Engedjük szabadjára az WatsonX-et
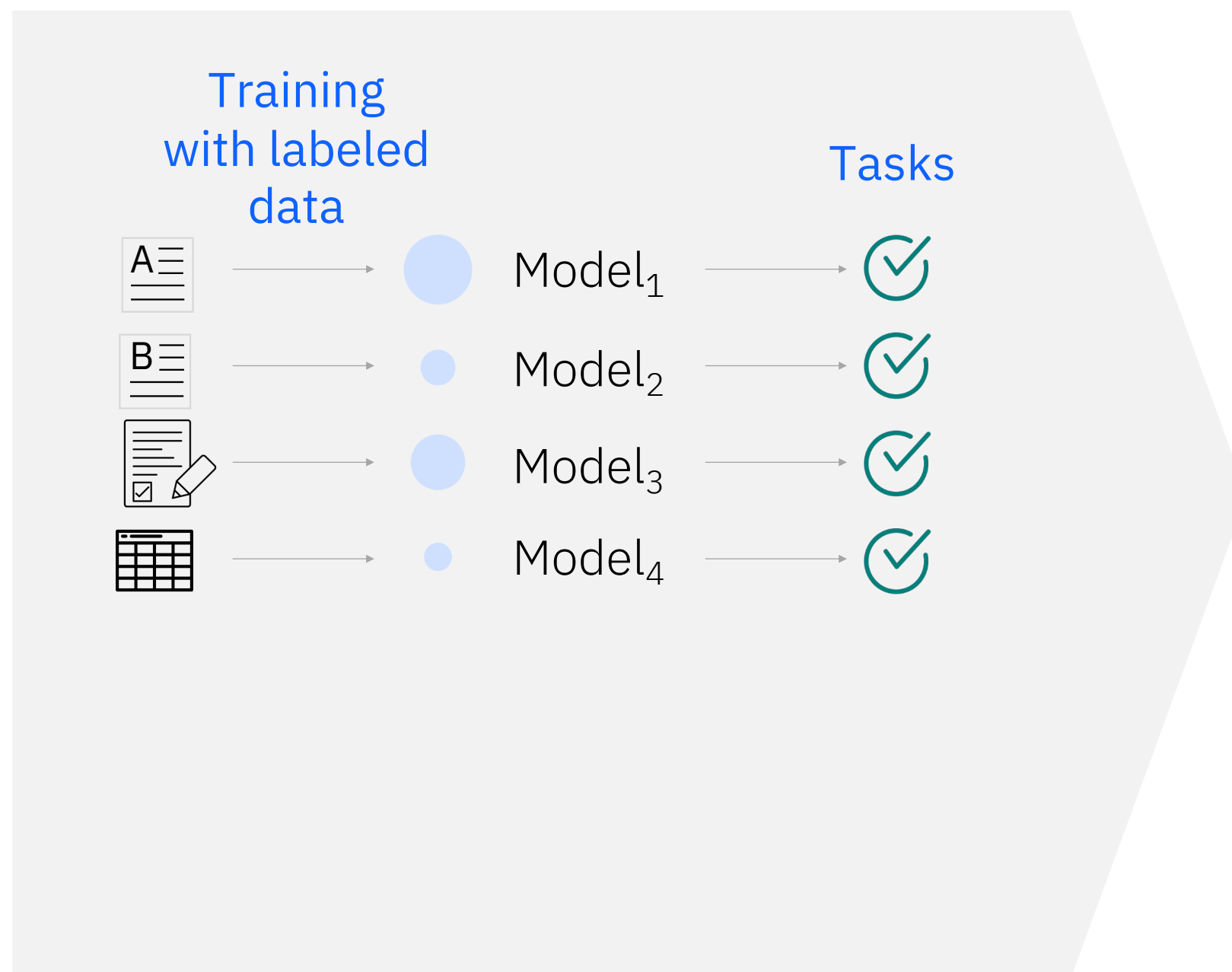
Huszti Dániel
daniel.huszti1@ibm.com
+36208235737

IBM

A ConText konferencia minden évben ...
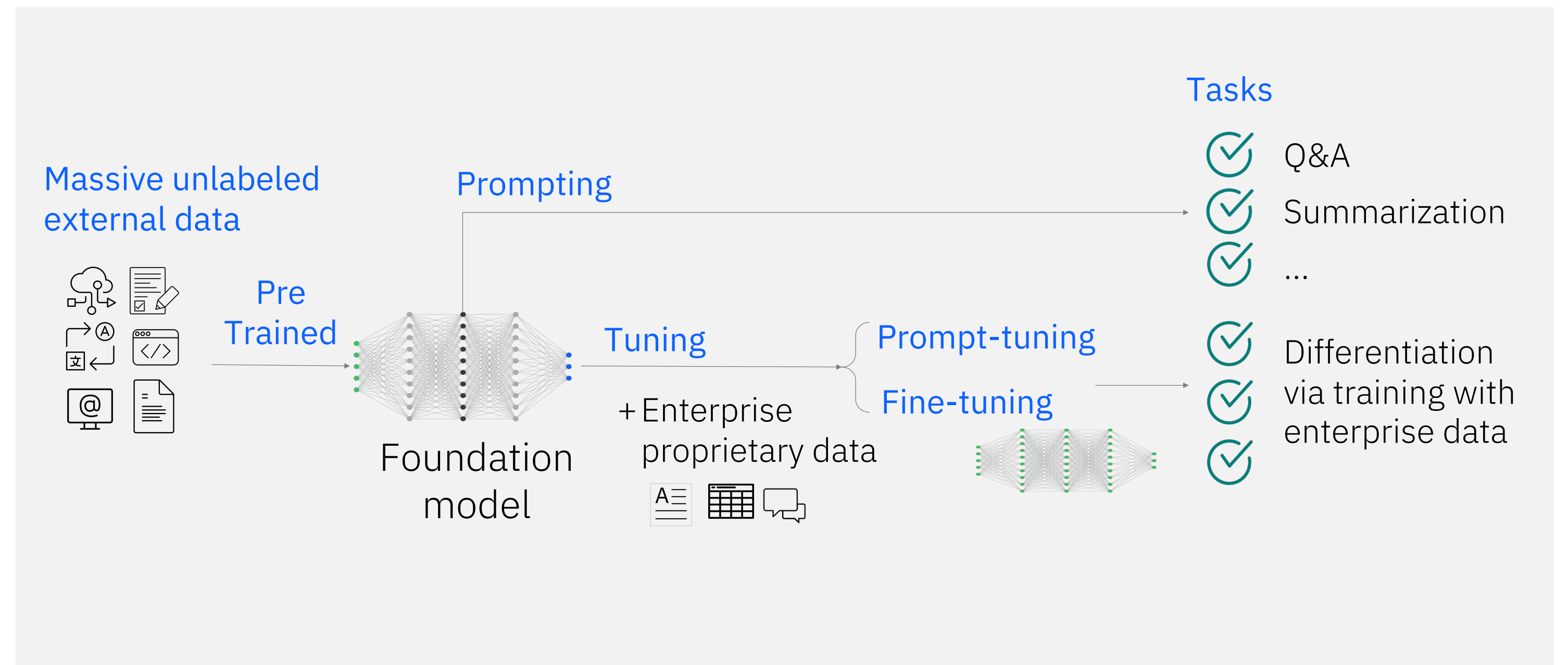
Foundational models enable a new paradigm of data-efficient AI development – generative AI

Traditional AI models

Training with labeled data

Tasks

$Model_1$
$Model_2$
$Model_3$
$Model_4$

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

Foundation Models

Massive unlabeled external data

Prompting

Pre Trained

Tuning

Prompt-tuning

Fine-tuning

Foundation model

+Enterprise proprietary data

Tasks

Q&A
Summarization
...

Differentiation via training with enterprise data

- Rapid adaptation to multiple tasks with small amounts of task-specific data
- Pre-trained unsupervised learning

Reinventing how work gets done

The most common generative AI tasks implemented today

Retrieval-Augmented Generation

Based on a documents or dynamic content, create a chatbot or question-answering feature.

Building a Q&A resource from a broad knowledge base, providing customer service assistance

Summarization

Transform text with domain-specific content into personalized overviews that capture key points.

Conversation summaries, insurance coverage, meeting transcripts, contract information

Content Generation

Generate text content for a specific purpose.

Marketing campaigns, job descriptions, blog posts and articles, email drafting support

Named Entity Recognition

Identify and extract essential information from unstructured text.

Audit acceleration, SEC 10K fact extraction

Insight Extraction

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

Medical diagnosis support, user research findings

Classification

Read and classify written input with as few as zero examples.

Sorting of customer complaints, threat and vulnerability classification, sentiment analysis, customer segmentation

Business leaders face challenges in scaling AI across the enterprise with trust
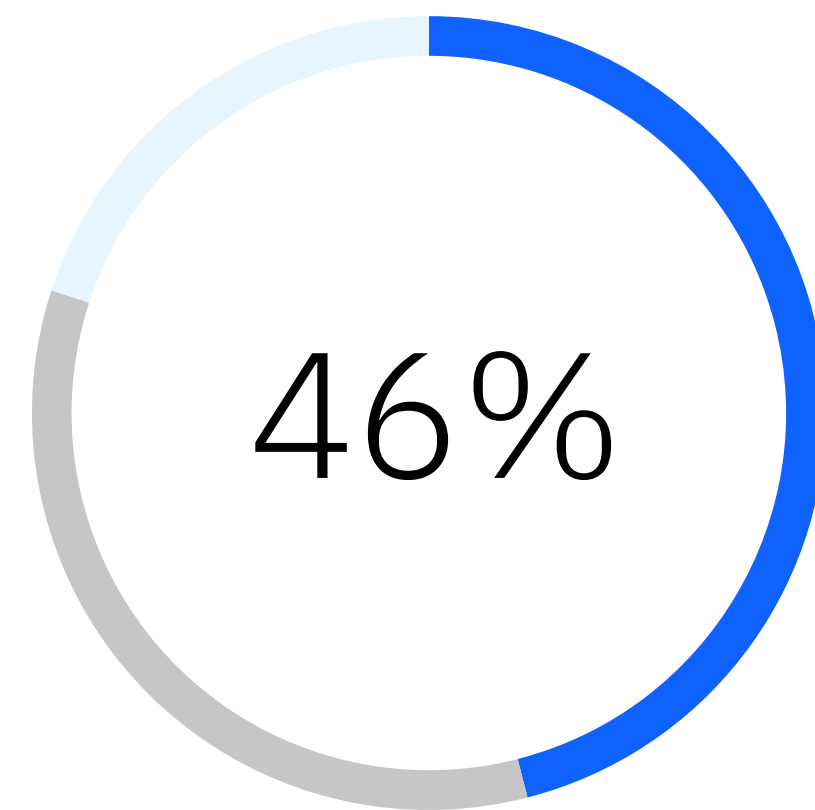
80% of business leaders see at least one of these ethical issues as a major concern

Explainability

48%
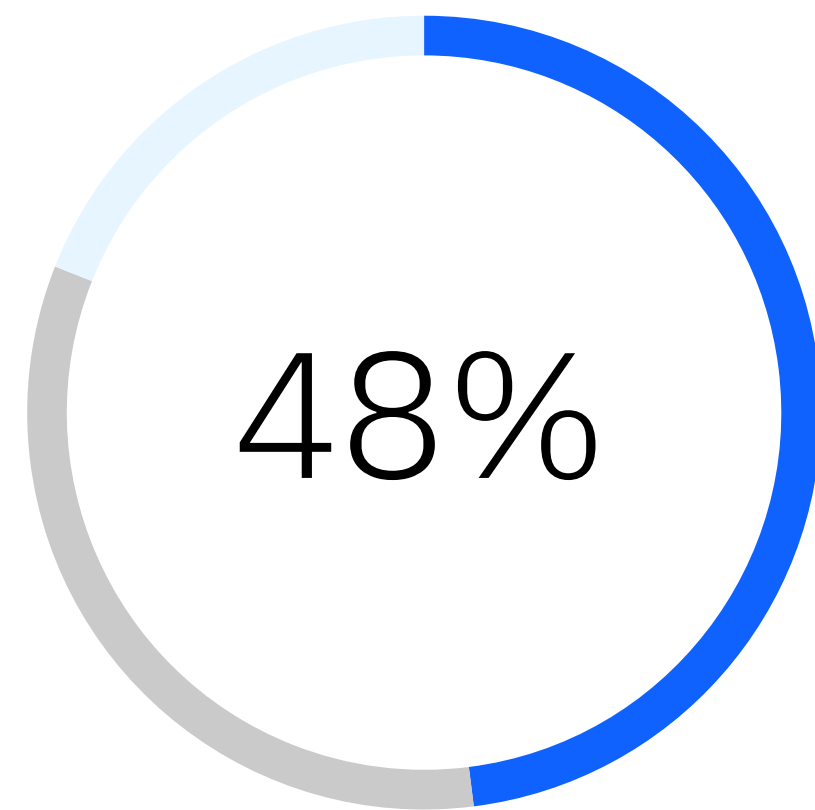
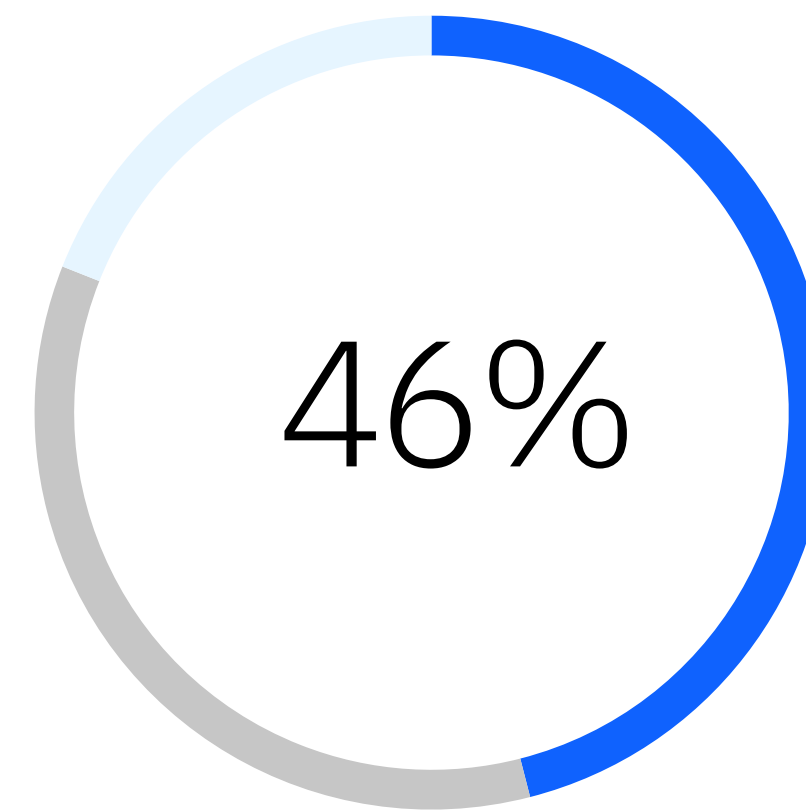Believe decisions made by Generative AI are not sufficiently **explainable**.
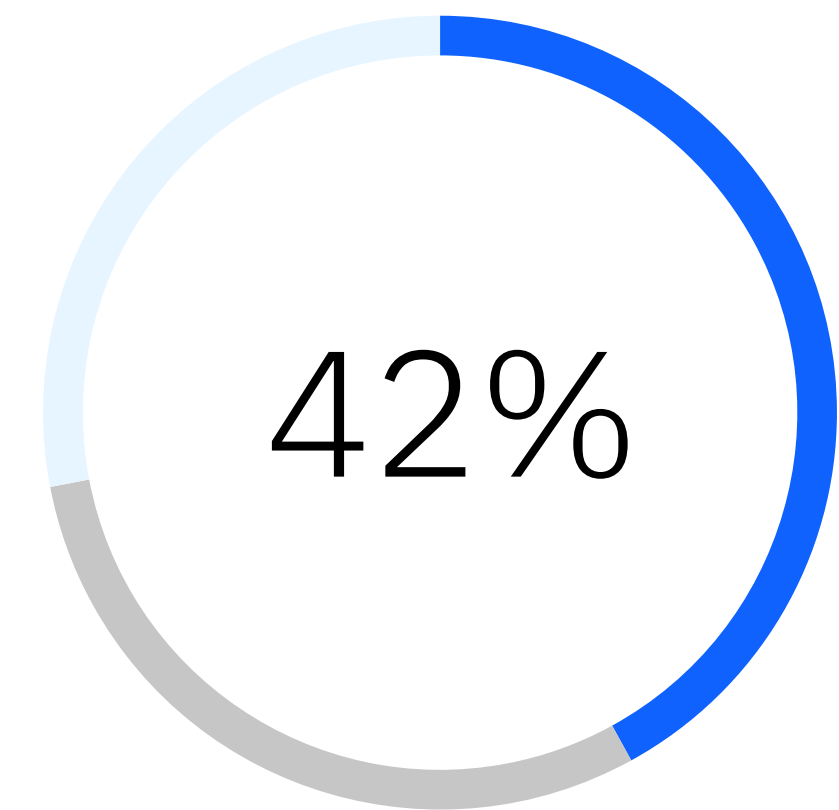
Ethics

46%

Concerned about the safety and **ethical** aspects of Generative AI.

Bias

46%

Believe that Generative AI will propagate established **biases**.

Trust

42%

Believe Generative AI cannot be **trusted**.

Source: IBM IBV "Generative AI: The state of the market", June 2023

■ Agree   ■ Neutral   ■ Disagree

Put AI to work with **watsonx**

Scale and accelerate the impact of AI with trusted data

Leverage foundation models to automate data
search, discovery, and linking in watsonx.data

watsonx.ai	**watsonx.data**	**watsonx.governance**
Train, validate, tune and deploy AI models	Scale AI workloads, for all your data, anywhere	Accelerate responsible, transparent and explainable data and AI workflows
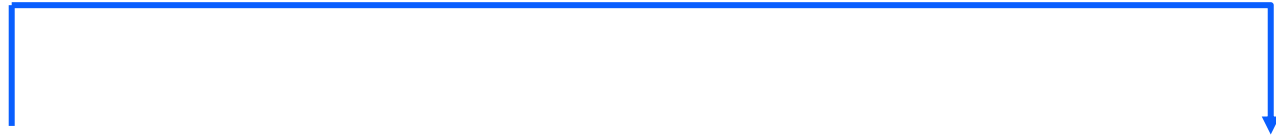
Leverage governed enterprise data in watsonx.data
to seamlessly train or fine-tune foundation models

Enable fine-tuned models to be managed through market
leading governance and lifecycle management capabilities

Most look something like this.

You need something like this.

First try

End-user

Self Service App

Another App

S/4 HANA

Success Factors
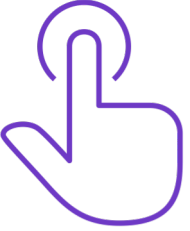
Concur

More Content

Content

Support

More Support

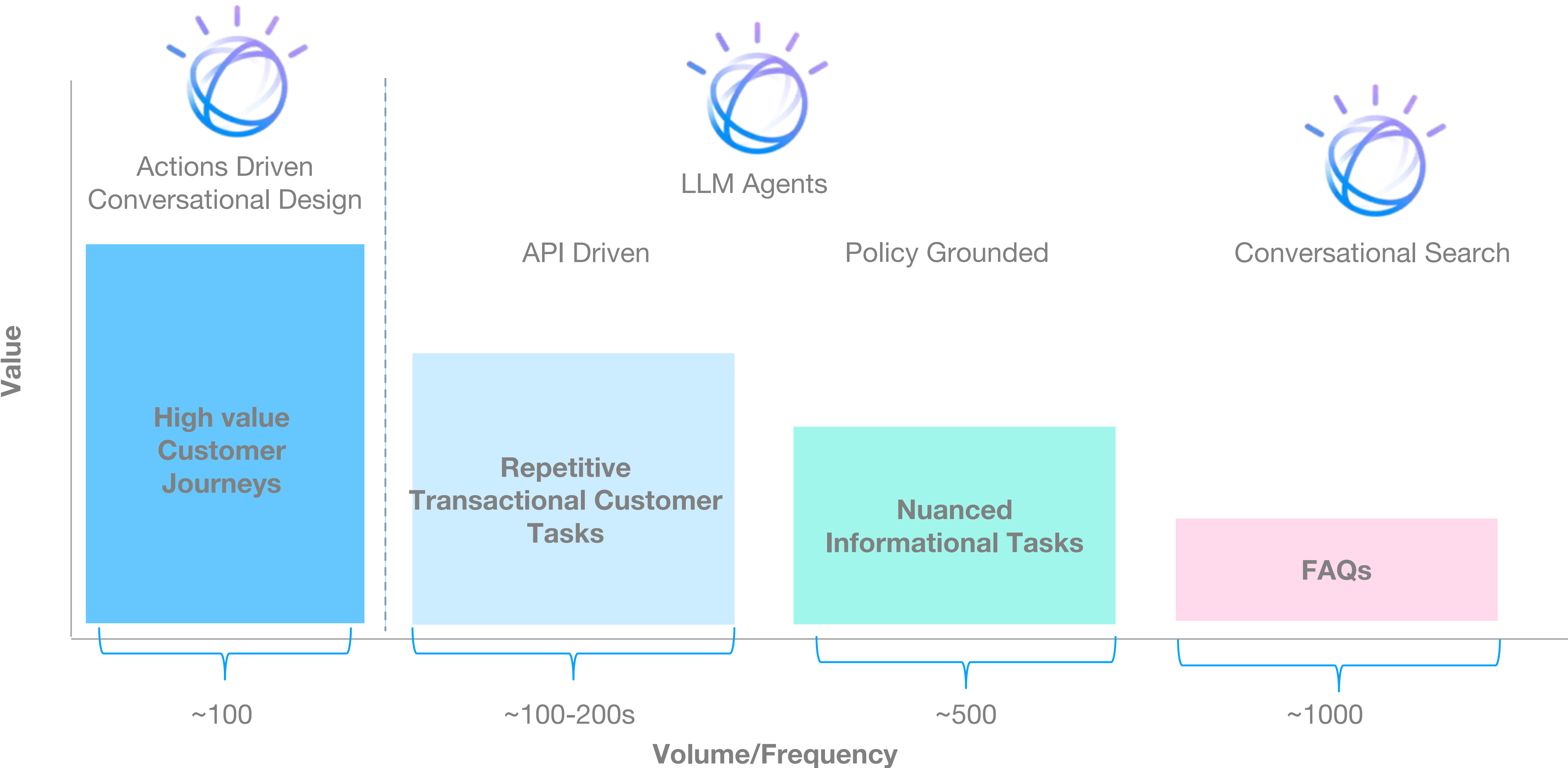| 30-40% | Operating Cost Reduction | 15+% | Increase in NPS |
| 3x-5x | ROI within 6 Months | 20%+ | Increase in Workflow Efficiency |

IBM **Watson**

IBM watsonx Assistant is an *end-to-end conversational AI platform designed* to deliver exceptional customer care across all entry points

- Easier to build, easier to improve
- Frictionless experience out of the box
- Integrate with anything
- Built for the Enterprise

Telephony & Messaging Channels

Speech

Smart Routing Decisions

Conversation Infrastructure
- State Management
- Routing/Handoff
- Orchestration via Webhooks
- History

Orchestration Callouts
- Pre-message
- Post-message
- Conversation Logs

Insights
- Assistant Performance
- Call Logging/Auditing
- Recommends
- Auto-Learning

Watson Language Translator Watson Tone Analyzer

Resolution Methods

Actions/Dialog
- NLP
- Disambiguation
- Autocorrect
- Languages

Task-Specific Callouts

Watson Discovery
AI Search
- NLP
- SDU
- Reading Comprehension

Content Crawler

Hand-off
Human Agent

Handoff Integrations

Integrations

Other App(s)
- Task 1
- Task 3
- Task 2
- Task 4

Content Repositories

box

Bring Your Own...

Contact Center Tools

salesforce NICE inContact

GENESYS Bring Your Own...

Analytics

Cloud Pak for Data
- Watson Studio
- Cognos
- Watson Knowledge Catalog

IBM DB2

Watson Assistant + LLMs enable dramatically more use cases

Value

Actions Driven Conversational Design

High value Customer Journeys

~100

API Driven

Repetitive Transactional Customer Tasks

~100-200s

LLM Agents

Policy Grounded

Nuanced Informational Tasks

~500

Conversational Search

FAQs

~1000

Volume/Frequency

14

Várunk titeket a november 22-23-ai TechXChange rendezvényen!

IBM