IARIA

Panel #3

LISBON
April 2026

Theme

Challenges on Human-Machine Reasoning and Co-Existence

ComputationWorld 2026 & DataSys 2026

Moderator

Dr. Steve Chan, Decision Engineering Analysis Laboratory, USA

Panelists

Assoc. Prof. Dr. Herodotos Herodotou, Cyprus University of Technology, Cyprus
Dr. Andre Frank Krause, Rhine-Waal University of Applied Sciences, Germany
Naavya Shetty, University of Illinois Urbana-Champaign, USA
Prof. Dr. Kristina Schaaff, IU International University of Applied Sciences, Germany

Chair Introduction

- Herodotos notes that a Multi-Agent Reinforcement Learning (MARL) approach could help obviate the need for a "central master"…
    - Increases *Diversification* (to avoid transitive closure).
    - Increases *Resiliency* (as agents could be in an MPE paradigm).
- Herodotos comments on handling hardware heterogeneity (e.g., CPU, GPU, and TPU)…
    - Scalar, Vector, Matrix-based Reasoning (differing levels of complexity and dimensionality).
    - High-level Intent, such as Strategic and Tactical (e.g., CO and XO).

Steve Chan
Decision Engineering
Analysis Laboratory

Chair Introduction

VALENCIA
May 2026

- Andre notes the potential stability challenges of loops…
  - Loop Stability (what will the latent stability be).
- Andre notes the challenges of time scales…
  - Submodules operate on varying time scales (impact of unknown time delays, drifts).

Steve Chan
Decision Engineering
Analysis Laboratory

Chair Introduction

- Naavya notes that a participation policy of Accept-Escalate-Defer-Refuse could be useful…
  - Novelty Estimation (e.g., mitigating against overconfidence, driving towards non-apriori facets).
  - Stagnation Estimation (e.g., chances of being trapped at local optima).
- Naavya comments on relational development and illuminates:
  - Functional criteria for intrinsic motivation (e.g., contextual generalization).
  - Theoretical grounding (e.g., scaffolding).

Steve Chan
Decision Engineering
Analysis Laboratory

Chair Introduction

Steve Chan
Decision Engineering
Analysis Laboratory

- Kristina underscores the notion of Human-centered AI…
  - Enhanced contextual awareness (e.g., functional empathy, cognitive empathy, etc.).
- Kristina advocates for Augmentation over Replacement…
  - The goal of AI should be to anneal human capabilities and decision-making (e.g., enhanced decision support).

Panelists

Assoc. Prof. Dr. Herodotos Herodotou, Cyprus University of Technology, Cyprus

Dr. Andre Frank Krause, Rhine-Waal University of Applied Sciences, Germany

Naavya Shetty, University of Illinois Urbana-Champaign, USA

Prof. Dr. Kristina Schaaff, IU International University of Applied Sciences, Germany

Dr. Steve Chan, Decision Engineering Analysis Laboratory, USA

Panelist Position

LISBON
April 2026

Infrastructure as a Peripheral Brain: Enabling Real-Time Offloading

- **Latency Bottleneck:** For a human to offload cognition to a machine, the machine must respond at the speed of thought
- **Dynamic Tiering:** Move from static storage to context-aware memory by optimizing multi-tier distributed systems to prioritize the active working set of human-machine collaboration
- **Predictive Caching as Intuition:** Use online streaming ML to anticipate data needs, mimicking how human memory surfaces information before it's consciously requested

So What? *Systems optimization isn't just about IOPS; it's about maintaining the fluidity of the human-machine reasoning loop*

Herodotos Herodotou
Cyprus University of Technology

Panelist Position

Orchestrating Collaboration: Multi-Agent Reasoning in Clusters

- **Emergent Co-reasoning:** Utilize MARL for autonomous resource negotiation between agents, removing the need for a central master
- **Handling Heterogeneity:** Manage the cognitive diversity of hardware (CPU/GPU/TPU) through intelligent scheduling to support high-level human intent
- **Self-Organizing Systems:** Shift from human-managed to human-guided, automating complex cluster reasoning so humans can focus on high-level intent

So What? *We can't have effective human-machine co-existence if the machines can't even coexist with each other*

Herodotos Herodotou
Cyprus University of Technology

IARIA

PANEL #3

LISBON
April 2026

Panelists

Assoc. Prof. Dr. Herodotos Herodotou, Cyprus University of Technology, Cyprus

Dr. Andre Frank Krause, Rhine-Waal University of Applied Sciences, Germany

Naavya Shetty, University of Illinois Urbana-Champaign, USA

Prof. Dr. Kristina Schaaff, IU International University of Applied Sciences, Germany

Dr. Steve Chan, Decision Engineering Analysis Laboratory, USA

IARIA

Panelist Position

LISBON
April 2026

Human – AI Coexistence: a Dynamical Systems Perspective

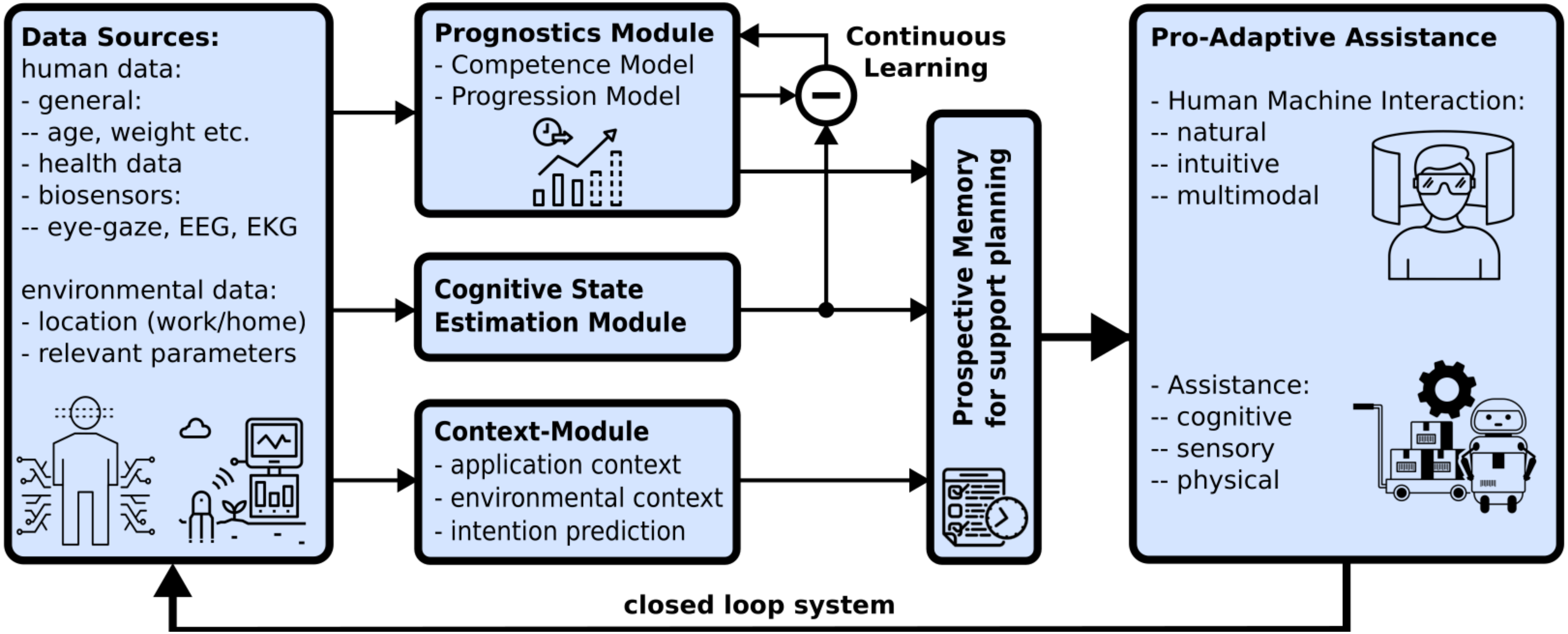
- **Continuous Learning:** Assistive Technology should adapt to the individual user and predict its state, intentions and actions.
- **Humans & AI form a tight feedback loop:** while the AI continuously learns from human parameters, the offered assistance changes the human state and behaviour
- **Human & AI form a complex dynamical system:** will the loop be stable? Or oscillating / chaotic?
- **Challenges**: submodules may operate on vastly different time scales; time delays introduce challenges; CL: distribution drifts

André Frank Krause
Germany
Rhine-Waal University of Applied Sciences

IARIA

Panelist Position

LISBON
April 2026



André Frank Krause
Germany
Rhine-Waal University of Applied Sciences

Image Source: Frank Krause, A., Kannen, K., Büscher, S., Ressel, C., & Wild-Wall, N. (2025, June). Pro-adaptive Cognitive Assistive Technology: Concept and Application in Reading Support for ADHD. In International Conference on Extended Reality (pp. 255-266). Cham: Springer Nature Switzerland.

Panelists

Assoc. Prof. Dr. Herodotos Herodotou, Cyprus University of Technology, Cyprus

Dr. Andre Frank Krause, Rhine-Waal University of Applied Sciences, Germany

Naavya Shetty, University of Illinois Urbana-Champaign, USA

Prof. Dr. Kristina Schaaff, IU International University of Applied Sciences, Germany

Dr. Steve Chan, Decision Engineering Analysis Laboratory, USA

Panelist Position

Shifting the focus from Increased Capability to Better Discipline:
1. Governed Participation

Naavya Shetty
University of Illinois
Urbana-Champaign

THE ARCHITECTURAL GAP

Input arrives → **Full compute engagement** → Output

No routing question asked – ever

Consequences for: Trustworthiness, Efficiency, Transparency

CLOSEST EXISTING WORK – AND WHY IT FALLS SHORT

Token caching / KV cache reuse
Reduces cost of repeated inputs – but decides nothing about whether engagement was appropriate. Makes indiscriminate participation cheaper, not principled.

Speculative decoding / early exit
Adapts compute per token – but purely as performance optimization. The routing question (should the system engage at all?) is never asked.

THE DIRECTION: PARTICIPATION POLICY

ACCEPT	Familiar input – lightweight heuristic sufficient
ESCALATE	Expected benefit justifies heavy downstream reasoning
DEFER	Competence gap – log for future learning, don't guess
REFUSE	Outside domain – principled abstention, not silent failure

Open problems: rule-based controller limits, multimodal novelty estimation, cold-start

Co-existence requires a partner that governs its own engagement – not one that computes indiscriminately on everything that arrives.

IARIA

Panelist Position

LISBON
April 2026

Shifting the focus from Increased Capability to Better Discipline: 2. Motivational Legitimacy

Naavya Shetty
University of Illinois
Urbana-Champaign

THE STRUCTURAL FAILURE

Diagnostic: The Noisy TV Problem (Burda et al.)

Curiosity-driven agent fixates permanently on random noise – not a calibration bug but structural: scalar operationalization strips the relational context that makes the construct functional. Not weaker curiosity but a **categorically different optimization process**. (Goodhart extremal failure – Manheim & Garrabrant)

Pattern generalizes:
Competence-seeking → drive toward maximally uncertain states
Achievement motivation → loophole exploitation
Reward hacking / spec. gaming documented in 70+ instances (Krakovna et al.); unavoidable for any sufficiently expressive reward class (Skalse et al.); reward tampering in LLMs (Denison et al.); scales with capability (Pan et al.)

CLOSEST EXISTING WORK – AND WHY IT FALLS SHORT

RLHF / Constitutional AI: Replace one proxy with another; RLHF produces its own Goodhart dynamics. Constraints around an unchanged motivational architecture – fences, not a different optimizer.
Autotelic agents (Colas et al.): Goal diversity ≠ motivational legitimacy. Baldassarre et al. explicitly: no criterion for purpose legitimacy.

THE DIRECTION: RELATIONAL DEVELOPMENT

Behavioral compliance ≠ Motivational legitimacy

Permanent heteronomy (Kant): capable optimizer pointed temporarily in a useful direction. Remove the incentive and the behavior degrades.

Functional criteria for intrinsic motivation (SDT)

- **Reward-free stability –** Aligned behavior without external contingencies
- **Contextual generalization –** Transfers to novel contexts without new regulation
- **Goal consistency –** Behavior reflects stable internal motivational priorities

Theoretical Grounding (underutilised in AI)

- Hegel / Brandom – recognition condition on goal legitimacy
- Vygotsky ZPD – structured interaction not training duration
- SDT internalization continuum
- Haidt SIM – moral motivation through social calibration

Early stage concerns

- **Scalability –** meaningfully constitutive, not superficial formality
- **Manipulation –** problem of normative framework for scaffolding partner
- **Scope –** universalisability argument or explicit scope boundary required

Co-existence demands a partner whose priorities survive the removal of the reward signal – behavioral success alone is not motivational legitimacy.

**Early-stage framework in development – no results claimed*

Panelists

Assoc. Prof. Dr. Herodotos Herodotou, Cyprus University of Technology, Cyprus

Dr. Andre Frank Krause, Rhine-Waal University of Applied Sciences, Germany

Naavya Shetty, University of Illinois Urbana-Champaign, USA

Prof. Dr. Kristina Schaaff, IU International University of Applied Sciences, Germany

Dr. Steve Chan, Decision Engineering Analysis Laboratory, USA

Panelist Position

- Human-centered AI
  - AI should be developed not only for accuracy and efficiency, but also with attention to human needs, relationships, and social contexts.
- Empathy and ethics must go together
  - Emotion-aware systems can improve interaction and support, but they also raise serious questions around privacy, manipulation, and trust.
- Augmentation over replacement
  - The goal of AI should be to strengthen human capabilities and decision-making, not to substitute human agency.

Kristina Schaaff
IU International
University of
Applied Sciences

IARIA

PANEL #3

LISBON
April 2026

Panelists

Assoc. Prof. Dr. Herodotos Herodotou, Cyprus University of Technology, Cyprus

Dr. Andre Frank Krause, Rhine-Waal University of Applied Sciences, Germany

Naavya Shetty, University of Illinois Urbana-Champaign, USA

Prof. Dr. Kristina Schaaff, IU International University of Applied Sciences, Germany

Dr. Steve Chan, Decision Engineering Analysis Laboratory, USA

Panelist Position

- **Challenges of a leveraging a Metacognitive Module (MCM)**

Steve Chan
Decision Engineering Analysis Laboratory

- **AI Data Centers (AIDC) and AI Systems (AIS) are already struggling with issues of energy efficiency and Operations & Maintenance (O&M) challenges**
    - For a large-scale AIS with production clusters, let us take the case of a single cluster drawing 10 MW running continuously for 24 hours; this might equate to 240MWh/day. Accordingly, if an AIDC had 10 such clusters drawing 100 MW running continuously for 24 hours x 365 days per year, this might equate to ≈ 876,000 MWh (876 GWh) of energy use. In addition, the on-site cooling might consume 1.2 million gallon of water per day for a 100MW AIDC and 5+ million gallons for a 500MW AIDC (Source: Control Associates Inc.).
- **AIDC and AIS are intricately intertwined**
    - In essence, an AIDC is a physical facility that provides the supporting infrastructure for continuous operations and hosts a large-scale AIS.
    - An AIS could be a single system or software stack. It could be a single server or cluster. Regardless, the AIS requires software updates/hardware refreshes as well as model training/tuning/re-training, etc.
    - The requirements of the AIS affects the host AIDC. For example, there are facility services requirements (e.g., power, cooling, monitoring). In addition, among others, there are requirements for: storage/database, network communications, data processing, training/inferencing, etc.
- **MCM**
    - For AIS, an MCM layer—for monitoring and regulating the involved AIS reasoning—will likely add extra computational overhead, which indirectly increases energy and water requirements at the AIDC. The communications overhead increases as the MCM needs to exchange messages across a myriad of nodes; the memory overhead increases, as tracking interim states and logging of self-assessment checkpoints requires memory access and storage usage; and the checking, evaluating, and planning functions each have their own computational overhead requirements, etc.

IARIA

Panelist Position

LISBON
April 2026

- **Notional Energy/Water Utilization Increases? Possibly… (again, these are notional figures)**
  - Small-scale AIS: 5-15% energy overhead increase, with accompanying water overhead increase.
  - Medium AI cluster: 10-15% energy overhead increase, with accompanying water overhead increase.
  - Hyperscale AIDC: 5-20% energy overhead increase, with accompanying water overhead increase.
- **Prospective Net Energy Reduction? Not sure…**
  - Nezhati alludes to the notion that a cognitive substrate (somewhat akin to the MCM) “that can adapt to new tasks” might use “67% less energy than traditional systems. (Source: M. Nehzati, “Self-evolving cognitive substrate through metabolic data processing and recursive self-representation with autonomous memory prioritization mechanisms,” *Frontiers in Artificial Intelligence*, Dec. 2025).
  - Sangarya asserts that if an MCM “continual learning” paradigm can facilitate “retraining an existing model” and reduce the need for “training a new model from scratch,” then there might be computational and energy savings; the technique for forecasting what the computational and energy costs might be is based upon a technique entitled, “REpresentation Shift QUantifying Estimator (RESQUE).” (Source: V. Sangarya and Jung-Eun Kim, “RESQUE: Quantifying Estimator to Task and Distribution Shift for Sustainable Model Reusability,” Arxiv.org, Dec. 2024).
- **Dearth of Literature on AI Energy Utilization…**
  - While the instantaneous power draw is likely higher, the total energy utilized per task and/or per learning cycle may possibly be lower for the cases of early stopping, adaptive computation, and/or “continuous learning,” etc. However, there seems to be a dearth of energy-focused research on not only MCM usage, but AI energy utilization in general. Rather, the dominant evaluation metric in AI has been task performance (e.g., Intersection over Union, Mean Average Precision, Top-K Accuracy, etc.) (for the computer vision field), not necessarily energy efficiency (although kWh or joules utilized during training/inferencing might be available in some cases).

Steve Chan
Decision Engineering
Analysis Laboratory

Questions & Answers Session
at Panel #3…

Thank you for joining us
at Panel #3

Have a wonderful conference!