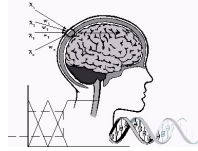




International

Innovation in Knowledge Based and Intelligent
Engineering Systems



INVITED SESSION SUMMARY

Title of Session:

***Information Disorder Challenges in the Era of AI-Generated Contents Manipulation:
Security, Trustability and Faithfulness of Shared Knowledge***

Name, Title and Affiliation of Chairs:

Fiammetta Marulli, Prof., fiammetta.marulli@unicampania.it
Gennaro Ragucci, PhD candidate, gragucci@unisa.it
Clemente Galdi, Prof., cgaldi@unisa.it
Luigi Catuogno, Prof., luigi.catuogno@uniparthenope.it

Details of Session:

The lexicon surrounding distorted information has shifted from "**fake contents**" (fake news, most of all) to "**information disorder**" a more comprehensive term, that encompasses three distinct categories:

- **Misinformation, characterized by the inadvertent dissemination of falsehoods;**
- **Disinformation, defined by the deliberate spreading of false information with malicious intent;**
- **Malinformation, the disclosure of authentic information to inflict harm, often involving the exposure of private data.**

"Fake news" can be placed within either the misinformation or disinformation categories. Propaganda, a related concept, denotes the strategic dissemination of information, regardless of veracity, to influence an audience, frequently with a political agenda.

The World Health Organization's introduction of the term "infodemic" underscores the critical implications of information disorder, particularly during public health emergencies, where an overabundance of information can undermine effective response efforts.

In this so challenging scenario, the rise of AI, along with the Generative Models have led to a surge in manipulated media, making it harder to know and recognize what's real and where is the truth.

This Invited Session is designed with the aim to address this challenge by bringing together AI experts. The goal is to create a playground for discussing about how to develop technological solutions, using AI, but not limited to it, how to identify and counter fake content and protect the integrity of digital contents, and evaluate the negative effects on several aspects of real human life, including manipulation of social behaviours and information literacy levels of current and next people generations. This session seeks research contributions that leverage AI and other technological patterns, as Federated Learning, Blockchains and Edge and Cloud Computing to assess media authenticity in a world where misinformation is rampant and proposals about metrics and evaluation methods of what is called with the comprehensive term of **Information Disorder**.

Main Topics for the session are:

- Deepfakes: creation, detection, and mitigation technologies AI-driven content manipulation detection and prevention systems Machine learning models for detecting AI-generated misinformation
- Advanced natural language processing (NLP) techniques for detecting fake news and misinformation
- Sentiment analysis for detecting fake news and deceptive content
- Multimodal AI approaches for detecting manipulated media, text, images, and video Cross-modal detection of manipulated media
- Real-Time Detection of Generative AI Manipulated Media
- Explainable AI methods for transparency in media verification processes Adversarial AI: attack vectors, security challenges, and defense mechanisms Generative Adversarial Networks (GANs) for Media Creation and Security AI-based content verification, validation, fact-checking, and authentication tools Blockchain technology for securing digital media authenticity
- Federated learning for decentralized detection of manipulated content
- Edge computing solutions for distributed AI-based content validation and media integrity Computer vision methods for identifying synthetic and manipulated visuals
- AI-based image and video forensics for media authenticity verification

Misinformation, Disinformation and Information Disorder Evaluation.

Further Related Topics to this session:

- **Sentiment Analysis and Opinion Mining:** Developing advanced deep learning models to accurately identify and classify sentiment expressed in social media text.
- **Topic Modeling and Trend Detection:** Employing deep learning techniques to discover emerging trends, identify relevant topics, and track their evolution over time. **User Behavior Analysis:** Leveraging deep learning to analyze user interactions, preferences, and social network structures to understand user behavior patterns. **Event Detection and Tracking:** Developing deep learning models to detect and track real-world events as they unfold on social media platforms.
- **Fake News Detection and Misinformation Mitigation:** Utilizing deep learning to identify and mitigate the spread of false information and misinformation.
- **Social Network Analysis:** Applying deep learning to analyze complex social networks and extract valuable insights from their structure and dynamics.
- **Controversial Content Mitigation:** modeling approaches (deep learning, statistical modeling etc.) and tools to identify and mitigate the spread of false information and misinformation. .

We encourage submissions that present novel deep learning architectures, innovative methodologies, and practical applications for social media mining. By fostering collaboration and knowledge sharing, this special session aims to advance the state-of-the-art in this exciting and rapidly evolving field.

Main Contributing Researchers and Research Centres:

University of Campania "Luigi Vanvitelli", Dept. of Maths and Physics, Italy
University of Salerno, Dept. of Political and Social Studies, Italy
University of Napoli Parthenope, Dept. of Economics, Law, Computer Science, and Sports Science, Italy
University of Molise, Dept. of Engineering, Italy
University of St.Polten, Fachshole St.Polten, Austria
University of Paris, Sorbonne Paris Nord (XIII), Paris, France
University of Foggia, Dept. of Applied Science, Italy
University of Zaragoza, Computer Science and Engineering Dept, Spain

Email & Contact Details:

Prof. Fiammetta Marulli, fiammetta.marulli@unicampania.it