



# 2024 IEEE CSR Workshop on Synthetic Data Generation for a Cyber-Physical World (SDG)

## Call for Papers

### Important dates

Paper submission deadline:

**June 3, 2024 AoE**

Authors' notification:

**July 3, 2024 AoE**

Camera-ready submission:

**July 14, 2024 AoE**

Early registration deadline:

**July 20, 2024 AoE**

Workshop dates:

**September 2–4, 2024**

### Workshop chairs

Samira Maghool (IT)

Ernesto Damiani (AE)

Faiza Allah Bukhsh (NL)

### Technical program committee

Juba Agoun (FR)

Jeewanee J. Arachchige (NL)

Rob Bemthuis (NL)

Nicola Bena (IT)

Filippo Berto (IT)

Paolo Caravolo (IT)

Elena Casiraghi (IT)

Marco Cremonini (IT)

Maya Daneva (NL)

MohammadReza Fani Sani (DE)

Sanja Lazarova-Molnar (DE)

Afshin Montakhab (IR)

Azzam Mourad (LB)

Ehsan Ullah Munir (PK)

Anastasija Nikiforova (EE)

Mirela Riveni (NL)

### Contact us

[samira.maghool@unimi.it](mailto:samira.maghool@unimi.it)

[ernesto.damiani@ku.ac.ae](mailto:ernesto.damiani@ku.ac.ae)

[f.a.bukhsh@utwente.nl](mailto:f.a.bukhsh@utwente.nl)

Synthetic datasets that reflect the statistical properties of authentic data allow to share research insights and findings without compromising privacy or proprietary interests. This approach not only promotes transparency and reproducibility in research but also encourages interdisciplinary collaboration and knowledge sharing. Artificial Intelligence (AI) is one of the main areas utilizing generated synthetic data. Privacy issues arise once the dataset contains sensitive features playing a role in training AI systems. Due to the fact that data collection is expensive and time-consuming, given some shortcomings such as low volume of data, non-compliance with regulations, and bias, we not only may achieve biased and low-performance models but also violate privacy principles. Synthetic data generation can facilitate analysis, the need for data augmentation, or prevent data breaches in highly sensitive domains, rather than weak anonymization approaches. Generative Adversarial Networks (GAN), Variational Autoencoders (VAE), and Agent-based modeling (ABM) are among the most common synthetic data generation algorithms.

However, it is critical to recognize the limitations of synthetic data generation, particularly in capturing the intricacies and interdependencies present in real-world systems. While synthetic datasets can mimic statistical distributions and patterns, they may struggle to replicate the nuanced relationships and contextual nuances inherent in complex phenomena. By leveraging advances in artificial intelligence, machine learning, and computational modeling, researchers can strive to bridge the gap between synthetic and authentic data, unlocking new opportunities for insight and innovation in fields as diverse as healthcare, finance, social sciences, and beyond.

The workshop is held in conjunction with the IEEE CSR 2024 conference as a **hybrid event**, during September 2–4, 2024. Prospective authors are encouraged to submit previously unpublished contributions from a broad range of topics, which include but are not limited to the following:

- › Generating synthetic data compliance with regulations
- › Privacy preserving in healthcare data
- › Algorithms for debiasing dataset (in the preprocessing phase of ML modeling)
- › Algorithms for debiasing the ML models' results
- › Uncovering and mitigating synthetic data algorithmic bias
- › Assurance and certification of the dataset and ML models
- › Synergy of ABM with ML focusing on the rule extraction
- › Domain dependent/independent synthetic data generation challenges and opportunities
- › FAIR (findability, accessibility, interoperability, and reuse) and ethical synthetic data generation
- › Explainability and interpretability aspects in synthetic data generation

The CSR SDG workshop will accept high-quality research papers presenting strong theoretical contributions, applied research and innovation results obtained from funded cyber-security and resilience projects, and industrial papers that promote contributions on technology development and contemporary implementations.

Submitted manuscripts should not exceed 6 pages (plus 2 extra pages, being subject to overlength page charges) and should be of sufficient detail to be evaluated by expert reviewers in the field. Accepted papers will be submitted for inclusion into IEEE Xplore subject to meeting IEEE Xplore's scope and quality requirements. Detailed information about paper submission and guidelines for authors can be found at the workshop's website <https://www.ieee-csr.org/sdg>.