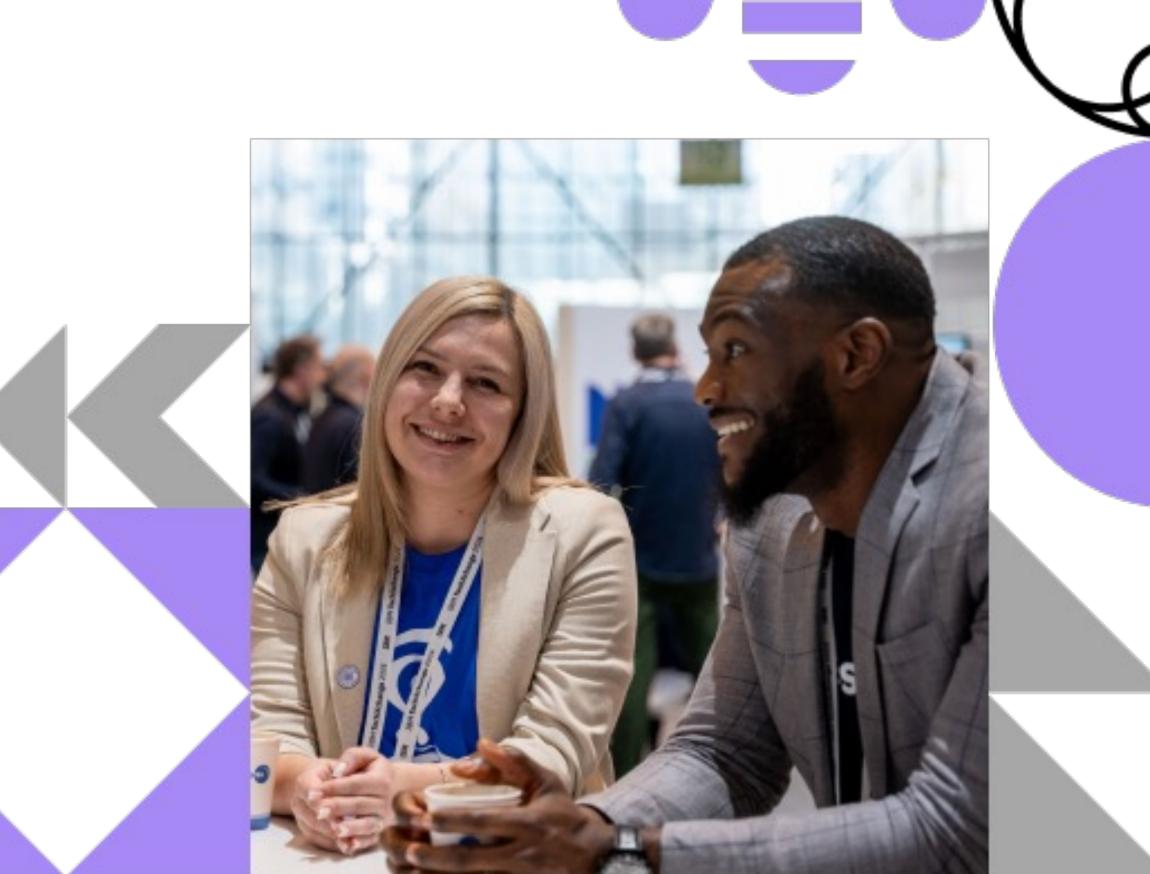
watsonx Day: Challenge





What is the watsonx Day: Challenge?

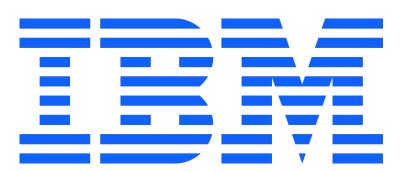
- A no-code/low-code event open to professionals interested in building their AI skills through an enterprise-focused challenge
- A learning opportunity with real world use-cases presented in a fun and engaging format
- Engagement of both technical and non-technical employees to raise awareness on watsonx capabilities

What to expect

- On April 10th the worldwide challenge will start at
 9:00 AM JST (GMT +9) and will end on April 11th at
 12:00 PM JST (GMT +9)
 - All required info to complete the challenge will be provided at 9:00 AM JST (GMT +9)
- You can work virtually or join us at an in-person event
- You can complete any time during the challenge hours
- Support will be available during the event if you have any questions

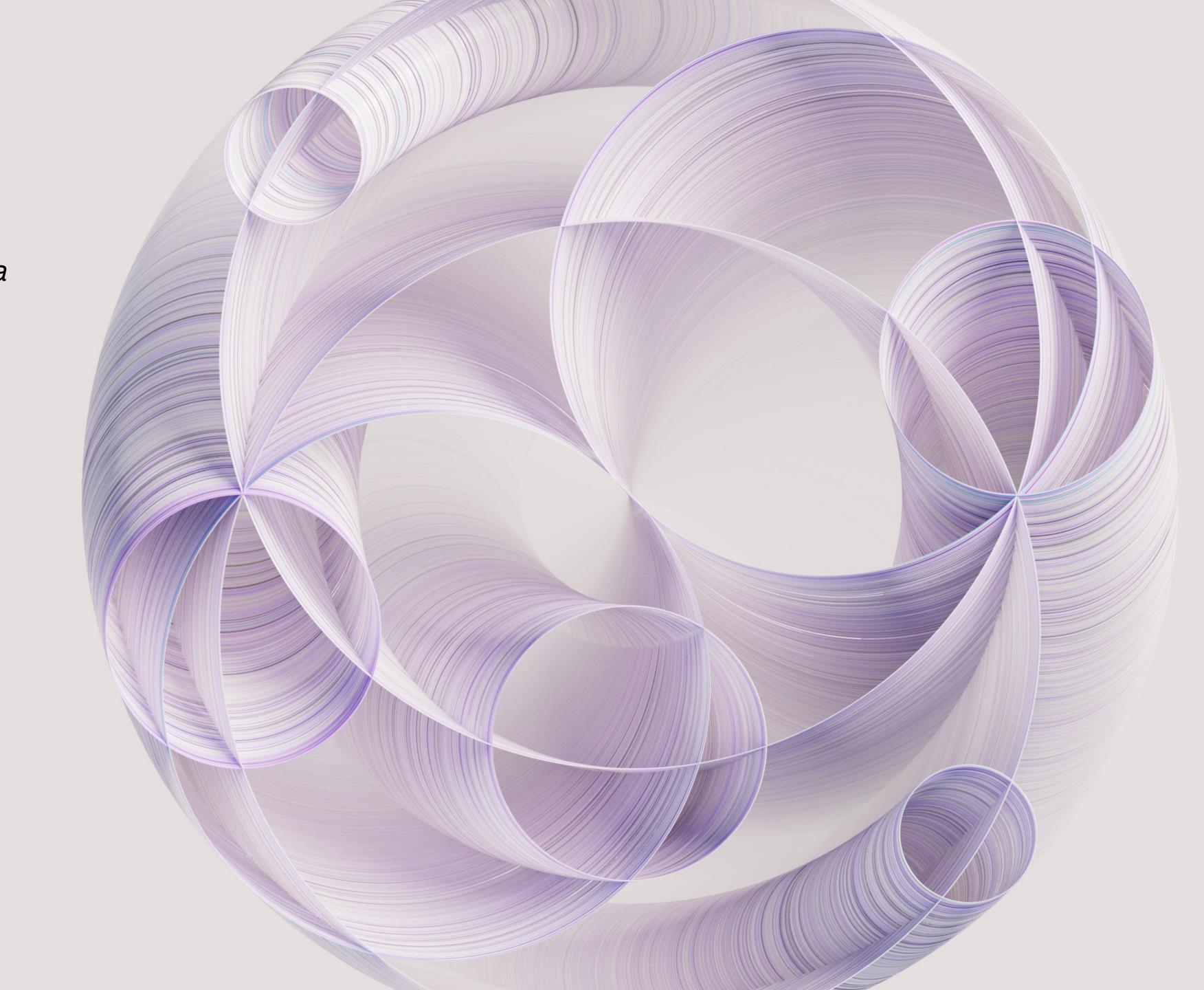
What will you do?

- You will be able to work individually or join a team of up to 4 people
- Complete a challenge track exercise
- Submit your track results and a recommendation
- Top submissions will win prizes



watsonx.ai

Explore the Power of Gen Al and gain a competitive edge using IBM's watsonx platform



Agenda

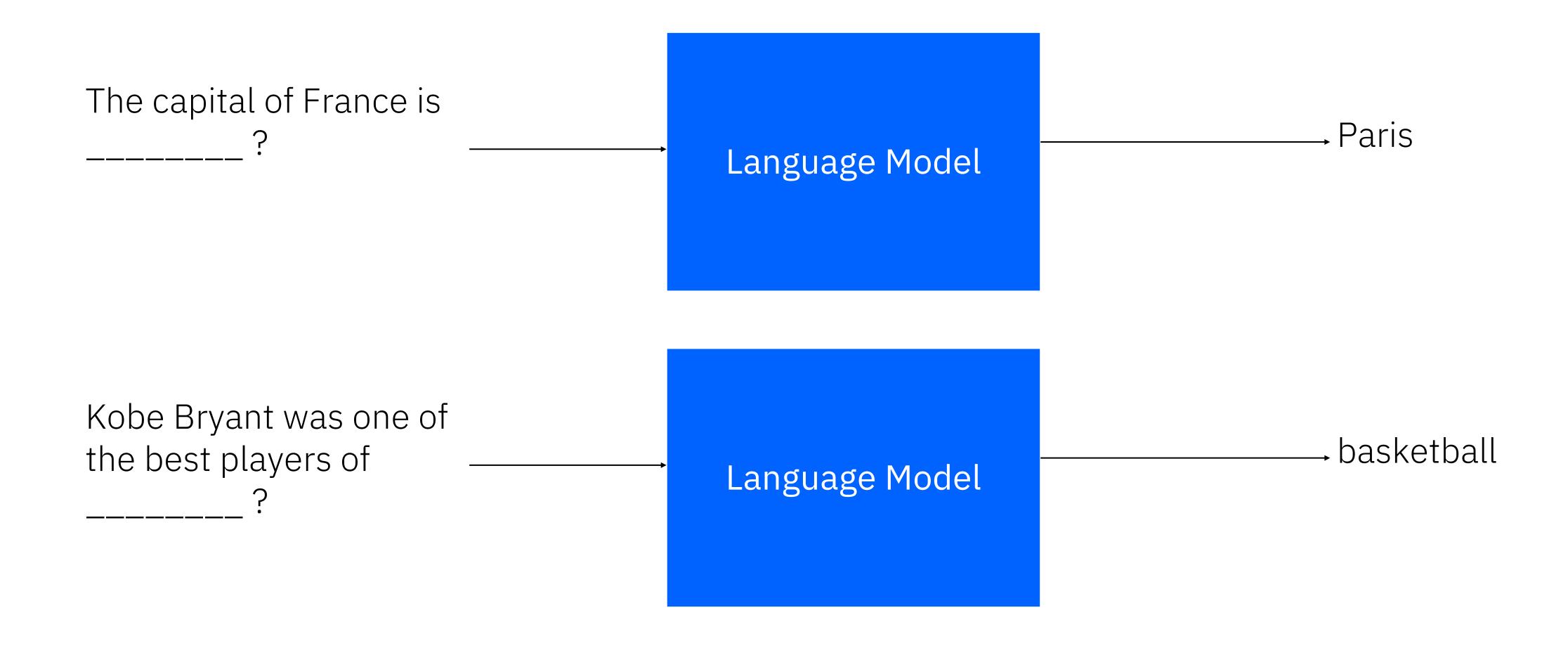
- 1. Understanding Gen AI and Foundation Models
- 2. Introduction to watsonx.ai
- 3. Integrating watsonx Assistant with Watson Discovery and watsonx.ai
- 4. Demo Prompt Lab, Assistant, Discovery
- 5. watsonx Orchestrate

Foundation Models and Generative Al are bringing an inflection point in Al...

...but how enterprises adopt and execute will define whether they unlock, create value, unleash innovation at scale and with speed

What is a Language Model?

Language models predict the next word in a sequence given the words have appeared before.



Generative AI and traditional AI

Both traditional AI and generative AI are useful for enterprises.

Neither replaces the other, generative AI opens new possibilities

Generative AI

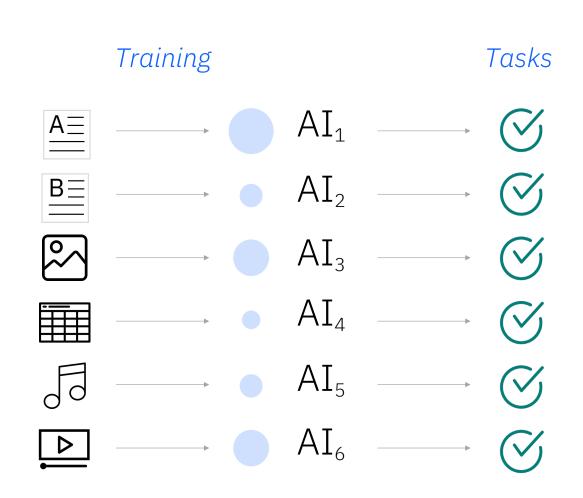
- Foundation models trained with unlabeled data
- Unsupervised
- Trained on very big data sets
- No specific task
- Transferable
- Works well for general tasks and can improve for specific tasks with less training
- Need to monitor bias and drift

Traditional AI

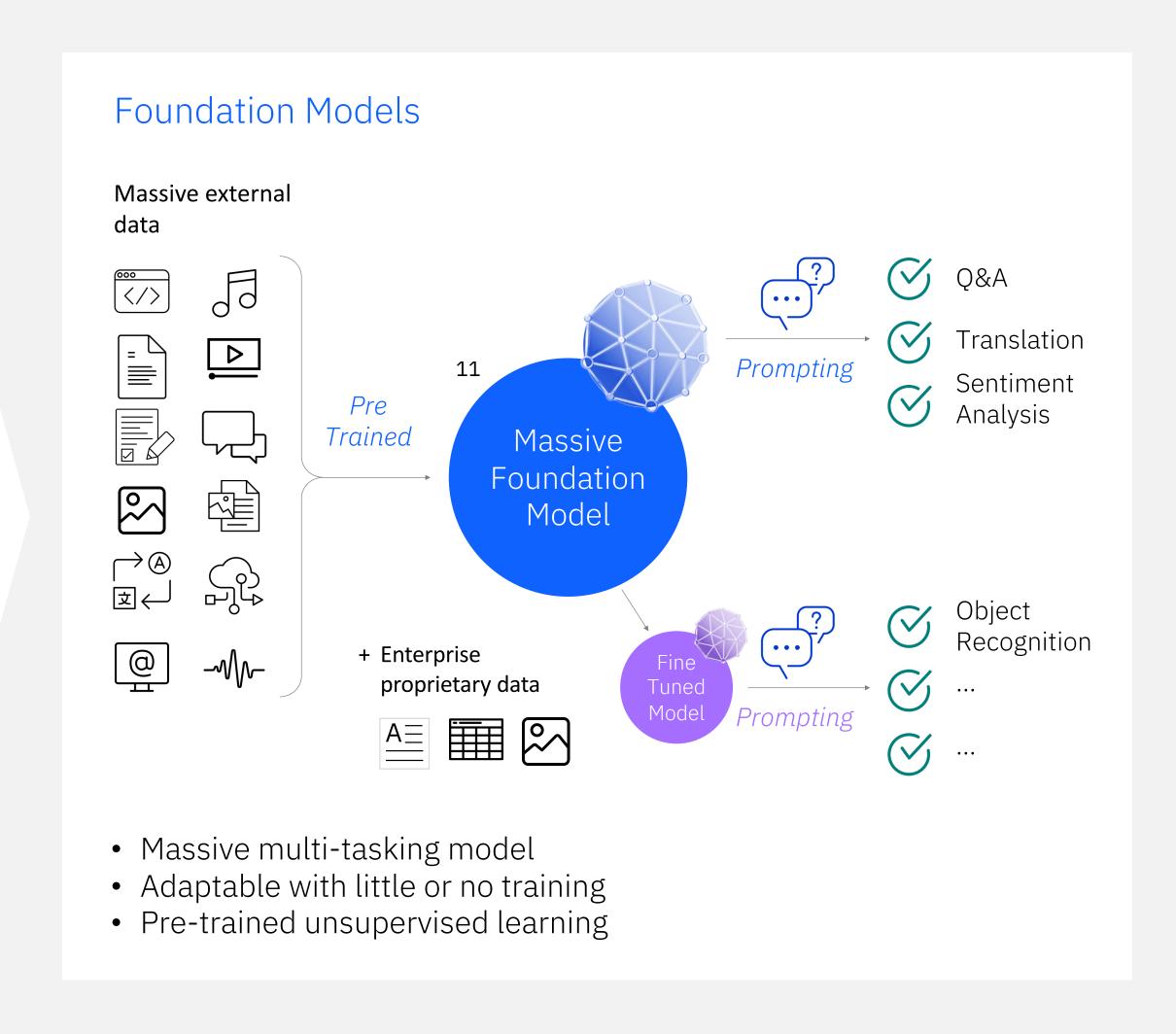
- Traditional Machine learning (ML/AI)
 model trained with "labeled" data
- Training is supervised
- Trained on proper, large data sets
- Trained for a specific task
- Does not transfer well to other tasks
- A tuned model can be very efficient for the specific task it was designed for
- Need to monitor bias and drift

Foundation models establish a new paradigm for AI capabilities

Traditional AI models



- Individual siloed models
- Require task specific training
- Lots of human supervised training



Enhanced capabilities

- Summarization
- Conversational Knowledge
- Content Creation
- Code Co-Creation

Key advantages

- Lower upfront costs through less labeling
- Faster deployment through fine tuning and inferencing
- Equal or better accuracy for multiple use cases
- Incremental revenue.
 through better performance

up to 70% reduction in certain NLP tasks

Put AI to work with watsonx.

The platform for AI and data

watsonx

Scale and accelerate the impact of AI across your business

watsonx.ai

Build, train, validate, tune and deploy AI models

A next generation enterprise studio for AI builders to build, train, validate, tune, and deploy both traditional machine learning and new generative AI capabilities powered by foundation models. It enables you to build AI applications in a fraction of the time with a fraction of the data.

watsonx.data

Scale AI workloads, for all your data, anywhere

Fit-for-purpose data store, built on an open lakehouse architecture, supported by querying, governance and open data formats to access and share data.

watsonx.governance

Accelerate responsible, transparent and explainable AI workflows

End-to-end toolkit for AI governance across the entire model lifecycle to accelerate responsible, transparent, and explainable AI workflows

watsonx

Scale and accelerate the impact of AI across your business





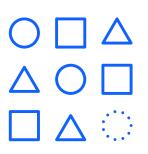
Select and tune AI models for better performance and TCO

- IBM models trained in targeted domains
- Curated library of third-party, open-source models
- Studio to train, validate, tune, and deploy models



Scale AI use cases faster

- Data store to access and share all your data
- Choice of infrastructure, from on premises to the cloud
- AI assistants to accelerate adoption
- Expertise to take AI from pilot to production

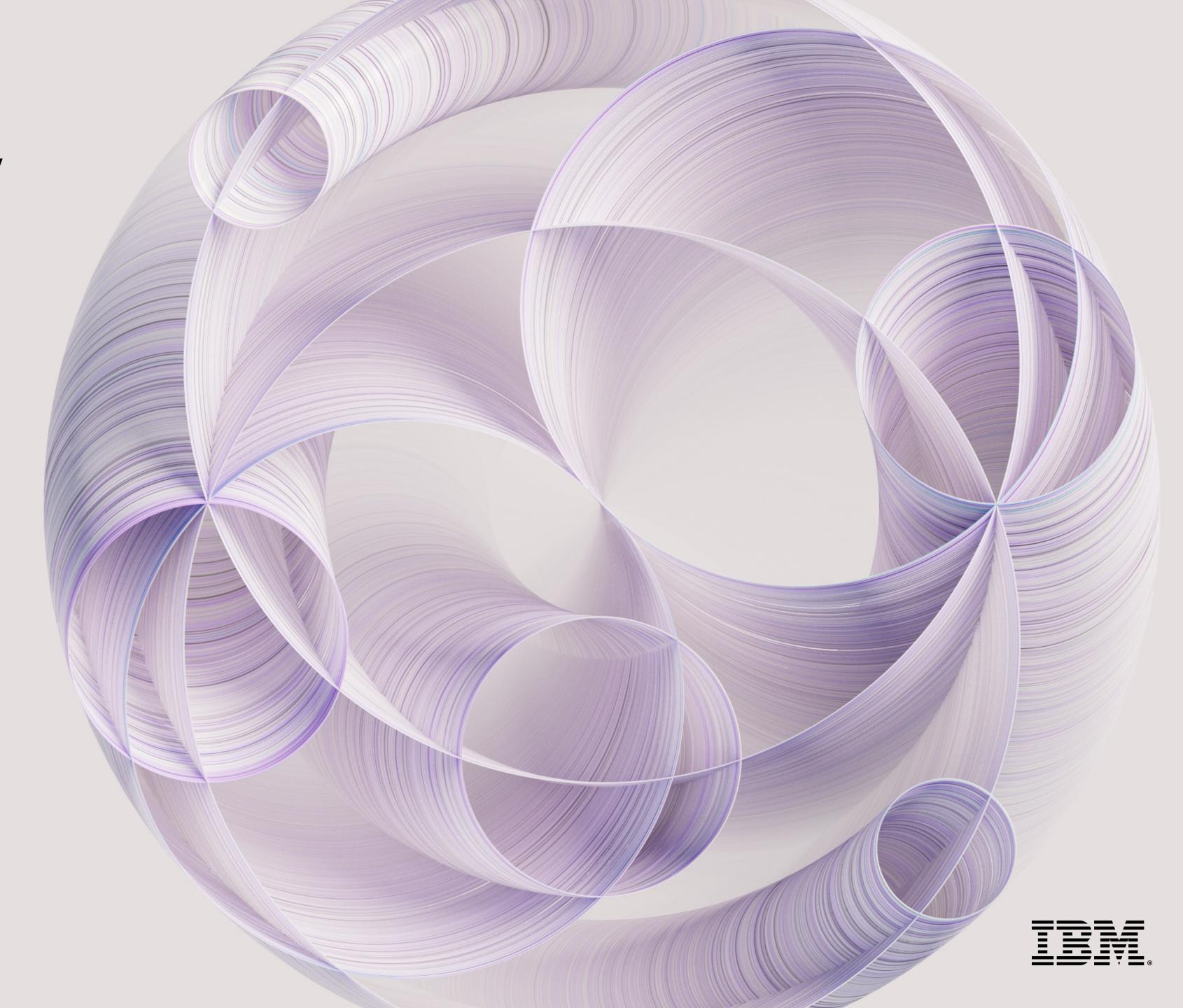


Implement AI with trust

- AI lifecycle governance, from training and tuning to deployment
- Model guardrails and IBM model indemnification
- Data governance, from quality to privacy and compliance

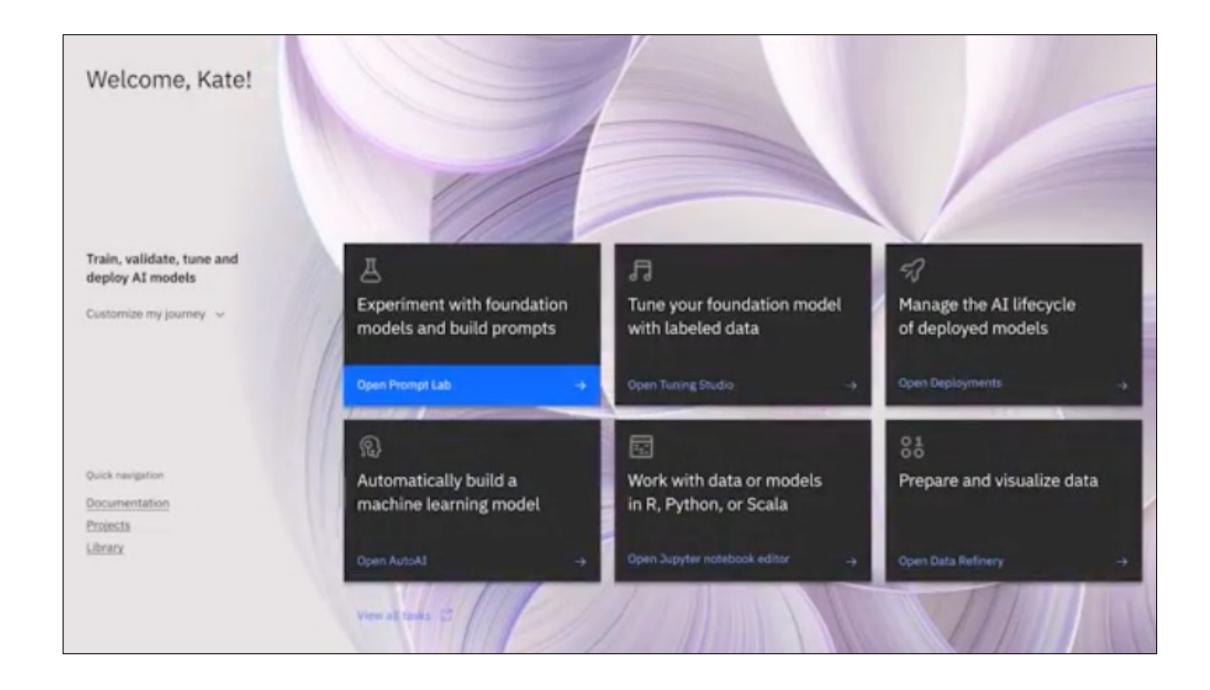
Train, validate, tune and deploy AI models

A next generation enterprise studio for AI builders to train, validate, tune and deploy both traditional ML and new generative AI capabilities, powered by foundation models



watsonx.ai

Build, train, validate, tune, and deploy AI models

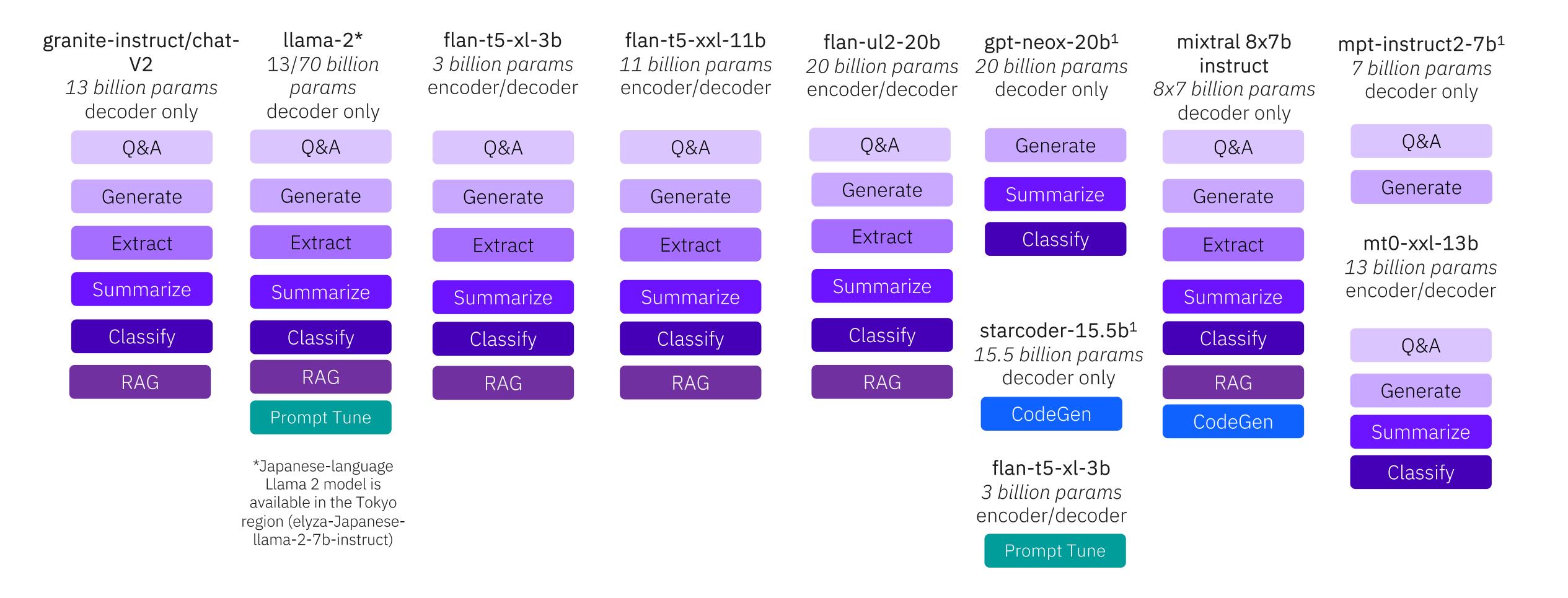


A next generation enterprise studio for AI builders to train, validate, tune, and deploy generative AI, foundation models, and machine learning capabilities.

The watsonx.ai components include:

- Foundation Model Library with IBM and open-source models
- Prompt Lab to experiment with foundation models and build prompts for various use cases and tasks
- Tuning Studio to tune your foundation models with labeled data
- Data Science and MLOps to build machine learning models automatically with model training, development, visual modeling, and synthetic data generation

watsonx.ai – Models available



watsonx.ai and its own foundation models

Model architectures	Architecture name	Use case
Encoder-only	Slate (use through a Watson Studio notebook)	Best cost performance trade-off for non-generative use cases but require task-specific labeled data for fine tuning.
Decoder-only	Granite	Designed explicitly for generative AI use cases; represents the architectures used in GPT-3 and other popular LLMs.
Encoder-decoder	Sandstone To be released	Support both generative and non-generative use cases. Best cost performance trade-off for generative use cases when input is large but generated output is small.

The initial focus is on two categories for each model: language and code. A model's name will identify its characteristics

Model naming convention: <category>/<architecture><opt sub arc>.<size>.<opt info> where the 2 initial categories are fm.code and fm.language or fm.code/<architecture> where fm is the name of the foundation model

Examples:

- fm.language/sandstone.3b is the Encoder-decoder model on language-related tasks with 3b parameters.
- fm.code/granite.350m.ansible is the Decoder model on code-related tasks with 350m parameter for ansible

watsonx.ai: Prompt Lab

Experiment with foundation models and build prompts

Interactive prompt builder

Includes prompt examples for various use cases and tasks

Experiment with different prompts, save and reuse older prompts, use different models and vary different parameters

Experiment with zero-shot, one-shot, or few-shot prompting to get the best results

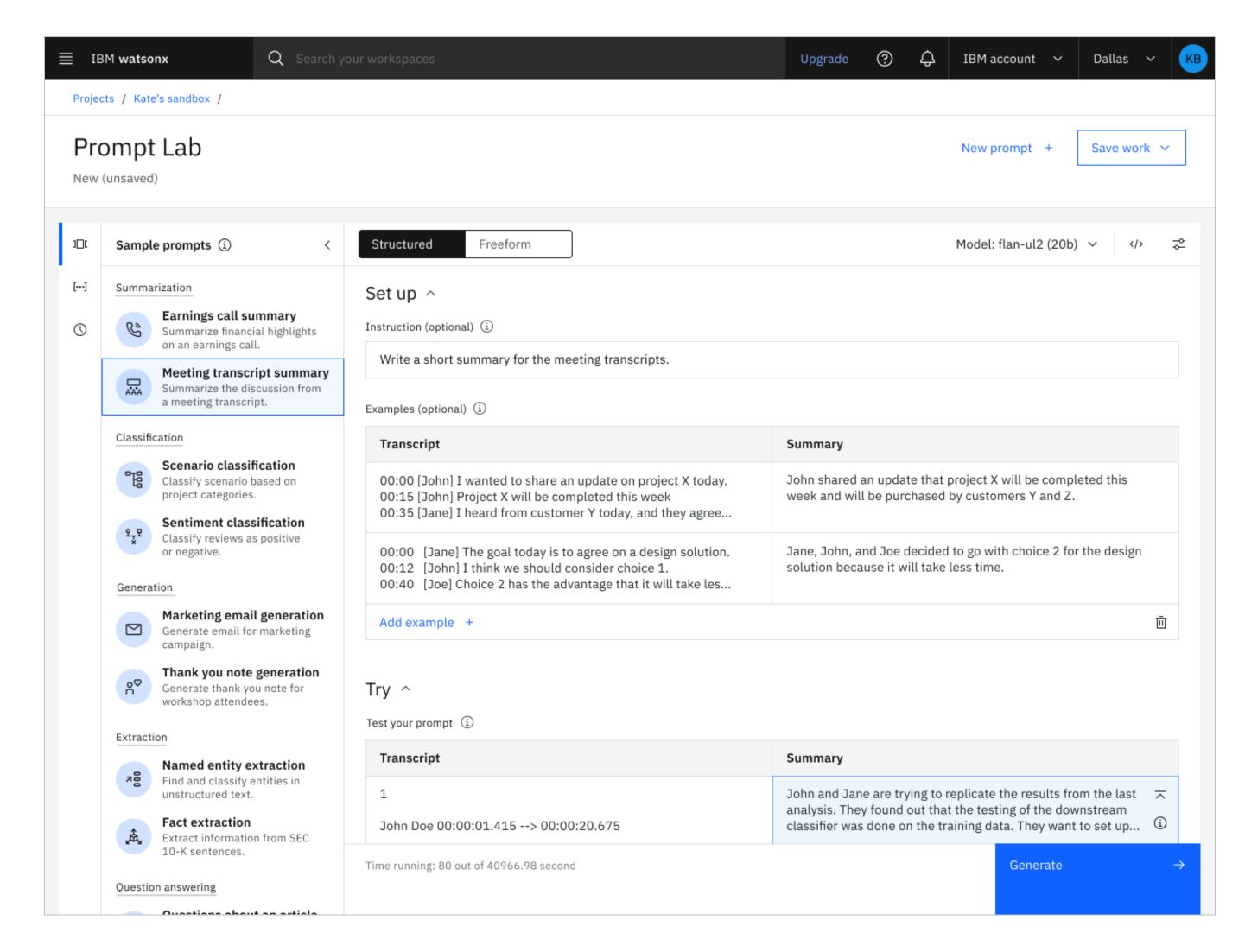
Experiment with prompt engineering

Choice of foundation models to use based on task requirements

Prevent the model from generating repeating phrases

Number of min and max new tokens in the response

Stop sequences – specifies sequences whose appearances should stop the model



Zero/One/Few Shot learning / Prompt Engineering

Foundation models are adaptable. One way to adapt the model is using prompt engineering.

What is prompt engineering?

- Prompt engineering is a new discipline for finding the optimal prompt to use with a foundation model for the best performance
- A prompt is a way to communicate with the foundation model to:

Pass additional instructions on how the model should respond

Feed task-specific context to the model

Zero-shot

```
1 Translate English to French:
2 cheese =>
```

One-shot

```
1 Translate English to French:
2 sea otter => loutre de mer
3 cheese =>
```

Few-shot

```
Translate English to French:

sea otter => loutre de mer

peppermint => menthe poivrée

plush girafe => girafe peluche

cheese =>
```

watsonx.ai Use cases

Retrieval-Augmented Generation (RAG)
Based on a set of documents or dynamic content, create a chatbot or a question-answering feature grounded on specific content. E.g., building a Q&A resource from a broad knowledge base, providing customer service assistance
*note: only possible with watsonx.ai and Watson discovery

2 Summarization Transform text with domain-specific content into personalized overviews, capturing key points E.g., sales conversation summaries, insurance coverage, meeting transcripts, and contract information

Generate text content for a specific purpose.
E.g., content creation for marketing campaigns, job descriptions, blog posts and articles, and email drafting support

4 Named Entity Recognition

Identify and extract essential info

Identify and extract essential information from unstructured text

E.g., audit acceleration, SEC 10K fact extraction

5 Insight Extraction

Analyze existing unstructured text content to surface insights in specialized domain areas.

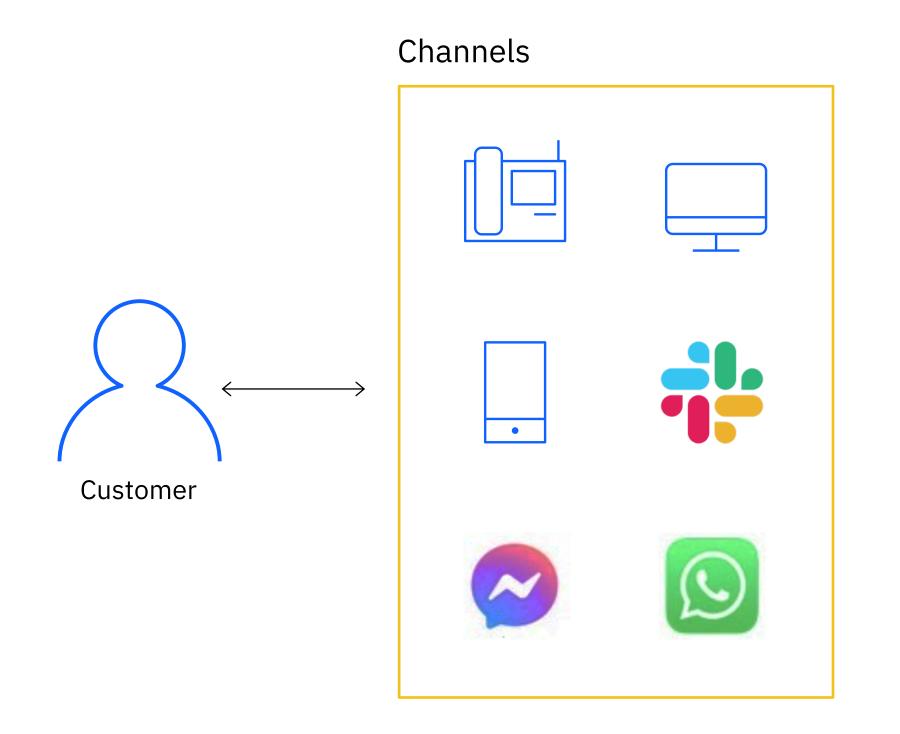
E.g., medical diagnosis support, user research findings

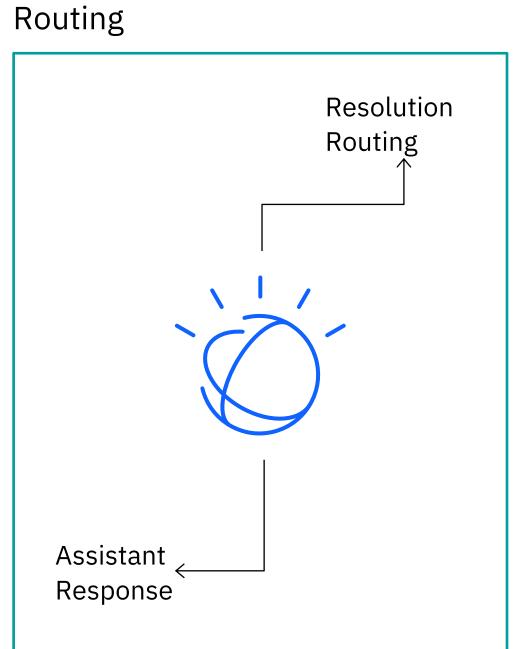
6 Classification

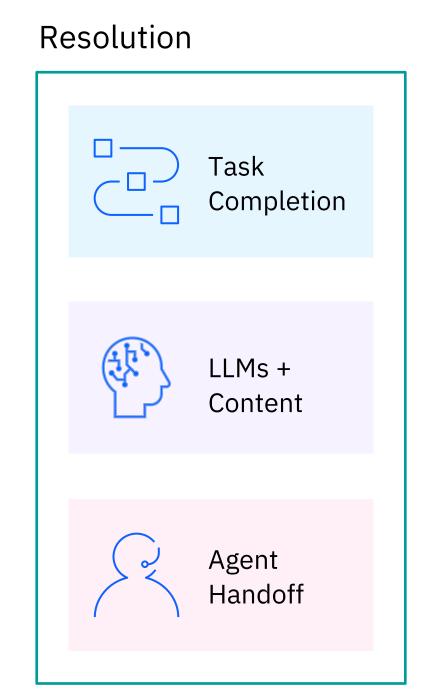
Read and classify written input with as few as zero examples

E.g., sorting of customer complaints, threat & vulnerability classification, sentiment analysis, and customer segmentation

Watson Assistant is IBM's Conversational AI Platform

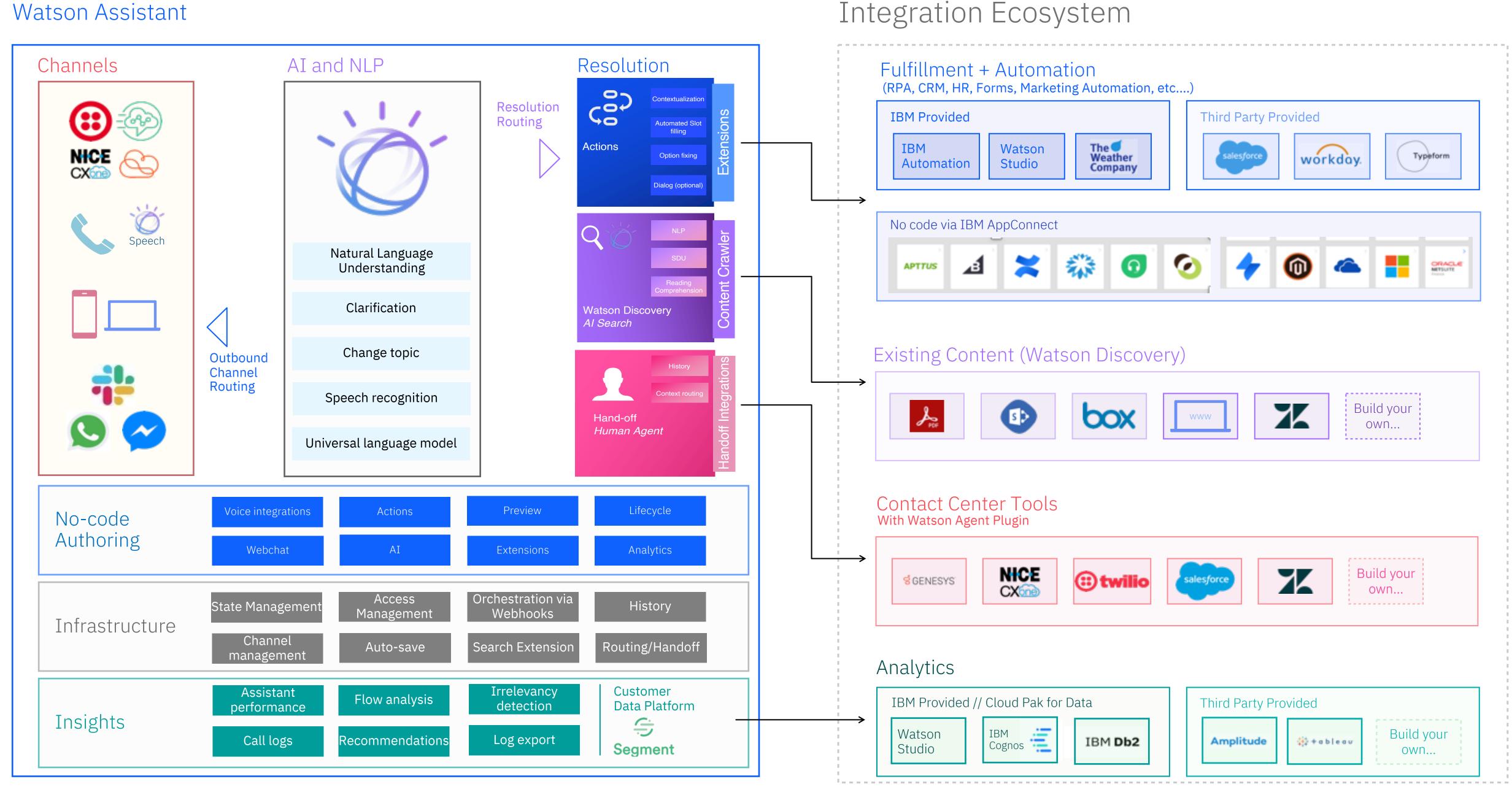




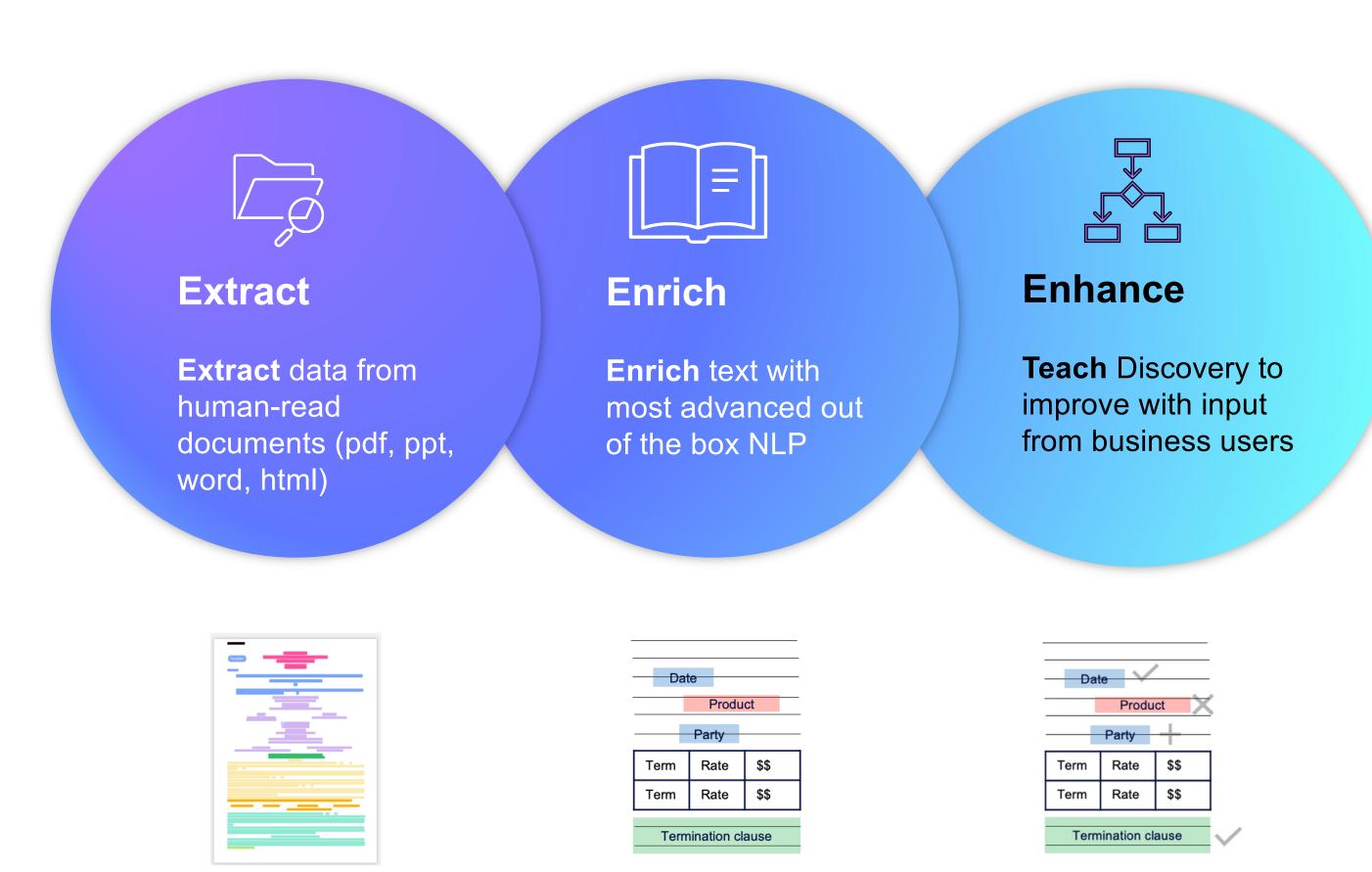


Architecture and Extensibility

Watson Assistant



Watson Discovery

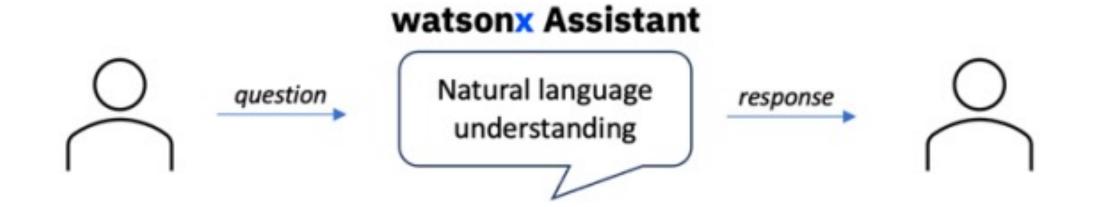


Analyze

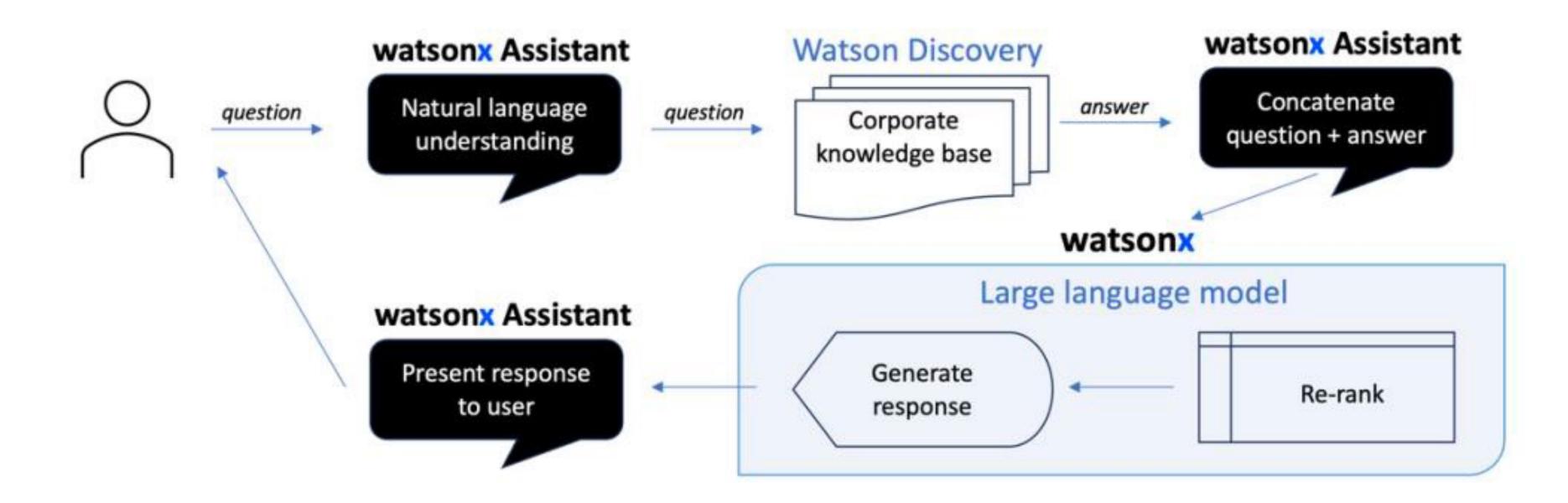
Navigate

Automate

Integrating watsonx Assistant with Watson Discovery and watsonx.ai



No large language model needed.



Data Fabric / March, 2022 / © 2022 IBM Corporation

Call to action

Challenge yourself to deliver a watsonx briefing for each of your clients

- –Join <u>watsonx</u> platform
- —Leverage the <u>watsonx.ai</u> presentation, <u>generative AI whitepaper</u> and <u>watsonx.ai video</u>

Get Prepared - Complete required learning & earn your watsonx.ai sales foundation badge

- AI Fundamentals
- AI for Business
- watsonx.ai Level 2
- watsonx.data Level 2





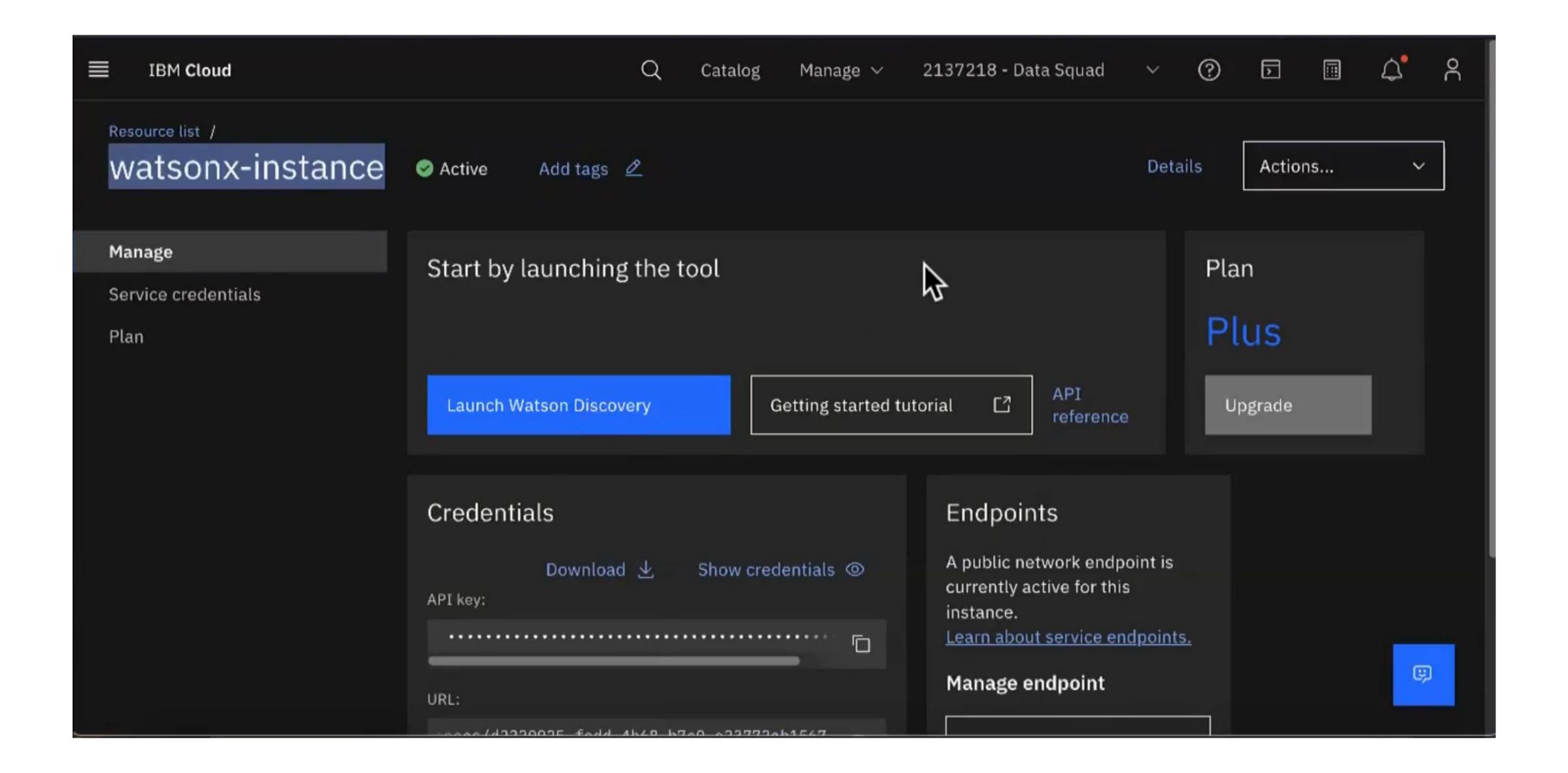
Learn AI Hands-on - Generative AI

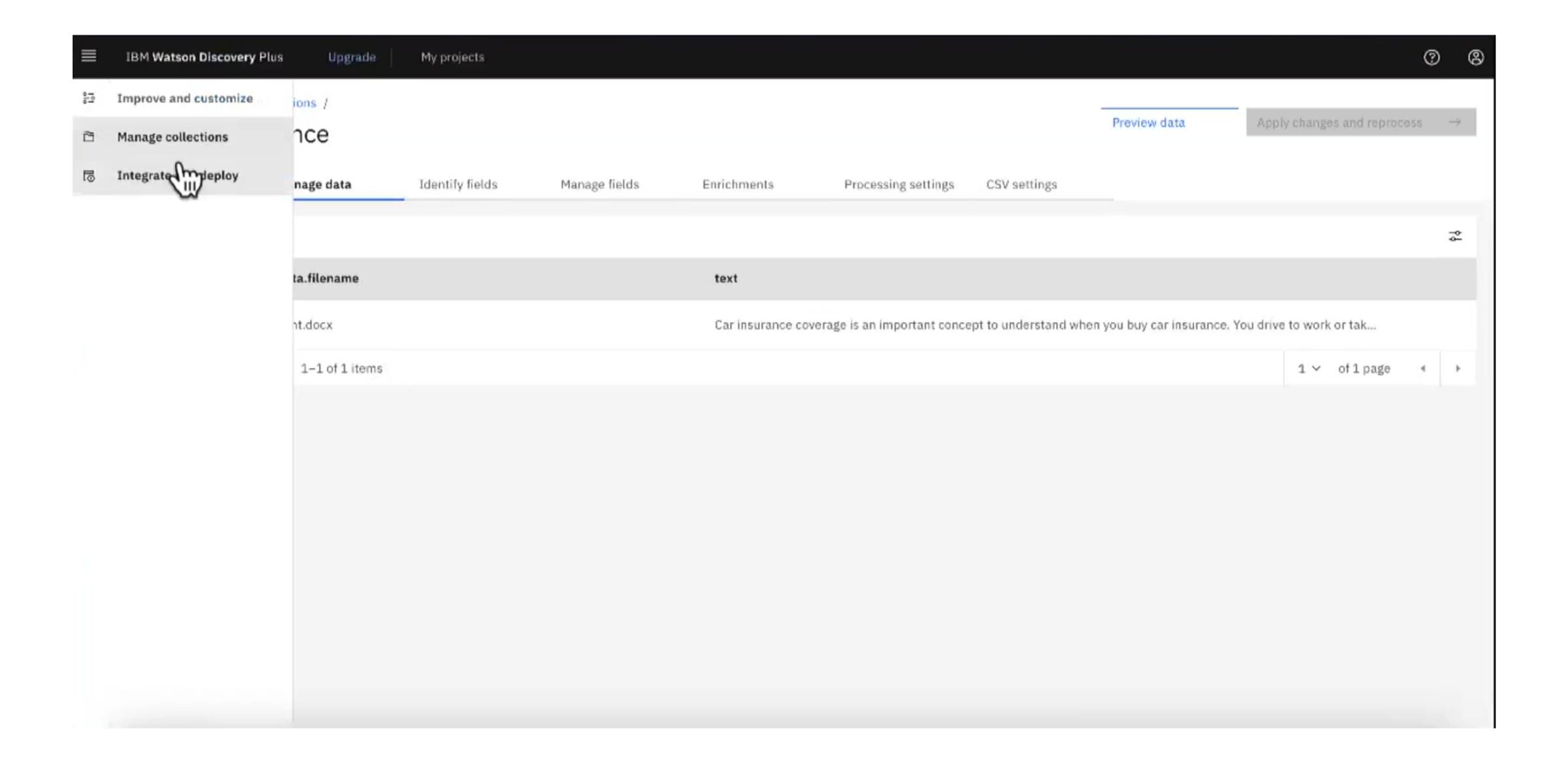
Prompt Engineering for Everyone

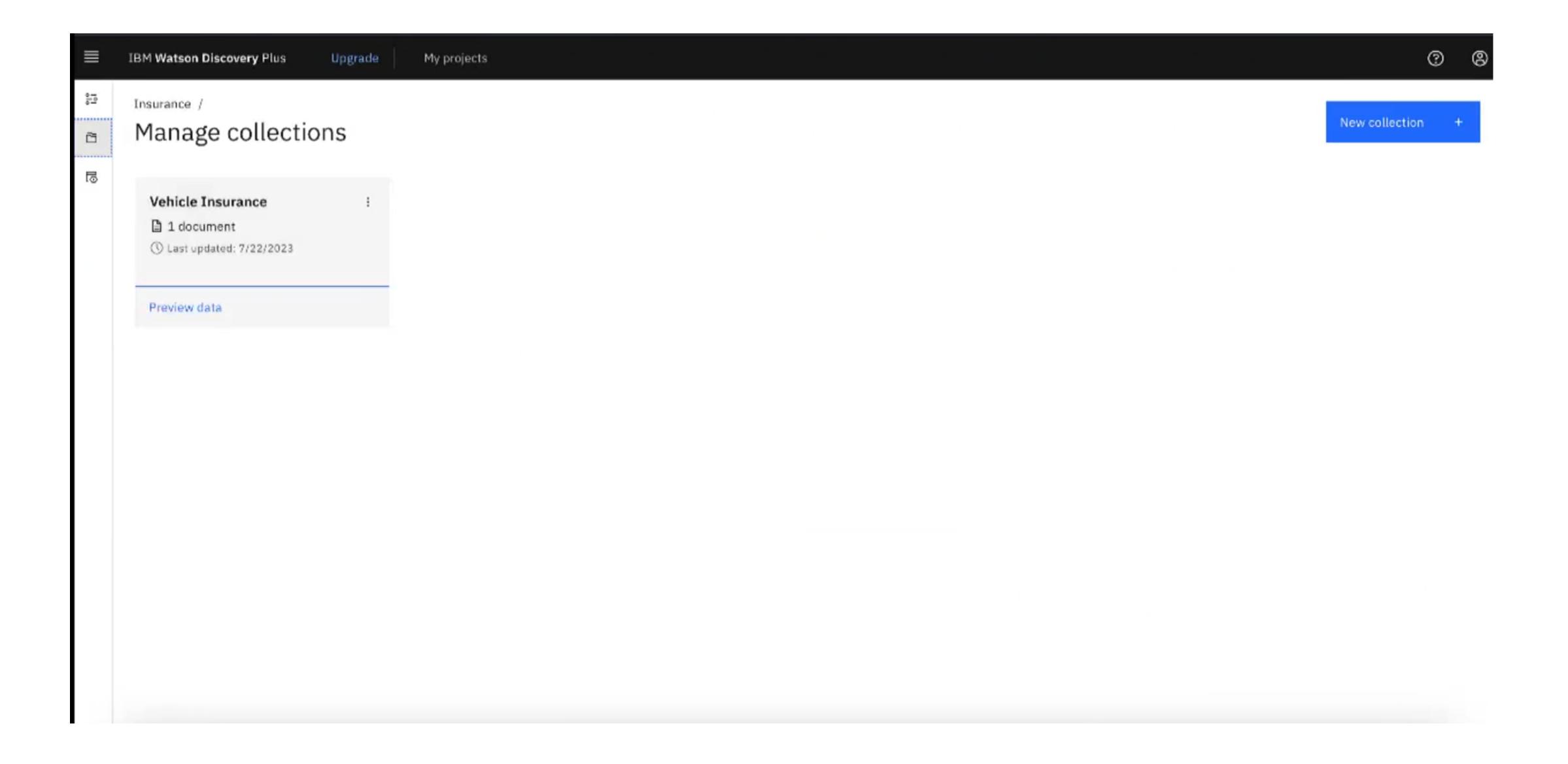
Integrate watsonx Assistant with watsonx.ai

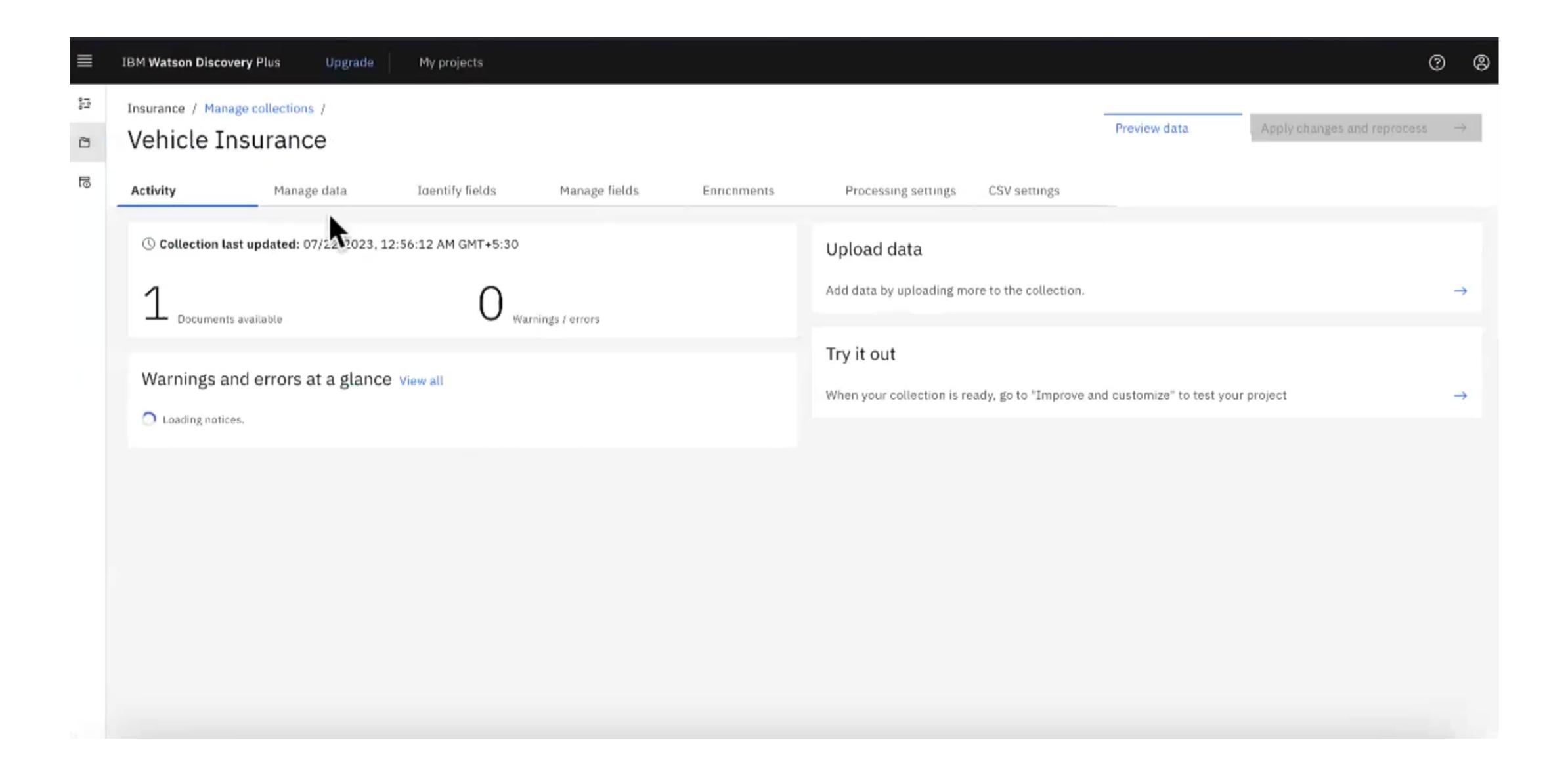
Demo - Backup

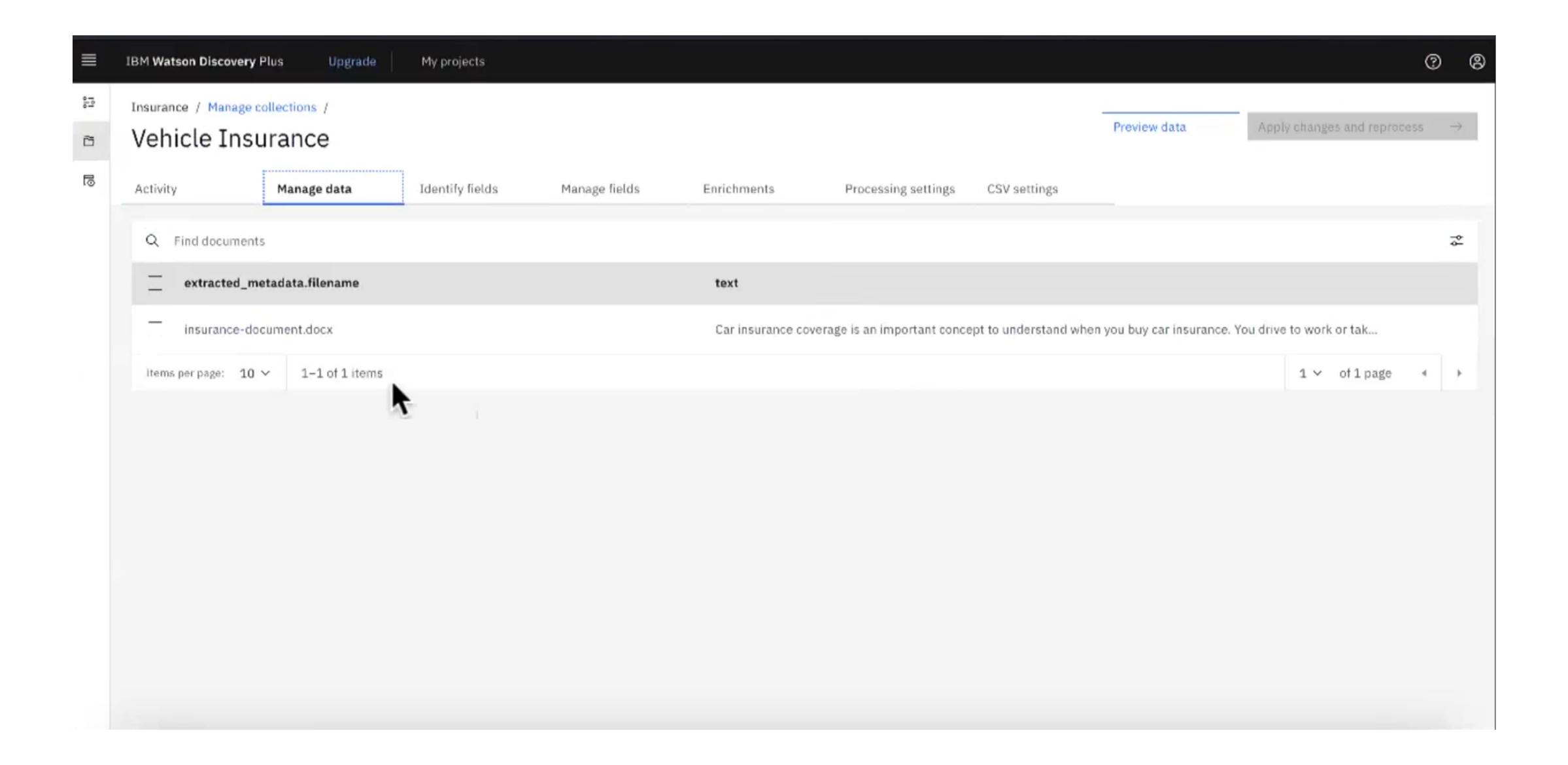
watson Discovery

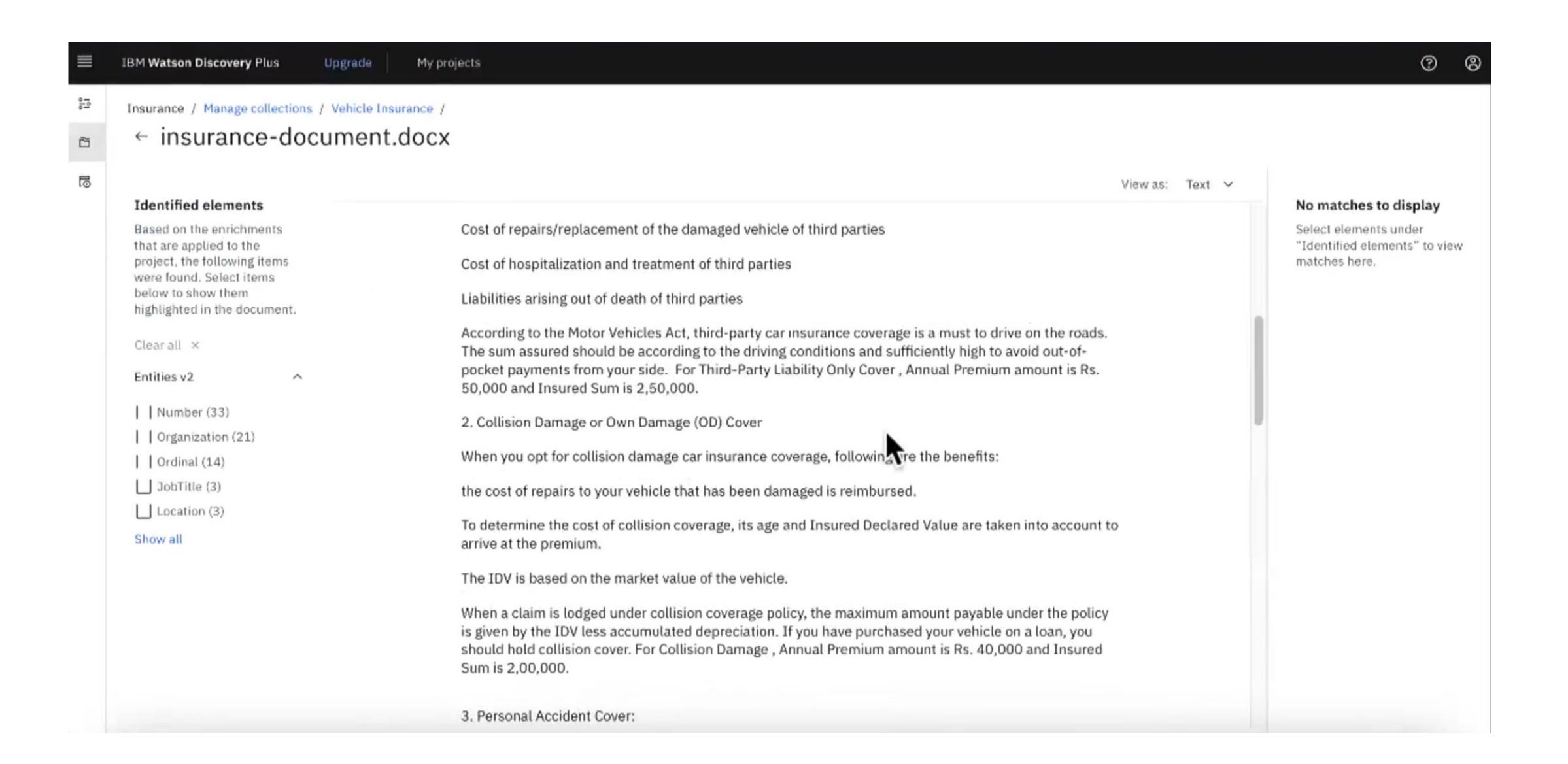




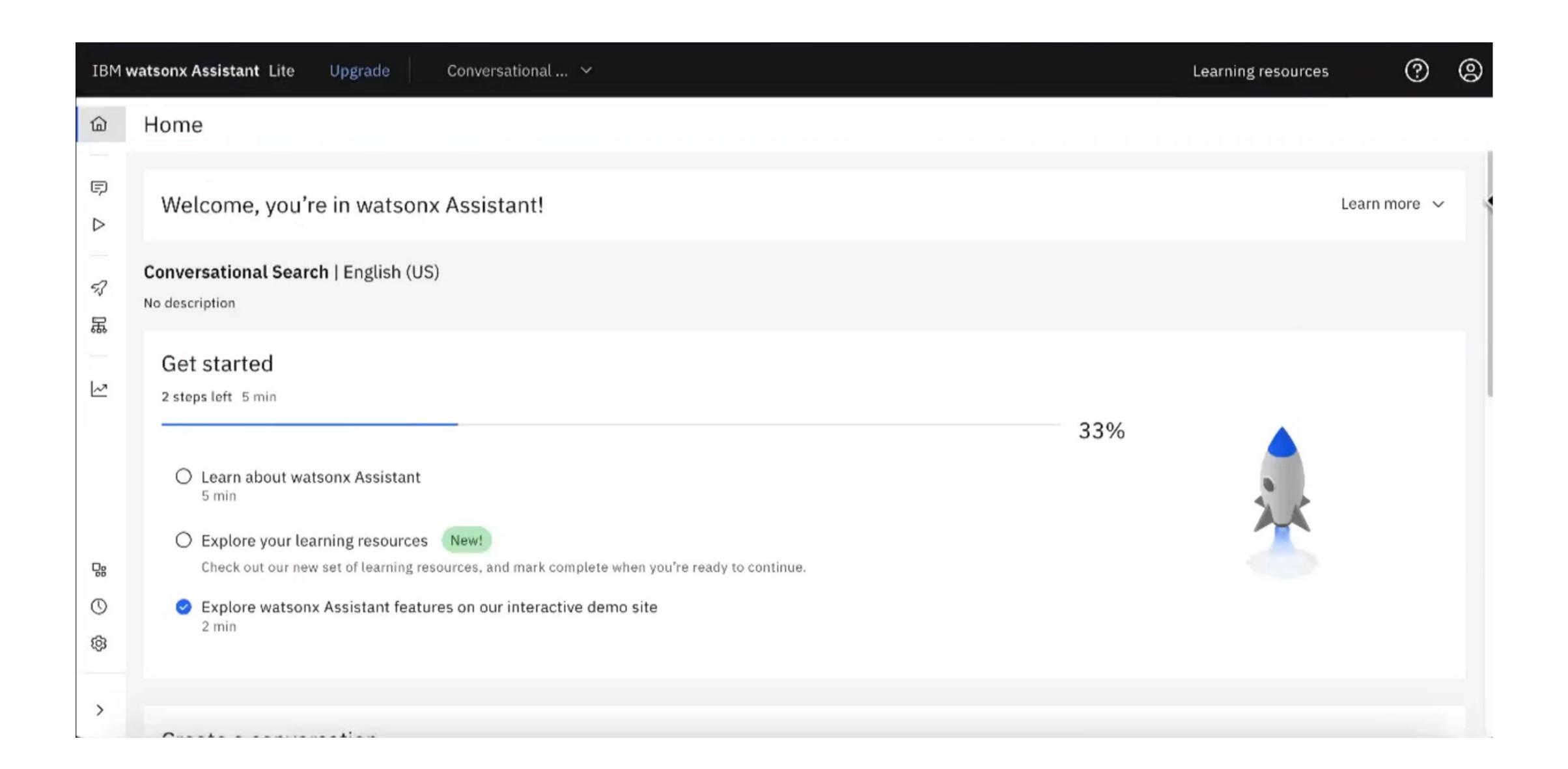


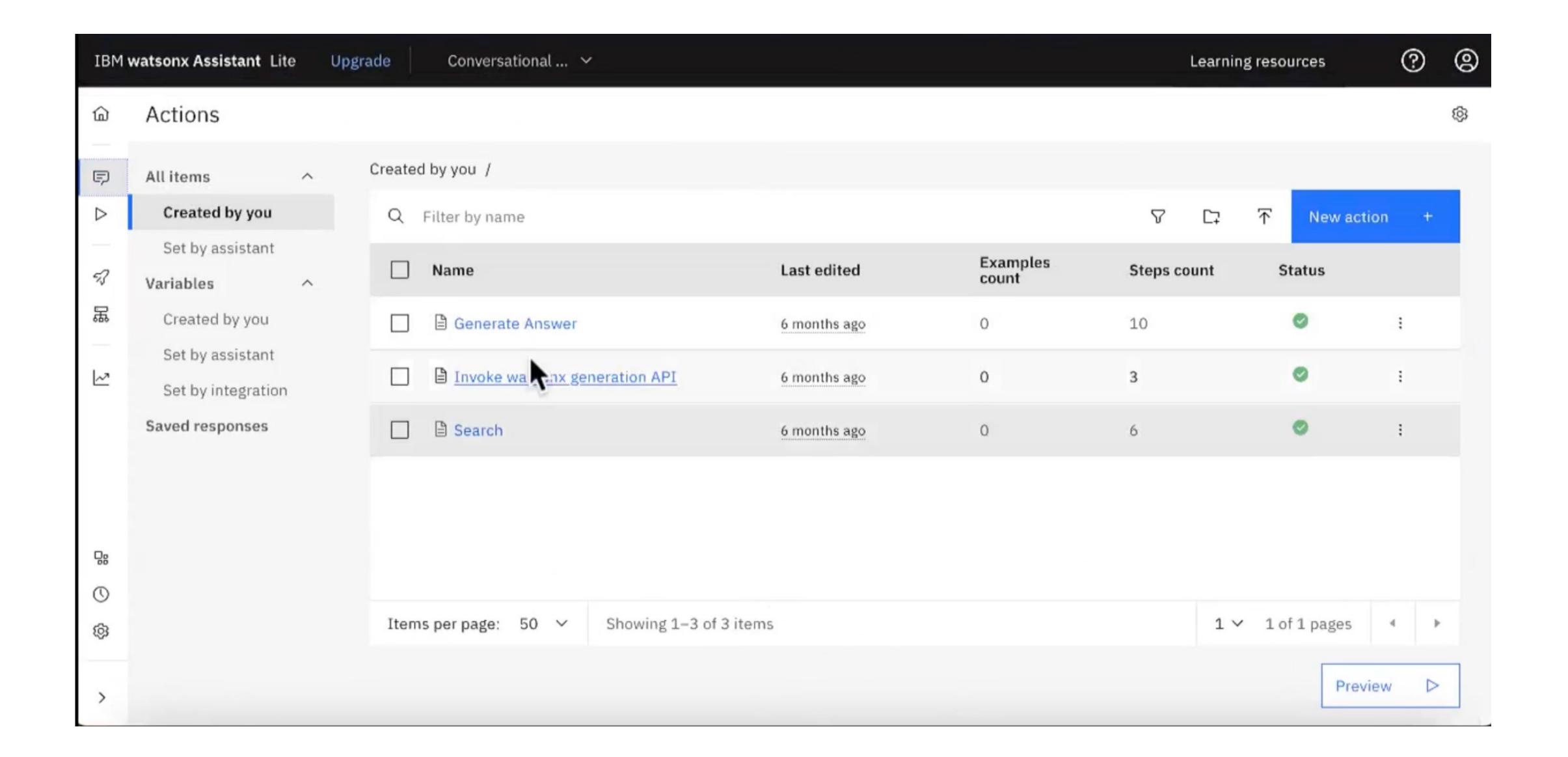


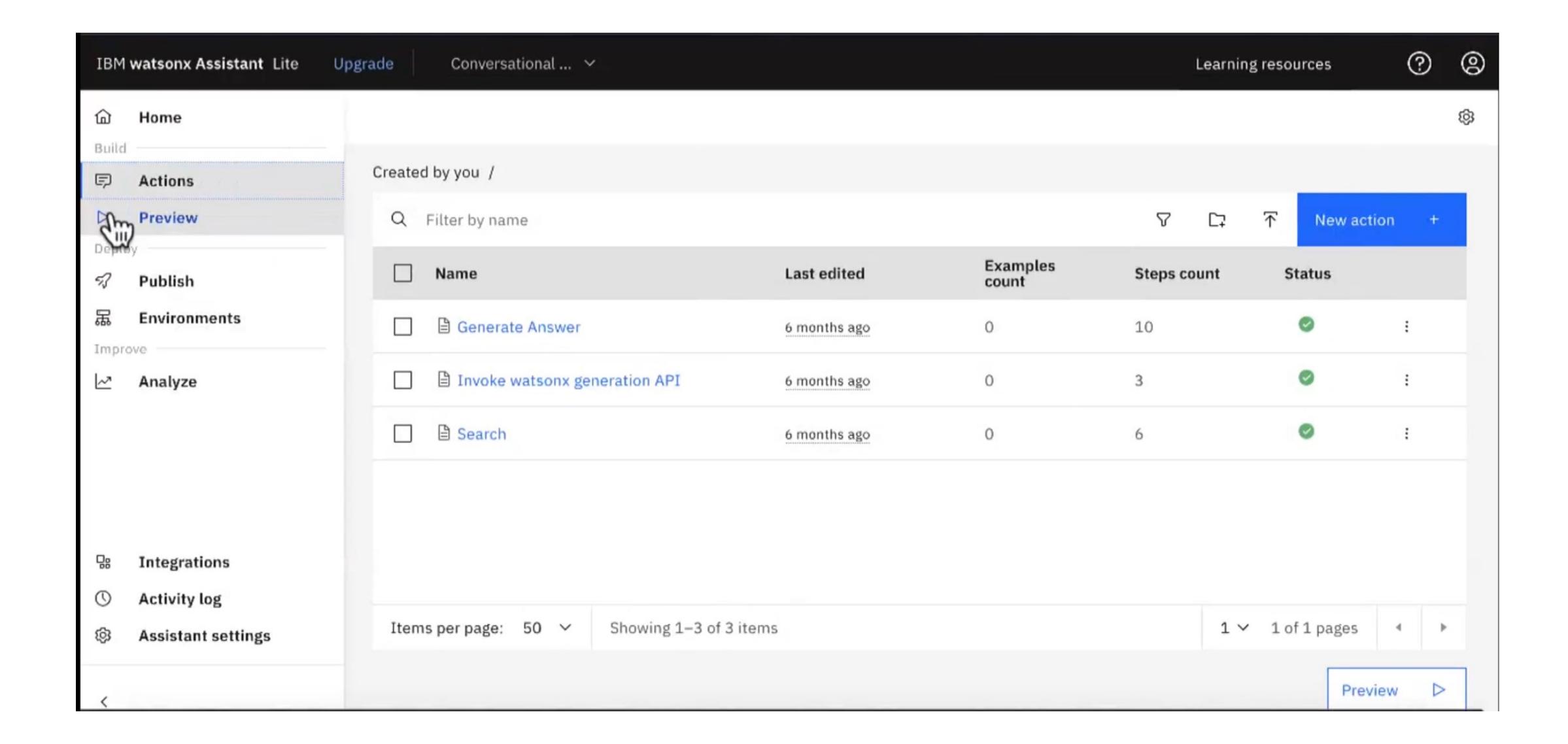


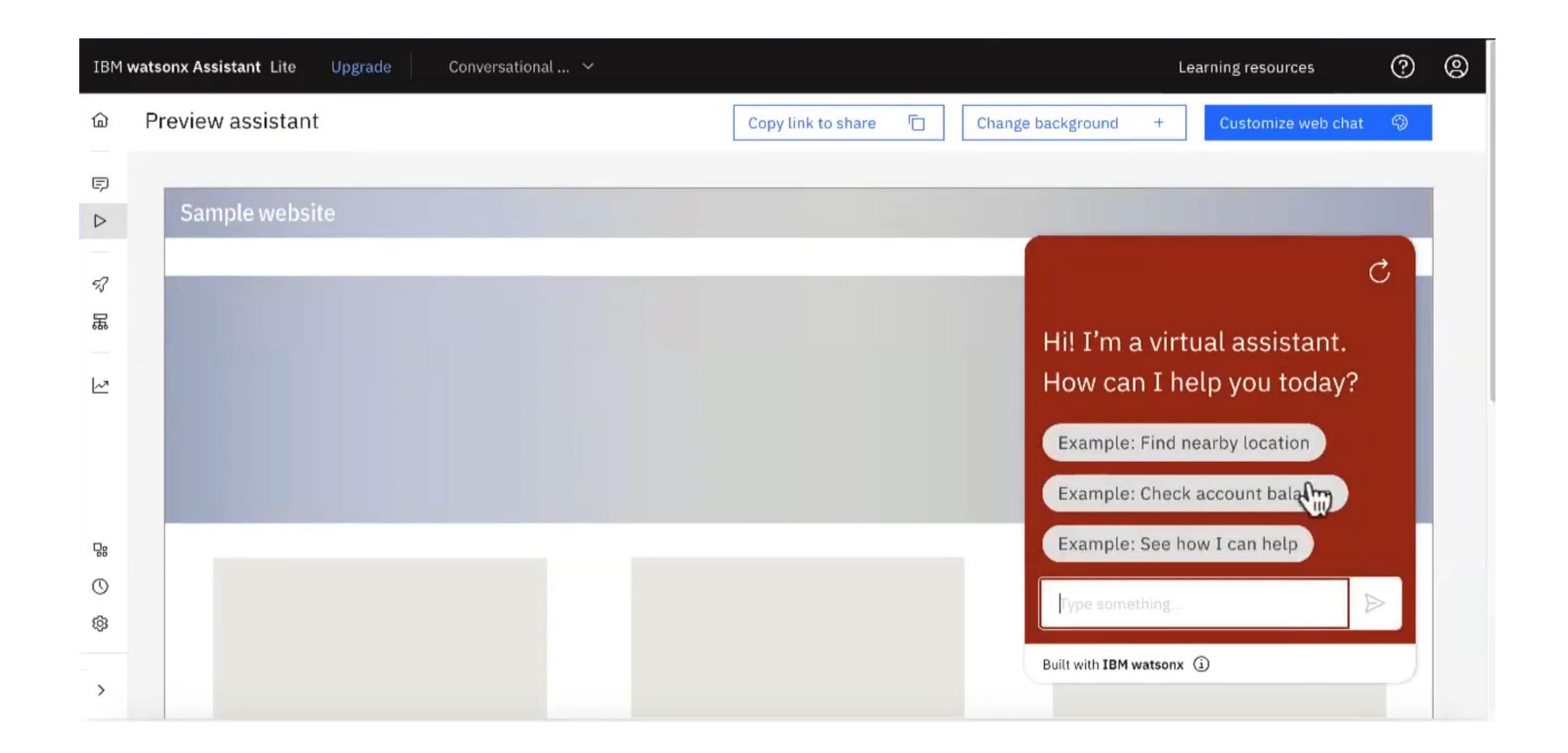


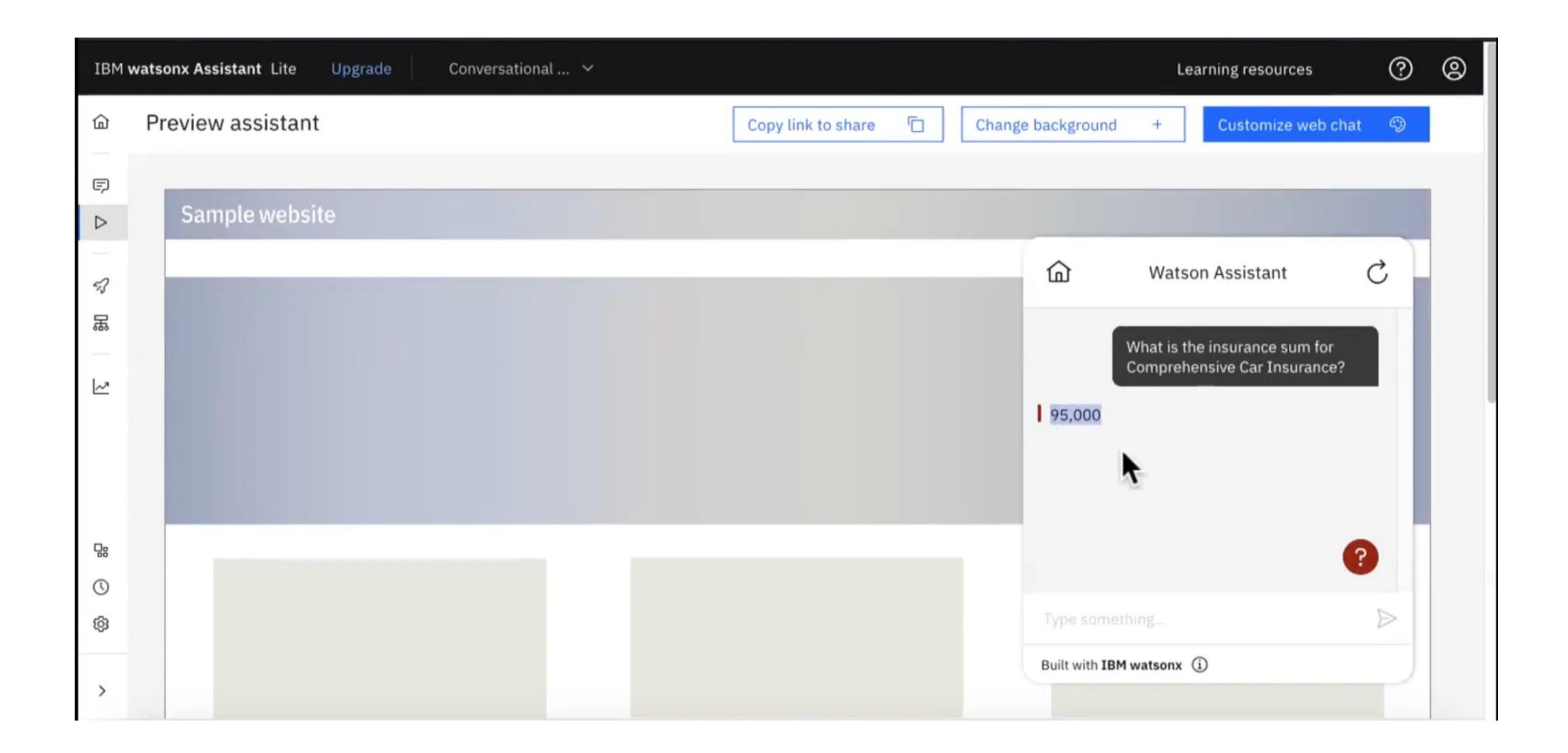
watsonx Assistant

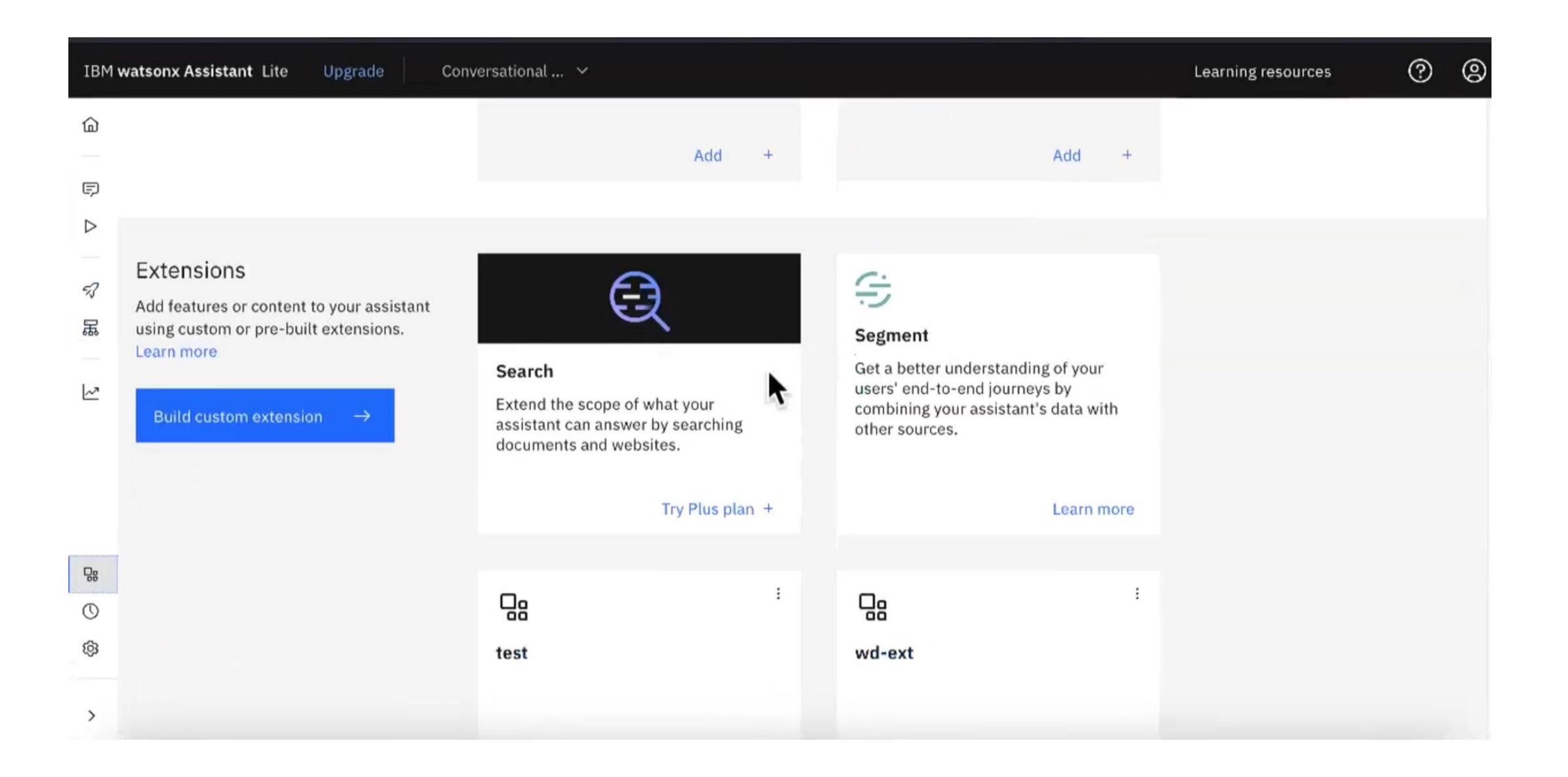


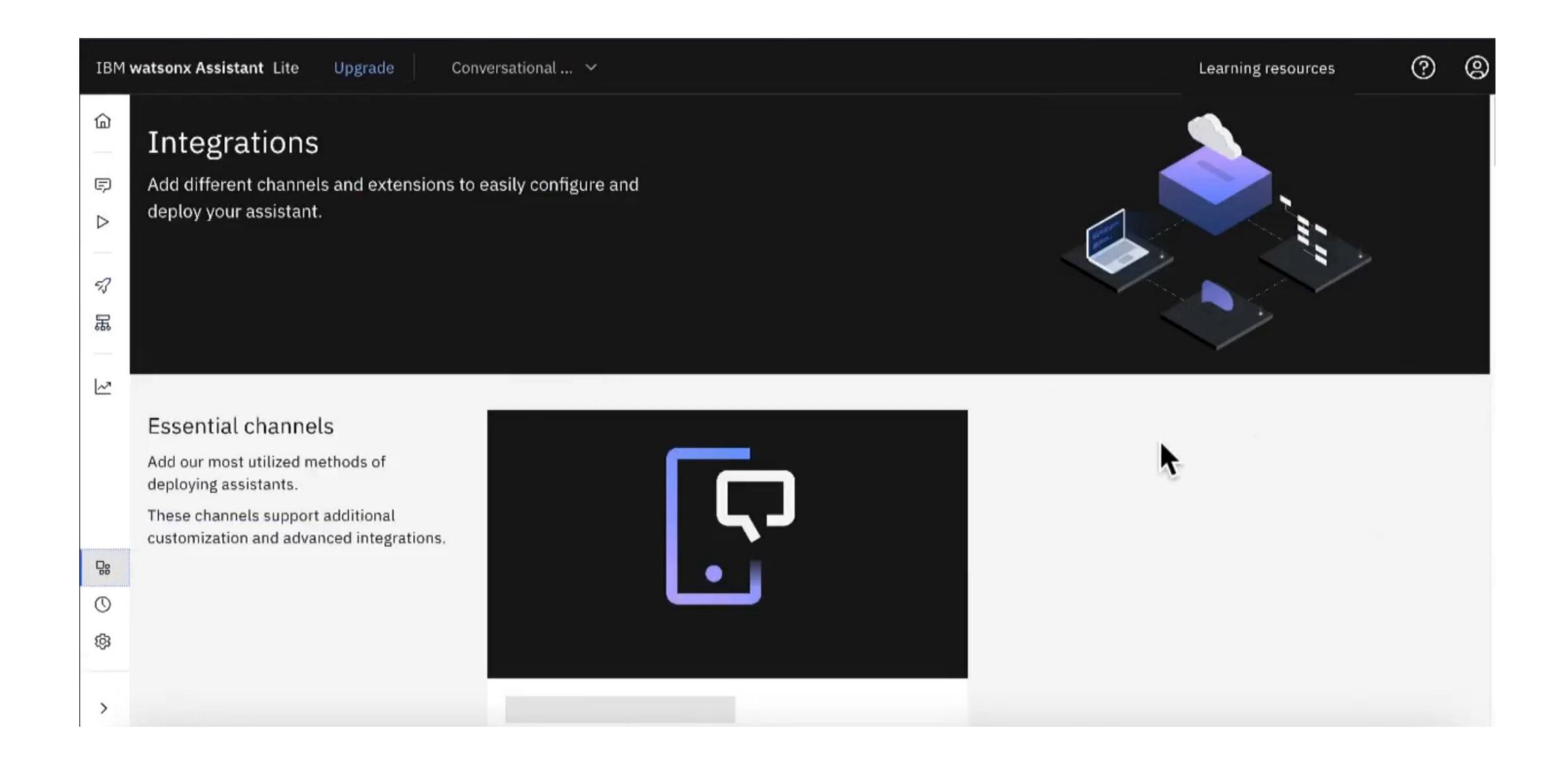












IBM watsonx Orchestrate

```
\nearrow \nearrow \nearrow \nearrow \nearrow \nwarrow \nwarrow \nwarrow \nwarrow \leftarrow \leftarrow \leftarrow
\uparrow\uparrow\uparrow\uparrow \\ \uparrow\uparrow\uparrow
```

After years of investments in automation, employee productivity is still declining

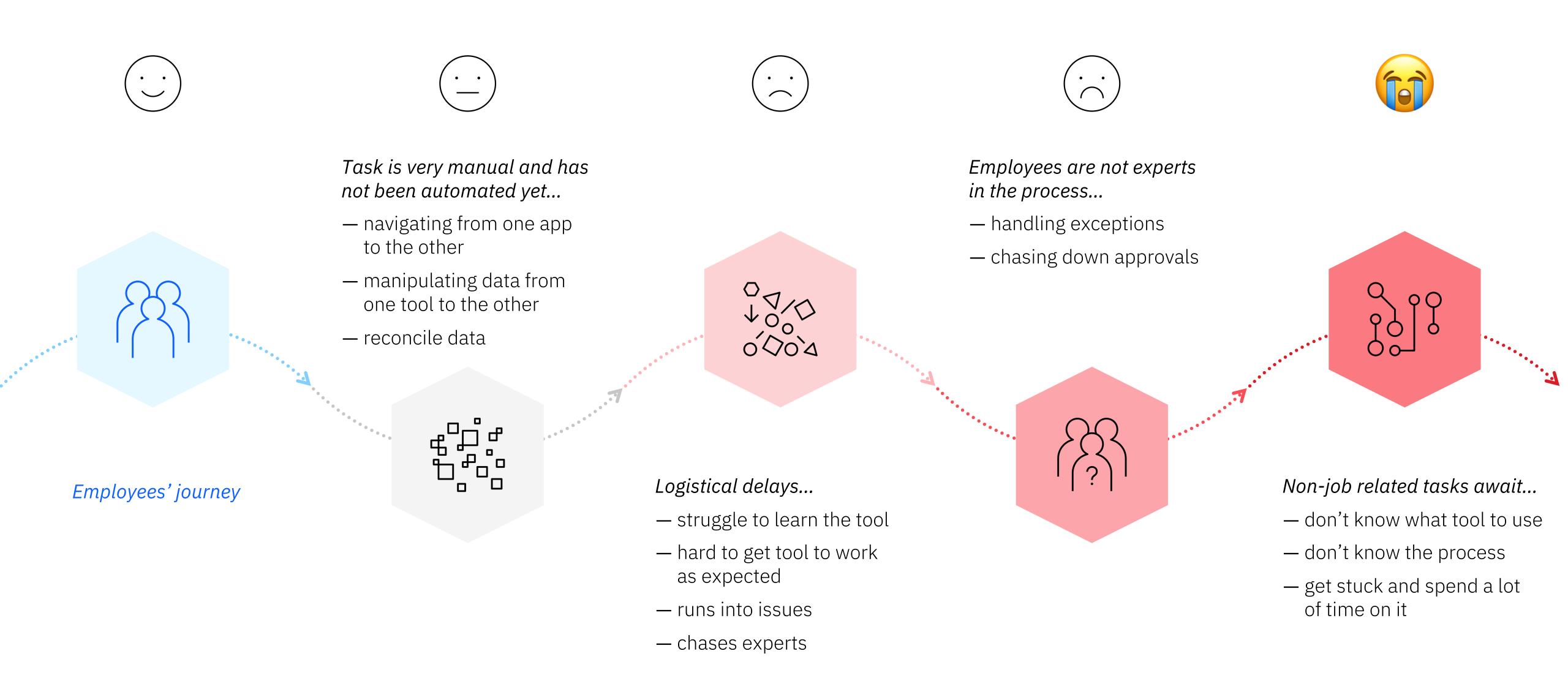


~3 hrs

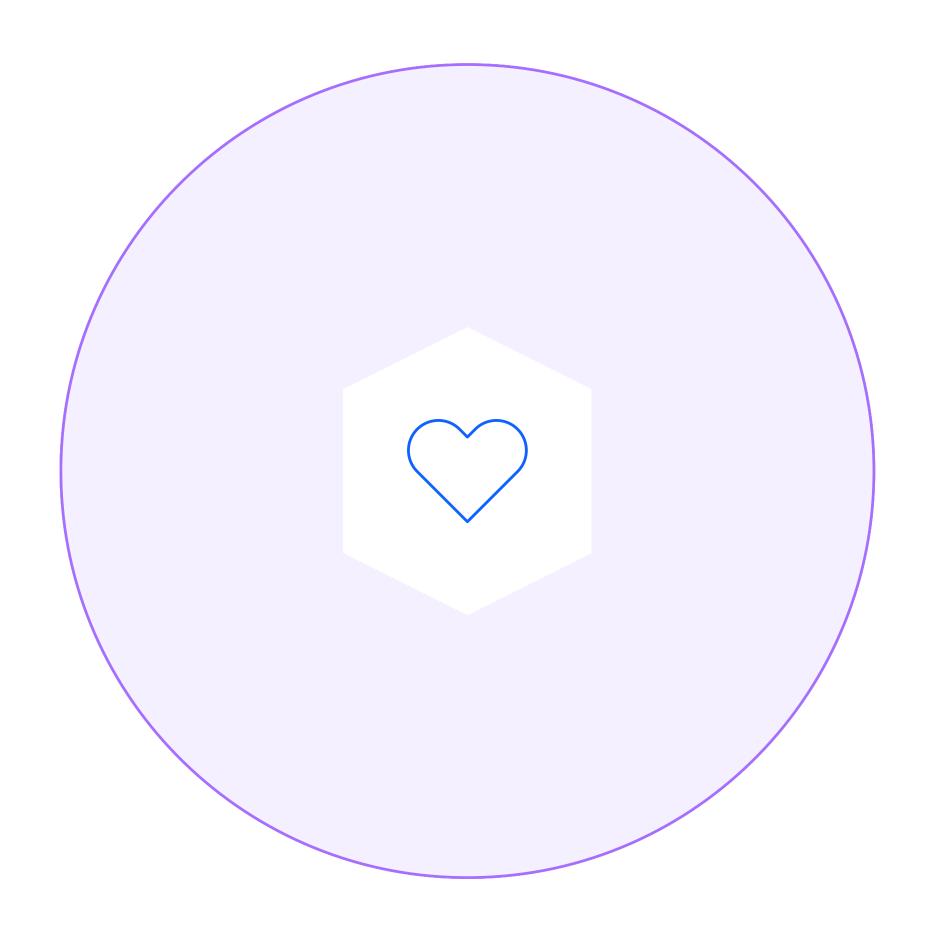
average time an employee is productive 13%

employees satisfied with their experience

Employee productivity *is hindered* by having siloed and disconnected tasks, systems, and processes



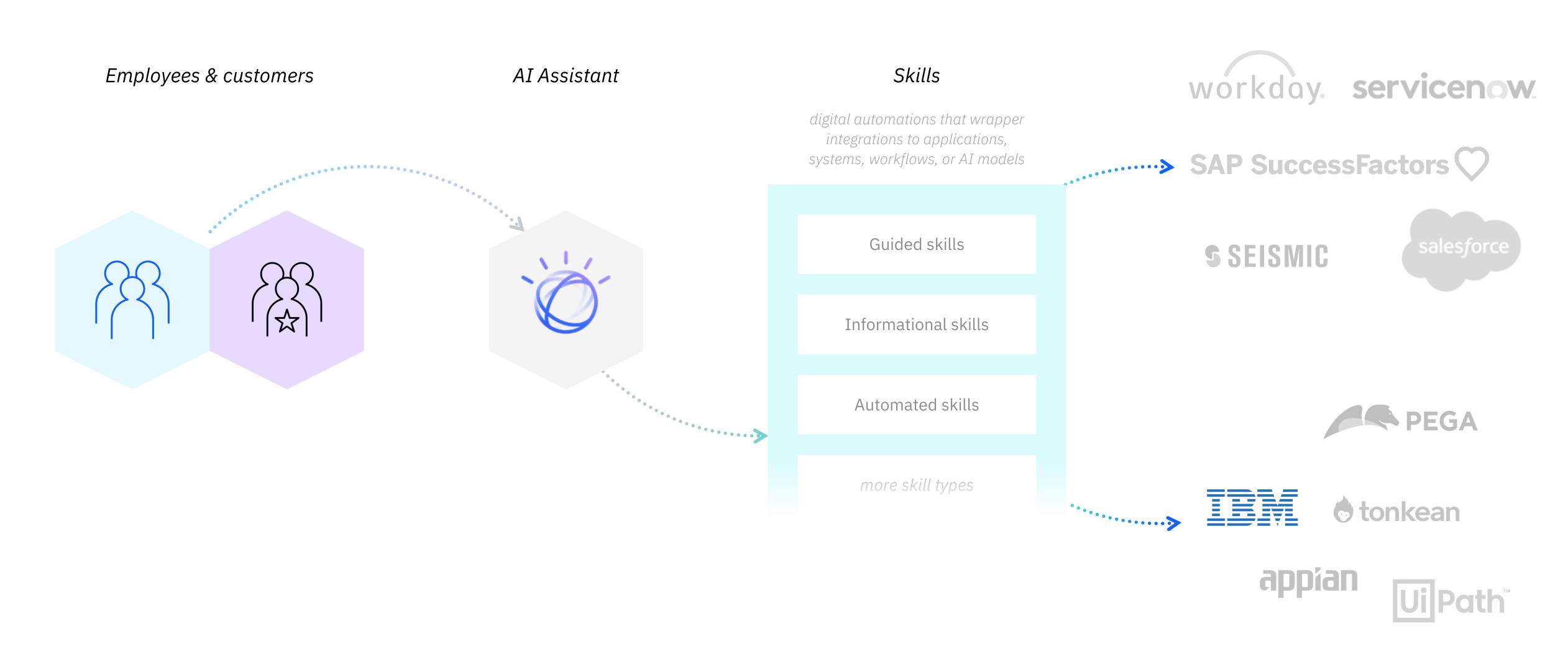
watsonx orchestrate: Unlock productivity with AI assistants



Harness the power of Gen AI and automation to *free up* individuals across the enterprise from tedious daily tasks

Work, reimagined.

IBM's AI assistants use Foundation Models and Automation to orchestrate skill execution and reduce time and effort



watsonx Orchestrate: platform to build AI assistants



Customers



Employees



SMEs











Conversational experiences

☼ Invoke advanced skills

- → Create job requisition
- → Send out offer
- → Identify upsell accounts

- → Schedule interview
- → Promote employee
- → Recommend new plans

- → Generate offer email
- → Transfer employee

Gen AI skills CLASSIFY | GENERATE | SUMMARIZE | EXTRACT **watson**x[™]



Execute automations















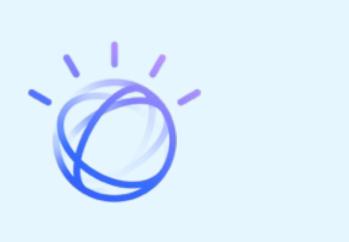


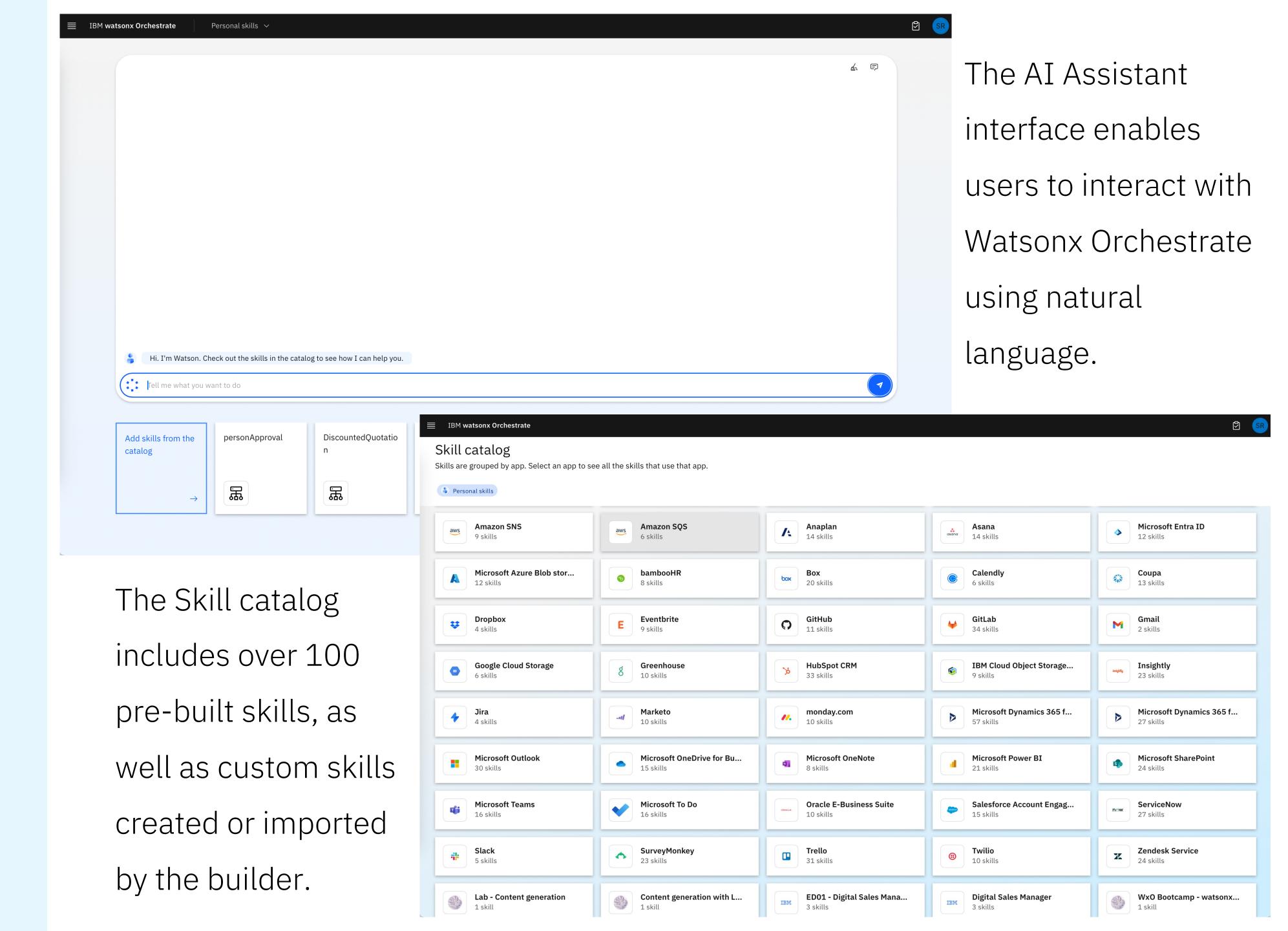
Democratize the availability of automation through natural language across multiple channels

Invoke advanced skills that leverage Gen AI to accomplish net new tasks

Power AI Assistant skills leveraging existing Automation

Al Assistant interface & skill catalog

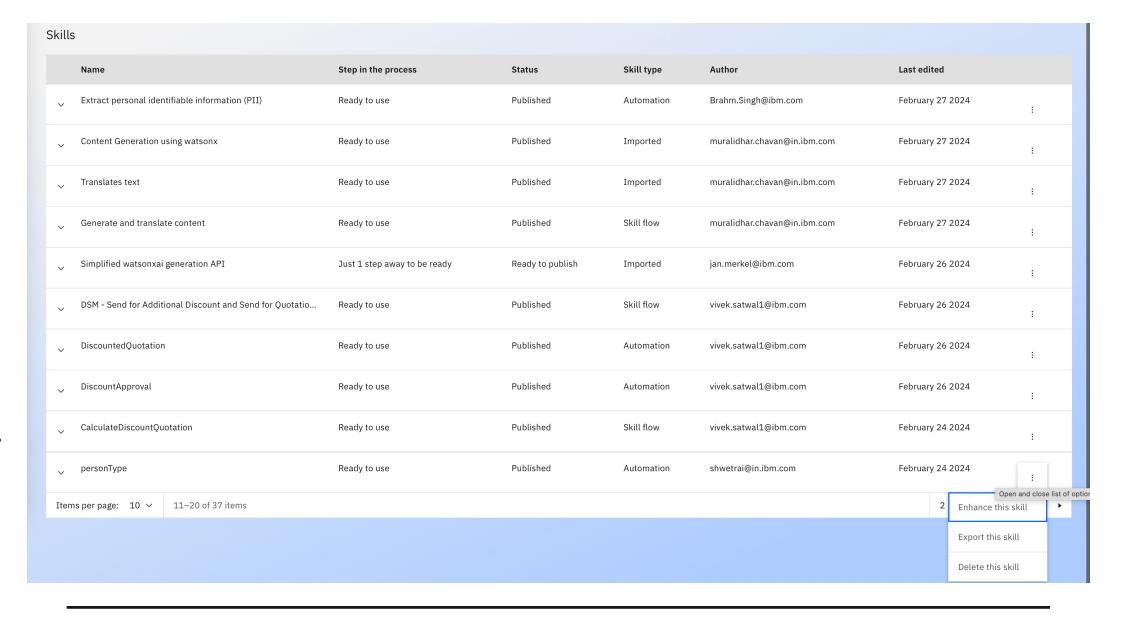




You can now customize and enhance your skills through a UI-based experience.

Training your skills to drive greater clarity is easier than ever before with the low code skill enhancing experience.

Custom annotations are no longer required to define inputs, outputs, phrases, and next best skill recommendations.



Home / Manage skills / Enhance this skill

Enhance the "Translates text" skill

Add details that will make people want to use this skill.

Name	Input	Output	Security	Phrases	Next best skills
Name and desc	ribe this skill in a way tha	at tells users how it's used	d and why they would wan	t to use it.	
Name*					
Translates te	xt				
Description					
Translates te	ext from one language to	another			
				<u>//</u>	
API version*					
3.0.0				*	
Categories					
Add categori	es				



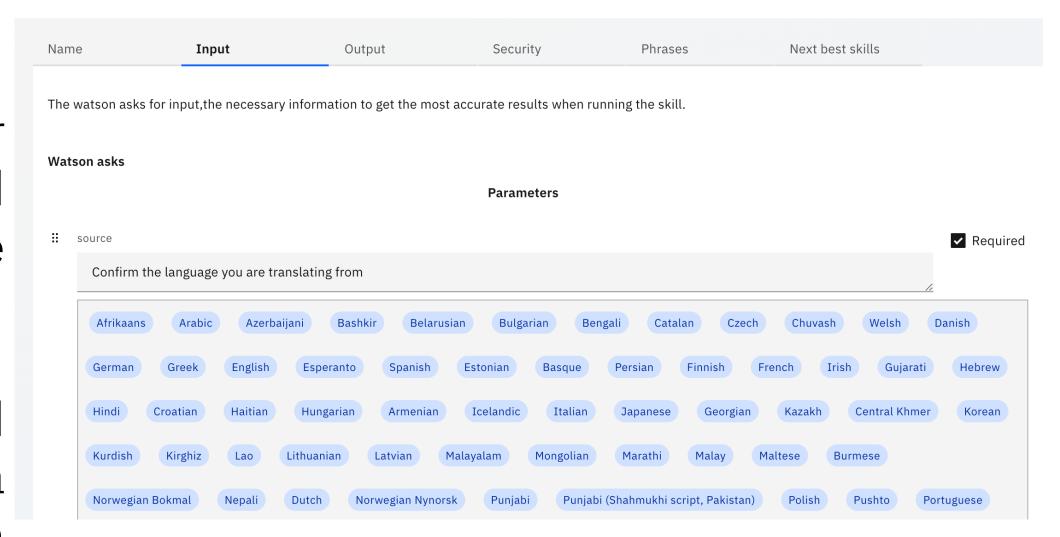
Inputs and Outputs

How the inputs and outputs are displayed to the skill users can now be defined through the UI. These are typically shown in a form format and are inherited from the OpenAPI spec

Parameter requirement, shown order to the user, and how the inputs and outputs are displayed can also be fully customized

Examples of these are options and text. Options present the user with a series of predefined choices for them to select from. The text options present open text fields of varying sizes: single-line or multi-line.





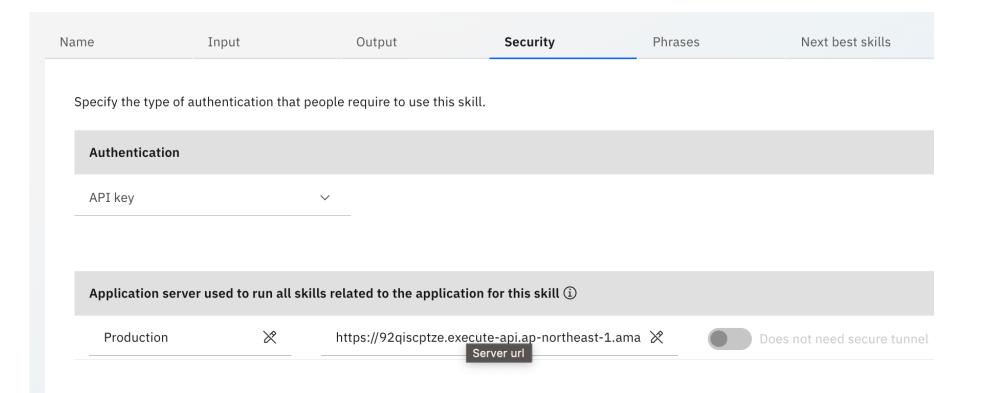


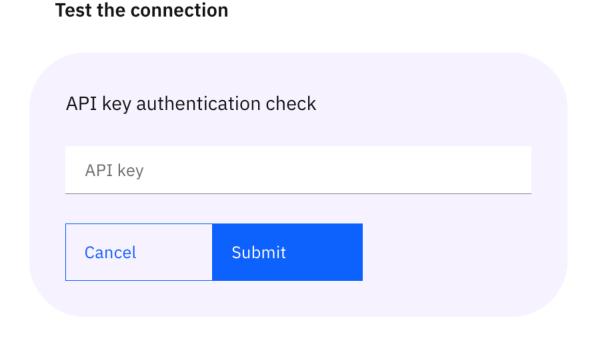
Security

On the Security tab, you can review the authentication type and the server. These are inherited from the OpenAPI spec.

On the same page, the connection can be tested.

Skills will still need to be connected to their respective application (RPA, SAP, TWG, etc)





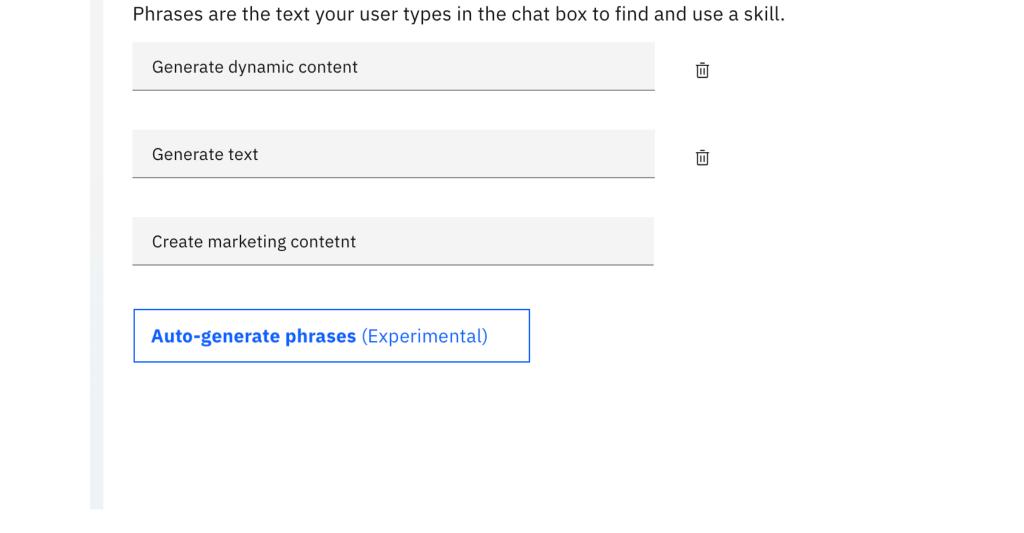


Phrases

Phrases can now be added and used to train the AI model used to trigger the skill from the chat interface. The custom annotation extension can still be used in the OpenAPI spec.

Phrases can be manually entered during skill enhancement or they can be defined in the OpenAPI spec before skill import.

We now introduce an experimental feature to generate phrases for the users to accelerate the build.



Output

Security

Phrases

Name

Input

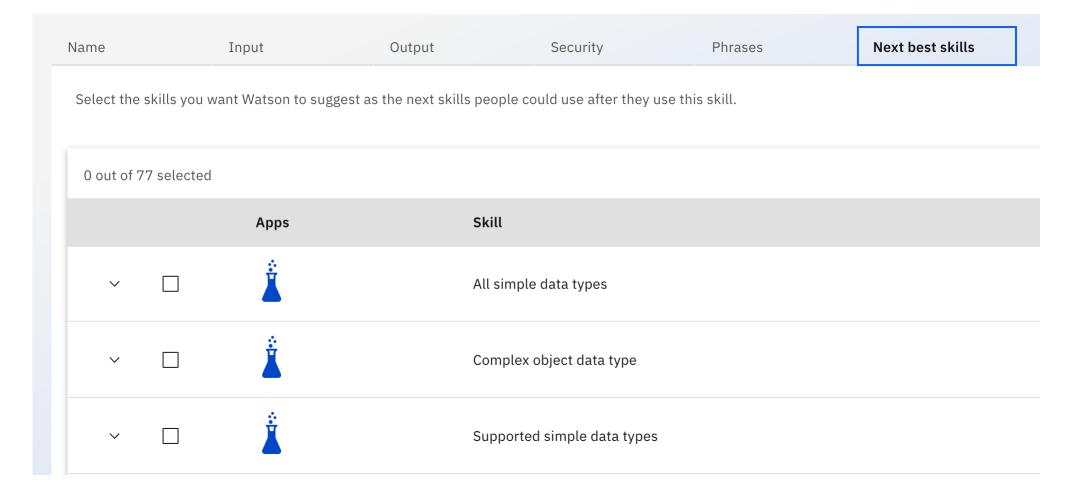


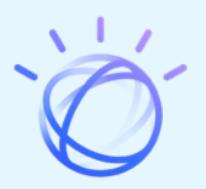
Next Best Skill Recommendations

Next best skills, which are skills suggested to the users after their current skill run returns results, can now be selected from the enhancement screen.

Suggested skills help users find skills to complete tasks that naturally follow the work currently being done.

Up to six skills can be added to the next best skill list.





Creating Skill Flows

Skills are often more valuable when combined with other skills.

Sequences of skills can be created using a skill flow. Structure your skill flow to use two or more skills to complete a task.

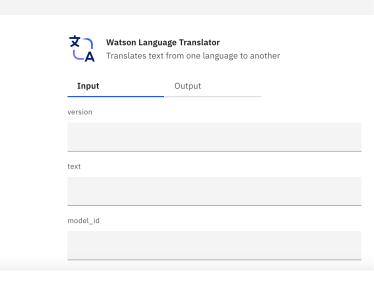
These skills can be from different applications. Using a skill flow, skills can be combined to carry out multiple functional calls using a single workflow

Each skill follow has its own utterance and next best skill enhancement



News Summaries

+ - End



ome / Manage skills / Enhance this skill

Enhance the "Test Skill Flow" skill

Summarize news and translate it to another language

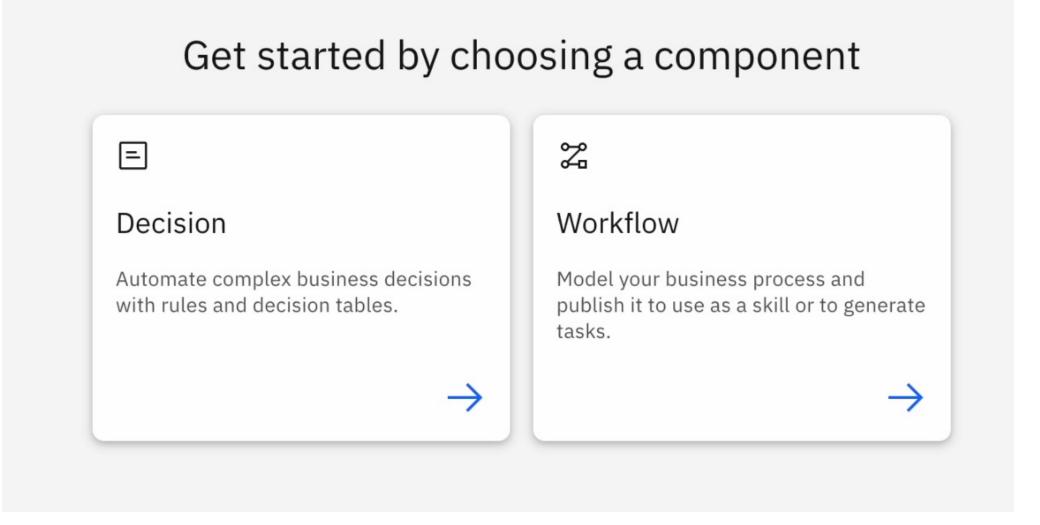
dd details that will make people want to use this skill.

Name	Phrases	Next best skills	
Name and decaribe	a this skill in a way the	t talla usara haw it's usad and why they would we	unt to uso it
	e this skill in a way tha	tt tells users how it's used and why they would wa	ant to use it.
Name*			
Test Skill Flow			
Description			
Enter the descrip	ption		
API version*			
1.0.0			×
Catagoria			
Categories			
Add categories			

Building new automations

In addition to tapping into existing investments, building RPA bots as a bridge to systems without an API, watsonx Orchestrate now empowers business teams to build new automations in low-code to fill the digital gaps and further reduce manual work.

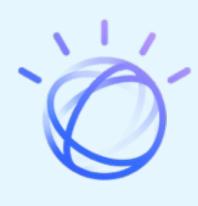
The automation builder is a brand new unified experience to define business decisions and workflows (to start with) from scratch or starting from industry templates.





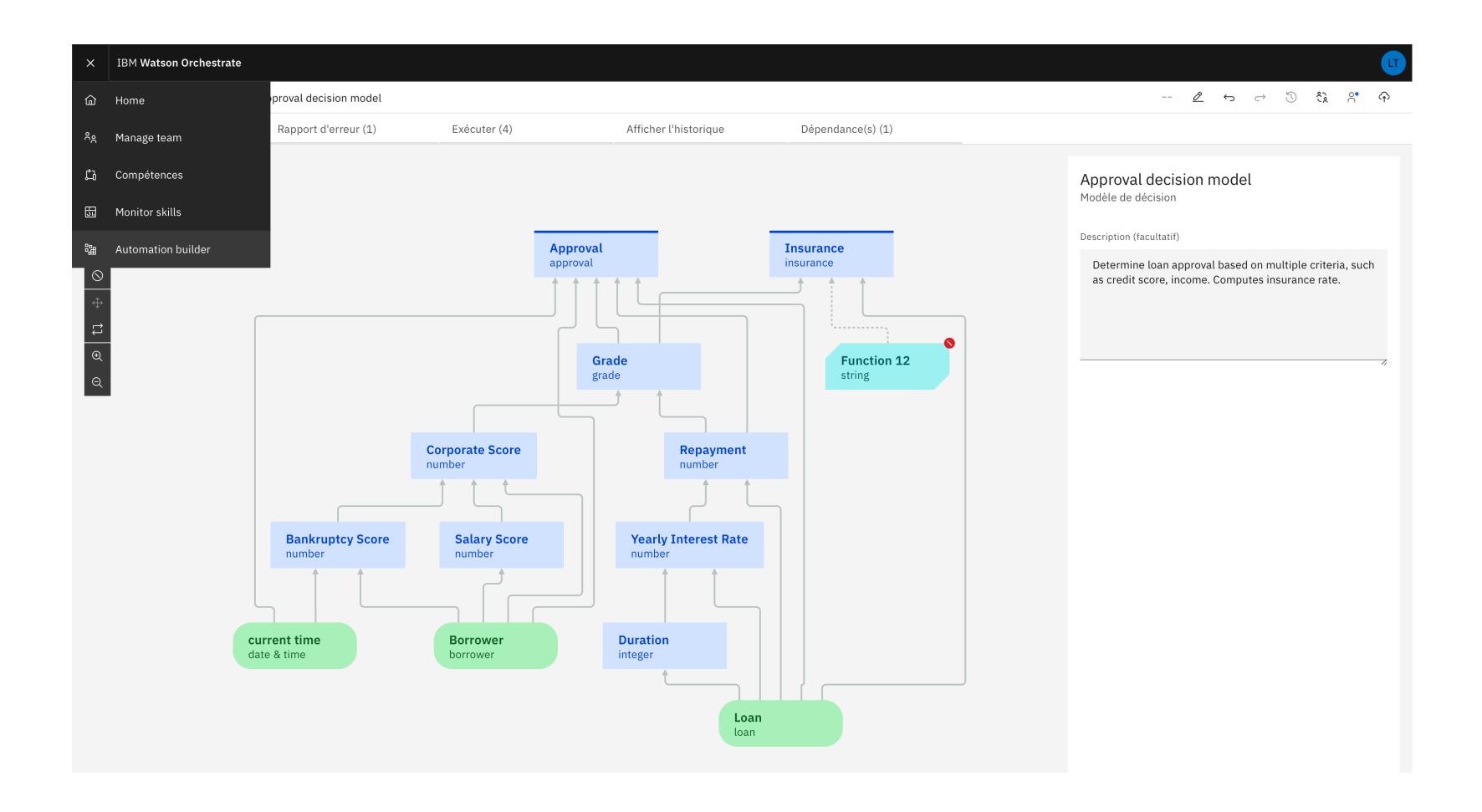
Building new automations

Decision



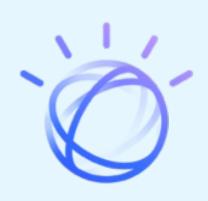
Empower business to codify & maintain complex decisions:

- Model a decision in graphical tree
- Write the business rules in natural language and decision tables
- Test right within the editor any



Building new automations

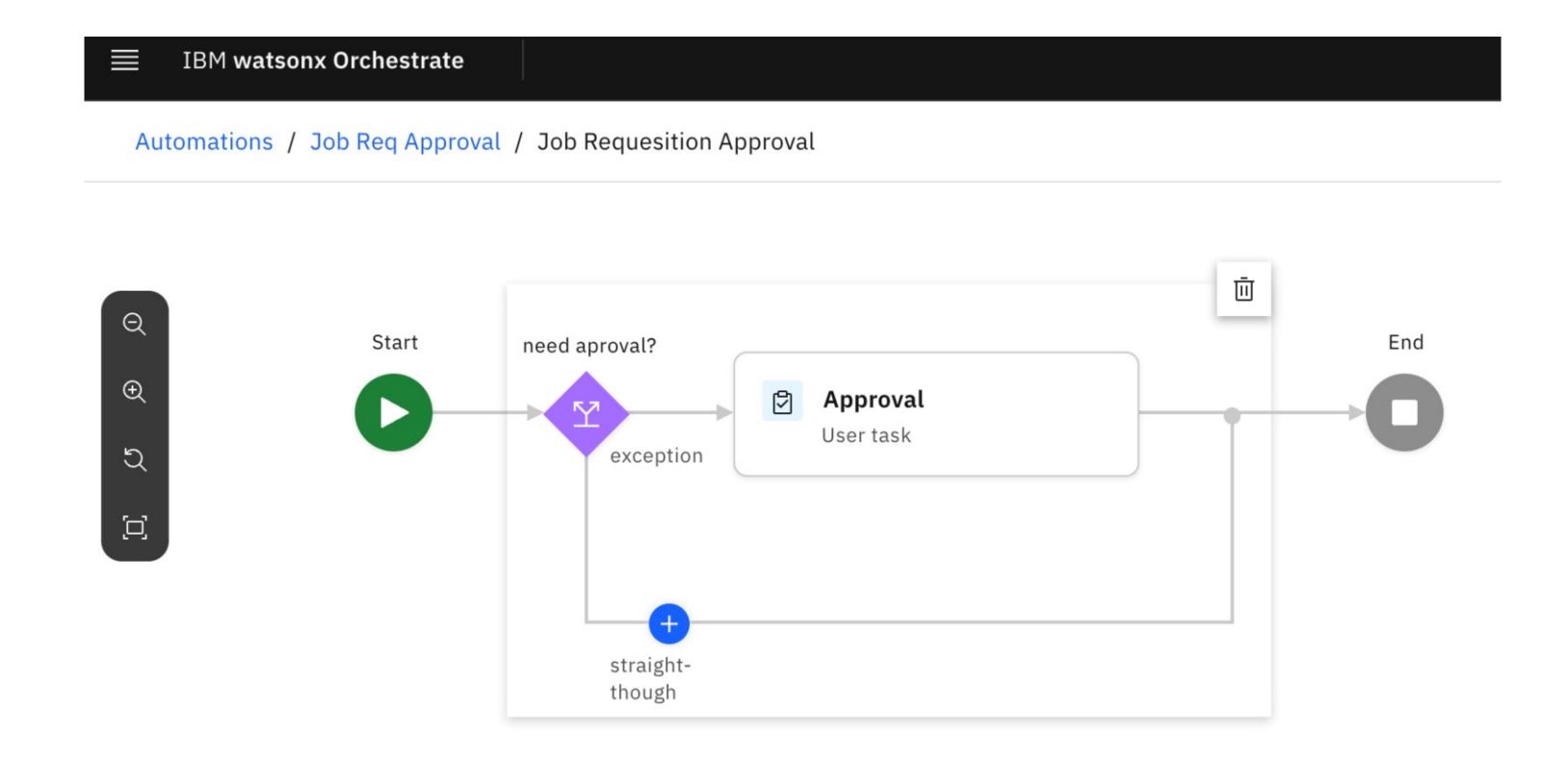
Workflow



Empower business to manage long-running workflows:

Model the diagram of activities and define UI interactions graphically.

- End-user to track & assign active work items & claim them in **Tasks**.
- Playback right within the editor any time



watsonx.ai + watsonx
Orchestrate: Generative
Al Recipe

This recipe connects watsonx Orchestrate to watsonx.ai. In this recipe you will create a project on watsonx.ai, connect your project to a Watson Machine Learning instance, and use generative AI to create dynamic content.



- watsonx Orchestrate Builder role
- IBM Cloud access
- IBM watsonx.ai access

