



Building AI Workloads locally

Raul Leite

Principal Solution Architect

r



What's Podman?

What is Podman?

A seamless way to work with containers (& Kubernetes!)



Fast and light

Daemonless, using the fastest technologies for a snappy experience.



Secure

Rootless containers allow you to contain privileges without compromising functionality.



Open

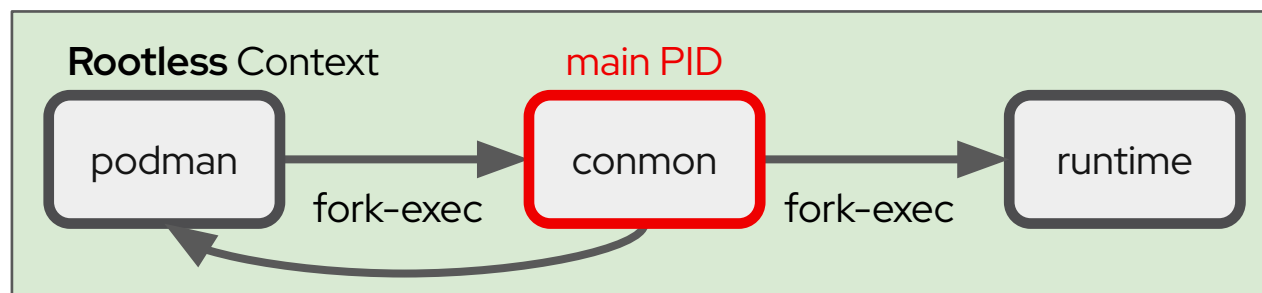
Podman is open source first and won't lock you in. Podman Desktop even supports Docker as an engine!



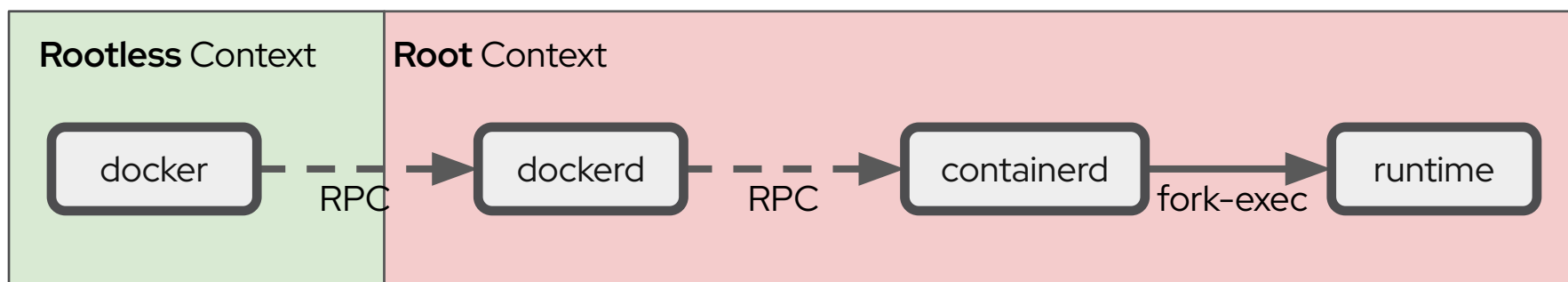
Compatible

Compatible with other OCI compliant container formats including Docker, as well as docker-compose files.

Container Engine Architectures



Compared to...



A refresh of the Podman 101's



Secure Containers

User namespace mapping, seccomp profiles, and SELinux support

No Daemon Process

Podman launches containers as child processes, fork/exec model

Kubernetes Support

Supports Kubernetes YAML & testing before cluster deployment

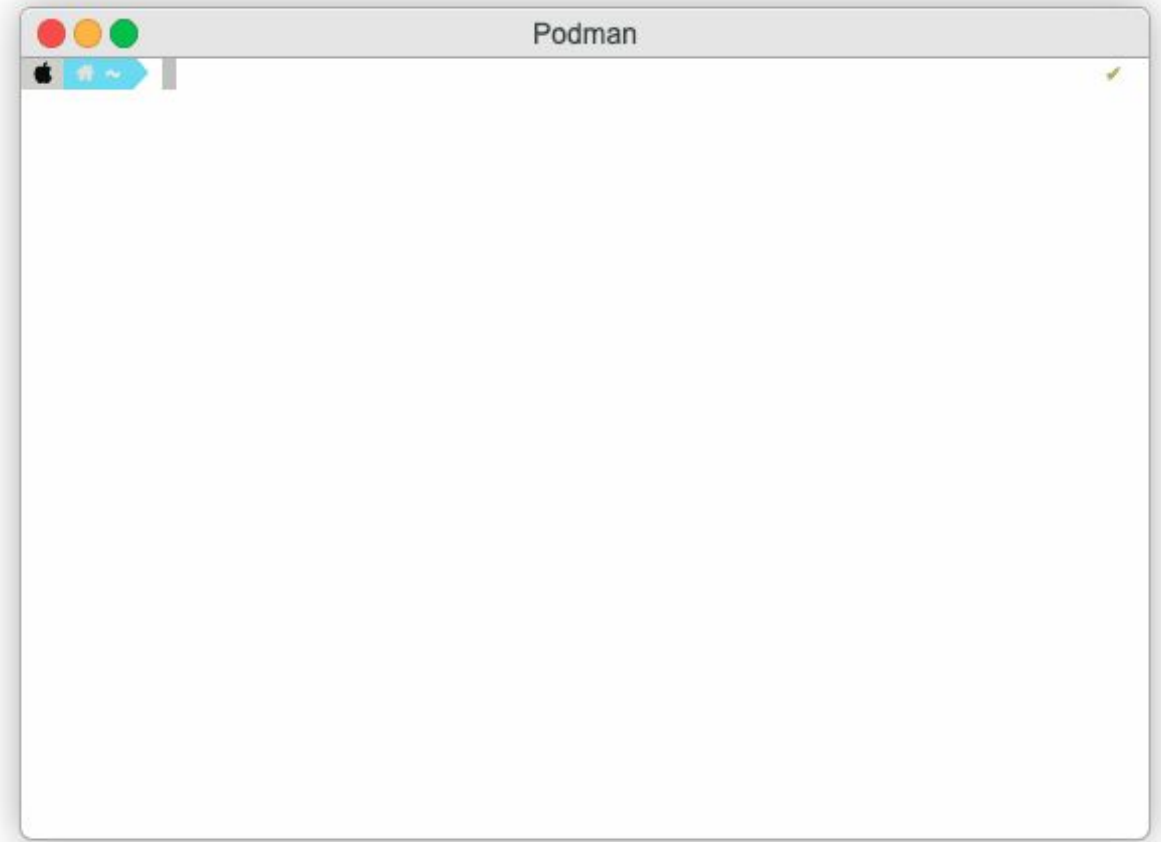
Platform Compatible

Runs on Linux, as well as Windows, Mac OS, FreeBSD

Podman in Action

It's **easy** to get started with Podman to manage containers and container images!

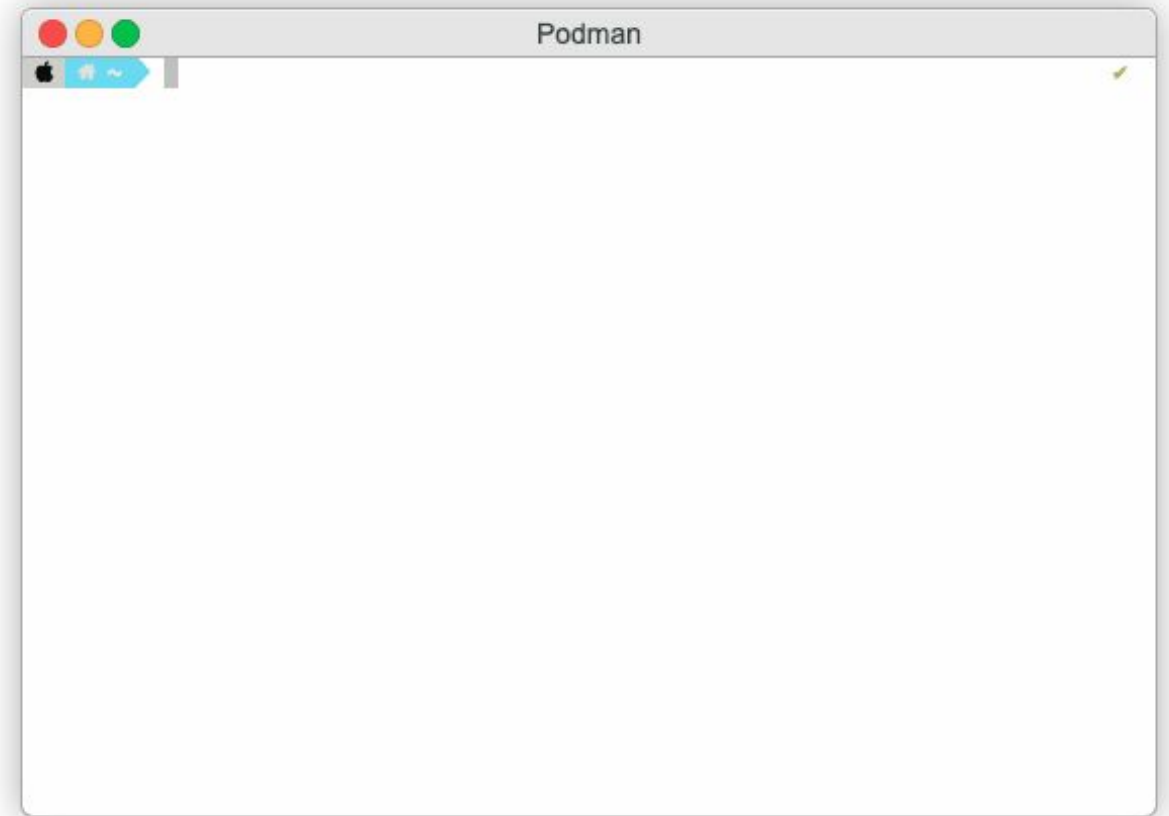
- ▶ Pulling, building, pushing images
- ▶ Running containers & debugging
- ▶ Working with Kubernetes
- ▶ & much more!



Podman in Action: Images

Container Images: Lightweight and isolated packages that encapsulate an application & it's dependencies.

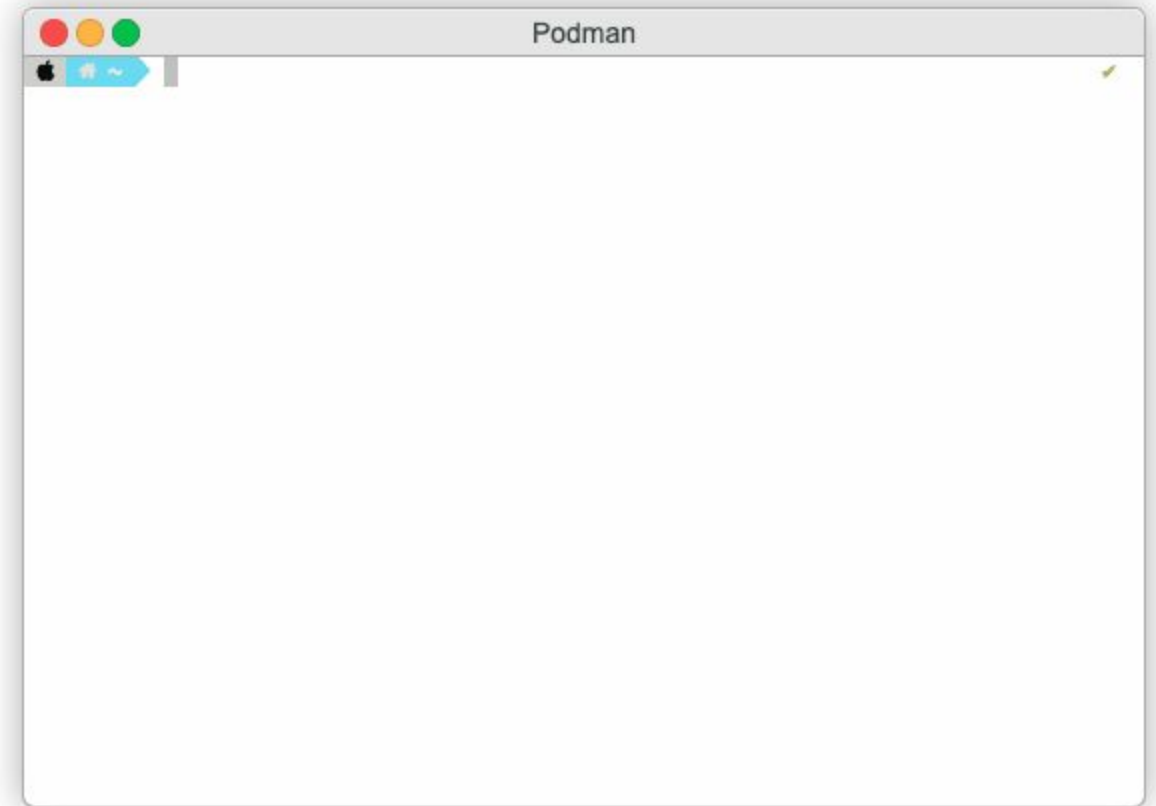
- ▶ `podman search <image>`
- ▶ `podman pull <image>`
- ▶ `podman images`
- ▶ `podman build .`
- ▶ `podman push <image>`



Podman in Action: Containers

Containers: Running instance of a container image, sharing host OS kernel and is ephemeral.

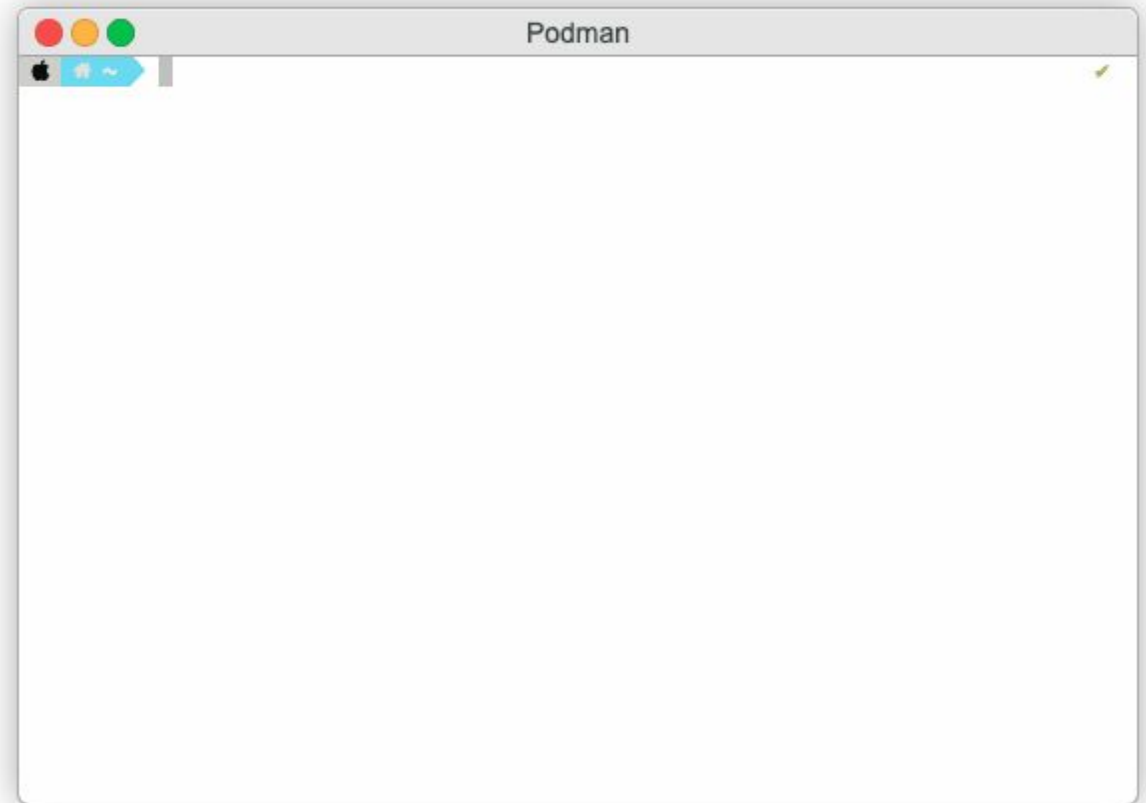
- ▶ `podman run <image>`
- ▶ `podman ps`
- ▶ `podman attach`
- ▶ `podman logs <container>`



Podman in Action: Pods

Pods: Group of containers that run together and share the same storage & network as a single unit.

- ▶ `podman pod create <name>`
- ▶ `podman generate kube`
- ▶ `podman run --pod`
- ▶ `podman play kube`



What's Podman Desktop?

Best open source developer
tool to work with **Containers**
and **Kubernetes!**



Free, Open and Extensible — By Default



Podman Desktop simplify workflows and experiences working with containers when targeting production on Kubernetes / OpenShift by:

- Bringing Kubernetes closer to the developer
- Minimizing the discrepancies between the desktop and target Kubernetes environments.



- Start Using

[Podman-desktop.io](https://podman-desktop.io)

[Podman.io](https://podman.io)

- Contribute, report issues:

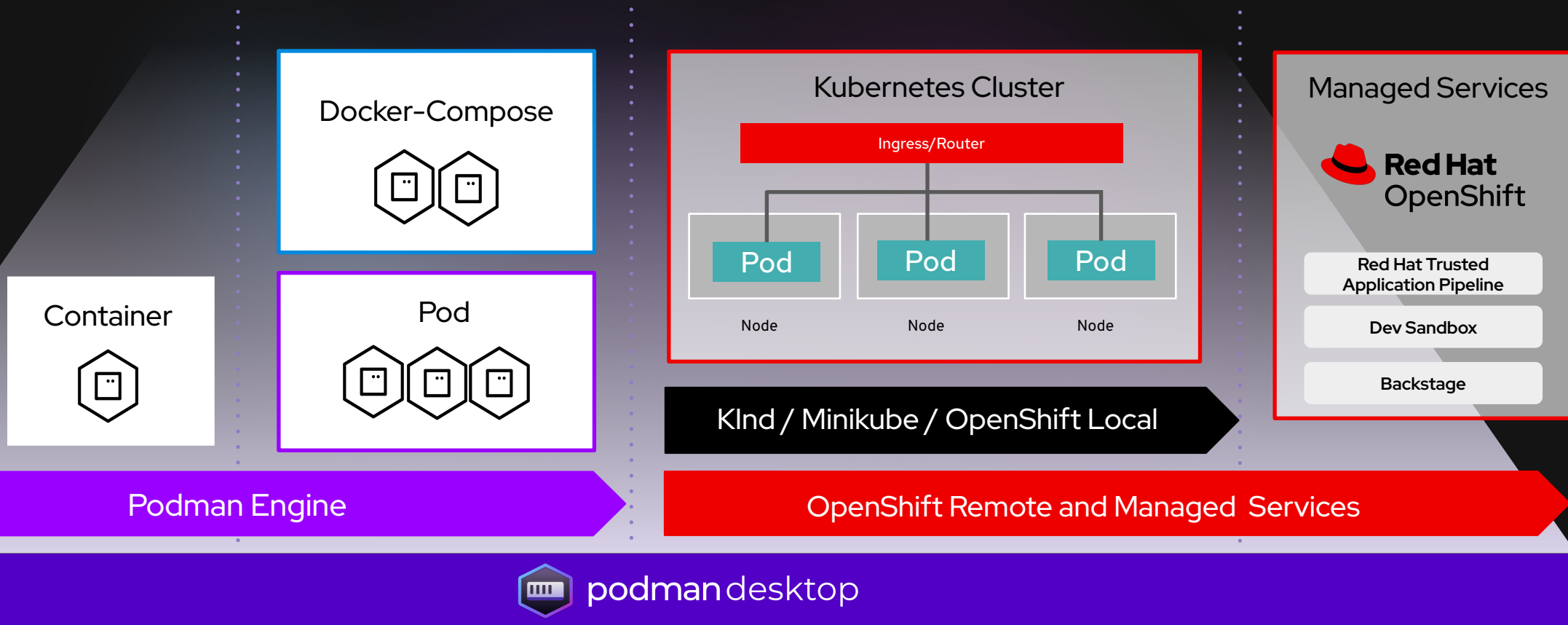
github.com/containers/podman-desktop



podman desktop

Simplistic onboarding.

From applications to containers, to pods, to Kubernetes.

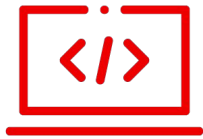


Podman AI Lab

*Your pathway to Gen AI, bridging
with containers and Kubernetes*

Why Podman AI Lab Matters!

& how it accelerates the AI application development workflow.



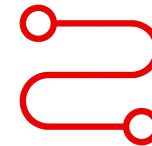
Run LLM's Locally

Simplify the environment for developing and debugging AI-enabled applications in your inner loop.



Cost Efficiency

Running AI models locally reduces cloud computing costs and vendor locking.



Gen AI Education

Learn how to infuse AI in your applications through sample use cases and best practices.

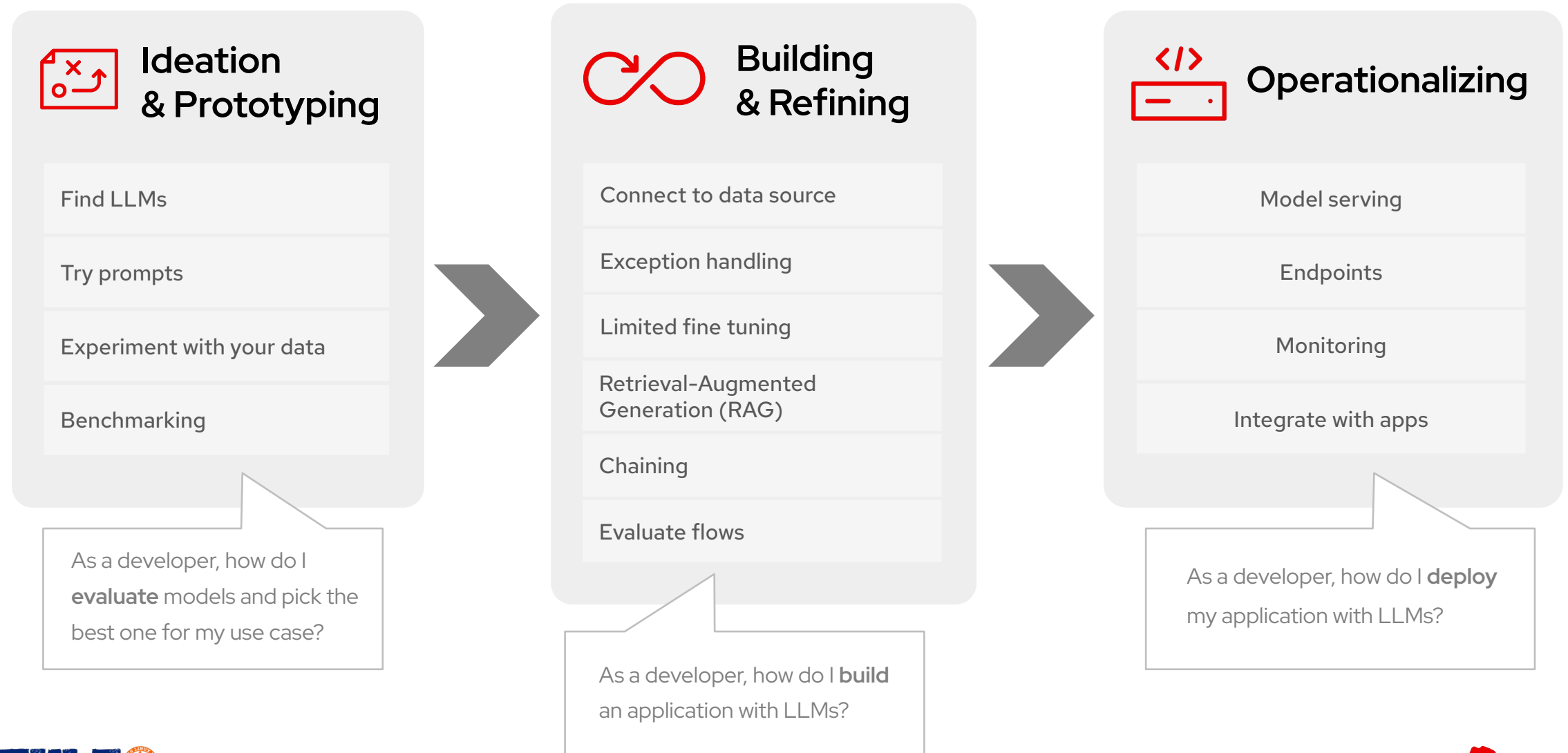


Data Ownership

Maintain ownership and control over your sensitive data from your local environment.

Adopting Generative AI for an **Application Developer**

CONFIDENTIAL designator



Why run a model locally?

CONFIDENTIAL designator

Take advantage of total AI customization and control



For Developers



For Organizations

Direct Access to Hardware

Convenience & Simplicity

Familiarity with the Development Environment and adherence of the developers to their "local developer experience" in particular for testing and debugging

Ease of Integration

Simplify the integration of the model with existing systems and applications that are already running locally.

Customization & Control

Easily train or fine-tune your own model, from the convenience of the developer's local machine.

Data Privacy and Security

Data is the fuel for AI, and a differentiator factor (quality, quantity, qualification) Keeping data on-premises ensures sensitive information doesn't leave the local environment → crucial for privacy-sensitive applications

Cost Control

While there is an initial investment in hardware and setup, running locally can potentially reduce ongoing costs of cloud computing services and alleviate the vendor-locking played by Amazon, MSFT, Google

Regulatory Compliance

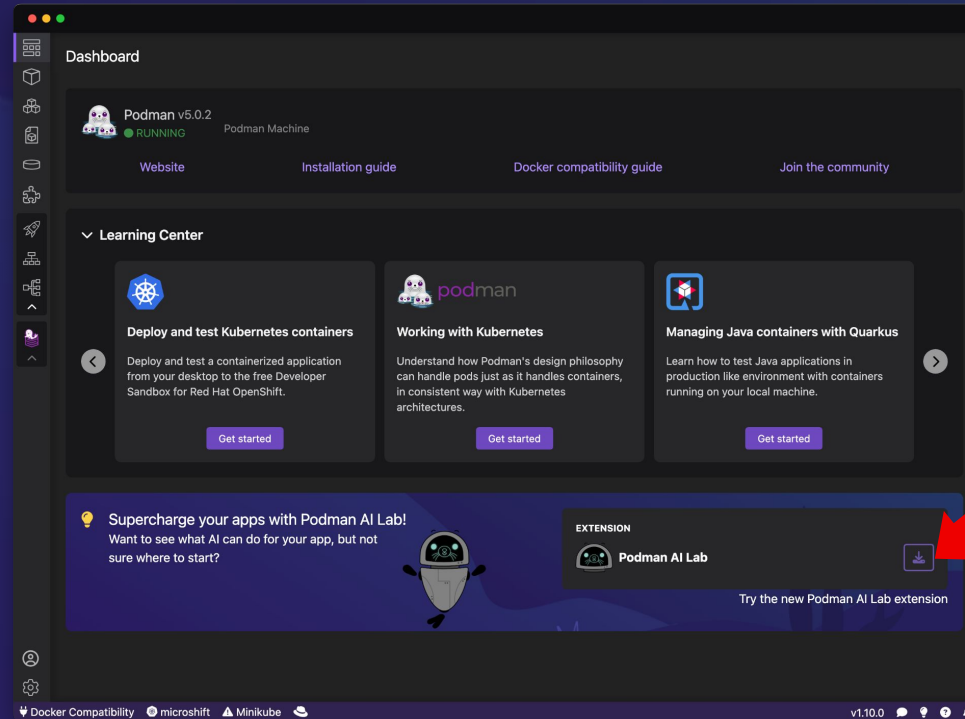
Some industries have strict regulations about where, how data is processed

It all starts with **enabling**
the developers to use
models



Podman AI

Lab Run LLMs locally and build AI applications



[github.com/containers/podman
-desktop-extension-ai-lab](https://github.com/containers/podman-desktop-extension-ai-lab)

Supported platforms:




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

 facebook.com/redhatinc

 youtube.com/user/RedHatVideos

 twitter.com/RedHat

