**Foundations of Trustworthy AI – Integrating Reasoning, Learning and Optimization**
**TAILOR**
**Grant Agreement Number 952215**

# Connectivity Fund

| | |
|---|---|
| **Document type (nature)** | Report |
| **Deliverable No** | D10.5 |
| **Work package number(s)** | WP10 |
| **Date** | Due M23 |
| **Responsible Beneficiary** | TUE, ID 12 |
| **Author(s)** | Joaquin Vanschoren |
| **Publicity level** | Public |
| **Short description** | This deliverable repeats every year and presents regular updates about the outcomes of the open call of the connectivity fund. |

| **Document History** | | | |
|---|---|---|---|
| **Revision** | **Date** | **Modification** | **Author** |
| 1.0 | 16/11/2022 | first version | Joaquin Vanschoren |

| **Document Review** | | |
|---|---|---|
| **Reviewer** | **Institution** | **Date and result of Review** |
| Fredrik Heintz | Linköping University | 2022-11-28 |
| Peter Flach | Bristol University | 2022-11-28 |

**Table of Contents**



## Useful links

Hosted website for the connectivity fund:
https://tailor-network.eu/connectivity-fund/

# Introduction

The TAILOR network includes many of Europe's top AI labs. However, we also want to reach out to the many other excellent labs and organizations across Europe to work together and create new breakthroughs in AI. The Connectivity Fund is a key instrument in this mission. To establish a truly vibrant network, the Connectivity Fund provides funding to AI researchers from across Europe for research visits or workshops that bring together researchers from TAILOR labs and non-TAILOR labs. It especially aims to support young researchers to gain valuable experience and nurture the next generation of AI researchers. The goals, scope, organization, proposal evaluation, and legal framework have all been described in earlier deliverables D10.1-10.4. These remain unchanged.

This series of deliverables provides updates on the status of the connectivity fund. Since September 2021, these updates will be done yearly. It will detail the number of submissions, the evaluation process, and the outcomes, funded visits, and workshops in a transparent way.

Since the connectivity fund operates using a continuous open call, with cut-offs every 4 months, this report covers the results of the following cut-off rounds:

- 15th of November, 2021
- 15th of March, 2022
- 15th of July, 2022

# Overview

The connectivity fund has funded 14 research visits and 2 workshops. A geographical overview is shown below.

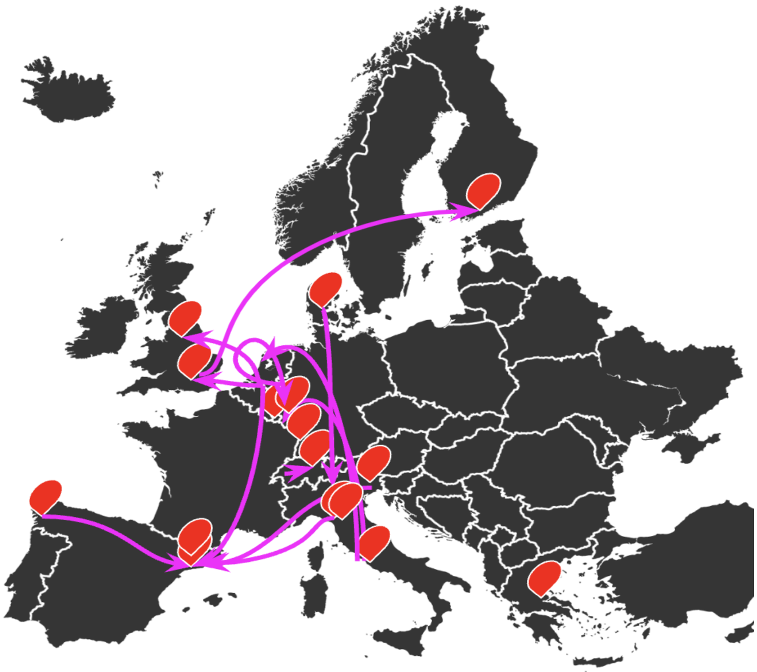

Figure 1. Geographical overview of Connectivity Fund visits.

# Dissemination activities

In this reporting period, we significantly stepped up our dissemination activities to ensure that the Connectivity Fund is widely known in the European AI community.

## Connectivity Fund Website

Following recommendations from the project reviewers, the connectivity fund website was seamlessly integrated in the TAILOR project website, as shown in the image below.


Figure 2. New Connectivity Fund website.

This website has all the information about the aims of the fund, how to apply, the evaluation procedure, and other useful information for applicants. It also contains a gallery of all funded projects. For each project, we show the abstract of the research proposal and a short bio of the researcher who won the award. A screenshot of this gallery is shown below.

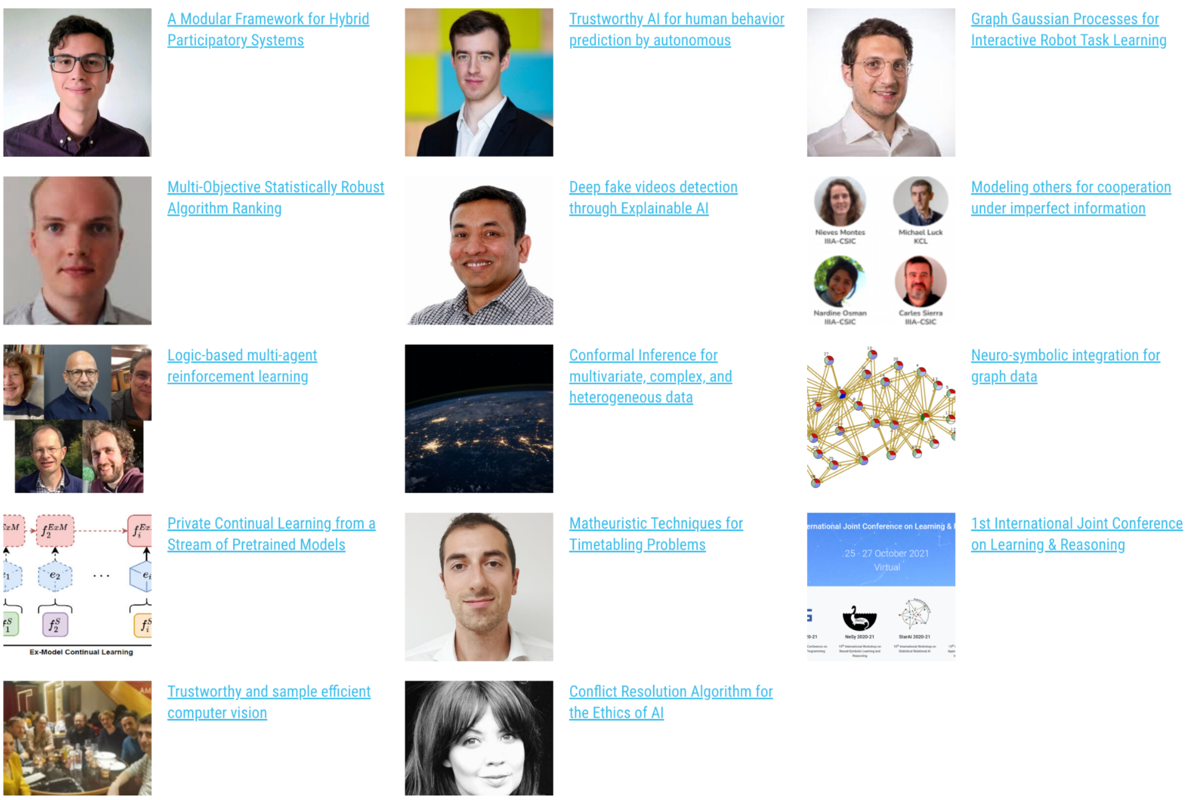


Figure 3. Gallery of successful Connectivity Fund projects

# General dissemination

We have spread the word about the connectivity fund through various communication channels:

- Regular announcements on the upcoming Connectivity Fund deadlines in the TAILOR newsletters
- News items on the AI on-demand platform
- Social media (e.g. Twitter). Examples are shown below.
- The TAILOR Conference in Prague and the TAILOR booth at the IJCAI conference included an information stand where the audience could learn more about the fund. This included an information banner, also shown below.
- During the TAILOR Open Monthly Meetings, we also invite Connectivity fund beneficiaries to tell the audience more about their experiences. In a recent edition, Roberto Rosati (Univ. Udine) and Antonio Cart (Univ. Pisa) told us about the impact that the connectivity fund had on their research and careers, and shared tips on applying and preparing for research visits.

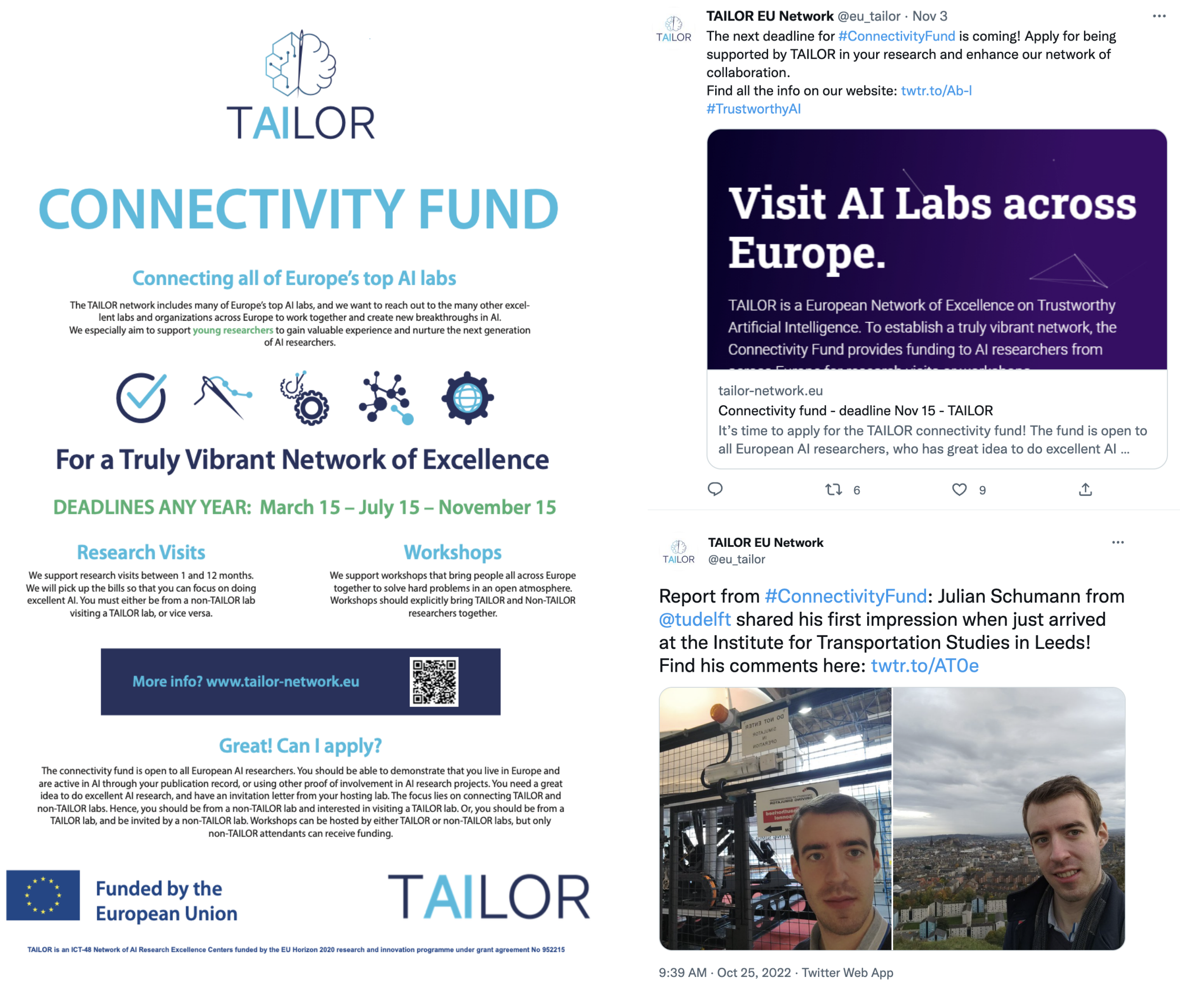


Figure 4. Connectivity Fund information sheet (left) and Twitter activity (right)

# New applications received

This deliverable focuses on the period of September 2021 to September 2022. There were many COVID19 restrictions still in place during this period. Due to the impossibility or uncertainty around research visits and in-person events, the response to the call in this period was quite low. We received the following eligible applications in this reporting period. The institute shown in **bold** is the hosting lab, hosting the researcher(s) during their visit.

| ID | Title | TAILOR lab | non-TAILOR lab (fund recipient) | status |
|---|---|---|---|---|
| 6 | **Neuro-symbolic integration for graph data** *Manfred Jaeger (U. Aalborg), Andrea Passerini (U. Trento)* | **Univ. Trento** | Univ. Aalborg | done |
| 7 | **Conformal Inference for multivariate, complex, and heterogeneous data** *Marcos Matabuena (CiTIUS, Univ. Compostela)* | **U.Pompeu Fabra** | CiTIUS | done |
| 8 | **Private Continual Learning from a Stream of Pretrained Models** *Joost van de Weijer (U.A. Barcelona)* | Univ. Pisa | **U.A. Barcelona** | ongoing |
| 10 | **Logic-based multi-agent reinforcement learning** *Natasha Alechina, Mehdi Dastani, Brian Logan and Giovanni Varricchione (U. Utrecht)* | **La Sapienza, Roma** | Univ. Utrecht | ongoing |
| 11 | **Modeling others for cooperation under imperfect information** *Nieves Montes, Nardine Osman, Carles Sierra (IIIA Barcelona)* | IIIA, Barcelona | **King's college London** | ongoing |
| 12 | **Deep fake videos detection through explainable AI to combat disinformation on social media** *Nadeem Qazi (U. East London)* | **TIETO Finland** | Univ. East London | ongoing |
| 13 | ***Workshop*: IAIL 2022 - Imagining the AI landscape after the AI act** *Francesca Naretto (Scuola Normale Superiore)* | CNR Italy | Scuola Norm. Superiore | done |
| 14 | **Trustworthy AI for human behavior prediction by autonomous vehicles: Towards a comprehensive benchmark -** *Julian Schumann (TU Delft), Gustav Markkula (U. Leeds)* | TU Delft | **Univ. Leeds** | to start |
| 15 | **A Modular Framework for Hybrid Participatory Systems** *Enrico Liscio (TU Delft), Maite Lopez Sanchez (U. Barcelona)* | TU Delft | **Univ. Barcelona** | ongoing |
| 16 | **Multi-Objective Statistically Robust Algorithm Ranking** *Jeroen Rook (U Twente), H. Hoos (RWTH Aachen), H. Trautmann (U. Munster)* | **RWTH Aachen** | Univ. Twente | ongoing |
| 17 | **Graph Gaussian Processes for Interactive Robot Task Learning** *Giovanni Franzese (TU Delft), Marc Deisenroth (UC London)* | TU Delft | **Univ. College London** | to start |

*Table 1. Overview of TAILOR Connectivity fund applications.*

# Review process and outcome

Proposals in this reporting period were collected via the Easychair platform. First, all applications submitted within the deadline were evaluated for formal eligibility by the call management. The eligibility criteria are specified in deliverable D10.1. All proposals except one passed the eligibility test.

Next, each proposal was reviewed by one or two members of the scientific board. All members of the board have been actively involved. We checked for any conflicts of interest in the assignment. The final results of the evaluation are summarized in Table 1. All proposals were evaluated according to 5 criteria (AI excellence, scientific track record of the candidate, scientific step-up, the suitability of the host, and appropriateness of the activity duration). Further details on these can also be found in deliverable D10.1.

While we originally required one review per application, we increased this to two reviewers since June 2022, following feedback from the TAILOR project reviewers. The table below summarizes all the scores as well as the involved labs and requested funding. From proposal #14 on, we show the scores of both reviewers.

| ID | AI Excellence | Science track record | Science step-up | Host lab | Visit length | Final Score | Planned start date | Requested fund (EUR) |
|---|---|---|---|---|---|---|---|---|
| 6 | 8 | 8 | 9 | 8 | 9 | 8.4 | 14/3/2021 | 3,735 |
| 7 | 9 | 9 | 9 | 9 | 8 | 8.9 | 1/2/2022 | 6,500 |
| 8 | 8 | 8 | 9 | 8 | 7 | 8.1 | 1/6/2022 | 5,000 |
| 10 | 9 | 10 | 9 | 9 | 8 | 9.1 | 1/3/2022 | 10,200 |
| 11 | 8 | 7 | 10 | 10 | 9 | 8.6 | 1/6/2022 | 15,000 |
| 12 | 7 | 7 | 8 | 7 | 7 | 7.3 | 1/8/2022 | 15,000 |
| 13 | 8 | 8 | 8 | 9 | 8 | 8.1 | 13/6/2022 | 10,000 |
| 14 | 8/8 | 8/7 | 8/10 | 9/10 | 9,9 | 8.4 | 1/10/2022 | 7,000 |
| 15 | 8/6 | 9/6 | 9/7 | 9/6 | 9/6 | 7.5 | 1/10/2022 | 6,870 |
| 16 | 9/8 | 8/9 | 9/7 | 9/10 | 9/8 | 8.5 | 15/9/2022 | 4,617 |
| 17 | 8/8 | 8/9 | 9/9 | 8/9 | 8/9 | 8.5 | 1/9/2022 | 11,000 |

*Table 1. Overview of TAILOR Connectivity fund applications and their evaluation.*

Scores per criteria are on a scale from 0-10:

- **0-1** Application fails to address the criterion or cannot be assessed due to missing or incomplete information
- **2-3 Poor** – criterion is inadequately addressed or there are serious inherent weaknesses
- **4-5 Fair** – application broadly addresses the criterion, but there are significant weaknesses
- **6-7 Good** – application addresses the criterion well, but a number of shortcomings are present
- **8-9 Very good** – application addresses the criterion very well, but a small number of shortcomings are present
- **Excellent** – application successfully addresses all relevant aspects of the criterion. Any shortcomings are minor.

The final score is a weighted average of all scores, using the weighting described in deliverable D10.1.

As per the Connectivity Fund rules, proposals must achieve a minimum of 70% of the maximal score to receive funding. All eligible applications passed this threshold. The total requested funding is also well within the budget earmarked for the second year of the connectivity fund (500k EUR divided over 4 cut-offs). Based on these outcomes, all these applications were accepted for funding.

# Organizational streamlining

To streamline the management of the connectivity fund, a number of changes were made which will take effect in the next reporting period.

## Collaboration Exchange Fund

Next to the Connectivity fund, which only supports third party funding, TAILOR will start a Collaboration Exchange Fund for research exchanges within the TAILOR network. This fund will be described in more detail in an upcoming Work Package 9 deliverable. It will generalize the goals of the connectivity fund and enable research visits which are currently not yet possible. To build on current experience, it will mirror the submission and evaluation procedures of the Connectivity fund, and depend on a shared scientific board to evaluate the applications.

## Scientific Board Extension

The scientific board will be extended with 17 new members. All 13 current members have also agreed to extend their tenure on the scientific board, bringing the total to 30 members. This will allow the connectivity fund to scale, and help us deal with the sharp growth in the number of applications, the doubling of the reviewing load (since we now require two reviews per application), and the additional applications coming from the new Collaboration Exchange Fund.

The new scientific board members are listed below. This extension will still need to be confirmed in the next General Assembly.

- Kim Baraka, VU Amsterdam
- Francisco Chicano, University of Malaga
- Miguel Couceiro, University of Lorraine, CNRS, LORIA
- Jérôme Euzenat, INRIA & University Grenoble Alpes
- Andreas Herzig, IRIT-CNS
- Manolis Koubarakis, National and Kapodistrian University of Athens
- Agnieszka Ławrynowicz, Poznan University of Technology
- Nikolaos Matragkas, CEA
- André Meyer-Vitali, DFKI
- Wico Mulder, TNO
- Isabel Neto, University of Lisbon
- Eva Onaindia, Universitat Politècnica de València
- Hossein Rahmani, Lancaster University
- Fabrizio Riguzzi, CINI
- Raul Santos-Rodriguez, University of Bristol
- Silvia Tulli, Instituto Superior Técnico
- Neil Yorke-Smith, Delft University of Technology

## Application streamlining

We currently use Easychair to collect proposals. This is not ideal since Easychair is not designed for open calls with multiple deadlines, so we can't use most of its useful reviewing features. For the next cutoff (in March 2023) we will move to an online form which will allow us to automate the collection, review, and tracking of the applications through simple scripts.

# Impact evaluation

The Connectivity fund is a key mechanism to foster collaboration and allows researchers all over Europe to work on the core research problems addressed by the TAILOR network. It also allows TAILOR to open up to a wider section of the AI community. To measure its impact, for all proposals since 2022, we ask all participants to send a structured scientific report containing:

- A summary of the research objectives
- Technical approach, findings, and future work
- A self-assessment of the impact of the research visit on AI excellence and their own careers, as well as the suitability of the host and the visit length.
- A list of publications and other outcomes of the visit.

These reports are attached with this deliverable.

# Outlook

Due to the COVID pandemic, the total amount of spending is far below our original estimates. However, as travel restrictions are being lifted, we are seeing a significant increase in interest. In the latest November call (not yet reported here) the number of applications has more than doubled. We will continue to monitor usage of the available funds and will take action to ensure that the connectivity fund has a maximal impact. We will also explore coordinated actions to involve other ICT-48 networks.

The extension of the TAILOR project to four years will allow us to recover from the restricted circumstances caused by the Covid pandemic and catch up with our original goals.

# Basic Information

**Project title**: Rotation-invariant vision transformers for trustworthy and sample efficient computer vision
**Period of project**: 10.2022-12.2022
**Period of reporting**: full project
**Author(s)**: Mohammadreza Amirian
**Organisation**: Zurich University of Applied Sciences (ZHAW)
**Host organisation**: Swiss Federal Institute of Technology Lausanne (EPFL)

# Public summary

After the breakthrough of transformers in natural language processing, these models are now adapted for computer vision and image classification tasks. Transformer-based models showed at least equal descriptive properties compared to convolutional models; however, initial specimens required more data for generalisation than convolutional models. Furthermore, these models didn't produce translation, rotation, and scale equivariant features. In a recent study, researchers introduced rotation equivariant transformers for a discrete rotation group, although these models' generalisation properties and sample efficiency were not well investigated yet. During this project, we aimed to extend the concept of rotation invariance for continuous rotation to improve the trustworthiness of the decisions and robustness of the vision transformer models. However, computer vision models demonstrated vulnerability towards variations in the angle and scale of the input images. This weakness leads to reduced trust in models' decisions in some circumstances that we addressed through furthering reliability and robustness using rotation invariance. Furthermore, gains in sample efficiency and improvement in generalisation are expected as the features show consistency over different variations in the rotation of the original image.

# Research objectives

## Objectives

Inductive biases such as translation-invariance undeniably accelerated the rapid advances of modern vision models through parameter sharing and improving sample efficiency. However, state-of-the-art models can only partially incorporate rotation-invariance. Recent attempts to develop rotation-invariant techniques mainly face the challenge of high memory requirements or limiting the original model capacity. This research proposes an embedding layer for vision transformers to leverage the invariance of self-attention layers to the order of tokens and train robust models against local and global rotation. The proposed image embedding technique requires negligible memory overhead to train rotation invariance models on large datasets such as ImageNet. The paper presents the proposed method's merit in improving the sample efficiency and robustness of vision transformers on small and larger datasets on classification and segmentation tasks.

### Impact

Reproducible research work with source code for rotation-invariant vision transformers has been developed. The developed method provides more robust vision models against rotation variations of the input images, contributing to training more robust and trustworthy vision models. Furthermore, the models aim at rotation invariance, and the proposed transformers can cope with a range of transformed input images; thus, we expect a higher sample efficiency in future experiments. The developed method will be published in a Ph.D. dissertation as a proof of concept with the possibility of application in future research.

## Technical approach

### Detailed description

Deep vision models demonstrated vulnerability against rotation and scaling, starting from early research works. Many researchers attempted to tackle this problem using data-driven techniques or by adding rotation invariance inductive bias to models. Although data-driven approaches turned out to be more straightforward to implement, rotation-invariant models show their merits regarding sample efficiency and small data applications. The memory required for current state-of-the-art vision transformer techniques aiming at rotation invariance shows a linear increase with the size of the rotation group. The memory requirement of these models based on lie-groups increases linearly with the size of the group and only applies to discrete rotations with a limited number of angles. This project resulted in a novel patch embedding method for transformers that is robust to the continuous global rotation of the input, shows minimal memory overhead, and can achieve acceptable performance on several benchmark datasets for image recognition with a minor drop in accuracy.

### Scientific outcomes

The overview of the present literature and techniques for rotation invariance, together with the insights found in rotation invariant and equivariant transformers, was presented both at the machine learning and optimization (MLO) lab group meeting at EPFL and computer vision, perception, and cognition (CVPC) lab at ZHAW.

### Future plans

This research work might be published in future publications, or there are chances for collaboration between participants at the postdoc level.

# Self-assessment

Please provide your own final assessment of the effective progress against the goals stated in the proposal, according to the following points:

- **AI Excellence**: The visit helped me to have a better understanding of vision transformers and rotation invariant and equivariant models. We found a tokenization technique for vision transformers to gain rotation invariance in object recognition. The proposed method achieved robustness against rotation in object recognition tasks with competitive performance with a minor compute overhead and drop in accuracy.
- **Scientific step-up**: I worked in a fully functional and productive team with a positive dynamic and vibe. I tried my best to learn leadership and research skills from Martin. The visit contributed to my personal development by learning new scientific content, working on a new problem, and getting in touch with scientists at the summit of their niche.
- **Suitability of the host**: I felt welcomed in the MLO lab and experienced an excellent work environment. In addition, I had good access to Martin and worked closely with his Ph.D. student. Hence, I enjoyed high-level and low-level support for concept development and hands-on work. I also had access to MLO's compute resources.
- **Suitability of the visit length**: The visit was too short for the project considering the time for paperwork, settling down in Lausanne, and new years' holidays. However, the extension was not also an option due to the other projects' commitments and limitations of migration rules and regulations in Switzerland and universities of applied sciences.

# List of publications, meetings, presentations, patents,...

Presentations:

- Equivariant Neural Networks: machine learning and optimization lab (MLO), 4th of January 2022, Lausanne, Switzerland
- Equivariant Neural Networks: computer vision, perception and cognition group (CVPC), 23th of February 2022, Winterthur, Switzerland

# Additional comments

This is the report of the first project supported by TAILOR's connectivity fund. From initial communications, it wasn't clear that the funding doesn't support working hours. Hence, the budget from other projects, personal overtime, and holidays of the applicant filled the working hours for this project. Unfortunately, this was not healthy for the applicant and closed the opportunity to continue the project after the end of the stay. Therefore, it is highly recommended to check the existence of base funding in the future to guarantee fair working conditions for applicants and ensure the project's success without early stopping.


# 1ST INTERNATIONAL JOINT CONFERENCE ON LEARNING & REASONING

7 March 2022


## Nikos Katzouris

### National Center for Scientific Research "Demokritos"


The rapid progress in machine learning has been the primary reason for a fresh look in the transformative potential of AI as a whole during the past decade. A crucial milestone for taking full advantage of this potential is the endowment of algorithms that learn from experience with the ability to consult existing knowledge and reason with what has already been learned. Integrating learning and reasoning constitutes one of the key open questions in AI, and holds the potential of addressing many of the shortcomings of contemporary AI approaches, including the black-box nature and the brittleness of deep learning, and the difficulty to adapt knowledge representation models in the light of new data. Integrating learning and reasoning calls for approaches that combine knowledge representation and machine reasoning techniques with learning algorithms from the fields of neural, statistical and relational learning.

Aiming to address such challenges, the 1st International Joint Conference on Learning & Reasoning (IJCLR 2021), which was sponsored by TAILOR, took place as a virtual conference from October 25-27 2021. IJCLR 2021 brought together, for the first time, four international conferences and workshops, addressing various aspects of integrating machine learning and machine reasoning:

- The 30th International Conference on Inductive Logic Programming (ILP).
- The 15th International Workshop on Neural-Symbolic Learning & Reasoning (NeSy).
- The 10th International Workshop on Statistical Relational Artificial Intelligence (StarAI).
- The 10th International Workshop on Approaches and Applications of Inductive Programming (AAIP).

The conference featured presentation of cutting-edge research in a number of parallel sessions for each participating event, in addition to a number of joint invited talks from leading researchers in the field, tutorials, poster sessions and a panel discussion on "Future Challenges in Learning & Reasoning".

The virtual conference was organized by the Institute of Informatics of the National Center for Scientific Research (NCSR) "Demokritos", in Athens, Greece. The conference had more than 550 registrants and a very high participation overall.

The video recordings from IJCLR 2021 are available online for everyone to watch.

← Previous Post                Next Post →

Funded by EU Horizon 2020.


Contact: Coordinator Fredrik Heintz, fredrik.heintz@liu.se.

**Final Report**

for the research stay of dr. Réka Markovich in CNR, Pisa, 2021 titled
"Abstraction and Implementation: Toward a Context-Dependent Conflict Resolution Algorithm for the Ethics of AI"

The goal of dr. Markovich's research stay at the CNR in Pisa was to proceed her investigations about a context-dependent conflict resolution algorithm which can be used in Machine Ethics.

The problem of moral disagreement is a serious barrier to the advancement of ethical AI as such [Bostrom 2014, Brundage 2014, Etzioni and Etzioni 2017, Formosa and Ryan 2020, Gabriel 2020]. People agree that the soon-to-be-developed autonomous intelligent systems (AIS) should obey our moral rules, but do not agree about what these rules are. The most well-known results about this problem are those of the Moral Machine experiment [Awad et al., 2018] exemplifying the differences of what behavior of an autonomous system people from all over the world find morally acceptable/desirable. But the situation is not better with the professionals when deciding what norms these systems should follow: as Jobin et al. [2019] wrote after analyzing 84 ethical guidelines for AI, while there were some emerging values, "no single ethical principle appeared to be common to the entire corpus of documents". The shared opinion of ethicist seems to be that AIS's operation should be restricted to those occasional cases where stakeholders actually agree on the moral rules [Anderson, 2011]. The problem of moral disagreement, therefore, threatens the promising aspects of AIS. What is more, if we think of that the technological development cannot be stopped, the conclusion is even (much) more threatening: with moral disagreement we lose moral control over these systems, so the demand for a solution is even more urgent.

Dr. Markovich has been working on a conflict-resolution algorithm which is based on a branch of law's model, the so-called Conflict of Laws. It is part of Private International Law, and its task is exactly to provide rules in international situations where more than one nation's laws could be applied: the Conflict of Laws rules tell which one's should be. Dr. Markovich has already published a paper, "On the Formal Structure of Rules in Conflict of Laws", in which she proposes a formalization of the rules of Conflict of Laws using an approach, language, and semantics taken from the so-called Input/Output Logics [Markovich, 2019]. Input/Output Logic is a rule-based system, a theoretical framework for reasoning about conditional norms and investigating normative systems. [Makinson and van der Torre, 2000]. It is a formalism that, modelling the essential structure of rule-based reasoning, allows the definition and analysis of different possible kinds of normative reasoning. However, it does not really take into consideration the expressivity needed for particular applications, and, consequently, it does not consider computational costs and ease of implementation.

The direct goal of dr. Markovich's research stay was to investigate a family of formal languages that are more expressive and more implementation-oriented: the OWL family of languages (https://www.w3.org/OWL/), and to see how it could be used for the modelling of Conflict of Laws and then later to the conflict-resolution algorithm to be applied in Machine Ethics. OWL (Web Ontology Language) is the most popular formalism in formal ontologies. Ontological modelling tools are very popular in many areas, even in legal informatics [Casanovas et al., 2016]. The OWL languages have a solid logic background, represented by the Description Logics (DLs), a family of logical systems that has allowed the development of efficient reasoners (http://owl.cs.manchester.ac.uk/tools/list-of-reasoners/). As in the case of normative reasoning, a big part of the DLs formalism (the TBox part) is of conditional nature in the language and in the reasoning processes. However, classical DLs rely on classical logical semantics. These are usually considered as not appropriate for modelling normative reasoning which exhibits properties that are not shared by classical logics: defeasible conclusions, different priorities among the norms, context-dependent interpretations… [Rotolo et al., 2017].

Dr. Markovich has chosen to visit ISTI-CNR because of the experts in the areas of Knowledge Representation and Reasoning and in formalisms for the Semantic Web. In particular, she was collaborating with
- Dr. Umberto Straccia (http://www.umbertostraccia.it/cs/)
- Dr. Giovanni Casini (https://www.isti.cnr.it/en/about/people-detail/85/Giovanni_Casini)
These researchers have provided a main contribution to the area of extending DLs with non-classical reasoning [Casini and Straccia 2013, Casini et al., 2019]. Actually, the Defeasible DL system developed by the researcher at ISTI-CNR has already been used to define an architecture to solve potential conflicts between the different l egal codes (national, regional, and so on) valid in Brazil [Rodrigues et al. 2019].

During the research stay, the researchers investigated the applicability of the methods used in [Rodrigues et al. 2019] to the given problem, identifying some limitations in such a proposal, and rethinking a possible approach in the framework of defeasible DLs toward a formalization of the CoL conflict resolution methodology. The desired properties of the mechanism are correctness, ease of implementation and relatively low computational costs.

The main goal of the research stay was to investigate whether the use of Defeasible DLs is a feasible approach toward the formalization of the CoL. Dr. Markovich and dr. Casini worked on an alternative model of Conflict of Laws, with the support of dr. Straccia, to proceed toward the conflict-resolution algorithm with the desiderata indicated above. The work focused on some portions of the Hungarian CoL, that dr. Markovich identified as particularly problematic from the point of view of formalization and reasoning. They identified an appropriate vocabulary (concepts, roles) and a first draft of an appropriate T-box (the part of the ontology containing all the general rules and constraints expressed, in this case, by a CoL code. We identified those norms in Conflict of Laws that do not seem expressible in description logic. We are still looking for a correct formalization in DLs, but if we do not find any, we are considering combining the DL ontology with a Datalog rule-based system, that would allow the formalization of such norms. Some complex constructors in the language turned out to be necessary in order to develop a proper ontology formalizing CoL rule: qualified number restrictions, inverse and transitive roles and nominals, among others. An expressive DL like SHOIQ [Horrocks and Sattler 2005] appears sufficiently expressive for the task.

On the one hand, as foreseen in the proposal's review, one month was enough only to start the process and identify some cornerstones and directions. Therefore, the researchers plan to submit a TAILOR proposal in 2022 too to continue the collaboration. On the other hand, however, the visit did already provide dr. Markovich with knowledge on the practical implementation of normative reasoning systems for conflict resolution and the identified aspects crucially contributed to the elaboration of her research plans about the algorithm, which she implemented in an ERC Starting Grant proposal submitted in January of 2022. Also, the researchers are preparing a submission to JURIX or its accompanying workshops on AI and Law, and will later a journal submission.

References:

[Anderson, 2011] Anderson, S. (2011). Philosophical Concerns with Machine Ethics. In M. Anderson & S. Anderson (Eds.) *Machine Ethics*. Cambridge: Cambridge University Press. pp. 162–167.
[Awad et al., 2018] E. Awad, S. Dsouza, R. Kim, J. Schulz, J. Henrich, A. Shariff, F. Bonnefon, and I. Rahwan. The Moral Machine experiment. *Nature*, 536:59., 2018
[Bostrom, 2014] Bostrom, N. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.
[Brundage, 214] Brundage, M. 2014. Limitations and Risks of Machine Ethics. *Journal of Experimental & Theoretical Artificial Intelligence* 26(3): 355–372.
[Casanovas et al., 2016] P. Casanovas et al., Special Issue on the Semantic Web for the Legal Domain. Editorial: The Next Step. *Semantic Web*, 7(3), 2016
[Casini and Straccia, 2013] G. Casini and U. Straccia. Defeasible Inheritance-Based Description Logics. *Journal of Artificial Intelligence Research*, 48, pp. 415-473, 2013.
[Casini et al., 2019] G. Casini, U. Straccia, T. Meyer. A polynomial Time Subsumption Algorithm for Nominal Safe ELO⊥ under Rational Closure. *Information Sciences*, 501, pp. 588-620, 2019.
[Casini et al., 2015] G. Casini, T. Meyer, K. Moodley, U. Sattler, I. Varzinczak. Introducing Defeasibility into OWL Ontologies. In *Proceedings of the International Semantic Web Conference* (ISWC - 2015), Vol. 2, pp. 409-426, Springer, 2015.
[Etzioni and Etzioni, 2017] Etzioni A. and Etzioni O. 2017. Incorporating Ethics into Artificial Intelligence. *The Journal of Ethics* 21: 403–418.
[Formosa and Ryan, 2021] Formosa, P., Ryan, M. Making moral machines: why we need artificial moral agents. *AI & Society* 36, 839–851 (2021)
[Gabriel 2020] Gabriel, I. 2020. Artificial Intelligence, Values, and Alignment. *Minds and Machines* 30: 411–437
[Horrocks and Sattler 2005] Ian Horrocks and Ulrike Sattler. A tableaux decision procedure for SHOIQ, *Proceedings of IJCAI 2005*, pp. 448-453, Professional Book Center, 2005.
[Jobin et al., 2019] Jobin, A.; Ienca, M.; and Vayena, E. 2019. The Global Landscape of AI Ethics Guidelines. *Nature Machine Intelligence* 1:389–399.
[Makinson and van der Torre, 2000] D. Makinson and L. van der Torre. Input/output logics. *Journal of Philosophical Logic*, 29(4), pp.383–408, August 2000.
[Markovich, 2019] R. Markovich. On the Formal Structure of Rules in Conflict of Laws. *Legal Knowledge and Information Systems. Frontiers in Artificial Intelligence and Applications, Vol 322*, pp. 199 - 204. IOS Press Books, 2019
[Rodrigues et al. 2019] C. M. O. Rodrigues, E. P. da Silva, F. Freitas, I J. da Silva Oliveira, I. Varzinczak. LEGIS: A Proposal to Handle Legal Normative Exceptions and Leverage Inference Proofs Readability. *FLAP*, 6(5), pp. 755-780, 2019.
[Rotolo et al., 2017] A. Rotolo et al., Deliverable D1.1 MIREL Project, Report, 2017, http://www.mirelproject.eu/publications/D1.1.pdf

# Report on the TAILOR connectivity fund supported research visit "Neuro-symbolic integration for graph data"

Manfred Jaeger
Aalborg University

June 13, 2022

## 1 Technical data

**Visitor:** Manfred Jaeger, Aalborg University
**Host:** Andrea Passerini, Trento University
**Duration:** March 14 - April 15, 2022

## 2 Activities during the visit

The objective of the visit was to serve as a seed event that initiates a longer term research program on neuro-symbolic methods for learning and reasoning with graph data. To this end, the activities during the visit consisted of the following main components:

**Development of research agenda:** Andrea and Manfred had nearly daily meetings for knowledge exchange and the development and initiation of concrete research plans. The discussions were mostly based on prior joint work [3, 2], and a recent tutorial article that exhibits the already existing close linkage between graph neural networks and the statistical relational learning framework of relational Bayesian networks (RBNs) [1]. Two specific lines of investigation were initiated:

- Neural network learning of multi-relational slotchain dependencies: in our earlier work [3] we have developed an approach to learn how a target label of an entity depends on attributes of other entities connected by a certain chain of relational connections (e.g., learn that the *thesis topic* of a student depends on the *research area* attribute of professors the student is connected to by the chain of *takes(student, course)*, *teaches(course,professor)* relations). During the visit we developed an adaptation of our method, which was originally developed in the context of statistical relational learning, to graph neural networks (GNNs).
- Integration of RBNs with GNNs: based on preliminary findings detailed in [1] we developed ideas and a proof-of-concept implementation for how GNNs trained in a neural network framework such as Pytorch geometric can be exported as an RBN, and thereby become amenable to a richer class of reasoning tasks than the specialized classification or link prediction tasks that GNNs usually are limited to.

**Recruitment of Master thesis students:** to support the two lines of research described above, two master students at Trento university were recruited who will write their master theses on these respective topics. Both students have procured an ERASMUS+ grant to conduct part of their thesis research as visiting students at Aalborg university. Both also are possible candidates for continuing this line of research as PhD students.

**Further collaboration initiatives:** apart from neuro-symbolic integration as the core area of interest, two other forms of collaborations took shape:

- During Manfred's visit it turned out that there was a strong common interest with Luciano Serafini and Sagar Malhotra from Fondazione Bruno Kessler on the topic of projectivity of statistical relational models. This common interest was explored in an extensive exchange of ideas with a view towards a continuing research collaboration.
- We explored the possibility of participating in a EU grant proposal on the topic of robustness and verification of neural networks in aviation.

# 3 Continuing activities

Since the main objective of the visit was to initiate a longer term research collaboration, the main outcomes of the visit are the continuing activities:

- The two master students working on the topics described above will visit Manfred at Aalborg in the periods June 15 - September 15, and September 15 - December 15, respectively. They will be co-supervised by Andrea and Manfred, and are expected to complete their Master theses in Trento within a short period after their return. It is envisioned that at least one of these students will continue under our co-supervision as a PhD student.
- Luciano, Sagar and Manfred continue to have regular online meetings on the subject of projective models. It is envisioned that Sagar may come to Aalborg as a visiting PhD student, and that this collaboration extends into a post-doc phase of Sagar.

# References

[1] Manfred Jaeger. Learning and reasoning with graph data: Neural and statistical-relational approaches. In *International Research School in Artificial Intelligence in Bergen (AIB 2022)*.

[2] Manfred Jaeger, Marco Lippi, Andrea Passerini, and Paolo Frasconi. Type extension trees for feature construction and learning in relational domains. *Artificial Intelligence*, 204:30–55, 2013.

[3] Marco Lippi, Manfred Jaeger, P. Frasconi, and A. Passerini. Relational information gain. *Machine Learning*, 83:219–239, 2011.

# New uncertainty quantification algorithms in metric spaces with strong theoretical guarantees and computational feasibility

*Short report,*
*Marcos Matabuena, CiTIUS, Universidad de Santiago de Compostela*

**Abstract**

Statistical and machine learning approaches transform industry, healthcare, and the digital marketplace. A key point in applying these data analysis strategies in decision-making is quantifying the limits of trustworthy statistical and machine learning models with an uncertainty analysis. Conformal inference maybe is the most dominant general framework for prediction intervals and obtains a predictive model outputs reliability measure. However, one of the significant open challenges of conformal inference is to build new general methods that remain valid with multivariate responses or even complex ones such as curves or graphs that may arise in modern medical applications of precision medicine. In addition, from a computational point of view, there are critical problems in building the prediction intervals in large and high-dimensional datasets. There is also a lack of solid theoretical results in many settings. The purpose of this article is to try to break this gap in the literature by proposing a new uncertainty quantification method when the response of the regression algorithm lives in a metrics space, and it can handle straightforwardly and efficiently large datasets with new theoretical strong results that it cannot reach with traditional conformal inference approaches. Two algorithm versions are proposed; the first is global and designed to handle the homoscedastic case, while the second is local and works in the heteroscedastic signal noise regime. The potential advantages of the new method are illustrated in the context of the global Fréchet regression model when to the best of our knowledge, any method exists to propose a level set of uncertainty. In this case, we analyze relevant medical examples with complex statistical objects as responses that appear in modern precision and digital medicine problems.

## 1 Introduction

A critical issue when we model many of these real world applications, is the considerable uncertainty in the outputs of a predictive model. In this context, to create trustworthy machine learning models, quantifying this uncertainty is crucial to determine the models' limits and and to specify when we can obtain reliable results [18].

In medicine, it is common that the patient's responses to the same clinical treatment can present a high variability at the individual level [19]. The same large variability arises in many machine learning models' output when we try to predict future patient status in the long term to improve the early diagnosis of diseases and screening campaigns. For example, our recent work on diabetes [11] tries to predict the A1C (glycosilated hemoglobin; the primary diabetes biomarkers to control and diagnose the disease) in five years. However, the considerable uncertainty in the problem does not allow our algorithm to obtain reliable results along the subset of patients, which is crucial to determine the development of diabetes disease accurately. Then, with our uncertainty analysis, we will provide a new stratification of subjects based on their glucose uncertainty, promoting new personalized patient follow-ups with more complex medical tests and new interventions that avoid the development of diabetes. In our work, to develop this modeling goal, we have extended conformal inference techniques in the setting of missing responses, allowing us to obtain prediction intervals of the A1C predictions.

Conformal inference techniques [20] constitute a general and unique framework for measuring the uncertainty of statistical and machine learning models with the estimation of prediction intervals in multiple situations. This research topic has become of great interest in the statistical and machine learning community in recent years. Multiple extensions of the usual techniques have appeared to model causality, missing data, and survival analysis problems [9, 11], and as even dependent-data situations [2]. Despite noteworthy advances in this area, as far we are concerned, there is only one work handling multivariate data [8]. In the case of classification problems, the first contributions to handle the multivariate responses have also been recently proposed (see for example [1]).

A critical aspect when increasing the realism of conformal inference techniques is to provide a local interval of predictions. Achieving this goal requires estimating local variability [7] or running traditional conformal inference algorithms locally [6, 10].

In personalized medicine applications, it is increasingly common to register patients' health with complex statistical objects such as curves to record patients' physiological functions and graphs, to measure the dynamic evolution of the patients' brain connectivity. For example, in our recent work, we have coined the concept of "glucodensity" [13], a new compositional distributional representation of a patient's glucose profiles that improve the existing methodology data analysis in the area. This type of representation was also helpful to obtain better results with accelerometer data [5, 12]. In this context, Fréchet's linear model has recently been proposed to analyze these predictors [15]. However, so far, there is no methodology to provide an interval of predictions in this setting. In addition, the disease definitions often depend on several biomarkers, and it is crucial to estimate the uncertainty about the evolution of patients' conditions, introducing the correlation structure of multiple biomarkers from a multivariate perspective.

There is hardly literature on conformal inference with stochastic processes, and this is restricted to the case of geometry of supremum norm, e.g., in the case of simulation models of epidemics [14] or with euclidean functional data objects [3].

Given the need for predicting several variables simultaneously and even complex objects such as curves and graphs, this research project aims to provide new uncertainity quantification stratifications which are valid for multivariate and even complex data in metrics spaces. A key point in the new uncertainity quantification strategies that we will propose is that they allow to obtain local prediction intervals.

In order to illustrate the need to quantify the uncertainty and provide predictions beyond the conditional mean, Figure 1 shows the levels set of the uncertainty of the linear regression model with the response to the distributional representations of glucose profiles [11]. We show that the conditional mean is constant with an increase in body mass index; however, the uncertainty increases dramatically with a variation of the body-mass index. Unfortunately, to the best of our knowledge, we do not know any prior methodology to address the uncertainty quantification for this problem.

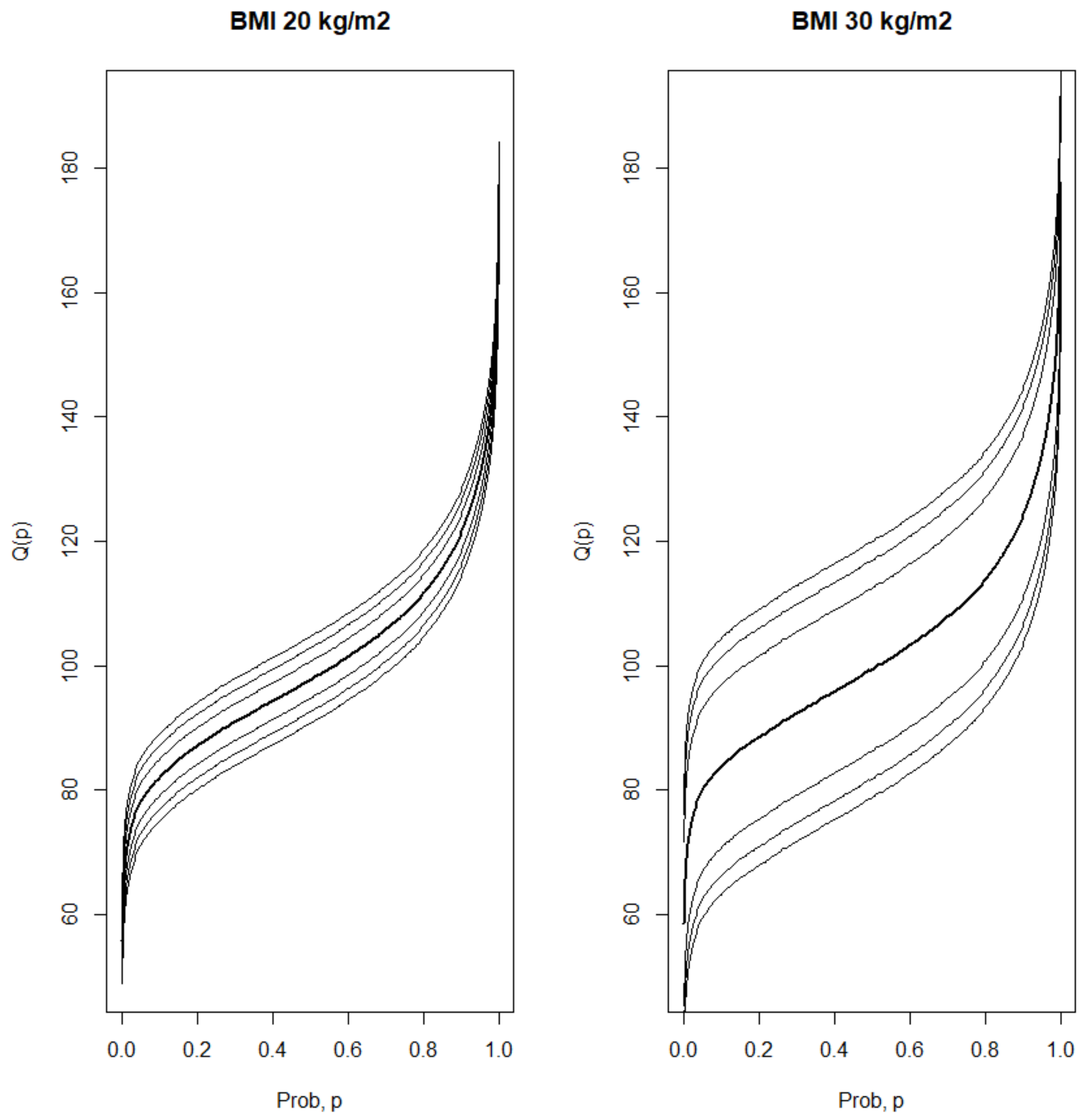


Fig. 1: Levels set of predicting distributional representations of glucose profiles according to our algorithm in two individuals: i) Individual with a body mass index of 20 mg/dL ii) Individual with a body mass index of 30 mg/dL..

## 1.1 Notation and problem definition

Suppose that we observe a random sample i.i.d $\mathcal{D}_n = \{(X_i, Y_i) \in \mathcal{X} \times \mathcal{Y} : 1 \leq i \leq n = n_1 + n_2\}$ from the random variable $(X, Y) \sim P = P_X \otimes P_{Y|X}$, where in our setting $\mathcal{X} = \mathbb{R}^p$, and $\mathcal{Y}$ denote a arbitrary separable metric space equipped with a specific metric $d : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}^+$.

In this paper, to connect the spaces $\mathcal{X}$ and $\mathcal{Y}$, we consider a function $m : \mathcal{X} \to \mathcal{Y}$, that in practice, is the regression function we consider to be estimated. Mathematically, we formally this problem as solve the following M-estimator problem:

$$m(\cdot) = \arg\min_{y \in \mathcal{Y}} M(\cdot, y), \tag{1}$$

where in practice, $m(\cdot) = \arg\min_{y\in\mathcal{Y}} M(\cdot, y) = E(d^2(Y, y)|X = \cdot)$ is the conditional Fréchet mean or the Fréchet median $m(\cdot) = \arg\min_{y\in\mathcal{Y}} M(\cdot, y) = E(d(Y, y)|X = \cdot)$, that constitute the natural mathematical notions of center in métric spaces.

Given a new random point $X_{n+1}$, our modeling goal is to estimate a prediction region of the random response variable $Y$, such that $C^{\alpha}(X_{n+1}) \subset \mathcal{Y}$, mimic the following property of the population bands,

$$C^{\alpha}(X_{n+1}) = \mathcal{B}(m(X_{n+1}), r(X_{n+1})) = \arg \min_{\mathcal{B}(m(X_{n+1}),r):\mathbb{P}(Y\in\mathcal{B}(m(X_{n+1}),r)|X=X_{n+1})\geq 1-\alpha} r, \tag{2}$$

where $\mathcal{B}(m(x), r(x))$ denote a ball of mean $m(x)$, and radius $r(x)$. $\mathbb{P}$ stands over the randomness of both the random pair $(X_{n+1}, Y)$ and the full procedure.

Our final estimator takes the following structure

$$\tilde{C}^{\alpha}(x) = \mathcal{B}(\tilde{m}(x), \tilde{r}(x)) \quad \forall x \in \mathcal{X}, \tag{3}$$

where $\tilde{m}(x)$, and $\tilde{r}(x)$ denote two estimators of the center and radius of the ball from the random sample $\mathcal{D}_n$, that hold the following property

$$\int_{\mathcal{X}} \mathbb{P}\left(Y \in C^{\alpha}(x) \triangle \tilde{C}^{\alpha}(x) | X = x, \mathcal{D}_n\right) P_X(dx) = o_p(1), \tag{4}$$

that is, that we recover in some uniform sense the right optimal region sets.

In the rest of this paper, we will split $\mathcal{D}_n$ in three disjoint subset, $\mathcal{D}_n = \mathcal{D}_{train} \cup \mathcal{D}_{test}$, where $|\mathcal{D}_{train}| = n_1$, $|\mathcal{D}_{test}| = n_2$. We denote the set of indexes $\mathcal{J}_{\mathcal{D}_{train}} := \{i \in \{1, \ldots, n\} : (X_i, Y_i) \in \mathcal{D}_{train}\}$, and $\mathcal{J}_{\mathcal{D}_{test}} := \{i \in \{1, \ldots, n\} : (X_i, Y_i) \in \mathcal{D}_{test}\}$ .

## 1.2 Paper motivation: Quantify the uncertainty with global Fréchet regression model in medical applications

Complex statistical objects such as graphs, strings, probability distributions, compositional data, or other functional data objects appear naturally in recording information in medical and related fields. For example, practitioners can register with new and more sophisticated profiles of patients' conditions using these more complex mathematical constructions, enabling improved and refined clinical decision-making using new, more advanced clinical models with these objects.

A general framework for predicting these complex objects is provided by the recent global Fréchet model, which would be the equivalent in metric spaces to the notion of the regression model in Euclidean spaces. Bellow, we introduce the mathematical details briefly.

Let $(X, Y) \sim P$ be a multivariate random variable, where $X \in \mathbb{R}^p$, and $Y \in \mathcal{Y}$ a separable bounded metric spaces.

For a fixed $X = x$, the population value of the Global-Fréchet model is given by solving the following optimization problem:

$$m(x) = \arg\min_{y \in \mathcal{Y}} E\left[\left[1 + (X - x)\Sigma^{-1}(x - \mu)\right]\left(d^2(Y, y)\right)\right], \tag{5}$$

where $\Sigma = Cov(X, X)$, and $\mu = E(X)$.

Suppose that a random sample i.i.d $\mathcal{D}_n = \{(X_i, Y_i)\}_{i=1}^n$ from a distribution $P$ is available, we can obtain a estimation of conditional mean as follows:

$$\tilde{m}(x) = \arg\min_{y \in \Omega} \frac{1}{n} \sum_{i=1}^{n} \left[1 + (x - X_i)\tilde{\Sigma}^{-1}\left(x - \overline{X}\right) d^2(y, Y_i)\right], \tag{6}$$

where $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$, and $\tilde{\Sigma} = \frac{1}{n-1}\sum_{i=1}^n \left(X_i - \overline{X}\right)^T \left(X_i - \overline{X}\right)$.

The following results characterize the ratios of convergence of a global Fréchet regression model.

**Proposition 1.** *Suppose that, for a fixed $x \in \mathbb{R}^p$ and the conditions s $2-4$ (see Supplemental Material for more details) are hold. Then*

$$d(\tilde{m}(x), m(x)) = O\left(n^{-\frac{1}{2(\beta-1)}}\right). \tag{7}$$

*Furthermore, for a given $B > 0$, if $5-7$ hold,*

$$\sup_{\|x\| \leq B} d(\tilde{m}(x), m(x)) = O\left(n^{-\frac{1}{2(\alpha-1)}}\right), \tag{8}$$

*for any $\alpha' > \alpha$.*

Recently a new random forest was proposed in this setting, extending this M-estimator theory to the context of infinite-order U-statistics [16].

In both cases, to the best of our knowledge, no statistical methodology exists to provide regions of predictions. Bellow, we introduce the three examples that we use to illustrate the potential of our proposal with the Global Fréchet-regression model.

### 1.2.1 Laplacian space graphs

Neuroimaging and related fields are another critical scientific branch to obtain relevant examples that analysis of metric spaces and, in particular, graphs connectivity brain structures can drive promising scientific progress about how the brain works and consequently draw new insights on how we can optimize brain behavior.

In this paper, we consider the space $\mathcal{Y} = (\Omega, d)$ where $\Omega$ is the set of networks with a fixed number, say $r$, of nodes. One can view networks as adjacency matrices, graph Laplacians equipped with the Frobenius metric $d_{\text{FRO}}(A, B) = \left(\sqrt{\sum_{i=1}^r \sum_{j=1}^r (A_{ij} - B_{ij})^2}\right)$ for all $A, B \in \mathcal{Y}$.

Recently, in the setting of the global Fréchet-regression model with this metric, an efficient projection strategy was proposed in [22] to estimate the conditional mean regression function.

Predicting brain graph structures is a fundamental problem in a medical image. Uncertain quantification is a significant step in mathematical modeling, for example, to determine whether cerebral connections vary over different test conditions or stimuli with a certain confidence level and as a consequence of reliability.

### 1.2.2 Probability distributions of physical activity data

In this work, we summarized the information of dynamic wearable time series in their specific representation in the space of univariate probability distributions. In particular, we use data from a continuous glucose monitor and consider the density function as a representation. Our recent work [11] showed the potential advantages of considering this representation in different predictive tasks concerning existing diabetes summary metrics to predict some diabetes biomarkers.

Mathematically, we consider the space $\mathcal{Y} = (\Omega, d)$ where $\Omega$ is the set of univariate probability distributions on a compact support in $T \subset \mathbb{R}$. We choices as metric popular $2$-Wasserstein distance $d_{\mathcal{W}_2}$ that is defined as $d^2_{\mathcal{W}_2}(F,G) = \int_0^1 \left(F^{-1}(t) - G^{-1}(t)\right)^2 dt$, for all $F, G \in \mathcal{Y}$.

The scientific challenge of predicting these representations allows us to elucidate which factors modify glucose profiles in the whole range of both high and low concentrations, i.e., hypo- and hyperglycemia. In this mathematical modeling task, determining the predictive limits of the algorithms is crucial to know how variable the glucose values are with the modification of some variables related to the patients' status.

### 1.2.3 Multivariate Euclidean data

Multivariate Euclidean data is another critical example in real applications, where sufficiently satisfactory methods that scale computationally efficiently and adapt adequately to the local geometry of the data have not yet been provided.

In this example, consider the space $\mathcal{Y} = (\Omega, d)$ where $\Omega = \mathbb{R}^p$, and we use as a metric the Mahalanobis distance that introduce the local geometry of the response variable on the randon variables $Y$ as follows, $d(x,y) = \sqrt{(x-\mu)\Sigma^{-1}(x-\mu)}$, where $\mu = E(Y)$ and $\Sigma = Cov(Y,Y)$.

The example introduced here is also related to Diabetes Mellitus Disease and $p = 2$. Using a large dataset, we try to predict the biomarkers used to control and diagnose the disease FPG together with the physical activity levels of individuals under multivariate response linear models. The new methods have the potential to determine the model's predictive limits in different clinical phenotypes of patients. In case of large uncertainty, we must use more complex models or, if necessary, resort use more complex diabetes biomarkers or medical tests to predict the mentioned variables.

## 1.3 Summary of methodological contributions

We will propose two general uncertainty quantification algorithms in the context of regression models that work in the setting of response $Y$ take values in separable metrics spaces $\mathcal{Y}$. More specifically, we propose two specific algorithms in basis on the signal-noise regime of the basic regression model specified by the function $m$.

1. Homocedastic set-up: We assume that $\mathbb{P}(Y \in \mathcal{B}(m(x), r) | X = x) = \phi(r)$, that is the probability mass of ball is invariant to the point $X = x$ selected.

2. Heterocedastic set-up: We assume that $\mathbb{P}(Y \in \mathcal{B}(m(x), r) | X = x) = \phi(r, x)$, that is the probability mass of ball depend locally point $X = x$ selected.

A primary characteristic of new algorithms proposed is that it works independently for any predictive learning algorithm that, in practice, is specified by the problem of estimating the function $m$.

Our theoretical results are summarized bellow:

1. We introduce a consistency results so in the homoscedastic and heteroscedastic case that we can recover the optimal regions set in some uniform sense:

$$\int_{\mathcal{X}} \mathbb{P}\left(Y \in C^{\alpha}(x) \triangle \tilde{C}^{\alpha}(x) \,|X = x, \mathcal{D}_n\right) P_X(dx) = o_p(1), \tag{9}$$

   The previous result are stronger than those presented in the literature for the univariate response set-up that not conditioned to the random sample $\mathcal{D}_n$ (see for example [4]).

2. In heterocedastic and homocedastic case, fixed $X = x$, according the convergence rate of estimator $\tilde{m}$ selected of function $m$, and the the radius $r(x)$

   we establish the following rate of convergence for the problem of estimated predictive region set

$$E\left(\left|\mathbb{P}\left(Y \in \tilde{C}^{\alpha}(x) \,|X = x, \mathcal{D}_n\right) - (1-\alpha)\right|\right) = \cdots \tag{10}$$

## 2 Our novel uncertainity quantification methods in metrics spaces

### 2.1 Homocedastic case

In this setting, to obtain $\tilde{C}_\alpha(x)$, we propose to use the following two-step algorithm:

**Algorithm 1** Uncertainty quantification algorithm homocedastic set-up

1. Estimate the function $m(\cdot)$, $\tilde{m}(\cdot)$ using the random sample $\{(X_i, Y_i) : i \in \mathcal{J}_{\mathcal{D}_{train}}\}$.
2. For all $i \in \mathcal{J}_{\mathcal{D}_{test}}$, evaluate $\tilde{m}(X_i)$ and define $\tilde{r}_i = d(Y_i, \tilde{m}(X_i))$.
3. Estimate the empirical distribution $\hat{F}_{n_3}(t) = \frac{1}{n_3}\sum_{i\in\mathcal{J}_{\mathcal{D}_{test}}} 1\{\tilde{r}_i \leq t\}$ and denote by $\tilde{q}_{1-\alpha}$ the empirical quantile of level $1-\alpha$.
4. Return as estimation of region band $\tilde{C}_\alpha(x) = \mathcal{B}(\tilde{m}(x), \tilde{q}_{1-\alpha})$

Bellow introduces some technical assumptions that guarantee that the method is asymptotically consistent.

**Assumption 1.** *Suppose that the following hold:*

1. $\{(X_i, Y_i)\}_{i\in\mathcal{J}_{\mathcal{D}_{test}}}$ *is* $iid$ *and* $|\mathcal{J}_{\mathcal{D}_{test}}| \to \infty$.
2. $\mathbb{P}(Y \in \mathcal{B}(m(x), r) | X = x) = \phi(r)$.
3. $\tilde{m}$ *is a consistent estimator in the sense that* $E(d(\tilde{m}(X), m(X)) | \mathcal{D}_{train}) \to 0$, *in probability as* $|\mathcal{J}_{\mathcal{D}_{train}}| \to \infty$, *that is, the next random variable hold in probability*

$$\int_{\mathcal{X}} d(m(x), \hat{m}(x)) P_X(dx) = o_p(1). \tag{11}$$

**Lemma 2.** *Assume that Assumptions 1 are hold. Then*

$$\frac{1}{|\mathcal{J}_{\mathcal{D}_{test}}|}\sum_{i\in\mathcal{J}_{\mathcal{D}_{test}}} |d(Y_i, m(X_i)) - d(Y_i, \tilde{m}(X_i))| = o_p(1) \tag{12}$$

*and*

$$\sup_{v\in\mathbb{R}^+} \left|\tilde{G}^*(v) - G^*(v)\right| = o_p(1) \tag{13}$$

*where* $\tilde{G}^*(v) = \frac{1}{|\mathcal{J}_{\mathcal{D}_{test}}|}\sum_{i\in\mathcal{J}_{\mathcal{D}_{test}}} 1\{d(Y_i, \tilde{m}(X_i)) \leq v\}$, $G^*(v) = \frac{1}{|\mathcal{J}_{\mathcal{D}_{test}}|}\sum_{i\in\mathcal{J}_{\mathcal{D}_{test}}} 1\{d(Y_i, m(X_i)) \leq v\}$.

*Proof.* Observe that

$$\frac{1}{|\mathcal{J}_{\mathcal{D}_{test}}|}\sum_{i\in\mathcal{J}_{\mathcal{D}_{test}}} |d(Y_i, m(X_i)) - d(Y_i, \tilde{m}(X_i))| \leq \frac{1}{|\mathcal{J}_{\mathcal{D}_{test}}|}\sum_{i\in\mathcal{J}_{\mathcal{D}_{test}}} |d(m(X_i), \tilde{m}(X_i))|, \tag{14}$$

where we infer trivially that

$$\frac{1}{|\mathcal{J}_{\mathcal{D}_{test}}|}\sum_{i\in\mathcal{J}_{\mathcal{D}_{test}}}|d(Y_i,m(X_i))-d(Y_i,\tilde{m}(X_i))|=o_p(1)$$

.

For the second part, define quantities $R_T=\sup_{v\in\mathbb{R}}|G^*(v)-G(v)|$ and $W=\sup_{x_1\neq x_2}\frac{|G(x_1)-G(x_2)|}{|x_1-x_2|}$.
Let $A=\{i\in\mathcal{J}_{\mathcal{D}_{test}}:|d(Y_i,\tilde{m}(X_i)-d(Y_i,m(X_i)))|\geq\delta\}$.
Fix $x\in\mathbb{R}$. Then

$$\begin{aligned}(|\mathcal{J}_{\mathcal{D}_{test}}|)\left[\tilde{G}^*(x)-G^*(x)\right]\leq&\left|\sum_{i\in A}(1\{d(Y_i,\tilde{m}(X_i)\leq x)\}-1\{d(Y_i,m(X_i))\leq x)\})\right|+ &(15)\\ &\left|\sum_{i\in A^c}(1(\{d(Y_i,\tilde{m}(X_i)\leq x))\}-1\{d(Y_i,m(X_i))\leq x)\})\right|\leq &(16)\\ &|A|+\left|\sum_{i\in A^c}(1(\{d(Y_i,\tilde{m}(X_i)\leq x))\}-1\{d(Y_i,m(X_i))\leq x)\})\right|, &(17)\end{aligned}$$

where the last inequality follows by the fact the difference of two indicators take values $\{-1,1,0\}$. We must note that for $i\in A^c$, $d(Y_i,m(X_i)\leq x)-\delta\leq d(Y_i,\tilde{m}(X_i)\leq x)\leq d(Y_i,\tilde{m}(X_i)\leq x)+\delta$. Therefore,

$$\sum_{i\in A^c}1\{d(Y_i,m(X_i))\leq x-\delta\}\leq\sum_{t\in A^c}1\{d(Y_i,\tilde{m}(X_i))\leq x\}\leq\sum_{i\in A^c}1\{d(Y_i,m(X_i))\leq x+\delta\} \tag{18}$$

Since, $\sum_{i\in A^c}1\{d(Y_i,m(X_i))\leq x-\delta\}\leq\sum_{t\in A^c}1\{d(Y_i,m(X_i))\leq x\}\leq\sum_{t\in A^c}1\{d(Y_i,m(X_i))\leq x+\delta\}$, it is follow that

$$\begin{aligned}&\left|\left(\sum_{i\in A^c}1\{d(Y_i,\tilde{m}(X_i))\leq x\}\right)-\left(\sum_{i\in A^c}1\{d(Y_i,m(X_i))\leq x\}\right)\right| &(19)\\ &\leq|(|\mathcal{J}_{\mathcal{D}_{test}}|)\left[\tilde{G}^*(x+\delta)-G^*(x-\delta)\right]||- &(20)\\ &\left(\left(\sum_{i\in A^c}1\{d(Y_i,m(X_i))\leq x+\delta\}\right)-\left(\sum_{i\in A^c}1\{d(Y_i,m(X_i))\leq x-\delta\}\right)\right) &(21)\\ &\leq(|\mathcal{J}_{\mathcal{D}_{test}}|(G(x+\delta))-G(x-\delta)+2R_T))+|A| &(22)\\ &\leq|\mathcal{J}_{\mathcal{D}_{test}}|(2\delta W+2R_T)+|A|. &(23)\end{aligned}$$

Using the previous inequality, we can show

$$|\mathcal{J}_{\mathcal{D}_{test}}|\left|\tilde{G}^*-\tilde{G}(x)\right|\leq|\mathcal{J}_{\mathcal{D}_{test}}|(2\delta W+2R_T)+|A| \tag{24}$$

Since the right-hand does not depend on $x$, we have that

$$\sup_{x\in\mathbb{R}}\left|\tilde{G}^{*}(x)-\tilde{G}(x)\right|\leq\frac{2}{|A|}\mathcal{J}_{\mathcal{D}_{test}}+(2\delta W+2R_T) \tag{25}$$

and we can bound the behavior $|A|$ using the fact that $\frac{1}{|\mathcal{J}_{\mathcal{D}_{test}}|}\sum_{i\in\mathcal{J}_{\mathcal{D}_{test}}}|d(Y_i,m(X_i))-d(Y_i,\tilde{m}(X_i))|\leq\frac{1}{|\mathcal{J}_{\mathcal{D}_{test}}|}\sum_{i\in\mathcal{J}_{\mathcal{D}_{test}}}|d(m(X_i),\tilde{m}(X_i))|$.

Using again the triangle inequality, we have:

$$\sup_{x\in\mathbb{R}}\left|\tilde{G}^{*}(x)-G(x)\right|\leq\sup_{x\in\mathbb{R}}\left|\tilde{G}^{*}(x)-\tilde{G}(x)\right|+R_T \tag{26}$$

and we conclude the results using the assumptions introduced.

□

**Lemma 3.** *For any $\alpha\in(0,1)$, as $|\mathcal{J}_{\mathcal{D}_{test}}|\to\infty$., $\tilde{q}_{1-\alpha}\to q_{1-\alpha}$.*

*Proof.* As $\tilde{G}^{*}(q_{1-\alpha})=G(q_{1-\alpha})+o_p(1)$, it is consequence of strong law of large numbers. □

**Theorem 4.** *Assume that Assumptions 1 are hold. Then, $\tilde{C}_{\alpha}(\cdot)$ estimated with the Algorithm 1 hold*

$$\int_{\mathcal{X}}\mathbb{P}\left(Y\in C^{\alpha}(x)\triangle\tilde{C}^{\alpha}(x)|X=x,\mathcal{D}_n\right)P_X(dx)=o_p(1).$$

*Proof.* For each $x\in\mathcal{X}$, we define $\tilde{C}^{\alpha}_{m}(x)=\mathcal{B}(m(x),\tilde{q}_{1-\alpha})$ and $\tilde{C}^{\alpha}_{q}(x)=\mathcal{B}(\tilde{m}(x),q_{1-\alpha})$.

For a fixed $x\in\mathcal{X}$ by the properties of the metric induced by the symmetric difference of two sets, it is hold that

$$\mathbb{P}\left(Y\in C^{\alpha}(x)\triangle\tilde{C}^{\alpha}(x)|X=x,\mathcal{D}_n\right)\leq \tag{27}$$

$$\mathbb{P}\left(Y\in C^{\alpha}(x)\triangle\tilde{C}^{\alpha}_{m}(x)|X=x,\mathcal{D}_n\right)+\mathbb{P}\left(Y\in\tilde{C}^{\alpha}_{m}(x)\triangle\tilde{C}^{\alpha}(x)|X=x,\mathcal{D}_n\right)+ \tag{28}$$

$$\mathbb{P}\left(Y\in C^{\alpha}(x)\triangle\tilde{C}^{\alpha}_{q}(x)|X=x,\mathcal{D}_n\right)+\mathbb{P}\left(Y\in\tilde{C}^{\alpha}_{q}(x)\triangle\tilde{C}^{\alpha}(x)|X=x,\mathcal{D}_n\right). \tag{29}$$

We will show:

$$\int_{\mathcal{X}}\mathbb{P}\left(Y\in C^{\alpha}(x)\triangle\tilde{C}^{\alpha}_{m}(x)|X=x,\mathcal{D}_n\right)P_X(dx)=o_p(1), \tag{30}$$

$$\int_{\mathcal{X}}\mathbb{P}\left(Y\in\tilde{C}^{\alpha}_{m}(x)\triangle\tilde{C}^{\alpha}(x)|X=x,\mathcal{D}_n\right)P_X(dx)=o_p(1), \tag{31}$$

$$\int_{\mathcal{X}}\mathbb{P}\left(Y\in C^{\alpha}(x)\triangle\tilde{C}^{\alpha}_{q}(x)|X=x,\mathcal{D}_n\right)P_X(dx)=o_p(1), \tag{32}$$

$$\int_{\mathcal{X}}\mathbb{P}\left(Y\in\tilde{C}^{\alpha}_{q}(x)\triangle\tilde{C}^{\alpha}(x)|X=x,\mathcal{D}_n\right)P_X(dx)=o_p(1), \tag{33}$$

To do this, we analyze the four terms separately.

**Case 1:**

We define $q^{m}_{1-\alpha}$ as the empirical quantile of the empirical distribution $G^{*}(v) = \frac{1}{|\mathcal{J}_{\mathcal{D}_{test}}|}\sum_{i\in\mathcal{J}_{\mathcal{D}_{test}}} 1\{d(Y_i, m(X_i)) \leq v\}$ and related with the ball $\tilde{C}^{\alpha}_{q_m}(x)$. Then

$$\mathbb{P}\left(Y \in C^{\alpha}(x) \triangle \tilde{C}_{m^{\alpha}}(x) \,|X = x, \mathcal{D}_n\right) \tag{34}$$
$$= \mathbb{P}\left(\{d(Y, m(x)) > q_{1-\alpha}, d(Y, m(x)) \leq \hat{q}_{1-\alpha}\} \,|X = x, \mathcal{D}_n\right) + \tag{35}$$
$$\mathbb{P}\left(\{d(Y, m(x)) \leq q_{1-\alpha}, d(y, m(x)) > \hat{q}_{1-\alpha}\} \,|X = x, \mathcal{D}_n\right) \tag{36}$$
$$\tag{37}$$
$$\leq \mathbb{P}\left(Y \in C^{\alpha}(x) \triangle \tilde{C}^{\alpha}_{q_m}(x) \,|X = x, \mathcal{D}_n\right) + \mathbb{P}\left(Y \in \tilde{C}^{\alpha}_{q_m}(x) \triangle \tilde{C}^{\alpha}_{m}(x) \,|X = x, \mathcal{D}_n\right) = \tag{38}$$
$$\mathbb{P}\left(\{d(Y, m(x)) > q_{1-\alpha}, d(Y, m(x)) \leq q^{m}_{1-\alpha}\}\big|\, X = x, \mathcal{D}_n\right) \tag{39}$$
$$+\mathbb{P}\left(\{d(Y, m(x)) \leq q_{1-\alpha}, d(y, m(x)) > q^{m}_{1-\alpha}\}\big|\, X = x, \mathcal{D}_n\right) \tag{40}$$
$$\mathbb{P}\left(\left\{d(Y, m(x)) > \hat{q}_{1-\alpha}, d(Y, m(x)) \leq q^{m}_{1-\alpha}\right\} |X = x, \mathcal{D}_n\right) \tag{41}$$
$$+\mathbb{P}\left(\{d(Y, m(x)) \leq \hat{q}_{1-\alpha}, d(Y, m(x)) > q^{m}_{1-\alpha}\} X = x, \mathcal{D}_n\right) \leq \tag{42}$$
$$2\left|G\left(q^{m}_{1-\alpha}\right) - G(q_{1-\alpha})\right| + 2\left|G(\hat{q}_{1-\alpha}) - G\left(q^{m}_{1-\alpha}\right)\right|. \tag{43}$$

Therefore

$$\int_{\mathcal{X}} \left|G\left(q^{m}_{1-\alpha}\right) - G(q_{1-\alpha})\right| + 2\left|G(\hat{q}_{1-\alpha}) - G\left(q^{m}_{1-\alpha}\right)\right| P_X(dx) = o_p(1). \tag{44}$$

**Case 2:**

$$\mathbb{P}\left(Y \in \tilde{C}^{\alpha}_{m}(x) \triangle \tilde{C}^{\alpha}(x) \,|X = x, \mathcal{D}_n\right) = \mathbb{P}\left(\{d(Y, m(x)) > \hat{q}_{1-\alpha}, d(y, \hat{m}(x)) \leq \hat{q}_{1-\alpha}\}|X = x, \mathcal{D}_n\right) \tag{45}$$
$$+\mathbb{P}\left(\{d(Y, m(x)) \leq \hat{q}_{1-\alpha}, d(y, \hat{m}(x)) > \hat{q}_{1-\alpha}\}|\, X = x, \mathcal{D}_n\right) \leq \tag{46}$$
$$\mathbb{P}\left(\{d(Y, \hat{m}(x)) \leq \hat{q}_{1-\alpha} < d(\hat{m}(x), m(x)) + d(Y, \hat{m}(x))\}|X = x, \mathcal{D}_n\right) \tag{47}$$
$$\mathbb{P}\left(\{d(Y, m(x)) \leq \hat{q}_{1-\alpha} < d(\hat{m}(x), m(x)) + d(y, m(x))\}|\, X = x, \mathcal{D}_n\right) \tag{48}$$
$$\leq G^{*}(\hat{q}_{1-\alpha} + d(\hat{m}(x), m(x))) - G^{*}(\hat{q}_{1-\alpha}) + G(\hat{q}_{1-\alpha}) + d(\hat{m}(x), m(x)) - G(\hat{q}_{1-\alpha}), \tag{49}$$

**Case 3:**

$$\mathbb{P}\left(Y \in C^{\alpha}(x) \triangle \tilde{C}^{\alpha}_{q}(x) \,|X = x, \mathcal{D}_n\right) = \mathbb{P}\left(\{d(Y, m(x)) > q_{1-\alpha}, d(Y, \hat{m}(x)) \leq q_{1-\alpha}\}|\, X = x, \mathcal{D}_n\right) \tag{50}$$
$$+\mathbb{P}\left(\{d(Y, m(x)) \leq q_{1-\alpha}, d(y, \hat{m}(x)) > q_{1-\alpha}\}|\, X = x, \mathcal{D}_n\right) \tag{51}$$

that it is similar to the prior case except that we exchange the role $\hat{q}_{1-\alpha}$ by $q_{1-\alpha}$. Therefore, we have

$$\mathbb{P}\left(Y \in C^{\alpha}(x) \triangle \tilde{C}^{\alpha}_{q}(x) \,|X = x, \mathcal{D}_n\right) \leq \tag{52}$$
$$\leq G^{*}(q_{1-\alpha} + d(\hat{m}(x), m(x))) - G^{*}(q_{1-\alpha}) + G(q_{1-\alpha} + d(\hat{m}(x), m(x))) - G(q_{1-\alpha}) \tag{53}$$

and as a consequence

$$\int_{\mathcal{X}} \left[G^* \left(q_{1-\alpha} + d\left(\hat{m}\left(x\right), m\left(x\right)\right)\right) - G^*\left(q_{1-\alpha}\right) + G\left(q_{1-\alpha} + d\left(\hat{m}\left(x\right), m\left(x\right)\right)\right) - G\left(q_{1-\alpha}\right) P_X\left(dx\right)\right] = o_p\left(1\right). \quad (54)$$

**Case 4:**

$$\mathbb{P}\left(Y \in \tilde{C}_q^{\alpha}\left(x\right) \triangle \tilde{C}^{\alpha}\left(x\right) | X = x, \mathcal{D}_n\right) = \mathbb{P}\left(\left\{y : d\left(y, \hat{m}\left(x\right)\right) > q_{1-\alpha}, d\left(y, \hat{m}\left(x\right)\right) \leq \hat{q}_{1-\alpha}\right\} | X = x, \mathcal{D}_n\right) \quad (55)$$
$$+\mathbb{P}\left(\left\{y : d\left(y, \hat{m}\left(x\right)\right) \leq q_{1-\alpha}, d\left(y, \hat{m}\left(x\right)\right) > \hat{q}_{1-\alpha}\right\} | X = x, \mathcal{D}_n\right), \quad (56)$$

that is similar to Step $1$, the main diference is consider the random variable $d\left(Y, \tilde{m}\left(X\right)\right) | X = x$ instead of random variable $d\left(Y, m\left(X\right)\right) | X = x$. Therefore we have

$$\mathbb{P}\left(Y \in \tilde{C}_q^{\alpha}\left(x\right) \triangle \tilde{C}^{\alpha}\left(x\right) | X = x, \mathcal{D}_n\right) \leq 2|G^*(q_{1-\alpha}^m) - G^*(q_{1-\alpha})| + 2|G^*(\hat{q}_{1-\alpha}) - G^*(q_{1-\alpha}^m)| = 4 o_p\left(1\right). \quad (57)$$

Combining the prior results, we have

$$\int_{\mathcal{X}} \mathbb{P}\left(Y \in C^{\alpha}\left(x\right) \triangle \tilde{C}^{\alpha}\left(x\right) | X = x, \mathcal{D}_n\right) P_X\left(dx\right) = o_p\left(1\right).$$

□

| | Age | BMXWAIST | Diastolic Blood Pressure | Systolic blood pressure | Glucose |
|---|---|---|---|---|---|
| 1 | 66 | 97.70 | 68 | 130 | 87 |
| 2 | 73 | 101.90 | 82 | 118 | 95 |
| 3 | 53 | 101.00 | 64 | 116 | 83 |
| 4 | 67 | 101.20 | 68 | 114 | 160 |
| 5 | 31 | 134.40 | 76 | 142 | 95 |
| 6 | 79 | 88.70 | 76 | 134 | 147 |

## 2.2 Heteroscedasticity case

In this setting, in order to obtain a estimator $\tilde{C}_\alpha(x)$, we propose to use the following two-step algorithm that, that unlike to Algorithm 1, we make a local approximation of the radius of the ball with k-nearest neighbors algorithm.

**Algorithm 2** Uncertainty quantification algorithm heterocedastic set-up

1. Estimate the function $m(\cdot)$, $\tilde{m}(\cdot)$ using the random sample $\{(X_i, Y_i) : i \in \mathcal{J}_{\mathcal{D}_{train}}\}$.

2. For a fixed $X = x$, denote by $X_{(1,n_2)}(x), \cdots, X_{(n_2,n_2)}(x)$ the orders elements in decreasing order by a particular distance to the point $x$. Evaluate $\tilde{m}(x)$, and for all $i \in \mathcal{J}_{\mathcal{D}_{train2}}$, define $\tilde{r}_i = d(Y_i, \tilde{m}(x))$. We denote by $\tilde{r}_{(i,n_2)}(x)$ the pseudo-residuals related with the $i$-order-observation. Define by $\hat{F}_{n_2,k}(x,t) = \frac{1}{k}\sum_{i=1}^{k} 1\{\tilde{r}_{(i,n_2)}(x) \leq t\}$ the empirical conditional distributional and denote by $\tilde{q}_{1-\alpha}(x)$ the empirical quantile.

3. For a fixed $k$, return as estimation of prediction region $\tilde{C}_\alpha^k(x) = \mathcal{B}(\tilde{m}(x), \tilde{q}_{1-\alpha}(x))$.

# 3 Our model in global Fréchet linear regresion model

## 3.1 Multivariate value-responses

In this example, consider the space $\mathcal{Y} = (\Omega, d)$ where $\Omega = \mathbb{R}^p$, and we use as a metric the Mahalanobis distance that introduce the local geometry of the response variable on the randon variables $Y$ as follows, $d(x,y) = \sqrt{(x-\mu)\Sigma^{-1}(x-\mu)}$, where $\mu = E(Y)$ and $\Sigma = Cov(Y,Y)$.

The example introduced here is also related to Diabetes Mellitus Disease and $p = 2$. Using a large dataset of more than 5000 patients, we try to predict the biomarkers used to control and diagnose the disease A1C (glycosylated hemoglobin) together with the physical activity levels of individuals under multivariate response linear models. For this purpose, we consider a bivariate regression model whose response is the mentioned A1C and TAC, and the predictors are age, waist, diastolic Blood Pressure, systolic blood pressure, and glucose. Figure 2 shows the levels set for six patients with basis protectors are found in Table 3.1- Generally, we observe that if the glucose is altered, the uncertainty increases.

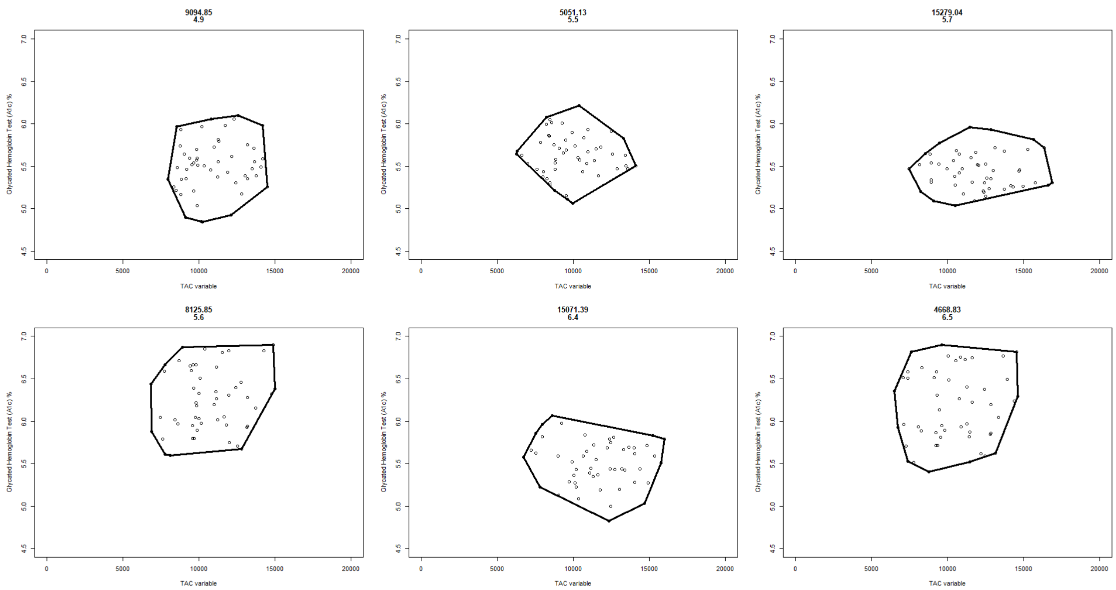


Fig. 2: Prediction region interval from six patients in the total activity time variable (TAC) and glycosilated haemoglobin (A1C).

## 3.2 Probability distribution with $2$-Wassertien metrics in diabetes example

Using non-diabetes individuals from [21], we estimate for the first time using the new distributional representations the expected glucose confidence at a level of 80 percent of confidence by age groups. Figure 3 shows that the mean glucose values do not modify. However, the width of the bands yes.

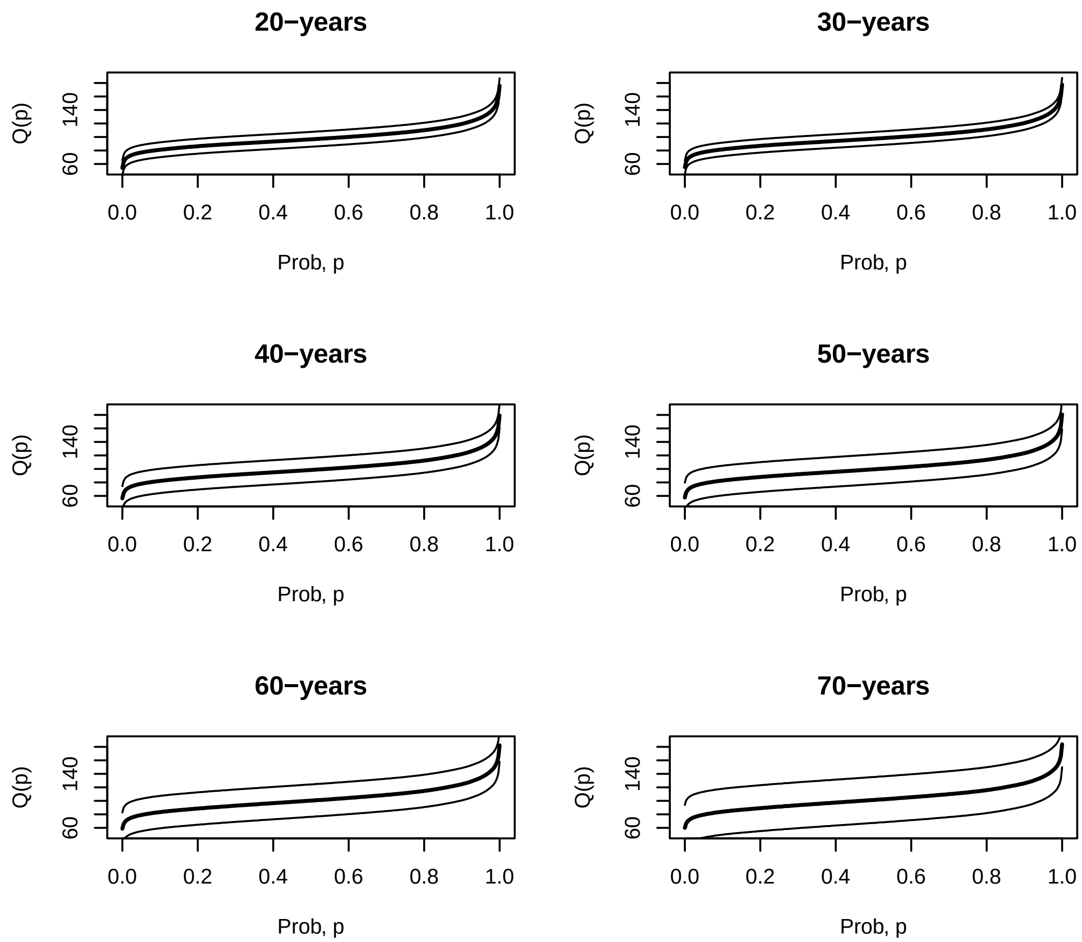


Fig. 3: Left: Expected glucose profiles by age

.

## 3.3 Graphs with Laplacian metrics

Using the first schizophrenia dataset from [17], we fitted a conditional fréchet model with a Laplacian metric by mental status and applied the uncertainty quantification algorithm. In this case, we use the homoscedastic algorithm. Figure 4 shows the Laplacian matrix's different conditional mean and uncertainty estimations of the mentioned matrix.

# 4 Discussion and final comments

In this work, to define the optimal uncertainty region set, we define the mathematical problem using balls covering specific probabilities and minimizing the diameter. However, many other geometrical sets for this purpose can be considered. The new methods provide more robust results than classical conformal inference algorithms with the cost that we do lose the theoretical guarantee of exact control with finite samples. In addition, the new methods are computational efficiently and can handle large datasets.

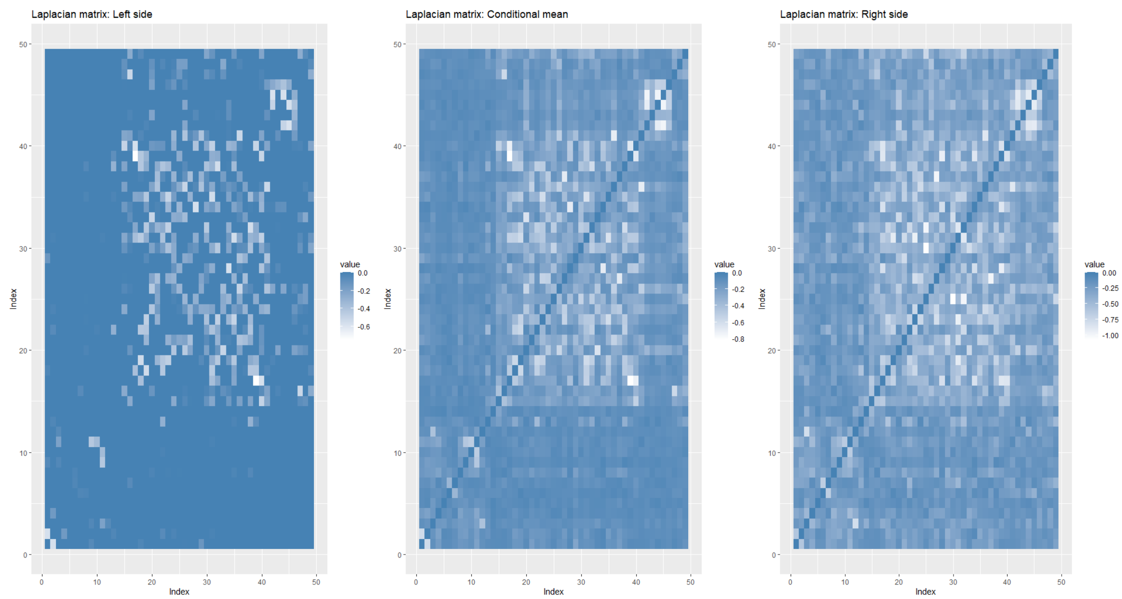


Fig. 4: Left: The CGM recording from a normoglycemic patient. Right: The corresponding glucodensity.

## Appendix A Technical conditions global Fréchet model

Define the following quantity:

$$M(\omega, x) := E\left(\left[1 + (X - x)\,\Sigma\widetilde{\Sigma}^{-1}(x - \mu)\right]\left(d^2(Y, \omega)\right)\right). \tag{58}$$

For a fixed $x \in \mathbb{R}^p$, in order to guarantee the existence of population conditional mean, the convergence of empirical estimators, and ratios of convergence, we require to introduce the following assumptions.

**Assumption 2.** *The objects $m(x)$ and $\tilde{m}(x)$ exists and are unique, the latter almost surely, that is, for any $\epsilon > 0$, $\inf_{d(\omega, m(x)) > \epsilon} M(\omega, x)$.*

**Assumption 3.**

*Let $B_\delta(m(x)) \subset \Omega$ be the ball of radius $\delta$ centered at $m(x)$ and $N(\epsilon, B_\delta(m(x)), d)$ be its covering number using balls of size $\epsilon$. Then*

$$\int_0^1 \sqrt{1 + log\left[N(\epsilon, B_\delta(m(x)), d)\right]}\,d\epsilon = O(1) \text{ as } \delta \to 0. \tag{59}$$

**Assumption 4.** *There exist $\eta > 0$, $C > 0$ and $\beta > 1$, possibly depending on $x$, such that, whenever $d(m(x), \omega) < \eta$, we have $M(\omega, x) - M(m(x), x) \geq C d(\omega, m(x))^{\beta}$.*

In order to establish results in a uniform sense, we must introduce more strong assumptions. Let $\|\cdot\|$ be the Euclidean norm on $\mathbb{R}^p$ and $B > 0$.

**Assumption 5.** *Almost surely, for all* $\|x\| \leq B$*, the objects* $m(x)$ *and* $\tilde{m}(x)$ *exists and are unique. Additionally, for any* $\epsilon > 0$*,*

$$\inf_{\|x\| \leq B} \inf_{d(\omega, m(x)) > \epsilon} M(\omega, x) - M(m(x), x) > 0, \tag{60}$$

*and there exist* $\gamma = \gamma(\epsilon) > 0$ *such that*

$$P\left(\inf_{\|x\| \leq B} \inf_{d(\omega, m(x)) > \epsilon} M(\omega, x) - M(m(x), x) \geq \gamma\right) = 1 \tag{61}$$

**Assumption 6.**

*With* $B_\delta(m(x))$ *and* $N(\epsilon, B_\delta(m(x)), d)$*,*

$$\int_0^1 \sup_{\|x\| \leq B} \sqrt{1 + log\left[N(\epsilon, B_\delta(m(x)), d)\right]} d\epsilon = O(1) \text{ as } \delta \to 0. \tag{62}$$

**Assumption 7.** *There exist* $\theta > 0$*,* $D > 0$ *and* $\alpha > 11$*, possibly depending on* $B$*, such that,*

$$\inf_{\|x\| \leq B} \inf_{d(m(x), \omega) < \theta} \{M(\omega, x) - M(m(x), x) - Dd(\omega, m(x))^\alpha\} \geq 0 \tag{63}$$

## References


[1] Maxime Cauchois, Suyash Gupta, and John C Duchi. Knowing what you know: valid and validated confidence sets in multiclass and multilabel prediction. *Journal of Machine Learning Research*, 22(81):1–42, 2021.

[2] Victor Chernozhukov, Kaspar Wüthrich, and Yinchu Zhu. An exact and robust conformal inference method for counterfactual and synthetic controls. *Journal of the American Statistical Association*, 0(0):1–16, 2021.

[3] Jacopo Diquigiovanni, Matteo Fontana, and Simone Vantini. Conformal prediction bands for multivariate functional data. *arXiv preprint arXiv:2106.01792*, 2021.

[4] Rina Foygel Barber, Emmanuel J Candes, Aaditya Ramdas, and Ryan J Tibshirani. The limits of distribution-free conditional predictive inference. *Information and Inference: A Journal of the IMA*, 10(2):455–482, 2021.

[5] Rahul Ghosal, Vijay R Varma, Dmitri Volfson, Jacek Urbanek, Jeffrey M Hausdorff, Amber Watts, and Vadim Zipunnikov. Scalar on time-by-distribution regression and its application for modelling associations between daily-living physical activity and cognitive functions in alzheimer's disease. *arXiv preprint arXiv:2106.03979*, 2021.

[6] L GYÖFI and Harro Walk. Nearest neighbor based conformal prediction, 2020.

[7] Jing Lei, Max G'Sell, Alessandro Rinaldo, Ryan J. Tibshirani, and Larry Wasserman. Distribution-free predictive inference for regression. *Journal of the American Statistical Association*, 113(523):1094–1111, 2018.

[8] Jing Lei, James Robins, and Larry Wasserman. Distribution-free prediction sets. *Journal of the American Statistical Association*, 108(501):278–287, 2013.

[9] Lihua Lei and Emmanuel J. Candès. Conformal inference of counterfactuals and individual treatment effects. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, n/a(n/a).

[10] Benjamin LeRoy and David Zhao. Md-split+: Practical local conformal inference in high dimensions. *arXiv preprint arXiv:2107.03280*, 2021.

[11] Marcos Matabuena, Paulo Félix, Carlos Meijide-Garcia, and Francisco Gude. Kernel methods to handle missing responses and their application in modeling five-year glucose changes using distributional representations. `https://github.com/mmatabuena/Marcos-Matabuena-website/blob/main/Matabuena21web.pdf`, 2021.

[12] Marcos Matabuena and Alex Petersen. Distributional data analysis with accelerometer data in a nhanes database with nonparametric survey regression models. *arXiv preprint arXiv:2104.01165*, 2021.

[13] Marcos Matabuena, Alexander Petersen, Juan C Vidal, and Francisco Gude. Glucodensities: A new representation of glucose profiles using distributional data analysis. *Statistical Methods in Medical Research*, 30(6):1445–1464, 2021. PMID: 33760665.

[14] Marcos Matabuena, Pablo Rodríguez-Mier, Carlos García-Meixide, and Victor Leborán. Covid-19: Estimation of the transmission dynamics in spain using a stochastic simulator and black-box optimization techniques. *Computer methods and programs in biomedicine*, 211:106399, 2021.

[15] Alexander Petersen and Hans-Georg Müller. Fréchet regression for random objects with euclidean predictors. *The Annals of Statistics*, 47(2):691–719, 2019.

[16] Rui Qiu, Zhou Yu, and Ruoqing Zhu. Random forests weighted local fr\'echet regression with theoretical guarantee. *arXiv preprint arXiv:2202.04912*, 2022.

[17] Jesús D Arroyo Relión, Daniel Kessler, Elizaveta Levina, and Stephan F Taylor. Network classification with applications to brain connectomics. *The annals of applied statistics*, 13(3):1648, 2019.

[18] Yaniv Romano, Evan Patterson, and Emmanuel Candes. Conformalized quantile regression. *Advances in Neural Information Processing Systems*, 32:3543–3553, 2019.

[19] André J Scheen. Precision medicine: the future in diabetes care? *Diabetes research and clinical practice*, 117:12–21, 2016.

[20] Glenn Shafer and Vladimir Vovk. A tutorial on conformal prediction. *Journal of Machine Learning Research*, 9(3), 2008.

[21] Viral N Shah, Stephanie N DuBose, Zoey Li, Roy W Beck, Anne L Peters, Ruth S Weinstock, Davida Kruger, Michael Tansey, David Sparling, Stephanie Woerner, et al. Continuous glucose monitoring profiles in healthy nondiabetic participants: a multicenter prospective study. *The Journal of Clinical Endocrinology & Metabolism*, 104(10):4356–4364, 2019.

[22] Yidong Zhou and Hans-Georg Müller. Dynamic network regression. *arXiv preprint arXiv:2109.02981*, 2021.

# Basic Information

**Project title**: Workshop proposal: Imagining the AI landscape after the AI Act
**Period of project**: 13 june 2022
**Period of reporting**: 13 June 2022
**Author(s)**: Francesca Naretto
**Organization**: Scuola Normale Superiore
**Host organization**: Vrije Universiteit Amsterdam ( HHAI 2022 Conference – Hybrid Human Artificial Intelligence conference)

# Public summary

The TAILOR Connectivity fund in analysis was used for organizing the workshop: "Imagining the AI landscape after the AI Act" at the first conference on Hybrid Human Artificial Intelligence.
The workshop started with a short welcome moment, in which the organizers presented the schedule of the day. Then, there were the two invited speakers: first Prof. Virginia Dignum and then Prof. Mireille Hildebrandt. Regarding Prof. Dignum, she presented different perspectives regarding the AI Act, focusing on several limitations that she still sees in this Act. Prof. Dignum started her presentation by asking an interesting question: what does the AI Act represent for people who work with AI? This is a particularly difficult question to answer since for now we do not even have an agreement on the definition of AI. Hence, Prof. Dignum started to reason about this question considering the fact that the AI Act focuses mostly on data and the users: in fact, the problems considered all originated from the data, which may be dirt and incorrect. From them, bias and discrimination may arise, creating several problems. Following this line of thinking, she categorized the prohibitions in the AI Act into three categories: (i) the use of social scoring, (ii) distortion of user behavior, and (iii) biometric identification. She also pointed out a wide number of limitations in the AI Act, such as the lack of sustainability and power consumption, as well as difficulty in allowing innovation and development.
The second talk was from prof. Hildebrandt. In this case, the keynote was a series of questions, all regarding the AI Act, collected from the participants. Involving them made the discussion even more interesting, raising several questions but also possible solutions. Overall, Prof. Hildebrandt has an optimistic view: she reckons that implementing all the requirements in the AI Act is a difficult task, but finding the best level of abstraction and, with time, it is feasible.
After the invited speakers, we moved to the presentation of the accepted papers. We split the presentations of the works into two sessions: Session 1, in the morning, was titled "Technical Aspects of the AI Act" in which the papers focused on the technicalities required by the AI Act, as well as the requirements for the individuals; while the Session 2, in the afternoon, was titled "Ethical and Legal Aspects about the AI Act" and hence included contributions focused more on the ethical and legal problems of the AI Act. At the end of every session we left some time for open mike, in which the participants were able to talk more about the contributions presented.
Between the two sessions of papers, we also organized a group activity held by Dr. Tommaso Turchi. The group activity is called "MiniCoDe", which stands for "MINImize algorithmic bias in COllaborative decision making with DEsign fiction". It is a project lead by Prof. Alessio Malizia, in which the aim is to tackle social injustice in future algorithmic-based decision-making applications. This project was already used in several different context, both

in academic settings and in companies. In our case, the participants were split into small groups and were asked to think about a possible scenario which involved Artificial Intelligence as well as ethical problems. At the beginning, the participants were asked to reason first alone and then as a group, with the goal of proposing a solution. In particular, the group activity was a structured brainstorming around how to implement a process/methodology to be compliant with Art.14 on Human Oversight. More specifically, participants were presented with a fictional narrative describing how postcode bias might lead to discrimination against the poor. This type of bias is more subtle compared to other types of biases such as gender or race bias, so enabling human oversight is more difficult. The discussions allowed participants to have a deeper understanding of the implications of the AI Act and EU digital policies.

The proposed workshop found great interest from researchers in a variety of fields, from computer science to law, psychology and economics. We received 17 submissions and we accepted, through a peer reviewed process, 11 contributions, with regular papers, short papers, and extended abstracts (more details about the accepted contributions can be found at http://iail2022.isti.cnr.it/#program). Among them, 3 were regular papers (i.e., 12+ pages), 6 short papers, and 2 abstracts. The abstract can contain preliminary or already published work, while papers must contain original work.

Overall, we are confident that our workshop was successful. The participants at the workshop were 30-35 we are particularly happy to have gathered 11 contributions, all of which were extremely interesting and multidisciplinary, but most of all to have initiated very interesting discussions on the topic, which we are confident will bring new contributions in this area. Due to the success of this first workshop, also demonstrated by the several grateful emails, we have received, we hope to have a workshop again in the coming year.

## Research objectives

*Maximum 1 page.*

### Objectives

Scientific high-level description of the scientific goals of this research visit or workshop. Which scientific challenge are you addressing and why is it important.

The workshop's main goal was to help the community understand and reason over the implications of an AI regulation: what problems does it solve, what problems does it not solve, what problems does it cause, and propose new approaches that solve the new challenges.We decided to conduct this workshop with the objective of collecting ideas regarding how new AI regulation will shape the AI technologies of the future. We aimed to collect contributions regarding how to operationalize the AIA requirements and how to guarantee privacy, fairness and explainability. In particular, our most significant concerns regarded how to guarantee individual rights while achieving the requirements stated in the AIA, and how to assess the AI risk. This is due to the fact that the AI Act considers different classes of AI systems, depending on how risky the Ai under analysis is. But the AIA does not clearly state how to assess an AI system's risk. In addition, in the AIA, several ethical values are cited and considered, but also in this case it is not clear how to guarantee them while preserving the performance of the AI model. In practice, we were afraid that the technologies available right now were not enough to fulfill the general and high-level requirements proposed in the AIA.

### Impact

Please describe the expected impact of this work on society.

Firstly, our goal was to bring together legal experts, tech experts and other interested stakeholders for constructive discussions. Hence, we aimed at having a multidisciplinary setting in which persons with different backgrounds and levels of expertise were brought together and encouraged to share different points of view on the subject of the AI Act. For this reason, in our workshop schedule, we planned several moments of discussion. Before the workshop, we collected the participants' questions to ask Prof. Hildebrandt, intending to conduct a discussion with her on the themes most important for the audience. Then, we encouraged people to ask questions at the end of each keynote speaker, as well as after every paper presentation. Lastly, we organized a moment of open mike and a group activity to encourage exchanging ideas and constructive discussion. In addition, we aimed at stakeholder and geographical balance. These goals dictated the choice of hosting the workshop in the Hybrid Human Artificial Intelligence conference. In fact, this conference was multidisciplinary, with special attention to the human in interaction with the AI systems and the ethical aspects related to this subject, such as guaranteeing fairness, protecting privacy, and ensuring accountability while providing secure and sustainable AI systems.

## Technical approach

*Maximum 1 page.*

### Detailed description

Please detail the technical approach that you followed during your research visit. Include a comparison with the state of the art. For workshops, describe the performed work as well as possible.

We received 17 submissions, and we accepted, through a peer-reviewed process, 11 contributions, with regular papers, short papers, and extended abstracts (more details about the accepted contributions can be found at http://iail2022.isti.cnr.it/#program). Among them, 3 were regular papers (i.e., 12+ pages), 6 short papers, and 2 abstracts. The abstract can contain preliminary or already published work, while papers must contain original work. These contributions will be published in the proceedings of Ceur (http://ceur-ws.org/). All accepted papers presented their work. Our workshop was in-person (as the main conference), but allowed a hybrid mode for those who had difficulty being present. Most of the papers were presented in person, creating an excellent dialogue among all participants. We split the presentation of the contributions into 2 sessions: the first one, titled "Technical aspects of AI Act" was more technical, while the second one, titled "Ethical and Legal Aspects about the AI Act" focused more on the several legal and ethical implications of the AIA. In the first session, we had 5 contributions.

The first one, titled "Using Sentence Embeddings and Semantic Similarity for Seeking Consensus when Assessing Trustworthy AI" was presented by Dennis Vetter and dealt with the difficult task of assessing the trustworthiness of the AI systems. This work addressed this problem by considering sentence embeddings and semantic similarity during the consensus phase of the assessment. The contribution was interesting and well presented; the audience asked several questions, which allowed for an in-depth discussion during the coffee break. The second contribution was titled "FutureNewsCorp, or how the AI Act changed the future of news" and was presented by the single author Natali Helberger. In this case, the paper focused more on the technical aspects of news, newspapers and journalists in the new era of AI. The author started her presentation by presenting a possible future scenario, which was quite intriguing and scary, thus providing several constructive discussions. Then, there was

“Federated Learning as an Analytical Framework for Personal Data Management”, a paper presented by Maciej Zuziak. This author has a legal background, but he is now collaborating with computer scientists to merge his knowledge with the more technical aspects of Federated Learning. The paper “The forgotten human autonomy in Machine Learning” was presented by Prof. Oriol Pujol. It considers several limitations of AI systems when dealing with human autonomy. This theme was also a key topic of the conference; thus, it allowed for several comments and discussions. Lastly, Prof. Francesca Carroccia presented the paper “AI Act and Individual Rights: A Juridical and Technical Perspective”, in which she tackled the problem of the gap between the requirements of the AIA and the technical capacities now available.
Moving to the second session of the paper presentation, we first had an extremely interesting presentation by Prof. Marc Anderson for his paper “Some Ethical Reflections on the EU AI Act”. This contribution considered several ethical aspects in relation to the requirements of the AIA, highlighting the limitations now present in this context from a legal point of view. Then, we had Jonne Mas presenting “A Neo-republican Critique of AI ethics”. This presentation considered the ethical problems from a broader perspective with respect to the works considered so far. The themes considered by Jonne were also linked to the next presentation from Prof. Jerome De Cooman of his work titled “Without Any Prejudice? The Antitrust Implication of the AI Act”. Following, Pietro Dunn and Giovanni De Gregorio presented “The Ambiguous Risk-Based Approach of the Artificial Intelligence Act: Links and Discrepancies with Other Union Strategies”: in this case, the AIA was compared against other European regulations. Then, the paper “The Artificial Intelligence Act. A Jurisprudential View” was presented. This paper focuses mostly on the legal aspects of the AIA, dealing with the actual actuation of all the requirements when considering the different European countries and their different needs. Linked to this talk, there was also the last one from Farhana Ferdousi Liza. She presented “Challenges of Enforcing Regulations in Artificial Intelligence Act” in which the problem of enforcing the regulation is tackled from the point of view of the problems that may arise.

## Scientific outcomes

Please detail the outcomes of this work. This includes any novel insight, discoveries, transfer of technology, or other types of knowledge sharing. Publications can be listed below.

The papers accepted to this workshop will be published as proceedings in Ceur (http://ceur-ws.org/).. In the following there is a list of the papers accepted:

1. Without Any Prejudice? The Antitrust Implication of the AI Act, Jerome De Cooman
2. Challenges of Enforcing Regulations in Artificial Intelligence Act — Analyzing Quantity Requirement in Data and Data Governance, Farhana Ferdousi Liza
3. Using Sentence Embeddings and Semantic Similarity for Seeking Consensus when Assessing Trustworthy AI, Dennis Vetter, Jesmin Jahan Tithi, Magnus Westerlund, Roberto V. Zicari and Gemma Roig
4. Federated Learning as an Analytical Framework for Personal Data Management – a proposition paper, Maciej Zuziak, Salvatore Rinzivillo
5. The forgotten human autonomy in Machine Learning, Paula Subías-Beltrán, Oriol Pujol and Itziar de Lecuona
6. AI Act and Individual Rights: A Juridical and Technical Perspective, Costanza Alfieri, Francesca Caroccia and Paola Inverardi
7. The Ambiguous Risk-Based Approach of the Artificial Intelligence Act: Links and Discrepancies with Other Union Strategies, Pietro Dunn and Giovanni De Gregorio
8. Some Ethical Reflections on the EU AI Act, Marc M. Anderson

9. The Artificial Intelligence Act: A Jurisprudential Perspective, Michał Araszkiewicz, Grzegorz J. Nalepa, and Radosław Pałosz

Future plans

Please detail future plans or new collaborations based on this work.
The proposed workshop greatly interested researchers in various fields, from computer science to law, psychology and economics. Overall, we are confident that our workshop was a success: we brought together people from different backgrounds, creating a constructive dialogue, which we are sure will lead to interesting works in the future. In addition, we were able to publish the contributions of the workshop in proceedings, hence more people, also outside the conference audience, will be able to see them and follow-up on these ideas and projects. We also plan to propose the workshop at the next HHAI 2022 conference to provide a follow-up on the several interesting ideas discussed in this workshop.

# Progress against planned goals

*Maximum 1 page. Only for intermediate reporting (projects running longer than 6 months)*

Please provide an indication of your progress towards the stated research goals. What have you already achieved? Are you on track to attain all goals? Did issues occur that made you adapt previous plans?

All the organizers of the workshop are happy to say that the workshop achieved the expected results: our main goal was to create a constructive discussion about the AI Act and the problems and limitations we are going to face when it will be actualized. In particular, we were interested into understanding what is missing in terms of technology for achieving a satisfying assessment of the AI risk and how to protect ethical requirements in this context. The workshop we organized tackled exactly these themes as well as all the publications produced. In addition, another important requirement for us was to have multidisciplinary both in terms of audience and of the publications gathered. This requirement for us was extremely important since the Ai Act will tackle everyday life, involving both legal and ethical experts, as well as different kinds of technicians to achieve the requirements stated in the regulation.

# Self-assessment

*Maximum 1 page. Only for final reporting (the project has finished)*

Please provide your own final assessment of the effective progress against the goals stated in the proposal, according to the following points:
The proposed workshop was a success: several people complimented us, both in person and via e-mails. The hybrid version of the event allowed several people to follow and to present even if they were not able to join us in presence. We found it a really good way to include more people, however, for the group activity we were not able to provide it online due to the impossibility of provide the necessary material to the people online. For this reason, for the next edition we plan to propose a full in-presence workshop. In addition, we would like to propose a two day workshop, so that we are going to have more time for the group activates.

In fact, at the beginning, we had several ideas about different group activities, but then due to the time constraints we were able to propose just one group activity.
To conclude, this workshop consider several topics related to Trustworthy AI.

## List of publications, meetings, presentations, patents,...

Please detail the outcomes of this work. This includes any produced deliverables, e.g. papers or technical reports, transfer of technology, or other types of knowledge sharing.
Please note that all dissemination, including publications, with financing from TAILOR needs to acknowledge this. Also, TAILOR is required to publish with open access. Also mention any intellectual property rights (IPR), such as patents, based on the performed work.

The papers accepted to this workshop are going to be published as proceedings. The list of all the papers is in section "Technical approach – scientific outcomes".

## Additional comments

Any additional comments that you'd like to share.