

CITIUS, USC

Santiago de Compostela, Spain

HYBRIDS JRS I

INITIAL TRAINING MEETING



18-22 SEPTEMBER 2023

Salón de Actos



Funded by
the European Union



UK Research
and Innovation

CITIUS
Centro Singular de Investigación
en Tecnologías Inteligentes

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

PROGRAMME

MONDAY, September 18th

09:00 - 09:10 h Opening

Paula López, CiTIUS Deputy Director
Pablo Gamallo, HYBRIDS Coordinator

09:10 - 10:45 h Presentation of the project and the WPs, Leader of each WP.

Detailed Presentation of the HYBRIDS project scope, goals, and specific work packages (WP) to the DCs by the leaders of each respective WP, including an overview of key objectives, strategies, and implementation approaches.

- Objectives and overview of the research programme, Pablo Gamallo, HYBRIDS Coordinator, CiTIUS- USC
- Use Cases Topics, Pablo Gamallo, HYBRIDS Coordinator, CiTIUS- USC
- WP1: Hybrid Intelligence Approach, Renata Vieira, UEVORA
- WP2: Public Discourse Analysis, Sara Tonelli, FBK
- WP3: Democracy Threats, Gael Dias, UNICAEN
- WP4: Training, Patricia Martín, UDC
- WP5: Impact, Dissemination and Outreach, Rubén Míguez, NEWTRAL
- WP6: Network Management, Raluca Tomoni, CiTIUS-USC
- WP6: Ethics Requirements, Raluca Tomoni, CiTIUS-USC
- Training Programme, Patricia Marín Rodilla, UDC
- Dissemination, Exploitation and Outreach Strategic Plan, Rubén Míguez Pérez, NEWTRAL
- Magnitude of Expected Impact, Pablo Gamallo, HYBRIDS Coordinator, CiTIUS- USC

10:45 - 11:15 h Coffee

11:15 - 13:00 h Presentation of all DCs

Each DC has a presentation of 10 minutes: background, interest in HYBRIDS project and in his/her specific PhD subject.

- DC1: Towards an automatic characterization of micro-level structures of discourse from social networks, Martial Pastor, RU
- DC2: Analysis and measurement of the impact of political discourse in the citizen's opinion, Davide Bassi, CiTIUS-USC
- DC3: Identifying the stance of argumentative opinions in the political discourse, Siddharth Bhargava, FBK
- DC4: Analysis of political propagation campaigns in social networking platforms, Erik Bran, UEVORA
- DC5: Analysis of the role of disinformation to automatically identify hate speech against immigrants on multilingual social environments, Katarina Laken, FBK
- DC6: Detection of toxic political bots in Twitter by analysing textual content, Søren Fomsgaard, UNICAEN
- DC7: Detection of hyperpartisan political news, Michele Joshua, CiTIUS-USC
- DC8: Hybrid data generation and retrieval for the detection of online harassment, Rabiraj Bandyopadhyay, GESIS
- DC9: Claim verification for fact-checking, Rafael Frade, NEWTRAL
- DC10: Detecting harassment on social media beyond isolated messages, Paloma Piot, UDC
- DC11: Cross-lingual claim detection for fact-checking, Rrubaa Panchendrarajan, QMUL

13:00 - 14:30 h Lunch

PROGRAMME

14:30 - 16:00 h Generalisation in social media research: From fact verification to abusive language detection, Arkaitz Zubiaga, QMUL

This talk aims to discuss a number of challenges that one encounters when working with social media data towards enabling generalisability, primarily in the downstream tasks of automated fact-checking and abusive language detection. The talk will cover two angles of issues hindering generalisability in today's social media research: (i) experimental stages from data collection to model development and evaluation, and (ii) dimensions impacting generalisability such as temporal changes in data, differences across social media platforms or language diversity. Recent efforts tackling these will be explored, setting forth open research avenues that can be of interest within the HYBRIDS consortium's objectives.

16:00 - 17:30 h Introduction to the detection and analysis of online hate speech, Sara Tonelli, FBK

The talk will be an introduction to the task of online hate speech detection and analysis. We will discuss what definitions and categorisations of online hate speech current NLP practitioners have adopted. We will also briefly summarise current state-of-the-art approaches for the automated detection of hate speech on social media, and what are the main issues related to this task. In particular, we will discuss annotators' (dis)agreement and the importance of contextual information. Finally, we will briefly mention the relationship between hate speech and misinformation.

TUESDAY, September 19th

09:00 - 10:45 h IAT/ML, a Methodology for Combined Discourse Analysis, Incipit-CSIC (part 1)

There are many approaches to the analysis of discourses, each one emphasising one aspect of the discourse at the expense of others. For example, argumentation analysis focuses on how speakers justify what they say, but pays little attention to their intentions or motivations. Critical analysis, on the other hand, aims to discover the beliefs and intentions of speakers, but avoids a systematic description of the things of the world that the discourse mentions. In this talk I will present IAT/ML (www.iatml.org), a methodology for combined discourse analysis that unites ontological (what the speakers talk about), argumentation (how they justify their claims) and agency (what their beliefs, desires and intentions are) analysis under a common conceptual infrastructure and with the support of a set of software tools. By using a combined approach methodology such as IAT/ML, we can obtain more and deeper insights into the target discourses, understand conflict much better, and support decision making or other tasks that rely on a thorough understanding of multiple points of view.

10:45 - 11:15 h Coffee

11:15 - 12:00 h IAT/ML, a Methodology for Combined Discourse Analysis, César González, Incipit-CSIC (part2)

12:00 - 13:00 h Navigating Rights and Responsibilities: A Guide for MSCA Hybrids Doctoral Candidates, Raluca Tomoni, HYBRIDS Project Manager-CiTIUS-USC

This informative session delves into the rights and responsibilities granted to Doctoral Candidates as part of the MSCA Hybrids project. Exploring both individual and collective aspects, we aim to empower Doctoral Candidates with a comprehensive understanding of their role within the project, fostering a thriving and mutually beneficial research community. Subsequently, the floor will be opened for a dynamic Q&A session, allowing participants to seek clarification and discuss any queries they may have.

PROGRAMME

13:00 - 14:30 h Lunch

14:30 - 17:30 h Separate meetings between each DC and their supervisors

This session comprises separate meetings between Doctoral Candidates (DC) and their respective supervisors. These meetings are designed to kickstart and align the planned work for each DC, providing a tailored approach to their research endeavors within the project.

On the first two days, the only expected attendees are the DCs and (co-)supervisors from HYBRIDS. However, from Wednesday onwards, DCs from the other Doctoral Network (NL4XAI) will join us.

WEDNESDAY, September 20th



08:45 - 09:00 h Opening NL4XAI & HYBRIDS

Senén Barro, NL4XAI Coordinator and CiTIUS Scientific Director
Pablo Gamallo, HYBRIDS Coordinator

09:00 - 10:45 h Science Communication-Expanding Knowledge and Practice, Ángela Muñiz y Tamara Rodríguez, FEUGA (part 1)

In this interactive session, we will explore techniques to effectively engage with key stakeholders and develop targeted communication strategies

- Stakeholders' Analysis: Identify target groups, understand the importance of engaging with them, and explore effective methods of engagement.
- Crafting Key Messages: Learn how to develop impactful messages that resonate with your target audience.
- Choosing the Right Channels: Discover the best communication channels to reach your stakeholders and strategies for maximizing their impact. Gain insights from different examples and lessons learned.
- Building Networks and Partnerships: Explore the power of collaborations and partnerships in science communication and how to leverage them effectively.

10:45 - 11:15 h Coffee

11:15 - 13:00 h Mentoring session (Part I: Project Presentations and Group Formation:)

This mentoring session serves as a collaborative platform for students within both the NL4XAI and HYBRIDS networks. It commences with brief project presentations by the Doctoral Candidates, allowing each to showcase their research endeavors within the two networks. Following these concise introductions (approximately 5 minutes each), participants will be grouped based on their research areas and interests.

13:00 - 14:30 h Lunch

PROGRAMME

14:30 - 16:30 h Mentoring session (Part 2: Group Work and General Discussion:)

In the second session, students will actively engage in group activities that promote knowledge exchange and collaboration. Working within their respective groups, participants will delve into meaningful discussions, sharing insights and leveraging the diverse expertise present in both networks. These interactions aim to enhance the learning experience and enrich the research pursuits of all involved.

Towards the conclusion of the mentoring session, the outcomes of these group discussions will be shared in a joint, comprehensive discussion. This final part provides an opportunity for participants to collectively reflect on the insights gained, fostering a deeper understanding of the various research areas and perspectives within the NL4XAI and HYBRIDS networks.

16:30 - 17:30 h , An overview of neuro-symbolic strategies for natural language processing, María Leonor Pacheco, Purdue University-USA

The goal of neuro-symbolic methods is to combine symbolic representations and neural networks to benefit from the complementary strengths of the two paradigms. On the one hand, symbols have inherent explanatory power, and they can help us express domain knowledge and enforce consistency across different decisions. On the other hand, neural methods are very expressive and can help us identify patterns in large-scale raw data. These ideas have a long history in AI and currently experience a resurgence of interest in several AI communities, including Natural Language Processing. In this talk, I will give an overview of neuro-symbolic strategies and how they have been applied to problems in Natural Language Processing and Computational Linguistics.

18:00 Cultural visit to the USC heritage

THURSDAY, September 21st



09:00 - 10:45 h Science communication-Expanding knowledge and practice, Ángela Muñiz y tamara Rodríguez, FEUGA (part 1)

In this interactive session, we will explore techniques to effectively engage with key stakeholders and develop targeted communication strategies

- Scientific Publications Academy: Understand the role of scientific publications in communication and learn techniques to make them accessible and engaging for broader audiences.
- Monitoring and Evaluation: Learn how to monitor and evaluate the effectiveness of your science communication efforts and make necessary adjustments for continuous improvement.

10:45 - 11:15 h Coffee

11:15 - 13:00 h Round table: Challenges of large language models and explainable artificial intelligence to deal with polluted information

PANELISTS: **Alisa Rieger**, NL4XAI PhD Resercher, University of Delft; **Corrado Mencar**, Associate Professor, Università degli Studi di Bari Aldo Moro; **Javier Díaz Cortés**, Data Scientist, INDITEX; **José Ramom Pichel Campos**, R&D Director, IMAXIN; **Katarina Laken**, HYBRIDS PhD Researcher, FBK; **Lina María Rojas Barahona**, Senior Research Scientist, ORANGE.

MODERATOR: **José María Alonso Moral**, Tenured Professor, CiTIUS-USC

PROGRAMME

13:00 - 14:30 h Lunch

21:00 - Dinner at A MOA, Rúa de San Pedro, 32, 15703 - Santiago de Compostela

FRIDAY, September 22nd



10:00 - 11:45 h , Corpus Annotation: Basic Steps to Plan an Annotation Campaign, Martín Pereira, USC

Robust, well-organised, and machine-readable annotated corpora are valuable resources for qualitative and quantitative inquiries in Social Sciences, Humanities and Computer Science. Nevertheless, the generation of such corpora is a time-consuming and expensive task, demanding a careful planitification for succesful completion. The main goal of this methodological presentation is to outline a set of basic steps, alongside additional tips and tricks, for developing an annotation campaign. We'll begin with the very firsts steps (e.g. selecting the text fro annotation, creating annotation guidelines, etc.) to the very lasts ones, including evaluation or making it publicly available.

11:45 - 12:15 h Coffee break

12:15 - 13:00 Final discussion and farewell