

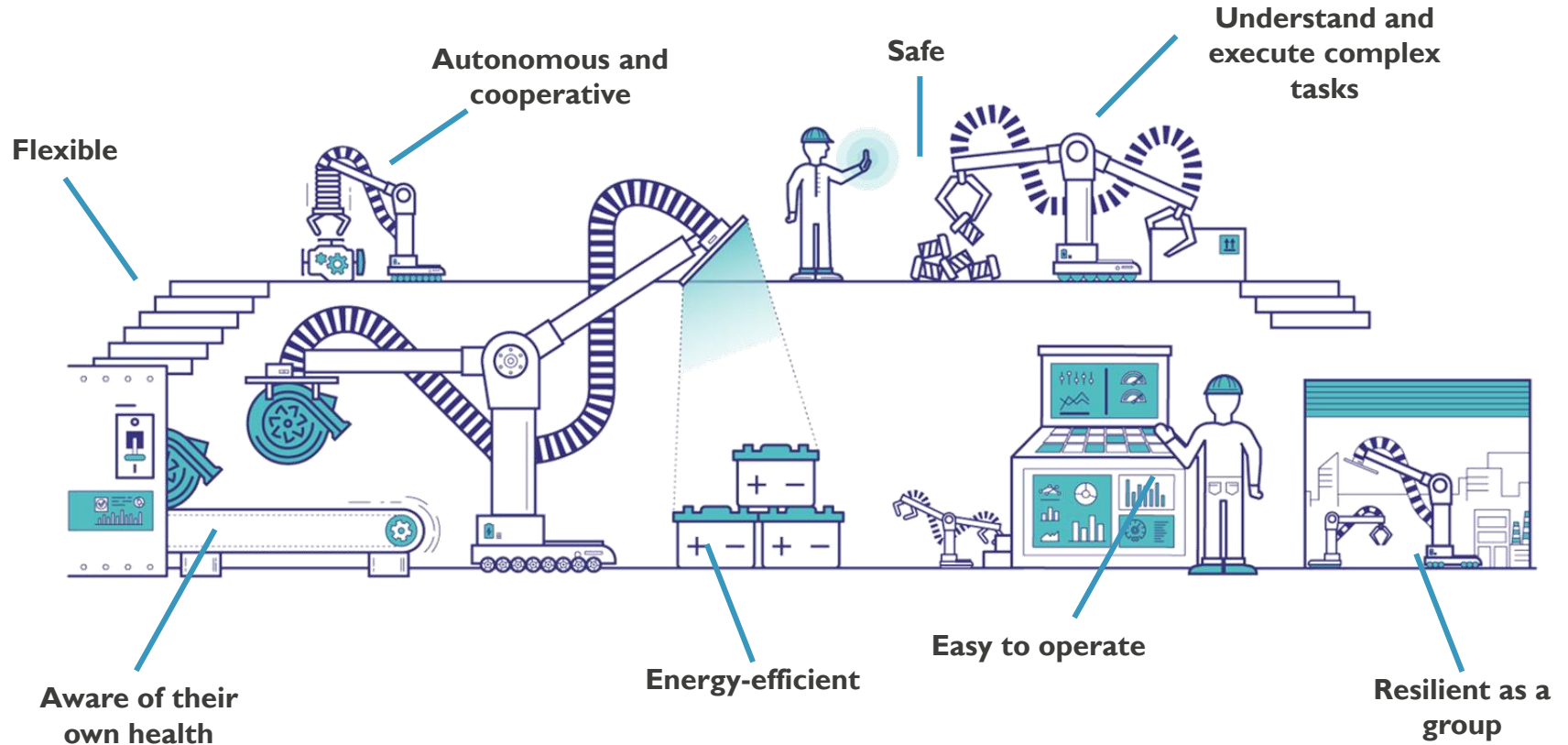


## Seamless AI: Hardware and Software

Axel Nackaerts, Program Manager Artificial Intelligence

# Seamless AI

## Example: Smart Industry



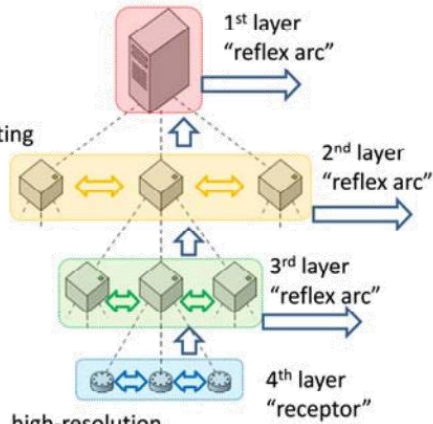
# Seamless AI Challenge

Data center on cloud  
(Behavior analysis)

Intermediate computing  
nodes  
(Spatial-temporal  
association)

Edge devices  
(Pattern recognition)

Sensing network  
(distributed, massive, high-resolution,  
parallel, heterogeneous)



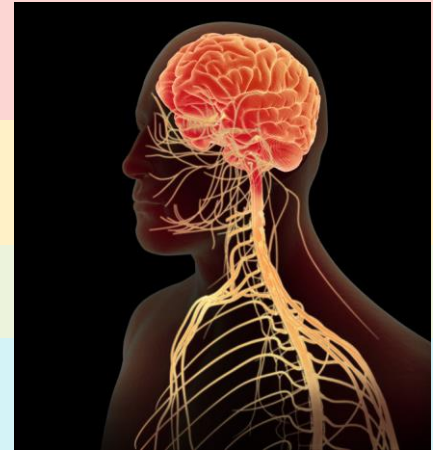
Voluntary action



Reflexes

Kane, Jason, Bo Tang, Zhen Chen, Jun Rong Yan, Tao Wei, Haibo He and Qing Yang. "Reflex-Tree : A Biologically Inspired Architecture for Future Smart Cities." (2014).

Conscious  
Unconscious

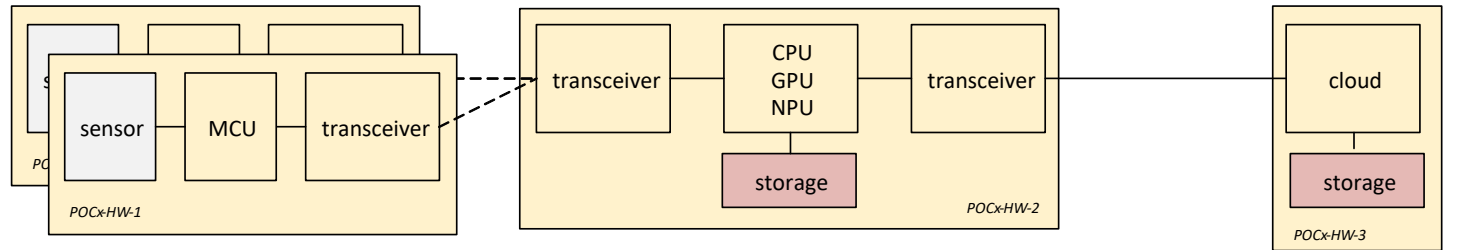
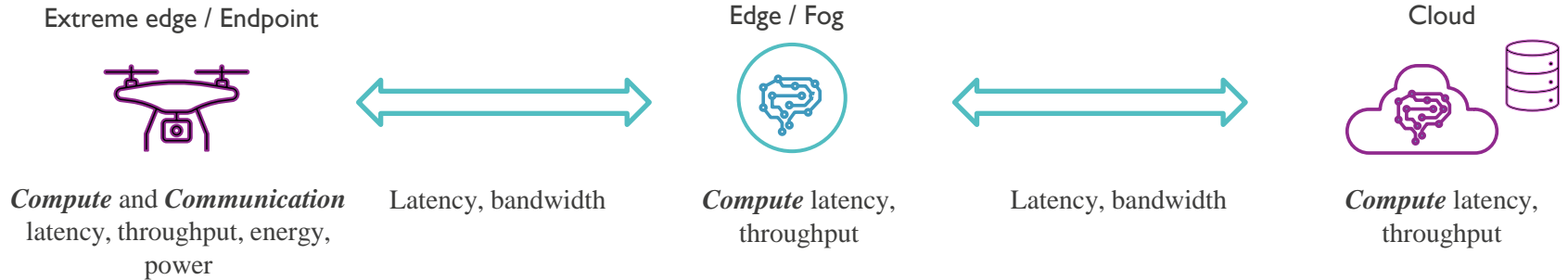


- Ethics
- Explainability
- Privacy and security
- Federated learning
- Distribution methods
- Safety
- Sensor fusion
- Communication
- Edge learning
- Energy efficiency

Distributed hardware platform

# Seamless AI

## System perspective



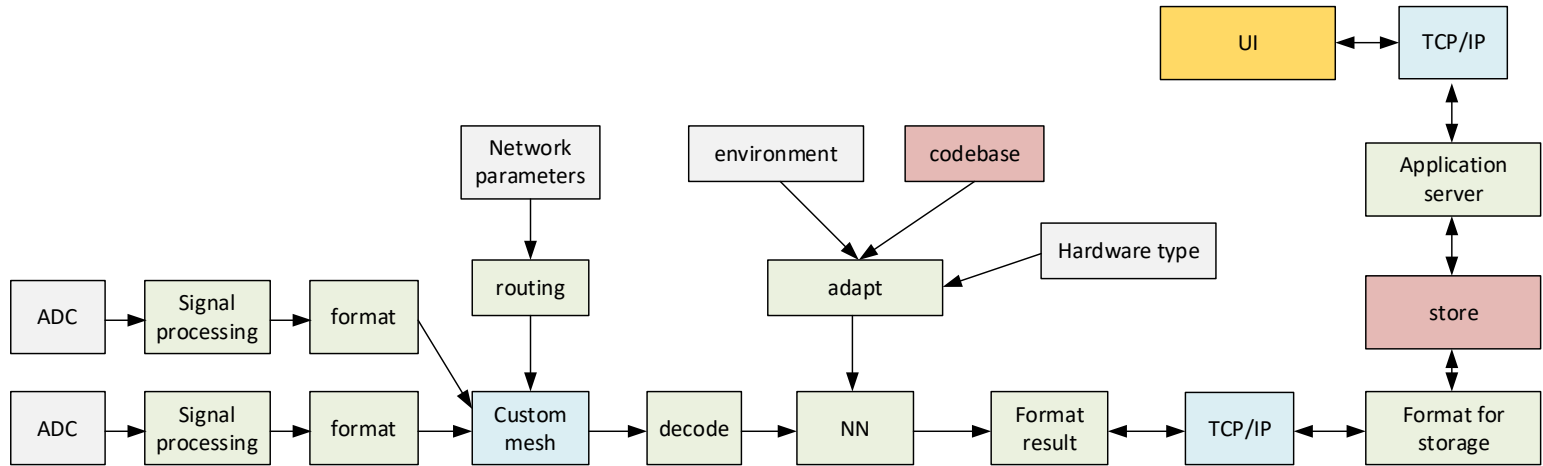
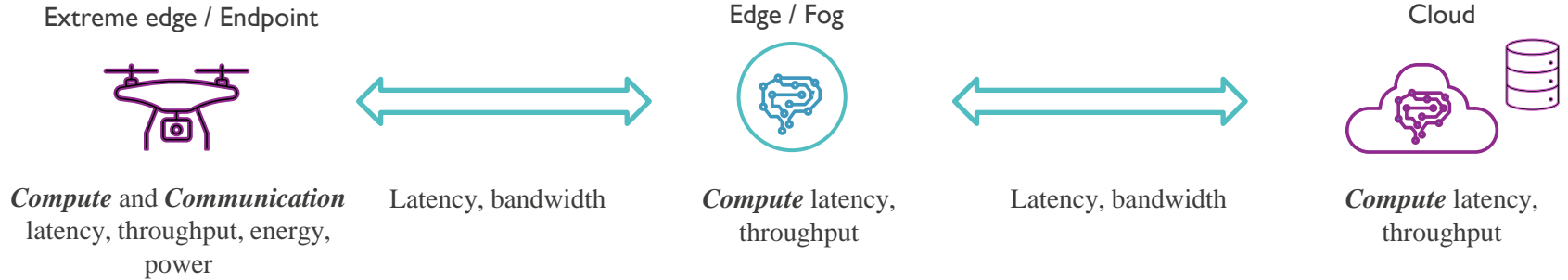
### AI acceleration

Conventional ALU  
Digital MAC array  
Approximating MAC array

Digital in-memory compute  
Analog in-memory compute  
Spiking Neural Network

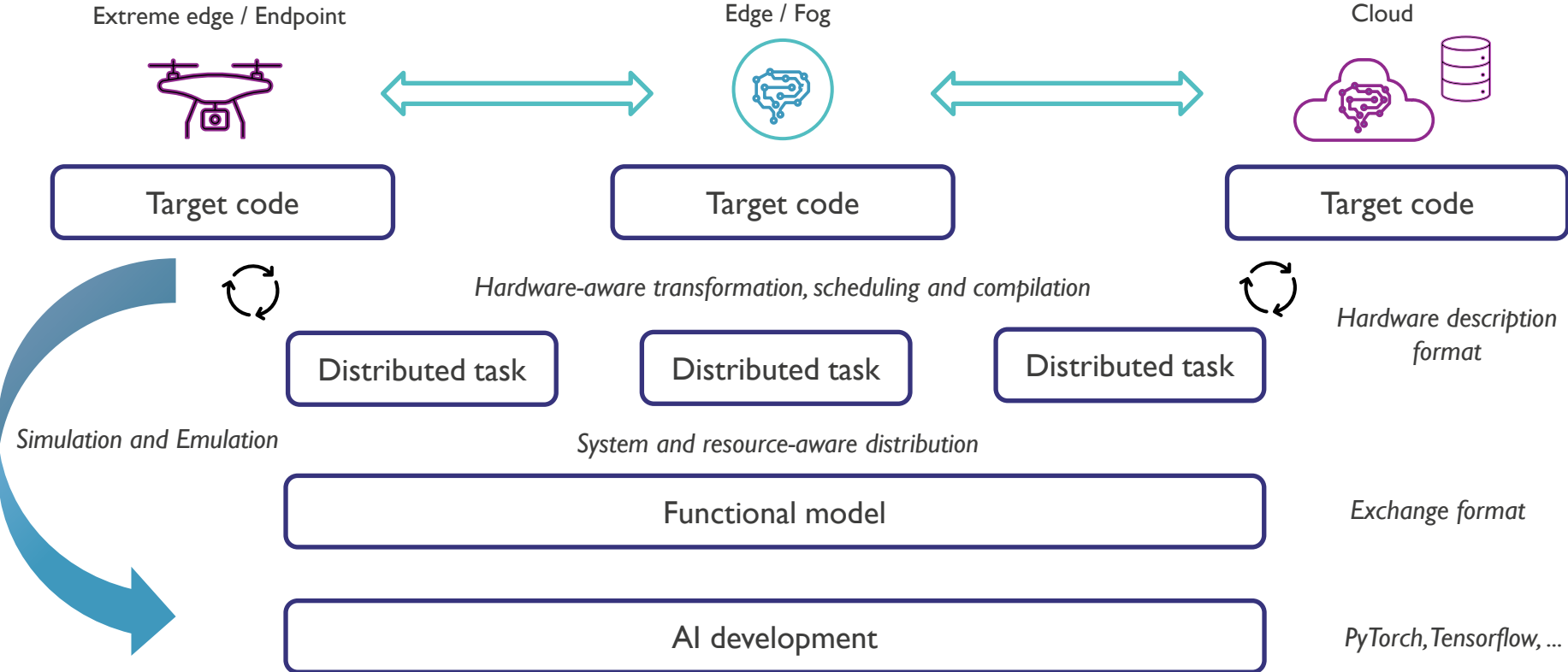
# Seamless AI

## System perspective



# Hardware-Software

Toolflows, API, standardized exchange formats



# Hardware technologies

*Integrate and prototype!*

Extreme edge / Endpoint



Edge / Fog



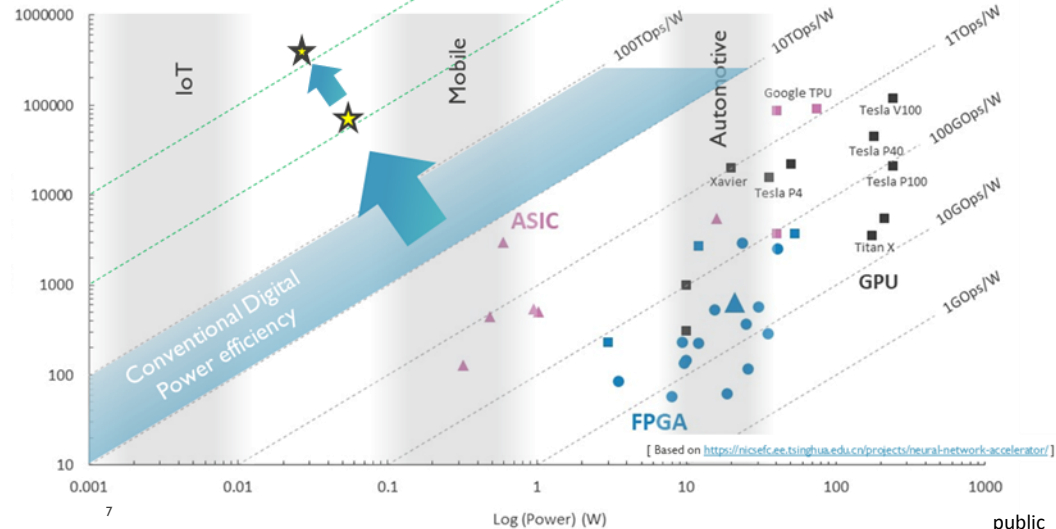
Cloud



Increased specialized computational load  
Application- and node-specific  
Fixed power envelope



- **Specialized** compute units
  - Digital
  - Analog
  - Photonic
- Flexible **integration** (e.g. 3D stacking)
- Increase **power-efficiency**
- Software-aware



# Software

Extreme edge / Endpoint



Edge / Fog



Cloud



- Flexible models
  - Resource-aware
  - Energy-efficient
  - Adapting to HW
  - Targetable to specific situations
- Fast or real-time
- Pre-processing
  - Local sensor fusion
  - Feature extraction
  - Bandwidth reduction

- Distribution engine
  - Remapping to different endpoints
  - Knowledge sharing and training
- Aggregation and sensor fusion
- Hybrid models
- User interaction

- Cognitive models
- Large-scale hybrid models
- Long-term models and data storage
- Training

**Ideally, AI development is agnostic of underlying hardware  
HW-aware, optimizing toolflow does the heavy lifting**



# Cross-disciplinary integration

- Hardware and software co-development
  - From semiconductor technology to AI network topology
  - Different application domains require different and modular hardware solutions
- Integration and Proof-of-Concept building
  - Hardware monolithic or 3D stacked
  - Software toolflows
- Prepared for a modular and flexible world
  - Link hardware modules and software flows
  - De-risk choices for partners



mec

embracing a better life