

Data Scientist, Gnyan (2023 – Present)

- Deliver Data Modernization Solutions, ETL Pipelines, Machine Learning Algorithms, Big Data Processing Engines, Artificial Intelligence Systems, MLOps Workflows, Computer Vision Capabilities, Natural Language Processing Modules, Neural Network Architectures, Evolutionary Programming Routines, and GenAI Agents
- Develop Cloud Ready IaaS, PaaS, SaaS, Serverless, Container, Infrastructure-as-Code Solutions

Data

Product Manager, Elsevier (2003 – 2024)

- Lead Product Experience Designer for Elsevier Science Digital and Mobile Solutions.
- Conduct market research, user research, solution discovery, iterate A/B designs, engineer solutions, participate in GTM planning and release milestones.
- Benchmark omnichannel KPIs and manage continuous evolution iterations. Administer accessibility review board charter and compliance.

Product

Certifications

- PhD (incomplete), Wright State University, 2004
- Masters in Human Factors Engineering, Wright State University, 2002
- Masters in Computer Science, Andhra University, 1999

Experience (2003-present)

Education (US)

Innovator, Founder

- Founder of Data Consultancy <http://data-bloom.com>
- Founder of AI Consultancy <http://gnyan.ai>
- 2002 DAGSI Eminent Scholar
- Patent Grantee
- 30+ Publications
- Program Review, GTM Chair

Recognition (Industry)

Python	R	AWS	Azure	GCP	Hadoop	Hive	Spark	Kafka	Flume	Sqoop	Flink	Hudi	EMR	Synapse	Snowflake
Databricks	Redshift	BigQuery	Looker	PowerBI	Qlik	Superset	Metabase	DBT	Airflow	Airbyte	Deequ	Great Exp	Alteryx	Nifi	Camel
Pandas	Polars	Storm	Datahub	Collibra	Alation	BigEye	Informatica	IaaS	PaaS	SaaS	Docker	IaC	Kubernetes	Terraform	AutoML
SageMaker	AutoML	XGBoost	Jupyter	MLFlow	Git	Jenkins	Docker	IPC/RPC	TensorFlow	PyTorch	Keras	Hugging Face	OpenAI	GPT	Ollama
LangChain	LlamaIndex	FastAI	LLaMA	Diffuser	Transformer	CNN	RNN	Clustering	Classification	Regression	Deep Learning	Gen AI	Forecasting	Anomaly Detection	Entity Recognition
Image Processing	Object Detection	Sentiment Analysis	Intent Mining	Translation	Dim Factorization	NLP	Computer Vision	NLP	Speech Processing	Neo4j	JanusGraph	Graph Frames	NetworkX	Sentence Xformers	Embeddings
Quantization	Linear Prog	A/B leaderboards	Hyper Params	ONNX/PB	Artifactory	Balsamiq	Figma	Miro	Splunk	Elastic	Grafana	Trino	Helm	Flutter	NextJS

Skills Table (Data, ML, BD, AI, Gen AI, UX)

Building Science, Driving Outcomes

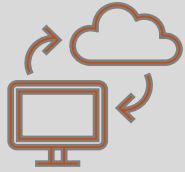
AI Gold Rush

We sell shovels

<http://gnyan.ai>



Our Data Services (1 Pager)



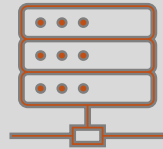
Cloud Modernization

Scalable Cloud first
FedRAMP Solutions



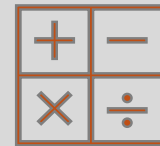
Generative AI

Modern Mixed
Intelligence
Outcomes



Analytics Engineering

Drive outcomes
from raw data



Data Science

Deploy Advanced
Machine Learning
Algorithms



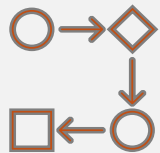
Edge/IoT Engineering

Situational In-Field
Solutions



Accessible Insights

24/7 Information;
every person, every
device



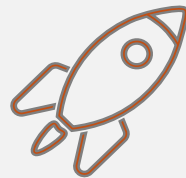
IT/OT/RPA Modernization

RPA Technology to
Support People &
Processes



NLP/Document Digitization

Mining structured
knowledge from
physical media



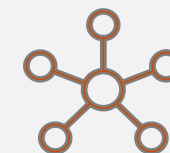
Rapid Innovation

Explore and
Engineer New
Innovations like
Blockchain,
Quantum



Optimization

Drive progressive
preemptive
strategies



Intelligence Research

Transitive
intelligence from
disparate geo-
temporal sources



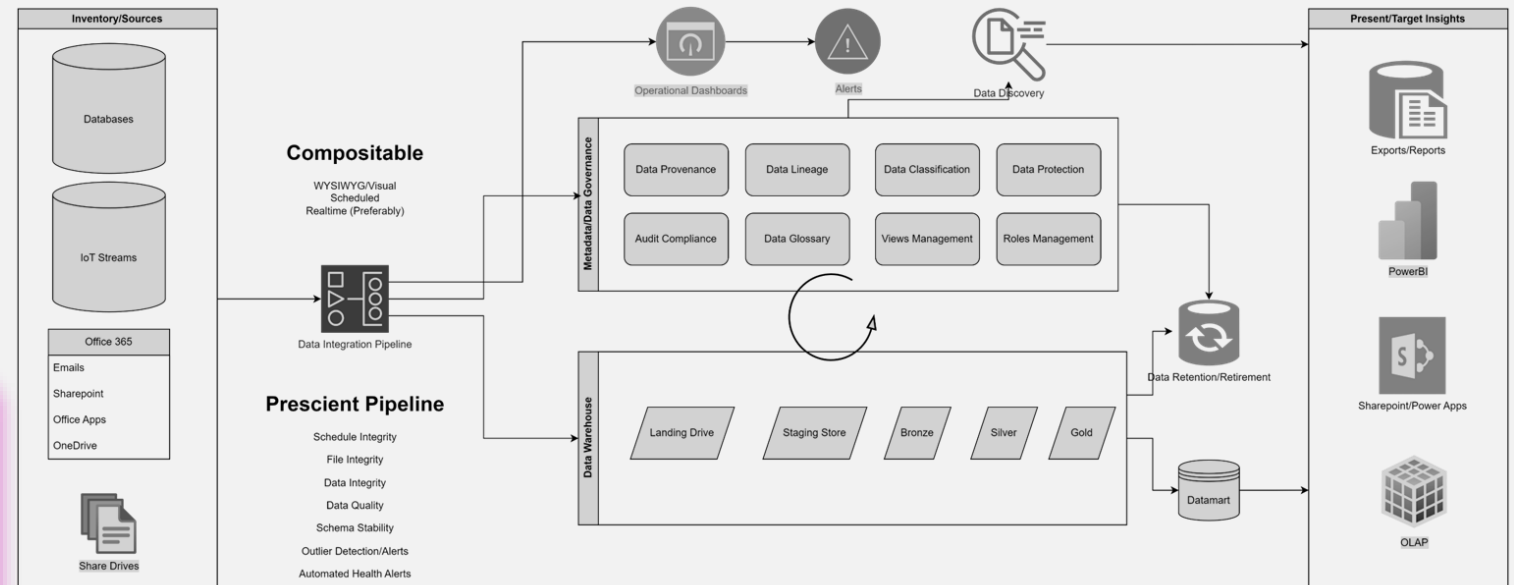
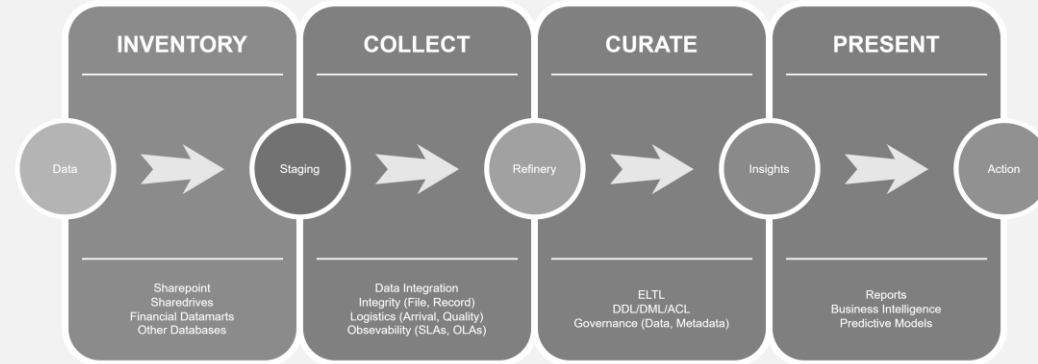
Search

Multimodal,
Personalized
Semantic Search &
Retrieval

Addressing Data Platform, Tooling Challenges

Integration
Quality
Exploration
ETL Refinery
Governance & Management
Reporting
Business Intelligence
Data Science
MLOps
Data Marketplaces

challenge: do your data professionals
have the right compute, ai, warehouse,
devops tooling?



Self-Service Analytics Appliance (Warehouse-on-the-Go)

demo: whirr analytics infrastructure for everyone in under 5 minutes

192.168.27.10:8510

Docker Composition Configuration

Project Name
Seshu Appliance

Select Project

Services

- airbyte
- airflow
- cdap
- es01
- grafana
- jupyter

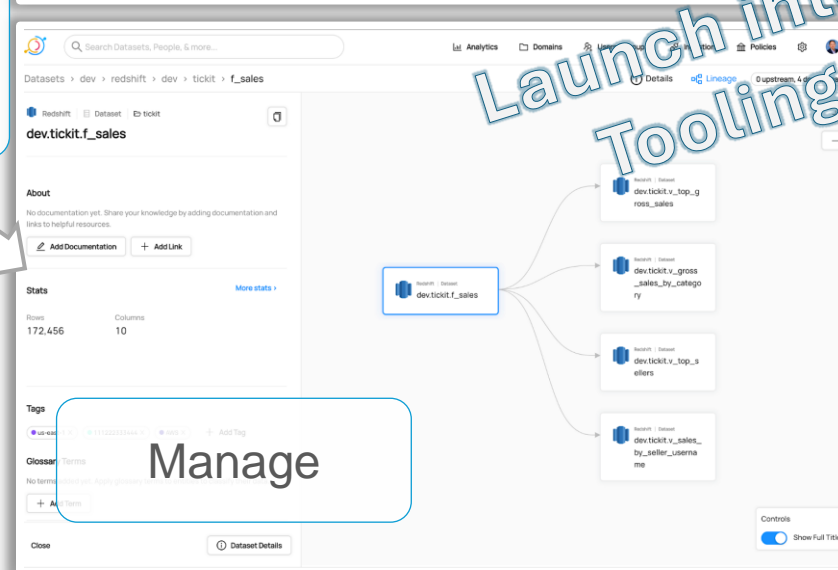
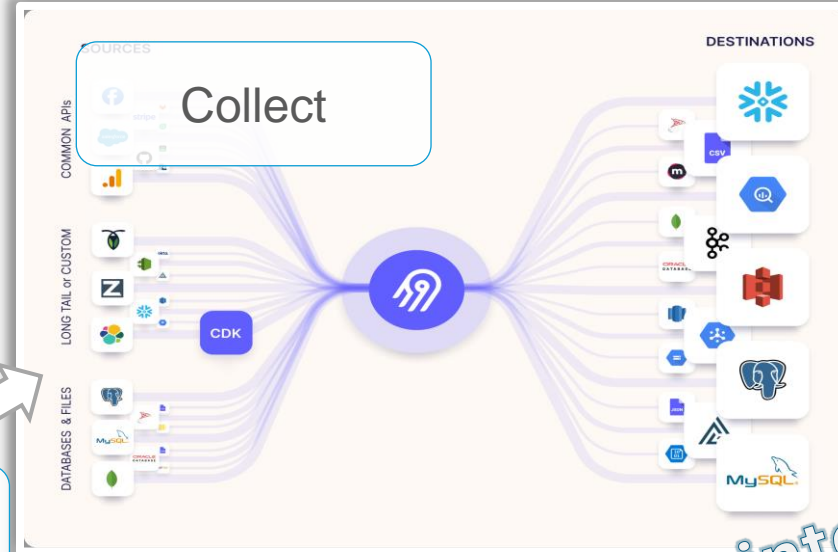
☒ Enable jupyter

Configuration for jupyter

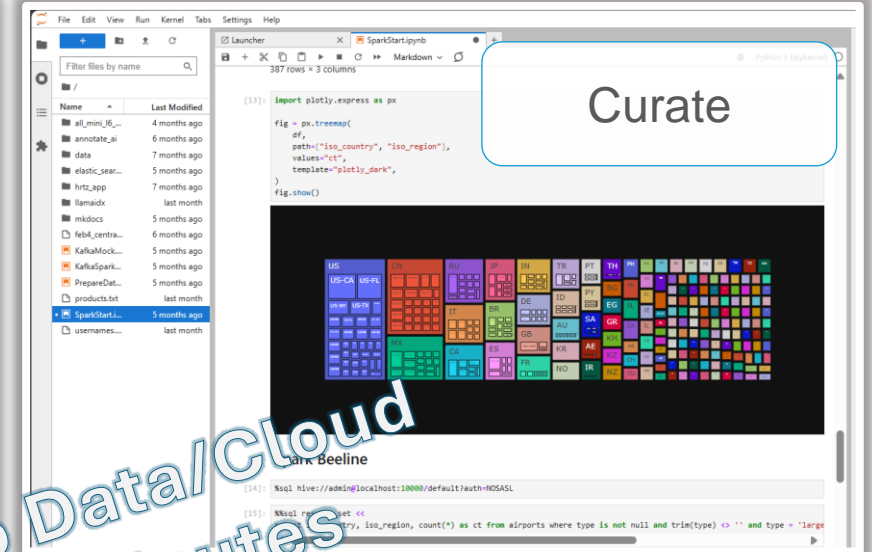
Base Image for jupyter
quay.io/jupyter/all-spark-notebook

☒ Use GPU Acceleration

Manage (Concurrent, Isolated, and Durable) Appliances for Every User in **Minutes** with a GUI

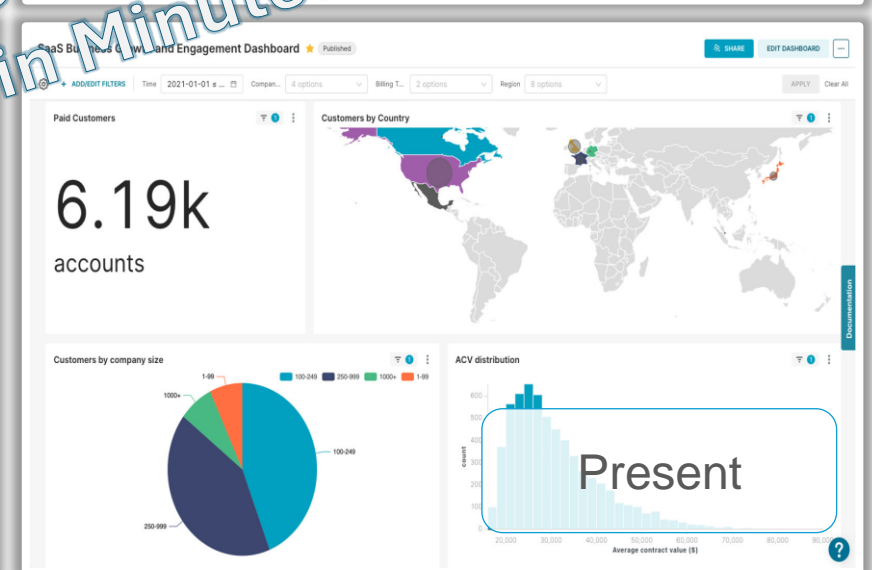


Manage



Curate

Launch into Data/Cloud Tooling in Minutes



Present

From Data to Solutions (Swift MLOps)



MLOps Demo.zip

The collage features several screenshots and code blocks:

- Top Left:** MLflow Experiments interface showing a model named 'gregarious-mouse-368'.
- Top Center:** A code snippet for setting up an experiment and tracking URI.
- Top Right:** MLflow Experiments interface showing a table of runs with columns like Run Name, Created, F1, Duration, User, Source, and Models.
- Bottom Left:** A terminal window showing Docker commands for setting up a container environment.
- Bottom Center:** A file explorer showing a folder named 'mlflow_experimentation' with files like 'handwritten_test' and 'handwritten_test.pkl'.
- Bottom Right:** MLflow Models interface showing a model named 'mnist_recognizer' with its version and schema.

Overlaid text boxes include:

- Continuous Artifacts** (top left)
- Auto Hyperparameter Tuning** (top center)
- A/B Champion Challenger Selection** (top right)
- Oneliner Configuration** (center)
- challenge: fragmented ml lifecycle** (center)
- demo: incorporating seamless ml + ops lifecycle with mlflow workbench** (center)
- Algo Selection** (bottom left)
- Scale Deployment** (bottom center)
- Visual DevOps Promotion** (bottom right)

Catalyze Continuous Evolution with ML Leaderboards

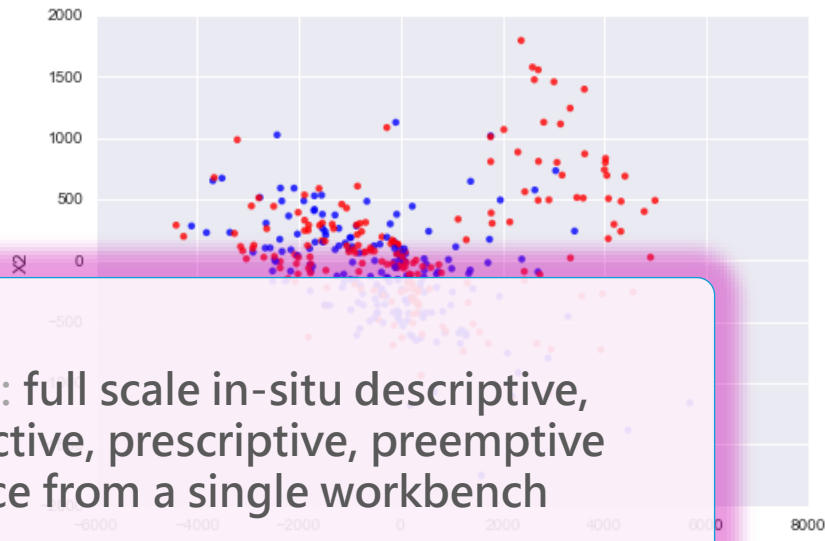
Descriptive & Prescriptive Modeling (ML)

```
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis
pca = PCA(n_components=2)
pipeline = Pipeline([('pca', pca)])
```

```
In [20]: reduced_x = pipeline.fit_transform(numeric_x)
X_r2 = pd.DataFrame(reduced_x, columns=['X1', 'X2'])

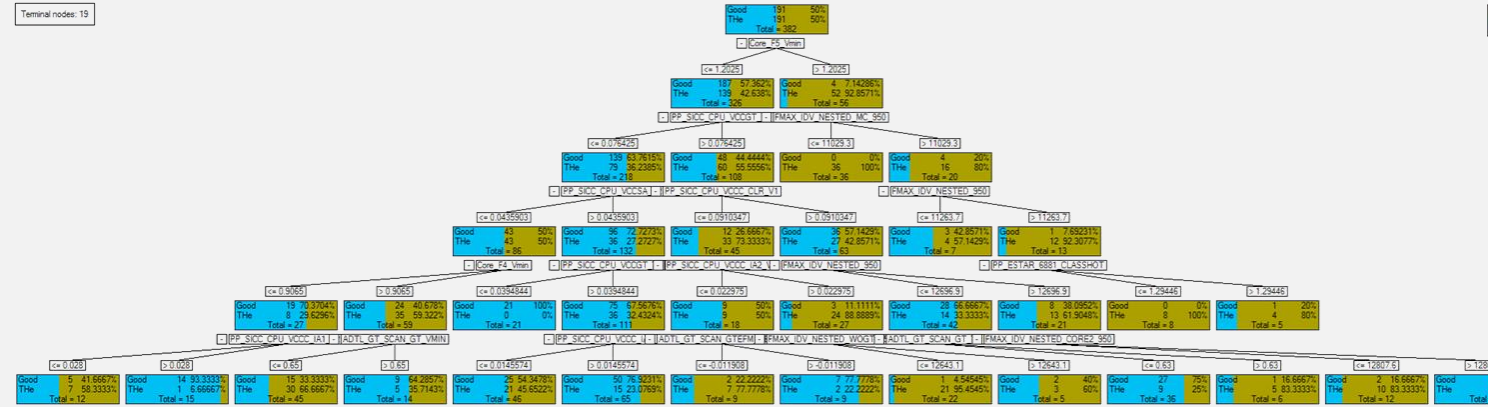
# Percentage of variance explained for each components
print('Explained variance ratio (first two components): %s'
      % str(pca.explained_variance_ratio_))
ax = sns.regplot(data=X_r2, x='X1', y='X2', x_jitter=10**2, y_jitter=10**2)
plt.show()
```

Explained variance ratio (first two components): [0.8605536 0.0550536]



demo: full scale in-situ descriptive, predictive, prescriptive, preemptive science from a single workbench

Terminal nodes: 19

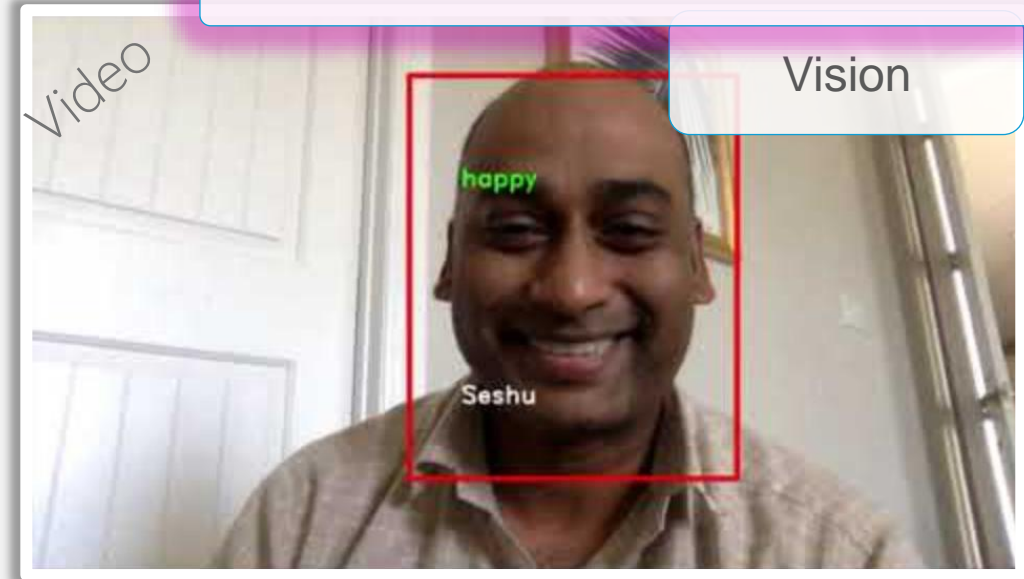


 Cluster+LSD+Pictures (2).html

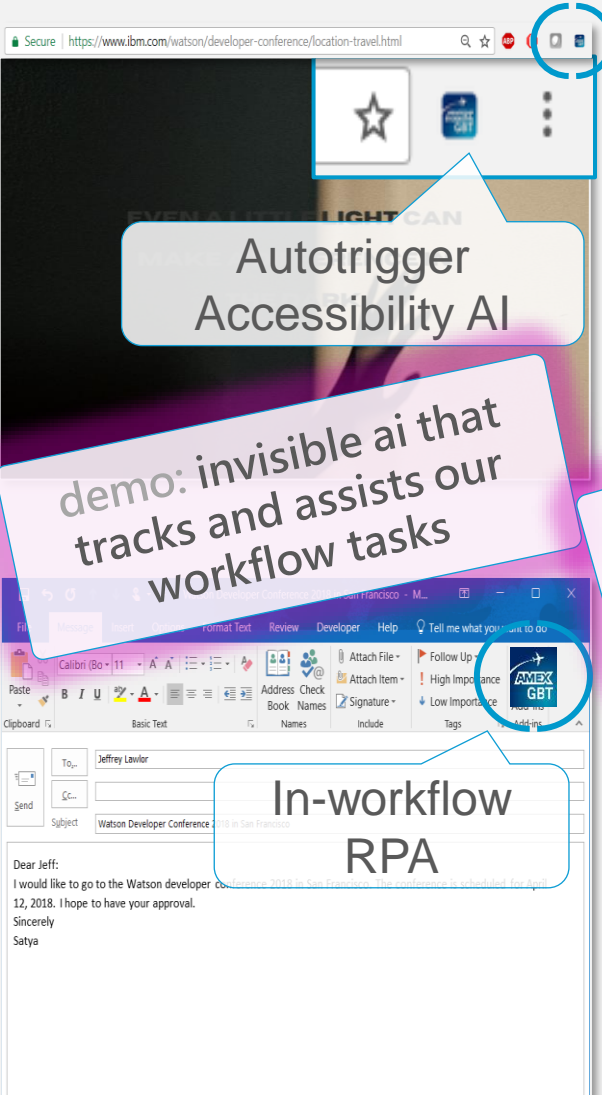
 DPMCRelationsSpecificThe.html

AI Sensory Intelligence (AI)

demo: circa 2016, ai doing sensory actions:
senses that belonged to human realm



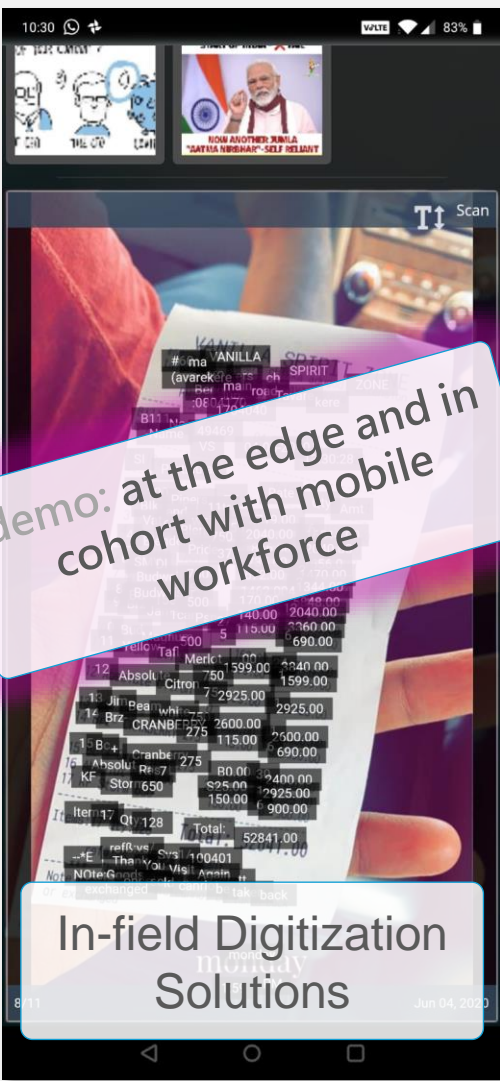
Accessibility & In-Situ Integration (Edge RPA)



Autotrigger Accessibility AI

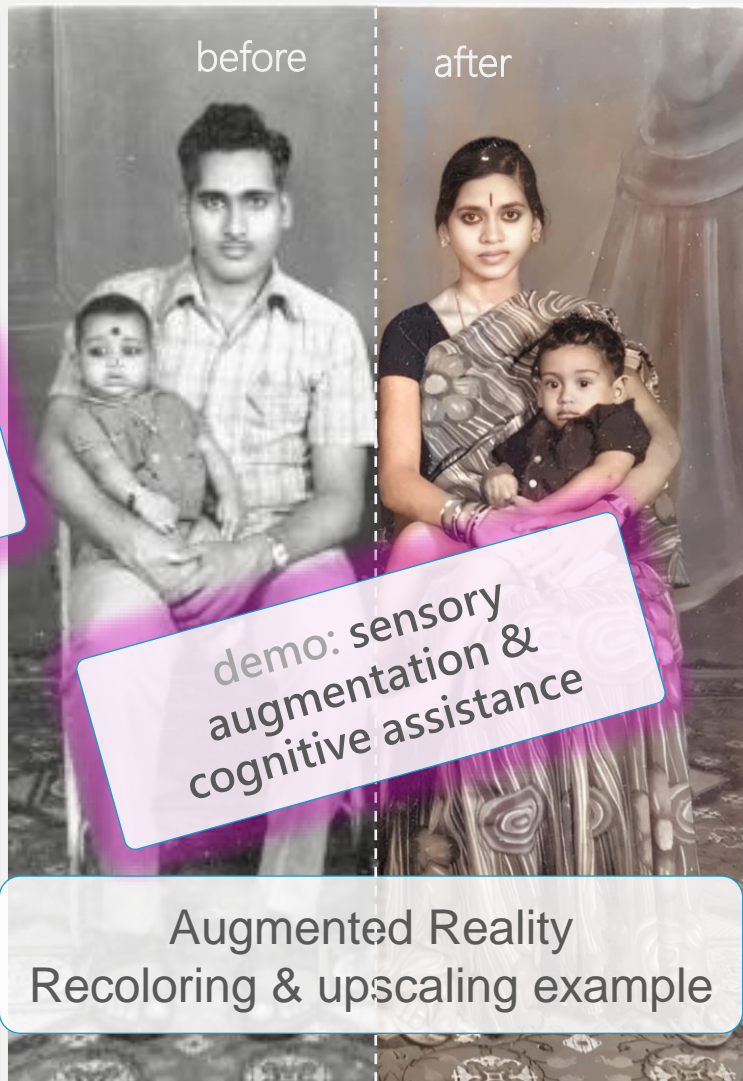
demo: invisible ai that tracks and assists our workflow tasks

In-workflow RPA



In-field Digitization Solutions

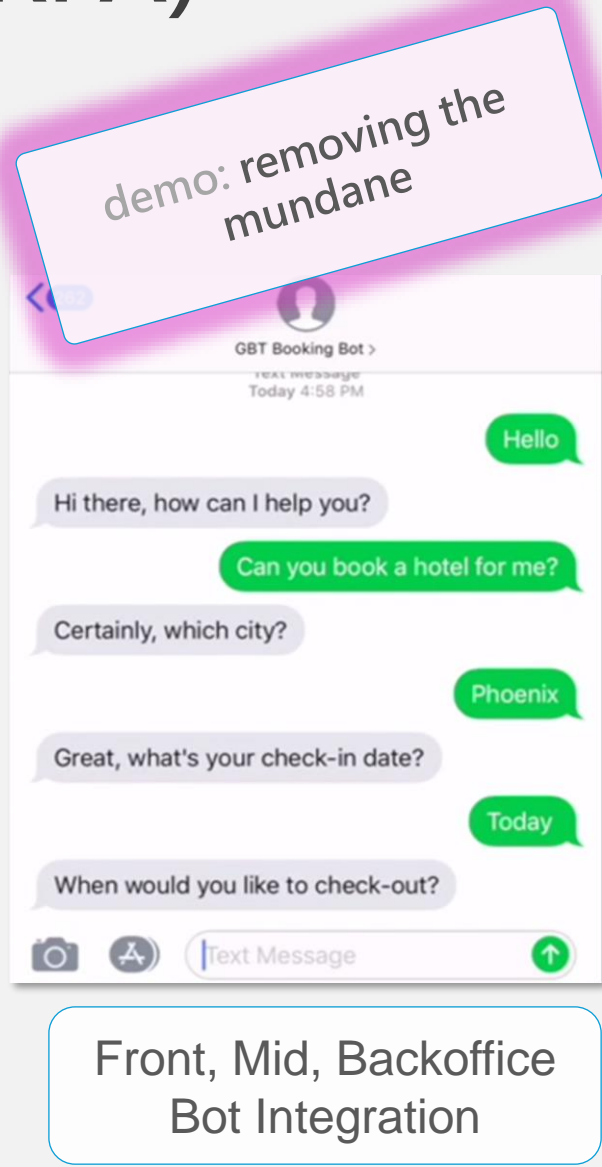
demo: at the edge and in cohort with mobile workforce



before after

Augmented Reality Recoloring & upscaling example

demo: sensory augmentation & cognitive assistance



demo: removing the mundane

GBT Booking Bot

Today 4:58 PM

Hello

Hi there, how can I help you?

Can you book a hotel for me?

Certainly, which city?

Phoenix

Great, what's your check-in date?

Today

When would you like to check-out?

Text Message

Front, Mid, Backoffice Bot Integration

Generative AI

Bespoke Agent Central
for the Enterprise

Human-on-the-loop

10 second slides

Research Copilot

demo: 360° wholistic human-on-
the-loop outcomes

not every ai is a chatbot

Iterate Outputs
Quickly (in seconds)

Code Dev

Governance
Steward

Digital Twin
Agents

Language
Translation

Research
Librarian

Business
Anomaly
Monitor

TLDR
Summarizer

RFP Bid Writer

Infographic
Analyst

Office Research
Toolbar

Course
Instructor

NAICS/HS
Classifier

Accessibility
Helper

Code
Conversion
Agent

CloudOps Agent

Front Office
Chatbot

Brand Sentiment

Code
Vulnerability
Patching

Data PCI/PII
Watchdog

Directory Router

Data Quality
Curator


Spiritual
Wellbeing Agent

BI SQL Assistant

Competitive
Intelligence
Agent

Keyword Coding

Misses the context

 **HTS**.USITC.GOV

snakes and ladders

2024 HTS Revision

User Query

Search Results: 9

Rates

Chapter Notes

Section Notes

Download Chapter 1

0106.20.00.00 Reptiles (including snakes and turtles)	Heading/ Subheading	Stat Suffix	Article Description	Unit of Quantity	RATES OF DUTY	
					1	
					General	Special
0208.50.00.00 Of reptiles (including snakes and turtles)	0106.31.00	00	Birds of prey	No.	1.8%.1/.	BH, CL, CO, D, E, IL, JO, KR, MA, OM, P, PA, PE, S, SG)
0210.93.00.00 Of reptiles (including snakes and turtles)						
3926.90.99.30 Ladders			Psittaciformes (including parrots, parakeets, macaws and cockatoos)	No.	1.8%.1/.	Free (A, AU, BH, CL, CO, D, E, IL, JO, KR, MA, OM, P, PA, PE, S
7326.90.86.60 Ladders	0106.32.00	00				

challenge: classic term search does not work

Intent Understanding

Search

snakes and ladders

Search

board game strategy rolling dice snakes and ladders

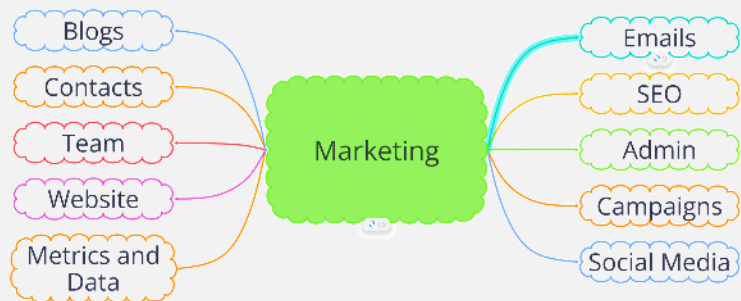
Generate, Augment, Retrieve (NORAG)

_id	Code	HierarchicDescription	combined_score
6110	950440	parlour games including pintables billiards special tables for casino games and automa	75%
6112	950490	parlour games including pintables billiards special tables for casino games and automa	74%
6107	950400	parlour games including pintables billiards special tables for casino games and automa	74%
6111	950450	parlour games including pintables billiards special tables for casino games and automa	73%
6109	950430	parlour games including pintables billiards special tables for casino games and automa	72%
6113	950460	parlour games including pintables billiards special tables for casino games and automa	72%

demo: modern embedding-based semantic search (using llms) to pivot from RAG to GAR yields better outcomes

Captures Gameboard Correctly

```
{"code": "950430", "description": "parlour games include special tables for casino games and automatic bowling machines operated by coins banknotes bank cards tokens of payment >> other games operated by coins banknotes any other means of payment other than automatic bowling"}  
"explanation": "Although the passage doesn't specifically mention strategy and rolling dice terms, it is related to parlour games which include various types of games. This code covers a wide range of games including those operated by coins or tokens."}
```

challenge: everyone has a data warehouse. but is it cataloged and annotated for good use?

demo: let ai dig deep and annotate databases, tables, columns, and contents to help the stewards

1) Delve deep into Lakehouses

Progress Indicator ⓘ

5 / 29 processed.
fleet_prd1_us.analytics.sales_full_history...

Abbreviation Settings ⓘ

Upload your abbreviations CSV file

Drag and drop file here
Limit 200MB per file • CSV

Browse files

Wipe Abbreviations

≡ Abbreviation	≡ Full_Form
A	AMOUNT
ACCPT	ACCEPTED
ACCS	ACCESS
ACCT	ACCOUNT
ACCIDENT	ACCIDENT

Feed business nomenclature (opt)

2) Let mixed intelligence annotate thoroughly

`datalabs.lab_sfmc_extracts.lists`

This table stores lists created by clients. Each list has a unique Client ID, List ID (which could be an identity), Name and Description. It also tracks the DateCreated when it was made, its Status indicating if active or inactive, as well as the type of List - which can be categorized under 'ListType'.

```
[
  ("ClientID", "Customer Client ID", "Unique identifier for a customer."),
  ("ListID", "IDENTITY (LIST)", "Identity or unique reference number of the list."),
  ("Name", "List Name", "The name given to the created list by client."),
  ("Description", "List Description", "A brief description about what is included in this particular list."),
  ("DateCreated", "Creation Date", "Shows when a specific list was made or generated."),
  ("Status", "Active/Inactive Status", "Indicates if the list is currently active or not, used to manage lists et"),
  ("ListType", "Category of List Type", "Categorizes different types of lists created by clients.")
]
```

AI generated descriptions

`datalabs.lab_sfmc_extracts.sendjobs`

This table stores information about jobs to send emails. Each job has a unique identity (ClientID) and is associated with another ID for the actual email message (SendID). It includes details like sender name, email address, scheduled time, sent time, subject line, recipient list name, triggering event external key, template external key, status of the job, URL to preview the content before sending, whether it's a multipart or single-part mail and any additional information. The table helps manage email campaigns efficiently by tracking their progress from creation to delivery.

Locate Data

Search Query

Industry average price of technology stocks in 2017

Search

Clear history

Query: Industry average price of technology stocks in 2017

Vector Results:

	_id	_score	table_catalog	table_schema	table_name	table_description
0	minio_default_s_and_p_mapping_name:varchar	9.7345	minio	default	s_and_p_mapping	This table maps stock
1	minio_default_s_and_p_mapping_sector:varchar	9.6782	minio	default	s_and_p_mapping	This table maps stock
2	minio_default_s_and_p_mapping_symbol:varchar	9.6336	minio	default	s_and_p_mapping	This table maps stock symbols
3	minio_default_s_and_p_5_years_name:varchar	6.274	minio	default	s_and_p_5_years	This table tracks stock and price
4	minio_default_s_and_p_5_years_high:double	6.2627	minio	default	s_and_p_5_years	This table tracks stock and price
5	minio_default_s_and_p_5_years_close:double	6.2572	minio	default	s_and_p_5_years	This table tracks stock and price
6	minio_default_s_and_p_5_years_open:double	6.2416	minio	default	s_and_p_5_years	This table tracks stock and price
7	minio_default_s_and_p_5_years_date:date	6.2142	minio	default	s_and_p_5_years	This table tracks stock and price
8	minio_default_s_and_p_5_years_low:double	6.2017	minio	default	s_and_p_5_years	This table tracks stock and price
9	minio_default_s_and_p_5_years_volume:integer	6.1546	minio	default	s_and_p_5_years	This table tracks stock and price

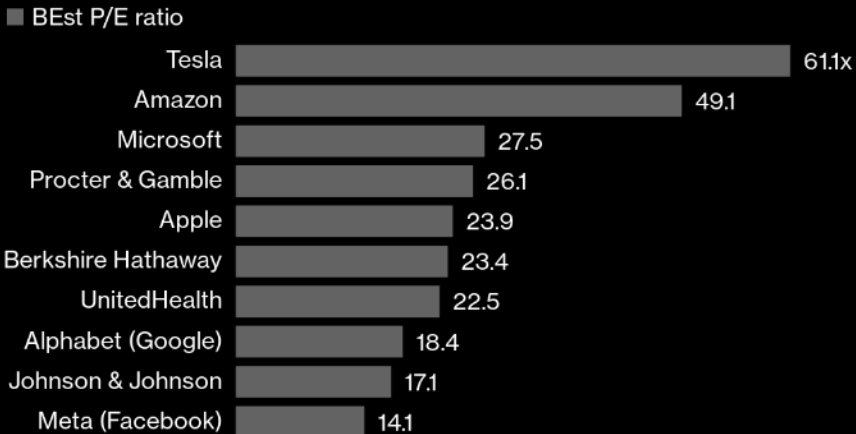
SQL Query:

```
SELECT AVG(close) AS avg_price
FROM minio.default.s_and_p_5_years ssp, minio.default.s_and_p_mapping sm
WHERE ssp.symbol = sm.symbol AND sm.sector LIKE '%Technology%';
```



3) Locate assets & compose SQL
with ease

Longtime growth stocks Google and Facebook are now cheaper than many value stocks



4) Elicit insights from English

challenge: sql and bi skills are not universal
demo: let ai discover, author, and deliver insights

Machine Generated Insights

←

↺

🏠

⚠️ Not secure | 192.168.27.10:34000

🔊

☆

⚙️

|

📄

⌵

📄 New Chat

🔗 Workspace

🔍 Search

Previous 30 days

Innovative Fed2Fed Use Cases: Dig

June

Roman Empire Facts

Internal Data Services Commerciali

Simple Demo Title

📄

neural-chat:latest

Set as default

Private, Free Models

📄

Hello, Seshu

How can I help you today?

⚡ Suggested

Give me ideas
for what to do with my kids' art

Prompt

↑

Tell me a fun fact
about the Roman Empire

Prompt

↑

Grammar check
rewrite it for better readability

Prompt

↑

Show me a code sn
of a website's sticky he

Prompt

↑

challenge: chatgpt is great, but is it private & safe?

demo: deploy a chatgpt functionality within your
enterprise

Host Private ChatGPT

+ Send a Message

🎤

↑

LLMs can make mistakes. Verify important information.

← ↻ 🏠

⚠ Not secure | 192.168.27.10:34000/c/4f88e290-91f0-400b-ad93-78716312e45d

🔍 ⭐ ⚙ | 📄 ⭐

📄 New Chat

🔗 Workspace

🔍 Search

Today

Building location ⋮

Judge Bio (3 words)

Previous 30 days

Innovative Fed2Fed Use Cases: Di


June

Roman Empire Facts

Internal Data Services Commerciali

Simple Demo Title

📄 llava:latest ▾ +



Where is this building?

Omni Models

Private, Secure Outcomes

Who presides here?

📄 llava:latest

< 2/2 >

challenge: can open models beat commercial performance?

demo: yes, plug-n-play open models easily

+ Send a Message

🔍

Parse Triples

Compose & Test AI Tools on a canvas

+

LangChain

POST API Chain

Conversation Chain

Conversational Retrieval QA Chain

LLM Chain

Multi Prompt Chain

Drag n Drop

challenge: langchain programming skills are challenging

demo: deliver visual composing capability

Ollama

Cache

Base URL

http://192.168.27.14:11434

Model Name

neural-chat

Temperature

0.0

Additional Parameters

Output

Ollama

LLM Chain

Language Model

Prompt

Output Parser

Input Moderation

Chain Name

Entity Extraction Assistant

Output

LLM Chain

Hi there! How can I help?

John Doe was caught with 10gms of meth in his trousers by Insp. Menendez on Jun 22, 2024 at El Paso TX PoE.

["John Doe", "", "2024-06-22", "El Paso, TX", "31.77", "-106.46", "10gms", "Meth", "", "Trousers"]

Building Knowledge Graph Services with Ease

Type your question...

1) Point-n-click AI Apps & Services

2) Driving customer outcomes...

Build & Operate in Minutes

challenge: managing gen ai apis/swarms is challenging

demo: build and scale gen ai apis, apps, and services easily from same canvas

Text to Code & Code to Cloud

1 Feature/Requirement

Articulate
requirements

2 Promptware

Generate the code

Feature/Requirements

Seizure Folio

Video Recoloring

ML Flow

Docker Compose

Sales Trends

What is the requirement?

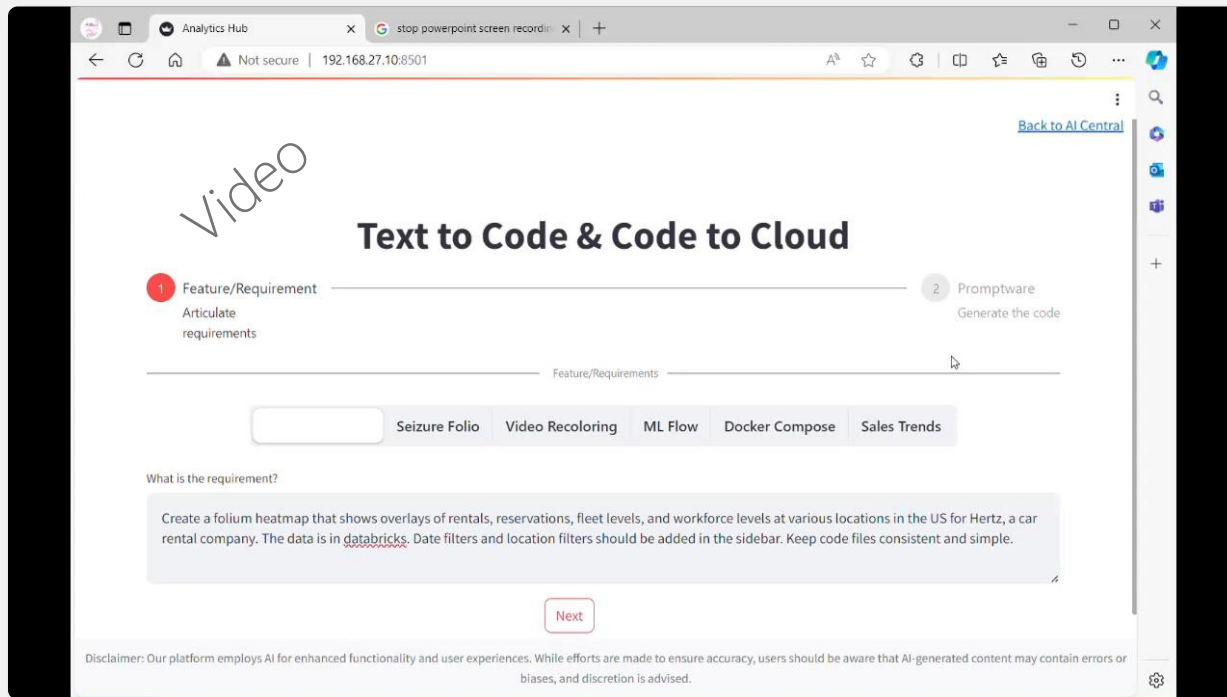
Create a folium heatmap that shows overlays of rentals, reservations, fleet levels, and workforce levels at various locations in the US for Hertz, a car rental company. The data is in databricks. Date filters and location filters should be added in the sidebar. Keep code consistent in fewest files.

Next

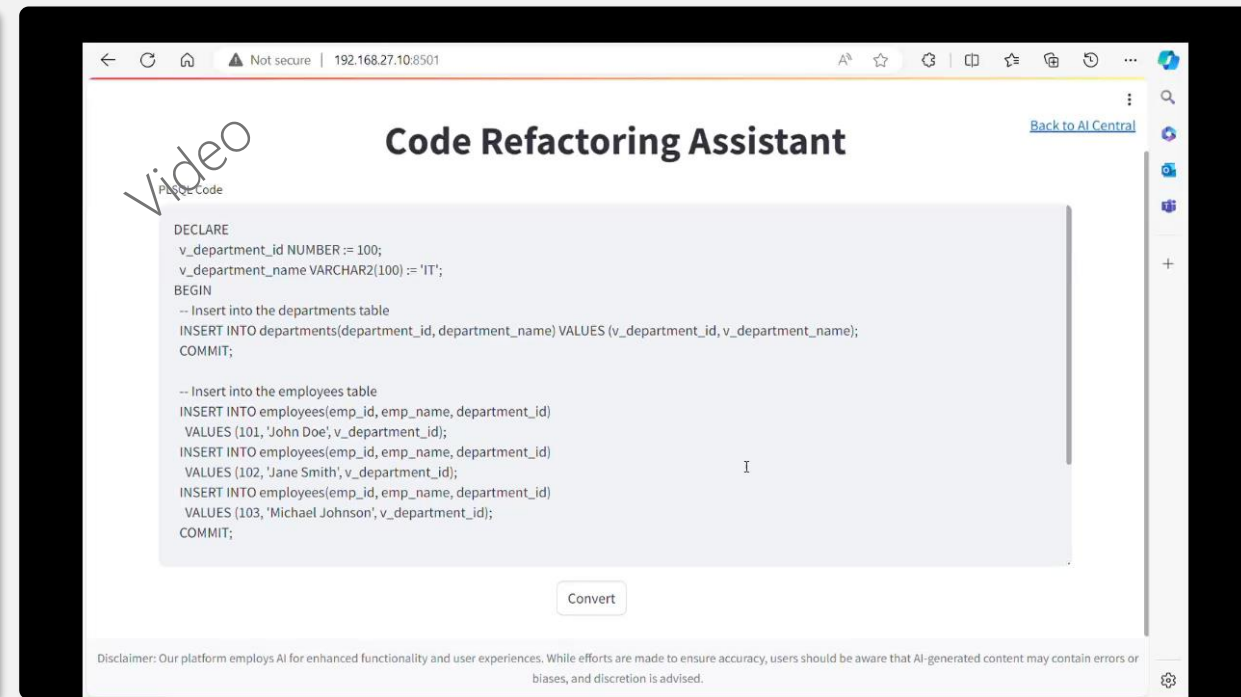
From JIRA to Cloud in Hours, not sprints

challenge: requirements → code → cloud is slow

For advanced coding & devops



Mundane lift-n-shift to the cloud



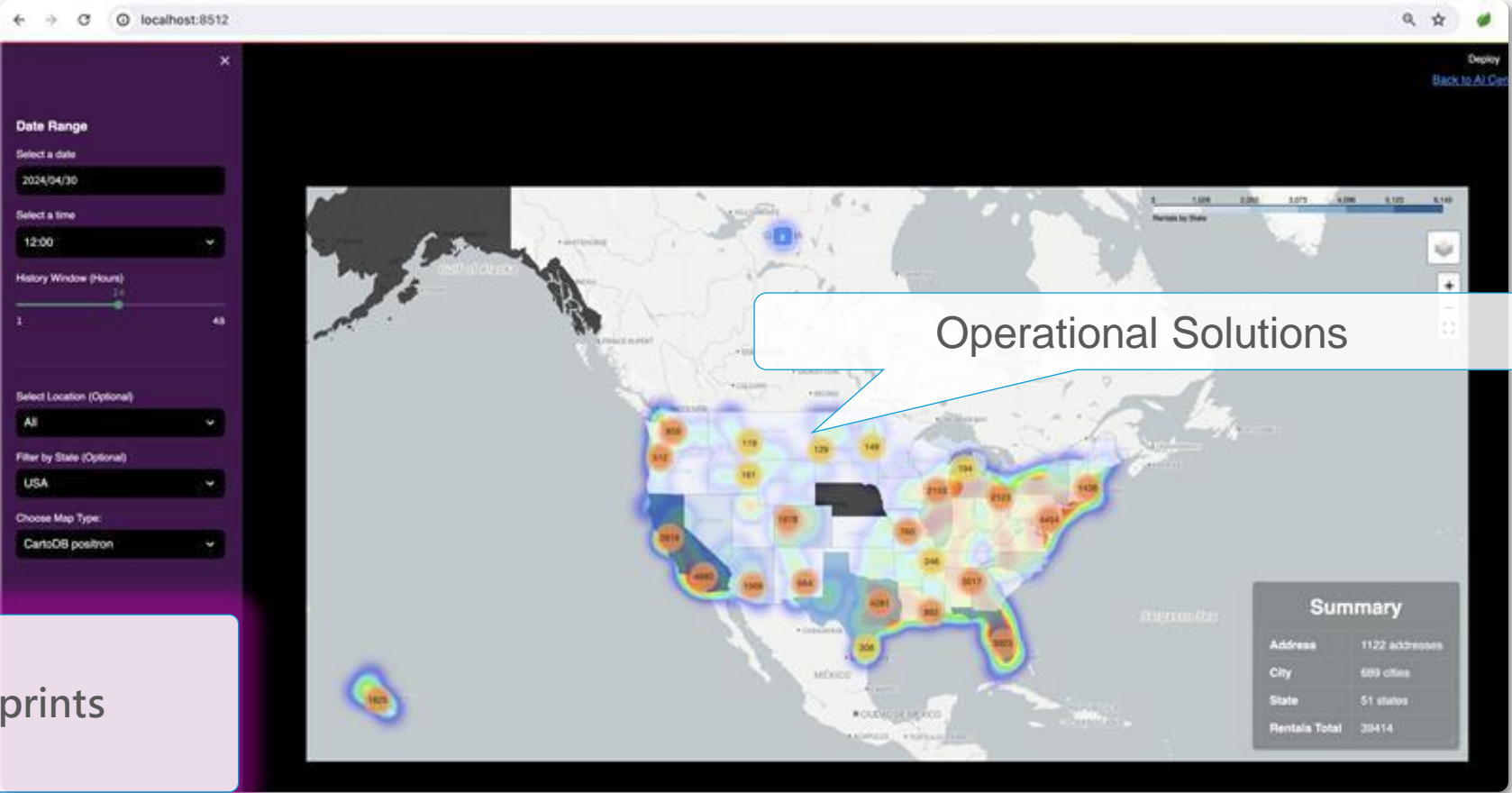
demo: ai developed promptware bootstraps archetypes quickly

Swift Promptware Releases in Hours

Create a folium heatmap that shows overlays of rentals, reservations, fleet levels, and workforce levels at various locations in the US for Hertz, a car rental company. The data is in databricks. Date filters and location filters should be added in the sidebar. Keep code consistent in fewest files.

10AM ●

3PM ●



demo: results in hours... not sprints