Data running “after” Red Hat OpenShift

Jan Cordtz,
Senior Cloud Solution Architect
Microsoft

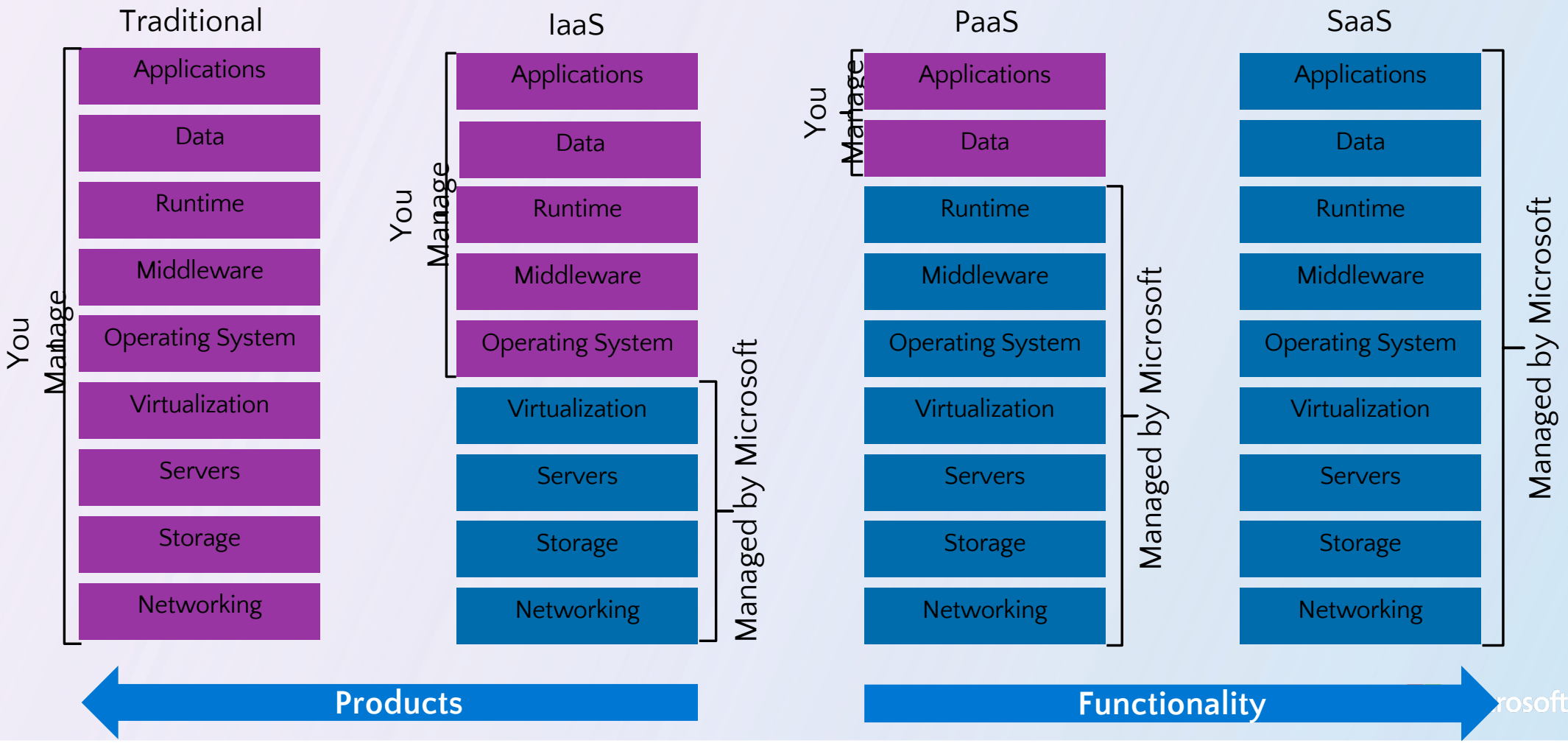
The Cloud

Red Hat

Azure

Microsoft

Cloud Models



Denmark East

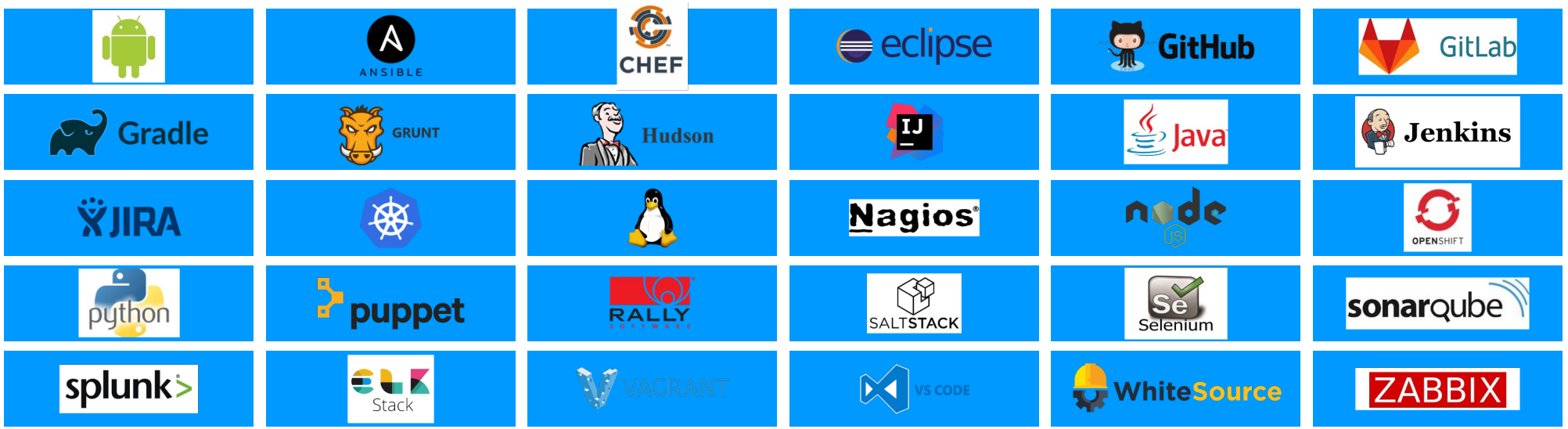
- One region with three underlying physical locations at Roskilde, Køge and Høje Tåstrup
  Three locations ensure Microsoft can deliver high SLA
  Under construction and going live in 2025

Why use Danish region?

- Hyper-scale cloud
- Security and compliance
- Local innovation and growth
- Green IT operations ESG
- Data at Danish location
- Very low latency

Red Hat

Azure runs what you need



Microsoft is one of the biggest OpenSource contributors

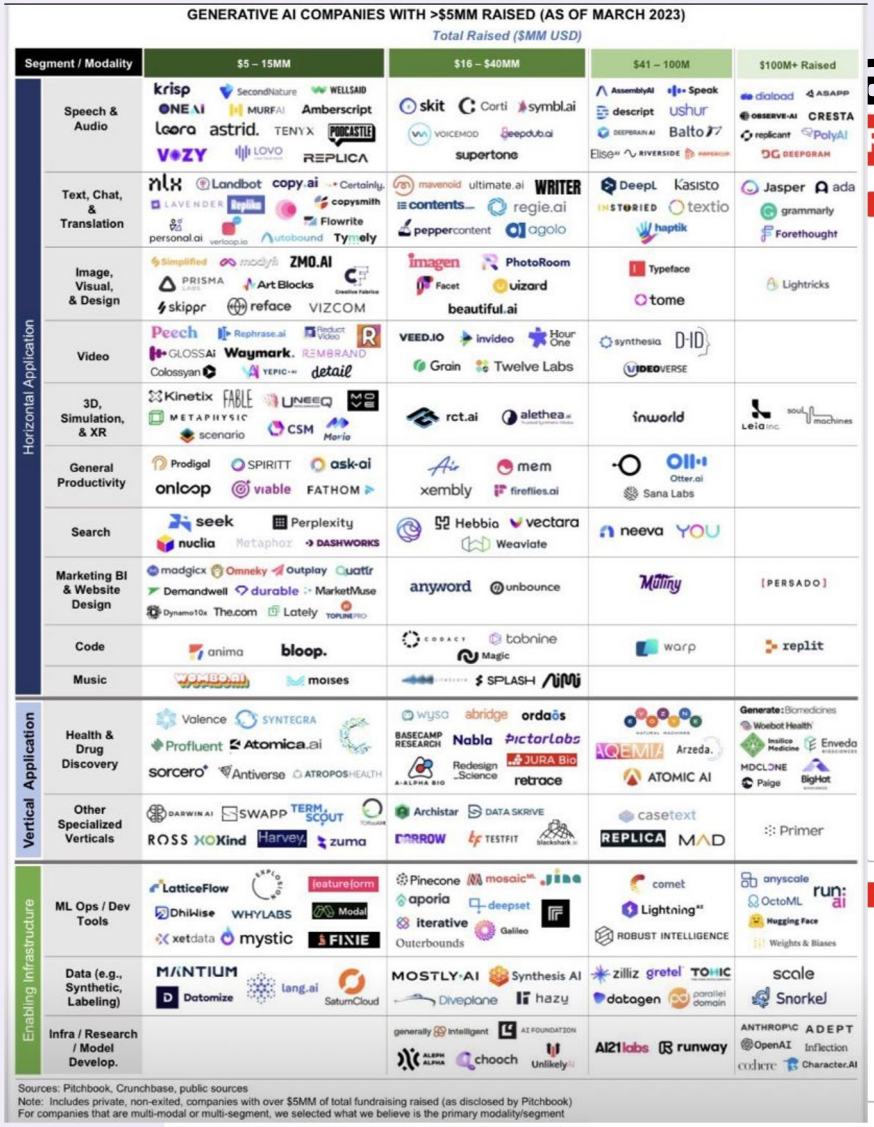
Microsoft

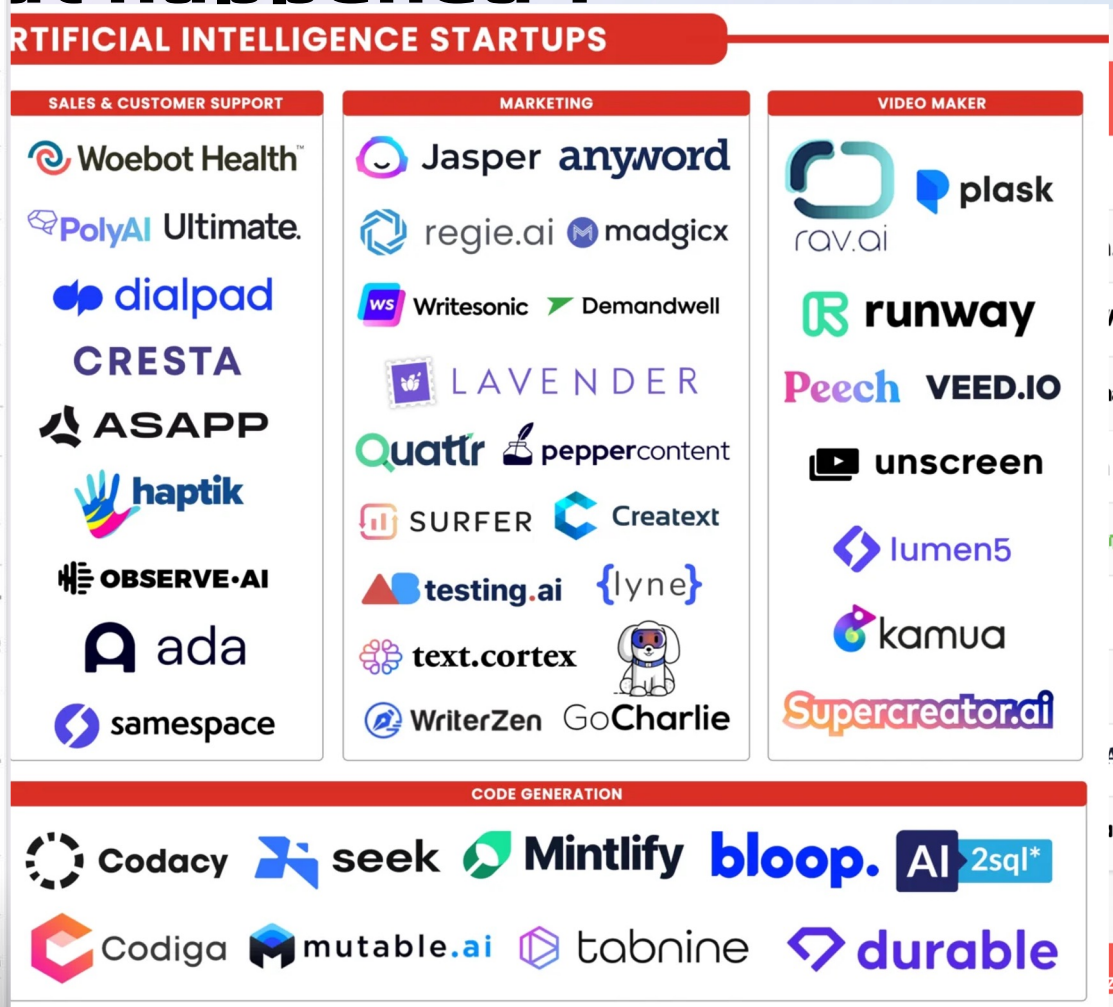
What is going on ?

GenAI is.....

Red Hat

at happened ?





Generative AI has been called the “next frontier in machine learning and AI”

- Until recently, machine learning was largely limited to predictive models, used to observe and classify patterns in content. Generative AI changed that.
  - Rather than simply *perceive* and *classify* a photo of a cat, machine learning is now able to *create* an image or text description of a cat on demand.
- Text applications are most advanced right now, but other modes, like code, speech, and images, are rapidly maturing.
- There’s a lot of tech industry excitement these days about generative tools like DALL-E and ChatGPT.

1. New York Times access notes

Note: DALL-E generated this image from a command for “cats playing chess.”

Image source: New York Times[1]

Red Hat

Best Friends

OpenAI

Ensure that artificial general intelligence (AGI) benefits humanity

Microsoft

Empower every person and organization on the planet to achieve more

Azure OpenAI Service

Language	Multi-Modal	Fine Tuning	Images	Transcription & Translation
GPT-4, 3.5-Turbo	*GPT-4o* GPT-4-Turbo with Vision	*GPT-4* GPT-3.5-Turbo	DALL·E 3	Whisper

On Your Data GA | *Azure AI Studio GA* | *Assistants API Preview* | *Provisioned Throughput*

Microsoft

Red Hat

Microsoft Copilot

Microsoft Security

Microsoft Windows

GitHub

LinkedIn

Xbox

Azure AI unlocks intelligence
across Microsoft products

and empowers every industry
to build custom AI apps

Chat with your data

Intelligent call centers

Your own copilots

Information discovery

Hyper-personalization

Microsoft

Is data being left behind ?

The Challenges

- Application Teams are applying modern DevOps/MLOps principles to deliver business value quicker
- Data teams are being left behind – DATA QUALITY !!!
- The modern data platform is just as sophisticated as the application platform
- They don't need to be different, the same DevOps/MLOps principles we've been using for the application platform must be used for the data platform – called DataOps
- Data teams must, and should, be 1-to-1 working with the application teams

Red Hat

Application TEAMS are MODERNIZING WITH KUBERNETES

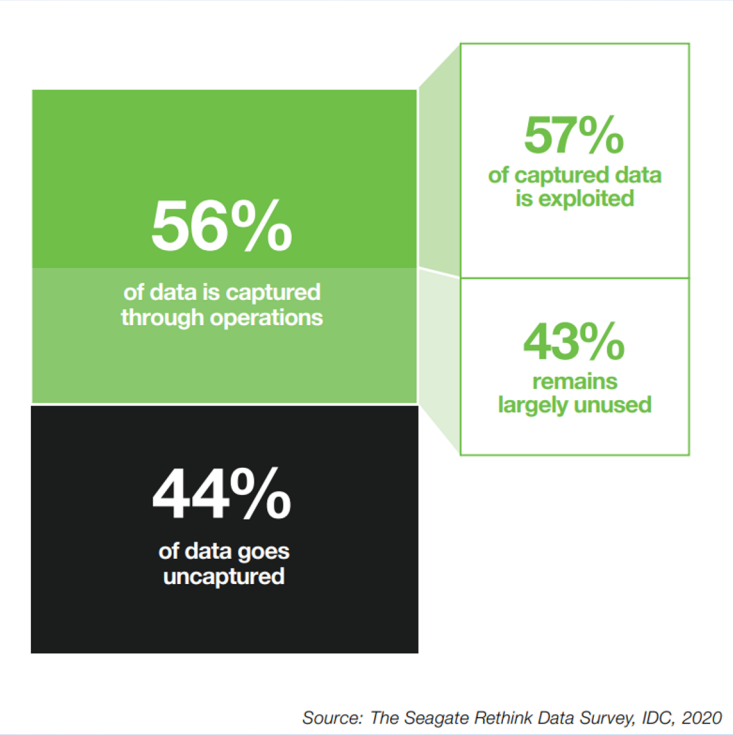
Application Developer

Microsoft

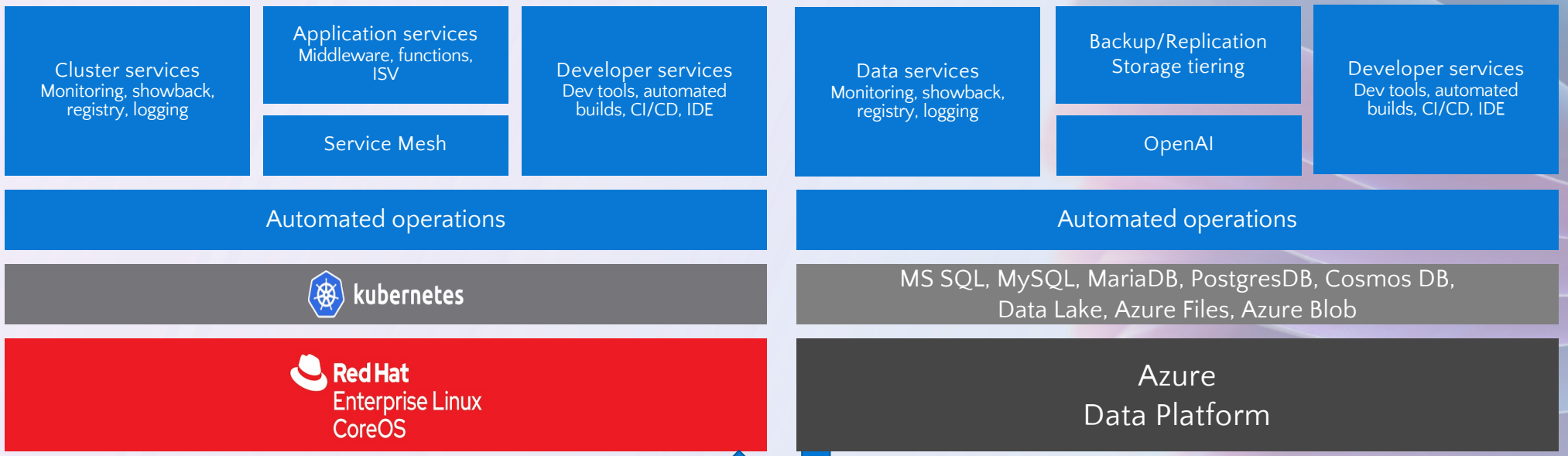
Applications and Data

- Applications are the main driver for capturing data
- Data needs to be used – ”to be data driven”
- New disciplines occurs like AI, ML and GenAI
- So, the application landscape will continue to change
- And Data Platforms needs to be modernized too

- Application and data development must go “hand in hand”



The ideal solution could be Azure Red Hat Openshift and a Data Platform



Let Red Hat and Microsoft ...

Manage all your clusters/databases

Secure your nodes

Monitor and operate your VMs

Manage environment patches

Backup/recovery

Defend your environment

Data : Capabilities

Capability	Description
Consistency	A consistent solution design to ensure simple operation and further development of solutions. Replacing resources must be flexible.
Modulated	Solutions in the data infrastructure are developed modularly with a well-defined interface, so it is easy to replace subcomponents.
Technology independency	The architecture must be independent of the technology being used. This means that no matter what technology is being used, the different processes, functionalities and layers are the same.
Scalability	Scalability (up/down) is part of the solution design from the beginning, so that implementation and operations are not affected by bottlenecks, downtime or a possible mandatory purchase of unforeseen licenses.
Agility	The focus is on MVP (minimum viable product) and ongoing feedback to previous steps in the data flow. It is OK to fail, and solution designs must be tested as soon as possible during the development process.
Security	Security is incorporated into the general architecture as well as in the concrete solution designs, both in terms of information security and privacy.
Data encapsulation	Data in the data infrastructure is accessed through an interface that enables control of who has access to what. The interface allows you to make changes to the data infrastructure without affecting external systems.
Reuse	Solutions are being developed for reuse. The architecture will consist of templates for solution designs that shorten time-to-market and ensure standardization.
Feedback	The architecture is continuously adapted and improved based on feedback from the different users of data.

Data : Layers

New systems

Legacy systems

Ingest

Characteristics

- Consume raw data
- High degree of integration points
- Very technical
- No transformation

Transform

Characteristics

- Technical useable data
- Getting data from ingest layer - only
- Very technical
- Place for technical transformations
- Adhere to company chosen standards

Publish

Characteristics

- Business useable data
- Data placed in whatever suitable technology for the end usage
- “Lives” as long as being used
- Can be recreated (by self-service)

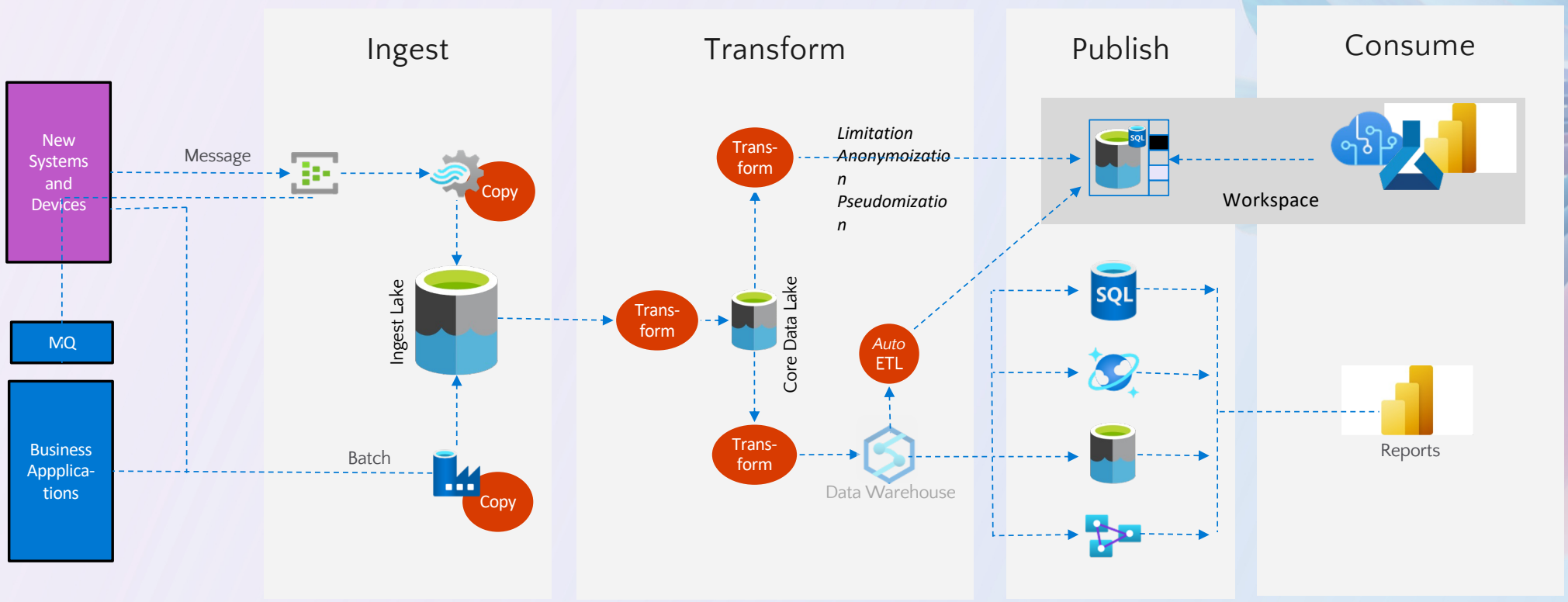
Consume

Characteristics

- End user tools
- Business Intelligence
- Reporting
- Machine Learning
- Artificial Intelligence

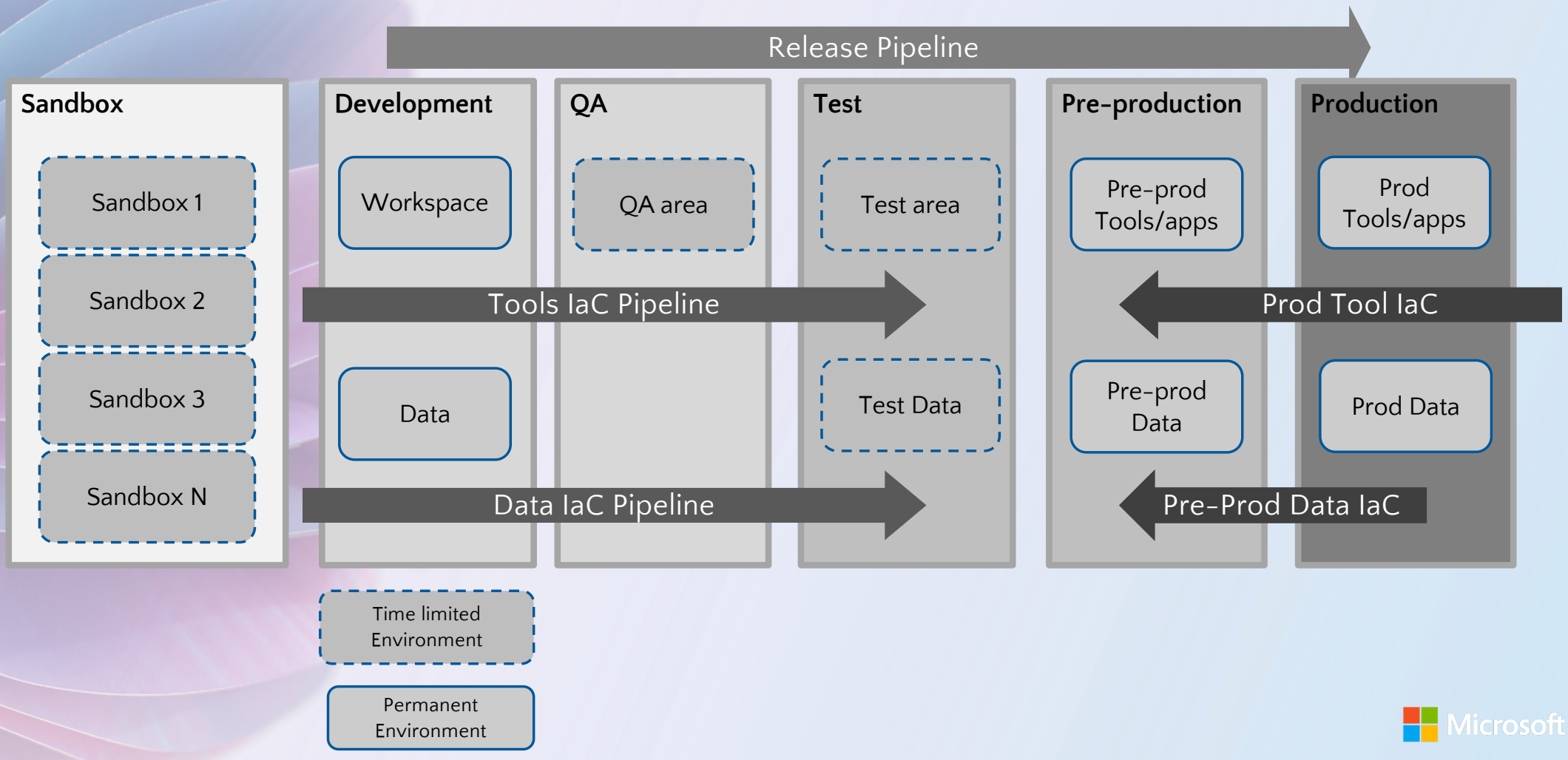
Supporting Data Services Data Catalog | Common Data Model | Data Discovery | Data Lineage | Data Governance | Master Data

Data : Process



Important abbreviations: IaC, IaD, Medalion

Data : Environments / Pipelines



Red Hat

Join us for the
Azure Red Hat OpenShift AI Workshop
November 14th at Microsoft in Kgs. Lyngby

red.ht/aro-ai-cph

Thank you

Jan.Cordtz@microsoft.com

https://github.com/jcordtz/a_data_platform

Microsoft