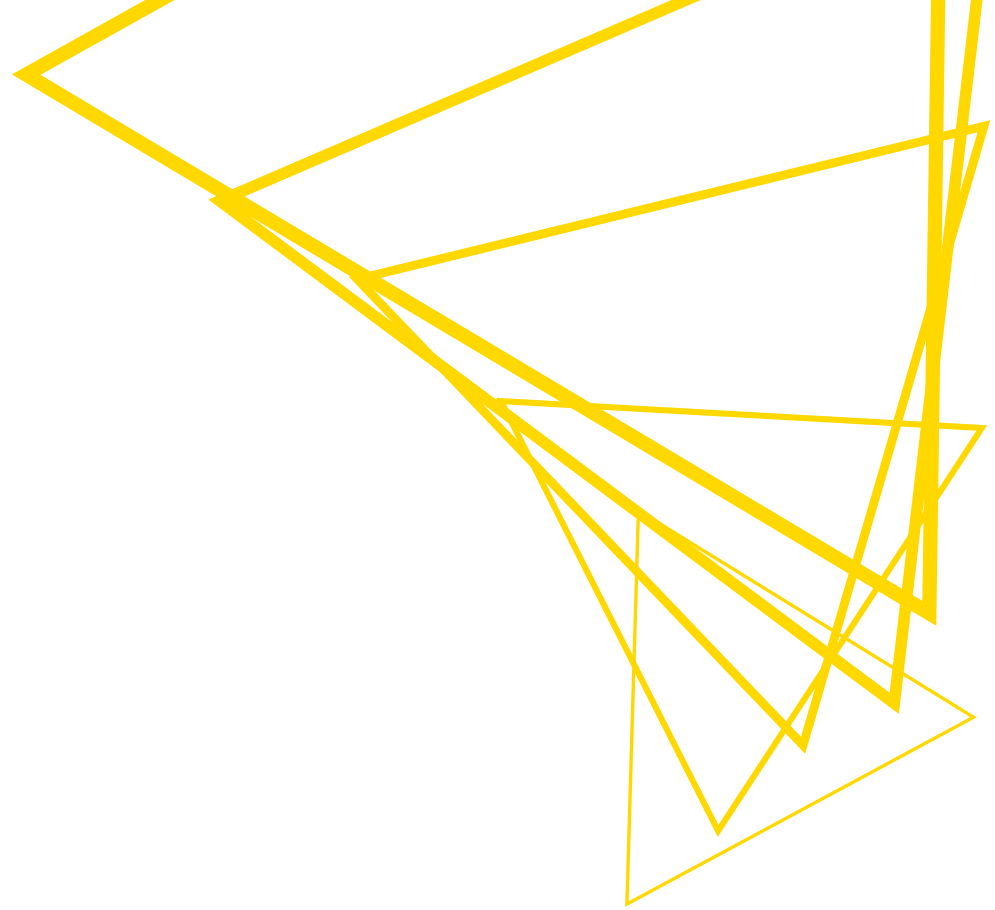


# Using, Customizing, and Governing GenAI

Michael Berthold

DataHop Munich, 14 March 2024



# GenAI - Opportunities and Risks

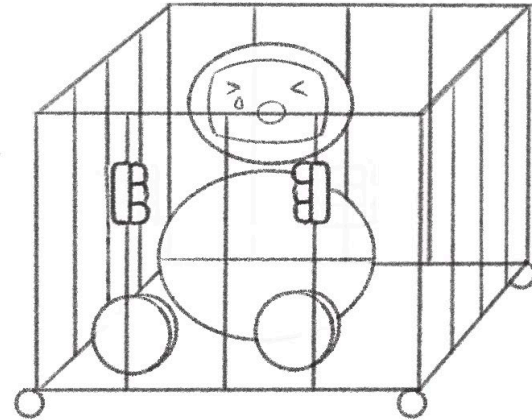
- + Massive Impact on “mundane” tasks
  - + Create new Content
  - + Optimize Content
- + Gather & Summarize Information
- + Broad User Base & Applicability
  - + Experiment with different models
  - + Explore data & information
  - + Apply large scale models easily
  - + Integrated into daily work everywhere

...inhibits potential positive impact

- Need for (Quality) Checking
  - Copyright
  - Sloppy/faulty quality
  - Unethical Outcomes
  - Fake Outcomes
- Governance Requirements
  - Which AI?
  - Privacy / Confidential Data Leakage
  - Skyrocketing Consumption
  - Regulation & Risk Management

Fear of all of the above...

If (Gen)AI can't be trusted...  
...what can we do?



# Agenda

KNIME and GenAI

GenAI to Assist & Explain

Using & Customizing GenAI

Managing and Governing GenAI

Summary

# GenAI and Data Science

## Assist & Explain

- Help Workflow Builders
- Teach Data Science Skills

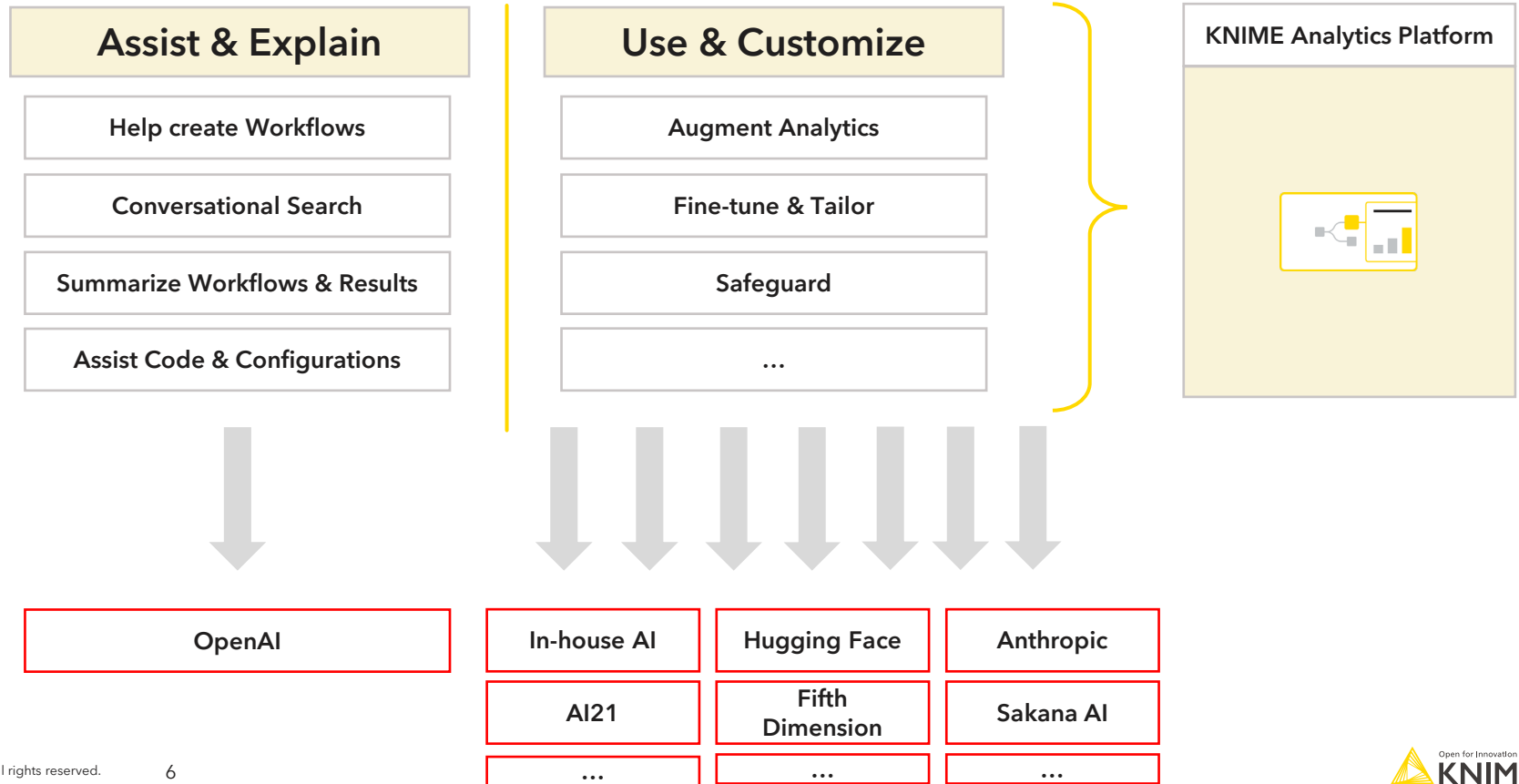
## Use & Customize

- Augment Analytics with GenAI
- Augment GenAI with Analytics

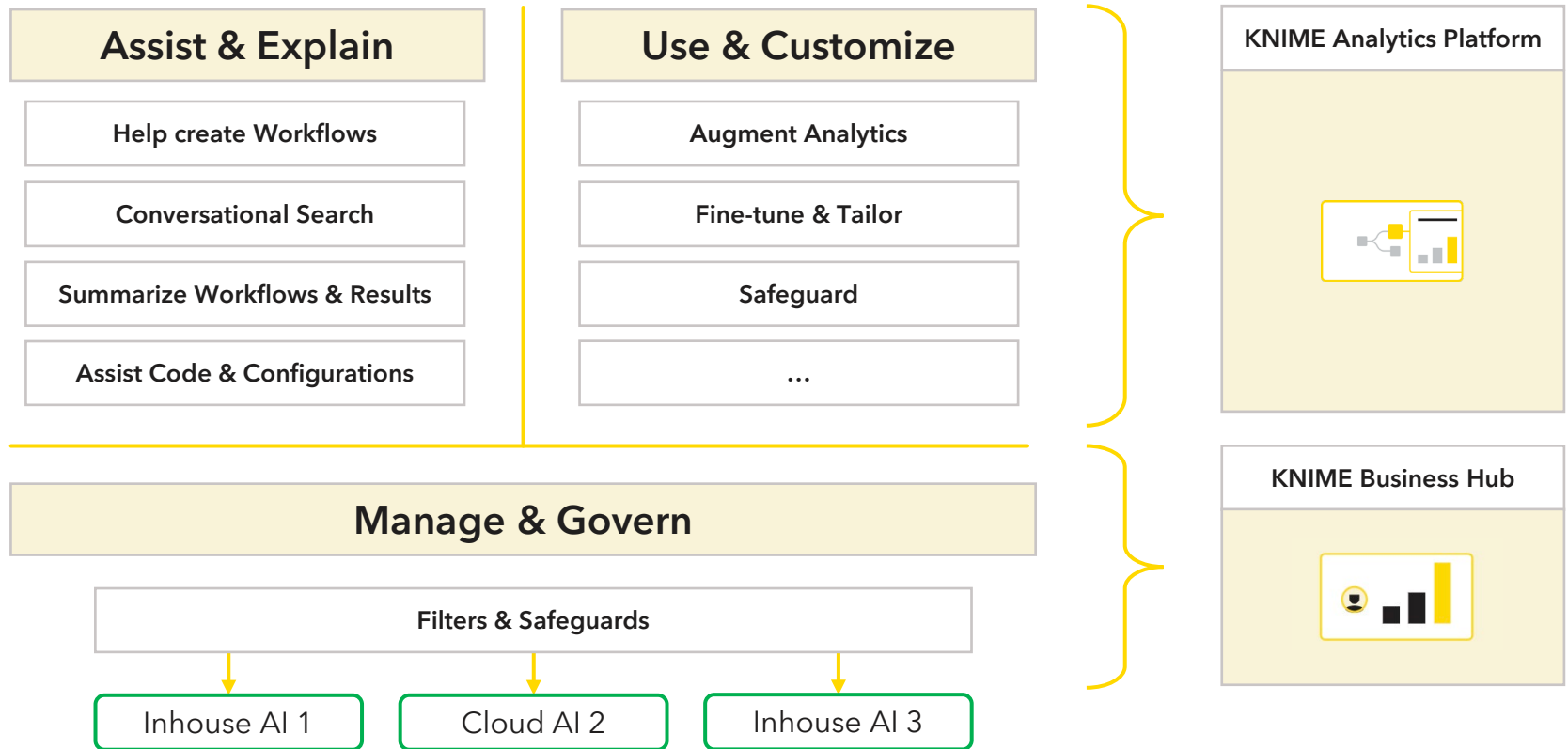
## Manage & Govern

- Protect Data & Conversations
- Govern Usage

# GenAI for everybody: Flexibility



# GenAI in the Enterprise: Flexibility & Governance



# Agenda

KNIME and GenAI

GenAI to Assist & Explain

Using & Customizing GenAI

Managing and Governing GenAI

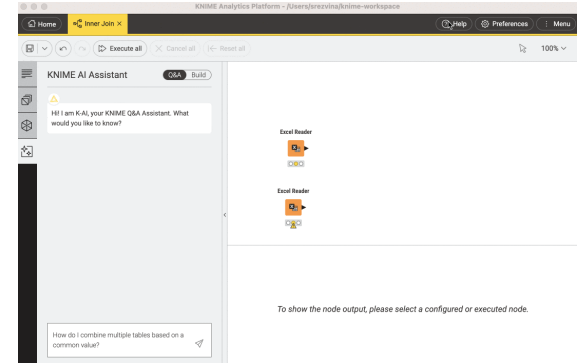
Summary



# Assisting and Explaining

K-AI: Build Q&A and Creator Mode

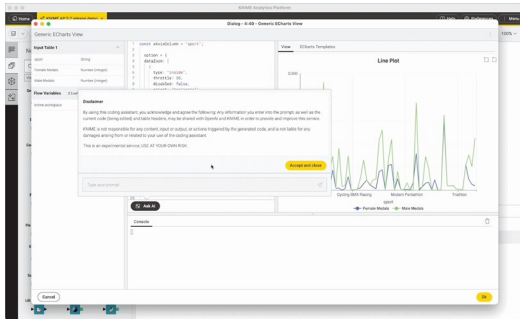
- Assist & Explain
- Help create Workflows
- Conversational Search
- Summarize Workflows & Results
- Assist Code & Configurations



[ tbdev: adding Context: Forum, Documentation, Components, Blueprints, ... ]

Data Apps...

Copilots for Python, eCharts  
[ tbdev: R, SQL, ... ]



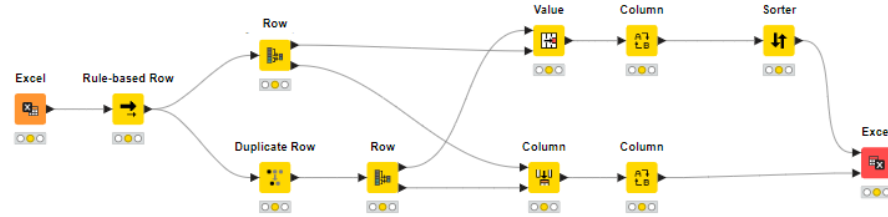
# Explaining Workflows

## GenAI Workflow Explainer

Summarize workflows using

### Workflow

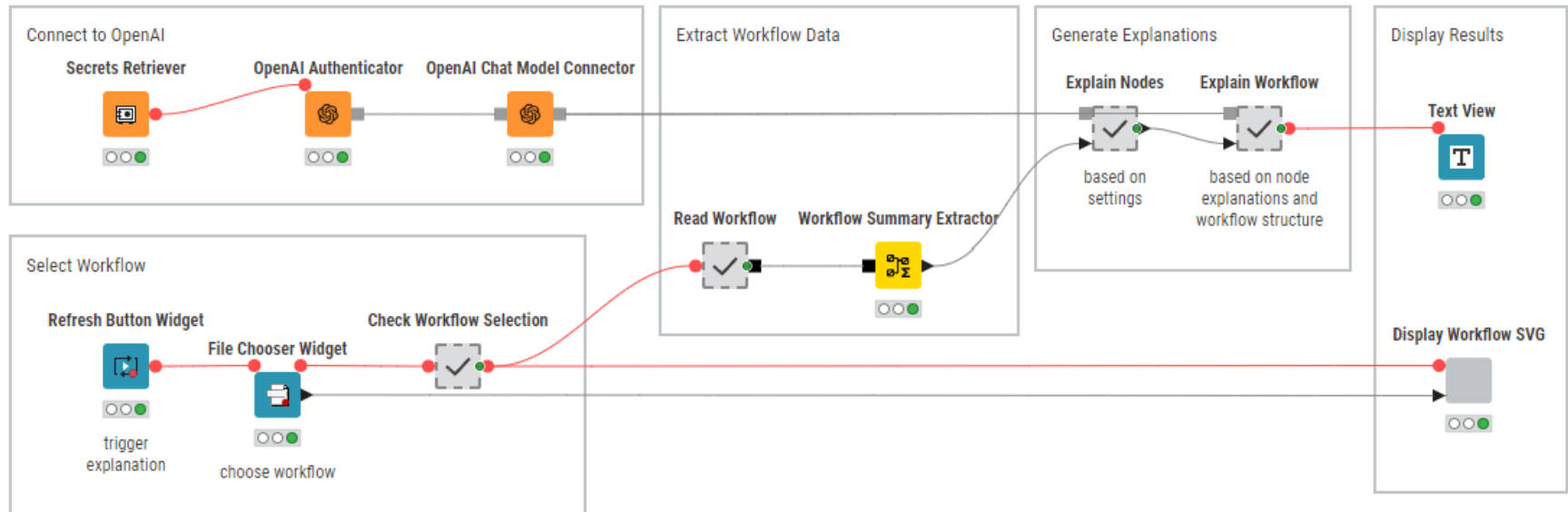
- Building a Simple Classifier
- Combine Clean and Summarize Spreadsheet Data
- Countif and Sumif**
- Data Blending
- Non-standard format Spreadsheets
- Visual Analysis of Sales Data



The workflow reads data from an Excel file containing Olympic athlete information, filters rows based on specific criteria, aggregates data, removes duplicate rows, performs value lookup, appends columns, renames columns, sorts the data, and finally writes the processed data to a new Excel file.

The use-case of this workflow is data transformation or ETL (Extract, Transform, Load). It involves cleaning, filtering, aggregating, and restructuring the data to prepare it for further analysis or reporting. The workflow aims to process and organize the Olympic athlete data for better understanding and analysis.

# Explaining Workflows ... Using Workflows



# Agenda

KNIME and GenAI

GenAI to Assist & Explain

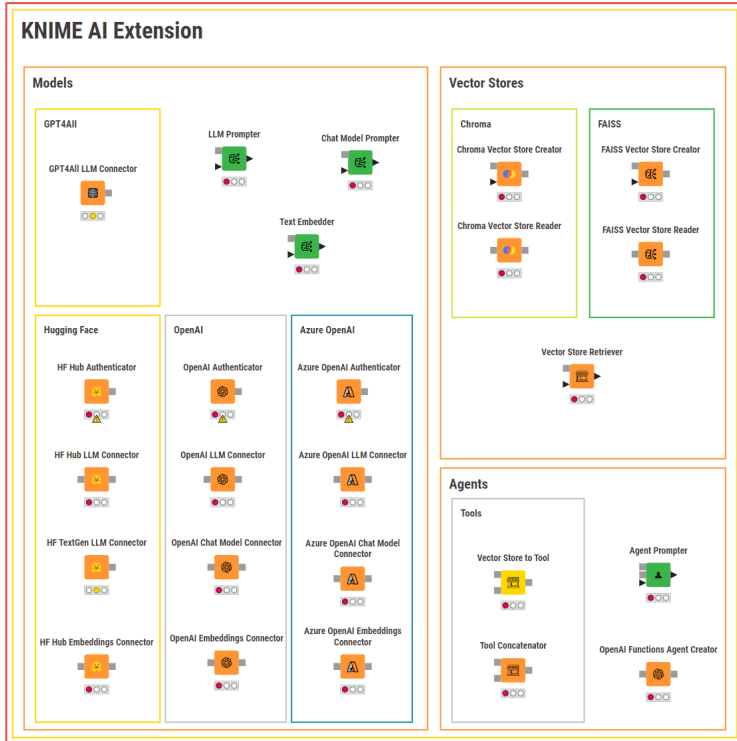
Using & Customizing GenAI

Managing and Governing GenAI

Summary

# Using and Customizing

## GenAI Extensions



## Use & Customize

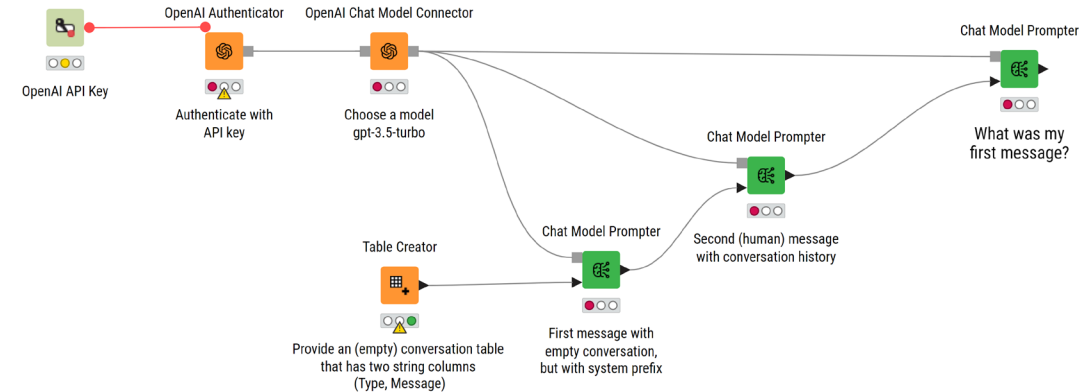
Augment Analytics

Fine-tune & Tailor

Safeguard

...

Credentials Configuration



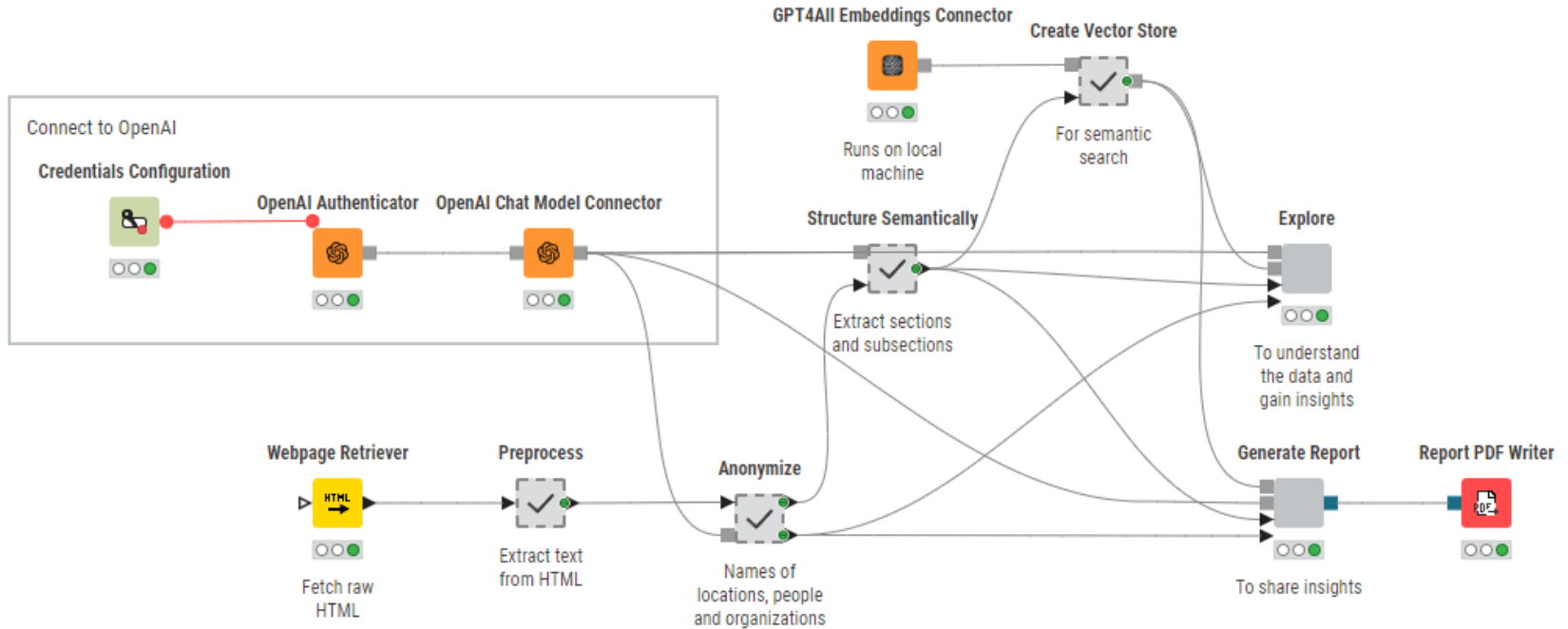
Public space

## AI Extension Example Workflows

Home

- 1) Large Language Models
- 2) Chat Model
- 3) Vector Stores
- 4) Agents

# Using the AI Extension



# Agenda

KNIME and GenAI

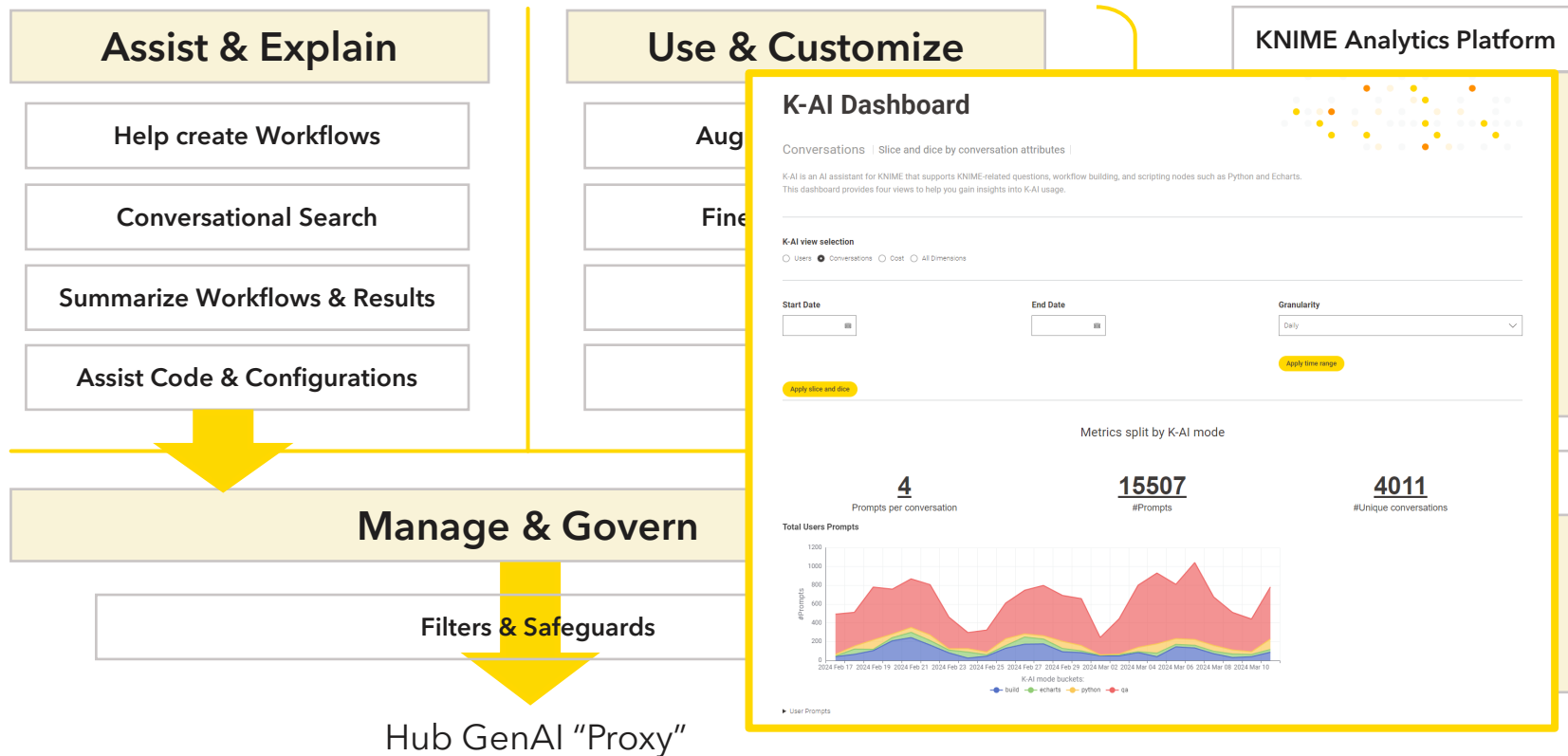
GenAI to Assist & Explain

Using & Customizing GenAI

Managing and Governing GenAI

Summary

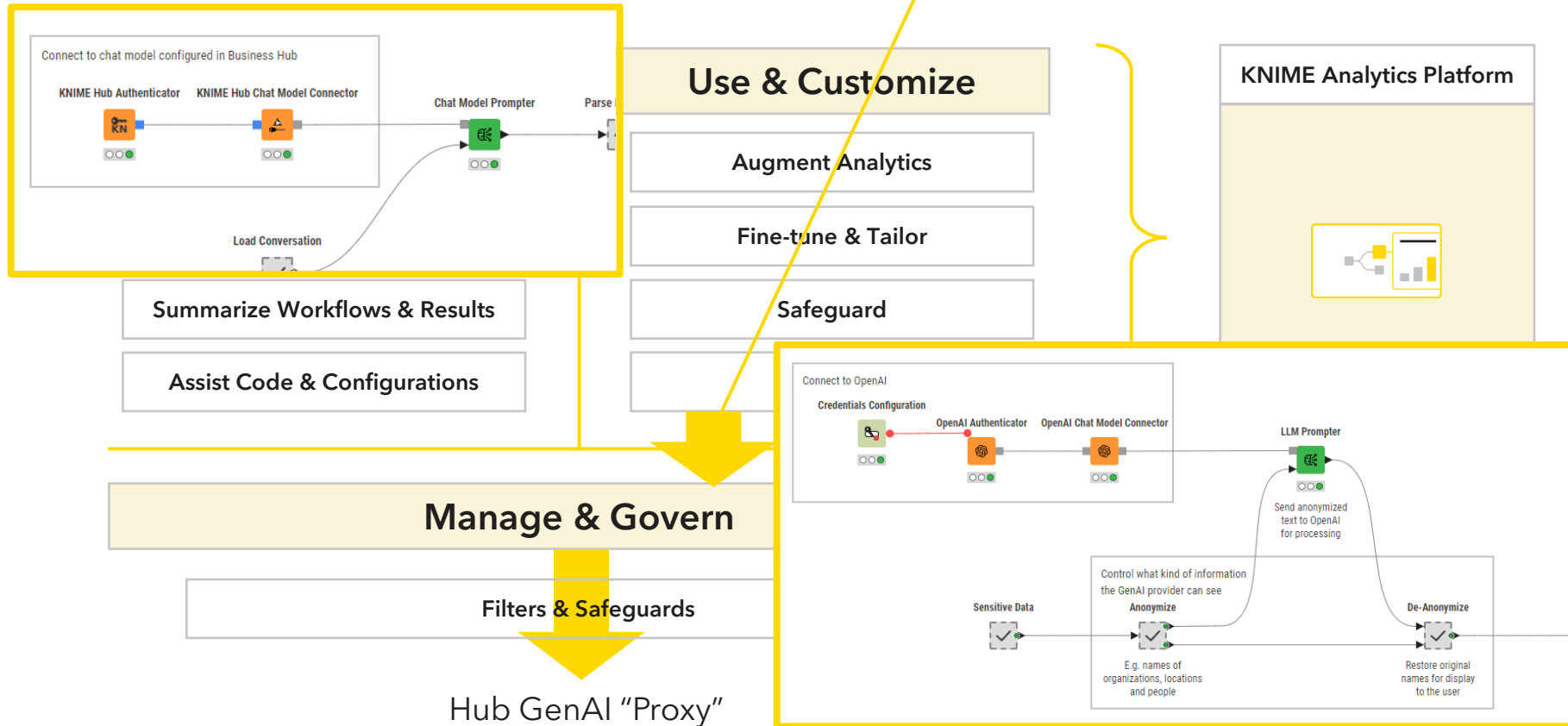
# Managing and Governance





# Managing and Governance

Hub GenAI Connector



# Managing and Governance

## Assist & Explain

Help create Workflows

Conversational Search

Summarize Workflows & Results

Assist Code & Configurations

## Use & Cust

Augment Ana

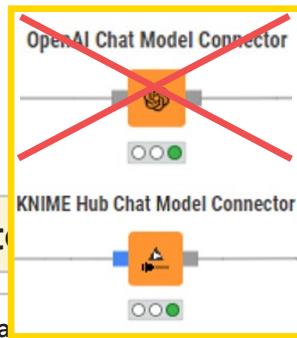
Fine-tune & Tailor

Safeguard

...

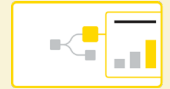
## Manage & Govern

Filters & Safeguards



- Hub AP Customizations
- Shared Components

## Anonymized Customer Data



## KNIME Business Hub



# Agenda

KNIME and GenAI

GenAI to Assist & Explain

Using & Customizing GenAI

Managing and Governing GenAI

Summary

# Data Science and GenAI

Proper GenAI Orchestration...

...enables Flexible Use for Everybody and

...provides Governance

- GenAI for Everybody: KNIME Analytics Platform
  - helps Workflow Creators
  - helps (KNIME and) Data Science Learners
  - augments KNIME Workflows
  - ...and KNIME Workflows augment & safeguard GenAI
- Governance Layer: KNIME Hub
  - low-code workflows provide documentation and auditability
  - workflow validation enables compliance checks
  - enforces use of approved AIs (inhouse / cloud)
  - ensures compliance with privacy, ethical, and other requirements



“Did you try Natural Intelligence?”

(sometimes smart engineering can help avoid (Gen)AI altogether)