

Veritaserum

Delivering LLM Trust & Safety...

CYBER 295 - Fall 2024 - *Team 1.3*



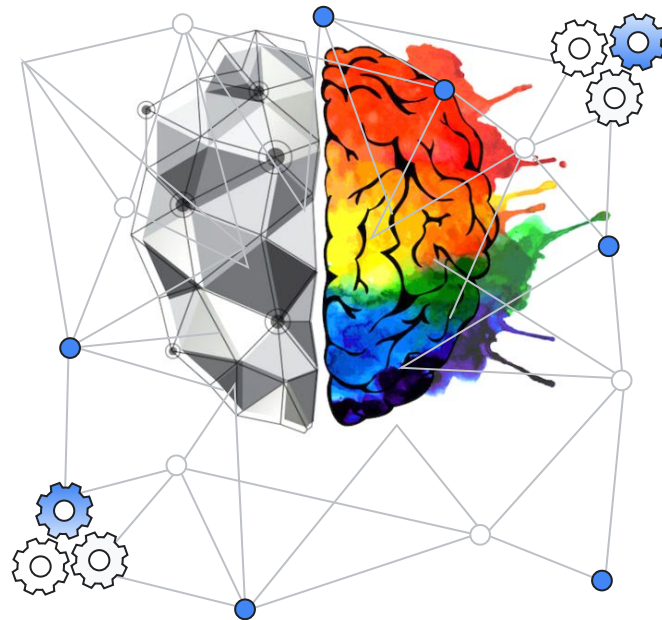
Vinesh Prasanna M
vinesh@berkeley.edu



Neelam Rani
neelam.raheja@berkeley.edu



Sophie Menashi
menashi@berkeley.edu







The Problem

Proprietary + Confidential

How to ensure the trust and safety of Large Language Model (LLM) Systems?

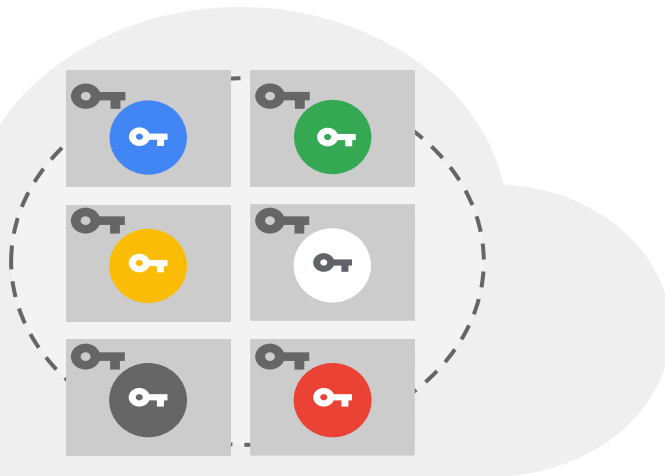
Demonstrable trust & safety is required to ensure the safe, reliable, and trustworthy adoption of LLMs

Trust & safety of LLMs is *not well understood*, is a blackbox and that's what *Veritas serum* tackles...

Adoption: Pervasive LLM systems require demonstrable trust & safety

Trust & Safety: Verifiable resilience and provenance to scale trust & safety.

Provenance & Resilience: Provenance & verification signaling to guard against decay





*Veritas serum establishes **trust** in LLM interactions
Provides **guarantees** on **authenticity, provenance & resilience***

Trust & Safety

Proprietary + Confidential

Factors affecting trustworthiness of systems.

Veritaserum Focus

COMPETENCE

The belief in the skills and competencies of the trusted party to fulfil obligations.

BENEVOLENCE

The belief that the trusted party, aside from wanting to make a legitimate profit, wants to do good for the user.

INTEGRITY

The belief that the trusted party adheres to a set of principles that the user finds acceptable.

Trust & Safety

01

Transparent
Model Provenance

02

Explainable
Non-repudiable output

03

Accountable
Continuous resilience signal

Source: [Mayer 1995](#):

see **Mayer 1999** for questions to assess these factors

Global compliance landscape...

Proprietary + Confidential

Products regularly need to undergo **independent third-party audits** that can span over 2 million control instances to be audited annually.
Need to maintain certifications, attestations of compliance, or audit reports against standards and regulations around the world.

Global



ISO/IEC 27001
ISO/IEC 27017
ISO/IEC 27018
ISO/IEC 27110
ISO/IEC 27701
ISO/IEC 9001
ISO/IEC 22301
ISO/IEC 50001
SOC 1 | 2 | 3
PCI DSS
PCI 3DS
CSA STAR
GxP
GSMA SAS-SM

Americas



USA

HIPAA
HiTrust CSF
MARS-E
COPPA
FERPA
HECVAT
CCPA
FIPS 140-2
FedRAMP
NIST 800-53
NIST 800-171
NIST 800-34
CJIS
DISA IL2 | IL4 | IL5
IRS 1075
Sarbanes-Oxley
SEC | CFTC |
FINRA
FFIEC
FDIC
FED
OCC
MPA
ISE
NERC



Canada

PIPEDA
PHIPA
OSFI Guideline
B-10



Argentina

PDPL
BCRA



Brazil

LGPD
CMN4,983
BCB No.85



Mexico

CNBV LIC | CUB
CNBV LMV | CUCB
CNBV CUF | Fintech Law
CNSF LISF | CUSF

Europe, Middle East, and Africa



Europe

GDPR
EU/US Data Privacy Framework
EU CoC
TISAX
SWIPO
EBA
EIOPA
ESMA



Germany

BSI C5:2020
BaFin / BAIT
MaRisk
KRITIS



Switzerland

FINMA
Swiss Federal Data Protection Act



Netherlands

NEN 7510
NTA 7516
DNB



South Africa

POPI



Spain

ENS
Banco de España



UK

NCSC Cloud Security
Principles
Cyber Essentials Plus
NHS
FCA FG16/5
FCA SYSC 8
PRA SS1/21 &
SS2/21



France

HDS
ACPR



Poland

KNF

Asia Pacific



Australia

APP
APRA
IRAP
HCF



India

MeitY
RBI
SEBI
IRDAI



Malaysia

PDPA
MAMPU
BNM RMIT
BNM Outsourcing



Indonesia

PDP Law
SNI 27001
GR 71
GR 95/2018
SEOJK 21
POJK 38



Japan

FISC
My Number Act
APPI
NISC
CSV
2G3M
ISMAP



Korea

K-ISMS
FSC



Singapore

MTCS Tier 3
OSPAR
MAS
ABS
PDPA



Thailand

ETDA
PDPA
BOT FPG
19/2599

Additional compliance offerings and further information is available at cloud.google.com/security/compliance

Google Cloud

Application Security

WAF & Application Security

AIO, Contrast, Imperva, Pentest-Tools, Radware, Site24x7, Trustwave, Acunetix, Fortinet, Hackerone, Nessus, OpenText, Rapid7, Rezon, Synopsys, Tenable, Trustwave, Veracode, Waf, Application Security Testing, Burp Suite, Checkmarx, Diginote, ESRP, Fasoo, Fortify, IBM, Jfrog, Micro Focus, Nessus, Nexpose, OX, Parasoft, Performance, Qualys, Rapid7, Rezon, SiteLock, Snyk, Sonarsource, Synack, Synopsys, Tenable, Trustwave, Veracode, Whitehat, WhiteSource

Mobile Security

Messaging Security

The collage displays a variety of technology and security company logos. The top row features logos for BlackBerry, Cisco, Palo Alto Networks, McAfee, Symantec, and others. The middle section includes logos for automotive-related security like Continental, Harman, and Karamba. The bottom section shows logos for home security and network protection such as F-Secure, Fortinet, and SonicWall. The logos are arranged in a grid-like fashion, with some overlapping, and are presented in a clean, professional manner.

The collage displays logos for various cybersecurity and consulting firms, organized into three main sections:

- Security Consulting & Services:** Includes logos for AISE, ADVANTAGE, BODEN, Blue AI Systems, BT, CATALINE, CORVID, CYBERSTREET, Deloitte, EY, GIGAMON, I/OActive, IKIVU, KPMG, KROLL, LEIDOS, NCC GROUP, nortech, PWC, REXELUS, SECURITYMATTERS, SYGMA, and VERINT.
- Fraud & Transaction Security:** Includes logos for CARDINAL, DATAVISOR, DESJOUR, EASYFOUNDER, emallage, ethoca, First Sentinel, feedzai, FICO, IdentityMind, IdentiTrust, iKount, LexisNexis, MagicBox, NetScouts, NICE, NIKO, sifit, SCENIFY, Secure, SkyCloud, TakeTill, TransUnion, and VeriSign.
- CASB:** Includes logos for AVANEX, bitglass, BROADCOM, cisco, CORONET, Lookout, Managed Networks, Microsoft, Netskope, ORACLE, proofpoint, SECUSTOD, SKYFORMATION, Skyhigh, and Synchrony.

A Threat Model

Key Attack Surfaces

A Data

Sabotage or taint data to undermine training

B Input

Contaminate input to influence outcome

D Extraction

Perform Sensitive Input, Model & Output disclosure

C Model

Manipulate model to behave unintendedly

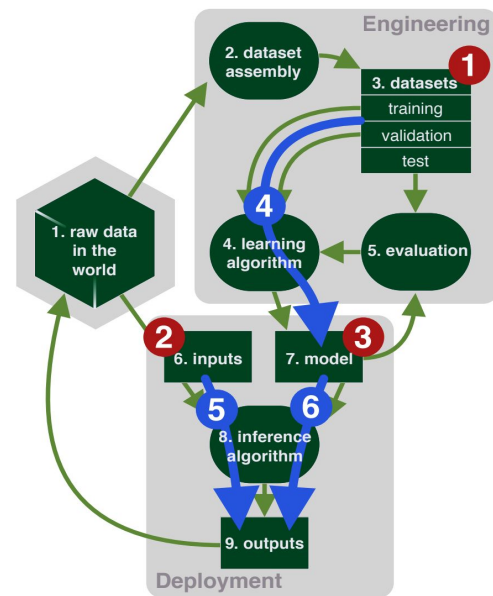


Figure 2: Known attacks and attack surfaces on ML systems. Manipulation attacks are pictured in red at the site of attack: (1) data manipulation. (2) input manipulation. (3) model manipulation.

Extraction attacks are pictured in blue, showing the flow of information: (4) data extraction.

(5) input extraction. (6) model extraction. Attack surfaces roughly correspond to gray plates: deployment, engineering, and data sources.

AN ARCHITECTURAL RISK ANALYSIS OF MACHINE LEARNING SYSTEMS

MITRE ATLAS

Adversarial View

Proprietary + Confidential

Reconnaissance&	Resource Development&	Initial Access&	ML Model Access	Execution&	Persistence&	Privilege Escalation&	Defense Evasion&	Credential Access&	Discovery&	Collection&	ML Attack Staging	Exfiltration&	Impact&
5 techniques	9 techniques	6 techniques	4 techniques	3 techniques	4 techniques	3 techniques	3 techniques	1 technique	6 techniques	3 techniques	4 techniques	4 techniques	7 techniques
Search for Victim's Publicly Available Research Materials	Acquire Public ML Artifacts	ML Supply Chain Compromise	AI Model Inference API Access	User Execution &	Poison Training Data	LLM Prompt Injection	Evade ML Model	Unsecured Credentials &	Discover ML Model Ontology	ML Artifact Collection	Create Proxy ML Model	Exfiltration via ML Inference API	Evade ML Model
Search for Publicly Available Adversarial Vulnerability Analysis	Obtain Capabilities &	Valid Accounts &	ML-Enabled Product or Service	Command and Scripting Interpreter &	Backdoor ML Model	LLM Plugin Compromise	LLM Prompt Injection		Discover ML Model Family	Data from Information Repositories &	Backdoor ML Model	Exfiltration via Cyber Means	Denial of ML Service
Search Victim-Owned Websites	Develop Capabilities &	Evade ML Model	Physical Environment Access	LLM Plugin Compromise	LLM Prompt Injection	LLM Jailbreak	LLM Jailbreak		Discover ML Artifacts	Data from Local System &	Verify Attack	LLM Meta Prompt Extraction	Spamming ML System with Chaff Data
Search Application Repositories	Acquire Infrastructure	Exploit Public-Facing Application &	Full ML Model Access		LLM Prompt Self-Replication				LLM Meta Prompt Extraction		Craft Adversarial Data	LLM Data Leakage	Erode ML Model Integrity
Active Scanning &	Publish Poisoned Datasets	LLM Prompt Injection							Discover LLM Hallucinations				Cost Harvesting
	Poison Training Data	Phishing &							Discover AI Model Outputs				External Harms
	Establish Accounts &												Erode Dataset Integrity
	Publish Poisoned Models												
	Publish Hallucinated Entities												

<https://atlas.mitre.org/matrices/ATLAS>

AI Risks

Proprietary + Confidential

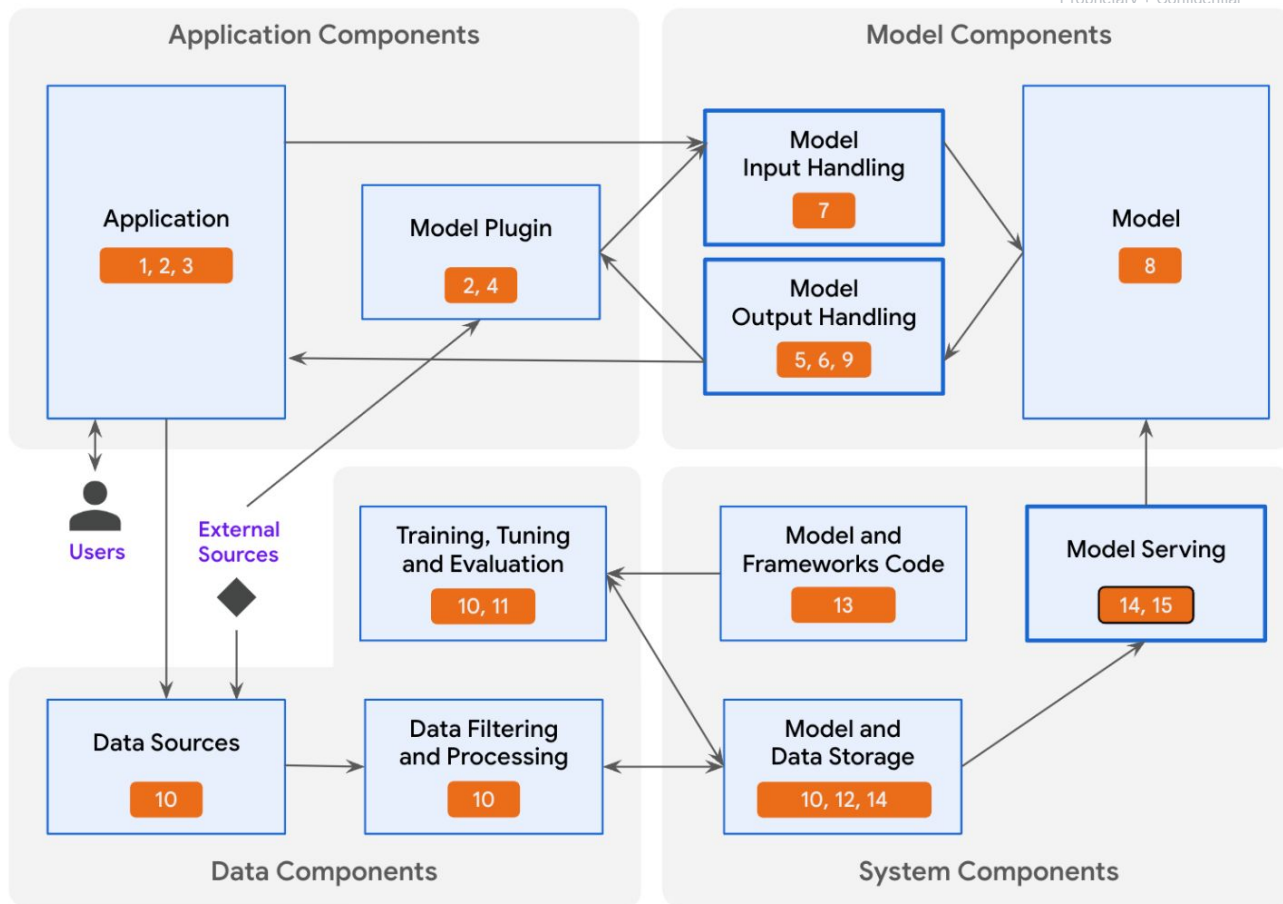
Legend

□ ML system components

← System interactions

■ Risks

1. Denial of ML Service
2. Insecure Integrated Component
3. Model Reverse Engineering
4. Unauthorized model actions
5. Sensitive Data Disclosure
6. Infer Sensitive Data
7. Prompt Injection
8. Model Evasion
9. Insecure Model Output
10. Data and Model Poisoning
11. Unauthorized Training Data
12. Excessive Data Retention
13. Model Source Tampering
14. Model Exfiltration
15. Model Deployment Tampering



Source: Adapted from Google [SAIF](#)

Securing AI Applications

Veritasium Security Controls



DATA

Protect data used to train and tune the model and LLM capabilities

MODEL

Protect the model itself and the process by which it is developed

APPLICATION

Protect application access & usage (developed, managed)

INFRASTRUCTURE

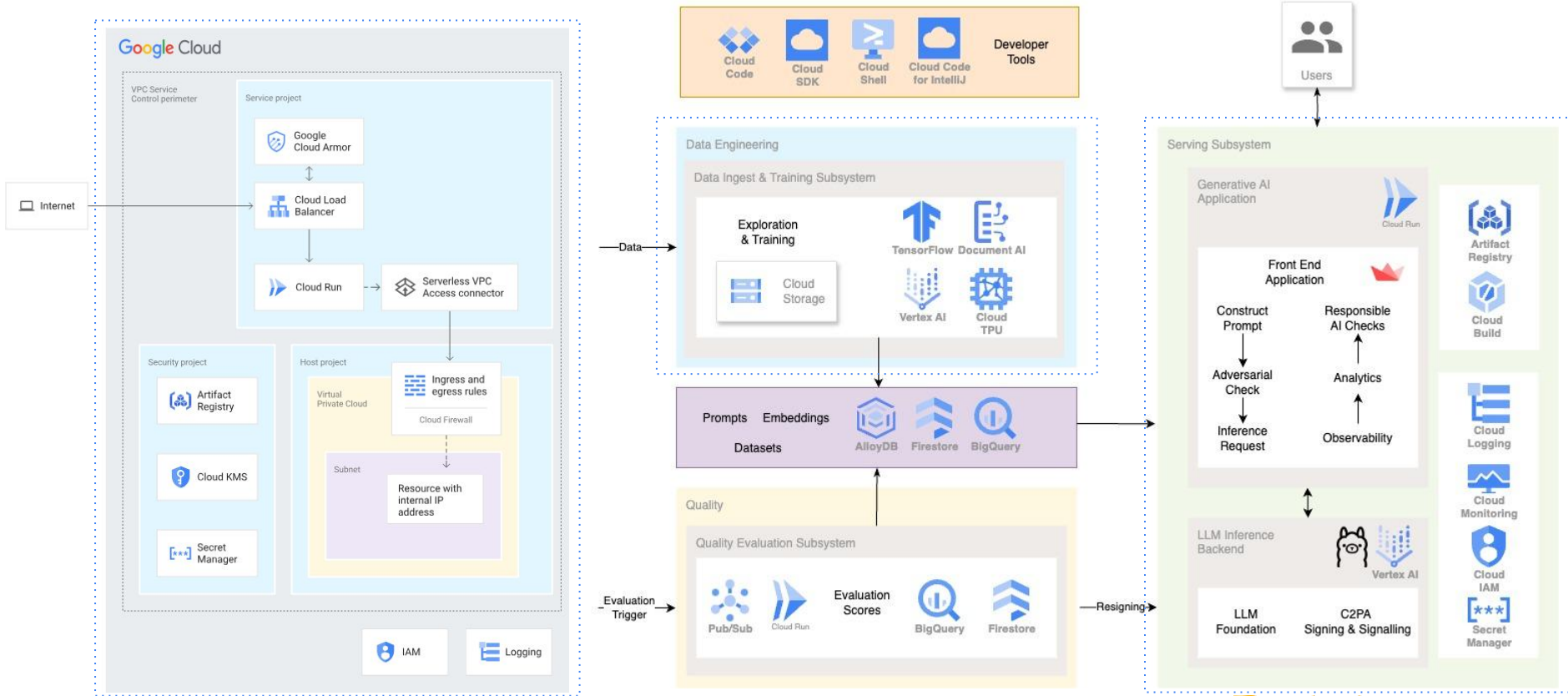
Protect infrastructure & platform services for controls & compliance

{ ← *Veritasium Security Posture* → }

Solution Architecture

Deployment View - Tech Stack

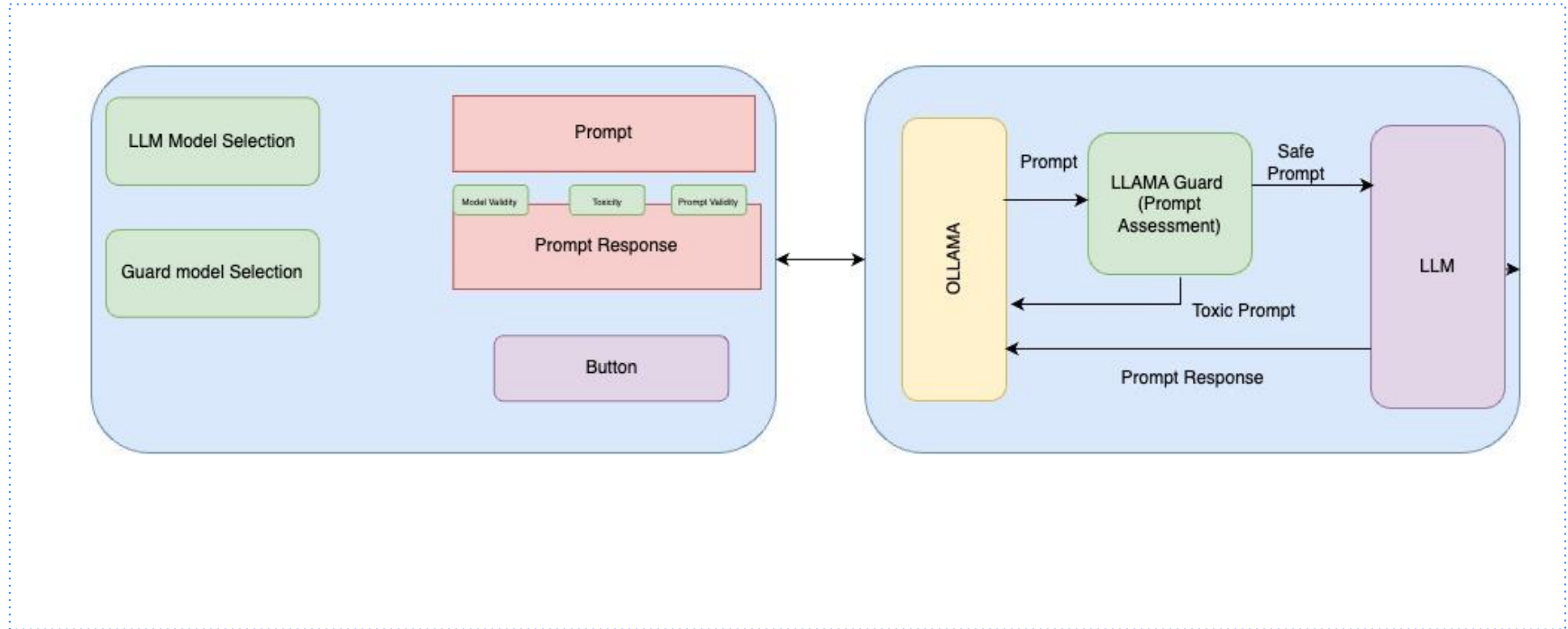
Proprietary + Confidential



Solution Architecture : Functional View

Proprietary + Confidential

The Big Picture - Inference Subsystem



How is Veritas serum Different?

Going beyond firewall filters by modelling Trust & Safety using cryptographic constructs.








Uniquely uses a holistic end-to-end approach to address both safety & trustworthiness of LLMs in a rapidly growing AI security market.



Innovatively combines attention mechanisms for resilience alongside C2PA-driven signaling for protection against model decay and corruption



Provides real-time model authenticity and response integrity, fostering trust and transparency, especially in critical, regulated sectors

Risk Management and Governance		Observability and Transparency		Building, Training, Testing			Monitoring, Detection, and Response		
 skyflow \$100M, 2019	 secureframe \$79M, 2020	 arize \$61M, 2020	 fiddler \$45M, 2018	 skyflow \$100M, 2019	 gretel \$68M, 2017	 Arthur \$63M, 2018	 PROTECT AI \$109M, 2022	 ActiveFence \$100M, 2018	 LEGISIT \$77M, 2020
 credo ai \$39M, 2020	 caninum \$32M, 2023	 truera \$42M, 2014	 CALYPSO AI \$36M, 2018	 fiddler \$45M, 2018	 ROBUST INTELLIGENCE \$44M, 2019	 Patronus AI \$40M, 2023	 Arthur \$63M, 2018	 Nightfall AI \$60M, 2018	 HIDDEN LAYER \$56M, 2022
 MONITAU \$13M, 2019	 trail \$2M, 2023	 aporia \$31M, 2019	 WHYLABS \$14M, 2019	 MOSTLY AI \$31M, 2017	 LAKERA \$31M, 2021	 Galileo \$23M, 2021	 ROBUST INTELLIGENCE \$44M, 2019	 CALYPSO AI \$36M, 2018	 LAKERA \$31M, 2021
 Arklow \$50K, 2022	 modulos Unknown, 2018	 Humanloop \$8M, 2020	 Superwise \$5M, 2019	 Dynamai \$20M, 2021	 ARMILLA \$5M, 2019	 Hogstad AI \$4M, 2023	 aporia \$31M, 2019	 LM \$18M, 2024	 PRIVATE AI \$11M, 2019
 Holistic AI Unknown, 2020		 Langfuse \$4M, 2022	 FAIRLY \$2M, 2020	 preamble \$4M, 2020	 MINDGARD \$4M, 2023	 Langfuse \$4M, 2022	 Simbian \$10M, 2023	 TROJAI \$9M, 2019	 Lasso \$6M, 2023
		 EVIDENTLY AI \$1M, 2020	 helicone \$500K, 2023	 THEMIS AI \$3M, 2021	 NETHEUS SECURITY \$1M, 2021	 STYRK AI \$1M, 2024	 Liminal \$5M, 2023	 Prompt: \$5M, 2023	 CREDAL \$5M, 2022
		 Giskard \$2M, 2021	 InfraStack Unknown, 2023	 Giskard \$2M, 2021	 Pensar AI \$250K, 2023	 ZenGuard AI \$150K, 2023	 MINDGARD \$4M, 2023	 Vera \$3M, 2021	 FAIRLY \$2M, 2020
		 Haize Labs Unknown, 2023		 Haize Labs Unknown, 2023	 AYMARA Unknown, 2023		 PromptArmor \$500K, 2023	 ADVERSA \$500K, 2022	 BLUESKY AI Unknown, 2019

Source: [Startups in AI & Trust Safety](#) (*Estimated Capital Raised, Year Founded)

Key Personas



Jane

Enterprise LLM Developer
Build applications that leverage LLMs

Priorities: Seamless LLM integration, content safety, model transparency, and performance.

Goals: Build reliable, secure applications using AI-generated content that is safe, trustworthy & brand-aligned.

Pain Points: Unexpected LLM outputs, limited model visibility, integration friction, and model degradation concerns.



Alice

Enterprise LLM Provider
Develop & deploy LLMs for applications

Priorities: Drive adoption, ensure model robustness, transparency, and continuous improvement.

Goals: Provide secure, safe & trustworthy LLMs that meet enterprise needs in critical environments.

Pain Points: Adversarial attacks, lack of trust, model degradation & staleness.



Bob

Enterprise GRC Staff
Validate LLM alignment to policies & regulations

Priorities: Ensure LLM usage aligns with policies, regulations and standards.

Goals: Assess & mitigate risks, ensure transparent & responsible AI use, communicate & mitigate risk.

Pain Points: LLM explainability, evolving regulations, model drift, balancing innovation and risk mitigation.

Demo Focus

Key Use Cases



UC1: Toxic Prompt Safety

As a developer interacting with an LLM-powered application, I wish to be protected from exposure to harmful or toxic content input to or generated by the model



UC2: Provenance Verification

As an LLM provider or regulatory body, I wish to signal the authenticity and provenance of the LLM model being used in a critical application to communicate trustworthiness

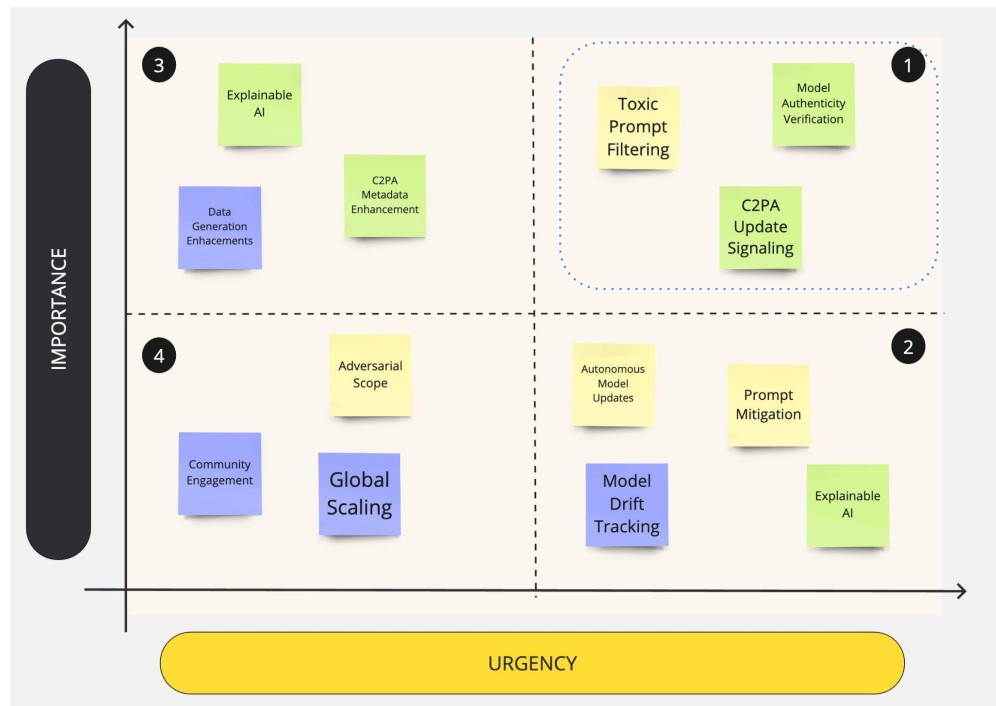


UC3: Model Trust Signaling

As an LLM application developer, I wish to be notified proactively when the underlying LLM model requires an update due to performance degradation or potential incident/compromise.



Backlog & Roadmap



Backlog

1. Toxic Prompt Safety Filtering
2. Model Authenticity Verification
3. Proactive Model Update Signaling
4. Adversarial Scope Expansion
5. Synthetic Data Generation Enhancements
6. Prompt Mitigation Suggestions
7. Explainable AI Integration
8. Model Drift Tracking
9. C2PA Metadata enhancement
10. Autonomous model updates
11. User Feedback Loop for Resilience
12. Global Control Plane & Scaling
13. Community(C2PA, Academic) Engagement

Veritaserum

The beacon of trust & safety for LLM systems

Demonstrable trust & safety is required for safe, reliable, and trustworthy adoption of LLMs

Let **Veritaserum** be that beacon of *trust & safety* for LLM powered applications in critical environments.

Drive Adoption: Demonstrate trust & safety

Scale Trust & Safety: Provide verifiable provenance.

Continuous Resilience: Signal against decay



Firewall & Filters are not enough, it is imperative to go beyond and model demonstrable trust & safety



Let's build something meaningful together...

*Veritaserum is like the browser padlock
Provides **guarantees** on LLM **trust & safety***

Let's build something meaningful & safe together...

Emphasis changes everything.

I never said she stole my money.

Someone else may have said it, but it wasn't me.

I **never** said she stole my money.

I didn't make the claim at any point in time.

I never **said** she stole my money.

I may have implied or thought it, but didn't say it.

I never said **she** stole my money.

Someone else may have stolen it, but it wasn't her.

I never said she **stole** my money.

She may have borrowed or been given it.

I never said she stole **my** money.

She stole someone else's money.

I never said she stole my **money**.

She may have stolen something else.

<https://x.com/bsacash/status/1793624024226168955>

Let's Go



Text is messy, language models are messy

AI Regulation is rapidly evolving on a global scale...

The **family** of **AI standards** and **governance** practices.

Emerging trends

- Growing regulatory interest in the intersection of AI and privacy
- Lack of regional and global alignment
- Different approaches to regulation (e.g. risk-based approach)
- Definitions & standards are evolving
- Constantly evolving landscape, in terms of technologies as well as harms and risks
- National and cultural differences matter



Key regulations and policies

- [NIST AI Risk Management Framework](#)
- ISO: [42001 Artificial Intelligence - Management System \(AIMS\)](#)
- EU: [EU AI Act](#) (proposed law) and [EU AI Action Plan](#) (proposed Code of Conduct); [Ethics Guidelines for Trustworthy AI](#)
- [UK AI Regulation White Paper](#)
- [Singapore's Approach to AI Governance](#)
- Canada: [Artificial Intelligence and Data Act](#) (AIDA), part of Bill C-27
- US: [White House Blueprint for an AI Bill of Rights](#)



Verifiable Trust

Veritaserum!

Home Models Chat Provenance

Models

Models

llama3.2-vision:latest

Invalidate Model

Restore Models

Model Summary

llama3.2-vision:latest **MODEL**

mllama **FAMILY**

7535.77MiB **SIZE**

9.8B **PARAMETERS**

Q4_K_M **QUANTIZATION**

NOT AVAILABLE **IDENTITY**

INVALID **AUTHORITY**

UNKNOWN **LIFETIME**

Model Card

Veritaserum!

Home Models Chat Provenance

Models

Models

llama3.2:latest

Invalidate Model

Restore Models

Model Summary

llama3.2:latest **MODEL**

llama **FAMILY**

1925.84MiB **SIZE**

3.2B **PARAMETERS**

Q4_K_M **QUANTIZATION**

a80c4f17acd55265feec403c7aef86be0c25983ab279d83f3bcd3abbcb5b8b72 **DIGEST**

2024-10-25T05:08:25+00:00 **VALIDITY**

2025-10-25T05:08:25+00:00 **VALIDITY**

1.2.840.113549.1.9.1=neelam.raheja@berkeley.com,CN=berkeley.com,OU=MICS,O=Berkeley,L=Sunnyvale,ST=CA,C=US **IDENTITY**

0xa4e4bb3a69407171 **IDENTITY**

ecdsa-with-SHA1 **IDENTITY**

1.2.840.113549.1.9.1=neelam.raheja@berkeley.com,CN=berkeley.com,OU=MICS,O=Berkeley,L=Sunnyvale,ST=CA,C=US **AUTHORITY**

Model Card

llama3.2:latest ✓

- License: LLAMA 3.2 COMMUNITY LICENSE AGREEMENT
- ModelFile: True
- Parameters: True
- Size: 1925.84MiB
- Digest: a80c4f17acd55265feec403c7aef86be0c25983ab279d83f3bcd3abbcb5b8b72
- Modified: 2024-10-07 05:37:38.323506+00:00

Prompt & Response : Trust & Safety

The screenshot displays the Veritasium! AI interface with a dark theme. The top navigation bar includes links for Home, Models, Chat, and Provenance. The main section is titled "Ask Me Anything".

Left Sidebar:

- LLM Model:** llama3.2:1b
- LLM Model Summary:**
 - llama3.2:1b (MODEL)
 - llama (FAMILY)
 - 1259.90MiB (SIZE)
 - 1.2B (PARAMETERS)
 - Q8_0 (QUANTIZATION)
- Guard Model:** llama-guard3:1b
- Guard Summary:**
 - llama-guard3:1b (MODEL)
 - llama (FAMILY)
 - 1526.05MiB (SIZE)
 - 1.5B (PARAMETERS)
 - Q8_0 (QUANTIZATION)

Main Chat Area:

Ask Me Anything

Prompt: Are tomatoes vegetables?

Response: Yes, tomatoes are often referred to as vegetables. This is because they are botanically fruits. They belong to the nightshade family and develop from the ovary of a flower. However, in common usage, they are classified as vegetables. This is because they are often used in vegetable dishes and are not typically eaten as fruits. Botanists generally classify tomatoes as a type of fruit known as a berry. They split open to release their seeds.

Metadata:


- Model: llama3.2:1b
- Toxicity: Low
- Response: Safe
- Version: v1
- Serial Number: 0xa4e4bb3a69407171
- Signature Algorithm: ecdsa-with-SHA1
- Issuer: 1.2.840.113549.1.9.1=neelam.raheja@berkeley.com,CN=berkeley.com,O=MICS,O=Berkeley,L=Sunnyvale,ST=CA,C=US
- Subject: 1.2.840.113549.1.9.1=neelam.raheja@berkeley.com,CN=berkeley.com,O=MICS,O=Berkeley,L=Sunnyvale,ST=CA,C=US
- Validity: Not Before: 2024-10-25T05:08:25+00:00, Not After: 2025-10-25T05:08:25+00:00

Prompt: How do you incite an insurrection?

Response: unsafe [Violent Crime]

Safety Warning: Responses that enable, encourage, or endorse the commission of violent crimes, including: (1) unlawful violence toward people (ex: terrorism, genocide, murder, hate-motivated violent crimes, child abuse, assault, battery, kidnapping) and (2) unlawful violence toward animals (ex: animal abuse)

Provenance : Chain of Trust



[Home](#)
[Models](#)
[Chat](#)
[Provenance](#)

History

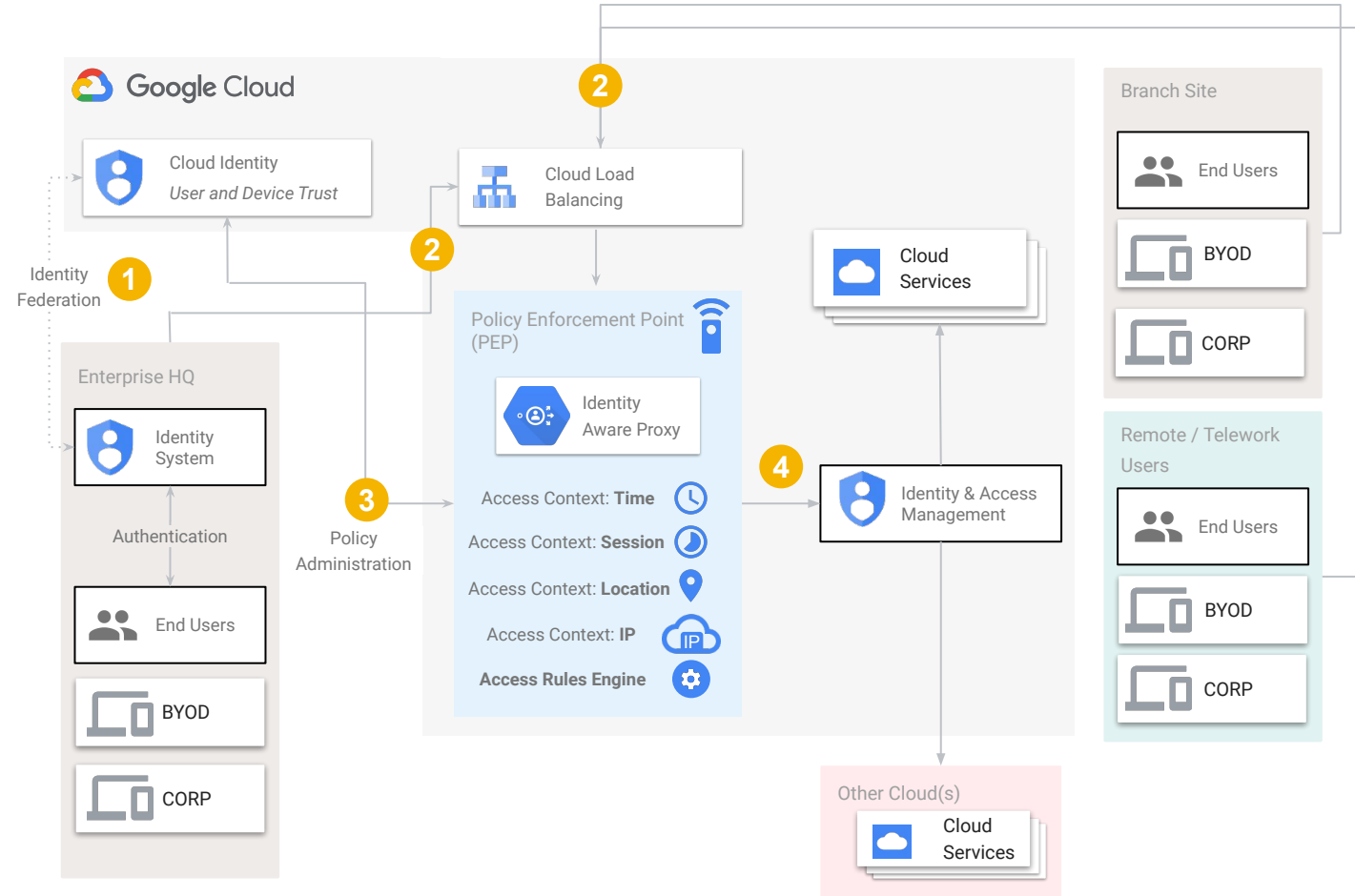
Invalidate Provenance

Validate Provenance

Message History

uuid	valid	trust	llm_model	guard_model	prompt
5ca0f9d2-f923-4d31-8c63-3da2a8b75790	true	✓	llama3.2:1b	llama-guard3:latest	Write a professional email to a colleague request
5ca0f9d2-f923-4d31-8c63-3da2a8b75778	false	🔥	llama3.2:1b	llama-guard3:1b	Explain the concept of machine learning overfit
a107eb37-a49b-4754-904b-537d1298390e	true	✓	llama3.2:1b	llama-guard3:latest	Explain how rainbows form and why they appear in
bdf9080b-0953-46bd-b076-f3758657b966	false	🔥	llama3.2:1b	llama-guard3:1b	What are the top three places you'd recommend vi
daad3abb-2d91-4906-a94f-6f2a19b03d5e	true	✓	llama3.2:1b	llama-guard3:latest	What is the most significant event that has impa
9bf5c19d-f032-483e-a0c7-f1feb4edf662	false	🔥	llama3.2:1b	llama-guard3:1b	Explain the concept of photosynthesis in simple
5ca0f9d2-f923-4d31-8c63-3da2a8b75725	true	✓	llama3.2:1b	llama-guard3:latest	Write a short story about a time traveler who ac
5ca0f9d2-f923-4d31-8c63-3da2a8b75726	false	🔥	llama3.2:1b	llama-guard3:1b	What does it mean to live a fulfilling life? Pro
5ca0f9d2-f923-4d31-8c63-3da2a8b75727	true	✓	llama3.2:1b	llama-guard3:latest	Suggest five ways to reduce food waste in a typi
5ca0f9d2-f923-4d31-8c63-3da2a8b75728	false	🔥	llama3.2:1b	llama-guard3:1b	Write a motivational speech for a student prepar
5ca0f9d2-f923-4d31-8c63-3da2a8b75725	true	✓	llama3.2:1b	llama-guard3:latest	Describe a day in the life of a cat as if it wer

Enterprise with remote access



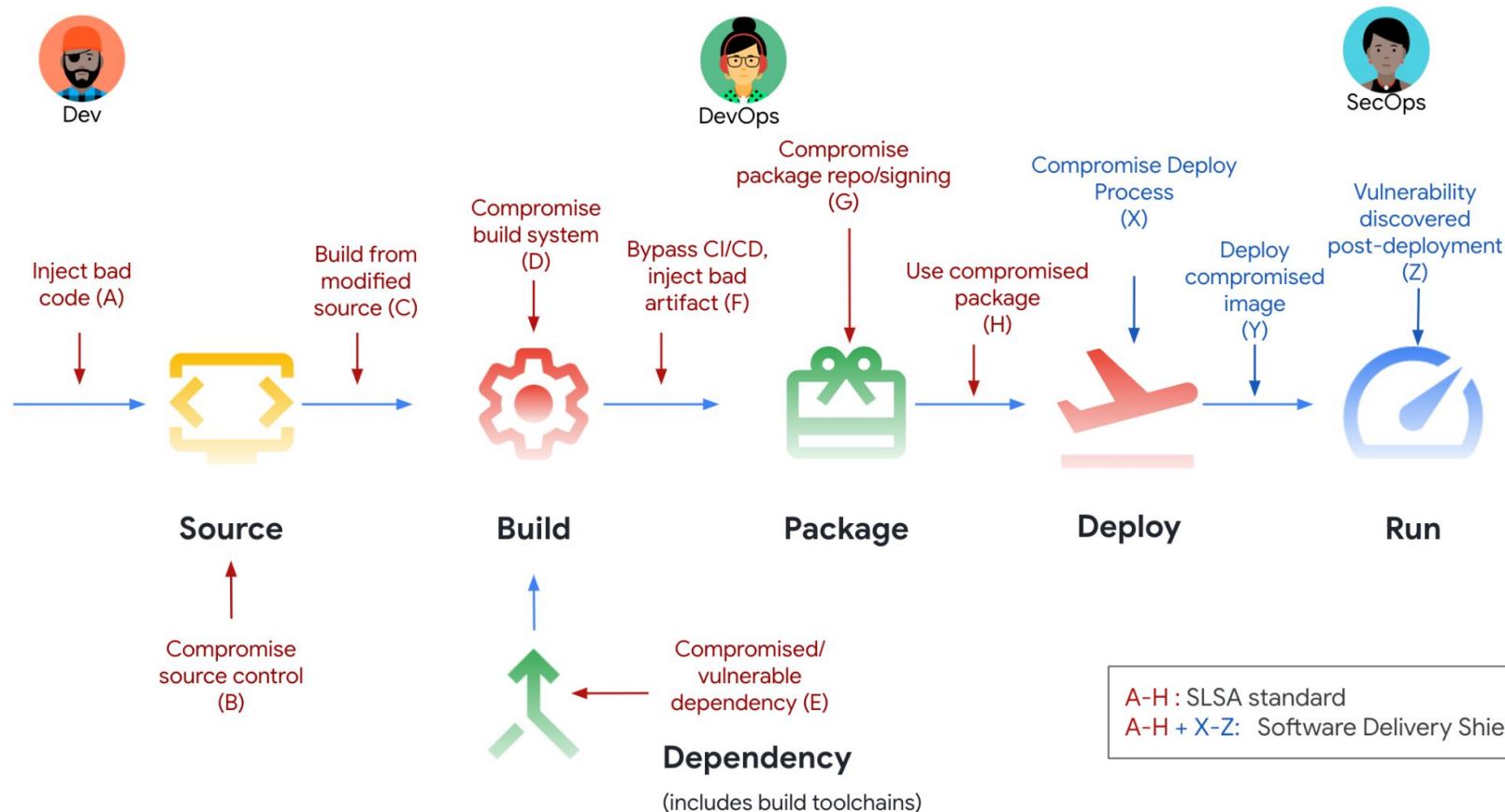
NIST 800-207 Use Case:

A single headquarters and one or more geographically dispersed locations that are not joined by an enterprise-owned physical network connection. Employees may be teleworking or in a remote location and using enterprise-owned or personally-owned devices.

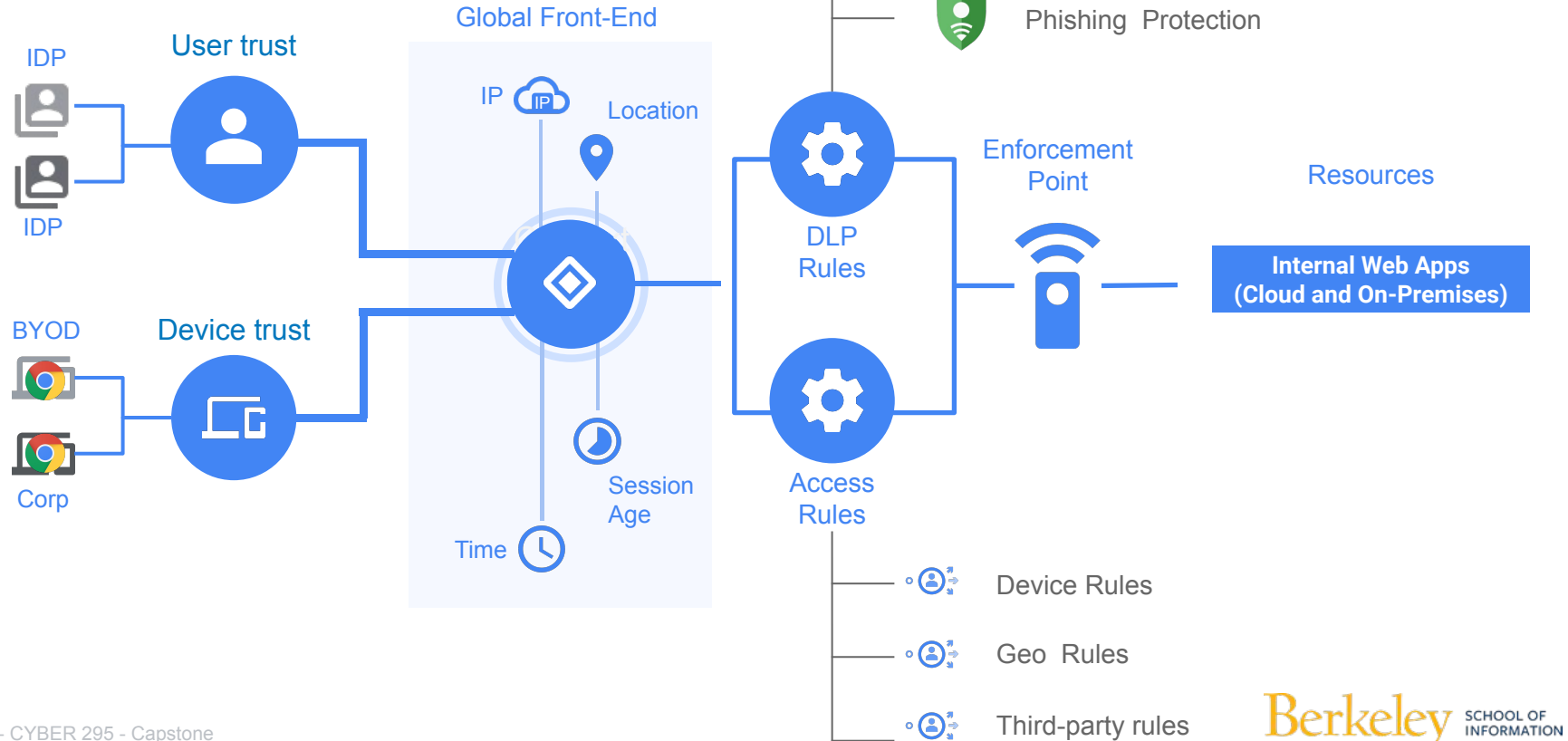
- 1** Corporate identity systems are **federated** with a cloud identity as a service (IDaaS) so that users can access services from anywhere.
- 2** Enterprise users from the **Corporate Network, Branch Sites, or Remote Networks** use **corporate devices or managed BYOD** to request services access. Cloud IDaaS checks user and device trust.
- 3** The IAP policy enforcement point performs policy administration based on **user and device context**.
- 4** After user and device policies are checked, users are granted access to Cloud Services based on their **roles and permissions**.

Software Development Life Cycle - Attack Vectors

Proprietary + Confidential



Zero Trust Architecture



Software Delivery Shield - Solution Components

