

# Theoretical and Scientific Data Science (TSDS) PhD Programme: Framework and Regulations (“Regolamento Didattico”)

Academic Year 2024-25

Approved by the TSDS Collegio Docenti on 30/08/2024;

17 September 2024

## Introduction

This document sets out the structure of the PhD programme in “Theoretical and Scientific Data Science” (TSDS) and the regulations for students.

The programme is under the responsibility of the Collegio Docenti of the Theoretical and Scientific Data Science group at SISSA, within the Physics Area. Academic oversight and quality assurance are provided by the TSDS Collegio Docenti, and, where appropriate, the Consiglio dell'Area di Fisica and SISSA's Senate.

These regulations complement the SISSA PhD regulations, “Regolamento Didattico dei Corsi di Philosophiae Doctor della SISSA”, Prot. n. 7374 - 1/3 del 21/03/2022 - UOR: 80925 - Rep. n. 209/2022, to which students are referred for further details.

## Overview of the Programme

In the last few years, the exponential growth both in the quantity and complexity of data sets has created an immense interest and an opportunity for data science: without novel machine learning and statistical methods able to compress, clean, interpret and ultimately extract knowledge and value from those vast data sets, our understanding of the world around and within us would be limited.

Data science today is a cross-disciplinary field that draws from computer science, mathematics, statistics, data analysis, programming and theoretical physics to advance our ability to understand phenomena ranging from the first instants after the Big Bang to the coronavirus pandemic.

The aim of this new PhD programme at SISSA is to equip students with the tools they need to develop and apply cutting-edge data science to problems in the natural sciences (hence the "scientific" adjective in the name of the programme). We are however not merely interested in application: our research encompasses theoretical aspects of data science and machine learning, aiming at understanding the how and the why of the unprecedented success of many of the tools that are shaping the world around us (hence the "theoretical" adjective in our name).

After 4 years, our students will have received an excellent education in machine learning (including deep learning, neural networks, reinforcement learning and information theory), statistical analysis (including Bayesian methods and algorithms such as Markov Chain Monte Carlo and nested sampling) and analytical problem solving (including advanced programming, high performance computing and software development). The highly cross-disciplinary programme is complemented by modules of a more humanistic flavour, focusing on ethical aspects of AI and machine learning, in order to broaden the students' understanding and appreciation of the ethical and moral questions associated with applications of machine learning and AI to society.

Graduates from our programme achieve competences and skills in data science that are in high demand today, both in academia and business. SISSA being one of the six Italian graduate schools, our training focuses on creating graduates who have an outstanding education at the frontiers of research, many of whom pursue an academic career. However, opportunities for a fulfilling career are plentiful in a variety of fields, given that data science is today deeply embedded in many aspects of society, industry and business. Potential careers paths range from the banking sector to transport, digital services, logistics, internet companies, market research, finance and many others.

#### Composition of TSDS Collegio Docenti (Academic Board) for AY 2024-25:

1. Roberto TROTTA (PhD Coordinator)
2. Jean BARBIER (ICTP)
3. Luca HELTAI
4. Nicoletta KRACHMALNICOFF
5. Alessandro LAIO
6. Guido SANGUINETTI
7. Mathew DIAMOND
8. Matteo VIEL
9. Stefano DE GIRONCOLI
10. Remo SANGES
11. Matteo MARSILI (ICTP)
12. Andrea SOTTORIVA (Human Technopole)
13. Sebastian Dominik GOLDT
14. Fernanda Pinheiro (Human Technopole; *adjunct*)
15. Alex RODRIGUEZ (Uni TS; *adjunct*)
16. Giulio CARAVAGNA (Uni TS; *adjunct*)

#### Programme Structure

The PhD programme lasts for 4 years, and it is structured in two parts: the first year is mostly in the form of taught modules, which lay the theoretical and practical foundations for the knowledge base that students will draw upon in their research.

Starting approximately 9 months after their enrolment, students undertake their PhD research work in one of the subjects offered across SISSA, with at least one supervisor from

the Collegio Docenti and staff at SISSA, and often a co-supervisor from another group - or another institute altogether. Students on the joint Human Technopole-SISSA programme will normally at this point in time start their research placement with Human Technopole.

We strongly encourage cross-disciplinary projects, offering the opportunity to pursue both a data science and a discipline-specific question. Fields range from astronomy to neurosciences, computer science to theory of neural networks, information theory to statistical physics, condensed matter and material sciences, bioinformatics and multi-omic data, and many more. Some projects might involve the opportunity of a research placement in industry. All students will be encouraged and supported to present their work at international conferences and meetings at the appropriate point in time of their studies.

We are part of the Italian National PhD in AI, of the Joint International ICTP/SISSA Ph.D (JIISP) Programme in Physics and Mathematics, the ELLIS PhD programme and we collaborate closely with the Human Technopole in Milan, with which we offer a joint SISSA-HT PhD stream (with courses at SISSA and extended placements in Human Technopole).

The TSDS group offers a series of core modules and some optional monographic courses. Additional modules from other PhD programmes within SISSA or other institutions may also be available to TSDS students, who need to discuss the matter (particularly if they plan to accrue credits) with the PhD coordinator.

All of our Data Science modules are open to students from other PhD programmes at SISSA. Credit can be accrued by such students with the written agreement of their PhD programme coordinator. Students from other institutions with which SISSA has a suitable agreement in place (e.g., University of Trieste, ICTP and others) can take exams and accrue credit as specified in the relevant agreements.

Auditors (students outside SISSA attending the course without credit and without taking the exam) are admitted at the discretion of the course lecturer. Their presence must be notified to the PhD coordinator for administrative reasons.

- YEAR 1:
  - First-year students from the TSDS group must take all core modules offered by the group, and take at least 1 optional module worth at least 2.5 credits; this could be one of the monographic courses offered in the third term, or can be chosen from the offering from other PhD programmes, upon agreement of the PhD coordinator.
  - Each first-year student will be assigned a personal tutor, usually drawn from the Collegio Docenti of the PhD programme. The role of the personal tutor is to act as a pastoral care point of contact, and to help the student with orienting their academic trajectory, including supporting the student in choosing the topic of their PhD research and offering career advice later on in the student's academic development.
  - To be admitted to year 2, TSDS first-year students must take and pass exams in all the core modules and in their chosen optional module(s) by September of the year following their enrolment. Marks are on a scale where 30 is the

maximum (with “lode” as an option to denote exceptional performance), and the pass mark is 18/30. Some exams are pass/fail, and this will be indicated by the lecturer.

- Additionally, students must achieve a weighted average of 27/30 in the core modules marks.
- By the end of June of the year following their enrolment, first-year students will select their PhD topic, which will be assigned together with a main supervisor and co-supervisor where appropriate. Either the main supervisor or the co-supervisor must be a member of the TSDS Collegio Docenti. Cross-disciplinary co-supervision between TSDS staff and researchers from other departments (and/or institutions) is encouraged.
- In late September of the year following their enrolment, first-year students submit a progress report and take a transfer exam (in the form of a viva) on the work done to date on their PhD topic. This is a pass/fail examination, which will be evaluated by the Collegio Docenti.
- YEARS 2-3:
  - In late September of each successive year, students present a Progress Report on their research project in the form of a viva and accompanying written document. This is a pass/fail examination, which will be evaluated by the Collegio Docenti.
  - Should a student fail to progress, the Collegio Docenti will discuss the individual circumstances and suggest an appropriate course of action.
- YEAR 4:
  - By September of the fourth year after their enrolment, students present their PhD thesis in accordance with SISSA regulations.

Students are expected to attend Journal Clubs and seminars, as an important part of the educational offer of the group, as well as other learning opportunities (e.g. group meetings) as indicated by their supervisors. An attendance register will be kept. Students who do not attend at least 90% of such events without a valid justification will be referred to the Collegio Docenti.

Further details and up-to-date offer of modules is published on our webpage: [datascience.sissa.it](https://datascience.sissa.it)

## Examinations

The module leader (usually the main lecturer) of each course will inform students of the timing, format and modality of the exam at the start of the module. Exams may take the form of oral examinations, open-book projects, group projects, essays, etc.

In case of insufficient or unsatisfactorily performance, exams may be re-taken **once** at the sole discretion of the examiner(s). Students who fail to attend scheduled exams without providing a valid reason and justification will be considered to have failed the exam.

Mitigating circumstances (e.g. illness, bereavement, mental health issues, etc) need to be notified promptly to the PhD Coordinator when they arise, and in any case before taking an

exam if the student feels their performance might be impacted. Mitigating circumstances disclosed after an examination cannot be taken into account.

## Feedback

During taught modules, students will have numerous opportunities to receive feedback on their work by the module leader in a number of ways. These include oral feedback on their work in labs, feedback on the exam, continuous assessment as part of class discussions, etc.

In the run-up to the first year transfer exam, students will also receive continuous feedback from their project advisor, and will be supported with regular meetings and guidance in preparation for the exam.

## Academic Supervision and Pastoral Care

Students are expected to meet at least once a term with their personal tutor; the frequency will be adjusted by mutual agreement between the student and the tutor according to the individual circumstances.

Students are expected to meet regularly with their PhD supervisor once they have begun their research project – a typical schedule is once a week. The supervisor (or the co-supervisors in case of shared supervision) also provides pastoral care for the students, in concert with the personal tutor.

In case of concerns involving the (co)-supervisor(s), students should contact the PhD coordinator, their personal tutor and/or another TSDS senior staff member if that would not be appropriate. If the matter cannot be resolved, they will escalate it in accordance with SISSA regulations.

## Code of Conduct

The TSDS group is committed to creating a work environment that is safe, professional and of mutual trust where diversity and inclusion are valued and where everyone is entitled to be treated with courtesy and respect.

As members of the SISSA community, students and staff pledge to treat all people with respect and provide a harassment- and bullying-free environment, regardless of sex, sexual orientation and/or gender identity, disability, physical appearance, body size, race, nationality, ethnicity and religion. In particular, sexual language and imagery, sexist, racist or otherwise exclusionary jokes are not tolerated.

We pledge to provide respectful criticisms and feedback in our daily interactions. We pledge to model behaviours that encourage productive discussion for better scientific outcomes, even when there is disagreement, while fully respecting individual differences and preferences.

## Teaching credits

Supervisors of TSDS students from Year 2 onwards are entitled to claim up to 80 hours of teaching credits per academic year, which will count against their teaching load.