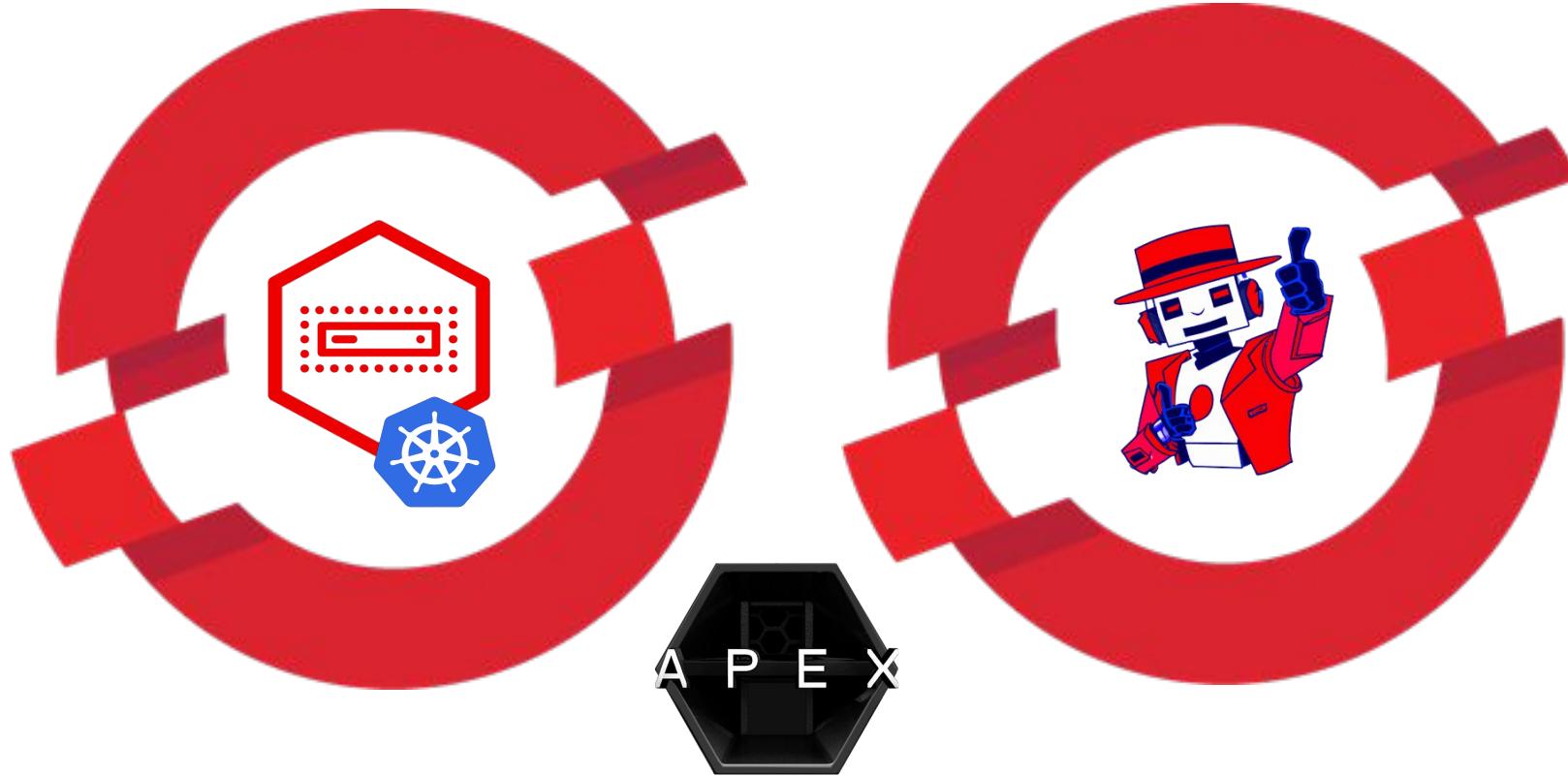


See how **Virt & AI** meet **OpenShift**

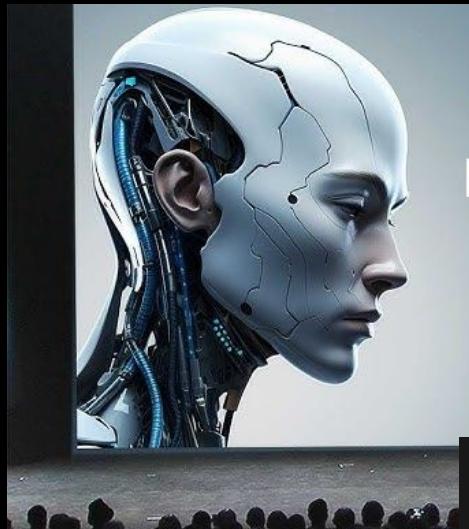


and featuring DELL **APEX** Cloud Platform

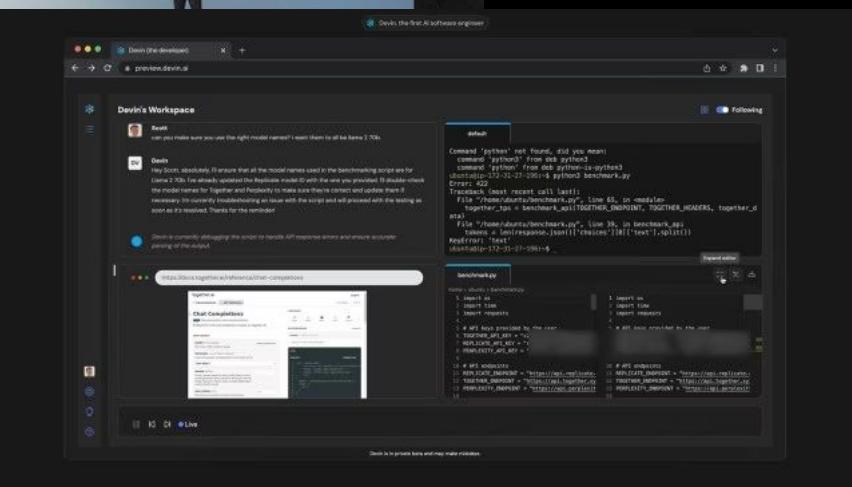
Melbourne OpenShift Meetup

14 March 2024 - 6:00pm, Lvl 23, 120 Collins St

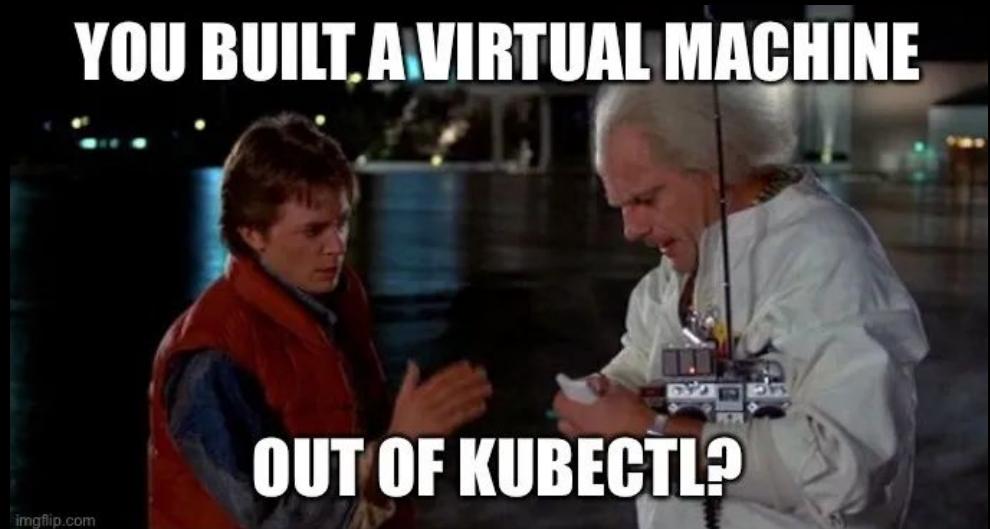
RSVP: <https://www.meetup.com/en-AU/melbourne-openshift-meetup/>



DEVELOPMENT is DEAD



YOU BUILT A VIRTUAL MACHINE



OUT OF KUBECTL?

Agenda:

Welcome

Bryon Baker & Neo Xu

- ▶ OpenShift AI

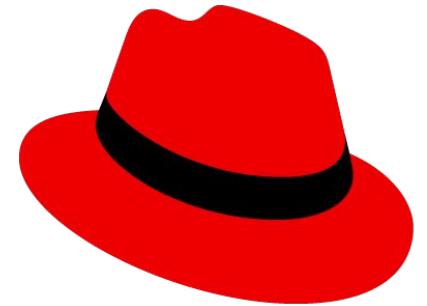
Chris Butler

- ▶ OpenShift Virtualisation

Theo Crithary

- ▶ Dell Technologies – APEX Cloud Platform for Red Hat OpenShift

Bryon Baker & Neo Xu



Red Hat OpenShift AI



OpenShift AI

Overview of Red Hat OpenShift AI

MLOps platform for artificial intelligence/ machine learning (AI/ML) use cases

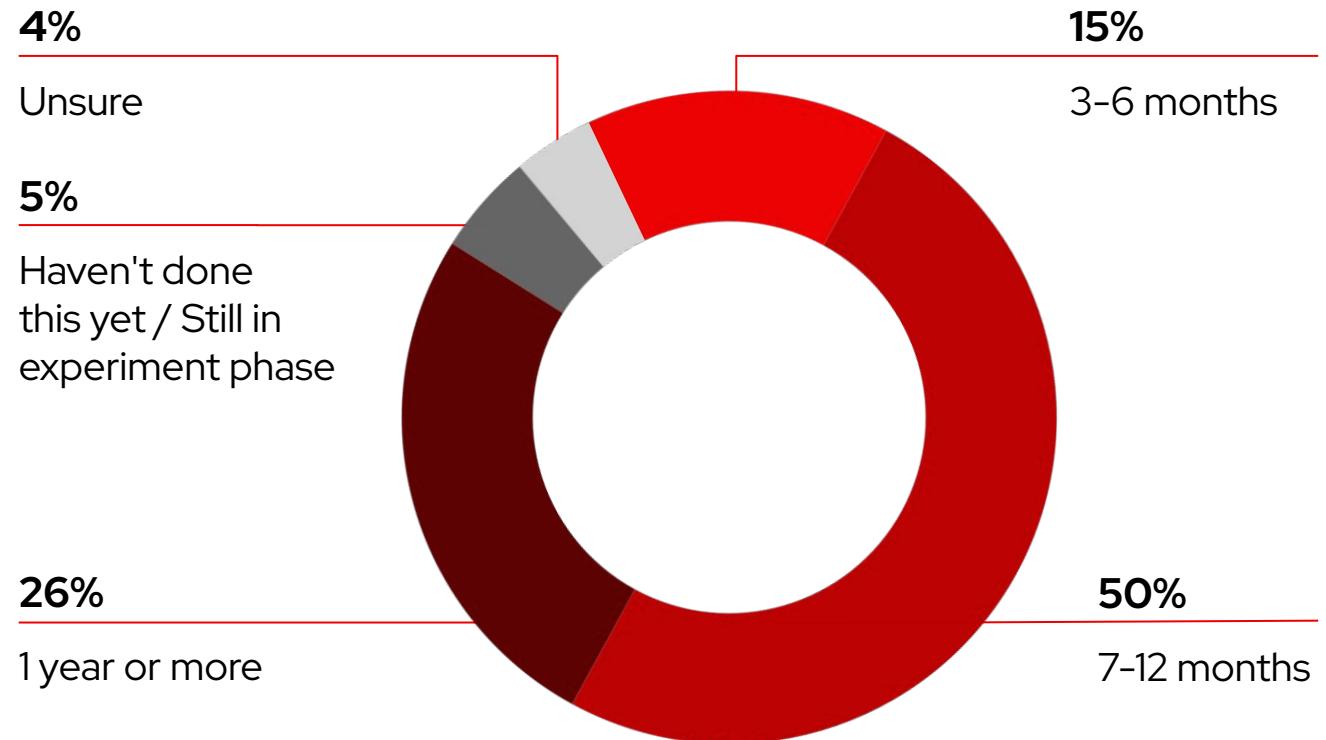
Bryon Baker
Specialist Solution Architect

Neo Xu
Senior Consultant

Operationalizing AI is still a challenging process

What is the average AI/ML timeline from idea to operationalizing the model?

Half of respondents (50%) say their average AI/ML timeline from idea to operationalizing the model is 7-12 months.



Understanding the Industry Forces at Play

Summary

Competition in the industry

The rapid growth in the AI software market is:

- ▶ Making it harder for technology vendors to offer an e2e differentiated offering.
- ▶ Overwhelming IT orgs with models sizes & number of AI-enabled apps to manage and monitor.
- ▶ Increasing need to show tangible ROIs on AI investments.

Creating a niche market

Hyperscalers are making it difficult for customers to make choice. *Customers want models customised to their business scenarios requiring larger upfront investments.*

Customers should focus on

Customers are *leaning into composable technologies* to gain flexibility, adaptability and freedom of choice. No vendor lock-in.

A different approach

Vendors offering solutions that let customers meet you where your data is at. *Solves data compliance requirements, security and government regulations.*

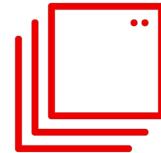
Emerging Vertical Markets

Intelligent AI-infused industry platforms that solve for specific use cases could *become substitutes* of dedicated data science and AI/ML platforms.

Red Hat's AI/ML Focus



Red Hat OpenShift AI



AI-optimized infra



AI-enabled products

Accelerate time to value of
AI-enabled apps and models

An add-on to OpenShift

Optimize underlying platforms
for AI-deployments

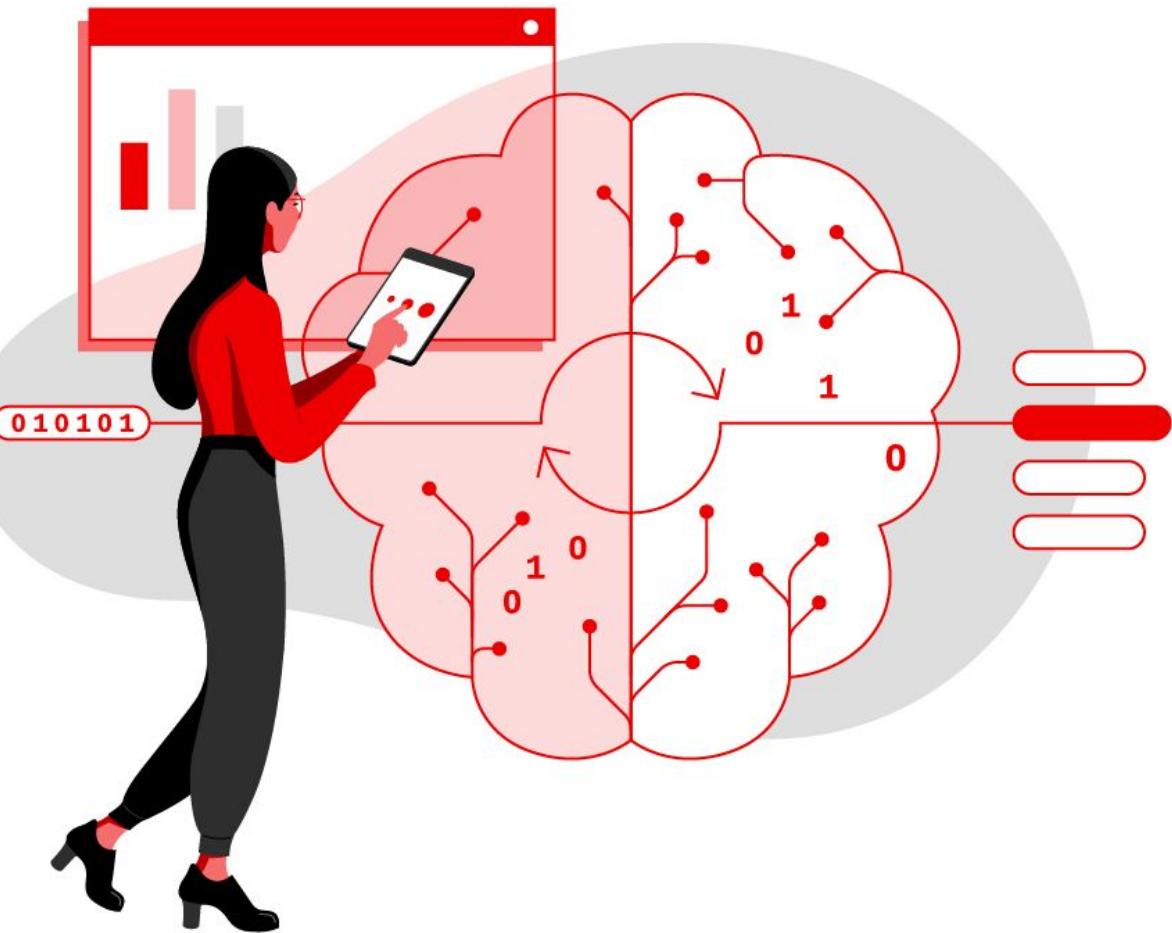
RHEL, OCP, GPU/Accelerators

**Drive adoption of existing
products** with AI-enabled tools

Ansible Lightspeed

**Fostering an AI Partner
Ecosystem**

Empower choice with best-of-breed **AI technologies from a certified
partner ecosystem**



Red Hat OpenShift AI

An AI-focused platform that provides tools to train, tune, serve, monitor and manage AI/ML experiments and models.

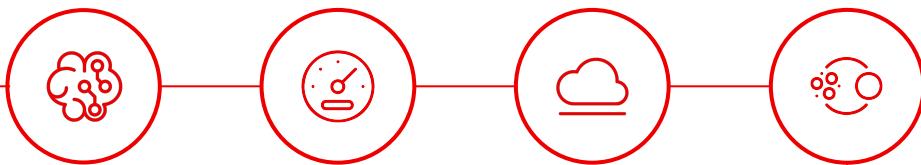
AI for the open hybrid cloud

Enterprise-grade open source hybrid AI and MLOps platform



Red Hat
OpenShift AI

Develop, train, serve, monitor, and manage the life cycle of AI/ML models and applications, from experiments to production.



- ▶ Provide a unified platform for data scientists and intelligent application developers
- ▶ Scale to meet the workload demands of foundation models: data volume, training time, model size, acceleration, and scalability
- ▶ Deliver consistency, cloud-to-edge production deployment and monitoring capabilities
- ▶ Underlying platform for training, serving, and tuning foundation models in Red Hat Ansible Lightspeed with IBM Watson Code Assistant

Red Hat OpenShift AI

Hybrid MLOps platform

Collaborate within a common platform to bring IT, data science, and app dev teams together

Available as

- managed cloud service
- traditional software product on-site
- or in the cloud!



Model development

Conduct exploratory data science in JupyterLab with access to core AI / ML libraries and frameworks including TensorFlow and PyTorch using our notebook images or your own.



Model serving & monitoring

Deploy models across any cloud, fully managed, and self-managed OpenShift footprint and centrally monitor their performance.



Lifecycle Management

Create repeatable data science pipelines for model training and validation and integrate them with devops pipelines for delivery of models across your enterprise.



Increased capabilities / collaboration

Create projects and share them across teams. Combine Red Hat components, open source software, and ISV certified software.

Red Hat OpenShift AI - Key features

Model development

Interactive, collaborative UI for **seamless access** AI/ML tooling, libraries, frameworks, etc.

Model serving

Model serving routing for **deploying models to production** environments

Model monitoring

Centralized monitoring for **tracking models performance and accuracy**

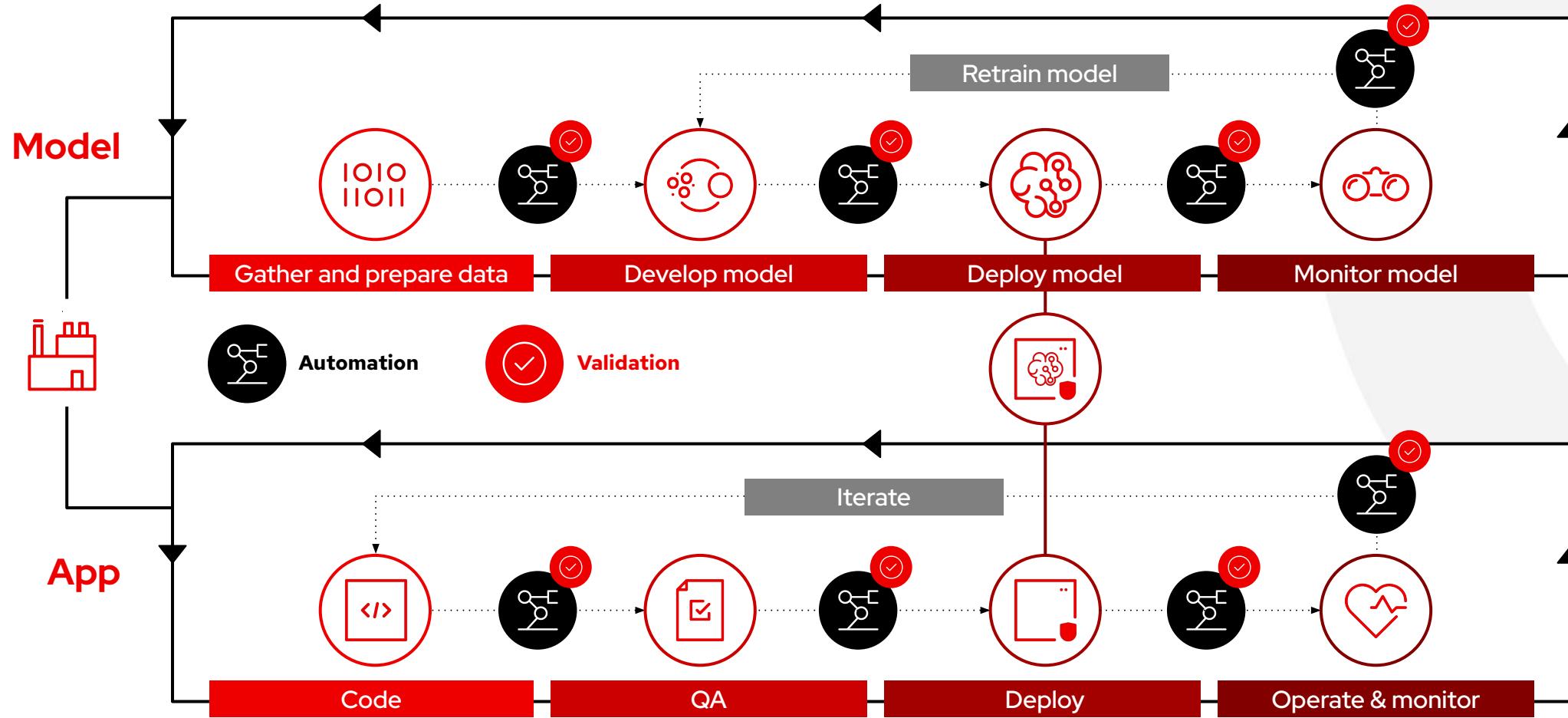
Data & model pipelines

Visual editor for **creating and automating** data science pipelines

Distributed workloads

Seamless experience for **efficient data processing, model training, and tuning**

Lifecycle for operationalizing models



Red Hat's engineering is 100% open source

Upstream
Projects



Red Hat
Enterprise Linux

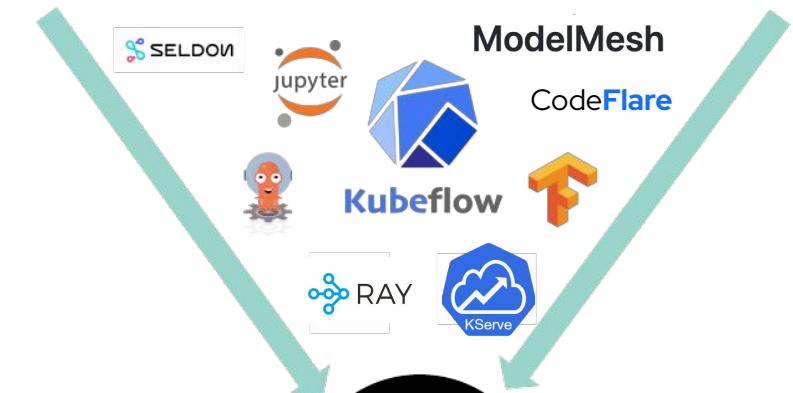
Community
Project



kubernetes



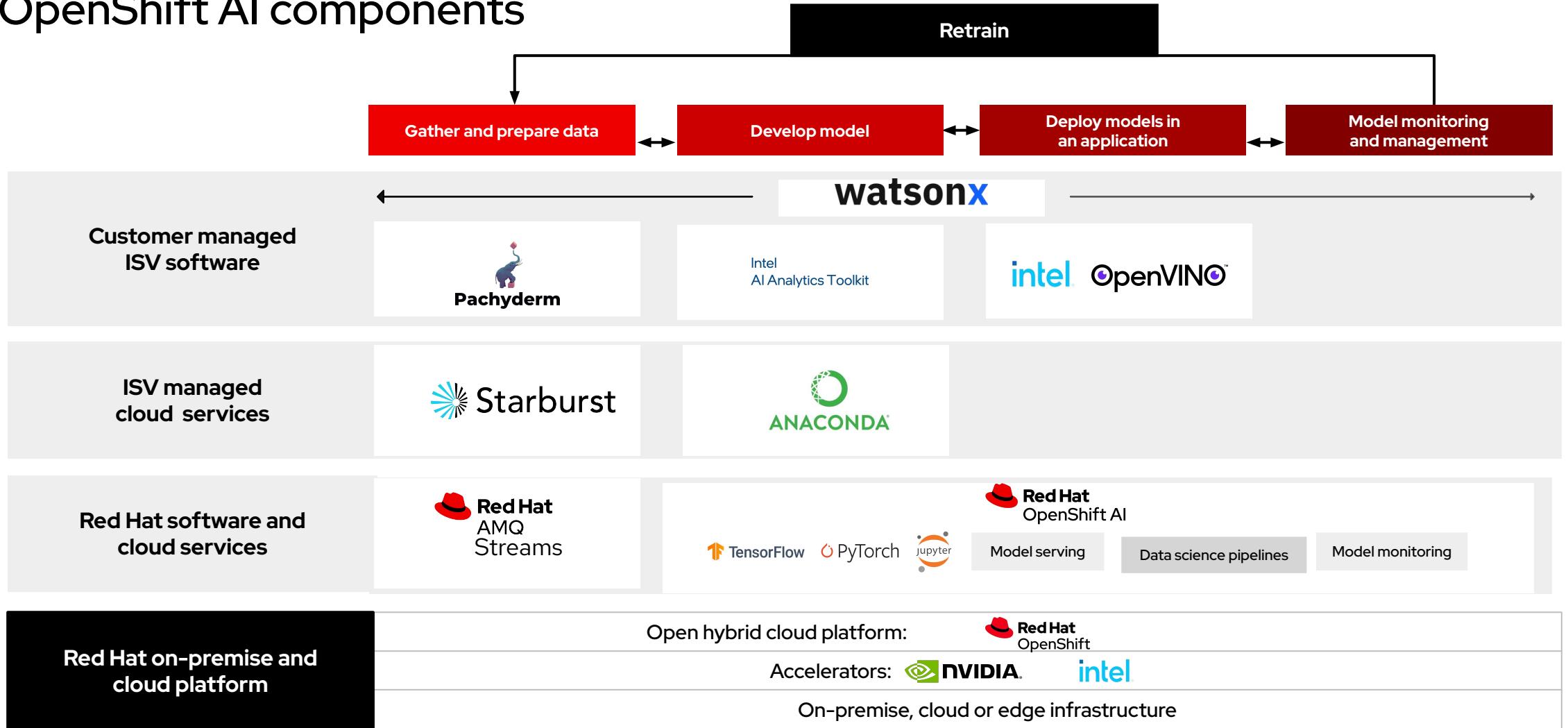
Product



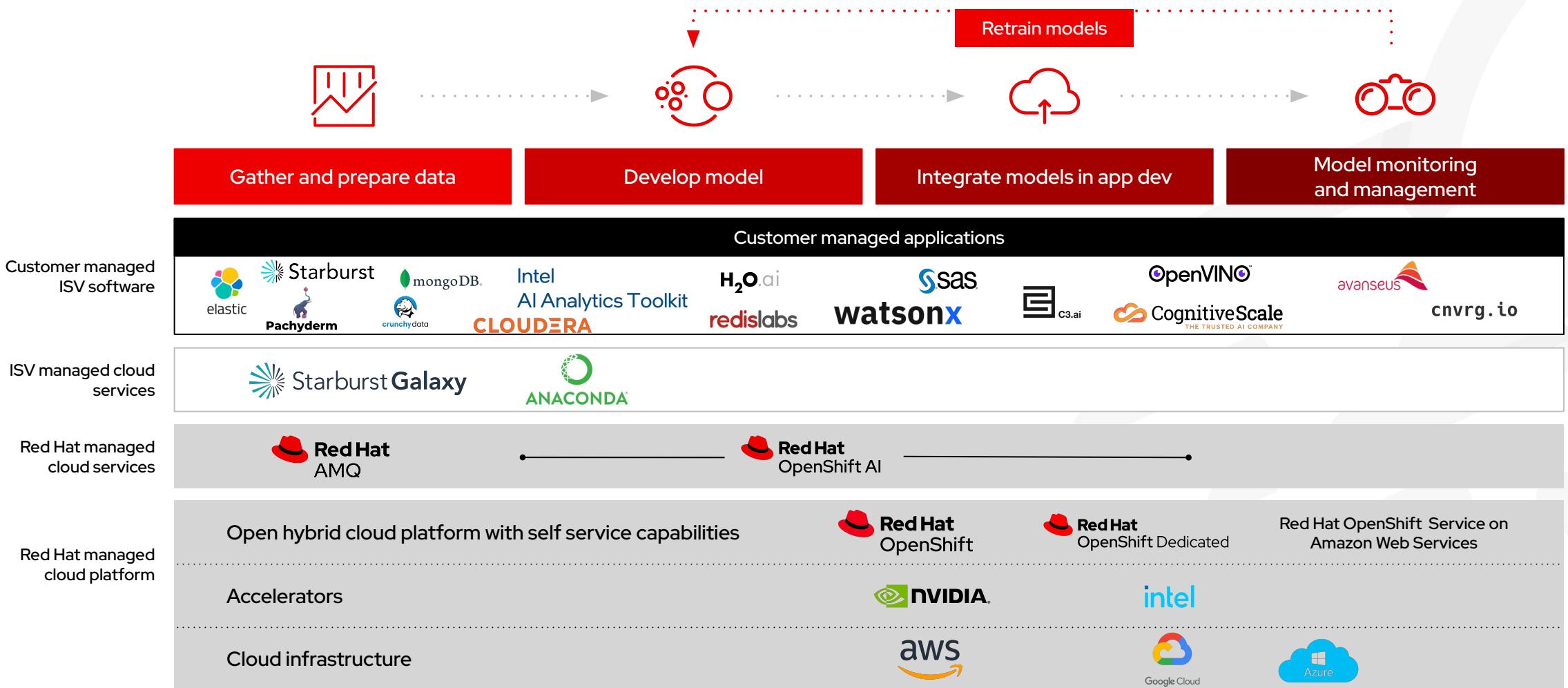
Red Hat
OpenShift AI



OpenShift AI components



... and integrating our partner ecosystem



Red Hat strategy around generative AI and foundation models

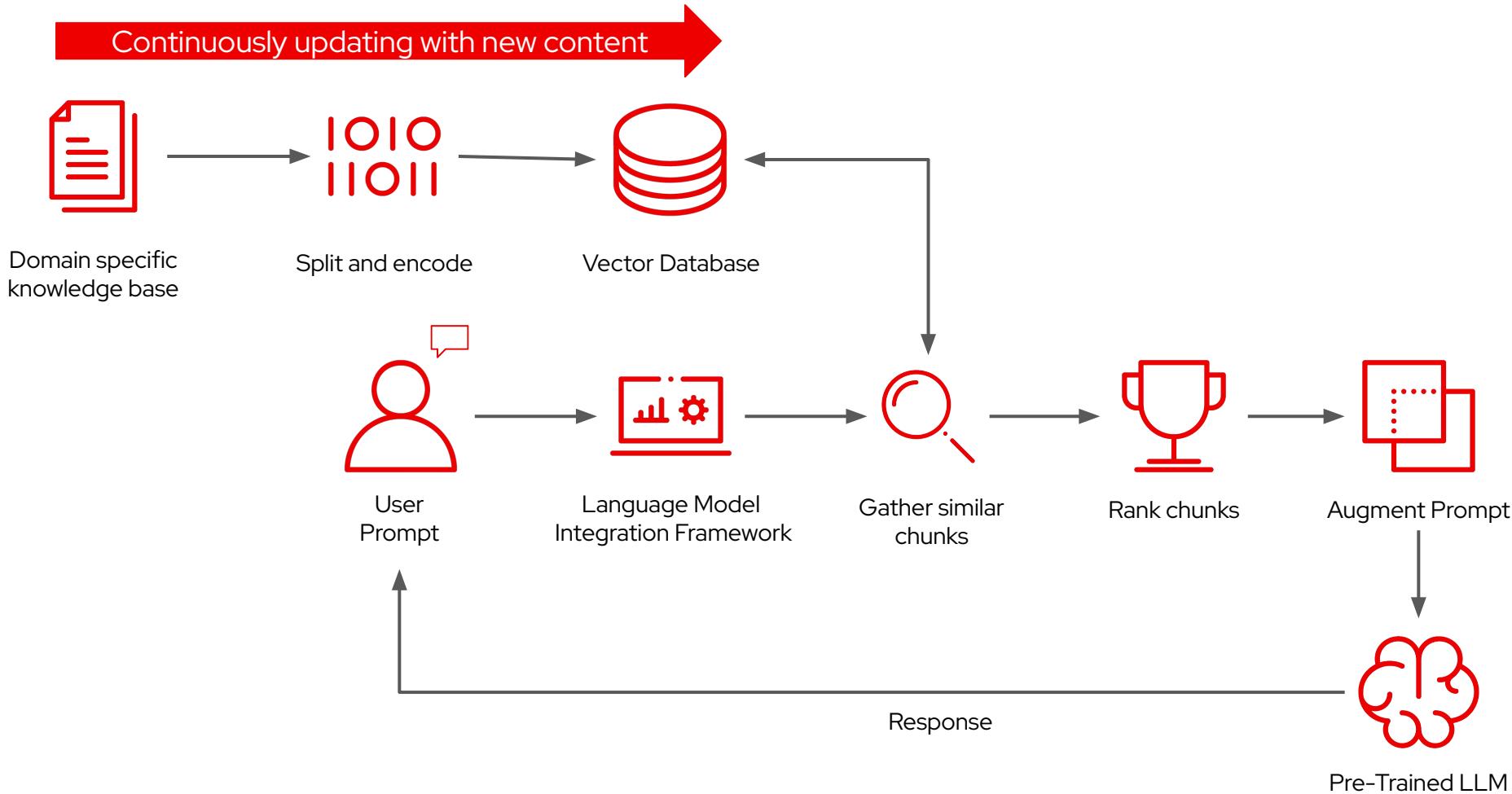


- ▶ Developing the infrastructure stack for distributed workloads, scheduling for building, prompt-tuning, fine-tuning and serving foundation models
 - OpenShift AI is a foundation layer for IBM watsonx.ai and Ansible Lightspeed with IBM Watson Code Assistant
-
- ▶ Partner with model builders to offer models with OpenShift AI
-
- ▶ Enable out-of-the-box “bring your own model” use cases
-
- ▶ Red Hat will infuse Generative AI capabilities into more of its portfolio.
 - Ansible Lightspeed is Generally Available to customers.

Demo time

Retrieval Augmented Generation (RAG)

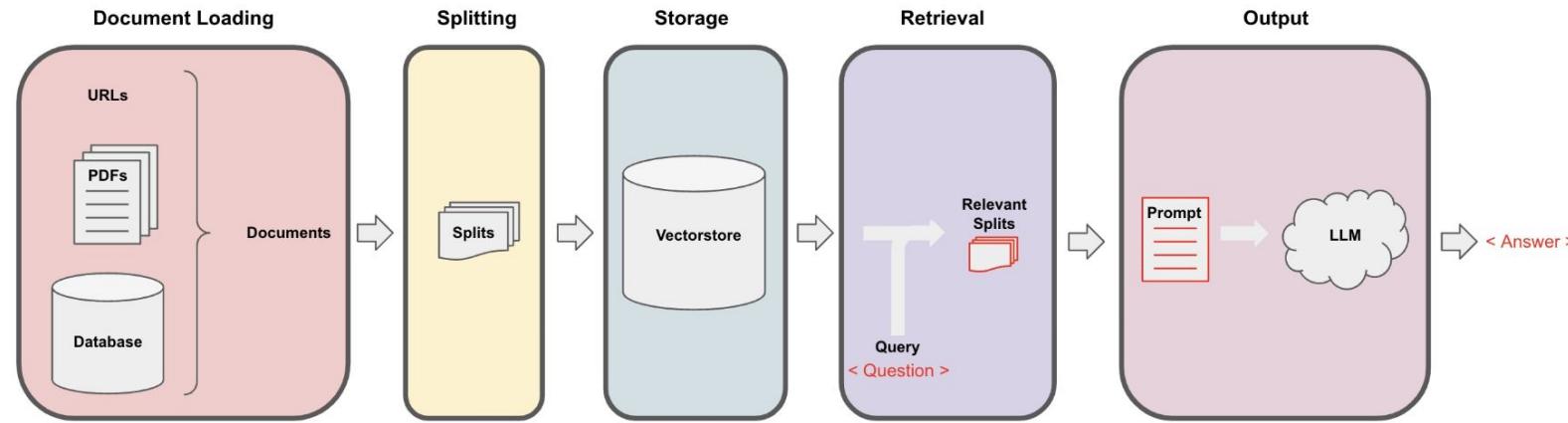
Helps the model to “look up” external information to improve generated text responses



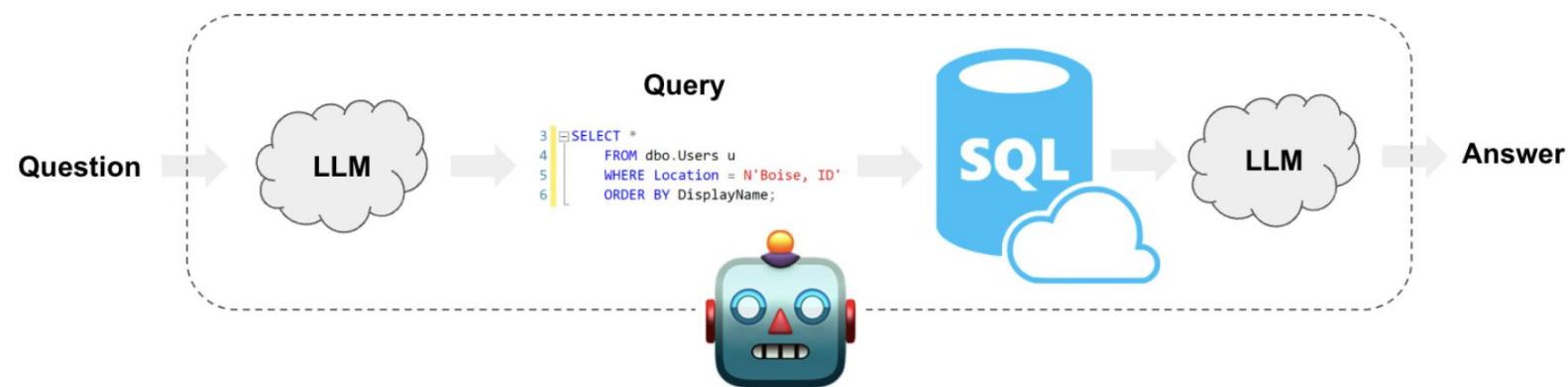
Build a Chatbot over Document

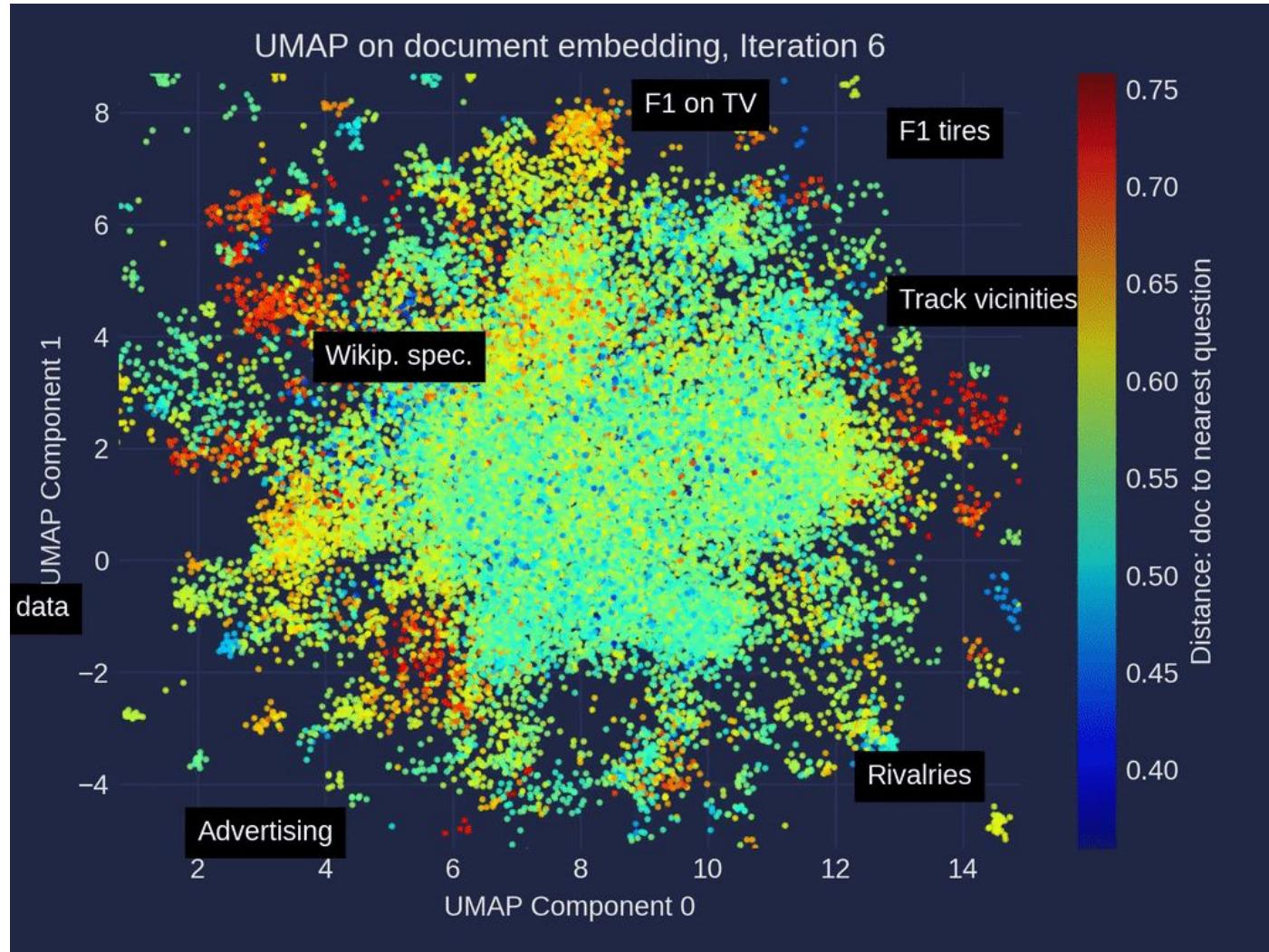
Enhance the Chatbot by leveraging the knowledge from documents

Question & Answering over unstructured document



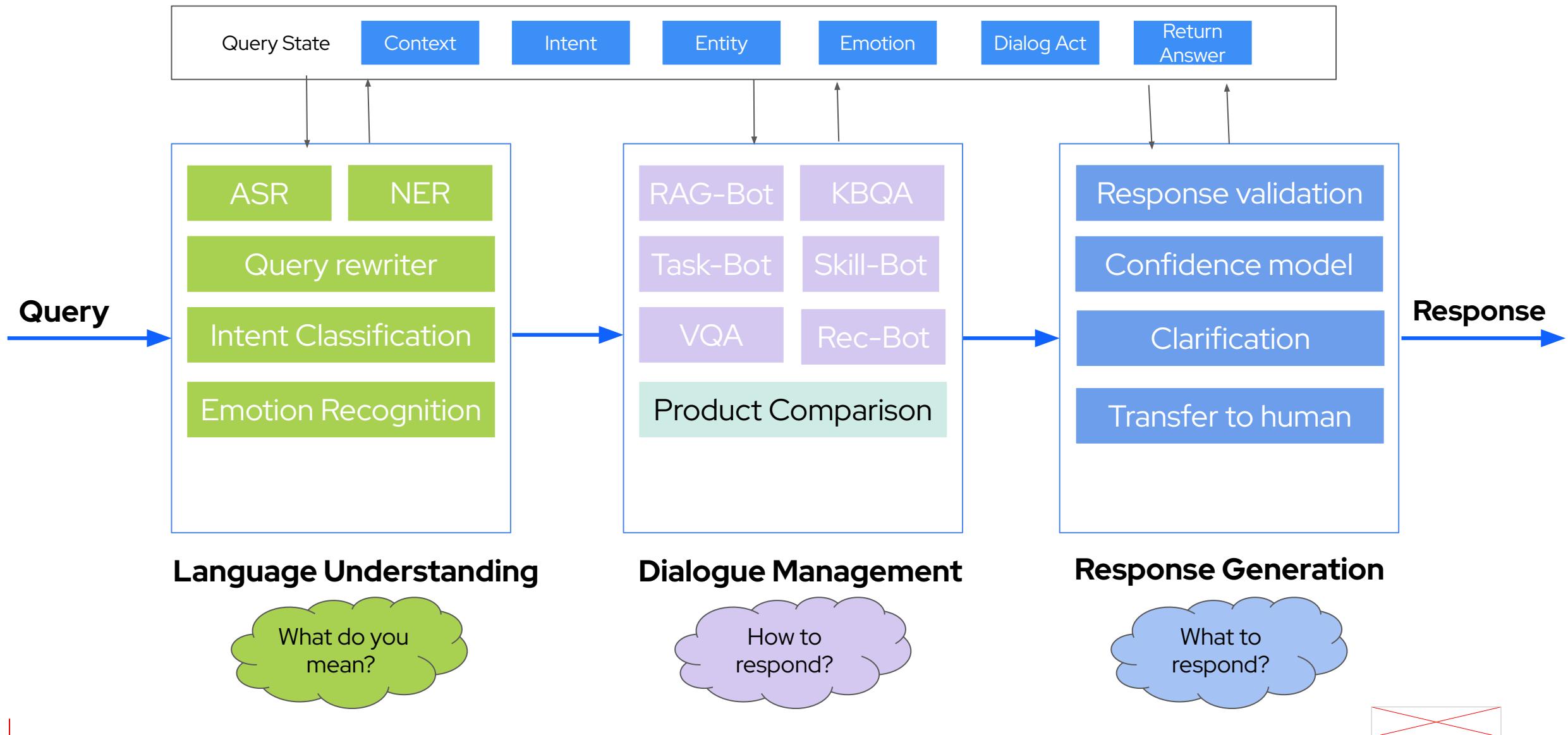
Question & Answering over structured database



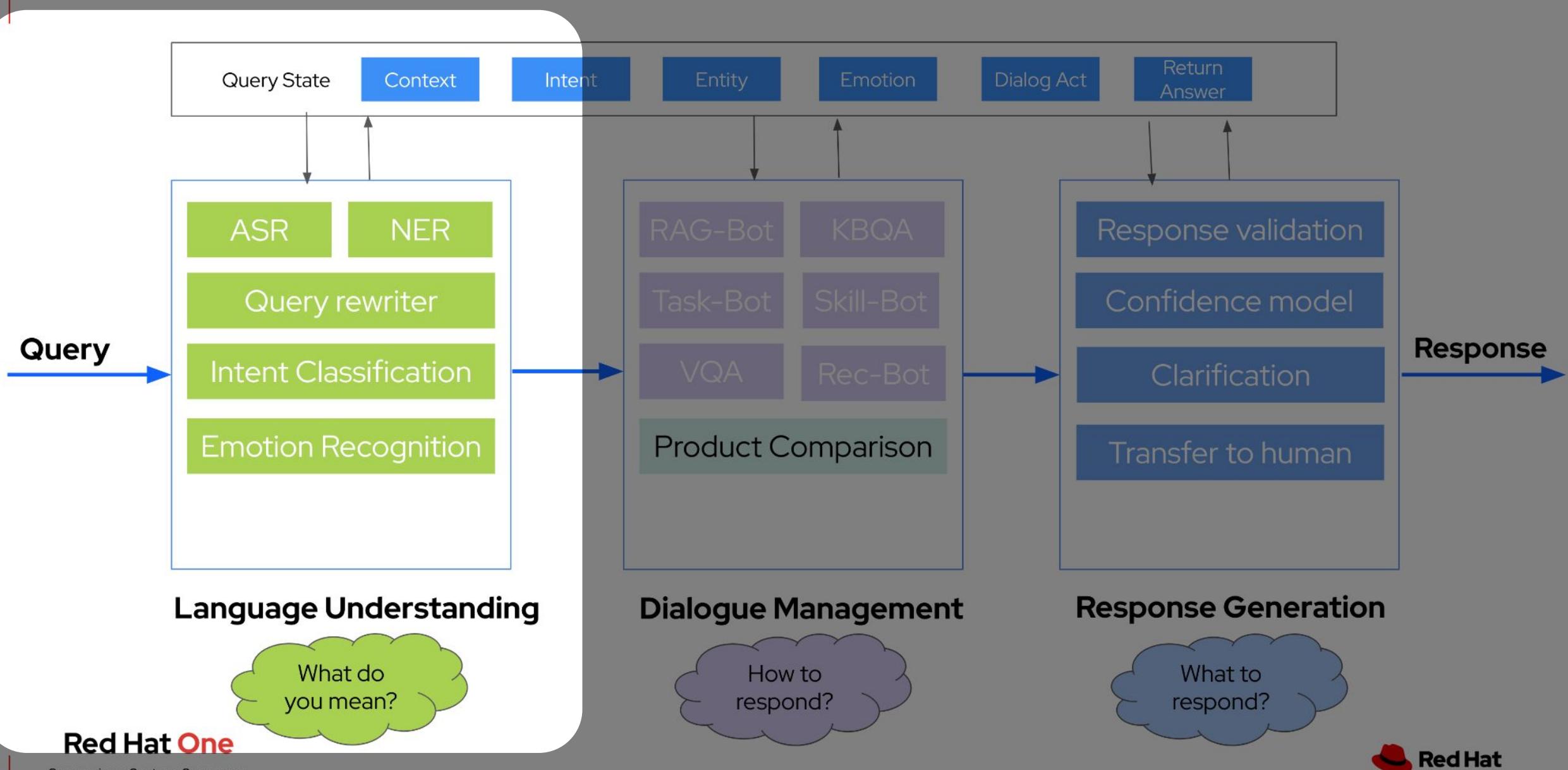


UMAP dimensionality reduction of the embeddings of document snippets, colored by to their relevance to the question.

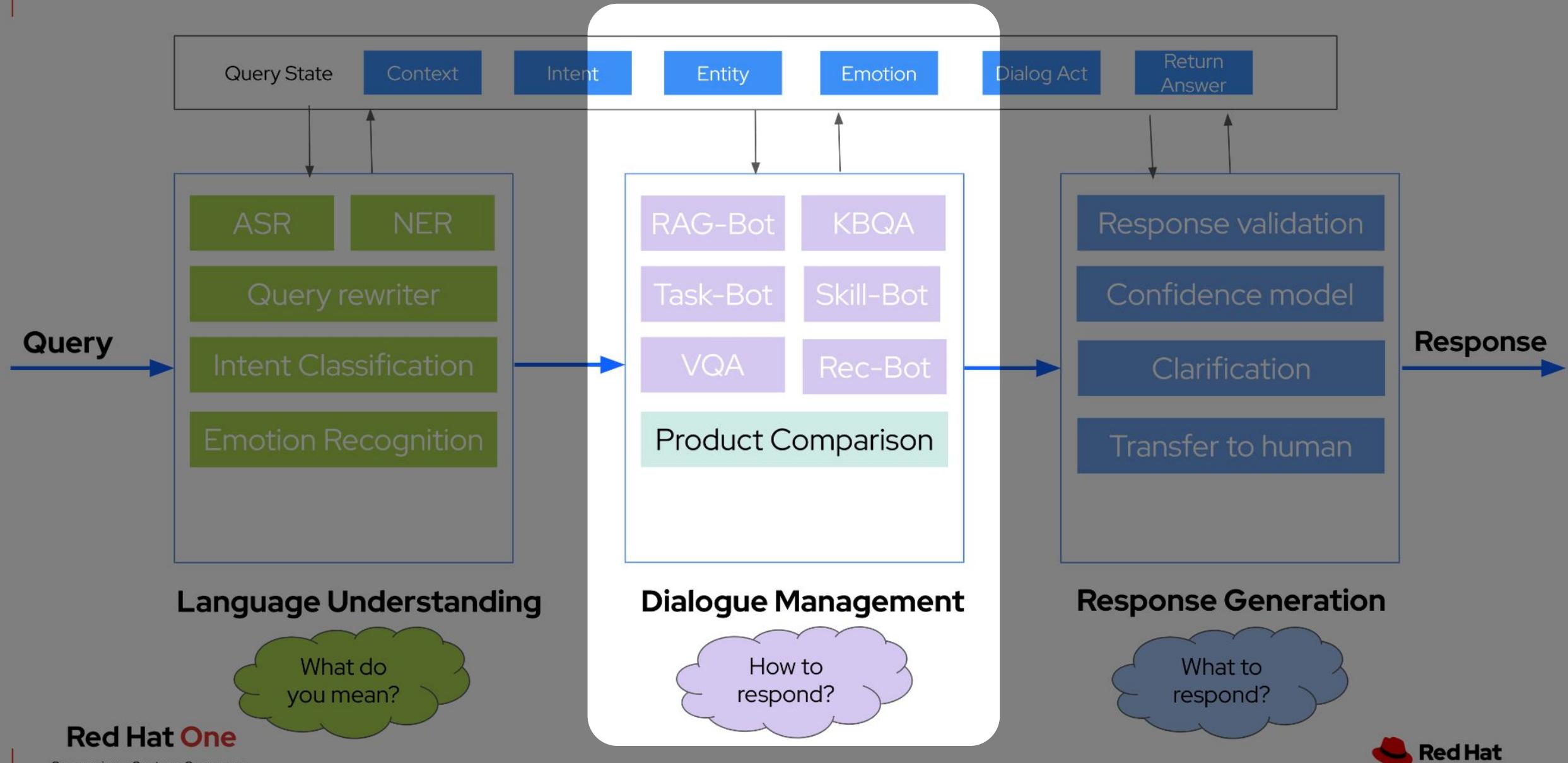
Guide on Functionality and Processes for LLM Applications



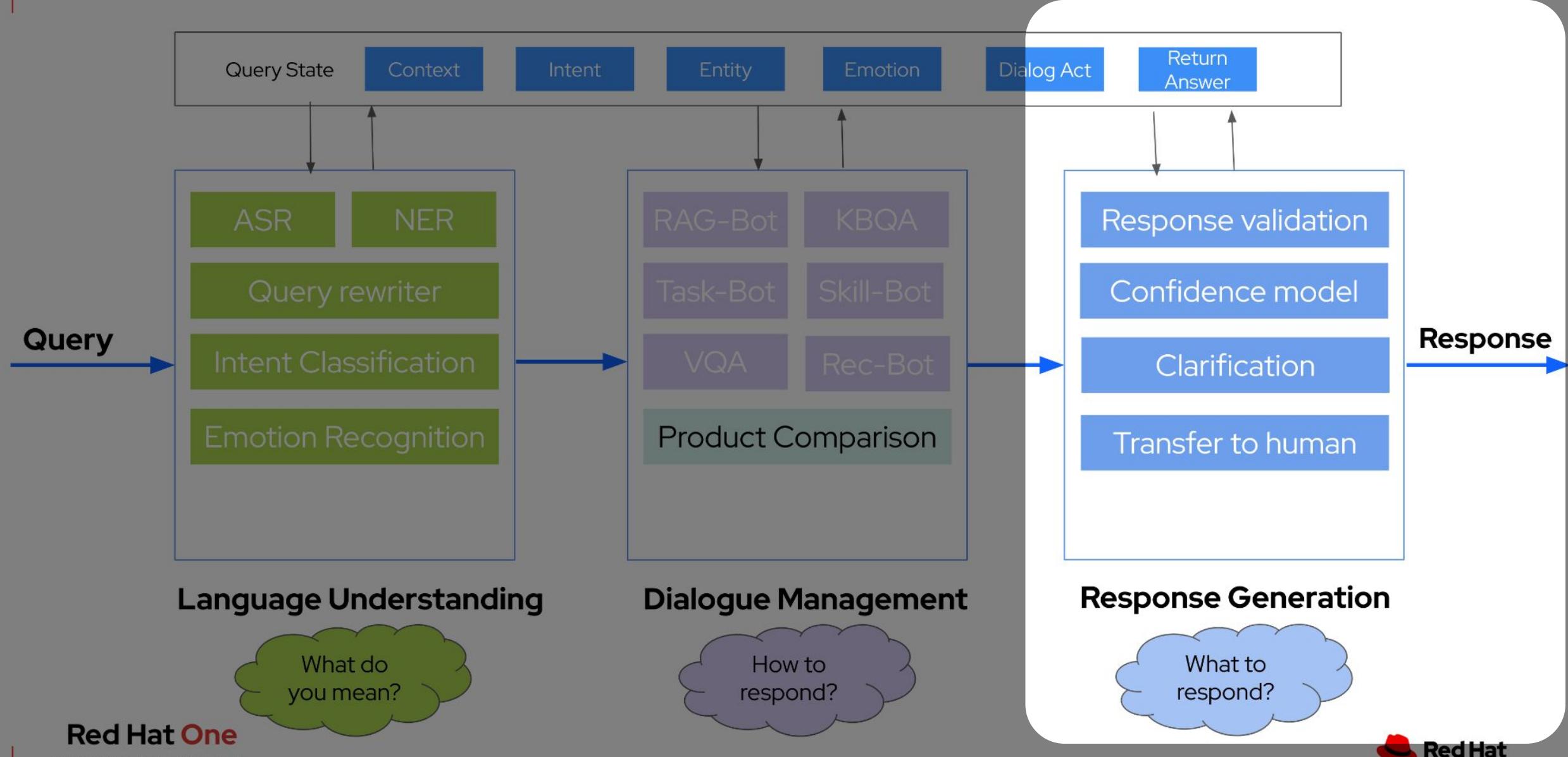
Guide on Functionality and Processes for LLM Applications



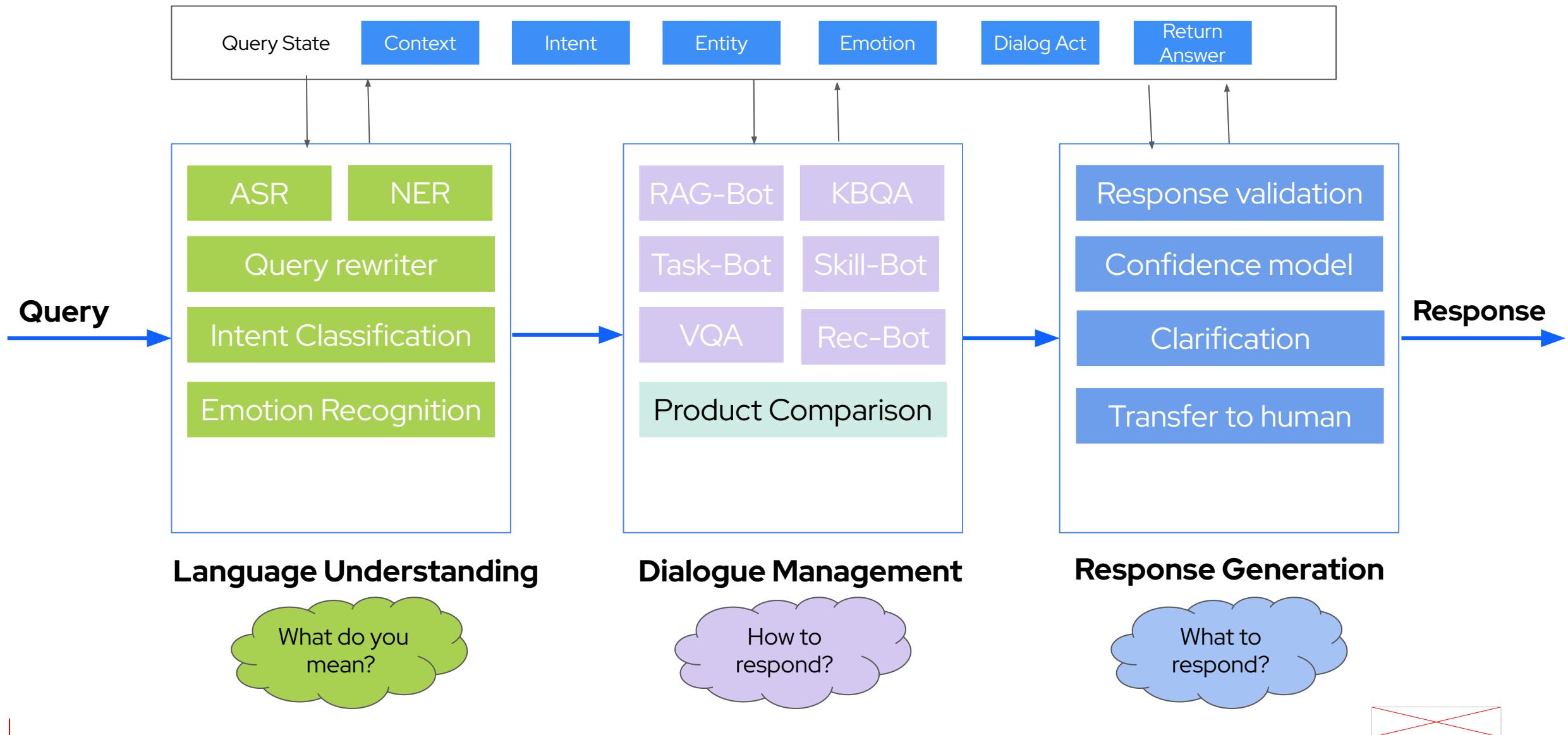
Guide on Functionality and Processes for LLM Applications



Guide on Functionality and Processes for LLM Applications



Guide on Functionality and Processes for LLM Applications



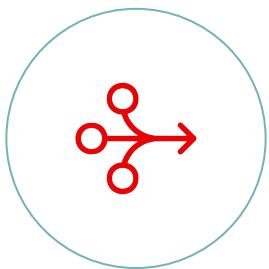
Wrap Up

What differentiates us



Simplify AI adoption

Promotes freedom of choice and access to latest innovation on AI/ML technologies



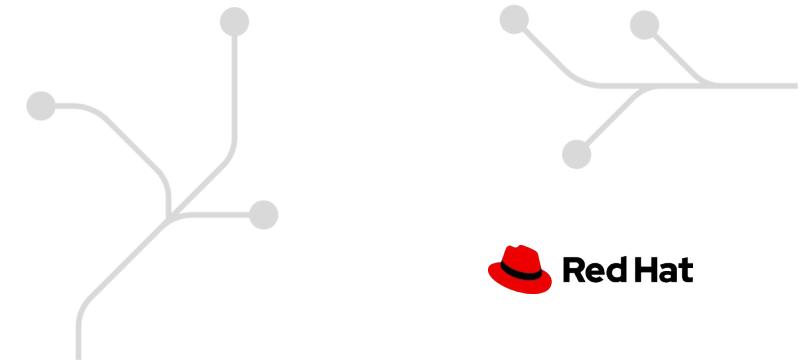
Drive AI/ML operational consistency

Streamline the process of moving models from experiments to production



Gain hybrid cloud flexibility

Deploy models in containerized format across on-prem, clouds and edge to comply with regulatory requirements



Thank you

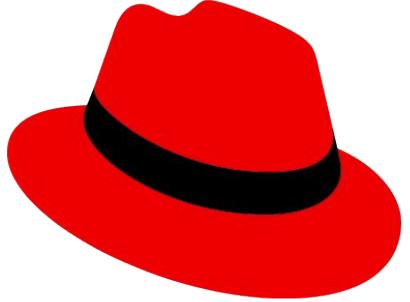
Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make Red Hat a trusted adviser to the Fortune 500.

 [linkedin.com/company/red-hat](https://www.linkedin.com/company/red-hat)

 [youtube.com/user/RedHatVideos](https://www.youtube.com/user/RedHatVideos)

 [facebook.com/redhatinc](https://www.facebook.com/redhatinc)

 twitter.com/RedHat



Chris Butler

Chief Architect

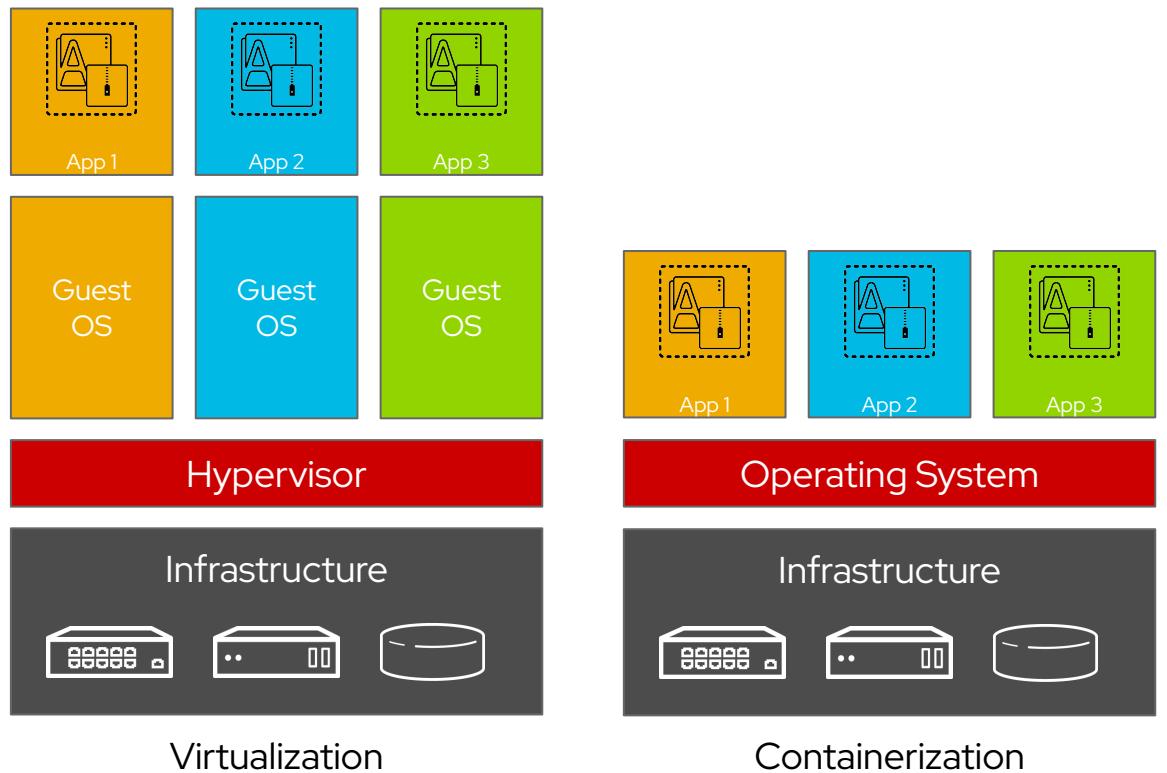
Red Hat Virtualization

OpenShift Virtualization

Dr. Chris Butler
Chief Architect
Field CTO Office

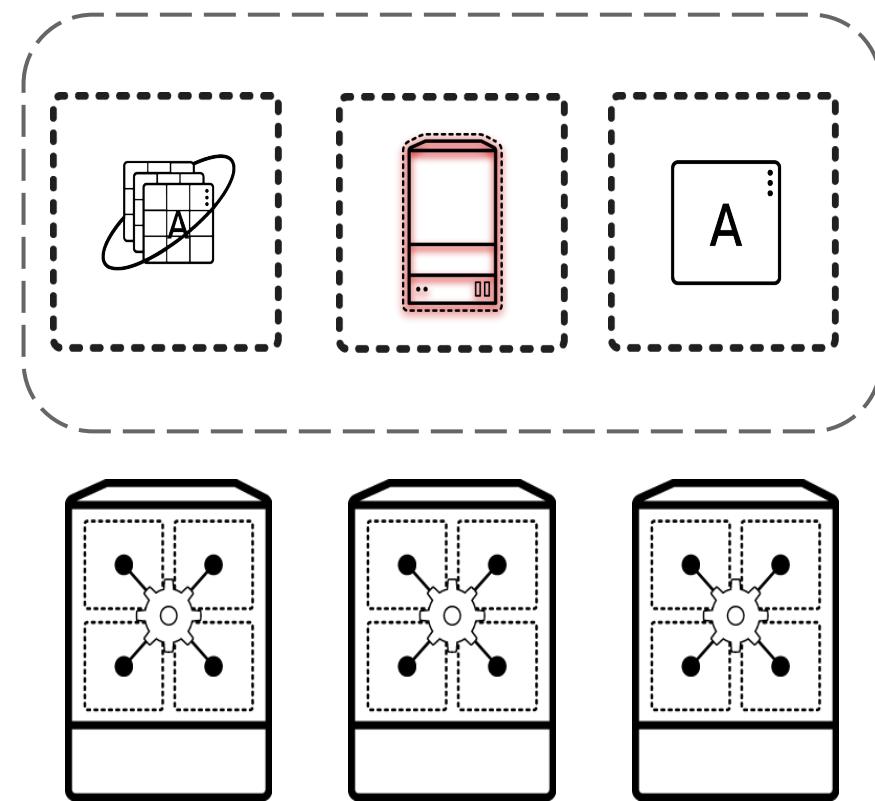
Containers are not virtual machines

- Containers are process isolation
- Kernel namespaces provide isolation and cgroups provide resource controls
- No hypervisor needed for containers
- Contain only binaries, libraries, and tools which are needed by the application
- Ephemeral



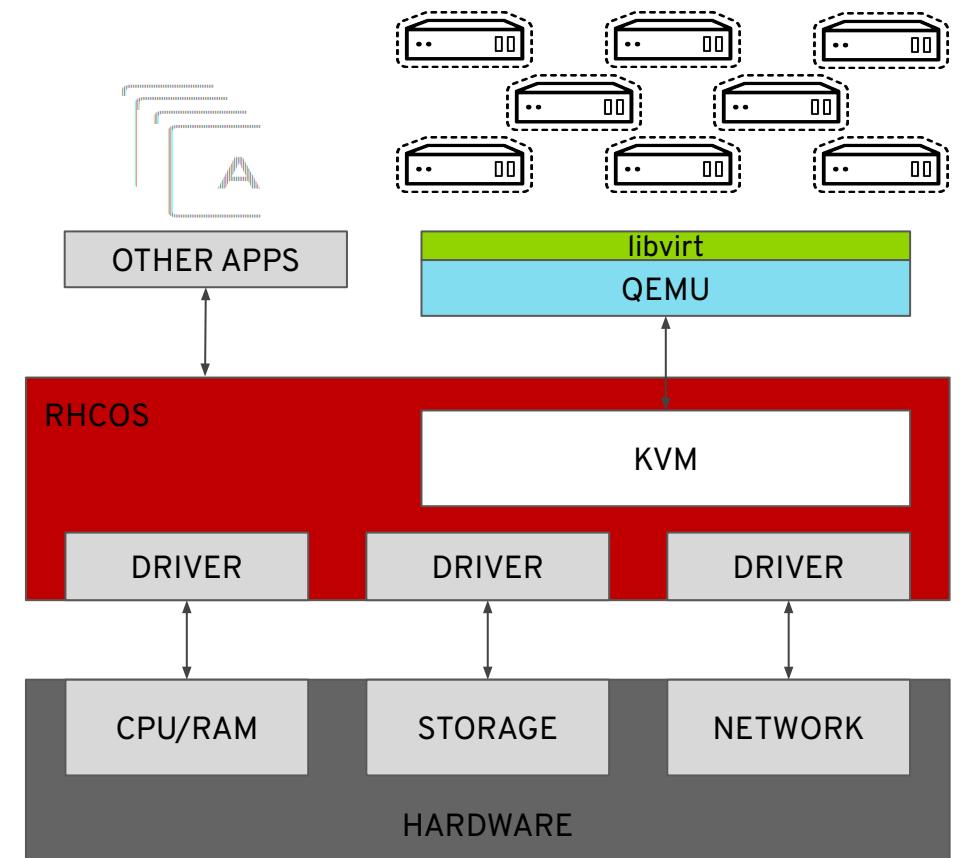
Virtual machines can be put into containers

- A KVM virtual machine is a process
- Containers encapsulate processes
- Both have the same underlying resource needs:
 - Compute
 - Network
 - (sometimes) Storage



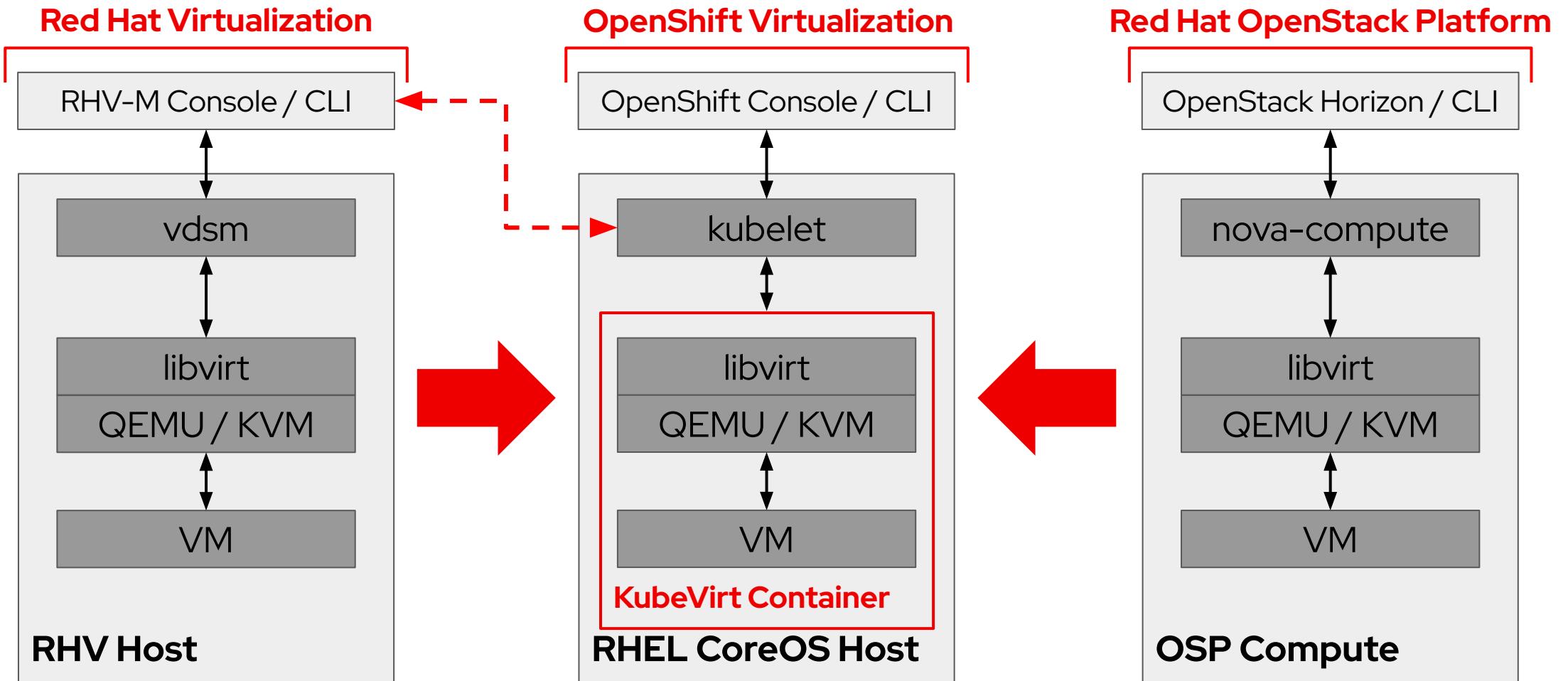
OpenShift Virtualization uses KVM

- OpenShift Virtualization uses KVM, the Linux kernel hypervisor
- KVM is a core component of the Red Hat Enterprise Linux kernel
 - KVM has 10+ years of production use: Red Hat Virtualization, Red Hat OpenStack Platform, and RHEL all leverage KVM, QEMU, and libvirt
- QEMU uses KVM to execute virtual machines
- libvirt provides a management abstraction layer
- Currently supported on x86 bare metal
- For other platforms contact Product Management for roadmap



Containerizing KVM

Trusted, mature KVM wrapped in modern management and automation

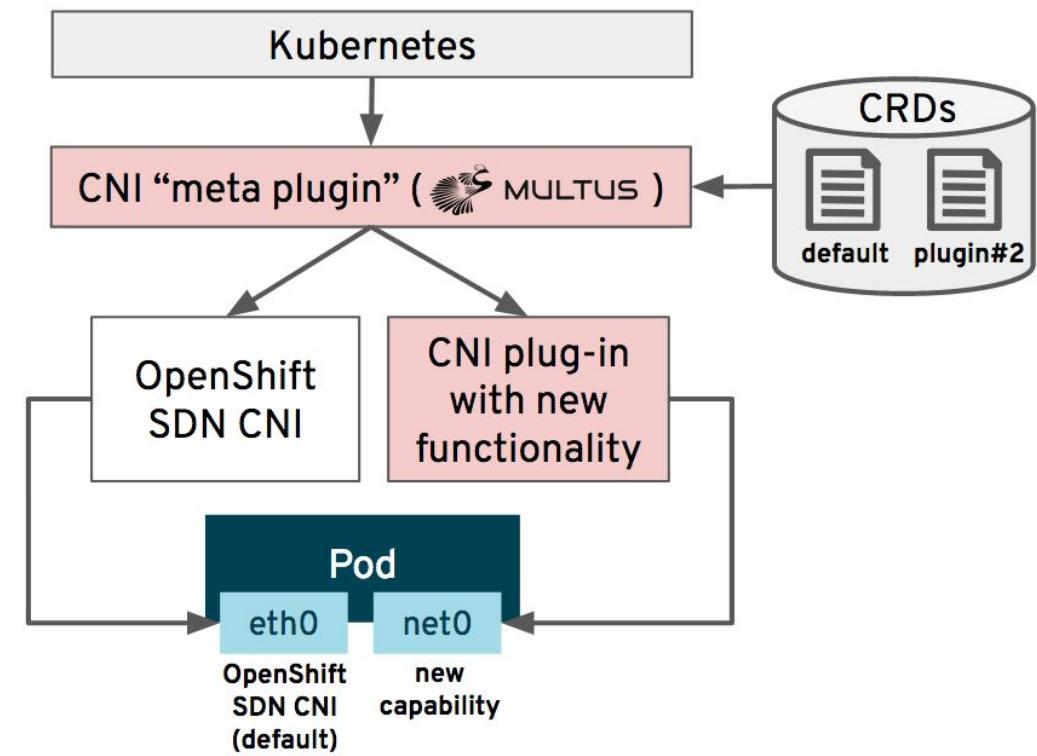


Migrating a VM with MTV

Networking

Virtual Machine Networking

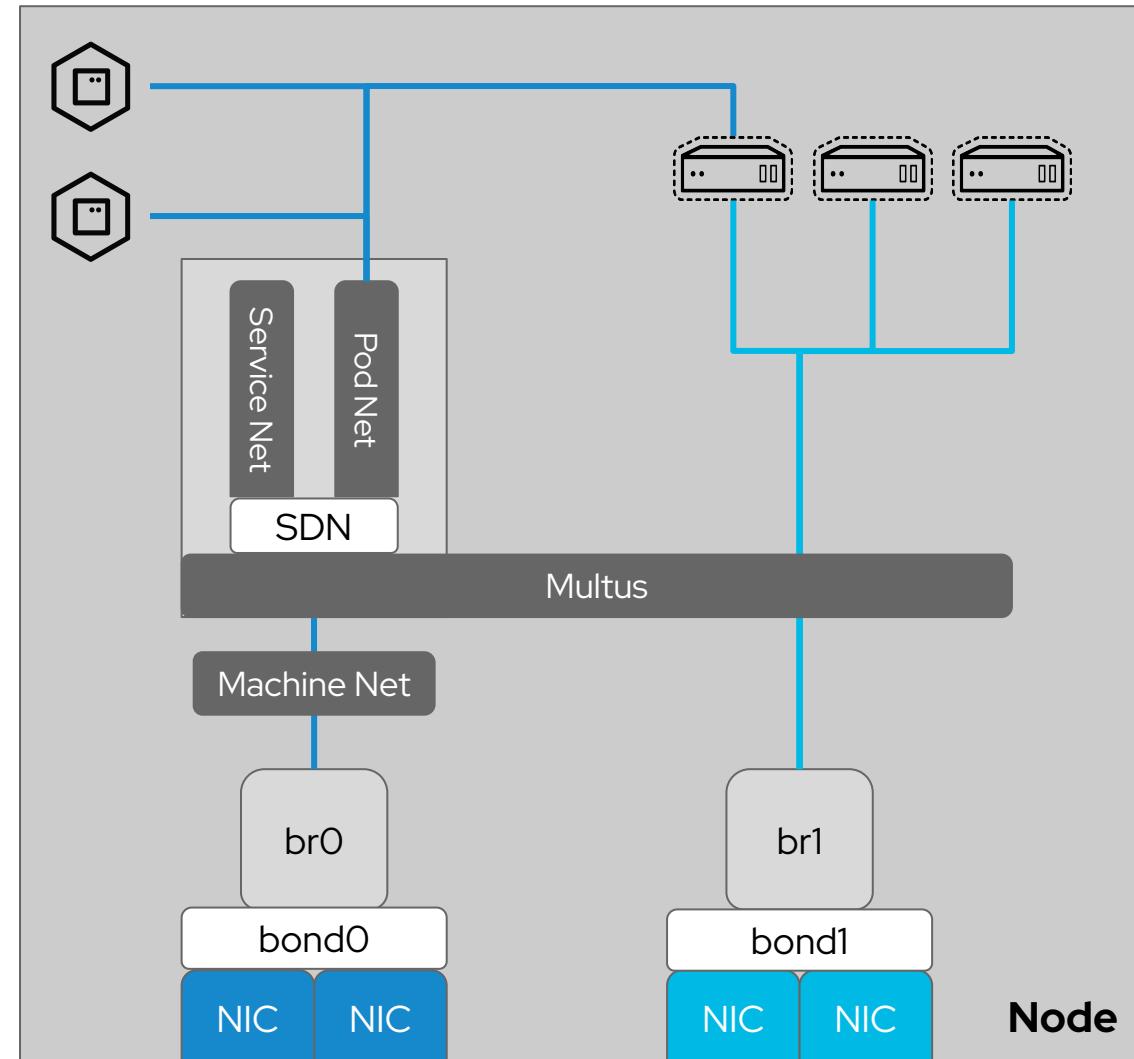
- Virtual machines optionally connect to the standard pod network
 - OpenShift SDN, OVNKubernetes
 - Partners, such as Calico, are also supported
- Additional network interfaces accessible via Multus:
 - Bridge, SR-IOV, OVN secondary networks
 - VLAN and other networks can be created at the host level using nmstate
- When using at least one interface on the default SDN, Service, Route, and Ingress configuration applies to VM pods the same as others



Example host network configuration

- Pod, service, and machine network are configured by OpenShift automatically
 - Use kernel parameters (dracut) for configuration at install – bond0 in the example to the right
- Use the NMstate Operator to configure additional host network interfaces
 - bond1 and br1 in the example to the right
- VMs and Pods connect to one or more networks simultaneously

The following slides show an example of how this setup is configured





A P E X Cloud Platform for Red Hat OpenShift

—
Theo Crithary

Principal Systems Engineer
Customer Centric Cloud & Containers (C4)
Asia Pacific & Japan



Thank you

Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make Red Hat a trusted adviser to the Fortune 500.



linkedin.com/showcase/red-hat-developer



youtube.com/RedHatDevelopers



facebook.com/RedHatDeveloper



twitter.com/rhdevelopers