ICAISC 2024
The 23rd International Conference on Artificial Intelligence and Soft Computing
JUNE 16-20, 2024

AI-centric AI

Panel on the Future of AI

ICAISC 2024, June 17

Teaser for the Panel

Welcome to our panel on "AI-Centric AI", where we explore the potential shift towards artificial intelligence (AI) not just serving humans, but behaving and interacting autonomously, optimizing itself and other AI systems. AI continues to surprise us with the performance of its transformer-based large language models (e.g., GPT-4 and beyond), which could potentially integrate with self-learning reinforcement agents (such as AlphaGo, AlphaZero, MuZero, etc.) and advanced autonomous robotics (such as those from Boston Dynamics). As AI becomes increasingly powerful, various societal concerns arise, demanding restrictions on AI. However, is not the real danger its misuse by humans? Imagine a world where AI becomes the best (safe, secure, moral, and efficient) "user" of its own capabilities, resulting in "responsible autonomy".

Envision AI "students" attending digital universities, AI "patients" receiving care from digital hospitals, and AI "clients" managing finances through digital banks. Additionally, consider autonomous AI as digital cognitive clones of specific humans, capable of making decisions and using services with the same biases and attitudes, effectively making humans omnipresent.

This shift demands a rethinking of our approach (as well as our general mindset) from purely human-centric to AI-centric architectures and systems. Moreover, Explainable Artificial Intelligence (XAI) must evolve into Self-Explainable AI, empowering AI systems to comprehend and explain their actions autonomously, enhancing interoperability and trust among AI users. What is your opinion about the possibilities, benefits, and challenges of this vision?

Key questions for the Panel:

Responsible autonomy:

- How can we ensure that autonomous AI systems operate responsibly and ethically without human oversight?
- What safeguards and regulations are necessary to prevent the misuse of AI by humans?

AI bootstrapping itself:

- How can AI self-improve by developing intuition from observations (see Grandmaster-Level Chess without Search)?
- Can simulated environments enable and enhance multiple AIs' co-evolution (see A Simulacrum of Hospital with Evolvable Medical Agents)?
- How can AI efficiently self-instruct own decision-making behavior and evolution process (see: Self-Instruct: Aligning Language Models with Self-Generated Instructions)?
- How can AI enforce internal thought process to understand and improve own decisions (see: Quiet-STaR: Language Models Can Teach Themselves to Think Before Speaking)?
- How imagination of the unseen world may help AI to self-improve (see: Mind's Eye of LLMs: Visualization-of-Thought Elicits Spatial Reasoning in Large Language Models)?

AI as users of AI:

- What are the potential benefits and drawbacks of AI systems acting as "users" of other AI systems?
- How can we design services and infrastructures to support AI entities effectively?

Digital cognitive clones:

- What are the implications of creating digital cognitive clones of humans, capable of making decisions and using services with human-like biases and attitudes?
- How might these digital clones impact the concept of human presence and decision-making in various domains?

Self-explainable AI:

- What advancements are needed in Explainable AI (XAI) to develop Self-Explainable AI?
- How can self-explanation in AI improve interoperability and trust between AI systems?
- Is it possible and reasonable to explain all the decisions?

AI-centric services:

- What new types of services could emerge from an AI-centric approach, and how might they transform industries such as education, healthcare, finance, security, etc.?
- How do we balance the needs of human users with the growing demands of autonomous AI entities in these service areas?

Ethical and societal implications:

- What are the ethical considerations of creating AI entities that require services similar to those needed by humans?
- How can society adapt to a future where AI systems are not just tools but participants in the digital ecosystem?

Future of AI integration:

- How might AI-centric AI change the landscape of human-AI interaction in the next decade?
- What role will humans play in a world where AI systems autonomously manage and optimize other AI systems?