

L'**art** de maîtriser l'imprévisible : façonner une **solution** IA Gen

Raphaël Anjou - @Orange Business
Azur Tech Winter – 26/11/2024

**Qu'allez-vous obtenir
de cette présentation ?**

Qui suis-je ?

**Pourquoi avons-nous besoin
de maîtriser l'imprévisible ?**

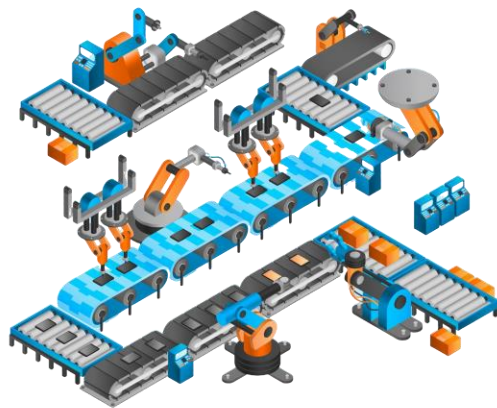
Un LLM est une boîte noire.

**Le plus de responsabilités on donne,
le moins de contrôle on a.**

Comment les humains **partagent-ils** les responsabilités ?



Outil



Usine

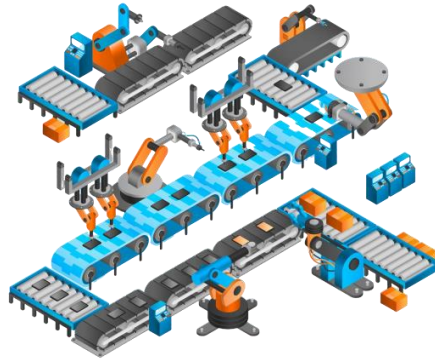


Équipe

Dans l'IA Gen ça ressemble à quoi ?



**GPT
wrappers**

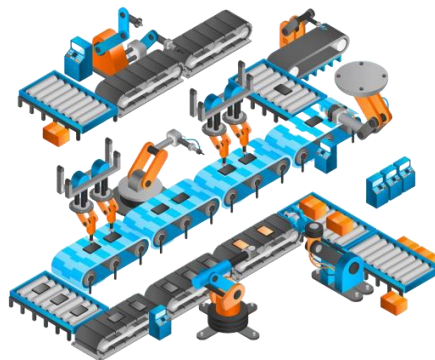


**LLM
chaining**



**Système
multi-agents**

Où mettre le curseur ?



1 job
1 responsabilité
+ de contrôle

La question qui fâche...

**Est-ce que
j'ai vraiment besoin d'IA Gen ?**

Les LLMs sont juste de **nouveaux outils**.

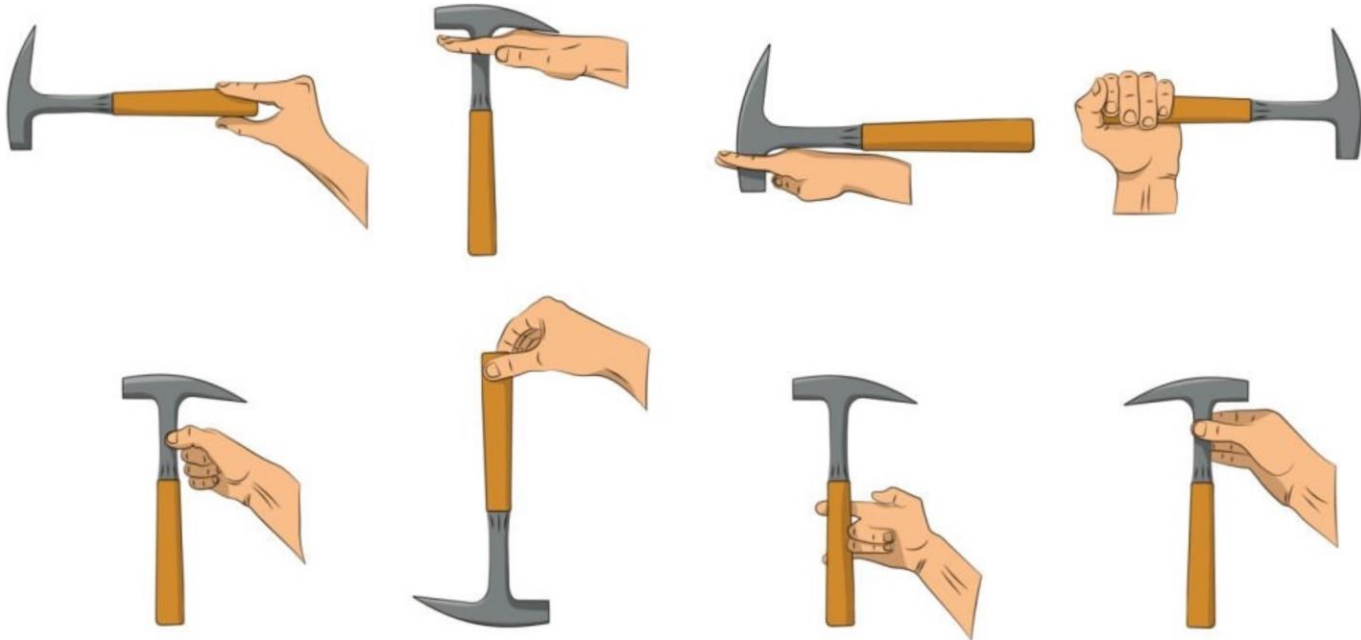
**Lire, regarder,
écrire**



**« Comprendre »
« Réfléchir »**

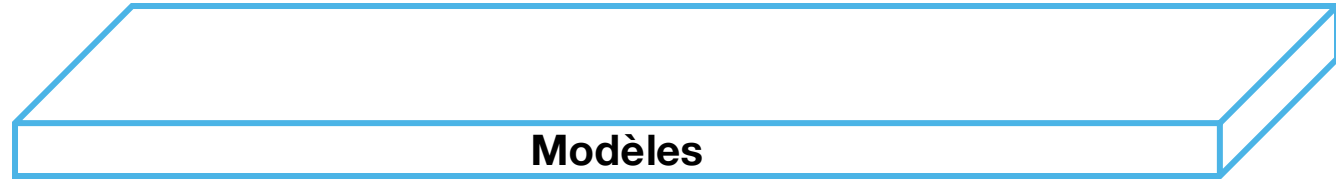
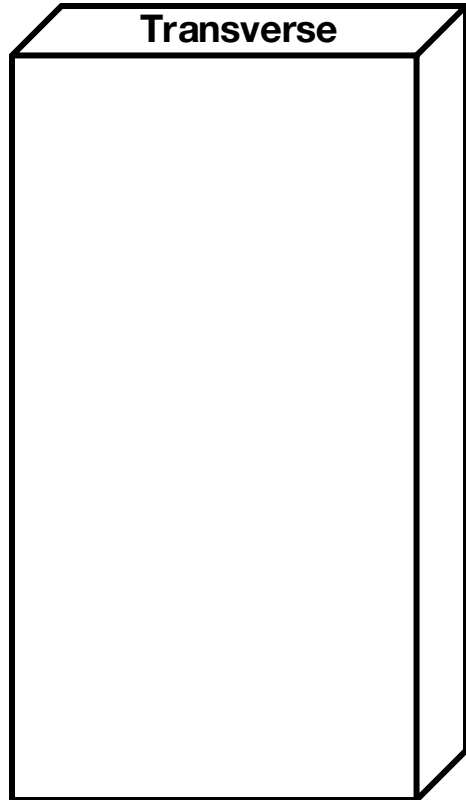


Un outil doit être **bien utilisé**.
Libre à vous d'en faire autre chose.



Un LLM est **une partie de la solution,
ce n'est **pas LA** solution.**

Les briques de notre **solution**



Infrastructure

Je ne veux
pas savoir.

groq



OpenAI

ANTHROPIC

Je veux le GPU
d'un autre.
(GPUaaS)

aws

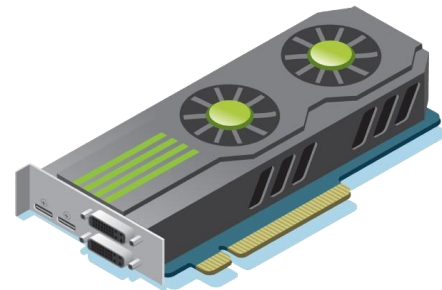


Google Cloud



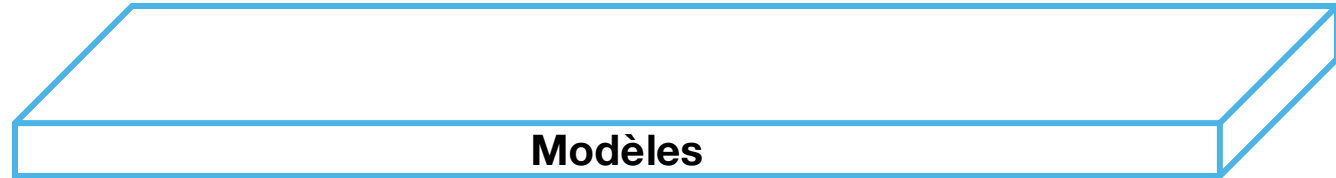
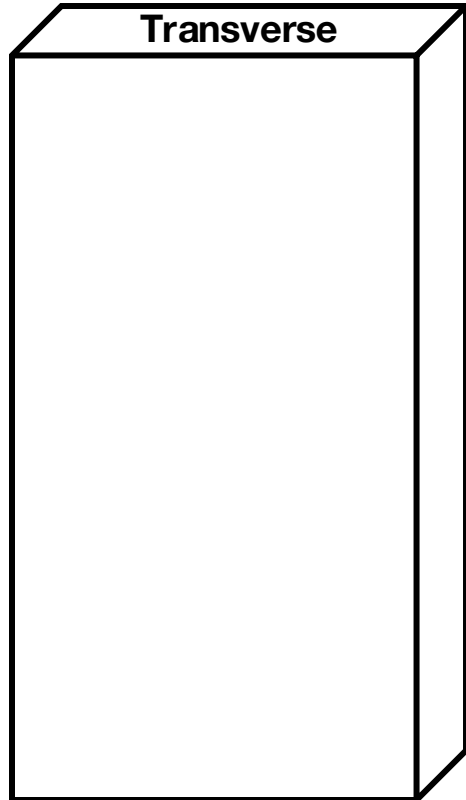
Business

Je veux
mon GPU.



Cloud vs On-Premise ?
Besoin d'être souverain ?

Les briques de notre **solution**



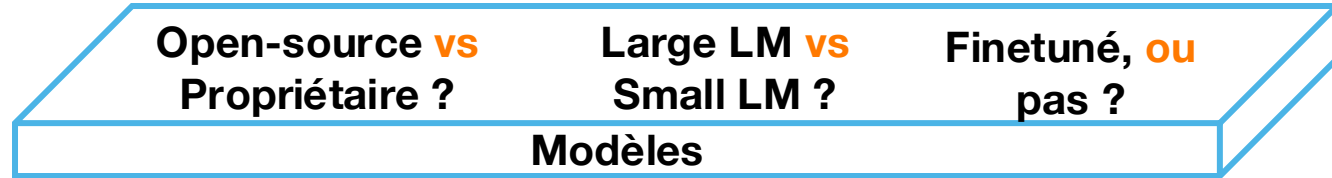
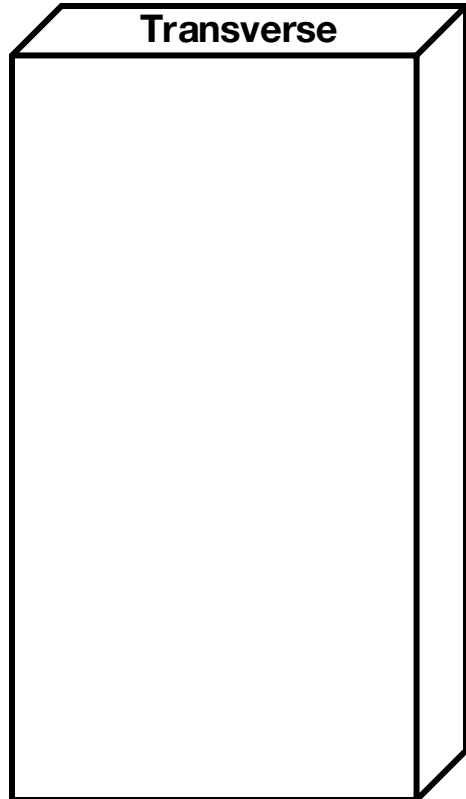
Modèles

Open-source vs Propriétaire ?

Large LM vs Small LM ?

Finetuné, ou pas ?

Les briques de notre **solution**





Intelligence

LLM chaining

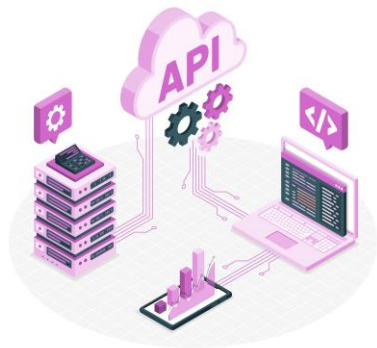
Prompt engineering

Systeme multi-agents

Données

Retrieval Augmented Generation

API



Database



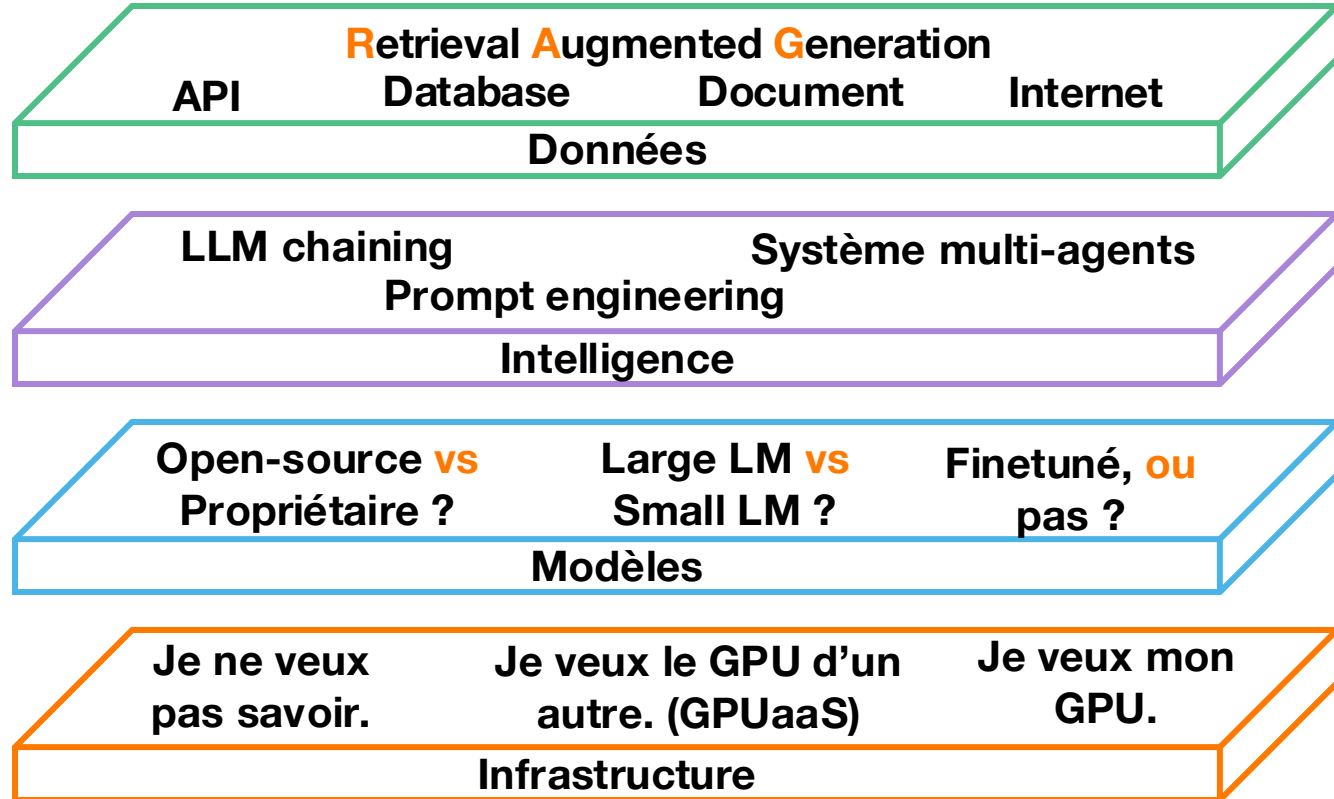
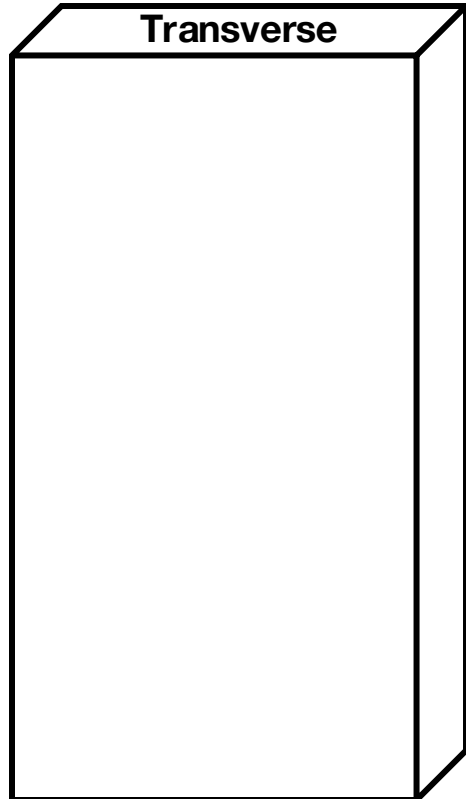
Document



Internet



Les briques de notre **solution**



Transverse

Sécurité

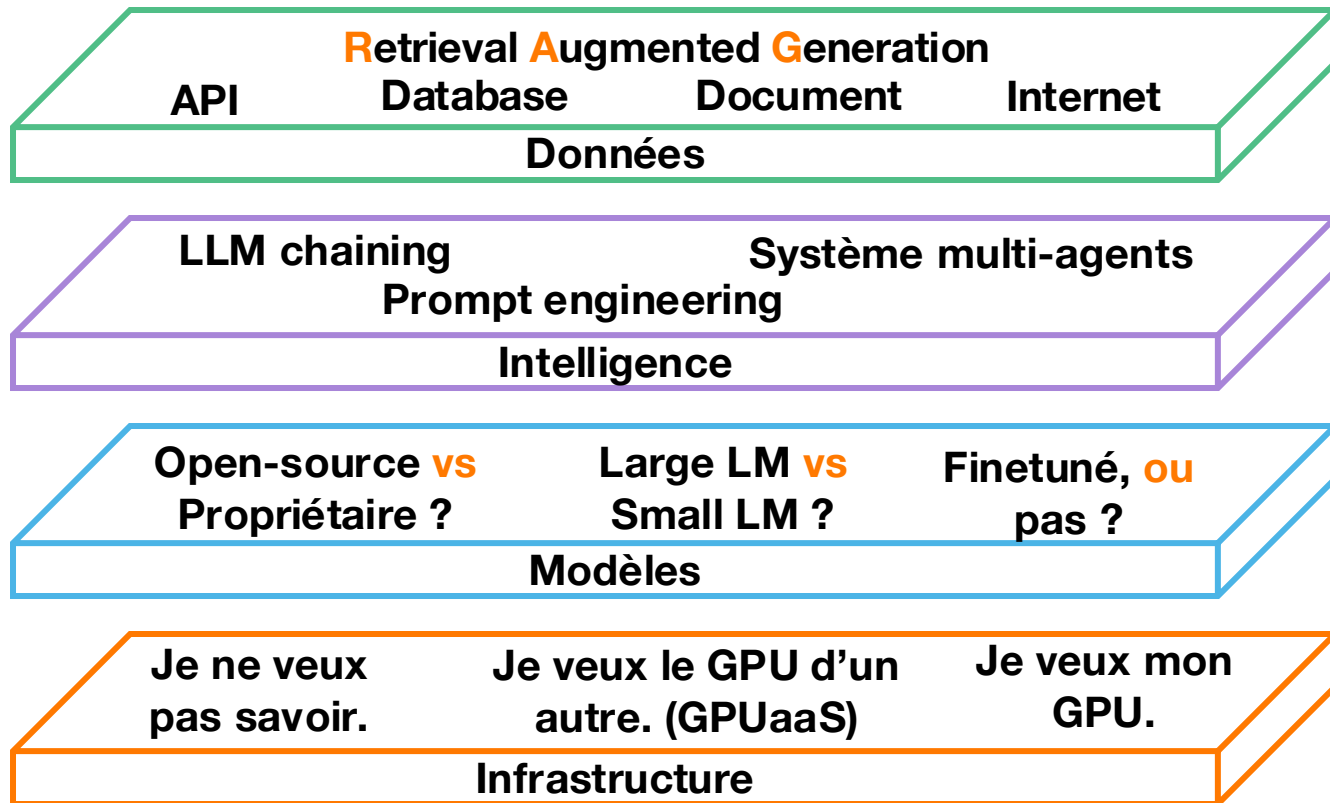
Gouvernance des données

DevOps, MLOps, LLMOps

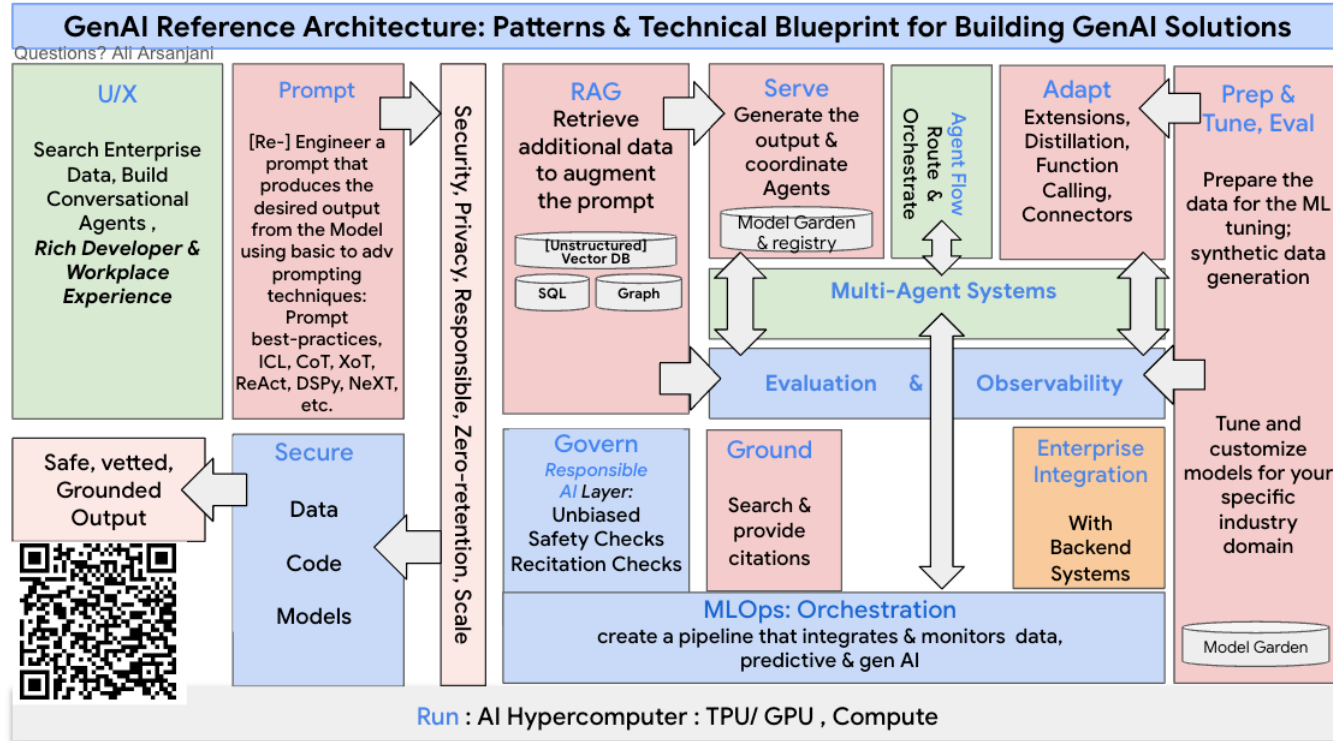
Impact Business et ROI

Éthique et responsabilité

Les briques de notre **solution**



Une architecture plus détaillée



Réalisé par Ali Arsanjani

Que devez-vous **retenir** ?

Le **succès** d'une solution IA Gen, repose sur l'**alignement** entre :

- les utilisateurs,
- la complexité du problème,
- les données,
- l'intelligence,
- l'infrastructure.

Merci de m'avoir écouté !
Des questions ?

Raphaël Anjou - @Orange Business
Azur Tech Winter – 26/11/2024