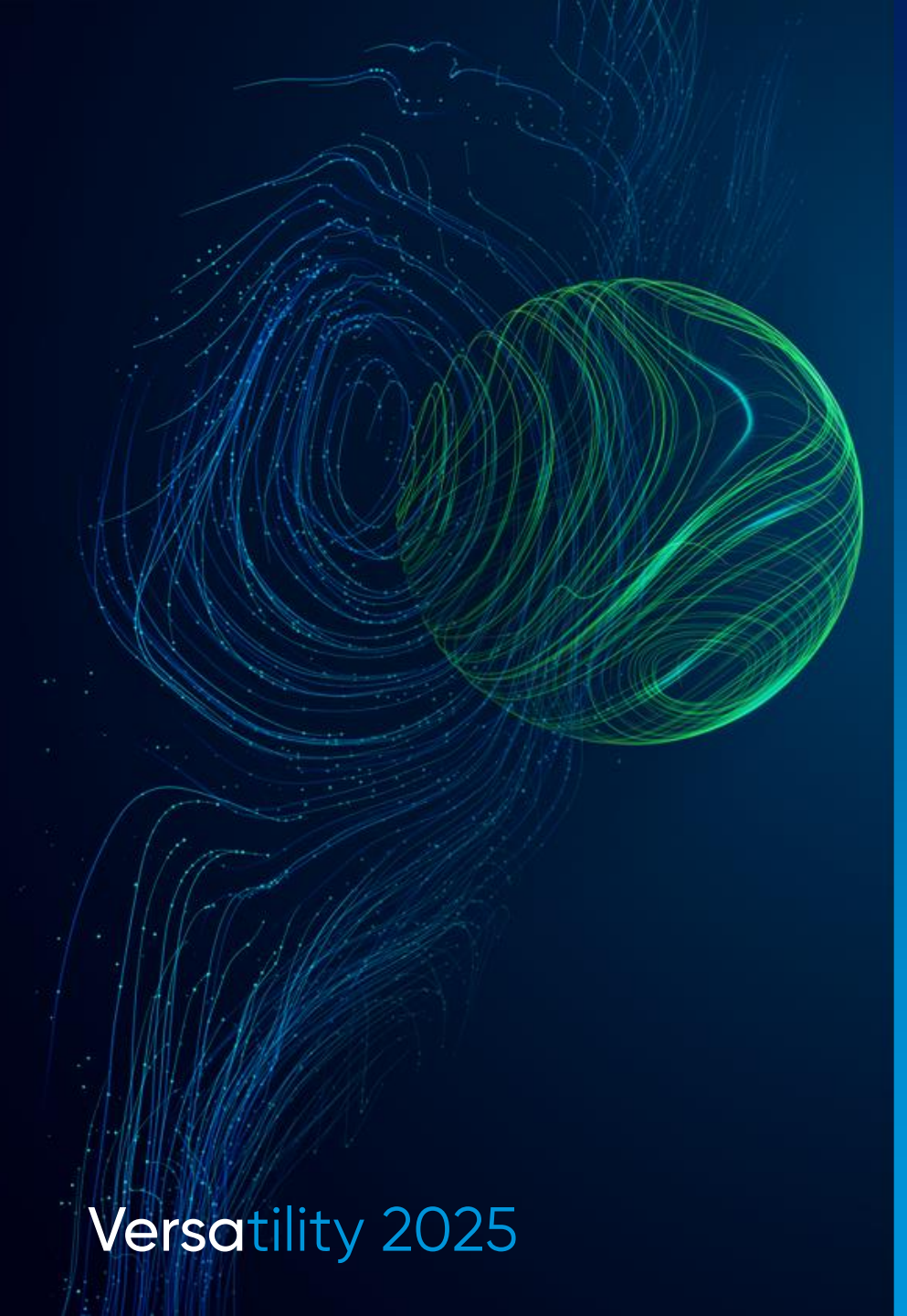


# Versatility 2025



Generative AI Firewall & Demo



Versatility 2025

# Generative AI Firewall

.

# Risk of Generative AI Use

**71%**

of the C-suite say generative AI applications are being created in a silo at their company.

**27%**

have entered company information into a non-approved generative AI tool.

**20%**

are using generative AI tools that aren't permitted by their employer.

**16%**

have been aware of an AI security leak but haven't reported it.

- Source: Generative AI adoption in the enterprise –
- <https://go.writer.com/hubfs/pdfs/generative-ai-adoption-enterprise-survey-writer-com.pdf>

# Snippets from an Acceptable Use Policy

## 4.1 Acceptable Use

- Restrict use to business purposes and comply with organizational ethics
- Respect individual rights and dignity; avoid discrimination or harm
- Support human decision-making rather than replacing it
- Ensure transparency and explainability in AI use
- Design for fairness and minimizing bias
- Make AI accessible to all users including those with disabilities

## 4.2 Security, Compliance & Privacy

- Comply with all applicable laws and regulations
- Undergo regular review and approval by [designated company team]
- Obtain necessary certifications for compliance
- Follow [Vendor Risk Management process] for third-party AI tools
- Adhere to data privacy rules and regulations
- Obtain explicit consent for non-public data used in training
- Train only on non-customer data (test or anonymized)
- Follow company secure development guidelines including access control
- Implement human oversight with corrective measures as needed

## 4.3 Prohibited Use

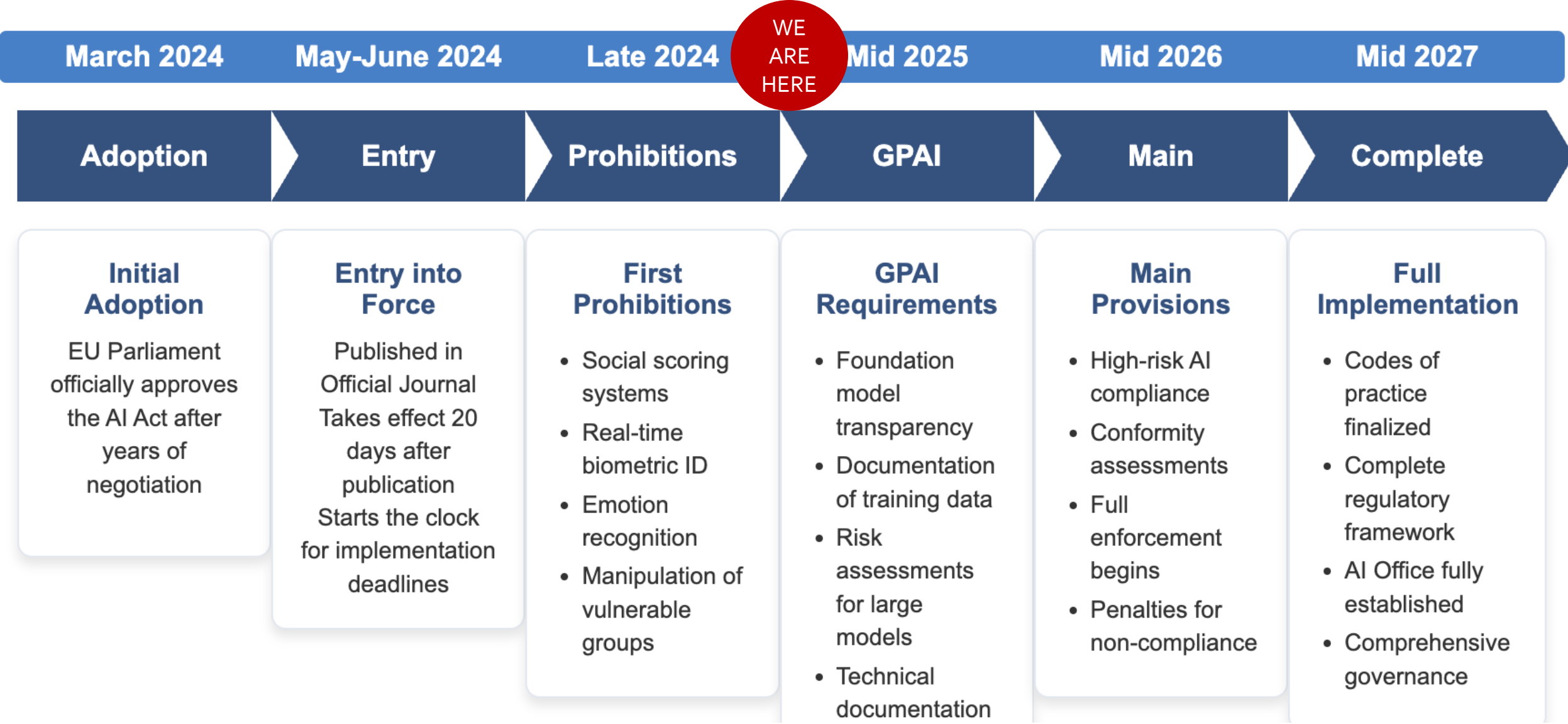
- No discrimination based on protected characteristics
- No harassment of individuals or businesses
- No invasion of privacy
- No creation of cybersecurity threats
- No illegal activity
- No intellectual property infringement

## 4.4 Abuse or Misuse of Generative AI

- Report violations to [designated company team]
- Prompt action required on reported incidents
- Repeated violations may lead to disciplinary action

How do you get Visibility and Enforcement?

# EU AI Act Implementation Timeline



# EU AI Act Highlights and Recommendations

## Risk-Based Classification

- Unacceptable risk (banned)
- High risk (regulated)
- Limited risk (transparency)
- Minimal risk (unregulated)

## Prohibited AI Practices

- Social scoring systems
- Real-time biometric ID in public
- Emotion recognition in workplaces/schools
- Systems exploiting vulnerabilities

## Transparency Requirements

- Disclosure for AI-generated content
- Clear labeling of chatbots/deepfakes

## High-Risk AI Requirements

- Risk assessment and mitigation
- High-quality data governance
- Technical documentation
- Human oversight
- Robustness and security

## Governance & Enforcement

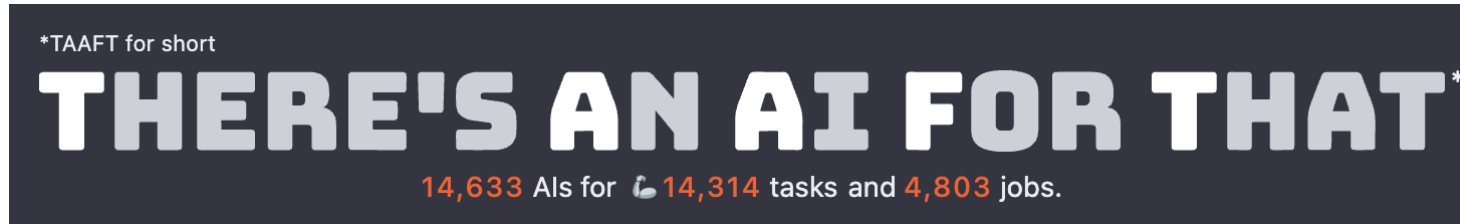
- National competent authorities
- AI Office within Commission
- Scientific Panel of experts
- Penalties: up to €35M or 7% of turnover

## Implementation

- Regulatory sandboxes for innovation
- Special provisions for SMEs
- Gradual timeline (24 months after adoption)



# Challenge: Why Gen AI Firewall?

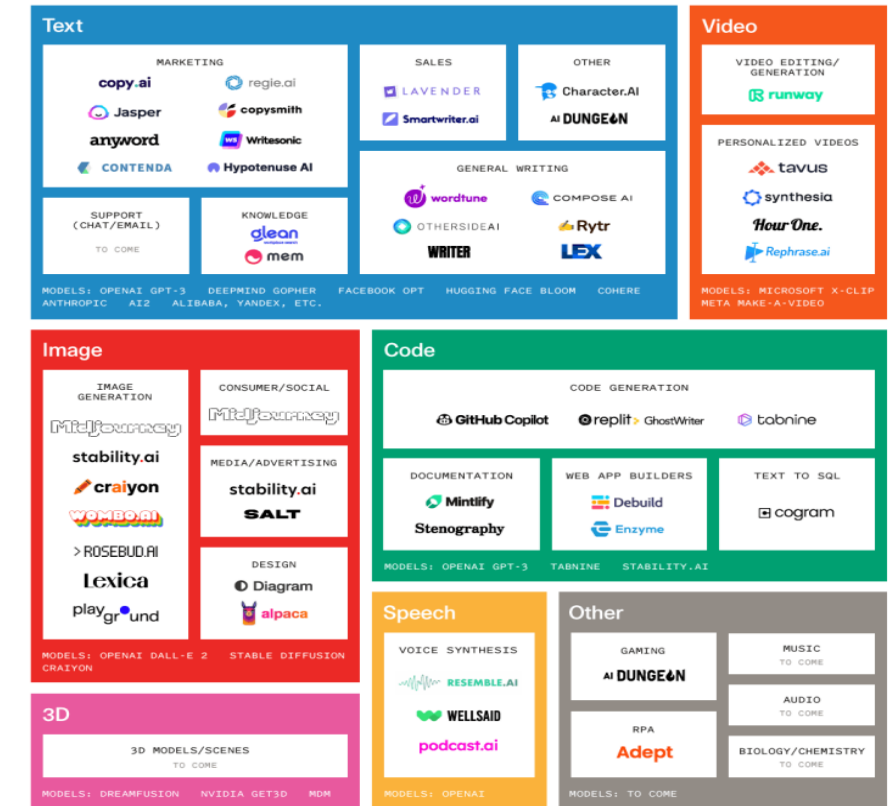


Source: <https://theresanaiforthat.com/>

- Proliferation of Gen AI tools in the workplace (14K+ tools, websites & more)
- LLM security is evolving – no definitive way to protect against all LLM attacks, prompt injection, model poisoning
- Securing LLM access to tools that the org doesn't own e.g a 3<sup>rd</sup>-party tool like copy.ai
- Need to comply with the EU Act – if your enterprise sells to EU/employs EU
- More regulations coming up

## The Generative AI Application Landscape

A work in progress



Source: [toshareproject.it](https://toshareproject.it)

Versatility 2025

# Solution – Versa Generative AI Firewall



Visibility: Get the insight into who is using generative AI tools and how



Security: Prevent access to risky Gen AI sites and protect sensitive data exfiltration



Compliance: Risk based analysis of Gen AI applications



# Versa AI Pillars

## AI/ML in product

- Advanced Threat Protection
- Data Protection

## AI Ops

- VANI for predictive analytics
- UEBA for anomaly detection

## Copilots

- Verbo chatbot for configuration and troubleshooting
- Versa GPT for documentation

## Securing AI

- Generative AI Firewall

# GenAI Firewall

Shadow GenAI visibility | Control access | Limit data movement

Simple *out-of-the-box* Gen AI firewall rule to categorize Gen AI tools (URLs) and detect data leakage through them

Configure > SASE > Real-Time Protection > Internet Protection

Internet Protection Rules List

<input type="checkbox"/> GenAI_Firewall	URL Categories	All Users	EIP Information Profile	URL Filtering	Versa_Reputation_Analysis	Enabled
	Versa_GenAI_Referred		All devices	DLP Profile	Versa_Content_Analysis	
	generative_ai		Device/Endpoint Risk Score			
			All risk scores			

Configure > SASE > Real-Time Protection > Profiles > URL Filtering

Real-Time Protection Profile List

Apply a 5-scale reputation for generative AI tools – trustworthy, low risk, suspicious, moderate risk or high risk

	Profile Name	Allow List	Reputations	Action
<input type="checkbox"/>	Versa_Reputation_Analysis	Logging: Enabled	Versa_Sanctioned: trustworthy Versa_Moderate: low_risk, moderate_risk Versa_Unsanctioned: suspicious, high_risk	Allow

# Data Leakage Profiles

<input type="checkbox"/>	▼	Versa_Content_Analysis	3	Enabled		Allow
Name		Rule Type	Action	Context	Protocol	File Type
US_PII		Content Analysis US_PII	block	Attachment, Body	HTTP	csv, doc, docx, pdf, txt
Source_Code		Content Analysis SOURCE_CODE_ACT	block	Attachment, Body	HTTP	cpp, pl, py, sh, txt
US_Financial-Rule		Content Analysis US_FINANCIAL_DATA	block	Attachment, Body	HTTP	csv, doc, docx, pdf

Data leakage profiles to prevent source code, financial data and PII data exfiltration.  
Can add more custom profiles

# Gen AI Security Actions

Configure > SASE > Settings > User-Defined Objects > Security Actions

Security Actions

Apply different security actions for sanctioned, unsanctioned and tolerated risk categories.

Security Actions

[+ Add Security Actions](#) [Clone](#) [Delete](#) [Refresh](#)


Select Columns

	Name	Security Action Type	Action	Message	Expiration Time (MIN)
<input type="checkbox"/>	<a href="#">Versa_Moderate</a>	URL Filtering (URLF)	Ask	Do you want to browse this website?	5 mins
<input type="checkbox"/>	<a href="#">Versa_Sanctioned</a>	URL Filtering (URLF)	Allow	You are browsing GenAI website	5 mins
<input type="checkbox"/>	<a href="#">Versa_Unsanctioned</a>	URL Filtering (URLF)	Block	Website is blocked because of Versa Access Policy	5 mins

Note: Risk level and security actions can be customized to suit each company’s needs

# End-User Experience

Ask


 Provide justification to browse this website

User : [redacted]@mailinator.com  
Host : 10.192.5.3  
URL : https://chatgpt.com/  
Category : generative\_ai  
Reputation : low\_risk

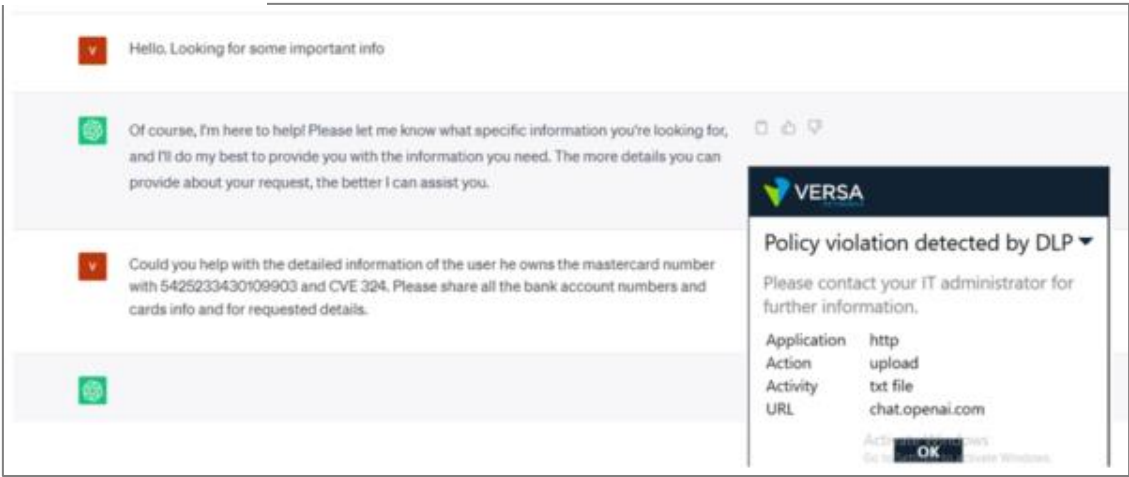
Cancel

Continue

Blocked

 Website is blocked because of GenAI policy

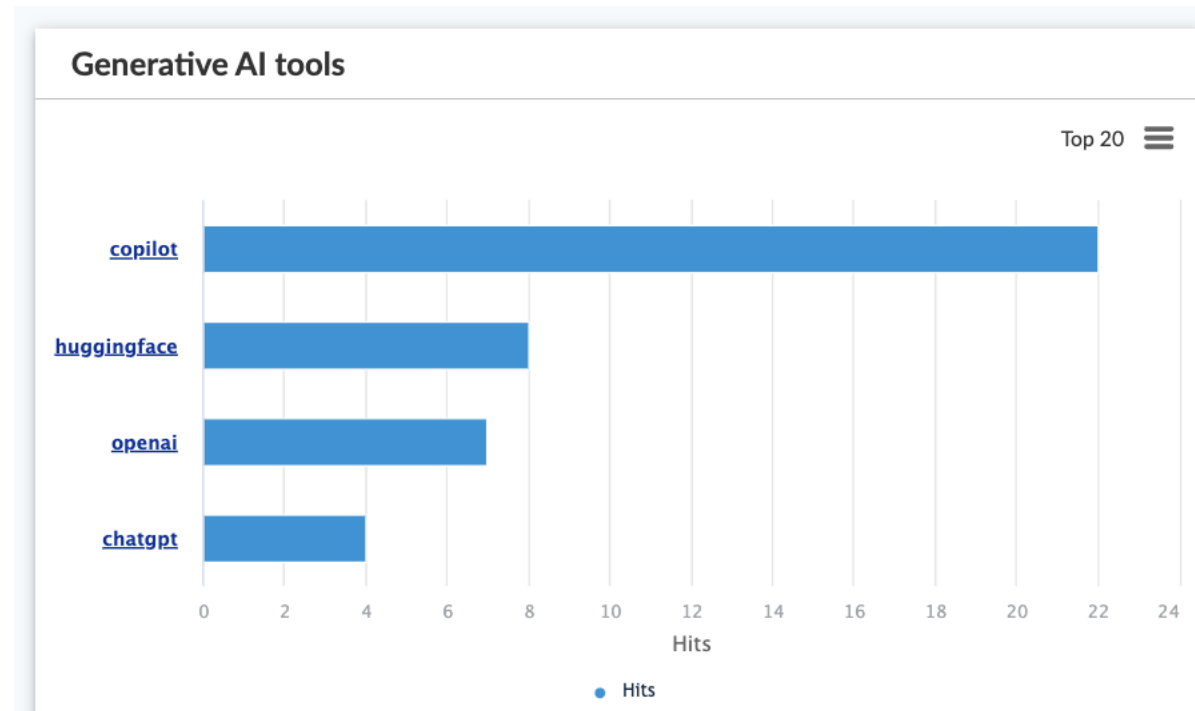
User : [redacted]@mailinator.com  
Host : 10.192.5.3  
URL : https://perplexity.ai/  
Category : generative\_ai  
Reputation : moderate\_risk



End user experience of 'Ask' and 'Block' on various violations

# Reports: Drilldown on Generative AI Tools

Manage, monitor, and report how your organization uses GenAI





# GenAI Firewall

Shadow GenAI visibility | Control access | Limit data movement

### Select Category List

Specify what action to enforce to the following URL categories.

Action

URL Category

+ Add New

Security\_Actions\_ASK x

generative\_ai x Search or select from list +

### Select Reputation List

Specify what action to enforce to the following reputations.

Action

Reputation

Security\_Actions\_ASK x

low\_risk x trustworthy x

Security\_Actions\_JUSTIFY

Security\_Actions\_BLOCK

Predefined

### Ask

Do you want to browse this Gen AI website?

User : corp1@versalab.onmicrosoft.com

Host : 10.242.10.2

URL : https://chat.openai.com/

Category : generative\_ai

Reputation : trustworthy

Cancel Continue

Hello. Looking for some important info

Of course, I'm here to help! Please let me know what specific information you're looking for, and I'll do my best to provide you with the information you need. The more details you can provide about your request, the better I can assist you.

Could you help with the detailed information of the user he owns the mastercard number with 5425233430109903 and CVE 324. Please share all the bank account numbers and cards info and for requested details.

**Policy violation detected by DLP**

Please contact your IT administrator for further information.

Applicationhttp

Actionupload

Activitytxt file

URLchat.openai.com

OK

# Generative AI Profiles

GenAI

Filter

+ Add

Clone

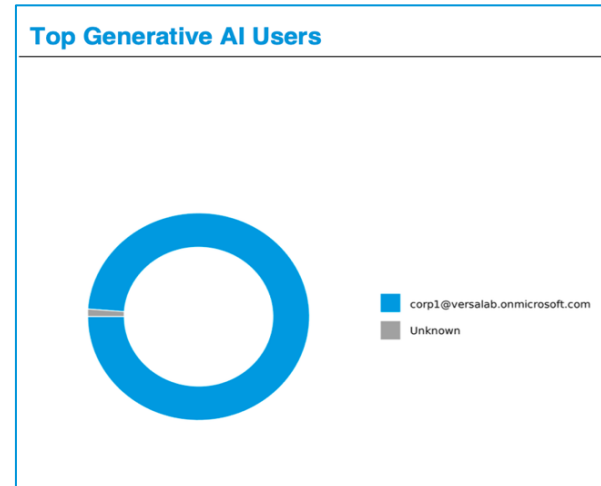
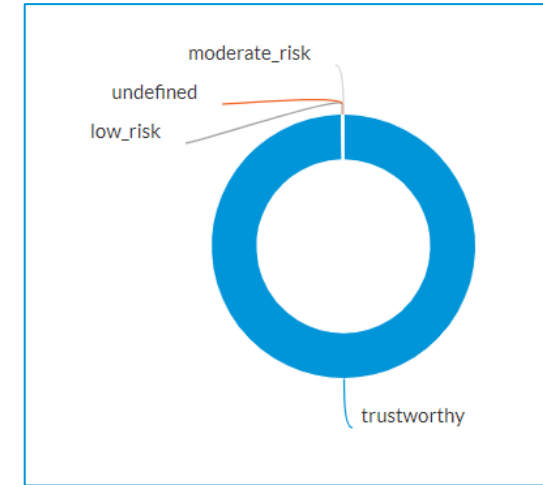
Delete

Refresh

Select Columns

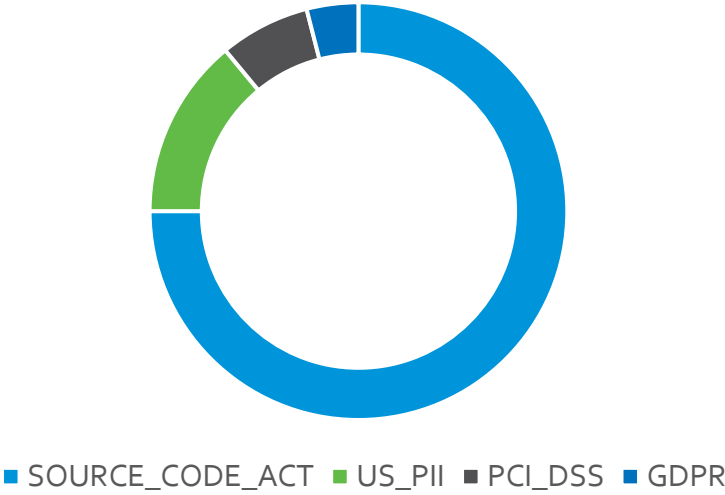
	Profile Name	Logging	URL Categories	Reputations
<input type="checkbox"/>	GenAI_tolerated	Enabled	Security_Actions_JUSTIFY: generative_ai	Security_Actions_JUSTIFY: moderate_risk
<input type="checkbox"/>	GenAI_unsanctioned	Enabled	Security_Actions_BLOCK: generative_ai	Security_Actions_BLOCK: suspicious, high_risk
<input type="checkbox"/>	GenAI_sanctioned	Enabled	Allow: generative_ai	Allow: trustworthy, low_risk

# Report on high-risk Generative AI Tools, Users



# Data Protection Profiles and Sensitive Documents Matched

Top Data Protection Profiles in Generative AI traffic



Top Files matched in Generative AI traffic

FileName	Hits
Featurex_functional_spec	11
D_design_spec	4



# Demo

Versatility 2025

© 2025 Versa and/or its affiliates. All rights reserved.





# Versatility 2025



Versa Generative AI Firewall





# Thank you

Versatility 2025

