Authorized Distributor

Red Hat AI PoC's

From Possibility to Proof: Turning AI Potential into Real-World Impact with Red Hat

This document is the property of TD SYNNEX and can only be used by the intended recipients.
Any unauthorized use, distribution, reproduction, or modification of this document or its contents is strictly prohibited.

Date	Author	Version
22.09.2025	Matthias Laetzsch Sejid Canoski	1.0

Todays Agenda

0 1

What is Red Hat AI?

A scalable, open and secure path to business AI

0 2

Red Hat AI PoC initiative

Red Hat AI PoCs for business partners

0 3

Identification & Qualification

What can a Red Hat AI PoC look like?

0 4

Q&A – Call to action

TD SYNNEX

01

What is Red Hat AI?

A scalable, open and secure path to business AI

This document is the property of TD SYNNEX and can only be used by the intended recipients.
Any unauthorized use, distribution, reproduction, or modification of this document or its contents is strictly prohibited.

Ai is a strategic enabler across industries

Predictive AI runs businesses today, Generative AI brings innovation to the enterprise

01

Revenue Generation

- Chatbots
- Campaign and Content Marketing
- Developer assistants
- Guided selling
- Drive product innovation

02

Cost Optimization

- Automated AI support
- Knowledgebase Search & Summarization
- Doc summarization
- AI-optimized logistics
- Augmented Product R&D

03

Risk Management

- Sentiment analysis
- Predict employee attrition
- Contract risk assessment
- Fraud detection
- AI-assisted Security Operations

Gen AI customer adoption challenges

0 1

Cost

Generative AI frontier model services are cost prohibitive at scale for most enterprise customer use cases.

0 2

Complexity

Tuning models with private enterprise data for customer use cases is too complex for non-data scientists.

0 3

Flexibility

Enterprise AI use cases span data center, cloud & edge and can't be constrained to a single public cloud service.

Red Hat AI accelerates the development and delivery of AI solutions across hybrid-cloud environments.

TD SYNNEX

Manage and monitor the lifecycle of predictive and generative AI models at scale

Deploy with flexibility and consistency wherever your data resides

Tune small, purpose-built models efficiently with enterprise-relevant data

Maximize your technology investments with operational guidance & advisory services

Making IT Personal™

Red Hat AI supports each stage of the AI adoption journey

TD SYNNEX

From single server deployments to highly scaled-out platform architectures

Generative Model

Data

Model Server

Trusted & Consistent foundation

Single server GPU

Physical

Virtual

Private

Public

Edge

Generative Model

Predictive Model

AI-enabled workloads

Data & Connections

Model servers (1, 2, …n)

Containers

Storage

Trusted, Consistent and Comprehensive foundation

Hardware Acceleration

IBM Cloud

aws

Microsoft Azure

Google Cloud

Making IT Personal™

TD SYNNEX

Hybrid cloud deployment for AI

Across different hardware accelerators, on-prem OEM servers, and cloud environments

1

0 1

Hardware Accelerators

NVIDIA GPU

AMD Instinct

intel Gaudi

Google TPU

aws Neuron

IBM AIU

2

0 2

On-prem OEM Servers

DELL

Lenovo

CISCO

Hewlett Packard Enterprise

3

0 3

Cloud Environments

IBM Cloud

aws

Azure

Google Cloud

Making IT Personal™

* NVIDIA GPUs fully supported in Red Hat AI. AMD Instruct $ Intel Gaudi in Preview, AWS Inferentia/Neuron, Google TPU, IBM AIU are on our roadmap

Red Hat AI overview

Red Hat AI

Red Hat AI Inference Server

Red Hat Enterprise Linux AI

Red Hat OpenShift AI

Trusted, Consistent and Comprehensive foundation

NVIDIA | AMD | intel | Hardware Acceleration | Google | aws | IBM

Physical

Virtual

Private Cloud

Public Cloud

Edge

* NVIDIA GPUs fully supported in Red Hat AI. AMD Instruct $ Intel Gaudi in Preview, AWS Inferentia/Neuron, Google TPU, IBM AIU are on our roadmap

TD SYNNEX

0 1

Increased efficiency

Access to enterprise-grade, open-source Granite models fully supported and indemnified by Red Hat.

0 2

Simplified experience

InstructLab model alignment tooling to support developers and domain experts to tailor models more efficiently.

0 3

Flexible deployment

Mitigate risks and generate value while controlling costs by owning the decision of where to train and deploy AI models.

0 4

Agile flexibility

Access to cutting-edge open-source innovations to keep up with a fast-moving market.

0 5

Choice and adaptability

Access to an open ecosystem of communities, technology providers, ISVs and customers.

0 6

Abstract the complexity

Reduce the complexity of switching and adapting to new technologies.

Making IT Personal™

02

Red Hat AI PoC initiative

Red Hat AI PoCs for business partners

This document is the property of TD SYNNEX and can only be used by the intended recipients.
Any unauthorized use, distribution, reproduction, or modification of this document or its contents is strictly prohibited.

What is a GenAI Proof of Concept (POC)?

TD SYNNEX

A short, low-risk engagement to explore how GenAI can address a business need using customer data.

The goal is to quickly demonstrate real-world value and build confidence in Red Hat's AI capabilities. Success is measured against customer criteria, often compared to GPT-4.

Execution Paths:

Fine-tune Granite with customer data

Use Retrieval-Augmented Generation (RAG) to personalize the model based on customer documents.

Typical duration: **14+ days**

Compare Granite to other models

Test performance and cost alongside other open source models like Llama or Mistral to find the right fit.

Typical duration: **10+ days**

Run advanced or custom use cases

Handle non-English content, complex documents, or custom requirements with extra engineering support.

Typical duration: **15+ days**

Scale with vLLM inferencing

Use vLLM for faster, more efficient model performance—great for batch processing or large-scale needs.

Typical duration: **30+ days**

Making IT Personal™

Drive New Gen AI POCs & Opportunities

TD SYNNEX

Provide evidence that Red Hat AI will meet or exceed GenAI requirements for a specified use case

Identify customer use case	Set evaluation criteria	Gather and prepare data	Develop and tune model	Serve model	Evaluate results
‣ Work with the customer to isolate a specific use case for GenAI in their business, and the challenges they face in attaining that use case (guided by AI Platform Sales Play content)	‣ Define success criteria in form of quality improvement measures ‣ Identify the benchmark for which quality will be judged (ex. GPT4)	‣ Collect data from the customer (private or public) aligned to the specific use case	‣ Fine tune model on customer data ‣ Deploy RAG based solution with customer data on model	‣ Expose the trained model's functions via an API	‣ Conduct a model evaluation with benchmark ‣ Share and discuss results with the customer
Driven by SSP, AE, ASA	*Driven by SSP, SSA*	*Driven by SSA*	*Driven by SSA*	*Driven by SSA*	*Driven by SSP, SSA*

Who Are the Target Customers?

TD SYNNEX

InstructLab POC

AI-invested organizations building or evaluating **chatbots or internal assistants**. Often in **regulated industries** (e.g., finance, healthcare) where **data privacy**, **model ownership**, and **compliance** are critical. Want **full control** and **customization**.

vLLM POC

Teams running **multiple GenAI models at scale** across **hybrid or edge** environments. Focused on **reducing cost**, **improving performance**, and **optimizing infrastructure**.

OpenShift AI

Existing **OpenShift users** looking to add **AI/ML** to their current platform. Value **tool consolidation**, **workflow consistency**, and a **trusted, unified environment**.

Not a Fit

Customers with **no defined use case**, **no internal champion**, or just beginning to **explore AI**. Not ready for **custom model development** or **platform-level investment**.

TD SYNNEX

03

Identification & Qualification

How to identify and qualify a customer for a PoC

This document is the property of TD SYNNEX and can only be used by the intended recipients.
Any unauthorized use, distribution, reproduction, or modification of this document or its contents is strictly prohibited.

Key Personas

TD SYNNEX

Decision makers

C-level
(CIO, CTO)

Drives AI strategy and approves large-scale investments tied to business impact and risk.

LoB
(VP or Director of marketing, sales, or service development)

Funds and prioritizes AI use cases that support business unit goals (e.g. chatbots, personalization).

AppDev ITDM
(Director/Manager of IT or Manager of a specialized department)

Leads AI project execution, approves tools, and drives app and platform decisions.

Key influencers

Data Scientist
(Data Manager or Data Analyst)

Defines AI use cases and influences investment in models, data tools, and platforms.

Developer
(Senior/Lead Software Engineer or Architect, Application Developer, or Principal Engineer)

Builds AI-powered apps and influences tooling to improve dev velocity and integration.

AI Engineer
(Machine Learning Engineer, AI/ML Engineer, Deep Learning Engineer)

Architects AI infrastructure; evaluates model fit, performance, and scalability.

Platform Engineer
(DevOps Engineer, Cloud Platform Engineer, Application Architect)

Ensures AI workloads run reliably and securely across infra; operationalizes pilots to production.

Learn more

Making IT Personal™

Filter Criteria

AI Maturity

Prioritize **accounts with a clear AI vision**—those with internal AI teams, published strategies, or a history of running GenAI or ML pilots.

Other AI Signals

Lean into **accounts showing GenAI interest**—those who've attended Red Hat AI events, downloaded InstructLab docs, or asked about tuning and trials.

Red Hat Presence

Focus on customers **already running OpenShift** who are modernizing applications and well-positioned to extend their platform into AI workloads.

Qualifying Questions

TD SYNNEX

AI Maturity & Readiness

Are you currently exploring or implementing generative AI solutions? If yes, in what capacity?
What are you trying to optimize or create by using Gen AI in your business processes?
Have you identified the data sources you'd require for your use cases?
Do you have a centralized AI strategy or are initiatives decentralized across teams?
What platform do you leverage for Predictive and/or Generative AI today?

Infrastructure Alignment

Which of the following tools/libraries are your teams using today (Python, Hugging Face, PyTorch, etc.)?
Are your teams leveraging any pre-trained models or APIs (e.g., OpenAI, Hugging Face, Cohere, Mistral)?
What is your current cloud/on-prem/edge strategy for AI/ML?
Do you require or use hardware acceleration (e.g., GPUs, TPUs)? If yes, what type of hardware?
Have you identified or procured infrastructure to support these use cases?
Do you scale using containerization? Do you create base images for DS teams? Where are these stored?

Organizational Readiness

Which departments own AI/ML initiatives in your organization?
What roles and expertise make up your AI/ML projects?
Are there any lines of business integrating AI/ML into current applications?
Do you partner with external consultants or System Integrators for AI/ML efforts?

Top Challenges & Priorities

What challenges are you facing with your current AI/ML efforts?
What technical or resource challenges are holding you back?
What governance or compliance concerns are you navigating?
What are you hoping to improve with GenAI in your business?
What are your biggest priorities when evaluating and choosing an LLM provider?

Execution Path Discovery Questions

TD SYNNEX

Paths	InstructLab + Granite	InstructLab + OSS Models	Customized InstructLab	vLLM Inferencing	OpenShift AI
Customer Characteristics	• New to GenAI • Have basic internal content (FAQs, SOPs, policy docs) • Looking for quick value or internal buy-in • Not focused on specific models	• Technical users or AI/ML teams • Already testing OSS models like Llama or Mistral • Interested in flexibility and benchmarking • Want to validate model performance before scaling	• Regulated industries, multilingual environments, or specialized workflows • Require domain alignment (legal, medical, technical content) • Want models to behave or respond in a specific, structured way	• Already running GenAI models (any model) • Bottlenecked by inference latency or cost • Using GPUs but can't scale effectively • Looking for a better way to run—not tune—models	• Already running OpenShift in production (on-prem, cloud, or hybrid) • Interested in extending platform value into AI/ML • May already be testing AI models but lack infrastructure to scale or manage them
Discovery Questions	"Would you like to test GenAI using your existing documents like FAQs, policies, or manuals?" "Do you want a fast and simple way to show value from GenAI internally?" "Do you care which model you use, or are you just looking for something that works?"	"Are you currently experimenting with open source models like Llama or Mistral?" "Would it help to compare performance across different models using your own data?" "Do you prefer using open models over vendor-controlled offerings?"	"Is your content domain-specific or hard for general models to understand?" "Do you want the model to generate responses in a certain tone, structure, or format?" "Is your content multilingual or written in a language other than English?"	"Are you already running GenAI models in production or at scale?" "Are you hitting any limits with inference performance, cost, or GPU capacity?" "Would you benefit from faster inference without changing or re-training your model?"	"Are your internal teams asking for AI or model deployment support on OpenShift?" "Would you benefit from managing both app and AI workloads on a single platform?" "Are you starting to think about how to monitor, scale, and retrain models on your existing platform?"

Learn more: Discovery Guide

Making IT Personal™

GenAI POC Objection Handling - Top 5

TD SYNNEX

O: "We're not ready to commit to AI yet."
R: That's completely fair – a lot of teams are still figuring out where AI fits. The nice thing about a proof of concept is that it's not a big commitment. It's more like a small test run, just to see what's possible in one specific area. If it doesn't make a difference, no pressure to move forward. But if it does, you'll have real data to help with next steps.

O: "We're concerned about the cost of even a trial PoC."
R: Totally get that – budgets are tight. We try to keep these PoCs simple and affordable by focusing on just one goal or workflow. If there's something you really want to try out but cost is a blocker, let's talk through it. Sometimes we can find a way to make it work without a big spend upfront.

O: "We don't know what use case to start with."
R: That's okay – you don't need to have it all figured out. We can help you narrow it down by looking at where your team is spending the most time or where there are repeat questions or manual steps. Even something small can be a great starting point.

O: "We're concerned about data privacy and governance."
R: That's a really important concern, and we hear it often. For the PoC, we can keep everything running in your environment so your data stays in your hands. We're also happy to loop in your security or compliance teams early on to make sure everyone's comfortable with how it works.

O: "What happens after the PoC? We don't want to get stuck in something long-term."
R: Good question – the PoC doesn't lock you into anything. It's designed to give you a clear picture of value without a long-term commitment. If it works, great – we can talk about the next steps. If not, there is no pressure to move forward.

Sample Sales Pitch

TD SYNNEX

Opening the conversation

Seller: "Thanks again for taking the time. A lot of our customers are starting to explore how AI can help their teams be more efficient—especially when it comes to answering repeat questions or dealing with large volumes of information. Is that something your team is thinking about right now?"
Customer: "Yeah, we've definitely been talking about it. We're still figuring out where it could help."

Introducing the idea

S: "One of the easiest ways we've seen teams get started is with a chatbot trained on their own content—things like internal FAQs, or support docs. It gives employees or customers instant answers to common questions, so your team spends less time replying to the same things over and over."
C: "Okay, that sounds helpful—but how much effort does it take to set up?"

Lowering the barrier

S: "Great question. We make it really simple. We run a short proof of concept—usually just a couple days—where we set up the chatbot using your documents. You get to see how it works, how accurate the answers are, and whether it could save your team time. It's low risk, doesn't require a big project, and it helps you make a decision based on real results."
C: "So we'd use our own content, not some generic training data?"
S: "Exactly—your content, your environment. That way it's actually useful to your business from day one."

Platform tie-in (if OpenShift is in play)

S: "And if you're already running OpenShift, we can do this right on your existing platform—no need to spin up anything new. It all stays in a setup your teams already know and trust."
C: "Nice. So we can try it out without a lot of extra work."

Closing with next steps

S:"Exactly. If this sounds like something worth exploring, I can help you kick off a POC. All we'd need is a few sample documents, and we'll walk you through the rest.
C: "Alright, let's give it a shot. Send me what you need to get started."
S: "Perfect—I'll send over the POC request info and a quick checklist. We'll make it easy."

Making IT Personal™

Next best steps you can take

Reach out, learn more and get hands-on experience

TD SYNNEX

0 1

Try RHEL AI

A single, 60-day,
self-supported subscription to
Red Hat Enterprise Linux AI

0 2

Try OpenShift AI

A single, 60-day,
self-supported subscription to Red
Hat OpenShift AI (Self-Managed)

0 3

Enablement

Red Hat Partner Connect:

- (Tech-) Sales Enablement
- Technical certification:
  - RHCSA (EX200)
  - OpenShift Admin (EX280)
  - OpenShift AI (EX267)

Reach out: TD SYNNEX Redhat CH
TD SYNNEX Redhat DE

0 4

Pricing & Sizing

Red Hat AI – Fragen zu
Subscriptions und Prozessen?
Ihr TD SYNNEX Red Hat
Vertriebsteam hilft Ihnen gerne!

Making IT Personal™

What is Next?

Once you have Identified your PoC. Please get with your TD SYNNEX Red Hat Engineer and we can get your customers PoC started today.

Links:

TD SYNNEX Academy

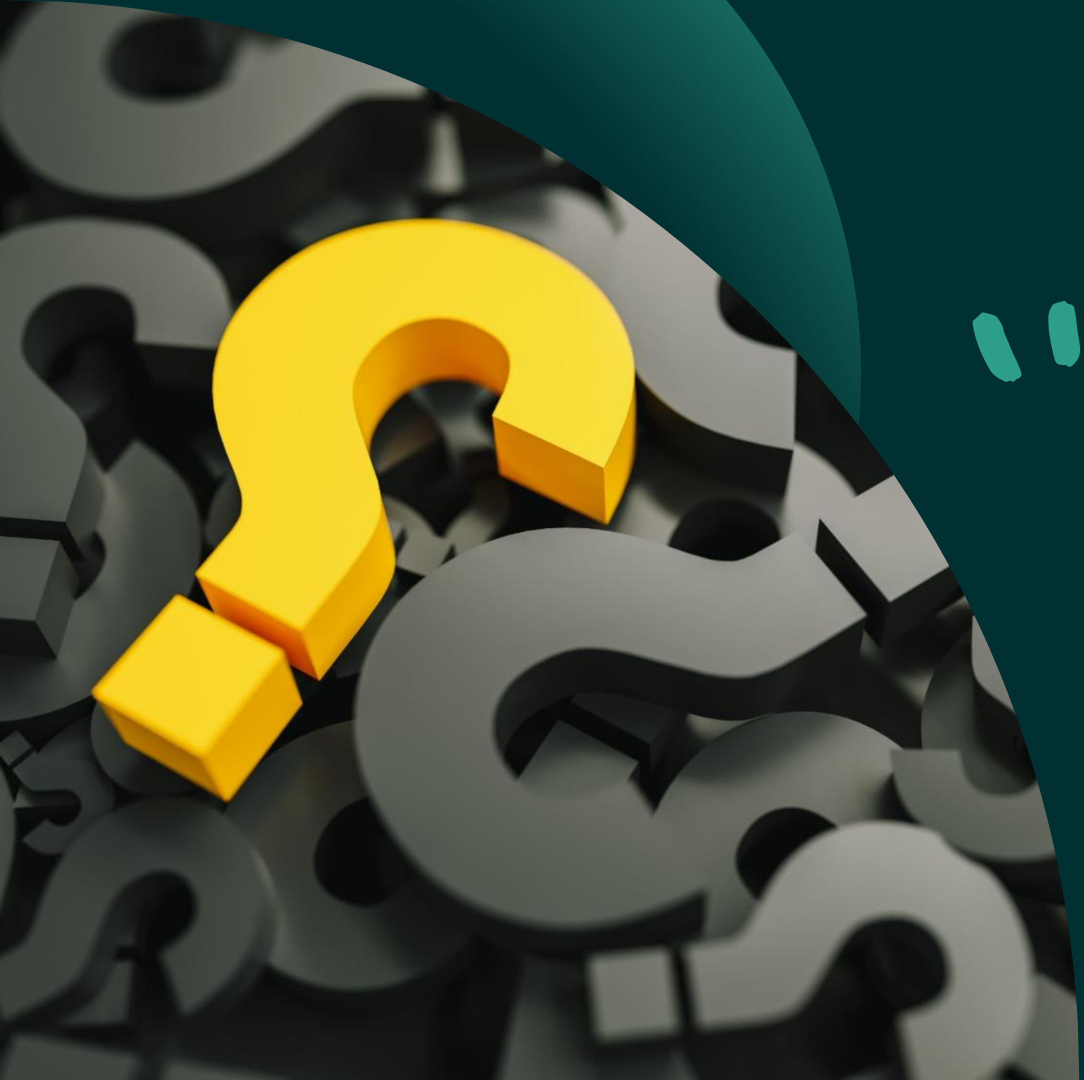
Red Hat AI learning hub | Red Hat Documentation

Interactive experiences | Red Hat products and solutions

Red Hat Interactive Learning Portal | Red Hat Developer

Red Hat AI | product page

Making IT Personal™



TD SYNNEX

Questions & Answers

Making IT Personal™

TD SYNNEX

Making IT Personal™