

# Data Privacy in Big Data Context

## Teacher

Domenico Desiato

## Course Website

Code: To be defined

## Course description (min 150, max 300 words)

Big data collections are available thanks to digitalizing business processes and public administrations. Data privacy needs to be guaranteed when data refers to individuals by permitting the application of big data processes without jeopardizing their privacy. In particular, analytical processes or the application of machine learning techniques to privatized data might still lead to the disclosure of sensitive and confidential information about data subjects, thanks to the power of current predictive models. To this end, it is necessary to enable big data processes to extract valuable knowledge and insights from data while avoiding disclosing sensitive information. Moreover, the trade-off between data quality and privacy needs to be managed to understand the relationship between privacy guarantees and usable knowledge.

The course will cover the main topics concerning privacy-preserving big data and discuss innovative research proposals in such a field. In particular, different application scenarios, such as privacy-preserving data analytics, data preparation, machine learning and social network domain, will be analyzed. Furthermore, the course will discuss privacy models and methodologies for privatizing data. Additionally, the theoretical lectures will be supported by practical applications of algorithms for privatizing data. Finally, the course will provide insights for evaluating the trade-off between data utility and privacy.

## Course period

December 2023 to February 2024

## SSD

INF/01

## Course References (optional)

[1] Das, P.K., et al. "Privacy and Security Issues in Big Data: An Analytical View on Business Intelligence." Springer, 2021, <https://books.google.es/books?id=nworEAAAQBAJ>.

[2] Li, J. et al. "Privacy-Preserving Machine Learning." Springer, 2022, <https://books.google.es/books?id=KhpkEAAAQBAJ>.

## Credits and Hours

3 CFU, 2 of lecture (16 Hours) and 1 of practice (15 hours), for a total of 31 hours.

**Exam Modality**

Paper presentation. Students present the content of one paper suggested by the teacher. No groups are allowed.

**Teacher(s) CV**

Please, see attachment.

**Teacher(s) Main Publications**

Please, see attachment.

# DOMENICO DESIATO

Department of Computer Science, University of Bari Aldo Moro, Via E. Orabona N. 4, 70125, Bari - Italy

✉ [domenico.desiato@uniba.it](mailto:domenico.desiato@uniba.it)

## Education

---

### University of Salerno, Italy

*Bachelor degree in computer science*

### University of Salerno, Italy

*Master degree in computer science*

### University of Salerno, Italy

*PhD in computer science and information engineering*

## Professional experience

---

### Research fellow at Department of Informatics (University of Salerno)

**July 2021 – June 2022**

*Fisciano (SA), Italy*

### Research fellow at Department of Informatics (University of Salerno)

**July 2022 – February 2023**

*Fisciano (SA), Italy*

## Current position

---

### Research at Department of Informatics (University of Bari Aldo Moro)

**February 2023 – Today**

*Bari (BA), Italy*

*Project: Next Generation EU (PE0000014 - "Security and Rights In the CyberSpace - SERICS" - CUP: H93C22000620001)*

## External collaborations

---

### Datonix Spa

*Development of Applications Privacy Preserving*

**April 2019 – March 2020**

*Avella (AV), Italy*

### Eindhoven University of Technology

*Research activity in the field of Privacy Preserving Machine Learning*

**October 2020 – April 2021**

*Eindhoven, Holland*

## Teaching assistance activities

---

### Course of Databases (Bachelor degree)

*Lecturer from September 2021 to September 2022 at Department of Informatics (University of Salerno)*

### Course of IoT Data Analytics (Master degree)

*Support to laboratory experimental sessions and evaluation of projects at Department of Informatics (University of Salerno)*

### Course on Fundamentals of Data Science and Machine Learning (Master degree)

*Support to laboratory experimental sessions and evaluation of projects at Department of Informatics (University of Salerno)*

## Professional services

---

### Reviewer of international journals

*Multimedia Tools and Application (Springer), IEEE Access (IEEE), Computer & Security (Elsevier), Big Data Research (Elsevier), Information Sciences (Elsevier)*

### Reviewer of ACM journals

*ACM Computing Surveys*

### Program committees

*DMS Conference on Visualization and Visual Languages*

## Publications

---

- Loredana Caruccio, Domenico Desiato, Giuseppe Polese, Genoveffa Tortora. GDPR compliant information confidentiality preservation in Big Data processing, IEEE Access, vol. 8, pp. 205034-205050, 2020.
- Domenico Desiato. A Methodology for GDPR Compliant Data Processing, in proceedings of the 26th Italian Symposium on Advanced Database System, Castellaneta Marina (Taranto), Italy, June 24-27, 2018.
- Loredana Caruccio, Domenico Desiato, Giuseppe Polese. Fake Account Identification in Social Networks, in 2018 IEEE International Conference on Big Data (big data), pp. 5078-5085, 2018.
- Stefano Cirillo, Domenico Desiato, Bernardo Breve. CHRAVAT – Chronology Awareness Visual Analytic Tool, in proceedings of the 23rd International Conference on Information Visualization (IV), Paris, France, July 02-05, 2019.
- Bernardo Breve, Loredana Caruccio, Stefano Cirillo, Domenico Desiato, Vincenzo Deufemia, Giuseppe Polese. Enhancing user awareness during internet browsing, in proceedings of the 4th Italian Conference on Cyber Security (ITASEC 2020): pp. 71-81.
- Bernardo Breve, Stefano Cirillo, Mario Cuofano, Domenico Desiato. Perceiving Space Through sound: mapping human movements into MIDI, accepted in the 26th International DMS Conference on Visualization Languages (DMSVIVA20), pp. 49-56, 2020.
- Bernardo Breve, Stefano Cirillo, Mario Cuofano, Domenico Desiato. Enhancing spatial perception through sound: mapping human movements into MIDI, Multimedia Tools and Applications, vol. 81, pp. 73-94, 2022.
- Giuseppe De Gregorio, Domenico Desiato, Angelo Marcelli, Giuseppe Polese. A Multi Classifier Approach for Supporting Alzheimer’s Diagnosis Based on Handwriting Analysis, International Conference on Pattern Recognition, pp. 559-574, 2021.
- Francesca Cerruto, Stefano Cirillo, Domenico Desiato, Simone Michele Gambardella, Giuseppe Polese. Social network data analysis to highlight privacy threats in sharing data, Journal of Big Data, vol. 9, pp. 1-26, 2022.
- Francesca Cerruto, Stefano Cirillo, Domenico Desiato, Simone Michele Gambardella, Giuseppe Polese. Cross-social network investigation to highlight privacy violations in data sharing activities, in proceedings of the 30th Italian Symposium on Advanced Database System, Tirrenia (Pisa), Italy, June 19-22, 2022.
- Loredana Caruccio, Domenico Desiato, Giuseppe Polese, Genoveffa Tortora, Nicola Zannone. A decision-support framework for data anonymization with application to machine learning processes, Information Sciences, vol. 613, pp. 1-32, 2022.
- Stefano Cirillo, Domenico Desiato, Michele Scalera, Giandomenico Solimando. A Visual Privacy Tool to Help Users in Preserving Social Network Data, Joint Proceedings of the Workshops, Work in Progress Demos and Doctoral Consortium at the IS-EUD 2023 co-located with the 9th International Symposium on End-User Development (IS-EUD 2023), Cagliari, Italy, June 6-8, 2023, vol.3408, 2023.
- Loredana Caruccio, Domenico Desiato, Giuseppe Polese, Genoveffa Tortora, Nicola Zannone. An Approach to Trade-off Privacy and Classification Accuracy in Machine Learning Processes, Proceedings of the 31st Symposium of Advanced Database Systems, Galzingano Terme, Italy, July 2nd to 5th, 2023, vol. 3478, pp. 420-429, 2023.
- Loredana Caruccio, Gaetano Cimino, Stefano Cirillo, Domenico Desiato, Giuseppe Polese, Genoveffa Tortora. Malicious Account Identification in Social Network Platforms, ACM Transactions on Internet Technology, 2023.

14/11/2023

Date



Signature

## List of Publications

---

1. Loredana Caruccio, Domenico Desiato, Giuseppe Polese, Genoveffa Tortora, Nicola Zannone. A decision-support framework for data anonymization with application to machine learning processes, Information Sciences, vol. 613, pp. 1-32, 2022.
2. Francesca Cerruto, Stefano Cirillo, Domenico Desiato, Simone Michele Gambardella, Giuseppe Polese. Social network data analysis to highlight privacy threats in sharing data, Journal of Big Data, vol. 9, pp. 1-26, 2022.
3. Loredana Caruccio, Domenico Desiato, Giuseppe Polese, Genoveffa Tortora. GDPR compliant information confidentiality preservation in Big Data processing, IEEE Access, vol. 8, pp. 205034-205050, 2020.
4. Loredana Caruccio, Gaetano Cimino, Stefano Cirillo, Domenico Desiato, Giuseppe Polese, Genoveffa Tortora. Malicious Account Identification in Social Network Platforms, ACM Transactions on Internet Technology, 2023.
5. Stefano Cirillo, Domenico Desiato, Michele Scalera, Giandomenico Solimando. A Visual Privacy Tool to Help Users in Preserving Social Network Data, Joint Proceedings of the Workshops, Work in Progress Demos and Doctoral Consortium at the IS-EUD 2023 co-located with the 9th International Symposium on End-User Development (IS-EUD 2023), Cagliari, Italy, June 6-8, 2023, vol.3408, 2023.
6. Loredana Caruccio, Domenico Desiato, Giuseppe Polese, Genoveffa Tortora, Nicola Zannone. An Approach to Trade-off Privacy and Classification Accuracy in Machine Learning Processes, Proceedings of the 31st Symposium of Advanced Database Systems, Galzingano Terme, Italy, July 2nd to 5th, 2023, vol. 3478, pp. 420-429, 2023.
7. Loredana Caruccio, Domenico Desiato, Giuseppe Polese. Fake Account Identification in Social Networks, in 2018 IEEE International Conference on Big Data (big data), pp. 5078-5085, 2018.
8. Stefano Cirillo, Domenico Desiato, Bernardo Breve. CHRAVAT – Chronology Awareness Visual Analytic Tool, in proceedings of the 23rd International Conference on Information Visualization (IV), Paris, France, July 02-05, 2019.
9. Bernardo Breve, Loredana Caruccio, Stefano Cirillo, Domenico Desiato, Vincenzo Deufemia, Giuseppe Polese. Enhancing user awareness during internet browsing, in proceedings of the 4th Italian Conference on Cyber Security (ITASEC 2020): pp. 71-81.
10. Domenico Desiato, Genoveffa Tortora. A Methodology for GDPR Compliant Data Processing, in proceedings of the 26th Italian Symposium on Advanced Database System, Castellaneta Marina (Taranto), Italy, June 24-27, 2018.

14/11/2023

Date



Signature