

20<sup>e</sup> édition

infoprodigital

# Forum **TERATEC 25**

21 & 22 MAI 2025

Parc Floral, Paris

**LE RDV des experts du supercomputing**  
**HPC – IA – Simulation – Quantique**

En partenariat avec

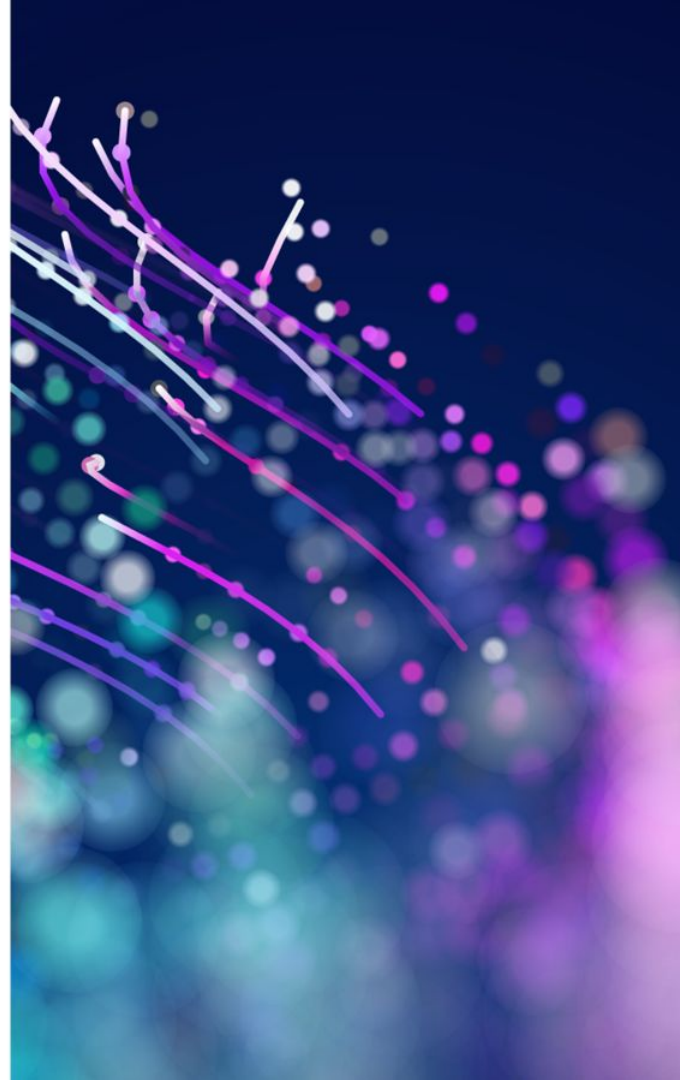
**L'USINE**NOUVELLE

**L'USINE**DIGITALE



# Enhancing Healthcare Support Systems with Agentic AI

Valentin Planes (MLOps Engineer)  
Alberto Garcia (HPC Architect)



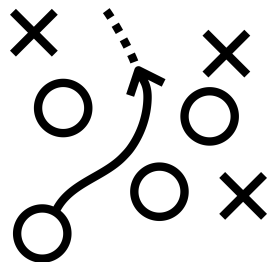
# A global HPC Services Company

Servicing HPC & AI customers world-wide

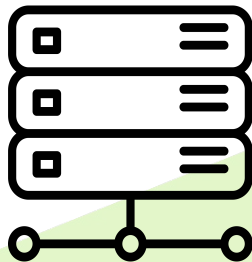
- 100% independent capital (Self-funded company)
- 140+ HPC & AI experts at your service
- 30+ years of experience in the EMEA HPC Market
- 200+ customers through all industry verticals (Aereospace, Automotive, Chemical, Energy, FSI, Life Sciences, Manufacturing, Oil & Gas)
- Services managed
  - 12k+ users/year
  - 150+ clusters/year
  - 200+ Training sessions/year
- Installation services
  - 5 clusters in the Top 500

## HPC & AI Services

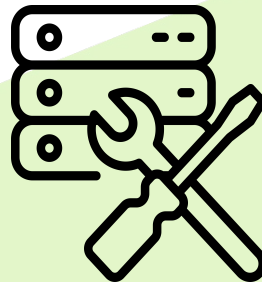
---



CONSULTING



INSTALLATION

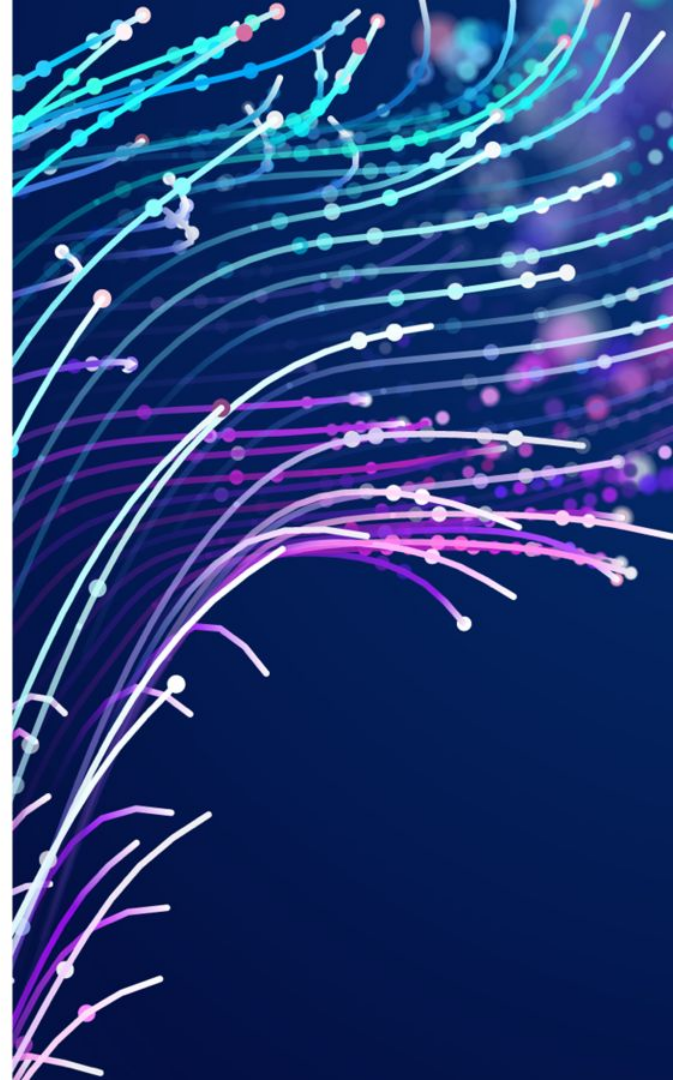


SUPPORT



R&D-as-a-Service

# **Fast-Track Information Retrieval and Streamlined Ticketing**



## Healthcare Context

---

Critical worldwide and 24/7 **Statistical and Development Platform** to:

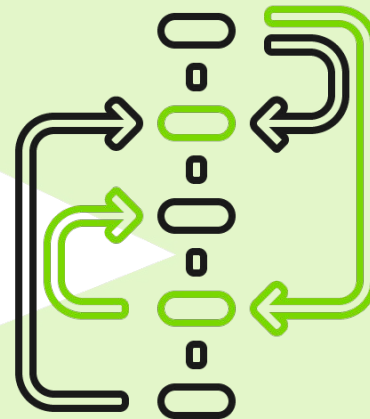
- Elaborate health authorities submission dossier.
- Validate clinical trials on drugs or vaccines.



## Key Points

---

- Complex and specialized workflows in pharma submissions.
- Users face documentation overload.
- Slow support response due to fragmented processes and responsibilities (ping pong effects).
- Need for fast and accurate information retrieval to be helpful.





## Agentic AI

---

Refers to systems with agency that can **perceive, reason, plan, and act** to achieve specific goals and it's built on top of GenAI Models.

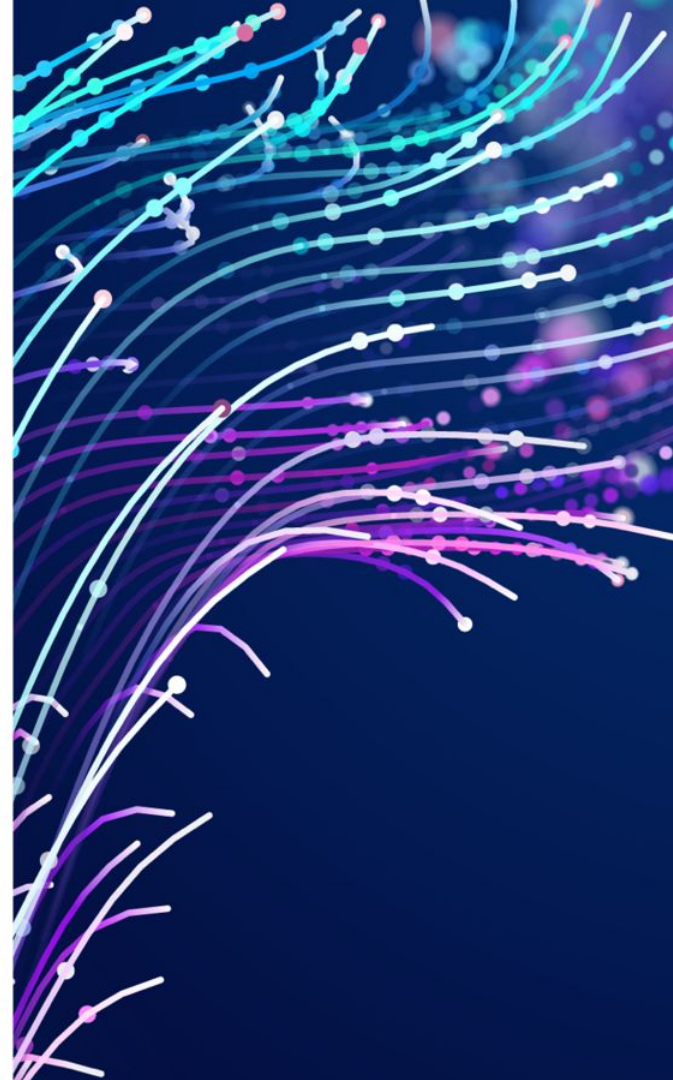
Capabilities:

- **Perception:** Understanding context and environment.
- **Reasoning:** Making informed decisions based on available data.
- **Planning and Action:** Executing tasks autonomously using tools and interacting with digital environments.

Enabling this level of intelligence requires longer generation cycles during inference, which significantly increases computational demands (x10 in tokens -> x150 in compute).



# Our Agentic AI Support Service



# Infrastructure

---

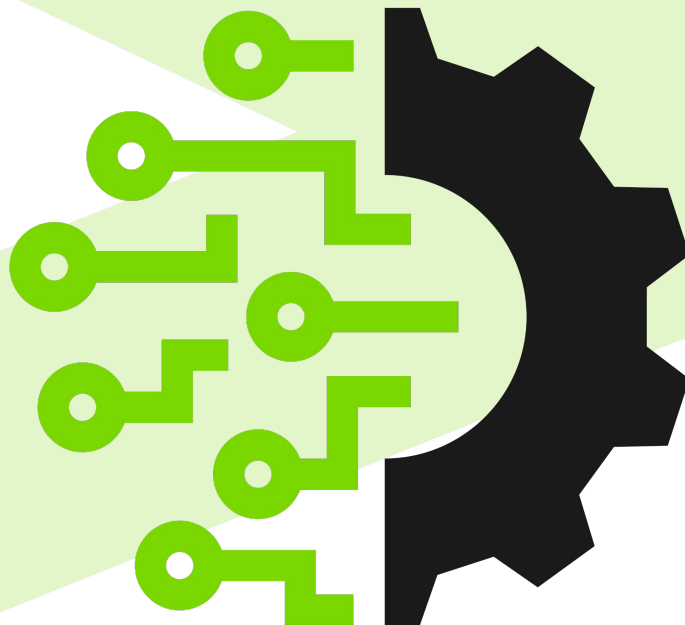
AWS Cloud

Multiple Region -> VPC peering

EKS cluster (kubernetes)

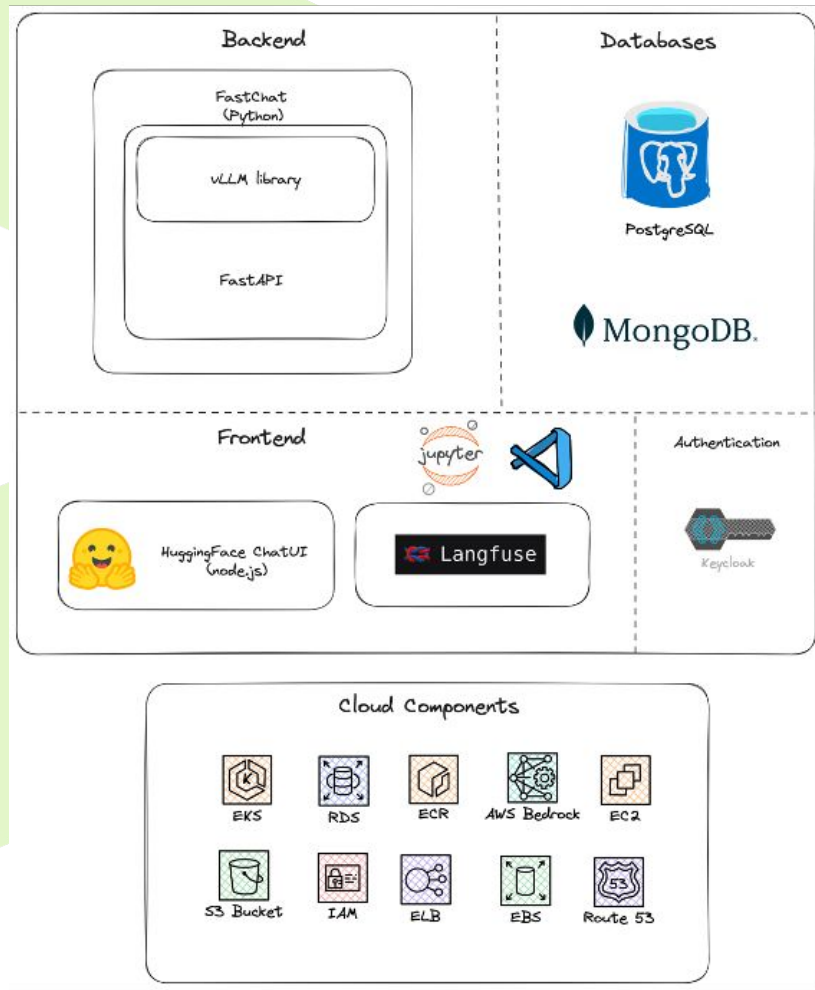
Instance type :

- P5.48XLarge ( 8x Nvidia H100 GPU )
- G5.48XLarge ( 8x Nvidia A10 GPU )



## Stack Overview

- Data Privacy preserved
  - End to end control
  - Based on Open Source components
  - Easy to customize
  - Cloud or On Prem compatible
  - Granular access control
- 
- Base LLM : Llama3 70B
  - Base embedding model : gte v1.5



## Our Agents

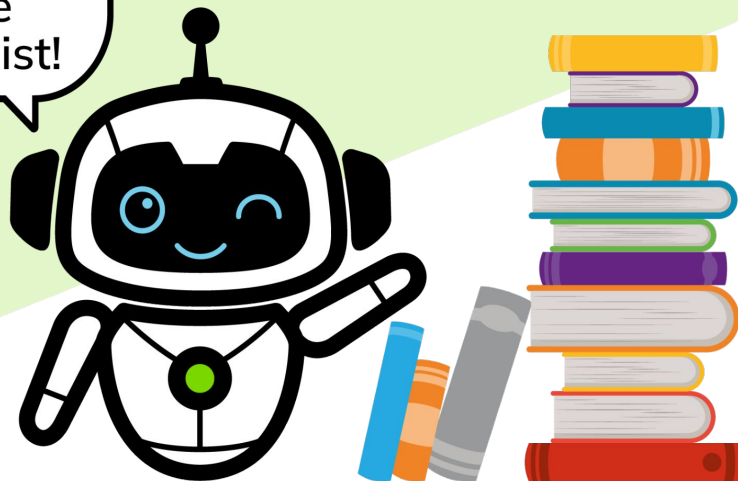
---

Agentic RAG (Retrieval Augmented Generation)

Can retrieve info from multiple sources of documentation:

- RAW Files
- Confluence Pages

Hi,  
I'm **MAX**, the  
Documentalist!

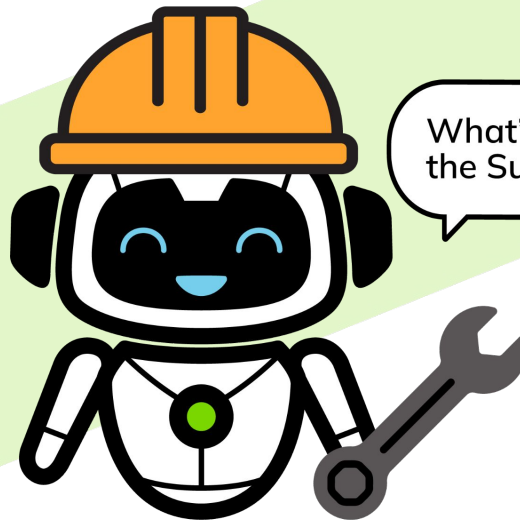


## Our Agents

---

### SAM Agent

- Can interact with ticketing system as ServiceNow!
- Help you to elaborate your ticket.



What's up? I'm **SAM**,  
the Support Intake Specialist!

## Agents in Action

I've an issue with my compute context can you help me?



Hello,

I'm excited to help with your request! I'm Sam, your digital assistant. My role is to help ensure your support ticket is clear and complete, so that the support team can resolve your inquiry as quickly and efficiently as possible.

Please provide additional information before I can forward your inquiry to the support team:

- Error raised: ?
- Was the SSO used for connection: ?
- URL of the service: ?

Thank you for your cooperation.  
Best regards.

error raised: "compute context failed with error 25552",  
sso used: i don't know, url: sasviya.com



Hey there, Sam checking back in with you!

Your ticket has been successfully verified and will be forwarded to our support team.

We appreciate your patience, and our team will work to provide you with a response as quickly as possible. In the meantime, I have provided an initial recommendation from the documentation:

1. Try restarting your SAS Viya Studio session.
2. Check if you have the latest version of SAS Viya Studio installed.
3. Verify that your compute context is properly configured.

Sources: [SAS Viya RGOS page](#), [SAS Viya 1.7.0 Webinar](#)

Stay tuned while our team is processing your request. If this answer was helpful, kindly let us know by marking the ticket as resolved. If not, please feel free to provide feedback

## Impact and Success Metrics

Some KPIs about SAM and MAX combined :

- Ticket resolution time ↓ by 50%
- Manual workload ↓
- 55% of tickets processed considered as ready

Capacity to retrieve information (MAX) :

- Relevant at 87% on our test set





## Conclusion

---

Agentic AI is reshaping support in healthcare, not by replacing humans, but by enhancing their ability to navigate and interact with complex systems.

- It acts as a new kind of collaborator, boosting team capacity, especially in terms of response time.
- Exceptionally effective at retrieving accurate information from vast knowledge databases.



This confirms the growing demand for computational power, so get ready!



valentin.planes@doitnowgroup.com

HPC & MLOps Engineer

<https://www.linkedin.com/in/vplanes/>



alberto.garcia@doitnowgroup.com

HPC Architect

<https://www.linkedin.com/in/alberto-garcia-fernandez-bb94449a/>

**Thank you  
for you  
attention!**