Rufin VanRullen

CerCo UMR5549

cnrs

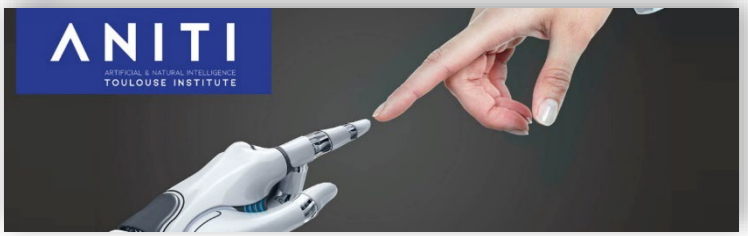
Funded by the European Union

erc

European Research Council

Established by the European Commission

Deep Learning & the Global Workspace Theory



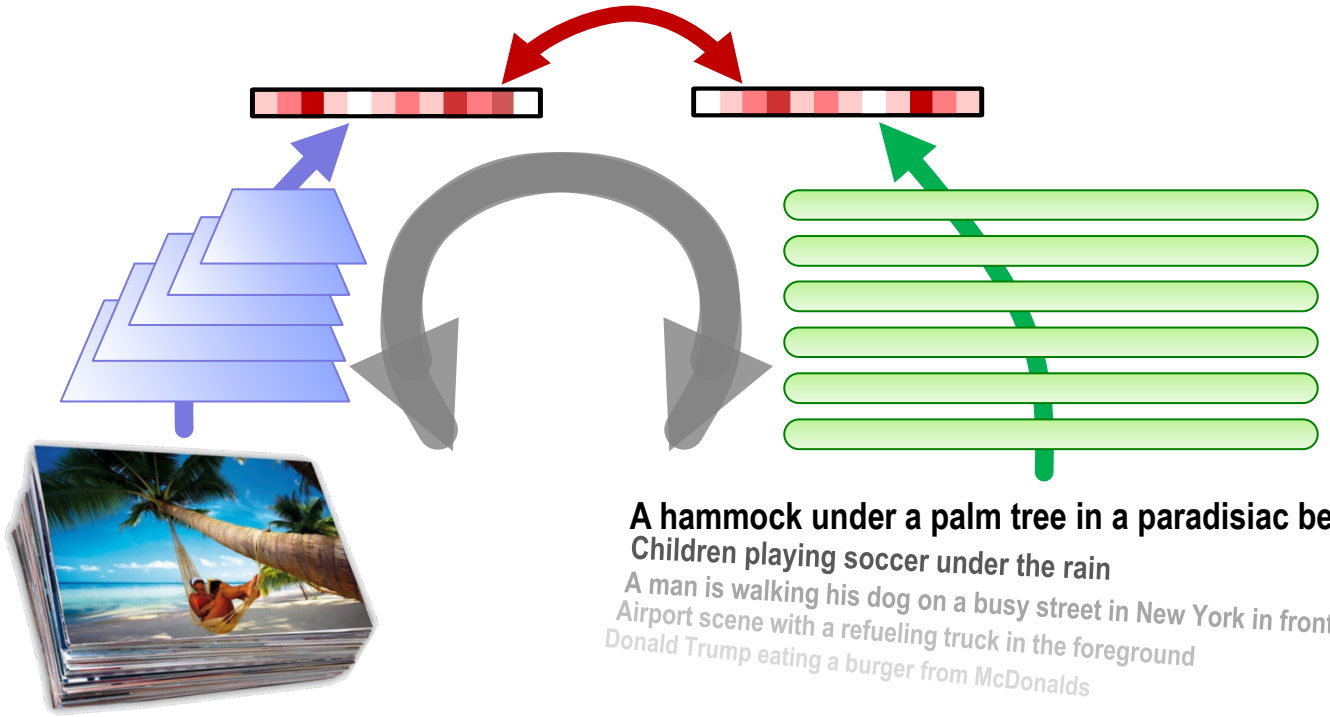
Toulouse Interdisciplinary Deep Learning Group (TIdDLe)

NeuroAi

Multimodal architectures: State-of-the-Art

Massively supervised learning (« brute-force »):

- **Representation alignment:** **CLIP (OpenAI)** **400M image-text pairs**
- **Image-to-Text (captioning):** **CoCa (Google)** **4B**
- **Text-to-Image:** **DALL-E3 (OpenAI)** **>5B**
  Imagen (Google)



Multimodal architectures: Limits

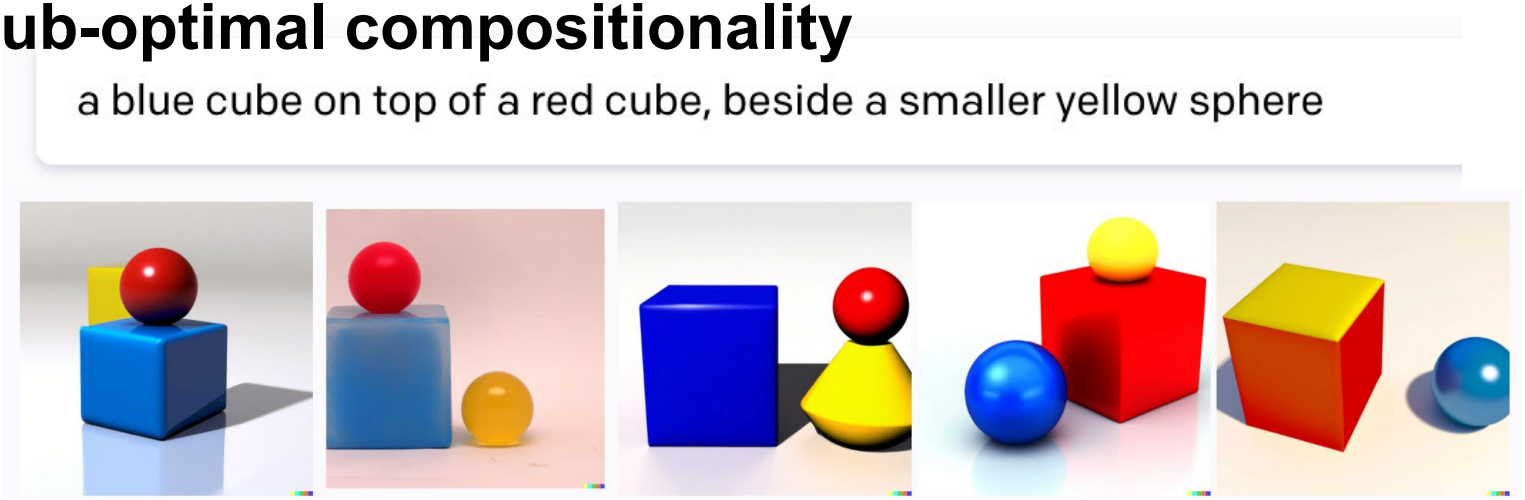
- **Hungry for data & computational resources**

  (a child probably gets <<1M of explicit supervision examples)

- **Sub-optimal multimodal grounding**

  Devillers et al (CoNLL 2021): *Does language help generalization in vision models?*

- **Sub-optimal compositionality**



➔ **Solution A: bigger models, trained with even more data?**

➔ **Solution B: change of paradigm?**

Multimodal architectures: Limits

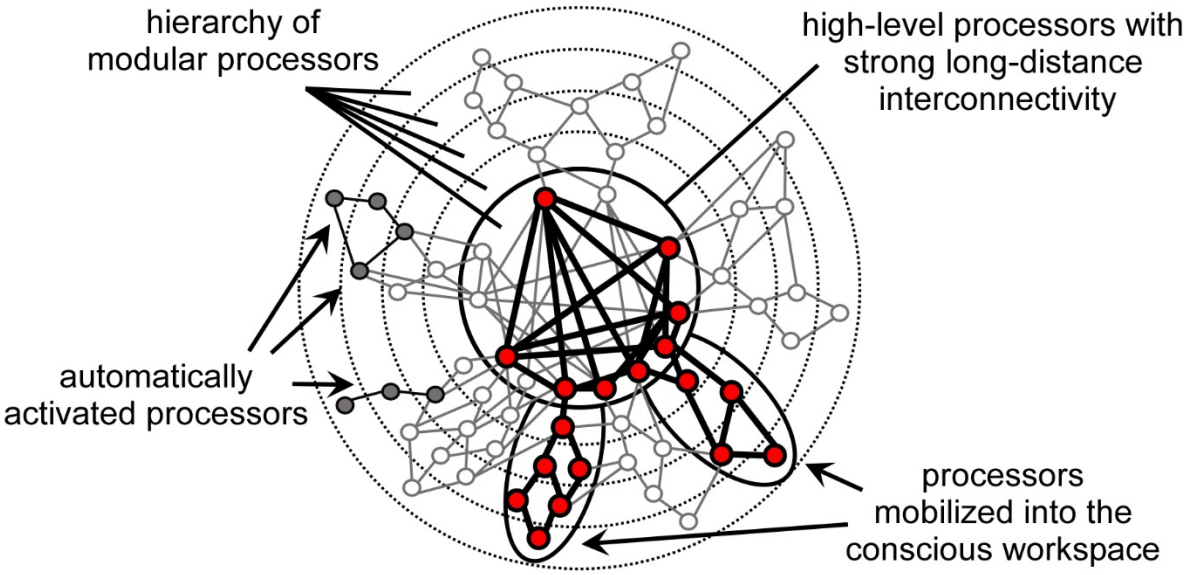
- **Hungry for … ces**
  - **(a child p… n examples)**
- **Sub-optim…**
  - **Devillers e… *…alization in vision models?***
- **Sub-optim…**

➔ **Solution A: bigger models, trained with even more data?**

➔ **Solution B: change of paradigm?**

The Global Workspace Theory

Baars (1993)

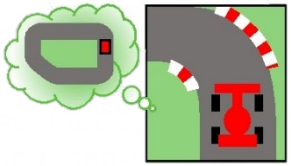


The Global Workspace Theory

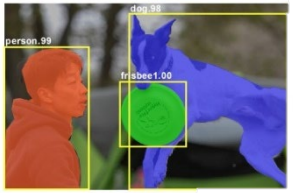
Baars (1993)

Motor outputs

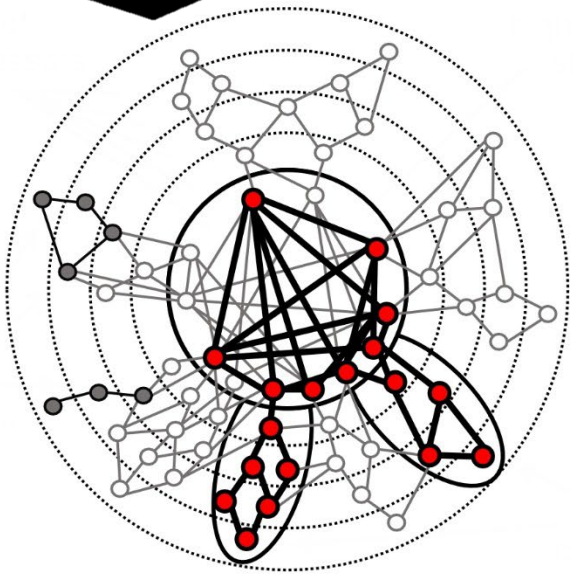
World models



Object detection



Object recognition



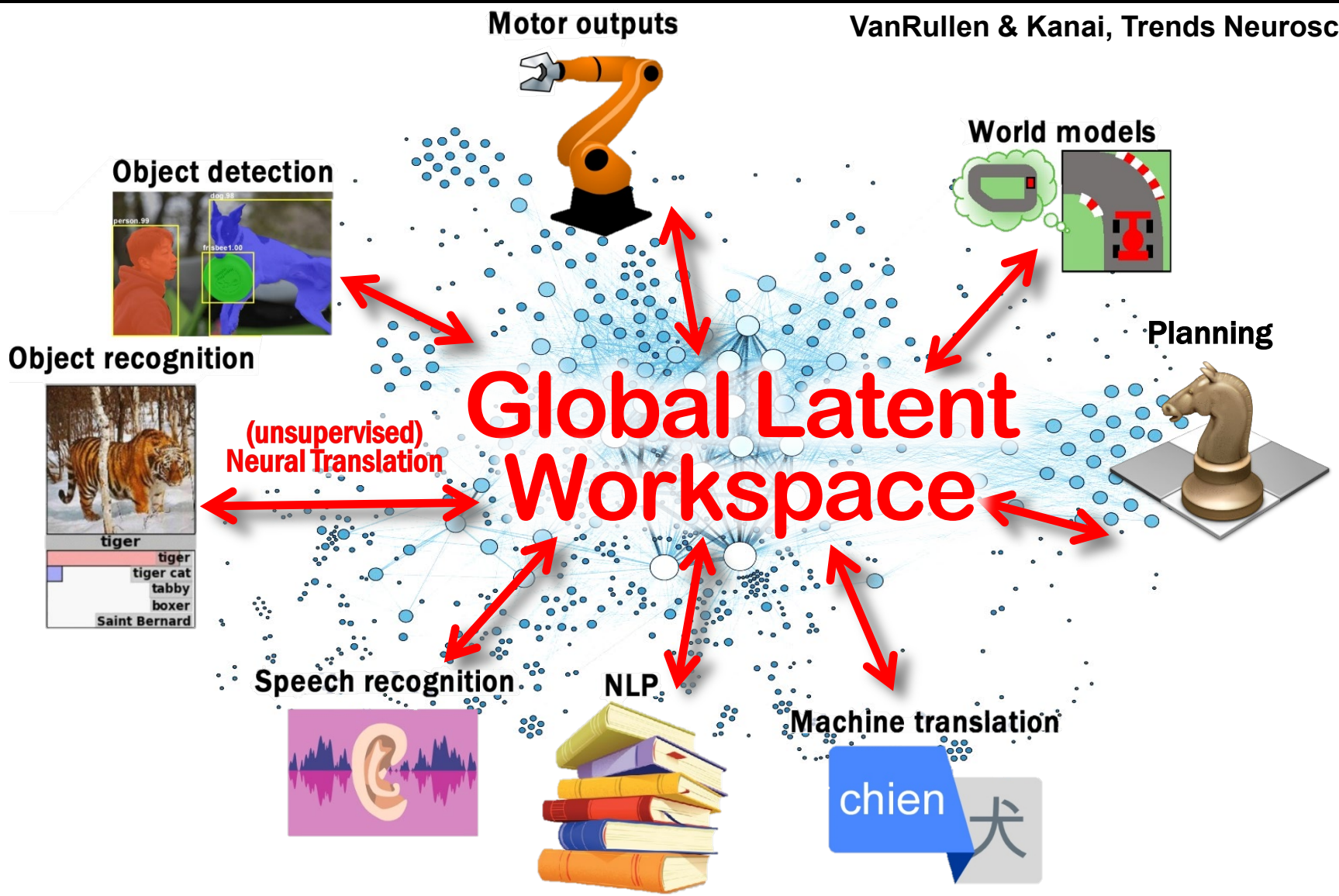
Planning

Speech recognition

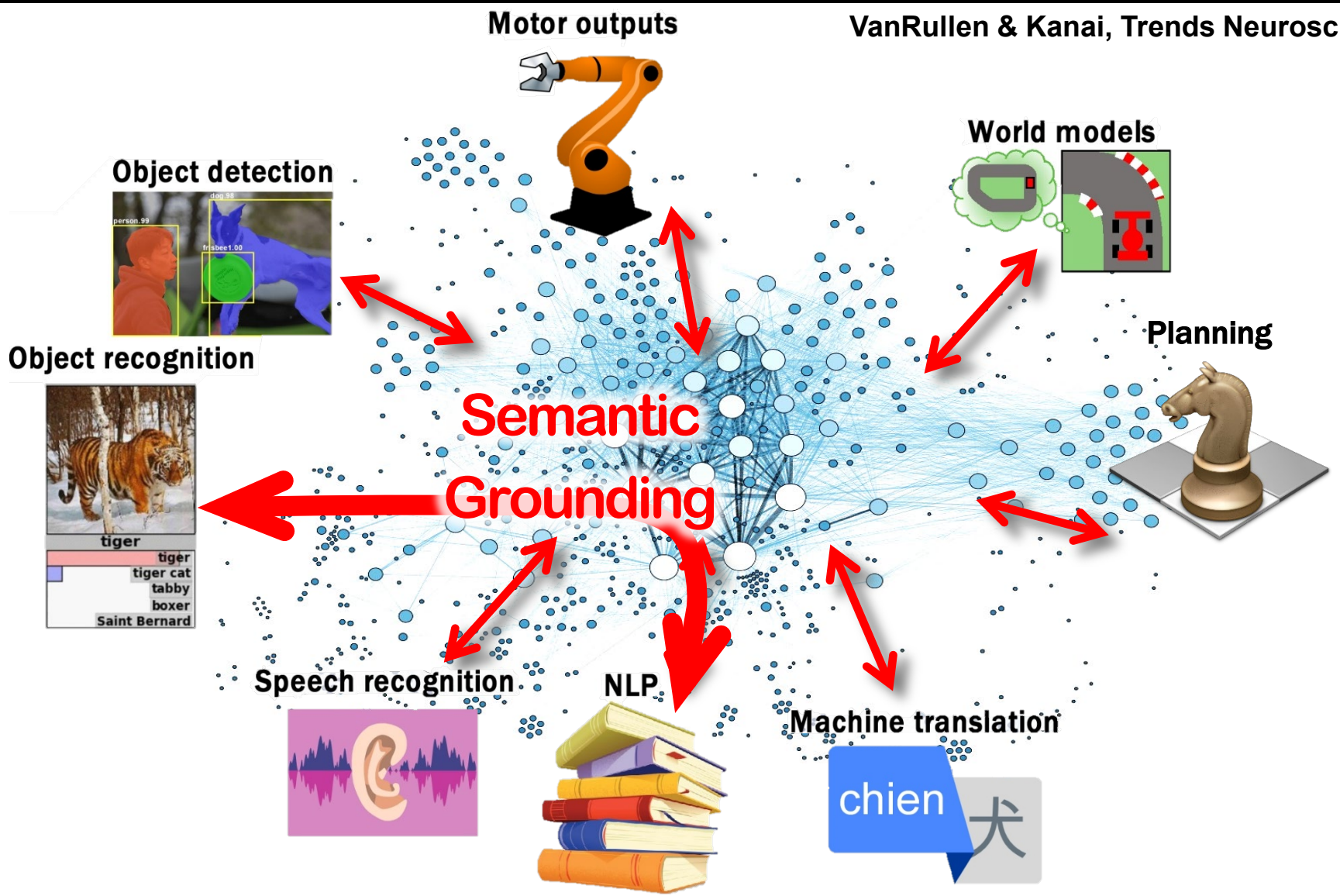
NLP

Machine translation

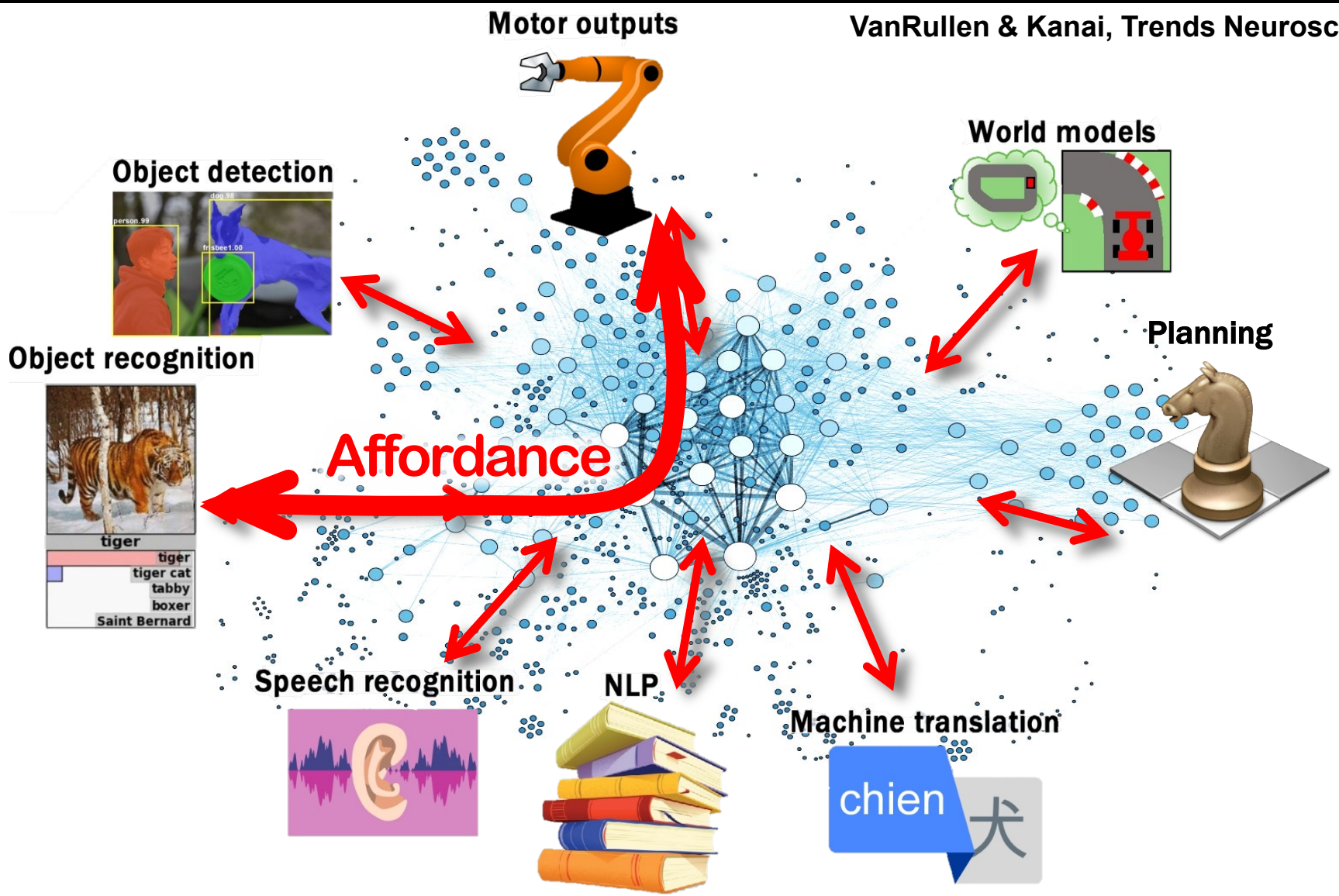
Towards implementing GWT



Towards implementing GWT



Towards implementing GWT



GLW: main concepts

◎Modules

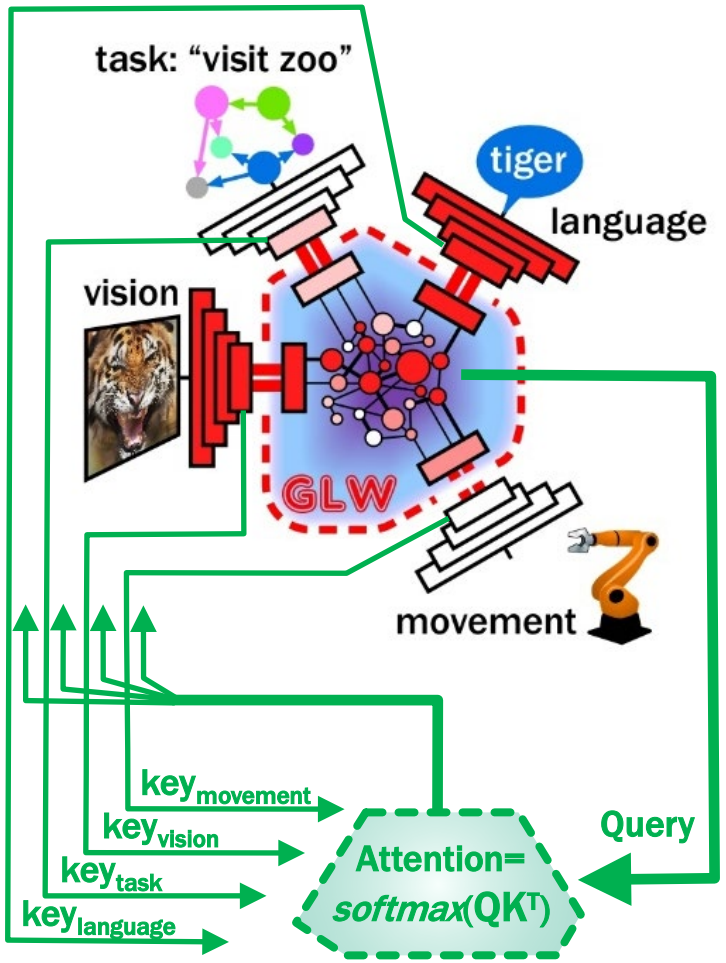
- ◎ (N≥2) Pretrained networks
  (hundreds of available choices!)
- ◎ Choice determines model functionality

◎Broadcast

- ◎ Unsupervised neural translation
- ◎ Trained via cycle-consistency objective

◎Attention

- ◎ Transformer: key-query matching
- ◎ Top-down & bottom-up control



Multimodal systems with Global Workspace

Our ecosystem…

Funded by the European Union

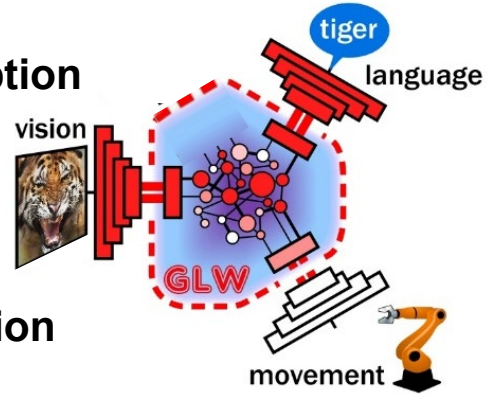
European Research Council
Established by the European Commission

ERC Advanced Project GLOW (2023-2028)

- Develop brain-inspired multimodal deep learning systems
- Evaluate their use and relevance for machine learning
- Advance our knowledge of the brain

ANITI Synergy Chair C3-PO (2024-2028)

- Cobots with Conversation, Cognition & Perception
- Chairs: R.VanRullen (CerCo), N. Asher (IRIT), T. Serre (Brown), O. Stasse (LAAS)
- Frugal multimodal robotic systems with grounded perception, language and action



So the story begins…

L. Maytié