Taming without Maiming
AI Governance for Data Infrastructure

Privacy and Security Academy
November 8–10

Speakers

Who are these guys anyway?

Bill Schaumann

Bill is an independent privacy consultant with twenty plus years of experience developing privacy and security systems and programs for a variety of fortune 100 clients in the financial services, healthcare, manufacturing, government, and insurance sectors. Bill has a deep understanding of the processes and related technologies needed to meet todays complex universe of regulatory requirements.

Priyadarshi "PD" Prasad

PD is the co-founder and chief product officer at LightBeam.ai, the pioneers in data security and privacy automation. PD is an experienced tech industry professional who studies the intersection of artificial intelligence and privacy. He is always keen to understand the interesting ways organizations secure and protect their customers' data, and looks for opportunities to replace complex tech stacks with 1-click simple solutions.

Shifting Landscape of Data Infrastructure

Data volumes & types growing

Structured and unstructured data continuously growing

+

Number of insertion points increasing

DBs, SaaS products, Kafka Streams, etc.

+

HIPAA COMPLIANCE GDPR PCI

Data Lakes are opaque

Data Provenance?
Data Identity??

AI processing enables new insights and automation

AI the promise. . . .

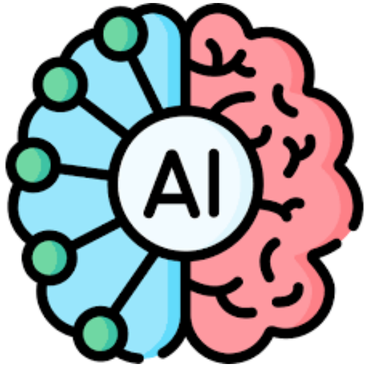
- …Potential to make the world a better place by solving a variety of pressing problems that humanity faces today.

Autonomous Vehicles

Security

Process Automation

HealthCare



Education

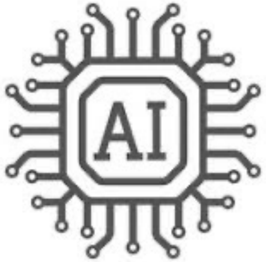
Privacy

Poverty

Finance

123H456E789L
Pattern Recognition
10112L131415O

Anomaly Detection

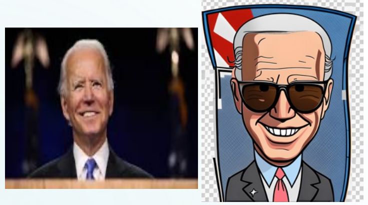


AI the worry.....

- However, there are risks associated with AI and many bad scenarios have been envisioned.

Ethical Dilemmas

Deepfakes and



Misinformation

Privacy Violations

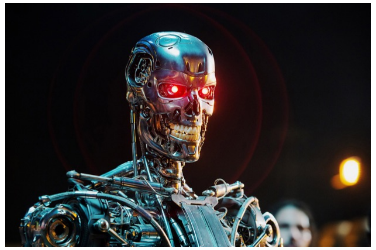
Addiction and Behavioral Manipulation

Job Displacement

Bias and Discrimination

Social Manipulation

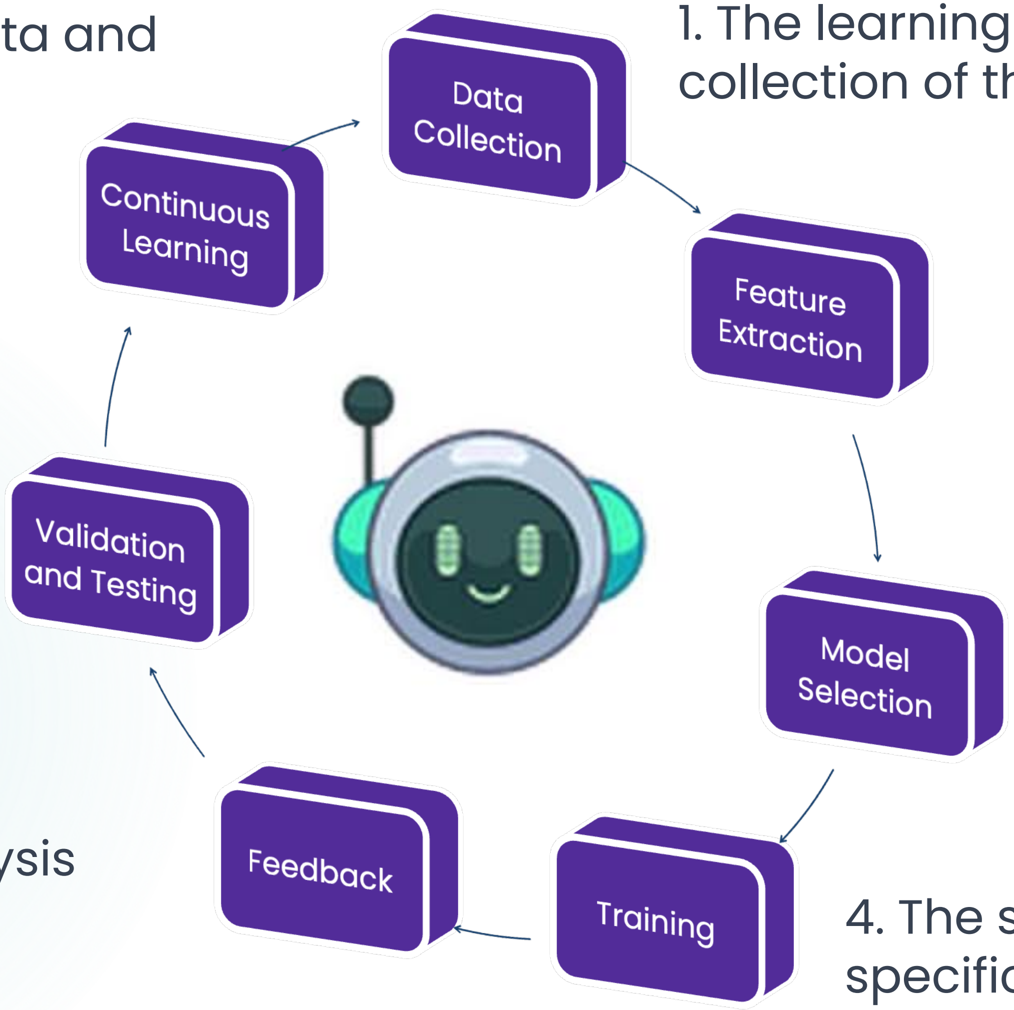
Healthcare Errors



Cybersecurity Threats

Autonomous Weapons

How do AI systems learn?



Data Collection

Feature Extraction

Model Selection

Training

Feedback

Validation and Testing

Continuous Learning

1. The learning process begins with the collection of the raw data

2. Relevant features & data attributes are identified.

3. A machine learning model is chosen based on the type of task

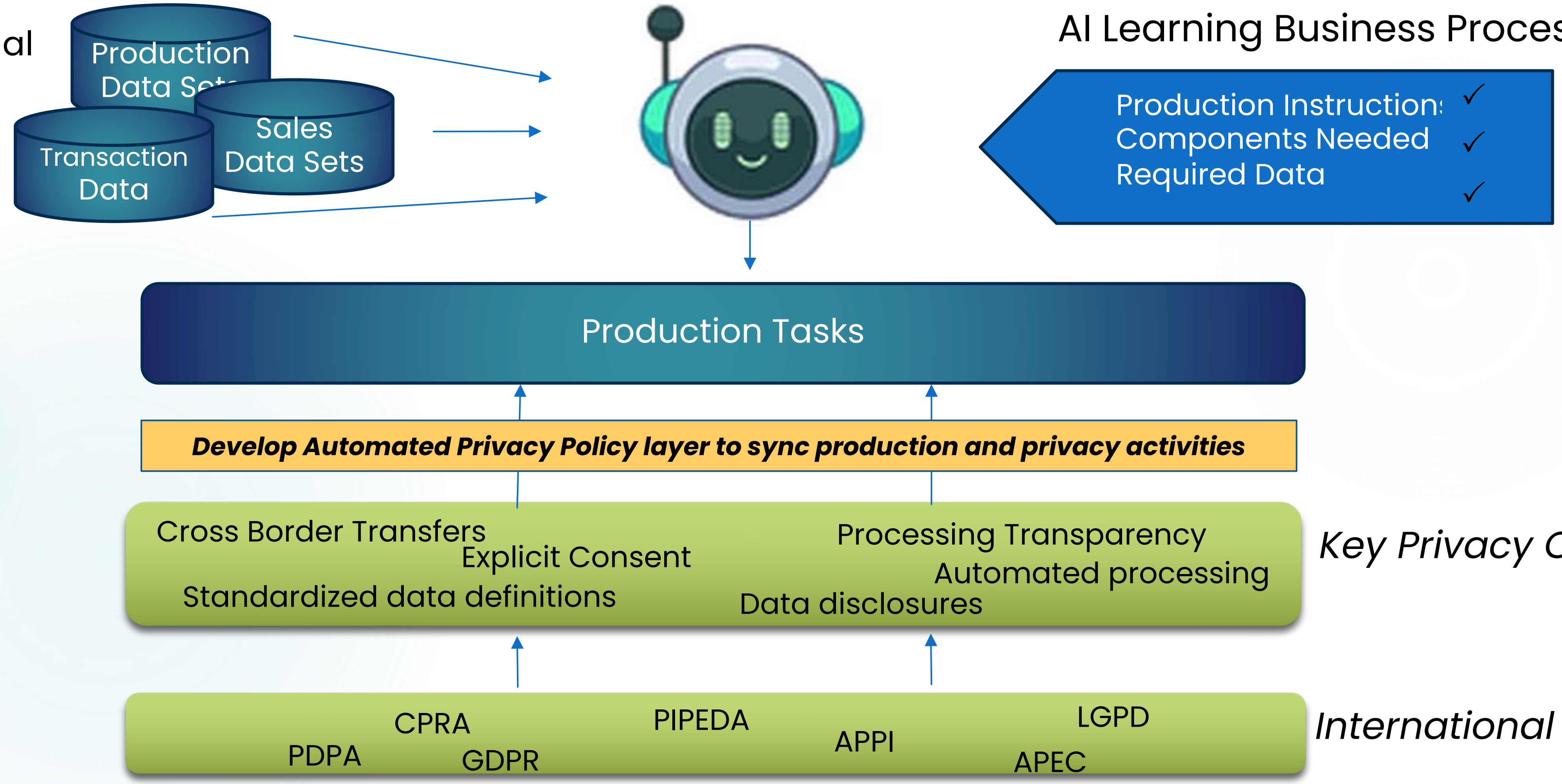
4. The selected model is "trained” on the specific data and model combinations.

5. Performance analysis feedback.

6. The AI model is validation testing

7. Adapt to changing data and environments through retraining.

But which content can AI understand?



Transactional production data feeds business processes

Production Data Sets

Sales Data Sets

Transaction Data

AI Learning Business Process

Production Instructions ✓
Components Needed ✓
Required Data ✓

Production Tasks

Develop Automated Privacy Policy layer to sync production and privacy activities

Cross Border Transfers
Explicit Consent
Standardized data definitions
Processing Transparency
Automated processing
Data disclosures

Key Privacy Concepts

CPRA
PIPEDA
LGPD
PDPA
GDPR
APPI
APEC

International Regulations

Can AI understand and apply key reglatory requirements?

Can AI models support guardrails

In the future AI services need to be aligned to our rules and regulations regarding the use of personal information.

- The EU released the **European Union's Artificial Intelligence Act** which is the world's first comprehensive effort to regulate AI.
- Through assessment processes, laws will categorized AI applications for:

  No risk
  Limited risk
  High risk
  Unacceptable risk

  - Any application that presents an unacceptable risk is prohibited by default and cannot be deployed in the EU.
- President Bidens Executive order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence

Turning reglatory requirements into technical controls will allow AI services to apply data protection to sensitive data processing.

Will guardrails slow us down ?

Which car would you drive fast?

Warning!
This car has no brakes or airbags

But this car has four wheel disk brakes and airbags

AI without governance, is like a car with no brakes

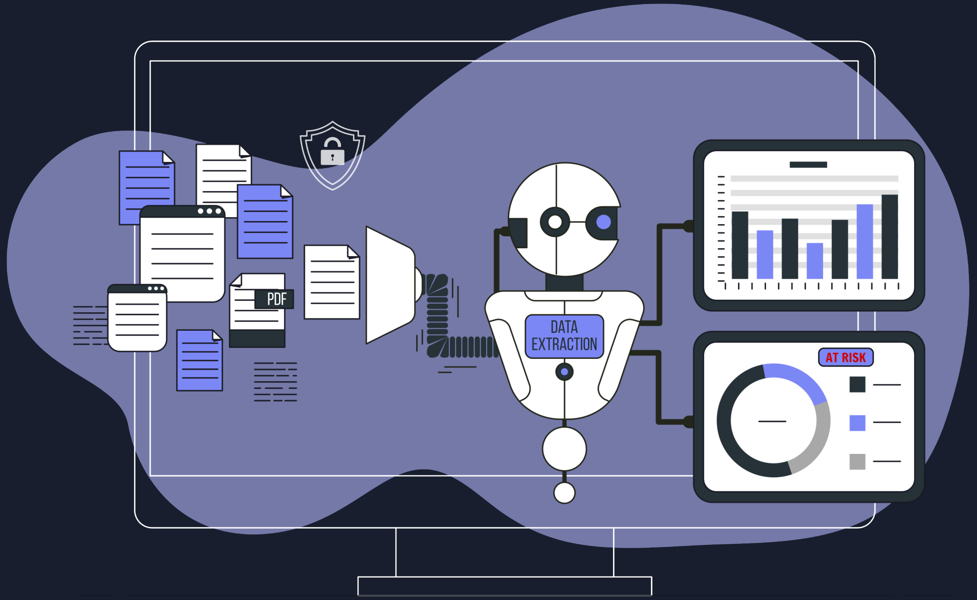
Enterprise Ready GenAI: Key Challenges

AI Governance beyond Privacy Impact Assessments

Exfiltration: All your data belongs to me…?
How don we Trust but Verify with GenAI usage?

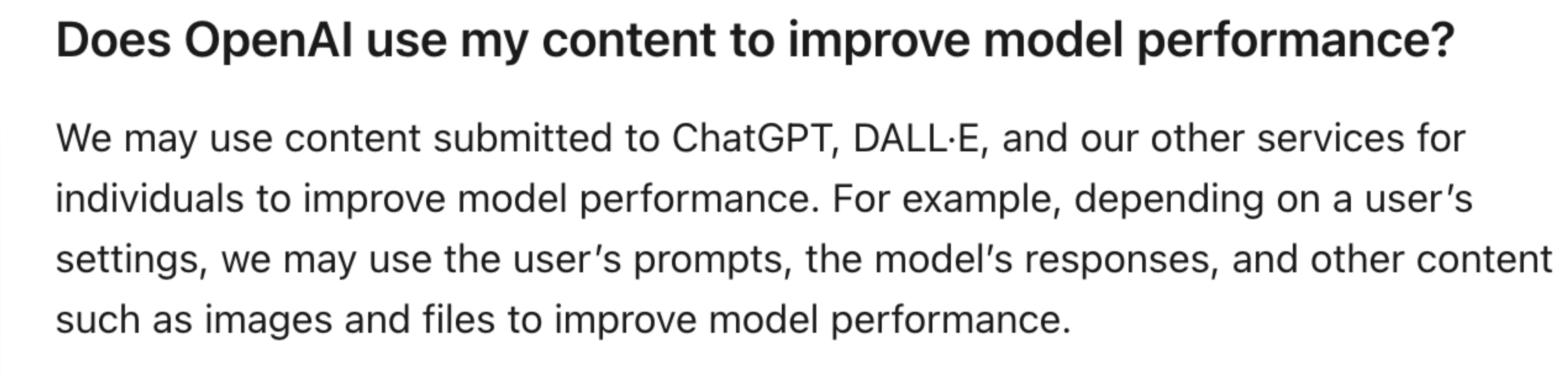
Sanitization: How do you train your ML models on sanitized data sets?

Bias: How do you detect and correct for bias in your data sets before ML model training?



Exfiltration

- Most data sent to SaaS AI apps may be used by them to further train their models.



Does OpenAI use my content to improve model performance?

We may use content submitted to ChatGPT, DALL·E, and our other services for individuals to improve model performance. For example, depending on a user's settings, we may use the user's prompts, the model's responses, and other content such as images and files to improve model performance.

https://help.openai.com/en/articles/7039943-data-usage-for-consumer-services-faq

Tame AI: Your Data, Your Premises

- No Exfiltration
- Contextual Learning
- Six-nines accuracy (no hallucination)

LEARN

Source of Truth

WHAT'S IMPORTANT TO YOU

Amazon RDS

amazon REDSHIFT

ORACLE

snowflake

salesforce

Microsoft SQL Azure

HubSpot

DISCOVER

Structured & Unstructured

IS IMPORTANT EXPOSED

Office 365

OneDrive

slack

box

Marketo

twilio

salesforce

zendesk

Looker

servicenow

teradata.

mongoDB

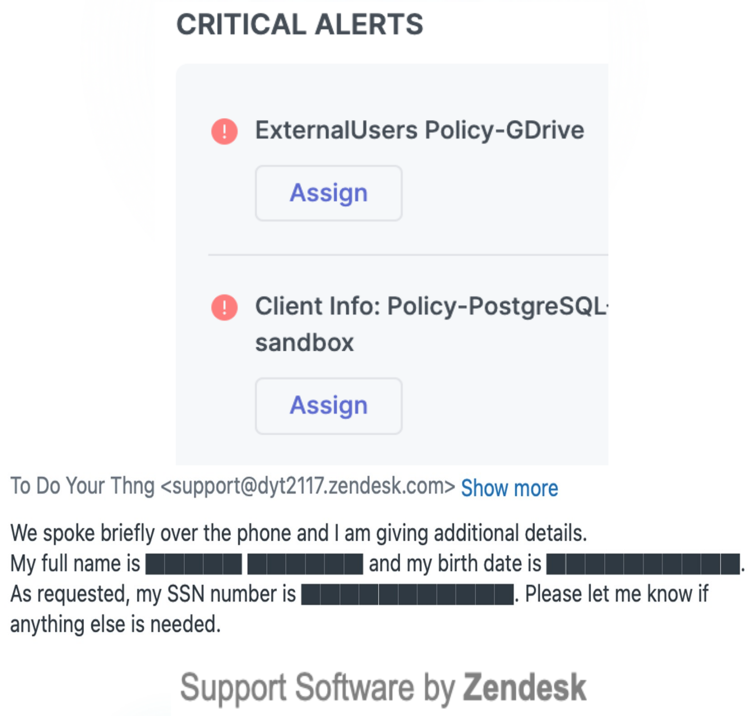
Jira Software

DATADOG

ENFORCE

Security & Privacy Policies

FLAG & PROTECT AUTOMATICALLY



CRITICAL ALERTS

ExternalUsers Policy-GDrive

Assign

Client Info: Policy-PostgreSQL-sandbox

Assign

To Do Your Thng <support@dyt2117.zendesk.com> Show more

We spoke briefly over the phone and I am giving additional details. My full name is [redacted] [redacted] and my birth date is [redacted]. As requested, my SSN number is [redacted]. Please let me know if anything else is needed.

Support Software by **Zendesk**

Transfer Impact Assessment

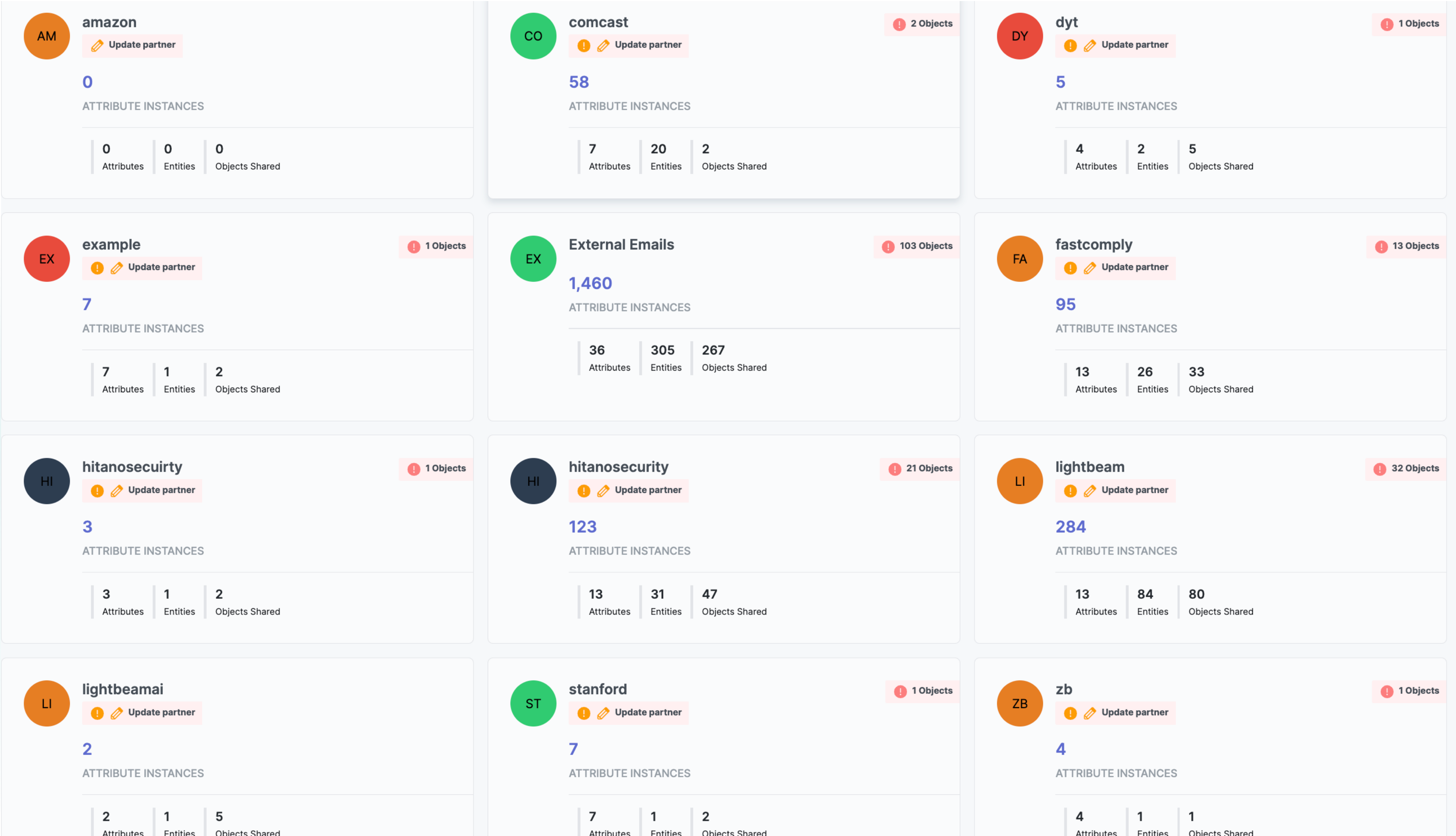
Automated Recording of Sensitive Data Transfers via Native API Integrations

What sensitive data is leaving your org?

Who is that sensitive data getting shared with (sub-processors)?

Are compliance and security controls in place?

Automated Monitoring of Data Transfers



AM **amazon**
Update partner
0
ATTRIBUTE INSTANCES
0 Attributes | 0 Entities | 0 Objects Shared

CO **comcast**
! 2 Objects
! Update partner
58
ATTRIBUTE INSTANCES
7 Attributes | 20 Entities | 2 Objects Shared

DY **dyt**
! 1 Objects
! Update partner
5
ATTRIBUTE INSTANCES
4 Attributes | 2 Entities | 5 Objects Shared

EX **example**
! 1 Objects
! Update partner
7
ATTRIBUTE INSTANCES
7 Attributes | 1 Entities | 2 Objects Shared

EX **External Emails**
! 103 Objects
1,460
ATTRIBUTE INSTANCES
36 Attributes | 305 Entities | 267 Objects Shared

FA **fastcomply**
! 13 Objects
! Update partner
95
ATTRIBUTE INSTANCES
13 Attributes | 26 Entities | 33 Objects Shared

HI **hitanosecuirty**
! 1 Objects
! Update partner
3
ATTRIBUTE INSTANCES
3 Attributes | 1 Entities | 2 Objects Shared

HI **hitanosecuirty**
! 21 Objects
! Update partner
123
ATTRIBUTE INSTANCES
13 Attributes | 31 Entities | 47 Objects Shared

LI **lightbeam**
! 32 Objects
! Update partner
284
ATTRIBUTE INSTANCES
13 Attributes | 84 Entities | 80 Objects Shared

LI **lightbeamai**
! Update partner
2
ATTRIBUTE INSTANCES
2 Attributes | 1 Entities | 5 Objects Shared

ST **stanford**
! 1 Objects
! Update partner
7
ATTRIBUTE INSTANCES
7 Attributes | 1 Entities | 2 Objects Shared

ZB **zb**
! 1 Objects
! Update partner
4
ATTRIBUTE INSTANCES
4 Attributes | 1 Entities | 1 Objects Shared

Sanitization with Anonymization

Tokenize, Anonymize, Redact PII before ML training

Train your AI on real world data sets without worrying about PII.

- Anonymization can be an effective way to stop PII proliferation.
    - Avoids training AI on protected categories.

- Completely fails when it comes to bias detection and control.

- E.g. avoiding race, but having addresses in a data set can still bias AI against certain communities.

Bias | Representative

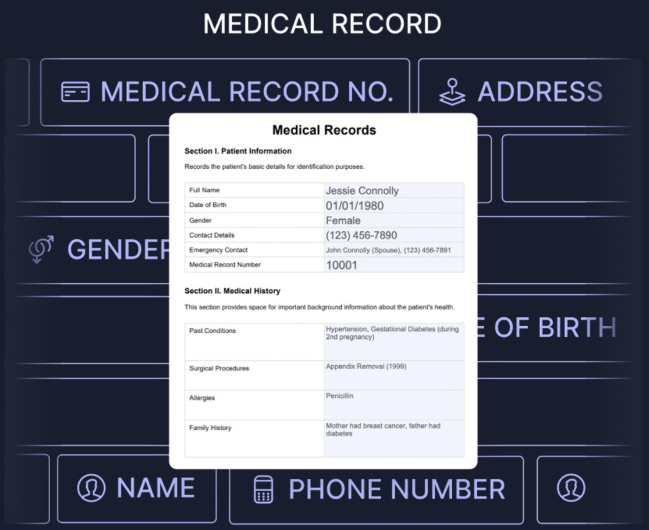
- Bias: prejudiced.
  - E.g. Your purchase history data set is exactly gender balance.

z

  - Is this a good thing?
- Key Question: Is your ML data set **Representative?**

Bias: Detect and Correct – before its too late

Is your ML data set representative of your user base



Medical Records

Section I. Patient Information

Records the patient's basic details for identification purposes.

Full Name	Jessie Connolly
Date of Birth	01/01/1980
Gender	Female
Contact Details	(123) 456-7890
Emergency Contact	John Connolly (Spouse), (123) 456-7891
Medical Record Number	10001

Section II. Medical History

This section provides space for important background information about the patient's health.

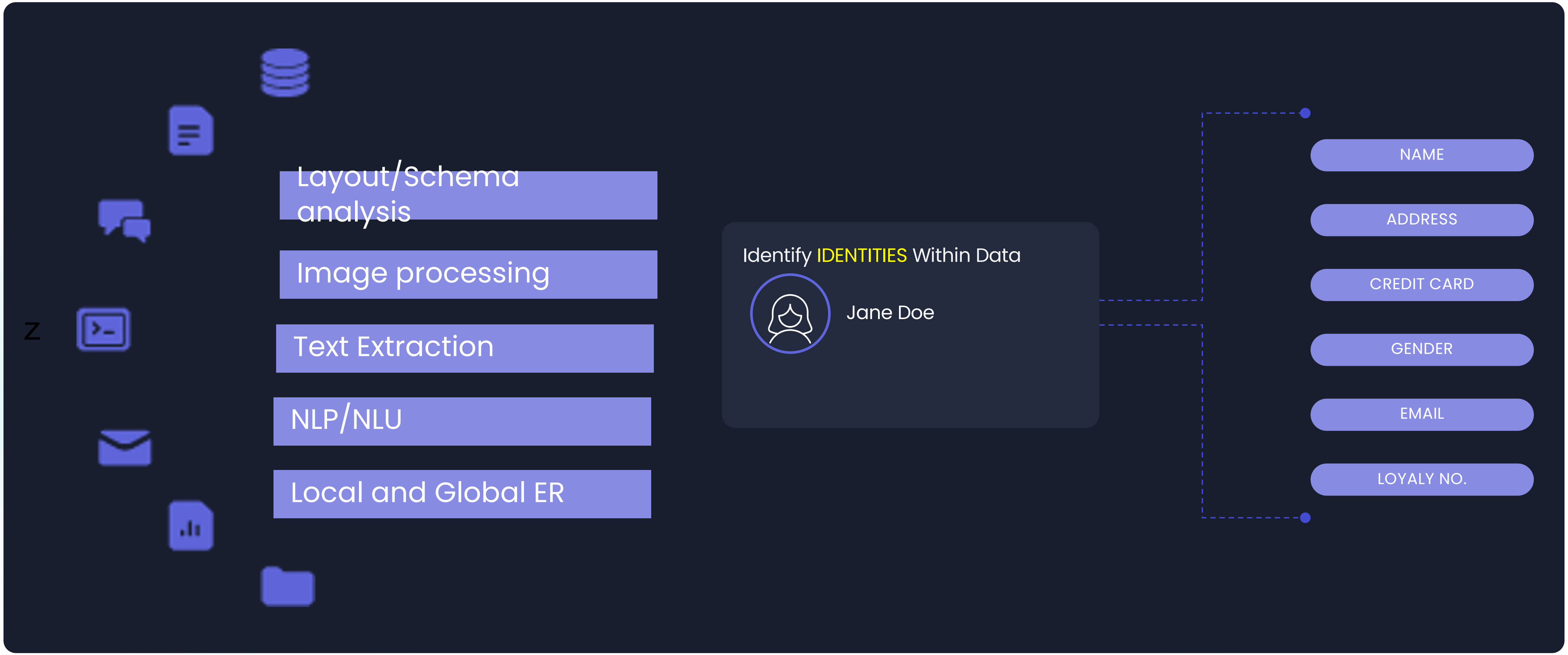
Past Conditions	Hypertension, Gestational Diabetes (during 2nd pregnancy)
Surgical Procedures	Appendix Removal (1999)
Allergies	Penicillin
Family History	Mother had breast cancer, father had diabetes

Define your target user profile mix.

Leverage Entity Resolution on your data set to understand exactly WHOSE data you have.

Validate ER results against your target user profile mix. Repeat until your data set become representative.

Entity Resolution (ER)



Your burden to bear

1. Don’t wait for that PIA request. Inject yourself in your org’s GenAI initiatives.
2. Help your data/AI teams understand the AI challenges:
    - Data Exfiltration
    - z Permitted (but surprising) Data Transfers
    - Personal Data in Training data sets
    - Biased Data Sets
3. Tame your AI without Maiming it!!!