

SAFEPROCESS 2027, Delft, June 29-July 2, 2027

<https://conferences.ifac-control.org/safeprocess2027/>

Invited Session: “Anomaly Detection and Explanation”

Organized by the AI-Enhanced Safeprocess Working Group
of the IFAC TC 6.4 Fault Detection, Supervision & Safety of Technical Processes-SAFEPROCESS

Organizers: Elodie Chanthery, LAAS-CNRS & ANITI, INSA Toulouse (elodie.chanthery@laas.fr)
Louise Travé-Massuyès, LAAS-CNRS & ANITI, Toulouse University (louise@laas.fr)

*Please send us your intention to contribute to this session before **September 14, 2026** with the following information: title, abstract, name of authors.*

*Papers must then be submitted following the standard procedure no later than **October 31, 2026** on PaperCept with the code of the session (that will be provided by the IFAC WC organizers after the acceptance of the session).*

Anomaly detection plays a central role in fault detection, supervision, and safety of socio-technical systems. Recent advances in machine learning and artificial intelligence have significantly improved the capability to detect abnormal behaviors, rare events, incipient faults, and previously unseen operating conditions. Anomaly detection plays a crucial role in certifying data as well as in identifying symptoms that can be used to drive diagnosis reasoning and health management.

Now, achieving high detection performance alone is not sufficient in many applications. Yet, pure machine learning methods rely on black box models that are opaque and do not justify their decisions. In this context, explainability has emerged as a key research challenge, implying a dual concept: the algorithm’s capacity to justify its outputs, i.e., why a specific result was produced, and the extent to which humans can comprehend those results within a specific task context and safely use them (sometimes referred to as interpretability). Explaining anomalies may thus reduce to identifying features that cause out-of-distribution behavior, whereas it is equally important to uncover their underlying business causes. In the latter case, diagnoses can shed light on *what* happened, and explanations should precisely clarify *why* it occurred.

Addressing explainable anomaly detection requires the integration of advances in explainable AI (XAI), causal reasoning, diagnosis, knowledge representation, digital twins, and hybrid model-based/data-driven approaches.

The objective of this invited session is to bring together researchers and practitioners working on anomaly detection, diagnosis, XAI, and trustworthy monitoring systems. The session aims to foster exchanges between communities developing advanced detection algorithms and those focusing on explanation, interpretation, causal analysis, and human-centered decision support.

Technical contributions, methodological advances, benchmark studies, surveys, and industrial applications are welcome.

A non-exhaustive list of topics is provided below:

- Explainable and interpretable anomaly detection,
- Root-cause analysis and fault localization,
- Causal reasoning for anomaly explanation,
- Counterfactual explanations for anomalies and faults,
- Hybrid model-based and data-driven anomaly detection,
- Physics-informed and knowledge-guided anomaly detection and diagnosis,
- Temporal, sequential, and event-based anomaly explanation,
- Graph-based anomaly detection and explanation,
- Foundation models and large language models for anomaly analysis and explanation,
- Evaluation metrics for explainability and interpretability,
- Trustworthy, transparent, and accountable monitoring systems.