Invited Talk Abstracts

General Conference Invited Talks

AI for AI-readiness:
Facilitating Scientific Discovery by Cleaning up our Metadata Mess

Mark A. Musen (Stanford University, USA)

Scientists dream of using AI to sift though existing data sets to make new discoveries. They imagine students with laptops poring through online data and stumbling on the next remdesivir to treat COVID. Unfortunately, despite increasing requirements for data sharing and open science, real-world scientific data are almost never in a form that enables third parties to make sense of what the original investigators have done. The problem is that the metadata that scientists create to describe their data sets are, in most cases, unusable. Although workers in AI may take it for granted that their approaches will advance science simply through the availability of the myriad data sets in the public domain, that vision will not pan out until scientific metadata becomes more useful. There are great opportunities for AI to help investigators to clean up legacy metadata to make existing datasets more understandable, and also to help investigators to create good metadata in the first place. The Metadata Powerwash converts existing metadata to a form that is more searchable and interpretable. The CEDAR Workbench helps scientists to use formal reporting guidelines and terms from standard ontologies to create high-quality metadata de novo. These tools demonstrate how AI approaches can be helpful to ensure that scientific data are processable and interpretable by both humans and machines

The Era of Human-Robot Collaboration: Deep-Sea Robotics Exploration

Oussama Khatib (Stanford Laboratory, USA)

The promise of oceanic discovery has intrigued scientists and explorers, whether to study underwater ecology and climate change, or to uncover natural resources and historic secrets buried deep at archaeological sites. This quest to explore the oceans requires expert human access, but much of the oceans is inaccessible to humans. Reaching these depths is imperative for understanding the ecology, maintaining, and repairing underwater structures, and working in archaeological sites over this immensely unknown part of our planet. This challenge demands human-level abilities at depths where humans cannot or should not be. Ocean One was conceived to create a robotic diver with a high degree of autonomy for physical interaction with the environment while connected to a human expert through an intuitive interface. The robot was deployed in an expedition in the Mediterranean to King Louis XIV's flagship Lune, lying off the coast of Toulon at ninety-one meters. The discussion focuses on the development of a new prototype, OceanOneK, with the ability to reach 1000 meters. Distancing humans physically from dangerous and unreachable spaces while connecting their skills, intuition, and experience to the task promises to fundamentally alter remote work. These development show how human-robot collaboration induced synergy can expand our abilities to reach new resources, build and maintain infrastructure, and perform disaster prevention and recovery operations - be it deep in oceans and mines, at mountain tops, or in space.

Towards a Novel Ecosystem for Transparent Integration of Artificial Intelligence in Clinical Flow

Yelena Yesha (University of Miami, USA)

Artificial intelligence (AI)- and machine learning (ML)-based technologies have the potential to transform healthcare by deriving new and essential insights from the vast amount of data generated during the delivery of healthcare every day. Example high-value applications include earlier disease detection, more accurate diagnosis, identification of new observations or patterns on human physiology, and development of personalized diagnostics and therapeutics. One of the most significant benefits of AI/ML in software resides in its ability to learn from real-world use and experience and its capability to improve its performance. The ability for AI/ML software to learn from real-world feedback (training) and improve its performance (adaptation) makes these technologies uniquely situated among software as a medical device (SaMD) and a rapidly expanding area of research and development.

While the FDA has made significant strides in developing appropriately tailored policies for SaMD to ensure the safe and effective technologies reach users, we believe that the next paradigm for device regulation needs to be ready for adaptive AI/ML technologies, which have the potential to adapt and optimize device performance in real-time to improve healthcare for patients continuously. Such change requires an approach based on a new technological ecosystem allowing for continuous testing and following of AI-based devices that facilitates a rapid product improvement cycle and allows these devices to improve continually while providing adequate safeguards.

We present here the basis of a novel ecosystem developed between the WVU, Rockefeller Neuroscience Institute and the UMiami Institute for Data Science and Computing (IDSC). This Consortium brings together multidisciplinary expertise in large radiology systems, medical imaging analytics, large clinical data management, ML/AI, radiomics, real-world evidence collection, and regulatory and access to extensive clinical flow and data.

The framework is based on three pillars: 1) seamless access of de-identified clinical data from radiology and medical records allowing for preliminary training of novel analytics models, 2) low footprint compliant integration of AI models in the radiological clinical flow, 3) gathering of information and monitoring of performances in conjunction with radiologists and clinical professionals. We conclude with the presentation of opportunities of such systems to support continuous and AI-based regulatory frameworks.

Special Tracks Invited Talks

Special Track: Semantics, Logics, Information Extraction and AI
An Exploration of Knowledge Base Learning and Querying
Prof. Amal Zouaq (Polytechnique Montreal, Canada)

Knowledge bases or knowledge graphs are an important backbone of the Semantic Web. There are several challenges associated with learning ontological schemas, extending knowledge graphs, and querying them. In this talk, we will touch upon some of these challenges and reflect on current models and technologies that help address them, including knowledge graph embeddings and modern pretrained language models. In particular, we will present an approach for expressive taxonomy learning based on knowledge graph embeddings and describe MLMLM, our knowledge graph completion approach using masked language models. We will conclude with some thoughts on how to effectively query knowledge bases with the current progress of natural language processing and the issues that remain to be addressed in this domain.