RESPONSIBLE DATA SCIENCE FOR SOCIETY

MODELS, ALGORITHMS, TRUSTWORTHY AI

Data Science and AI play an increasingly important role in our daily life. AI, with its applications, are essential tools for the ethical and responsible progress of our multicultural and interconnected society. The school for "Responsible Data Science for Society: Models, Algorithms, Trustworthy AI" introduces participants to selected topics to better understand the complexity of our world from the data scientist's perspective

TIME TABLE

Time		Monday	Tuesday	Wednesday	Thursday	Friday
08:30	09:30	Registration				
09:30	10:00	SoBigData RI	Personal data capture and platform monopolisation	Social AI	Exploratories Research Hihglights	Open Lab (Meet Tutors)
10:00	11:00	Responsible Data Scientists				
11:00	11:30	Coffee break				
11:30	12:30	Students Research Presentation	Privacy risks and harms	Explainable AI	Open Lab (Meet Tutors)	Students Final Presentation
12:30	13:00					
13:00	14:30	Lunch break				
14:30	15:00	Social Tour Salina	Privacy Preserving Techniques and Approaches (+ Hands on sessions)	Explainability techniques (+ Hands on sessions)	Free time	
15:00	16:00					
16:00	16:30					
16:30	17:00		Coffee break		Social Tour Vulcano + Party on the Beach	
17:00	18:00		Privacy Preserving Techniques and Approaches (+ Hands on sessions)	Explainability techniques (+ Hands on sessions)		
18:00	20:00					
20:00	23:30					

EXTENDED PROGRAM

MONDAY

09:30 - 10:00 **The SoBigData R.I** | *Roberto Trasarti (CNR) SoBigData RI Coordinator*

10:00 - 11:00 **Responsible Data Scientists**

Responsible data science: what is it and why do we want it? | *Juan Manuel Duran (TuDelft)*

The use of algorithms for assisting (or replacing) scientists (for instance, AI, ML, DNN, and computer simulations) pose ethical challenges that must be addressed. One particularly pressing stems from the responsibility of researchers in applying and using algorithms. How can such responsibility be conceived? What can researchers do to be more responsible with their findings? Why do we need responsible data scientists? This lecture will address these and other questions emerging in the intersection between responsible science and the use of algorithms for scientific purposes.

11:00 - 11:30 **Coffee break**

11:30 - 13:00 **Students Research Presentation** | *Beatrice Rapisarda (CNR)*
13:00 - 14:30 **Lunch break**
14:30 - 20:00 **Social tour on Salina Island**

TUESDAY

09:30 - 11:00 **Personal data capture and platform monopolisation**

The Trillion-dollar platform in your pocket? | *Mark Cotè (KCL)*

This session will address personal data capture and platform monopolisation through a socio-technical overview of mobile devices. We will focus on a somewhat obscure technical object, the Software Development Kit (SDK), a folder of tools used by developers to make apps work and to make money. The average person has more than 40 different apps on their phone and each app uses an average of 18 SDKs which harvest, share, and process our data. Users have little access to or understanding of this core technical data hub or of how it is supercharging profits for tech giants like Google or Facebook. We will open up SDKs using socio-technical methods to demonstrate how digital giants are controlling our data.

11:00 - 11:30 **Coffee break**

11:30 - 13:00 **Privacy risks and harms**

Everything Everywhere All at Once: Privacy Engineering Challenges of Software Developers | *Awais Rashid (REPHRAIN)*

Software developers play a critical role in implementing privacy features in applications utilised in a wide range of business and personal settings. But what are the challenges that they face when incorporating such features? Developers have to deal with a multitude of requirements and constraints - ranging from the need to deliver core application functionality, complying with regulatory requirements such as GDPR, complexity of third party services and application programming interfaces and monetisation models such as use of Ad Networks. In this lecture, I will discuss insights gained from multiple studies with developers — distilling the challenges they face and how these impact their privacy design choices. I will also discuss the importance of more usable privacy engineering methods and tools for developers including the need for more systematic privacy threat modelling as both applications and the adversarial landscape changes.

13:00 - 14:30 **Lunch break**

14:30 - 16:00 **Privacy Preserving Techniques and Approaches (+ Hands on sessions)**

Privacy Risk Assessment and Vulnerabilities: from theory to practice | *Josep Domingo Ferrer (URV), Alberto Blanco Justicia (URV), Roberto Pellungrini (SNS), Francesca Pratesi (CNR)*

In this lecture we will give an overview of methodologies for privacy evaluation, from empirical evaluation to modern privacy models. We will discuss the strengths and weaknesses of different privacy models, such as k-anonymity and differential privacy. We will show how it is possible to put the theoretical knowledge of privacy into practice, by showing Python implementations of the PRUDEnce privacy risk assessment framework, membership inference attacks and backdoor attacks for federated learning systems.

16:00 - 17:00 **Coffee break**

17:00 - 18:00 **Privacy Preserving Techniques and Approaches (+ Hands on sessions)**

Privacy Risk Assessment and Vulnerabilities: from theory to practice | *Josep Domingo Ferrer (URV), Alberto Blanco Justicia (URV), Roberto Pellungrini (SNS), Francesca Pratesi (CNR)*

In this lecture we will give an overview of methodologies for privacy evaluation, from empirical evaluation to modern privacy models. We will discuss the strengths and weaknesses of different privacy models, such as k-anonymity and differential privacy. We will show how it is possible to put the theoretical knowledge of privacy into practice, by showing Python implementations of the

PRUDEnce privacy risk assessment framework, membership inference attacks and backdoor attacks for federated learning systems.

WEDNESDAY

09:30 - 11:00 Social AI

Social Artificial Intelligence | *Dino Pedreschi (UNIPI), Salvatore Rinzivillo (CNR)*

The rise of large-scale socio-technical systems in which humans interact with artificial intelligence (AI) systems (including assistants and recommenders, in short AIs) multiplies the opportunity for the emergence of collective phenomena and tipping points, with unexpected, possibly unintended, consequences. For example, navigation systems' suggestions may create chaos if too many drivers are directed on the same route, and personalised recommendations on social media may amplify polarisation, filter bubbles, and radicalisation. On the other hand, we may learn how to foster "wisdom of crowds" and collective action effects to face social and environmental challenges. In order to understand the impact of AI on socio-technical systems and design next-generation AIs that team with humans to help overcome societal problems rather than exacerbate them, we propose to build the foundations of Social AI at the intersection of Complex Systems, Network Science and AI. In this lecture we discuss the main open questions and initial results in Social AI, outlining possible technical and scientific challenges and suggesting research avenues.

11:00 - 11:30 Coffee break

11:30 - 13:00 Explainable AI

Explainable Machine Learning for Trustworthy AI | *Fosca Giannotti (SNS), Andrea Beretta (CNR), Anna Monreale (UNIPI)*

Black box AI systems for automated decision making, often based on machine learning over (big) data, map a user's features into a class or a score without exposing the reasons why. This is problematic not only for the lack of transparency, but also for possible biases inherited by the algorithms from human prejudices and collection artifacts hidden in the training data, which may lead to unfair or wrong decisions. The future of AI lies in enabling people to collaborate with machines to solve complex problems. Like any efficient collaboration, this requires good communication, trust, clarity and understanding. Explainable AI addresses such challenges and for years different AI communities have studied such topic, leading to different definitions, evaluation protocols, motivations, and results. This lecture provides a reasoned introduction to the work of Explainable AI (XAI) to date, and surveys the literature with a focus on machine learning and symbolic AI related approaches and on the achievements of the ERC XAI project (Science and technology for the explanation of AI decision making – G.A. 834756). We motivate the needs of XAI in real-world and large-scale application, while presenting state-of-the-art techniques and best practices, as well as discussing the many open challenges. The final part of the lecture will focus on the interplay of explainable AI and trust calibration to improve human decision-making and privacy. In particular, the lecture provides an overview on how local and global explainers might jeopardize privacy protection.

13:00 - 14:30 Lunch break

14:30 - 17:00 Explainability techniques (+ Hands on sessions)

Explainability Techniques for Tabular Data, Images, Time Series and Graphs | *Anna Monreale (UNIPI), Carlo Metta (CNR), Riccardo Guidotti (UNIPI), Giovanni Stilo (UNIAQ), Mario Alfonso Prado Romero (UNIAQ)*

The emerging explosion of XAI techniques has recently brought to their design and usage for different data types ranging from tabular data to graphs and passing from images and time series. In this lecture we will provide a focused overview of some explanation methods illustrating with practical examples how certain types of explanation can be achieved with state-of-the-art explanation methods and which are the pre-processing operation necessary for the different data types. After the break In the second part, we will discuss in detail the counterfactual explainability on graphs (GCE) while we will show, by three fundamental use cases, how to extend GRETEL, an extensible-by-design GCE development and evaluation framework that promotes Open Science and reproducibility.

16:00 - 17:00 **Coffee break**

17:00 - 18:00 **Explainability techniques (+ Hands on sessions)**

Explainability Techniques for Tabular Data, Images, Time Series and Graphs | *Anna Monreale (UNIPI), Carlo Metta (CNR), Riccardo Guidotti (UNIPI), Giovanni Stilo (UNIAQ), Mario Alfonso Prado Romero (UNIAQ)*

The emerging explosion of XAI techniques has recently brought to their design and usage for different data types ranging from tabular data to graphs and passing from images and time series. In this lecture we will provide a focused overview of some explanation methods illustrating with practical examples how certain types of explanation can be achieved with state-of-the-art explanation methods and which are the pre-processing operation necessary for the different data types. After the break In the second part, we will discuss in detail the counterfactual explainability on graphs (GCE) while we will show, by three fundamental use cases, how to extend GRETEL, an extensible-by-design GCE development and evaluation framework that promotes Open Science and reproducibility.

THURSDAY

09:30 - 11:00 **Exploratories Research Hihglights** | Luca Pappalardo (CNR), Donia Kamel (PSE), Todor Galev (CSD), Angelo Facchini (IMT), Carolina Scarton (USFD)

The Bias of the Crowd: Bottom Up Discrimination in Football (co-author: Guillermo Woo Mora) | *Donia Kamel (PSE)*

At the heart of the football industry lies the billion-dollar German-based website, Transfermarkt. This website, which initially started as a source of and reflection on the transfer scene in football, has been, arguably ever since the 2006 World Cup, a focal point in defining it. The valuations reported on the website are initially reported by fans and adjusted by managers. At the end of the day, the system is entirely human. We are interested in studying bottom-up discrimination in football, from the fans to the player's valuation. We look into specifically skin-tone-based discrimination. Creating an algorithm that detects skin tone on a continuous scale, we study the effect of skin tone on player valuation. Just looking at correlations, we see that as skin color gets darker, valuations drop. For identification, we implement a geographical RDD using penalty data, with the two running variables being latitude and longitude surrounding the football goal post.

Studying the links between media capture and disinformation in Southeastern Europe: research challenges, data sources and investigation techniques | *Todor Galev (CSD)*

Over the past decade, foreign authoritarian regimes have intensified their disinformation and political interference campaigns across Europe, with a particular focus on South Eastern Europe (SEE). They exploit and reinforce existing governance vulnerabilities in SEE countries using instruments of state capture, including a specific focus on media capture, to influence decision-making and undermine public trust in democratic institutions. The SEE countries are particularly vulnerable to this threat, as some of them (e.g. Bulgaria and Serbia) show alarmingly high levels of cognitive capture among the general population, political elites and the media, swaying public opinion towards foreign authoritarian models and their aims. The current research analyses media capture as an enabler and amplifier of the creation and dissemination of disinformation. The methodology combines three research and investigative frameworks: (i) a research approach to studying media capture as consisting of four components - ownership, financial, regulatory and cognitive capture; (ii) a methodology to quantify foreign direct and indirect economic footprint in national economies (or particular sectors) and how it is used through the tools of hard, soft and sharp (covert) power to influence decision-making; and (iii) investigative and data science techniques to map disinformation production and amplification networks and their links to media capture.

<u>The Italian municipalities between vulnerability and the ecological transition</u> | *Angelo Facchini (IMT)*

We introduce the Municipality Transition Index, an original composite indicator, used to measure the digital, energy and ecological performance of every Italian municipality, collecting open data at municipality level on digitalization, infrastructures, mobility, environment, and waste. We explore the relative and cumulative impact of the digital and low carbon transition on the capacity to attract and partially revers the decline in population in peripheral location in Italy. The following figures show the distribution of the components of the MTI, highlighting significant disparities in energy and digitalisation, both at macro-regional levels (i.e., North-South divide) and at regional level.

<u>VaxxHesitancy: Studying Hesitancy towards COVID-19 Vaccination on Twitter</u> | *Carolina Scarton (USFD)*

Vaccine hesitancy has always been a concern, being the target of many disinformation campaigns in social media that aimed to cause confusion and make more people deny vaccines' efficacy. With the COVID-19 pandemic, narratives addressing the vaccines were rife in social media. Some people expressed their concerns regarding the fast development of the vaccines and their efficacy, some mentioned their support, while others expressed their denial. Understanding the reasons behind (COVID-19) vaccine hesitancy is important for policy makers, since it can support them in informing the population about the true facts behind vaccines, which can lead to an increase in vaccine take-up. In this talk, I will present our work on computational social science methods for analysing the hesitancy of Twitter users towards COVID-19 vaccination. I will present our novel dataset for addressing this task as a traditional multi-class classification problem, where the Twitter users' stance towards vaccines are the classes we want to predict. In addition, I will also discuss our work on a domain-specific language model that achieves the best results in our classification task.

11:00 - 11:30 **Coffee break**

11:30 - 13:00 **Open Lab (Meet Tutors)** | *Michela Natilli (CNR), Giulio Rossetti (CNR)*

13:00 - 14:30 **Lunch break**

14:30 - 16:00 **Free time**

16:00 - 22:00 **Social tour of Vulcano Island and party on the beach**

FRIDAY

09:30 - 11:00 **Open Lab (Meet Tutors)** | *Michela Natilli (CNR), Giulio Rossetti (CNR)*

11:00 - 11:30 **Coffee break**

11:30 - 13:00 **Students' Final Presentation**

13:00 - **End of the school**