SOBIGDATA++
PLUS PLUS

Deliverable D10.4

Exploratory activities report 3

SoBigData Plus Plus receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No. 871042.

DOCUMENT INFORMATION

PROJECT	
PROJECT ACRONYM	SoBigData Plus Plus
PROJECT TITLE	SoBigData++: European Integrated Infrastructure for Social Mining and Big Data Analytics
STARTING DATE	01/01/2020 (60 months)
ENDING DATE	31/12/2024
PROJECT WEBSITE	http://www.sobigdata.eu
TOPIC	INFRAIA-01-2018-2019 Integrating Activities for Advanced Communities
GRANT AGREEMENT N.	871042
DELIVERABLE INFORMATION	
WORK PACKAGE	WP10 JRA3 - Exploratories
WORK PACKAGE LEADER	KTH & UNIPI
WORK PACKAGE PARTICIPANTS	CNR, USFD, UNIPI, FRH, UT, IMT, LUH, KCL, SNS, AALTO, ETHZ, PSE, UNIROMA1, CNRS, CEU, URV, CSD, BSC, UPF, Eli, CRA, UvA
DELIVERABLE NUMBER and TITLE	D10.4 Exploratory activities report 3
AUTHOR(S)	Luca Pappalardo (CNR), Aris Gionis (KTH), Ilaria Barsanti (CNR), Marco Braghieri (KCL)
CONTRIBUTOR(S)	
EDITOR(S)	Valerio Grossi (CNR)
REVIEWER(S)	Valerio Grossi (CNR), Michela Natilli (CNR), and all the task leaders of the WP10
CONTRACTUAL DELIVERY DATE	31/12/2024
ACTUAL DELIVERY DATE	30/12/2024
VERSION	1.1
TYPE	Report
DISSEMINATION LEVEL	Public
TOTAL N. PAGES	33
KEYWORDS	Exploratory, artificial intelligence, social mining, research spaces

EXECUTIVE SUMMARY

This deliverable updates deliverables D10.1 “Initial Exploratory activities planning”, D10.2 “Exploratory activities report and planning for the next period 1” and D10.3 “Exploratory activities report and planning for the next period 2”. This document provides information about the activities performed since 01 January 2023 to 31 December 2024 for WP10 related to Exploratories - transformed in 2024 as research spaces.

DISCLAIMER

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 871042.

SoBigData++ strives to deliver a distributed, Pan-European, multi-disciplinary research infrastructure for big social data analytics, coupled with the consolidation of a cross-disciplinary European research community, aimed at using social mining and big data to understand the complexity of our contemporary, globally-interconnected society. SoBigData++ is set to advance on such ambitious tasks thanks to SoBigData, the predecessor project that started this construction in 2015. Becoming an advanced community, SoBigData++ will strengthen its tools and services to empower researchers and innovators through a platform for the design and execution of large-scale social mining experiments.

This document contains information on SoBigData++ core activities, findings and outcomes and it may also contain contributions from distinguished experts who contribute as SoBigData++ Board members. Any reference to content in this document should clearly indicate the authors, source, organisation and publication date.

The content of this publication is the sole responsibility of the SoBigData++ Consortium and its experts, and it cannot be considered to reflect the views of the European Commission. The authors of this document have taken any available measure in order for its content to be accurate, consistent and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated the creation and publication of this document hold any sort of responsibility that might occur as a result of using its content.

GLOSSARY

AI	Artificial Intelligence
EU	European Union
EC	European Commission
H2020	Horizon 2020 EU Framework Programme for Research and Innovation
MP	Micro-Project
PM	Person Month

TABLE OF CONTENTS

1 Relevance to SoBigData++

1.1 Relevance to project objectives

This document describes the activity carried out within the exploratories (transformed in 2024 as research spaces[1]) in the last period and the topics and activities planned for the next period. For each task of WP10, we report the results achieved for each topic and the activities carried out in terms of conferences/workshops, hackathons, data collection, and software development. The topics and activities described in this document are relevant to milestone MS5 "All the Exploratories reached technical and scientific maturity".

1.2 Relation to other work packages

Since in the document we also describe some activities made or planned for the next period, this deliverable is also related to work packages WP3 - Dissemination, Impact, and Sustainability (because of workshops and conferences have been made or planned), WP4 - Training (because hackathons have been made or planned), and WP7 - Virtual Access (because data sets and software have been made available on the infrastructure or planned).

1.3 Structure of the document

The research in WP10 is structured in vertical thematic environments (formerly known as exploratories) each associated with a task, aimed at creating new stories and new resources to be integrated within the SoBigData++ research infrastructure:

- Sections from 2 to 7 describe the scientific results each exploratory has achieved in the last period of the project and the topics and activities planned to investigate for the next period. For each exploratory, we also list the micro-project proposed and (if already terminated) the corresponding resources created.

[1] Research Spaces - The scientific SoBigData collaboration operates in specific research topics aimed to apply the data science to the real world - http://sobigdata.eu/research_spaces

2 T10.1 Societal Debates and Misinformation Analysis

This exploratory aims to develop methods and datasets for studying online public debates in (near) real-time and at scale, i.e., during election campaigns or on controversial topics such as vaccination, abortion, or discrimination. The central focus regards misinformation, with the purpose of developing new methods for detecting, analysing, and tracking online misinformation and propaganda across social media platforms, countries, and over time. A key aim is to improve the accuracy of the methods through collecting more data, experimentation with semi-supervised and unsupervised methods, and integrating the latest advances in deep learning. The exploratory also studies the effect of different social relationships when it comes to opinion formation.

2.1 Activities Report

Online hostility in politics
Partners involved: USFD

Online Abuse towards MPs in the 2024 UK election. We performed a study of online abuse on X towards MPs during the 2024 election period. The aim of this big data analytics is to investigate online abuse and sharing patterns of social media posts, exposing the ways social media can be used and abused to shape opinions about significant political events such as elections. NLP and AI tools were used to detect and characterise abuse, and to perform tasks such as topic detection, hashtag identification, and a variety of network analysis and related tasks such as measuring speed of response, patterns of abusive behaviour, and abuse triggers. After initially carrying out a broad analysis of tweets relating to UK election candidates and MPs, the study focused on 5 politicians who were deemed to be particularly threatened by abusive attack, including 2 women and 3 3 ethnic minorities. Collectively, they were sent more than 85,000 clearly abusive messages between 1 May and 30 July. Key findings included the fact that almost 20% of all abuse was sexist, misogynistic or sexually explicit; and over 450 messages were explicitly racist; all of which are exceptionally high. The work was reported in The Guardian, the London-based Evening Standard, the Evidence (a Gloria Media newsletter with 50k+ subscribers), and Jamaica's premier news site Nationwide News Network.

Dataset for studying hostility towards UK MPs. We performed a study to understand and address the hostility UK Members of Parliament (MPs) face on social media, particularly on X (formerly Twitter). Politicians use these platforms to engage with the public, often exposing them to hostile interactions that target their political roles and personal identities. Such hostility undermines public trust in government and can escalate to offline harm or violence. To address this issue, we compiled a dataset of 3,320 English-language tweets directed at UK MPs over a two-year period with expert annotations for hostility and the targeted identity characteristics (race, gender, religion, none), including individual annotations with confidence scores and gold labels. Each tweet was manually annotated for hostility and, when applicable, for identity-based targeting (e.g., race, gender, religion) by three annotators. This dataset served as the foundation for exploring how hostility manifests in UK political discourse. We conducted linguistic and topical analyses to identify key patterns. Through topic analysis, we demonstrate that political hostility is closely tied to contemporaneous events, shedding light on important ramifications for training models. We also tested the effectiveness of various pre-trained language models and large language models (LLMs) in two tasks: binary hostility identification and multi-class targeted identity type classification in flat and 2-level hierarchical classification settings. This study offers valuable data and lays the groundwork for future

research aimed at understanding and mitigating the impact of online hostility in political contexts specific to the UK. A paper is in preparation and will shortly be submitted for peer review at ICWSM Data Track. Upon acceptance, we will publish a blog post on the SoBigData++ website.
Dataset: https://zenodo.org/records/10809695
DOI: 10.5281/zenodo.10809694
Code: https://github.com/pandyamugdha/ohtukmp

The Effect of Label Aggregation Techniques on Minority Opinion Representation
Partners Involved: USFD
We performed a study examining annotator disagreements in the context of sexism detection, a subjective annotation task prone to varied interpretations. Disagreements commonly arise due to carelessness or differing, yet valid, perspectives influenced by social and cultural contexts. While label aggregation methods such as majority voting or expert opinion are widely used to resolve these disagreements, they often disregard minority opinions, potentially introducing bias into datasets and downstream models. We investigated various label aggregation strategies and their impact on minority opinion representation using two sexism detection datasets. We also conducted a qualitative analysis to categorise the nature of annotator disagreements and explored the downstream effects of these strategies on model behaviour. Our findings indicate that majority aggregation effectively captures broad consensus when sexism is overt, and expert aggregation can lead to the dataset reflecting the expert's biases. However, these approaches often underrepresent nuanced classes such as dehumanisation and mistreatment of women—categories that may hold significant importance due to their more harmful implications. Models trained on majority-aggregated labels tend to amplify this bias, further reducing the visibility of minority perspectives. This work highlights the importance of tailoring label aggregation strategies to the specific objectives and characteristics of the task. By incorporating minority voices, we can better address the complexities of subjective annotation tasks like sexism detection. A paper is in preparation and will shortly be submitted for peer review at ACL. Upon acceptance, we will publish a blog post on the SoBigData++ website.
Code: https://github.com/pandyamugdha/Aggregation-Bias

Analysis of self-supervised methods for hate speech detection
Partners Involved: USFD
We explored methods for detecting offensive text and hate speech on social media, addressing challenges posed by limited labeled datasets and data imbalances. The study leveraged self-training, a semi-supervised learning technique, to augment training data by weakly labeling large, unannotated datasets. We experimented with "noisy" self-training by incorporating three textual data augmentation methods: back translation, synonym substitution, and word swapping, tested across five pre-trained BERT models of varying sizes. Key findings demonstrated that self-training consistently improved performance by up to 1.5% in F1-macro scores over standard fine-tuning methods, making smaller models as effective as larger, resource-intensive ones. However, the incorporation of noise via augmentations did not universally enhance performance; instead, it often introduced semantic variations that affected classification accuracy. The study identified limitations in existing augmentation techniques for offensive language, particularly their inability to preserve key semantic elements critical for accurate classification. This research was published in the proceedings of Recent Advances in Natural Language Processing (RANLP 2023). A dataset and associated code have been made available to the research community.

Identifying false rumours with significant impact

Partners Involved: USFD

Malicious online rumours can spread rapidly with severe societal consequences. Early prediction of false rumour popularity is essential to complement automated detection and fact-checking efforts, enabling timely delivery of counter-information by social media platforms. To this end, we introduce: (i) a novel regression task to predict the future popularity of false rumours using post and user-level data, (ii) a publicly available Chinese dataset with 19,256 false rumours from Weibo, including user profiles and a popularity score based on shares, replies, and reports, and (iii) BERT-Weibo-Rumor, an open-source domain-adapted pre-trained language model. The best model, KG-Fusion, outperforms baselines with a 1.54 RMSE and 0.636 Pearson's correlation by leveraging textual and user information. Our analysis reveals linguistic differences between popular and less popular rumours, highlighting distinctive patterns in their content. This research was published in Expert Systems with Applications. Our dataset and model have been made available to the research community.

Weibo Dataset: https://zenodo.org/records/8374169 DOI: 10.5281/zenodo.8374168

BERT-Weibo-Rumo: https://huggingface.co/YidaM4396/BERT_Weibo_Rumor

Code: https://github.com/YIDAMU/Weibo_Rumor_Popularity

Exploring Prompt Complexity in Zero-Shot Classification Settings: Insights from Large Language Models for Computational Social Science

Partners Involved: USFD

Instruction-tuned Large Language Models (LLMs) have demonstrated remarkable capabilities in language understanding and prompt-driven response generation. However, due to the high computational costs of training these models, many applications adopt a zero-shot setting. We evaluate the zero-shot performance of two publicly accessible LLMs, ChatGPT and OpenAssistant, across six classification tasks in Computational Social Science, while systematically examining the effects of various prompting strategies. Specifically, we analyze the influence of prompt complexity by incorporating label definitions, using synonyms for label names, and exploring the role of foundational training memories. Our results show that current LLMs underperform compared to smaller, fine-tuned transformer models (e.g., BERT-large) in a zero-shot setting. Moreover, we find that prompting strategies significantly impact classification accuracy, with observed variations in accuracy and F1 scores exceeding 10%. These findings highlight the importance of prompt design in harnessing LLMs for domain-specific applications. This research was published in the Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024).

Understanding Temporal Shifts in Stance Detection on COVID-19 Vaccination

Partners Involved: USFD

Vaccination has been a vital strategy in controlling the spread of COVID-19, and it is important for policymakers to understand public attitudes toward vaccination at scale. However, opinions on COVID-19 vaccination, such as pro-vaccine support or vaccine hesitancy, have changed over time on social media. In this study, we explore how temporal shifts affect stance detection for COVID-19 vaccination on Twitter. We evaluate transformer-based models (e.g., BERT and XML-RoBERTa) using both chronological splits (training, validation, and test sets arranged by time) and random splits (randomly dividing these sets) of social media data. Our results show large performance differences between the two approaches, with chronological splits significantly reducing classification accuracy. These findings highlight the need to consider temporal factors in real-world stance detection models. This research was published in the Proceedings of the 2024 Joint

International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024).

Examining the Challenges of Rumor Detection Models Trained on Static Datasets
Partners Involved: USFD
A key challenge for rumour detection models is their ability to generalize and identify emerging, previously unknown rumors. Prior studies show that content-based models, which rely solely on the source post, often struggle with unseen rumors. Meanwhile, the potential of context-based models remains underexplored. In this work, we provide an in-depth evaluation of the performance gap between content-based and context-based models in detecting new rumors. Our findings reveal that context-based models still heavily depend on source post information, often neglecting the value of additional contextual clues. We also examine how different data split strategies influence classifier performance. Finally, we offer practical recommendations to mitigate the effects of temporal concept drift when training rumour detection models on static datasets. This research was published in the Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024).

2.2 Publications

João Leite, Carolina Scarton, and Diego Silva. 2023. Noisy Self-Training with Data Augmentations for Offensive and Hate Speech Detection Tasks. In Proceedings of the 14th International Conference on Recent Advances in Natural Language Processing, pages 631–640, Varna, Bulgaria. INCOMA Ltd., Shoumen, Bulgaria.

Yida Mu, Pu Niu, Kalina Bontcheva, and Nikolaos Aletras. "Predicting and analyzing the popularity of false rumors in Weibo." *Expert Systems with Applications* 243 (2024): 122791.

Yida Mu, Ben P. Wu, William Thorne, Ambrose Robinson, Nikolaos Aletras, Carolina Scarton, Kalina Bontcheva, and Xingyi Song. 2024. Navigating Prompt Complexity for Zero-Shot Classification: A Study of Large Language Models in Computational Social Science. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 12074–12086, Torino, Italia. ELRA and ICCL.

Yida Mu, Mali Jin, Kalina Bontcheva, and Xingyi Song. 2024. Examining Temporalities on Stance Detection towards COVID-19 Vaccination. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 6732–6738, Torino, Italia. ELRA and ICCL.

Yida Mu, Xingyi Song, Kalina Bontcheva, and Nikolaos Aletras. 2024. Examining the Limitations of Computational Rumor Detection Models Trained on Static Datasets. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 6739–6751, Torino, Italia. ELRA and ICCL.

Mugdha Pandya, Mali Jin, Kalina Bontcheva, Diana Maynard. 2024. Hostility detection in uk politics: A dataset on online abuse targeting mps. *arXiv preprint arXiv:*2412.04046

Mugdha Pandya, Nafise Sadat Moosavi, Diana Maynard. 2024. Exploring the influence of label aggregation on minority voices: Implications for dataset bias and model training. *arXiv preprint arXiv:2412.04025.*

3 T10.2 Demography, Economy & Finance 2.0

The aim of this exploratory is that of combining statistical methods and traditional economic data (typically at low-frequency) with high-frequency data from non-traditional digital sources (e.g., web, supermarkets), for monitoring economic, socio-economic and well-being indicators. Another purpose of this exploratory is studying traditional complex socio-economic and financial systems in conjunction with emerging ones, in particular, block-chain & cryptocurrency markets and their applications such as smart property, Internet of things (IoT), energy trading and smart contracts. In the field of finance, different aspects will be studied, such as risk and liquidity estimation, microstructure dynamics & market predictions as well as different connections to social media and news.

3.1 Activities Report

Complexity techniques in finance

Analysis of bank leverage via dynamical systems and deep neural networks [1].

Partners Involved: SNS

We consider a model of a simple financial system consisting of a leveraged investor that invests in a risky asset and manages risk by using value-at-risk (VaR). The VaR is estimated by using past data via an adaptive expectation scheme. We show that the leverage dynamics can be described by a dynamical system of slow-fast type associated with a unimodal map on [0,1] with an addi- tive heteroscedastic noise whose variance is related to the portfolio rebalancing frequency to target leverage. In absence of noise the model is purely deterministic and the parameter space splits into two regions: (i) a region with a globally attracting fixed point or a 2-cycle; (ii) a dynamical core region, where the map could exhibit chaotic behavior. Whenever the model is randomly perturbed, we prove the existence of a unique stationary density with bounded variation, the stochastic stability of the process, and the almost certain existence and continuity of the Lyapunov exponent for the stationary measure. We then use deep neural networks to estimate map parameters from a short time series. Using this method, we estimate the model in a large dataset of US commercial banks over the period 2001-2014. We find that the parameters of a substantial fraction of banks lie in the dynamical core, and their leverage time series are consistent with a chaotic behavior. We also present evidence that the time series of the leverage of large banks tend to exhibit chaoticity more frequently than those of small banks.

Unimodal maps perturbed by heteroscedastic noise: an application to a financial systems [2]

Partners Involved: SNS

We investigate and prove the mathematical properties of a general class of one-dimensional unimodal smooth maps perturbed with a heteroscedastic noise. Specifically, we investigate the stability of the associated Markov chain, show the weak convergence of the unique stationary measure to the invariant measure of the map, and show that the average Lyapunov exponent depends continuously on the Markov chain parameters. Representing the Markov chain in terms of random transformation enables us to state and prove the Central Limit Theorem, the large deviation principle, and the Berry-Esséen inequality. We perform a multifractal analysis for the invariant and the stationary measures, and we prove Gumbel's law for the Markov chain with an extreme index equal to 1. In addition, we present an example linked to the financial concept of systemic risk and leverage cycle, and we use the model to investigate the finite sample properties of our asymptotic results

Score-Driven Exponential Random Graphs: A New Class of Time-Varying Parameter Models for Dynamical Networks [3]

Partners Involved: SNS

Motivated by the increasing abundance of data describing real-world networks that exhibit dynamical features, we propose an extension of the exponential random graph models (ERGMs) that accommodates the time variation of its parameters. Inspired by the fast-growing literature on dynamic conditional score models, each parameter evolves according to an updating rule driven by the score of the ERGM distribution. We demonstrate the flexibility of score-driven ERGMs (SD-ERGMs) as data-generating processes and filters and show the advantages of the dynamic version over the static one. We discuss two applications to temporal networks from financial and political systems. First, we consider the prediction of future links in the Italian interbank credit network. Second, we show that the SD-ERGM allows discriminating between static or time-varying parameters when used to model the U.S. Congress co-voting network dynamics.

The public use of early-stage scientific advances in carbon dioxide removal: a science-technology-policy-media perspective [4]

Partners Involved: SNS

While Carbon Dioxide Removal (CDR) solutions are considered essential to meet Paris Agreement objectives and curb climate change, their maturity and current ability to operate at scale are highly debated. The rapid development, deployment, and diffusion of such methods will likely require the coordination of science, technology, policy, and societal support. This article proposes a bibliometric approach to quantify the public use of early-stage research in CDR. Specifically, we employ generalized linear models to estimate the likelihood that scientific advances in eight different carbon removal solutions may induce (i) further production of scientific knowledge, (ii) technological innovation, and (iii) policy and media discussion. Our main result is that research in CDR is of significant social value. CDR research generates significant, positive, yet heterogeneous spillovers within science and from science to technology, policy, and media. In particular, advances in Direct Air Capture spur further research and tend to result in patentable technologies, while Blue Carbon and Bio-energy with Carbon Capture and Storage appear to gain relative momentum in the policy and public debate. Moreover, scientific production and collaborations cluster geographically by type of CDR, potentially affecting long-term carbon removal strategies. Overall, our results suggest the existence of coordination gaps between science, technology, policy, and public support.

On nonlinear compression costs: when Shannon meets Rényi [5]

Partners Involved: IMT

In compression problems, the minimum average codeword length is achieved by Shannon entropy, and efficient coding schemes such as Arithmetic Coding (AC) achieve optimal compression. In contrast, when minimizing the exponential average length, Rényi entropy emerges as a compression lower bound. This paper presents a novel approach that extends and applies the AC model to achieve results that are arbitrarily close to Rényi's lower bound. While rooted in the theoretical framework assuming independent and identically distributed symbols, the empirical testing of this generalized AC model on a Wikipedia dataset with correlated symbols reveals significant performance enhancements over its classical counterpart, when considering the exponential average. The paper also demonstrates an intriguing equivalence between minimizing the exponential average and minimizing the likelihood of exceeding a predetermined threshold in codewords' length. An extensive experimental comparison between generalized and classical AC unveils a remarkable reduction, by several orders of magnitude, in the fraction of codewords surpassing the specified threshold in the Wikipedia dataset.

3.2 Publications

[1] Fabrizio Lillo, Giulia Livieri, Stefano Marmi, Anton Solomko, Sandro Vaienti, Analysis of bank leverage via dynamical systems and deep neural networks, SIAM Journal of Financial Mathematics 14, 598-643 (2023)

[2] Fabrizio Lillo, Giulia Livieri, Stefano Marmi, Anton Solomko, Sandro Vaienti, Unimodal maps perturbed by heteroscedastic noise: an application to a financial systems Journal of Statistical Physics 190, 156 (2023)

[3] Domenico Di Gangi, Giacomo Bormetti, Fabrizio Lillo, Score-Driven Exponential Random Graphs: A New Class of Time-Varying Parameter Models for Dynamical Networks, Chaos 34, 113101 (2024)

[4] Giorgio Tripodi, Fabrizio Lillo, Roberto Mavilia, Andrea Mina, Francesca Chiaromonte and Francesco Lamperti, The public use of early-stage scientific advances in carbon dioxide removal: a science-technology-policy-media perspective, Environmental Research Letters 19 114009 (2024)

[5] Andrea Somazzi, Paolo Ferragina, Diego Garlaschelli, *On nonlinear compression costs: when Shannon meets Rényi*, IEEE Access 12, pp. 77750-77763 (2024) http://dx.doi.org/10.1109/ACCESS.2024.3406912

4 T10.3 Sustainable Cities for Citizens

This exploratory focuses on the analysis of cities, the sustainability of their flows of energy and materials and people living in them. We analysed data from different spatial and temporal scales. On city-wide scales, we analysed energy and material flows to give insights on the sustainability of transformation processes occurring in cities (the so-called "urban metabolism") and point out the circularity of flows and main polluting/GHG emission sectors and factors. On a small scale, we analysed mobility in different cities, allowing the characterization of the demand of dynamic users and granting the derivation of models to estimation pollution and optimise the electric mobility charging and relocation service and minimise its impact on the power grid.

4.1 Activities Report

Optimal Planning of Regional Renewable energy sources

Partners Involved: IMT

Description: The data collection phase has been completed. Data have been filtered for outliers and non-compatible weather stations. Weather data have been converted in the generation profiles of 130 power stations in the macro-region Toscana and Marche. A portfolio of complementary generators has been produced using two methods: 1) Markowitz optimisation 2) Network science. Both methods lead to the definition of optima spatial location where the complementary of solar and wind sources minimise fluctuations and maximise electricity generation. The community detection algorithm implemented on the correlation matrix of the power plants led to the definition of specific communities distributed in both Toscana and Marche. Results show that the clustering methods based of network analysis lead to better results in terms of planning of renewable resources on the territory.

Urban metabolism and circularity of the municipality of Albavilla

Partners Involved: IMT

Data have been collected and verified. The Urban metabolism is ongoing without the support of EnelX, that decided to stop the activities. Data collected are in the availability of IMT.

The attractiveness of European Regions

Partners Involved: IMT

The activity focuses on eight distinct flow types across European NUTS2 regions from 2010 to 2018, employing a multilayer network approach. Notably, the multilayer approach unveils insights that would be missed in single-layer analyses. Community detection reveals complex structures that demonstrate the cohesive power of national borders and the existence of strong cross-border ties in specific regions. Results also highlight the heterogeneity of many EU regions that increased or decreased their attractiveness during the observation period. The study also contributes to a more nuanced understanding of regional attractiveness, with implications for targeted policy interventions in regional development and European cohesion.

Urban Inequalities

Partners Involved: CNR, SNS

We focused on further modeling some urban inequalities phenomena: Segregation, Gentrification and access to Point of Interests (POIs). We investigated urban segregation focused on various, possible, housing change policies in the Schelling model. The results demonstrated that when agents in the model were willing to select their new location based on both personal happiness and collective needs, the segregation levels in the simulated city reached their lower limits. We are developing a new agent-based model for the modeling of Gentrification. Our simulated results show that: 1) if super high-income do not move, gentrification is not triggered; 2) temporal network properties are a good early-warning of gentrification; 3) bigger cities are more likely to experience gentrification. We developed a new line of research on the role of Recommender Systems in producing inequalities in access to Points of Interest (POIs) in the city. The findings suggested that while these systems don't seem to affect the amount of mobility produced by users, they tend to favor more popular POIs and more habitual users.

From individual to collective impact of routing choices in cities

Partners Involved: ISTI-CNR, SNS

Our research explores how digital tools and urban policies might shape city dynamics. We looked at various interventions in urban mobility systems and their broader effects on the environment and city life. Initial investigations focused on examining potential outcomes of different traffic management approaches in an urban setting. Through various analytical methods, we found that certain strategic interventions could potentially lead to improvements in both traffic flow and environmental metrics. We then expanded our study to assess how popular navigation tools might influence urban movement patterns across several cities. Our analysis suggested that these technologies, while helpful for individual users, might have unexpected effects on overall traffic patterns, particularly during busy periods. Building on these observations, we explored how introducing some variability into route suggestions might affect overall traffic patterns. This led to the development of a new routing approach that showed promising results in reducing environmental impact compared to existing methods.

4.2 Publications

E. Calò, A. Facchini, Multidimensional Territorial Attractiveness: an Application to European Flows, https://arxiv.org/abs/2412.09178

Cornacchia, G., Lemma, L., & Pappalardo, L. (2024, October). Alternative Routing based on Road Popularity. In Proceedings of the 2nd ACM SIGSPATIAL Workshop on Sustainable Urban Mobility (pp. 14-17).

Cornacchia, G., Nanni, M., Pedreschi, D., & Pappalardo, L. (2024). Navigation services amplify concentration of traffic and emissions in our cities. arXiv preprint arXiv:2407.20004.

Baccile, S., Cornacchia, G., & Pappalardo, L. (2024). Measuring the Impact of Road Removal on Vehicular CO2 Emissions. In EDBT/ICDT Workshops.

Mauro, G., Pedreschi, N., Lambiotte, R., & Pappalardo, L. (2024). Dynamic models of gentrification. arXiv preprint arXiv:2410.18004.

Mauro, G., Minici, M., & Pugliese, C. (2024, June). A Preliminary Investigation of User-and Item-Centered Bias in POI Recommendation. In 2024 25th IEEE International Conference on Mobile Data Management (MDM) (pp. 277-282). IEEE.

Mauro, G., & Pappalardo, L. (2024). The Role of Relocation Policies in Urban Segregation Dynamics. In EDBT/ICDT Workshops.

Bontorin, S., Centellegher, S., Gallotti, R., Pappalardo, L., Lepri, B., & Luca, M. (2024). Mixing individual and collective behaviours to predict out-of-routine mobility. arXiv preprint arXiv:2404.02740.

Pappalardo, L., Ferragina, E., Citraro, S., Cornacchia, G., Nanni, M., Rossetti, G., ... & Pedreschi, D. (2024). A survey on the impact of AI-based recommenders on human behaviours: methodologies, outcomes and future directions. arXiv preprint arXiv:2407.01630.

5 T10.4 Migration Studies

This exploratory studies how big data can help understand the migration phenomenon. Our scientists will try to answer various questions about migration in Europe and the world. Several studies are ongoing, including developing economic models of migration, now-casting migration stocks and flows, identifying the perception of migration and effect on the leaving and the receiving communities. We will also study the effect of migrants' networks (through the ego network graph abstraction) on the different migration phases (i.e., migration choices as well as cultural assimilation and transnationalism).

5.1 Activities Report

Improving skin-tone detection algorithm for the paper: Skin Tone Penalties: Bottom-up Discrimination in Football

Partner: PSE

Description: The previous version of the paper included use of the luminosity component of the CIELAB color space to assess lightness or darkness of a skin color. We now are refining this measure to abstract from certain limitations. We specifically focus on the Individual Topology Angle which is a quantitative measure used to categorize human skin color and is commonly calculated using the CIELAB color space. Instead of just using the L components, other components are used (b, blue yellow) and expressed as an angle. The resulting angle (direction and size) would fall within a pre-determined range of skin color, balancing the two components. This eases interpretation and is more robust to lighting issues.

Creating a Name-Americanization index that is specific to Arabic names

Partner: PSE

Description: This is part of the work on the working paper Between Arab and White: Syrians and the Naturalization Law. We leverage the Syrian Business Directory (1908-1909) to create an Americanization Index specifically designed for Arabic names, accounting for phonological nuances. By extracting and manually transliterating Arabic names to English and matching them with their Americanized counterparts, the methodology incorporates measures such as normalized Levenshtein distances, phonetic similarity (NYSIIS), and a weighted edit distance that penalizes deviations from culturally significant Arabic sounds. This index captures name assimilation patterns, illustrating transitions like "Youssef" to "Joseph" or "Ilyas" to "Louis." Unlike traditional approaches using limited naturalization records, this method provides a detailed, culturally aware measure of naming conformity over time.

Digitizing historical documents

Partner: PSE

As part of the project Between Arab and White: Syrians and the Naturalization Law, several historical documents have been digitized. This includes Arab-American newspapers from the US from 1890 to 1930s, business directories, and books on participation in WWI.

Measuring population in parties' platforms

Partner: PSE

Our contribution to the measurement of populism starts with the recognition that populism is a continuum rather than a discrete concept. We attribute a 'populism score' to all parties running in a given legislative

election based on text analysis of their political platforms. These scores reflect the salience of the anti-elite and the commitment-to-protect stances in a party's political platform. Equipped with these scores, we can define a threshold (based on criteria extensively discussed in our paper) above which a party can be categorized as populist.

An algorithm that determines which Free Trade Agreements include provisions to facilitate the movement of business people and whether these are included in dispute settlement mechanisms

Partner: PSE

The number of trade agreements has steadily increased during the last forty years. As of October 1st 2019, the World Trade Organization reports that more than 300 Regional Trade Agreements are in force.3 Identifying provisions in those agreements is therefore far from trivial. In order to examine this vast amount of text, we develop an algorithm that combines machine learning and text analysis techniques. The algorithm is able to identify whether a topic is covered or not and determine then if it is included in the dispute settlement mechanisms specified in the agreement. This tool allows us to show, in a systematized way, the increasing role of trade agreements in the regulation of the movement of business people across the globe. The algorithm has wider applicability than just the movement of business people. To assess the algorithm's accuracy, we apply it to identify all the topics covered by the hand-coded "Content of Deep Trade Agreements" World Bank's database4 . We determine if an agreement was identified by the WB as covering a given topic5 and check then if the algorithm also identifies the agreement as covering that same topic. The results overlap in over 80% of cases. Of all 2012 international trade taking place between countries with an active FTA, the FTAs coded identically by the WB and the algorithm, for at least 20 different topics, accounted on average for 99% of trade under FTAs. Beyond 20 topics this percentage rate drops substantially (figure 7). This is mainly explained by topic definitions that could potentially be interpreted differently by multiple coders.

5.2 Publications

Mayer, Thierry, Hillel Rapoport, and Camilo Umana Dajud. Forthcoming (2024). "Free Trade Agreements and the Movement of Business People." Journal of Economic Geography.

Working Papers:

Coming soon is an updated version of Kamel, Donia and Woo-Mora, L. Guillermo, Skin Tone Penalties: Bottom-up Discrimination in Football (August 11, 2023). Available at SSRN: https://ssrn.com/abstract=4537612 or http://dx.doi.org/10.2139/ssrn.4537612

Kamel, Donia. Between Arab and White: Syrians and the Naturalization Law. November 2024. [Available upon request].

Docquier, F, S Iandolo, H Rapoport, R Turati and G Vannoorenberghe (2024), 'DP18822 Populism and the Skill-Content of Globalization', CEPR Discussion Paper No. 18822. CEPR Press, Paris & London. https://cepr.org/publications/dp18822

6 T10.5 Sports Data Science

This exploratory provides massive heterogeneous dynamic data describing several sports (e.g., soccer, cycling and rugby) to construct an interpretable and easy-to-use tool for a variety of stakeholders in sports: coaches and managers, athletes, scouts, journalists and the general public. Those studies will open an exciting perspective on how to understand and explain the factors influencing sports success and how to build simulation tools for boosting both individual and collective performance.

6.1 Activities Report

Blood sample profiles and injury forecasting

Partners involved: CNR, UNIPI

This study investigates whether adding players' blood sample profiles to machine learning models, which traditionally rely on external training workloads, can improve injury prediction accuracy. The study involved 18 elite soccer players from Serie B, whose blood sample parameters (e.g., Hematocrit, Hemoglobin, red blood cell count, ferritin, and testosterone) were collected over two seasons. Players were grouped using a non-supervised ML algorithm (k-means) and their blood profiles were incorporated into injury risk predictions alongside external workload data recorded during training and matches. The inclusion of blood sample characteristics improved the accuracy of injury prediction by approximately 15%, reaching 63%. The findings suggest that monitoring both external workloads and individual characteristics, like blood profiles, can provide a more comprehensive understanding of player well-being, helping coaches tailor training programs to minimize injury risk and maximize training effectiveness.

6.2 Publications

A Rossi, L Pappalardo, C Filetti, P Cintia, Blood sample profile helps to injury forecasting in elite soccer players, Sport Sciences for Health 19 (1), 285-296

7 T10.6 Social Impact of AI and Explainable Machine Learning

The exploratory investigates the foreseeable impact of AI and Big Data on society, developing analytical and simulation tools. It also integrates a vast repertoire of practical tools for explainable AI, in particular, methods for deriving meaningful explanations of black-boxes decision systems based on machine learning.

7.1 Activities Report

Generative Model for Decision Trees
Partners Involved: UNIPI
Decision trees are among the most popular supervised models due to their interpretability and knowledge representation resembling human reasoning. Commonly used decision tree induction algorithms are based on greedy top-down strategies. Although these approaches are known to be an efficient heuristic, the resulting trees are only locally optimal and tend to have overly complex structures. On the other hand, optimal decision tree algorithms attempt to create an entire decision tree at once to achieve global optimality. We place our proposal presented in [VMSG2024] between these approaches by designing a generative model for decision trees. Our method named GenTree first learns a latent decision tree space through a variational architecture using pre-trained decision tree models. Then, it adopts a genetic procedure to explore such latent space to find a compact decision tree with good predictive performance. We compare GenTree against classical tree induction methods, optimal approaches, and ensemble models. The results show that GenTree can generate accurate and shallow, i.e., interpretable, decision trees.

A Frank System for Co-Evolutionary Hybrid Decision-Making
Partners Involved: UNIPI
In [MGM2024] we introduce Frank, a human-in-the-loop system for co-evolutionary hybrid decision-making aiding the user to label records from an un-labeled dataset. Frank employs incremental learning to “evolve” in parallel with the user’s decisions, by training an interpretable machine learning model on the records labeled by the user. Furthermore, advances state-of-the-art approaches by offering inconsistency controls, explanations, fairness checks, and bad-faith safeguards simultaneously. We evaluate Frank by simulating the users’ behavior with various levels of expertise and reliance on Frank’s suggestions. The experiments show that Frank’s intervention leads to improvements in the accuracy and the fairness of the decisions.

Advancing Dermatological Diagnostics: Interpretable AI for Enhanced Skin Lesion Classification
Partners Involved: UNIPI, CNR
The paper focuses on enhancing the interpretability of AI models for diagnosing skin lesions through a method called ABELE (Adversarial Black Box Explainer generating Latent Exemplars). ABELE provides both exemplar and counter-exemplar images to explain the decisions made by a ResNet-based classifier, thereby improving transparency in critical medical diagnostics. The study demonstrates that ABELE-generated explanations increase trust and confidence in AI-assisted decision systems, especially among domain experts. Latent space analysis revealed separations among skin lesion classes, potentially aiding in identifying common misdiagnoses. The research involved a user survey, indicating that explanations generally enhanced confidence in the model's decisions. Limitations include the dataset's representativeness, as it lacks diverse skin lesions and demographics, and the absence of patient-friendly explanation modules. Future work aims to expand dataset coverage, enhance the user visualization module, and explore the applicability of ABELE to other medical domains.

FLocalX-Local to Global Fuzzy Explanations for Black Box Classifiers
Partners Involved: UNIPI
GLocalX introduced the "local to global" explanation problem, defined as the task of merging together local explanations over different instances into global explanations. Specifically, explanations are in the form of crisp decision rules. In FLocalX [fernandezFLocalXLocalTo2024], we expand upon GLocalX by devising a fuzzy merge. Explanations are provided as fuzzy decision rules, and merge is supported by an evolutionary algorithm. The resulting merge is a hierarchy of explanations at different granularities, and quantifying uncertainty among predictions. The paper introduces notions of fuzzy merge, explanation regularization, and as well as studying the algorithm ability in introducing fuzzy rules when necessary.

Explaining Siamese networks in few-shot learning
Partners Involved: UNIPI
Machine learning models often struggle to generalize accurately when tested on new class distributions that were not present in their training data. This is a significant challenge for real-world applications that require quick adaptation without the need for retraining. To address this issue, few-shot learning frameworks, which includes models such as Siamese Networks, have been proposed. Siamese Networks learn similarity between pairs of records through a metric that can be easily extended to new, unseen classes. However, these systems lack interpretability, which can hinder their use in certain applications. To address this we propose SINEX, a data-agnostic method to explain the outcomes of Siamese Networks in the context of few-shot learning. Our explanation method is based on a post-hoc perturbation-based procedure that evaluates the contribution of individual input features to the final outcome. As such, it falls under the category of post-hoc explanation methods. We present two variants, one that considers each input feature independently, and another that evaluates the interplay between features. Additionally, we propose two perturbation procedures to evaluate feature contributions. Qualitative and quantitative results demonstrate that our method is able to identify highly discriminant intra-class and inter-class characteristics, as well as predictive behaviors that lead to misclassification by relying on incorrect features.

Fast, Interpretable, and Deterministic Time Series Classification with a Bag-Of-Receptive-Fields
Partners Involved: UNIPI
The current trend in the literature on Time Series Classification is to develop increasingly accurate algorithms by combining multiple models in ensemble hybrids, representing time series in complex and expressive feature spaces, and extracting features from different representations of the same time series. As a consequence of this focus on predictive performance, the best time series classifiers are black-box models, which are not understandable from a human standpoint. Even the approaches that are regarded as interpretable, such as shapelet-based ones, rely on randomization to maintain computational efficiency. This poses challenges for interpretability, as the explanation can change from run to run. Given these limitations, we propose the Bag-Of-Receptive-Field (BORF), a fast, interpretable, and deterministic time series transform. Building upon the classical Bag-Of-Patterns, we bridge the gap between convolutional operators and discretization, enhancing the Symbolic Aggregate Approximation (SAX) with dilation and stride, which can more effectively capture temporal patterns at multiple scales. We propose an algorithmic speedup that reduces the time complexity associated with SAX-based classifiers, allowing the extension of the Bag-Of-Patterns to the more flexible Bag-Of-Receptive-Fields, represented as a sparse multivariate tensor. The empirical results from testing our proposal on more than 150 univariate and multivariate classification datasets demonstrate good accuracy and great computational efficiency compared to traditional SAX-based methods and state-of-the-art time series classifiers, while providing easy-to-understand explanations.

Enhancing Echo State Networks with Gradient-based Explainability Methods
Partners Involved: UNIPI
Recurrent Neural Networks are effective for analyzing temporal data, such as time series, but they often require costly and time-intensive training. Echo State Networks simplify the training process by using a fixed recurrent layer, the reservoir, and a trainable output layer, the readout. In sequence classification problems, the readout typically receives only the final state of the reservoir. However, averaging all states can sometimes be beneficial. In this work, we assess whether a weighted average of hidden states can enhance the Echo State Network performance. In this work, we propose a gradient-based, explainable technique to guide the contribution of each hidden state towards the final prediction. We show that our approach outperforms the naive average, as well as other baselines, in time series classification, particularly on noisy data.

Drifting explanations in continual learning
Partners Involved: UNIPI
Continual Learning (CL) trains models on data streams to learn new information without forgetting prior knowledge, yet often lacks interpretability, complicating the understanding of decision-making. This issue is exacerbated by the non-stationary nature of CL data streams, as strategies to mitigate forgetting affect learned representations. To address this, we introduce CLEX (ContinuaL EXplanations), a protocol to evaluate explanation changes in Class-Incremental scenarios, where forgetting is significant. Our findings reveal that models with similar predictive accuracy may generate divergent explanations. Replay-based strategies, effective in class-incremental settings, produce explanations aligned with offline-trained models, whereas naive fine-tuning leads to degenerate explanations. Interestingly, randomized recurrent models reduce explanation drift more effectively than fully-trained ones, highlighting this phenomenon's broader implications beyond predictive performance.

Causality-Aware Local Interpretable Model-Agnostic Explanations
Partners Involved: UNIPI
A main drawback of eXplainable Artificial Intelligence (XAI) approaches is the feature independence assumption, hindering the study of potential variable dependencies. This leads to approximating black box behaviors by analyzing the effects on randomly generated feature values that may rarely occur in the original samples. In [CG2024], we address this issue by integrating causal knowledge in an XAI method to enhance transparency and enable users to assess the quality of the generated explanations. Specifically, we propose a novel extension to a widely used local and model-agnostic explainer, which encodes explicit causal relationships within the data surrounding the instance being explained. Extensive experiments show that our approach overcomes the original method in terms of faithfully replicating the black-box model's mechanism and the consistency and reliability of the generated explanations.

Data-Agnostic Pivotal Instances Selection for Decision-Making Models
Partners Involved: UNIPI
Machine learning tools have become invaluable for addressing business and social challenges. However, many methodologies rely on complex models that lack interpretability for experts and end users. Since humans typically make decisions by comparing the case at hand with a few exemplary and representative cases stored in their minds, our goal was to design an approach capable of selecting such exemplary cases—referred to as pivots—in a data-driven manner and leveraging them to construct an interpretable predictive model. To achieve this, we proposed a hierarchical and interpretable pivot selection model inspired by decision trees, which relied on the similarity between pivots and input instances. Our approach was data-

agnostic, enabling its application to any data type by leveraging pre-trained networks for data transformations. In a work presented in [CSG2024], we detailed the implementation of this approach and demonstrated its effectiveness through experiments on various datasets, including tabular data, text, images, and time series. The results highlighted the superiority of our method compared to naive alternatives and state-of-the-art instance selection techniques, all while minimizing model complexity by reducing the number of pivots identified.

Counterfactual and Prototypical Explanations for Tabular Data via Interpretable Latent Space
Partners Involved: UNIPI
In [PBGGP2024] we propose CP-ILS, a comprehensive interpretable feature reduction method for tabular data capable of generating counterfactual and prototypical post-hoc explanations using an interpretable latent space. CP-ILS optimizes a transparent feature space whose similarity and linearity properties enable the easy extraction of local and global explanations for any pre-trained black-box model, in the form of counterfactual/prototype pairs. We evaluate the effectiveness of the created latent space by showing its capability to preserve pair-wise similarities like well-known dimensionality reduction techniques. Moreover, we assess the quality of counterfactuals and prototypes generated with CP-ILS against state-of-the-art explainers, demonstrating that our approach obtains more robust, plausible, and accurate explanations than its competitors under most experimental conditions.

User studies in algorithm-supported recidivism risk assessment
Partners involved: UPF
In [PCTKP2024] we study the effects of using an algorithm-based risk assessment instrument to support the prediction of risk of criminal recidivism. The instrument we use in our experiments is a machine learning version of RisCanvi, which is the main risk assessment instrument used by the Justice Department of Catalonia. The task is to predict whether a person who has been released from prison will commit a new crime, leading to re-incarceration, within the next two years. We measure, among other variables, the accuracy of human predictions with and without algorithmic support. This user study is done with (1) general participants from diverse backgrounds recruited through a crowdsourcing platform, (2) targeted participants who are students and practitioners of data science, criminology, or social work and professionals who work with RiskCanvi. Among other findings, we observe that algorithmic support systematically leads to more accurate predictions from all participants, but that statistically significant gains are only seen in the performance of targeted participants with respect to that of crowdsourced participants. We also run focus groups with participants of the targeted study to interpret the quantitative results, including people who use RisCanvi in a professional capacity. Among other comments, professional participants indicate that they would not foresee using a fully automated system in criminal risk assessment, but do consider it valuable for training, standardization, and to fine-tune or double-check their predictions on particularly difficult cases.

User studies in algorithmic recommendations and their long-term impact
Partners involved: UPF
In [PGC2024], we present the results of a 12-week longitudinal user study wherein the participants, 110 subjects from Southern Europe, received on a daily basis Electronic Music (EM) diversified recommendations. By analyzing their explicit and implicit feedback, we show that exposure to specific levels of music recommendation diversity may be responsible for long-term impacts on listeners' attitudes. In particular, we highlight the function of diversity in increasing the openness in listening to EM, a music genre not particularly known or liked by the participants previous to their participation in the study. Moreover, we demonstrate that recommendations may help listeners in removing positive and negative attachments towards EM,

deconstructing pre-existing implicit associations, but also stereotypes associated with this music. In addition, our results show the significant clout that recommendation diversity has in generating curiosity in listeners.

Analysis of medical data used to train machine learning models
Partners involved: UPF
In [BLC2024] we present an extensive survey of publicly available biomedical datasets, revealing four dozen databases connected to chronic diseases, such as cancer, diabetes, heart diseases, and COVID-19, among others. We analyze these datasets, highlighting commonalities and best practices among them, and to raise awareness about the wealth of data available to study chronic diseases, focusing on the importance of the sociodemographic data in biomedical research. The next step will be the analysis of fairness issues in models created from these datasets.

Explainability in Graph Neural Networks: The GRETEL Framework
Partners involved: UAQ
Graph Neural Networks (GNNs) have revolutionized applications across diverse domains, including social network analysis, molecular modeling, and community detection. However, their black-box nature poses significant challenges in interpretability, critical for trust, fairness, and transparency in machine learning systems. Addressing these issues, Graph Counterfactual Explanation (GCE) techniques provide a means to understand GNN predictions by generating counterfactual examples—minimal modifications to inputs that alter model outputs. Despite their potential, GCE methods face barriers due to a lack of standardization in methodologies, datasets, and evaluation metrics. To bridge this gap, the GRETEL framework was introduced as a comprehensive and extensible platform for developing, evaluating, and benchmarking GCE methods [GRTL, GRTL-DEMO, GRTL2]. GRETEL supports open science principles and provides a robust suite of pre-built components, including explainers, datasets, oracles, and evaluation metrics. Its object-oriented design emphasizes modularity and extensibility, allowing researchers and practitioners to seamlessly integrate custom components and assess GNN explainability across diverse domains. Building on the success of its predecessor, GRETEL 2.0 introduces significant enhancements in usability and extensibility [GRTL2]. It eliminates the need for specialized factory classes, introduces configurable and trainable classes for dynamic component generation, and expands support for learning-based explainers. The new framework accommodates various scenarios by supporting advanced features like synthetic dataset generation, ad-hoc metric implementation, and streamlined configuration via JSON files. These advancements empower users to evaluate counterfactual explanations more precisely while fostering innovation in XAI for graph data. Through rigorous evaluation of real and synthetic datasets, GRETEL 2.0 demonstrates its effectiveness in promoting transparency and fairness in GNN predictions [GRTL2]. By serving as a foundational tool for academic researchers and industry practitioners, GRETEL 2.0 advances the field of interpretable AI, ensuring that the benefits of GNNs are realized equitably and trustworthy.

Enhancing Privacy and Utility in Federated Learning: A Hybrid P2P and Server-based Approach with Differential Privacy Protection
Partners involved: UNIPI, SNS
The approach proposed in [CMP2024] introduces a hybrid Federated Learning (FL) framework designed to achieve the trade-off between privacy and model utility. FL allows distributed training of machine learning models without requiring participants to share raw data. However, privacy concerns often require the use of a privacy-enhancing technique such as the Differential Privacy (DP). DP effectively protects privacy but it tends to degrade the model's accuracy. The proposed approach combines Client-Server and Peer-to-Peer (P2P) architectures to mitigate this issue. The framework assumes that each client has access to two types of datasets: a low-privacy dataset with fewer privacy constraints and a high-privacy dataset requiring stricter

protections. Initially, clients collaborate within clusters using their low-privacy datasets to train a model via P2P without using DP. These trained models are then used as a starting point for server-based FL, which incorporates DP for training on high-privacy datasets. The experimental evaluation uses three widely recognized tabular datasets to demonstrate the effectiveness of the proposed approach in reducing accuracy degradation caused by DP. Results show that the hybrid approach can improve the model's utility while maintaining privacy guarantees, achieving up to a 32% reduction in accuracy loss in some cases.

Explainable Authorship Identification in Cultural Heritage Applications*
Partners Involved: UNIPI, ISTI-CNR
While a substantial amount of work has recently been devoted to improving the accuracy of computational Authorship Identification (AId) systems for textual data, little to no attention has been paid to endowing AId systems with the ability to explain the reasons behind their predictions. This substantially hinders the practical application of AId methods, since the predictions returned by such systems are hardly useful unless they are supported by suitable explanations. In the [SCM+2024], we have explored the applicability of existing general-purpose eXplainable Artificial Intelligence (XAI) techniques to AId, with a focus on explanations addressed to scholars working in cultural heritage. In particular, we have assessed the relative merits of three different types of XAI techniques (feature ranking, probing, factual and counterfactual selection) on three different AId tasks (authorship attribution, authorship verification and same-authorship verification) by running experiments on real AId textual data. Our analysis shows that, while these techniques make important first steps towards XAI, more work remains to be done to provide tools that can be profitably integrated into the workflows of scholars.

GLOR-FLEX: Local to Global Rule-Based EXplanations for Federated Learning
Partners Involved: UNIPI, URV
The increasing spread of artificial intelligence applications has led to decentralized frameworks that foster collaborative model training among multiple entities. One of such frameworks is federated learning, which ensures data availability in client nodes without requiring the central server to retain any data. Nevertheless, similar to centralized neural networks, interpretability remains a challenge in understanding the predictions of these decentralized frameworks. The limited access to data on the server side further complicates the applicability of explainers in such frameworks. To address this challenge, in [HN2024] we propose GLOR-FLEX, a framework designed to generate rule-based global explanations from local explainers. GLOR-FLEX ensures client privacy by preventing the sharing of actual data between the clients and the server. The proposed framework initiates the process by constructing local decision trees on each client's side to produce local explanations. Subsequently, by using rule extraction from these trees and strategically sorting and merging those rules, the server obtains a merged set of rules suitable to be used as a global explainer. We empirically evaluate the performance of GLOR-FLEX on three distinct tabular data sets, showing high fidelity scores between the explainers and both the local and global models. Our results support the effectiveness of GLOR-FLEX in generating accurate explanations that efficiently detect and explain the behavior of both local and global models.

Balancing Act: Navigating the Privacy-Utility Spectrum in Principal Component Analysis.
Partners Involved: UNIPI
A lot of research in federated learning has been ongoing ever since it was proposed. Federated learning allows collaborative learning among distributed clients without sharing their raw data to a central aggregator (if it is present) or to other clients in a peer to peer architecture. However, each client participating in the federation shares their model information learned from their data with other clients participating in the FL

process, or with the central aggregator. This sharing of information, however, makes this approach vulnerable to various attacks, including data reconstruction attacks. Our research specifically focuses on Principal Component Analysis (PCA), as it is a widely used dimensionality technique. For performing PCA in a federated setting, distributed clients share local eigenvectors computed from their respective data with the aggregator, which then combines and returns global eigenvectors. Previous studies on attacks against PCA have demonstrated that revealing eigenvectors can lead to membership inference and, when coupled with knowledge of data distribution, result in data reconstruction attacks. Consequently, our objective in the work presented in [KMN2024] is to augment privacy in eigenvectors while sustaining their utility. To obtain protected eigenvectors, we use k-anonymity, and generative networks. Through our experimentation, we did a complete privacy, and utility analysis of original and protected eigenvectors. For utility analysis, we apply HIERARCHICAL CLUSTERING, RANDOM FOREST regressor, and RANDOM FOREST classifier on the protected, and original eigenvectors. We got interesting results, when we applied HIERARCHICAL CLUSTERING on the original, and protected datasets, and eigenvectors. The height at which the clusters are merged declined from 250 to 150 for original, and synthetic versions of CALIFORNIA-HOUSING data, respectively. For the k-anonymous version of CALIFORNIA-HOUSING data, the height lies between 150, and 250. To evaluate the privacy risks of the federated PCA system, we act as an attacker and conduct a data reconstruction attack.

Decentralised Federated Learning on Complex Networks: Initialization and coordination issues
Partners Involved: IIT-CNR, CEU
Federated Learning (FL) is a well-known framework for successfully performing a learning task in an edge computing scenario where the devices involved have limited resources and incomplete data representation. The basic assumption of FL is that the devices communicate directly or indirectly with a parameter server that centrally coordinates the whole process, overcoming several challenges associated with it. However, in highly pervasive edge scenarios, the presence of a central controller that oversees the process cannot always be guaranteed, and the interactions between devices (i.e., the connectivity graph) are rarely centrally managed. Instead, these interactions are typically determined by the devices themselves, resulting in a connectivity graph with a complex network structure. Moreover, the heterogeneity of data and devices further complicates the learning process. This poses new challenges from a learning standpoint that we address by proposing a communication-efficient Decentralised Federated Learning (DFL) algorithm able to cope with them. In [VBPKKI2023], we propose a solution that allows devices communicating only with their direct neighbours to train an accurate model, overcoming the heterogeneity induced by data and different training histories. Our results show that the resulting local models generalise better than those trained with competing approaches (i.e., DecAvg, CFA and CFA-GE) and do so in a more communication-efficient way. Along an orthogonal research direction, our research in [BBVKK2024] highlights that the effectiveness of decentralised federated learning is significantly influenced by the network topology of connected devices and the learning models' initial conditions. We propose a strategy for uncoordinated initialisation of the artificial neural networks based on the distribution of eigenvector centralities of the underlying communication network, leading to a radically improved training efficiency. Additionally, our study explores the scaling behaviour and the choice of environmental parameters under our proposed initialisation strategy. This work paves the way for more efficient and scalable artificial neural network training in a distributed and uncoordinated environment, offering a deeper understanding of the intertwining roles of network structure and learning dynamics.

Enhancing Security and Privacy in Federated Learning Scenarios

Partners Involved: URV

Federated Learning (FL) allows multiple data owners to build high-quality deep learning models collaboratively, by sharing only model updates and keeping data on their premises. Although FL provides privacy by design, it is vulnerable to certain types of attacks, such as membership inference attacks (MIA), where an adversary tries to determine whether a sample was included in the training data. Existing defenses against MIA cannot offer meaningful privacy protection without significantly hampering the model's utility and causing a non-negligible training overhead. We analyzed the underlying causes of the differences in the model behavior for member and non-member samples, which arise from model overfitting and facilitate MIAs. Accordingly, in [ASHD2024] we proposed MemberShield, a generalization-based defense method for MIAs that consists of: (i) one-time preprocessing of each client's training data labels that transforms one-hot encoded labels to soft labels and eventually exploits them in local training, and (ii) early stopping the training when the local model's validation accuracy does not improve on that of the global model for a number of epochs. Extensive empirical evaluations on three widely used datasets and four model architectures demonstrate that MemberShield outperforms state-of-the-art defense methods by delivering substantially better practical privacy protection against all forms of MIAs, while better preserving the target model utility. On top of that, our proposal significantly reduces training time and is straightforward to implement, by just tuning a single hyperparameter. As mentioned earlier, FL introduces autonomy and privacy-by-design to participating peers. However, this same autonomy opens the door for malicious peers to poison the model by performing either untargeted or targeted poisoning attacks. The label-flipping (LF) attack is a targeted poisoning attack where the attackers poison their training data by flipping the labels of some examples from one class (i.e., the source class) to another (i.e., the target class). Unfortunately, this attack is easy to perform and hard to detect, and it negatively impacts the performance of the global model. Existing defenses against LF are limited by assumptions on the distribution of the peers' data and/or do not perform well with high-dimensional models. In this activity, we deeply investigate the LF attack behavior and we find that the conflicting goals of attackers and honest peers on the source class examples are reflected on the parameter gradients corresponding to the neurons of the source and target classes in the output layer. This makes those gradients good discriminative features for the attack detection. Accordingly, in [JDSB2024] we proposed LFighter, a novel defense against the LF attack that first dynamically extracts those gradients from the peers' local updates and then clusters the extracted gradients, analyzes the resulting clusters, and filters out potential bad updates before model aggregation. Extensive empirical analysis on three data sets shows the effectiveness of the proposed defense regardless of the data distribution or model dimensionality. Also, LFighter outperforms several state-of-the-art defenses by offering lower test error, higher overall accuracy, higher source class accuracy, lower attack success rate, and higher stability of the source class accuracy. Another well-known challenge of FL is to balance accuracy with privacy and security. On the one hand, good updates sent by honest participants may reveal their private local information, whereas poisoned updates sent by malicious participants may compromise the model's availability and/or integrity. On the other hand, enhancing privacy via update distortion damages accuracy, whereas doing so via update aggregation damages security because it does not allow the server to filter out individual poisoned updates. To tackle the accuracy-privacy-security conflict, in [JDBS2024] we proposed fragmented federated learning (FFL), in which participants randomly exchange and mix fragments of their updates before sending them to the server. To achieve privacy, we design a lightweight protocol that allows participants to privately exchange and mix encrypted fragments of their updates so that the server can neither obtain individual updates nor link them to their originators. To achieve security, we designed a reputation-based defense tailored for FFL that builds trust in participants and their mixed updates based on the quality of the fragments they exchange and the

mixed updates they send. Since the exchanged fragments' parameters keep their original coordinates and attackers can be neutralized, the server can correctly reconstruct a global model from the received mixed updates without accuracy loss. Experiments on four real data sets show that FFL can prevent semi-honest servers from mounting privacy attacks, can effectively counter poisoning attacks and can keep the accuracy of the global model. In another approach to solve the limitations of the existing privacy-preserving FL schemes (e.g., loss of accuracy, high communication/computation cost, failure to support dynamic users, and insecurity against collusion attacks), we also proposed a lightweight privacy-preserving FL scheme based on a dual-server architecture [ZWZD2024]. Our scheme involves only lightweight cryptographic operations, i.e., hash and symmetric encryption operations, and it has low communication overhead. Thus, it is computationally lightweight and round-efficient. Further, it allows users to join/quit an FL task and it is accuracy-lossless. We formally prove that our scheme remains secure even in case of collusion attacks. In particular, if an attacker colludes with one of the servers and all the users who participate in an FL task except two, the privacy of user gradients stays unviolated. The reported experimental results demonstrate that our scheme incurs only a marginal increase in total communication overhead compared to the FL scheme without any privacy protection. In terms of computation overhead, the cost per user remains stable as the number of users grows, while the cost for the server is comparable to that of the FL scheme without any privacy protection.

Federated learning-based natural language processing: a systematic literature review
Partners Involved: URV
As we mentioned earlier, Federated learning (FL) is a decentralized machine learning (ML) framework that allows models to be trained without sharing the participants' local data. As a result, FL thus preserves privacy better than centralized machine learning. Because textual data (such as medical records, posts in social networks, or search queries) often contains personal information, many natural language processing (NLP) tasks dealing with such data have shifted from the centralized to the FL setting. However, FL is not without problems, including convergence and security vulnerabilities (due to unreliable or poisoned data introduced into the model), communication and computational bottlenecks, and even privacy attacks orchestrated by honest-but-curious servers. In this activity we conducted a systematic literature review (SLR) of NLP applications in FL with a special focus on FL issues and the solutions proposed so far. Our review surveyed 36 recent papers published in relevant venues, which are systematically analyzed and compared from multiple perspectives. As a result of the survey, published in [KSD2024], we also identified the most outstanding challenges in the field.

Defending against backdoor attacks by layer-wise feature analysis
Partners Involved: URV
Training deep neural networks (DNNs) usually requires massive training data and computational resources. Users who cannot afford this may prefer to outsource training to a third party or resort to publicly available pre-trained models. Unfortunately, doing so facilitates a new training-time attack (i.e., backdoor attack) against DNNs. This attack aims to induce misclassification of input samples containing adversary specified trigger patterns. In [JDL2024], we conducted a layer-wise feature analysis of the behavior of benign and poisoned samples generated by attacked DNNs. We found that the feature difference between benign and poisoned samples tends to reach the maximum at a critical layer, which can be easily located based on the behaviors of benign samples. Based on this finding, we proposed a simple yet effective backdoor detection to determine whether a given suspicious testing sample is poisoned by analyzing the differences between its features and those of a few local benign samples. Our extensive experiments on benchmark datasets confirmed the effectiveness of our detection. We hope our work can provide a deeper understanding of

attack mechanisms, to facilitate the design of more effective and efficient backdoor defenses and more secure DNNs.

Privacy-preserving data sharing
Partners Involved: URV
IA models require a substantial amount of data to function effectively. Consequently, companies are leveraging a wide range of data sources to train and enhance their IA models. Therefore, it is crucial to safeguard the released data to ensure the privacy of individuals while maintaining the utility of the data.
In this activity we analyzed the use of differential privacy to protect the 2020 U.S. Decennial Census. The threat of reconstruction attacks led the U.S. Census Bureau (USCB) to replace in the Decennial Census 2020 the traditional statistical disclosure limitation based on rank swapping with one based on differential privacy (DP), leading to substantial accuracy loss of released statistics. Yet, it has been argued that, if many different reconstructions are compatible with the released statistics, most of them do not correspond to actual original data, which protects against respondent reidentification. Recently, a new attack has been proposed, which incorporates the condence that a reconstructed record was in the original data. The alleged risk of disclosure entailed by such confidence-ranked reconstruction has renewed the interest of the USCB to use DP-based solutions. To forestall a potential accuracy loss in future releases, we showed in [SJMD2024] that the proposed reconstruction is neither effective as a reconstruction method nor conducive to disclosure as claimed by its authors. We believe that Differential Privacy significantly deteriorates the analytical utility of the protected outcomes. To keep data utility at reasonable levels, practical applications of DP to data releases have used weak privacy parameters (large ϵ), which dilute the privacy guarantees of DP. In [SSDM2024], we tackled this issue by using an alternative formulation of the DP privacy guarantees, named ϵ-individual differential privacy (iDP), which causes less data distortion while providing the same protection as DP to subjects. We enforce iDP in data releases by relying on attribute masking plus a pre-processing step based on data microaggregation. The goal of this step is to reduce the sensitivity to record changes, which determines the amount of noise required to enforce iDP (and DP). Specifically, we propose data microaggregation strategies designed for iDP whose sensitivities are significantly lower than those used in DP. As a result, we obtain iDP-protected data with significantly better utility than with DP. We report on experiments that show how our approach can provide strong privacy (small ϵ) while yielding protected data that do not significantly degrade the accuracy of secondary data analysis. Finally, in [MRDS2024], we responded to an article published in 2023, in which the authors suggested assessing disclosure risk by using a counterfactual method to compare the posterior-to-posterior probability of an inference with and without the target record. We argued that this counterfactual method does not measure the risk to individual respondents.

7.2 Publications

[BLC2024] Ioannis Bilionis, Luis Fernandez-Luque and Carlos Castillo: A Survey on Public Data Sets Related to Chronic Diseases. Short paper to appear in International Symposium on Computer-Based Medical Systems (CBMS).

[PGC2024] Lorenzo Porcaro, Emilia Gómez, Carlos Castillo. Assessing the Impact of Music Recommendation Diversity on Listeners: A Longitudinal Study. ACM Transactions on Recommender Systems, 2024. DOI: 10.1145/3608487

[PCTKP2024] Manuel Portela, Carlos Castillo, Songül Tolan, Marzieh Karimi-Haghighi, Antonio Andres Pueyo. A Comparative User Study of Human Predictions in Algorithm-Supported Recidivism Risk

Assessment. Accepted for publication in Artificial Intelligence and Law. Springer, 2024. DOI: 10.1007/s10506-024-09393-y

[CSG2024] Cascione, A., Setzu, M., & Guidotti, R. (2024, August). Data-Agnostic Pivotal Instances Selection for Decision-Making Models. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases (pp. 367-386). Cham: Springer Nature Switzerland.

[VMSG2024] Guidotti, R., Monreale, A., Setzu, M., & Volpi, G. (2024, March). Generative Model for Decision Trees. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 38, No. 19, pp. 21116-21124).

[MGM2024] Mazzoni, F., Guidotti, R., & Malizia, A. (2024, April). A Frank System for Co-Evolutionary Hybrid Decision-Making. In International Symposium on Intelligent Data Analysis (pp. 236-248). Cham: Springer Nature Switzerland.

[PBGGP2024] Piaggesi, S., Bodria, F., Guidotti, R., Giannotti, F., & Pedreschi, D. (2024, November). Counterfactual and Prototypical Explanations for Tabular Data via Interpretable Latent Space. IEEE Access.

[CG2024] Cinquini, M., Guidotti, R. (2024). Causality-Aware Local Interpretable Model-Agnostic Explanations. In: Longo, L., Lapuschkin, S., Seifert, C. (eds) Explainable Artificial Intelligence. xAI 2024. Communications in Computer and Information Science, vol 2155. Springer, Cham.

[FGP2024] Fedele, A., Guidotti, R. & Pedreschi, D. Explaining Siamese networks in few-shot learning. Mach Learn 113, 7723–7760 (2024). DOI: https://doi.org/10.1007/s10994-024-06529-8

[fernandezFLocalXLocalTo2024] Guillermo Fernández, Riccardo Guidotti, Fosca Giannotti, Mattia Setzu, Juan A. Aledo, José A. Gámez, José Miguel Puerta. FLocalX - Local to Global Fuzzy Explanations for Black Box Classifiers. IDA (2) 2024: 197-209

[GRTL2] Prado-Romero Mario Alfonso, Prenkaj Bardh, Stilo, Giovanni, GRETEL 2.0: Generation and Evaluation of Graph Counterfactual Explanations Evolved. Machine Learning and Knowledge Discovery in Databases. Research Track and Demo Track - European Conference, ECML PKDD 2024, Vilnius, Lithuania, September 9-13, 2024, Proceedings, Part VIII, p.363 - 367, 2024, 10.1007/978-3-031-70371-3_21

[GRTL-DEMO] Prado-Romero Mario Alfonso, Prenkaj Bardh, Stilo Giovanni. Developing and Evaluating Graph Counterfactual Explanation with GRETEL. Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining, WSDM 2023, Singapore, 27 February 2023 - 3 March 2023. p. 1180 - 1183, 2023, 10.1145/3539597.3573026

[GRTL] Prado-Romero Mario Alfonso, Stilo, Giovanni. GRETEL: Graph Counterfactual Explanation Evaluation Framework. Proceedings of the 31st ACM International Conference on Information & Knowledge Management, Atlanta, GA, USA, October 17-21, 2022. p. 4389 - 4393, 2022, 10.1145/3511808.3557608

[CMP2024] Corbucci L., Monreale A., Pellungrini R. Enhancing Privacy and Utility in Federated Learning: A Hybrid P2P and Server-Based Approach with Differential Privacy Protection. Proceedings of the 21st International Conference on Security and Cryptography SECRYPT

[VBPKKI2023] Lorenzo Valerio, Chiara Boldrini, Andrea Passarella, János Kertész, Márton Karsai, Gerardo Iñiguez. Coordination-free Decentralised Federated Learning on Complex Networks: Overcoming Heterogeneity. https://arxiv.org/abs/2312.04504 Submitted.

[BBVKK2024] Arash Badie-Modiri, Chiara Boldrini, Lorenzo Valerio, János Kertész, Márton Karsai. Initialisation and Network Effects in Decentralised Federated Learning. https://arxiv.org/abs/2403.15855 Submitted.

[SCM+2024]Mattia Setzu, Silvia Corbara, Anna Monreale, Alejandro Moreo, Fabrizio Sebastiani: Explainable Authorship Identification in Cultural Heritage Applications. ACM Journal on Computing and Cultural Heritage 17(3): 44:1-44:23 (2024)

[HN2024] Rami Haffar, Francesca Naretto, David Sánchez, Anna Monreale, Josep Domingo-Ferrer: GLOR-FLEX: Local to Global Rule-Based EXplanations for Federated Learning. FUZZ 2024: 1-9

[KMN2024] Saloni Kwatra, Anna Monreale, Francesca Naretto: Balancing Act: Navigating the Privacy-Utility Spectrum in Principal Component Analysis. SECRYPT 2024: 850-857

[SSDM2024] Jordi Soria-Comas, David Sánchez, Josep Domingo-Ferrer, Sergio Martínez, Luis Del Vasto-Terrientes, "Conciliating privacy and utility in data releases via individual differential privacy and microaggregation", Transactions on Data Privacy, vol. 18, no. 1, pp. 29-50, 2025.

[DO2024] Josep Domingo-Ferrer and Melek Önen (eds.), Privacy in Statistical Databases-PSD 2024, Lecture Notes in Computer Science, vol. 14915, Springer, 2024. ISBN 978-3-031-69650-3

[ASHD2024] Faisal Ahmed, David Sánchez, Zouhair Haddi, and Josep Domingo-Ferrer, "MemberShield: a framework for federated learning with membership privacy", Neural Networks, vol. 181, art. 106768, 2025

[KSD2024] Younas Khan, David Sánchez, and Josep Domingo-Ferrer, "Federated learning-based natural language processing: a systematic literature review", Artificial Intelligence Review, to appear.

[ZWZD2024] Liangyu Zhong, Lulu Wang, Lei Zhang, Josep Domingo-Ferrer, Lin Xu, Changti Wu, and Rui Zhang, "Dual-server based lightweight privacy- preserving federated learning", IEEE Transactions on Network and Service Management, vol. 21, no. 4, pp. 4787-4800, 2024.

[MRDS2024] Krishnamurty Muralidhar, Steven Ruggles, Josep Domingo-Ferrer, and David Sánchez, "The counterfactual framework in Jarmin et al. is not a measure of disclosure risk of respondents", PNAS-Proceedings of the National Academy of Sciences, 121(11) e2319484121, 2024. https: //doi.org/10.1073/pnas.2319484121.

[JDSB2024] Najeeb Moharram Jebreel, Josep Domingo-Ferrer, David Sánchez, and Alberto Blanco-Justicia, "LFighter: Defending against the label-flipping attack in federated learning", Neural Networks, vol. 174, pp. 111-126, 2024.

[JDBS2024] Najeeb Jebreel, Josep Domingo-Ferrer, Alberto Blanco-Justicia, and David Sánchez, "Enhanced security and privacy via fragmented federated learning", IEEE Transactions on Neural Networks and Learning Systems, vol. 35, no. 5, pp. 6703-6717, 2024.

[SJMD2024] David Sánchez, Najeeb Jebreel, Krishnamurty Muralidhar, Josep Domingo-Ferrer, and Alberto Blanco Justicia, "An examination of the alleged privacy threats of confidence-ranked reconstruction of Census microdata", in Lecture Notes in Artificial Intelligence, vol. 14915, pp. 213-224. Vol. Privacy in Statistical Databases (PSD 2024), Antibes Juan- les-Pins, France, Sep. 25-27, 2024.

[JDL2024] Najeeb Jebreel, Josep Domingo-Ferrer, and Yiming Li, "Defending against backdoor attacks by layer-wise feature analysis (extended abstract)", in Proc. of the International Joint Conference on Artificial Intelligence (IJCAI 2024), Jeju, Korea, pp. 8416-8420, 2024

7.3 Events

Privacy in Statistical Databases (PSD 2024), Antibes Juan-les-Pins, France, Sep. 25-27, 2024. https://crises-deim.urv.cat/psd2024/

8 Conclusions

As witnessed by the vast amount of activities carried out and the variety of topics investigated, the second year and a half of WP10 has been very productive. The introduction of micro-project at the beginning of year two helped significantly organise and manage the activities in WP10, and to track the growth of the platform in terms of items (methods, datasets, experiments, applications) and stories.

We hope to observe an even higher improvement during the remainder of the project, also thanks to the transnational access program, allowing for strengthening the collaborations among the partners of the consortium as well as the collaborations with external institutions, organisations, companies, and researchers.