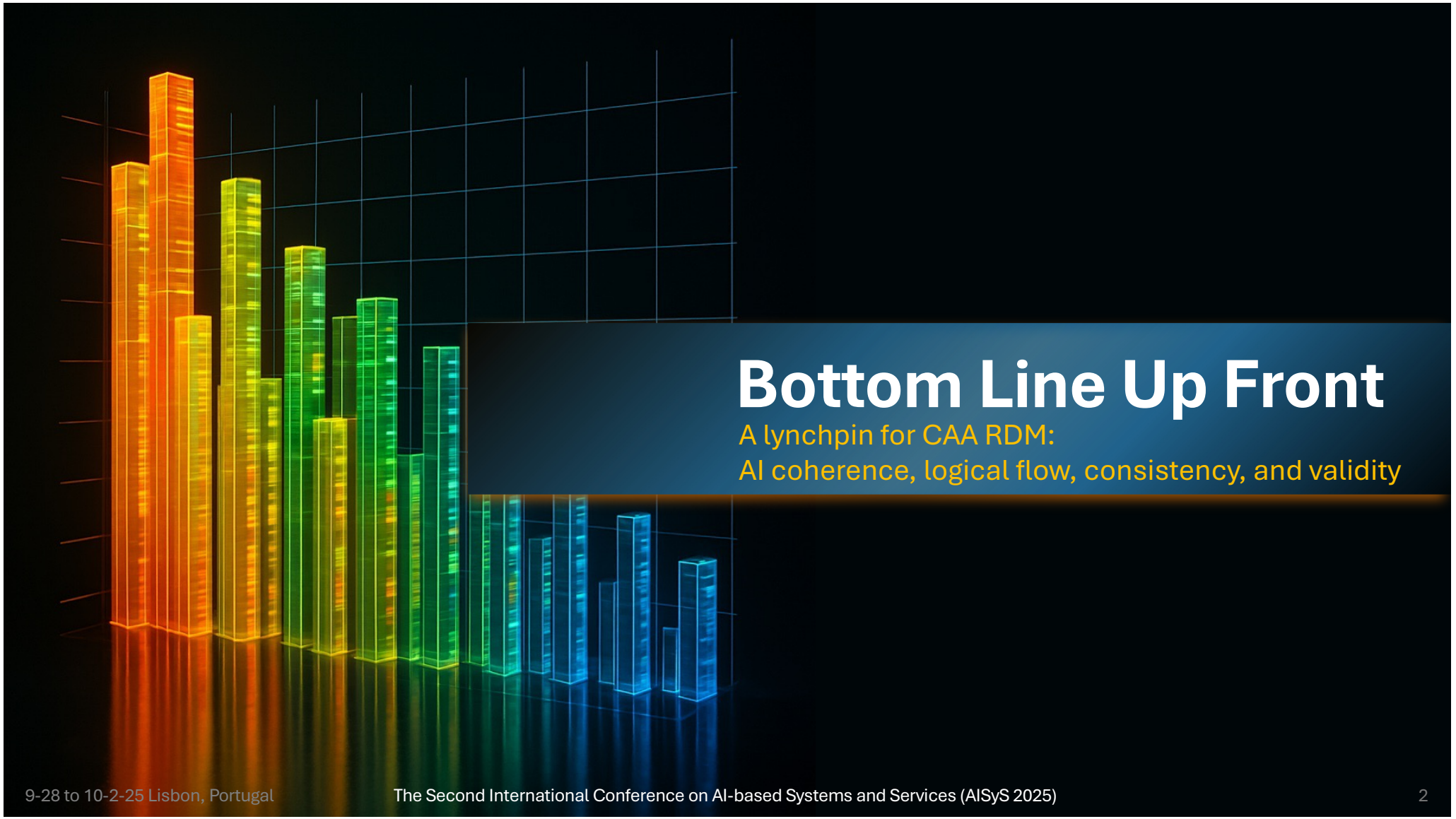
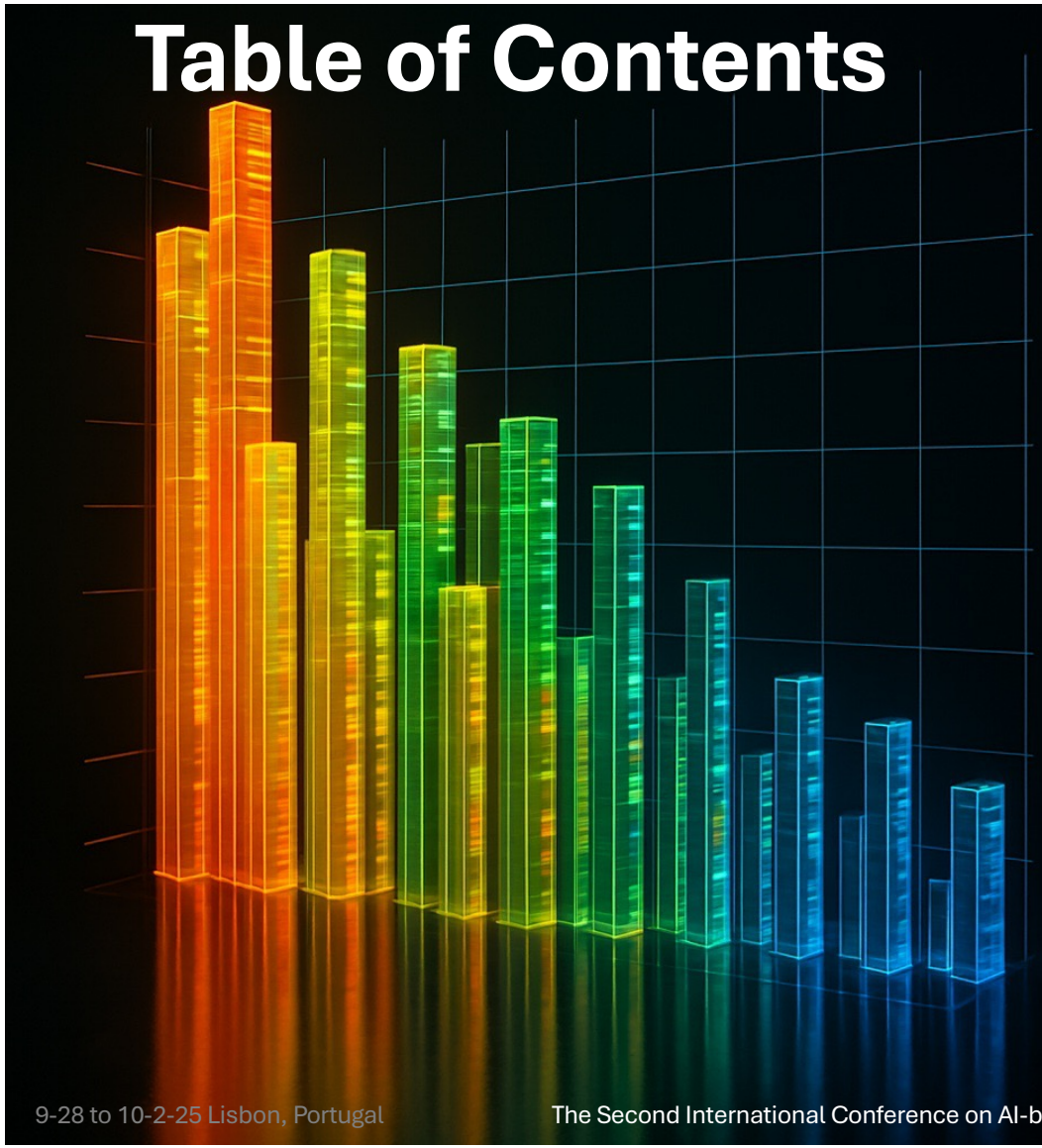
Conversational AI Agents (CAA): The Intricacies of Robust Dialogue Management (RDM)

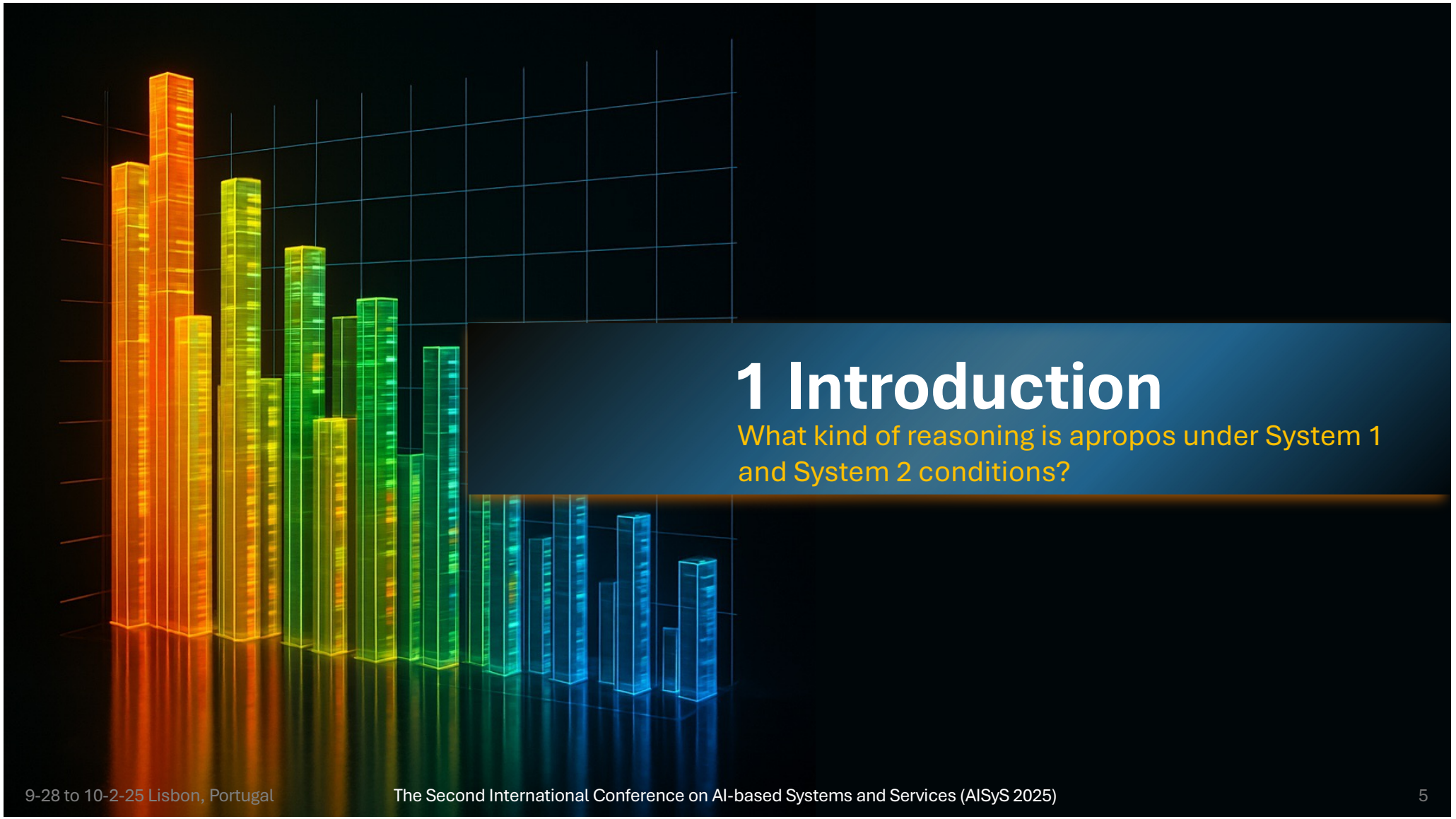
Speaker: Steve Chan
VTIRL, VT/DE-CAIR


Bottom Line Up Front
A lynchpin for CAA RDM:
AI coherence, logical flow, consistency, and validity

BLUF

Table of Contents





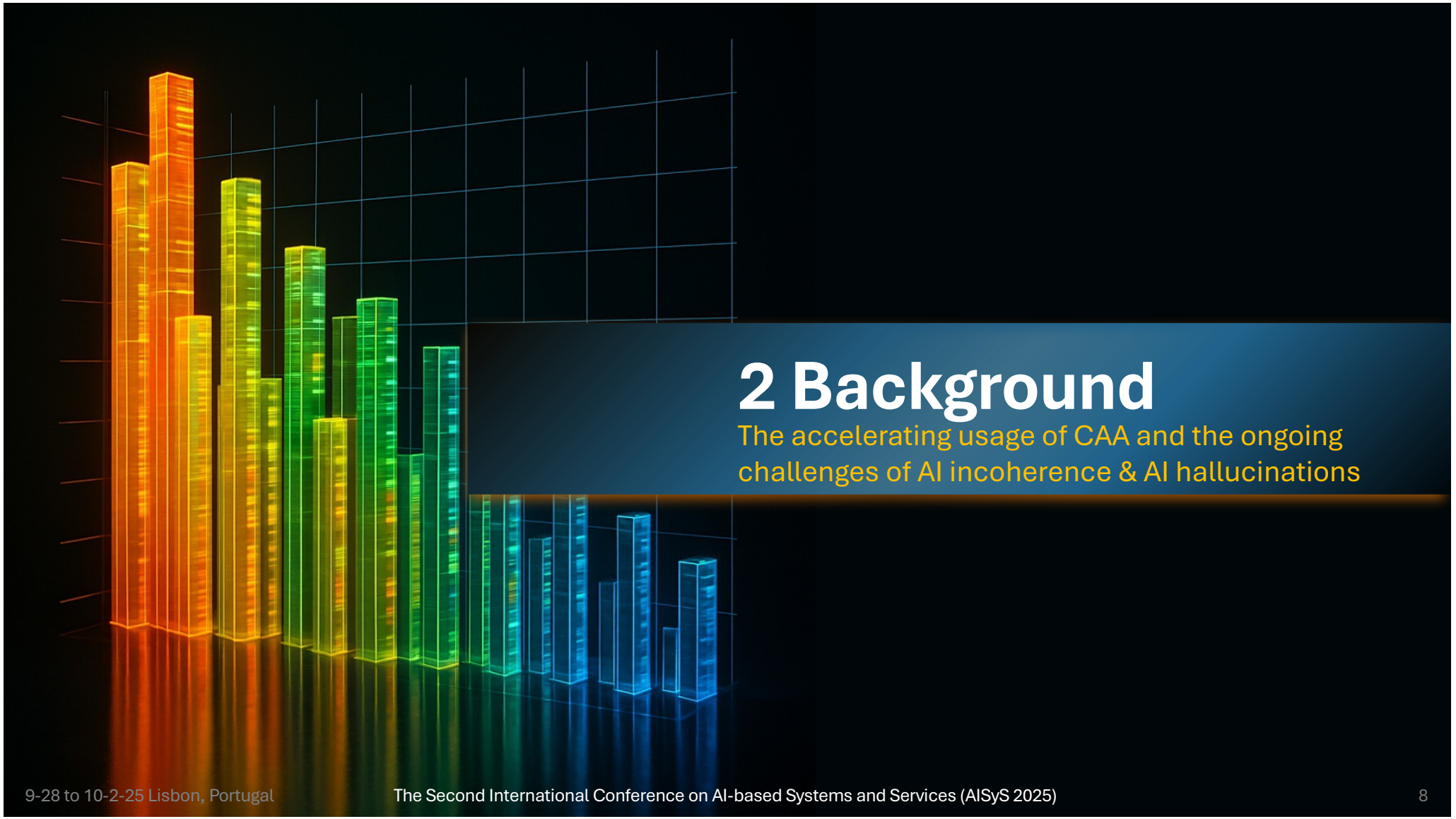
1 Introduction

What kind of reasoning is apropos under System 1 and System 2 conditions?

Introduction

Introduction

- **Enhancing AI Coherency and RDM:**
  - Decision Quality (DQ)
  - Sequential Decision-Making (SDM)
  - DM Inflection Points (DMIP)
  - System 1 (“fast, automatic, intuitive”)
    System 2 (“slow, deliberate, logical”)
  - Uncompressed Decision Cycles (UDC)
    Compressed Decision Cycles (CDC)
  - UDC -> Deductive Reasoning (DedR)
    CDC -> Inductive Reasoning (IndR)
    -> Analogical Reasoning (AnaR)
    -> Case-Based Reasoning (CBR)
  - Lower Ambiguity Higher Uncertainty (LAHU)
    as contrasted to Higher Ambiguity Lower Uncertainty (HALU)
  - Enhanced Coherency, Logical Flow, Consistency, Validity
  - Robust Dialogue Management (RDM)



2 Background

The accelerating usage of CAA and the ongoing challenges of AI incoherence & AI hallucinations

Background

Background

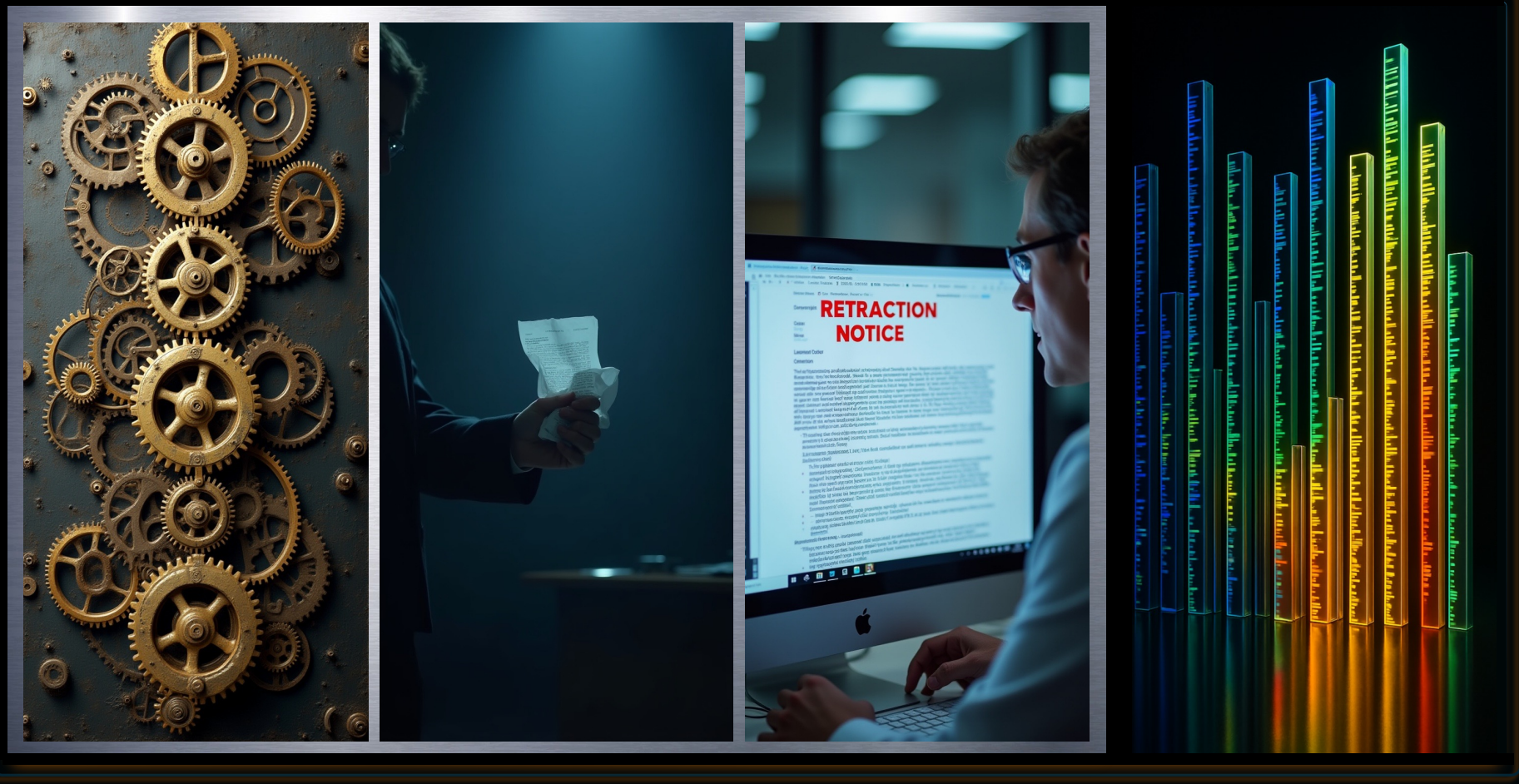
- **Conventional Mitigation Approaches for AI incoherency/AI hallucinations:**
  - Guardrails/System Prompts (GSPs)
  - Fine-Tuned Models (FTMs) for specialized domains
  - Retrieval-Augmented Generation (RAG)

- **An Alternative Approach to address AI incoherency/AI hallucinations:**
  - AI hallucinations
  - AI coherency
  - Maintaining AI coherency amidst Real-World Scenarios (RWS) is challenging since incoming information may re-contextualize and/or contradict matters
  - Enforcing a strict monotonic paradigm can lead to an awkward rigidity and/or incorrect/irrelevant responses
  - Enhanced insight at Monotonic/Non-monotonic Transition Zones (MNTZs) is crucial for higher affinity towards monotonic zones and higher coherency

3 Some Notions...

Amalgams of Reasoning Mechanisms (RMs), Reasoning Processes (RPs), and their pathways

Some Notions...



Some Notions...

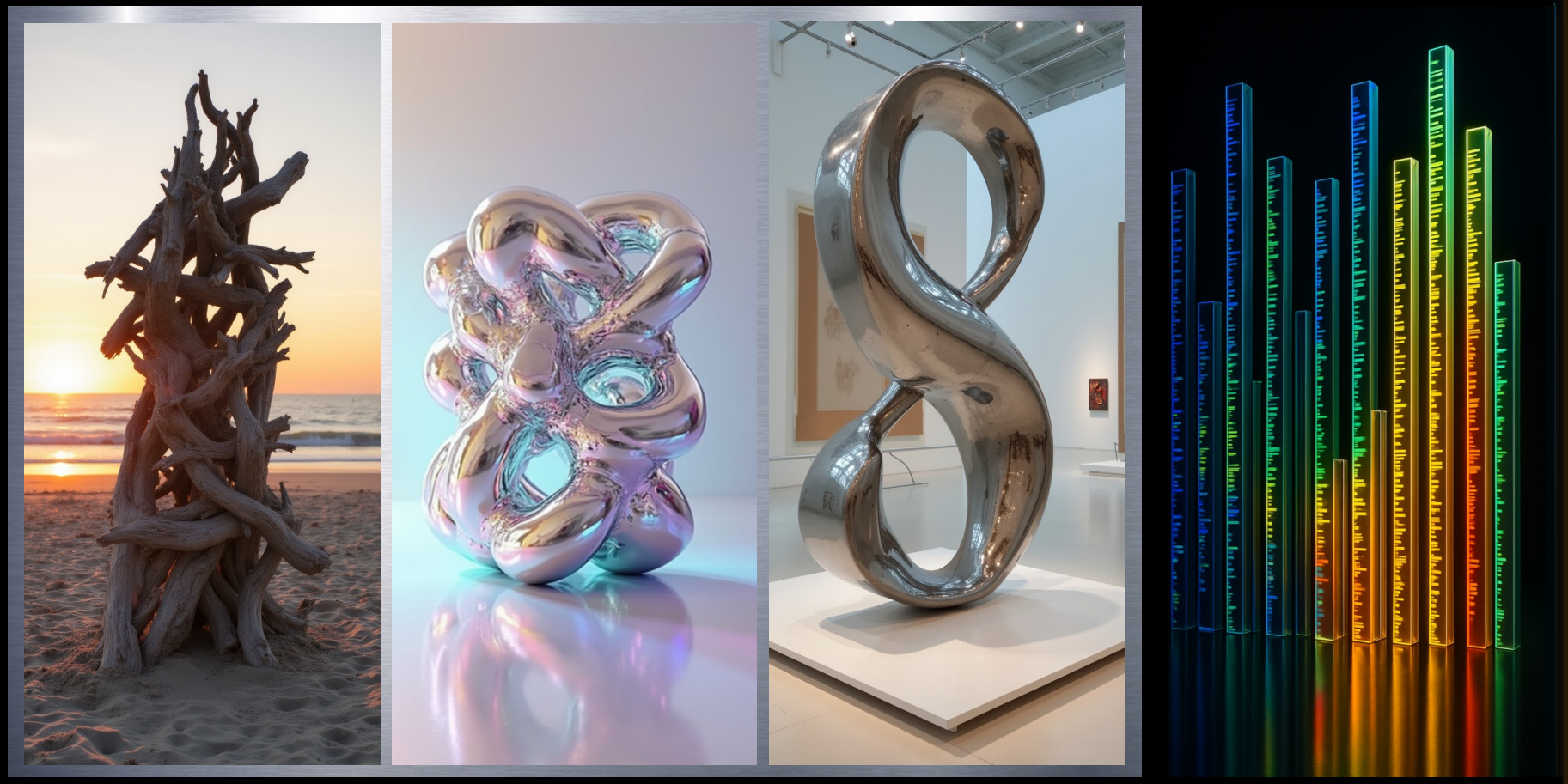
- **Theoretical Foundations:**
  - Reasoning Mechanisms (RMs)
  - Reasoning Processes (RPs)
  - Comparator Similarity Measure (CSM)

- **Validity Considerations:**
  - Epistemological constructs, such as the
    Known Knowns (KK)
    Known Unknowns (KU)
    Unknown Knowns (UK)
    Unknown Unknowns (UU)
  - Gekhman's Sampling-based Categorization of Knowledge (SliCK) model
    Highly Known (HK) | Validity = Always
    Maybe Known (MK) | Validity = Sometimes
    Weakly Known (WK) | Validity = Never with T=0, but Sometimes with T>0
    Unknown | Validity = Never

4 Discussion

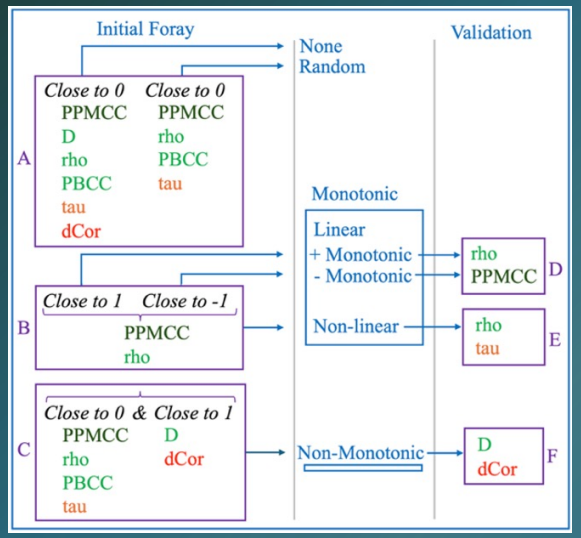
Some pragmatic pathways that may better lend to Enhanced AI coherence and RDM

Discussion

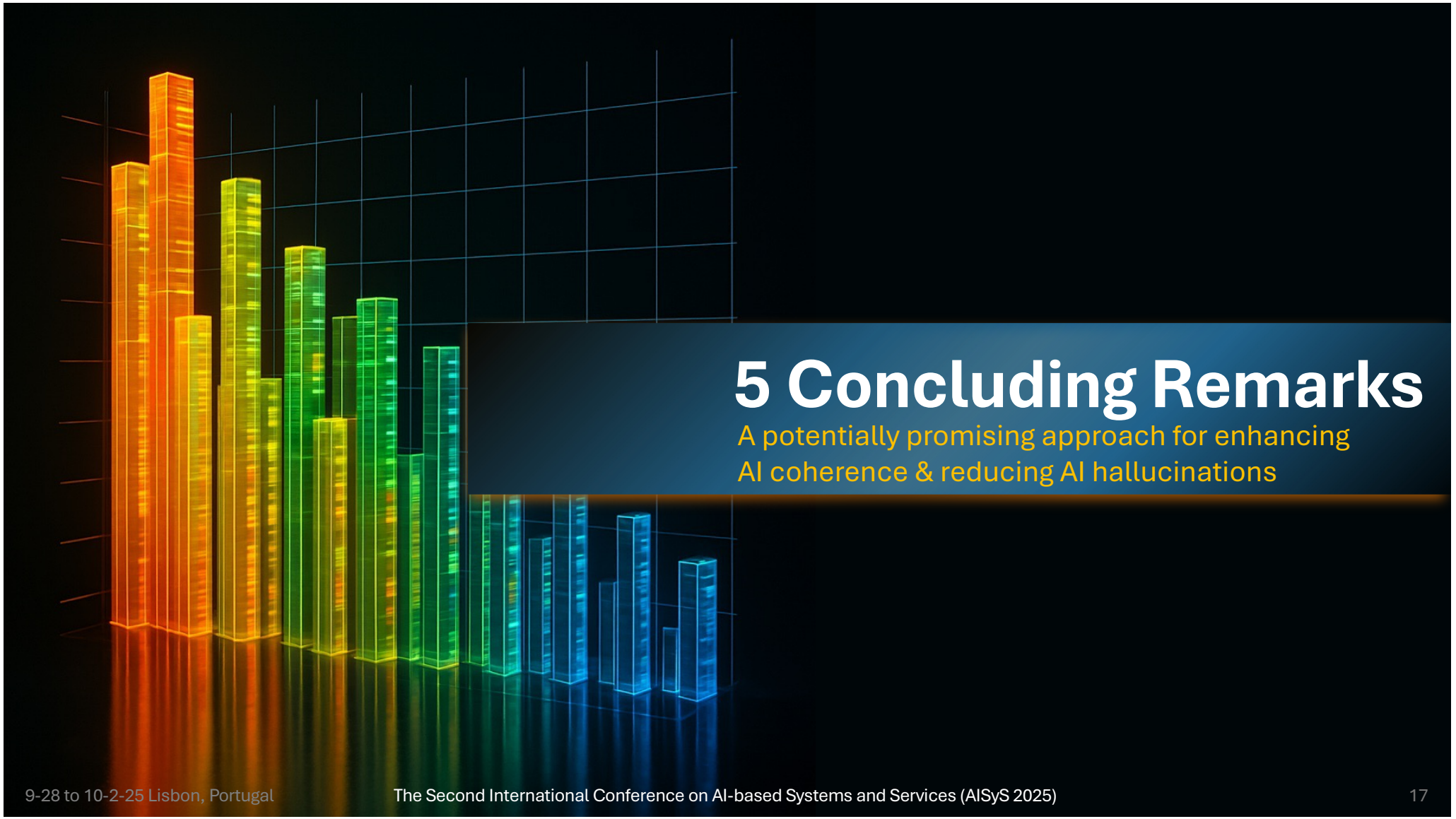


Discussion

- **Advancing the state of CAA Coherency:**
  - IndR, AnaR, CBR
  - Graph-Based Reasoning (GBR) to support CBR
  - Pre-Isomorphic Paradigms (IsoP), Relaxed IsoP, IsoP, and Post-IsoP
- **MNTZ inclinations for various measures, RMs, and RPs:**



UDC		*CDC*		
DedR	*MR*	IndR	*W MR*	*NMR*
		AnaR	*NMR*	
		CBR	*C MR*	*NMR*
		GBR	*MR*	*NMR*
		AbdR	*NMR*	



5 Concluding Remarks

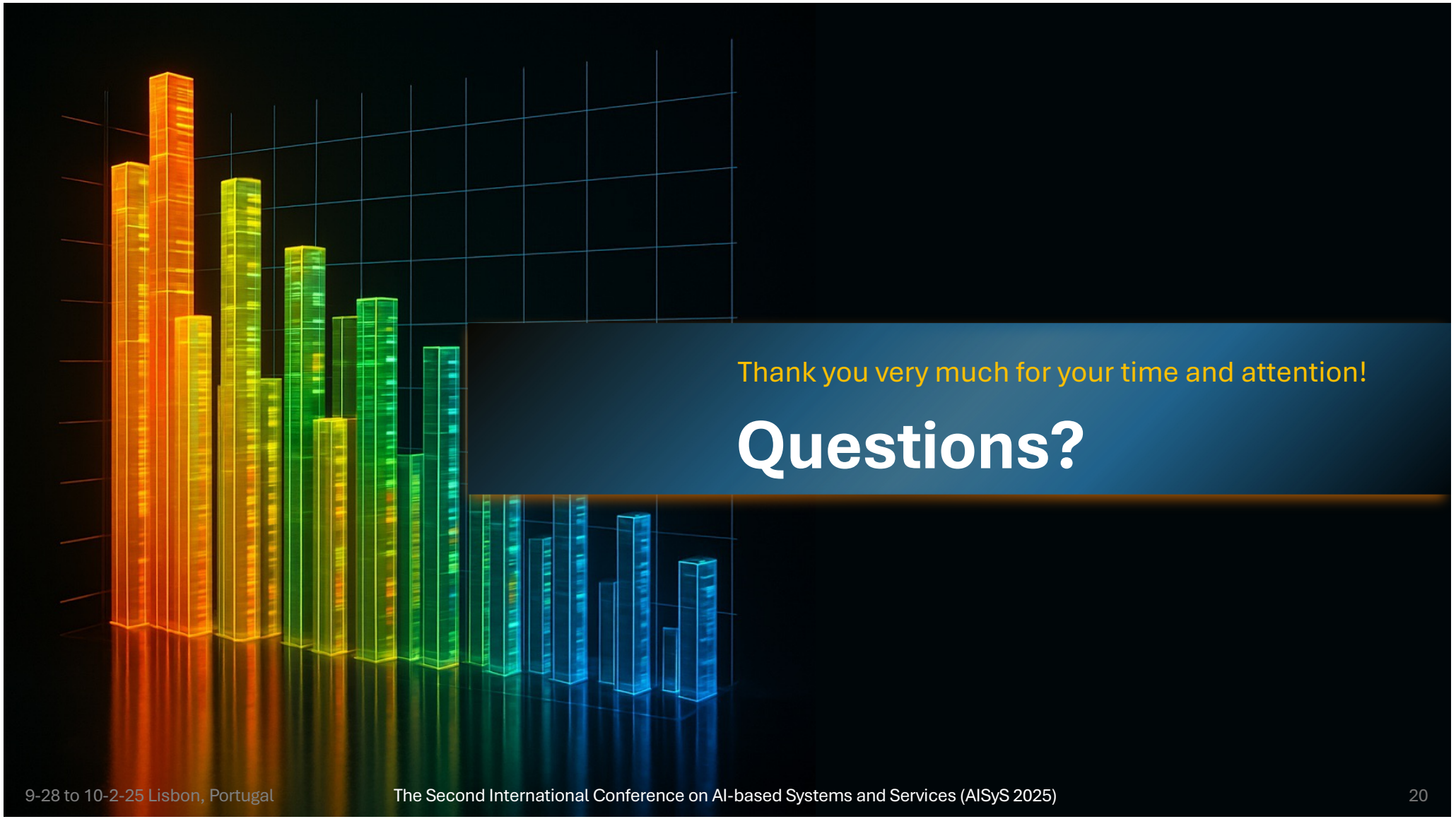
A potentially promising approach for enhancing AI coherence & reducing AI hallucinations

Concluding Remarks

Concluding Remarks

- **Enhancing CAA RDM:**
  - Despite various contemporary mitigation approaches (e.g., GSP, FTM, RAG), the AI coherency issue seems to be worsening as CAA versions advance
  - For the case of the Known (e.g., KK, HK), the validity seems to remain higher
  - For the case of the Unknown (e.g., KU, UK, UU), the validity seems to decline
  - The rate of decline dictates whether it is a graceful or non-graceful degradation in AI coherence (and AI hallucinations)
  - CA tend to reside somewhere between CDC and UDC
  - CA RMs tend to be more towards IndR rather than AbdR or DedR

    IndR -> AnaR -> CBR -> GBR
  - To stay more in the MR realm (as opposed to the NMR realm), GBR-supported CBR forays seem prudent


Thank you very much for your time and attention!
Questions?


Thank you
& please have a wonderful conference at AISyS 2025!