

Workshop #1 | 14:20 - 15:30

The NexusForum.eu Roadmap

#### **PROVIDE YOUR FEEDBACK!**

Thomas Ohlson Timoudas — Researcher, RISE

Massimo Neri — CTO, Martel Innovate

NexusForum Summit 2024, September 19, Brussels



## Agenda for Workshop #1



# 3 Research & Innovation Agenda Topics

- Cloud-Edge for Al
- AI for Cloud-Edge
- Telco Cloud-Edge

## **Discussion structure**

- Open Poll (Multiple-answer)
- Discussion of preliminary feedback
- Consultation and validation



**R&I Agenda** 



30/09/2024

#### Poll to kick off the discussion



#### 1. Cloud-Edge for AI

- Poll Question: "What do you consider the most critical requirement for developing European alternatives to hyperscaler AI ecosystems?"
  - a) Open-source development frameworks
  - b) Vendor-neutral AI platforms
  - c) Investment in RISC-V architecture
  - d) Data sharing and federation mechanisms
  - e) Regulatory support for EU-based AI services
  - f) Easier access to computing resources at scale in EU
  - g) Other (please specify in Q&A tab)





### Requirements for the EU Computing Continuum

- "The computing continuum must operate in a **sustainable** way: consume as little energy as possible and have a low carbon footprint." NGO/NPO respondent
- "Ease of **customisation** for deployment in specific application areas, privacy protection, ease of updates and adjustments." Startup/SME respondent
- "They need to deal with a huge amount of unstructured data to train **LLMs** Foundation models while at the same time **comply** with the Data Act." NGO/NPO respondent
- "... Compute must be where the data is rather than bringing the data to where the compute is ..." Large Enterprise

Sustainability, customization, and compliance with EU regulations are key requirements for the European Computing Continuum





#### **Development of EU AI Ecosystems**

- "...network AI Factories across Europe to create a unique collaborative AI framework" NGO/NPO respondent
- "Open source software and open source hardware" Startup/SME respondent
- "... the EU could bet on forming a big consortium composed of practitioners and researchers to 1) define a conceptual architecture to support such large ecosystems; and 2) propose an initial implementation of these components following open-source principles." University respondent
- "Open-source ... will not solve the current challenges [open standards and APIs are the way to go]. We need to push for alternative, competitive solutions, independent from the delivery model ..." Large Enterprise respondent

There's a strong emphasis on openness, open-source solutions and collaborative frameworks for developing EU AI ecosystems





#### **RISC-V and Open Hardware**

- "It is time that RISC-V starts to match Intel performance that challenges European cloud providers to install compute and storage resources based on RISC-V." NGO/NPO respondent
- "RISC-V is very suitable for for embedded/edge, but how about scaling it up to server-grade implementations" Startup/SME respondent
- "Concrete and complete visions/plans for RISC-V architecture in not only feasible technology design is explicated, but also the societal embedding: feasible business cases, feasible legislation" University respondent

While there's interest in RISC-V, respondents recognize the need for significant performance improvements and custom solutions for AI workloads



### Poll to kick off the discussion



#### 2. Al for Cloud-Edge

- Question: "What is the biggest challenge in creating a European federated computing continuum?"
  - a) Interoperability between different providers
  - b) Data sovereignty and security concerns
  - c) Latency and bandwidth limitations
  - d) Lack of standardization
  - e) Regulatory and compliance issues
  - f) Economic viability and business models
  - g) Other (please specify in Q&A tab)





#### Al for Optimizing the Computing Continuum

- "Al can help to predict where large compute is needed and aim for a redirection of workload and possibly perform advance ahead-of time scheduling of tasks." -NGO/NPO respondent
- "Compute optimization is a commodity in Hyperscaler data center for years. We just need to commoditize it for the rest of the world." Large Enterprise respondent
- "It is limited by definition. Not sure about the cost/benefit of using AI for management." – University respondent

There's different views on AI's potential to optimize resource management, but with an emphasis on maintaining human oversight





#### **Addressing Latency and Bandwidth Challenges**

- "Combine the computation power of the smart devices with the power of cloud technologies. Bring trained AI closer to the device to be applied with the data gathered and generate customized actions while keeping the Data sovereignty." NGO/NPO respondent
- "I don't think latency and bandwidth are our biggest challenge today, even (if) we talk a lot about it." Large Enterprise respondent
- "... invest in edge AI hardware." University respondent

There are differing views on the urgency of latency and bandwidth challenges, with some emphasizing edge computing solutions



### Poll to kick-off the discussion



#### 3. Telco Cloud-Edge

- Question: "Which new business model do you think has the most potential for private 5G/6G networks in industrial settings?"
  - a) Network-as-a-Service (NaaS) for factories
  - b) Edge computing marketplaces
  - c) Industry-specific AI/ML services
  - d) IoT data management and analytics platforms
  - e) Collaborative robotics and automation services
  - f) Augmented reality for industrial operations
  - g) Other (please specify in Q&A tab)





#### **Telco Cloud and Public Cloud Integration**

- "Development of the standard OPEN APIs that will allow CSP to access Telco cloud services" - NGO/NPO respondent
- "Telco companies has always very reliable in terms of control and sovereignty as they are dealing with critical infrastructures. The public and edge services should guarantee at least the sovereignty level demanded as in EUCS+ or Gaia-x Level 3." NGO/NPO respondent

Standardization and maintaining high levels of sovereignty are seen as key challenges in integrating telco and public cloud environments





#### **Federating Resources and Interoperability**

- "The EU providers need to see a real business benefit by federating their resources with other providers. So far, they see that is technically feasible but not provide added value to their current offer." NGO/NPO respondent
- "Interoperability among resources belonging to different entities. Need of defining clear and stable interfaces" Large Enterprise respondent
- "The EU providers need to see a real business benefit by federating their resources with other providers. So far, they see that is technically feasible but not provide added value to their current offer ..." NGO/NPO respondent

Clear business benefits and well-defined interfaces are seen as crucial for successful federation of resources.



### **Open questions**



#### **Cloud-Edge for Al**

- 1. What are the main requirements on the computing continuum to support the evolution of AI in EU? What are the requirements if we want to develop foundation models in Europe?
- 2. Hyperscalers have AI service ecosystems, mainly based on their own stacks. What will be needed to establish EU alternatives built on open source and vendor-neutrality?
- 3. What are the next steps in terms of research and innovation for creating an EU ecosystem built on the open RISC-V architecture?



### **Open Questions**



#### Al for Cloud-Edge

- 4. What are the main challenges and approaches to explore for creating a European computing continuum, in particular federating resources from multiple providers?
- 5. How can we address the challenges of latency and bandwidth in the computing continuum to support real-time AI applications?
- 6. How can Al optimize the use and management of the computing continuum, and what measures should be taken to balance these benefits with the need for human oversight and control?

### **Open Questions**



#### Telco Cloud-Edge

- 7. What are the main challenges in ensuring seamless interoperability between telco cloud and public cloud environments?
- 8. How can telcos maintain control and sovereignty over their networks while leveraging public edge and cloud capabilities?
- 9. What new business models might emerge from the deployment of private 5G/6G networks in industrial settings? What is the role of edge nodes in enabling them?





# THANK YOU FOR YOUR ATTENTION















