Conference Programme

18th International Conference on Statistical Analysis of Textual Data

8–10 July 2026
Palermo, Italy

Programme at a glance and detailed sessions

Programme may be subject to minor changes.

UNIVERSITÀ DEGLI STUDI DI PALERMO

dSEAS

SIS Società Italiana di Statistica

Regione Siciliana

Città di Palermo

StatPanormus

VADISTAT SIMONA BALBI

Programme at a glance

Wednesday, 8 July 2026

Time	Aula 9	Aula 12	Aula 10
08:30	Welcome		
09:15	Opening Session Rector of the University of Palermo, Prof. Massimo Midiri; Pro-Rector for Internationalization and Academic Planning, Prof. Fabio Mazzola; Head of the Department of Economics, Business and Statistics, University of Palermo, Prof. Marcantonio Ruisi; President of the Italian Statistical Society, Prof. Marcello Chiodi; JADT 2026 Conference Chairs, Prof. Antonella Plaia and Prof. Mariangela Sciandra.		
10:00 - 11:00	Keynote talk: Arjuna Tuzzi - *The Analysis of Textual Data at a Crossroads Between Tradition and Transformation: Are Statistics and Statisticians Still Crucial?* (Aula 9) Chair: Massimo Attanasio		
11:00 - 11:30	Coffee Break		
11:30 - 13:00	**S01** – GLASS Group: Approaches to Textual Data in the Social Sciences Chair: Vincenzo Genova	**C15** – Online communities, conflict and harmful discourse Chair: Pierre Ratinaud	**C12** – Corporate, finance and market intelligence Chair: Andrea Simonetti
13:00 - 14:30	Lunch		
14:30 - 16:00	**S02** – Emotional Text Mining Chair: Francesca Greco	**C09** – Transformers, embeddings and LLM evaluation Chair: Rosario D'Agata	**C14** – Political discourse and democratic communication Chair: Elio Mineo
16:00 - 16:30	Coffee Break		
16:30 - 18:00	**S03** – Embedding-based Approaches to Text Categorization Chair: Domenica Fioredistella Iezzi	**C04** – Institutional, literary and conceptual corpora Chair: Céline Poudat	**C10** – Multimodal, web and user-generated content analytics Chair: Irene Siragusa
20:00	Cocktail Talea - Orto Botanico, Università degli Studi di Palermo		

Thursday, 9 July 2026

Time	Aula 9	Aula 12	Aula 10
09:00 - 10:30	**S04** – Advanced Text Analytics for Official Statistics: Methods, Models, and Applications Chair: Orietta Luzi	**C08** – Topic modeling, classification and interpretability Chair: Laurent Vanni	**C17** – Translation, stylistics and literary variation Chair: Serge Heiden
10:30 - 11:00	Coffee Break		
11:00 - 12:00	Keynote talk: Christof Schöch - *Open Science in Computational Literary Studies* (Aula 9) Chair: Michelangelo Misuraca		
12:00 - 13:30	**S05** – Quantitative Text Analysis Methods for Scientific Literature Chair: Massimo Aria	**C16** – Andrea Camilleri and Sellerio corpora Chair: Giovanni Boscaino	**C11** – Science, technology and knowledge networks Chair: Gianluca Sottile
13:30 - 14:30	Lunch		
14:30 - 16:00	**S06** – Text Analysis in Economics Chair: Davide Furceri	**C03** – Public, political and media discourse Chair: Damon Mayaffre	**C02** – French NLP, lexicon and contrastive neology Chair: Jules Dilé-Toustou
16:00 - 16:30	Coffee Break		

16:30 - 18:00	**C01** – Textometry, corpus methods and visualization Chair: Bénédicte Pincemin	**C05** – Italian institutional and cultural corpora Chair: Andrea Sciandra	**C13** – Health, clinical and accessibility text analytics Chair: Chiara Di Maria
19:45	Social Dinner		

Friday, 10 July 2026

Time	Aula 9	Aula 12	Aula 10
09:30 - 11:00	**S07** – Finance and Textual Analysis Chair: Antonio Scalia	**C06** – Digital tools, annotation and corpus infrastructures Chair: Andrea Marletta	**C07** – Statistical methods and pattern extraction Chair: Vito Rosario Muggeo
11:00 - 11:30	Coffee Break		
11:30 - 12:30	Keynote talk: Guido Vetere - *Linguistic Competence and Large Language Models: Extracting Interpretable Lexical Knowledge* (Aula 9) Chair: Roberto Pirrone		
12:30 - 12:45	Award Ceremony: 2026 VADISTAT – Per Simona Balbi Young Award		
12:45 - 13:00	Closing Ceremony		
13:00	Lunch		

Detailed programme

Each session card reports the date, time and room, followed by the list of contributions.

Wednesday, 8 July 2026

08:30 Welcome

09:15 Opening Session
Rector of the University of Palermo, Prof. Massimo Midiri; Pro-Rector for Internationalization and Academic Planning, Prof. Fabio Mazzola; Head of the Department of Economics, Business and Statistics, University of Palermo, Prof. Marcantonio Ruisi; President of the Italian Statistical Society, Prof. Marcello Chiodi; JADT 2026 Conference Chairs, Prof. Antonella Plaia and Prof. Mariangela Sciandra.

10:00 - 11:00 Keynote talk - Arjuna Tuzzi - *The Analysis of Textual Data at a Crossroads Between Tradition and Transformation: Are Statistics and Statisticians Still Crucial? (Aula 9)*
Chair: Massimo Attanasio

11:00 - 11:30 Coffee Break

11:30 - 13:00

S01 – GLASS Group: Approaches to Textual Data in the Social Sciences

Wednesday, 8 July 2026 **11:30 - 13:00** **Aula 9**
Chair: Vincenzo Genova

Ageing, health and social isolation among older refugees in Italy: an exploratory textual analysis of in-depth interviews
Eleonora Miaci, Eralba Cela and Elisa Barbiano di Belgiojoso

Categorizing Authors in Co-Authorship Networks Through Textual Information Analysis: A Comparative Study of Two Scientific Communities
Francesco Santelli, Domenico De Stefano, Alice Tontodimamma, Francesco Capone and Luciana Lazzeretti

From Keywords to Networks: Analysing Mission Statements with Text Mining Methods
Kristijan Breznik

Integrating Synthetic and Human Data for Complex Psychosocial Categories in Text Classification
Valentina Rizzoli, Alessandro Meneghini and Andrea Sciandra

Robust community detection for topic identification in narratives of financial crisis
Valeria Policastro, Giuseppe De Luca, Maria Carmela Schisani and Giancarlo Ragozini

C15 – Online communities, conflict and harmful discourse

Wednesday, 8 July 2026 **11:30 - 13:00** **Aula 12**
Chair: Pierre Ratinaud

Engaging in interactions: conflicts and harmony on Wikipedia pages
Cyril Bruneau, Delphine Battistelli and Céline Poudat

Exploration contrastive de constructions préfabriquées en français et en allemand – les formules de clôture polémiques dans les discussions Wikipédia
Alexia Jingand, Céline Poudat and Anja Smith

Spherical analysis of communities and AI counter-narratives against online abusive content
Jai Jobe, Alex Cucco, Stefania Fensore, Annalina Sarra, Lara Fontanella and Marco Di Marzio

Dive Into Anything: A Framework for Wide-Scale Analysis on Reddit Based on a Case Study of Audience Discussions of Sexual Violence On Screen
Marceau Hernandez, Léa Andolfi and Virginie Julliard

Discours polémiques sur la vaccination sur YouTube : analyse textométrique comparative des vaccins humains et animaux
Lucie Loubere

C12 – Corporate, finance and market intelligence

Wednesday, 8 July 2026 **11:30 - 13:00** **Aula 10**
Chair: Andrea Simonetti

The use of carbon and GHG vocabulary by French banking groups: a comparative study from 2013 to 2024
Benoit Jamet

E.A.G.L.E: An LLM-Based Framework for Automated Assessment of Banks' Explanatory Notes in EU-Wide Stress Testing
Christoffer Kok, Anastasios Petropoulos, Javier Arranz Barquero and Alexandros Kouris

Assessing Corporate Sustainability Alignment through NLP: From Narrative Disclosure to Actionable ESG Scores
Nicolò Biasetton, Luigi Salmaso and Mahir Selek

Financial News as Textual Signals: A FinBERT-Based NLP Approach to Tesla Sales Forecasting
Francesco Altarocca, Domenica Fioredistella Iezzi, Roberto Monte and Daniele Pasquini

A Reinforcement Learning Meta Controller for Multimodal Stock Selection Using LLM Sentiment
Chaher Alzaman

Cassandra: an automated AI-based surveillance system to forecast bank distress from financial news
Christoffer Kok, Anastasios Petropoulos, Javier Arranz Barquero and Alexandros Kouris

13:00 - 14:30 Lunch

14:30 - 16:00

S02 – Emotional Text Mining

Wednesday, 8 July 2026 **14:30 - 16:00** **Aula 9**
Chair: Francesca Greco

Voluntary Return: A Choice, a Socio-Cultural Impact.
Simona Rita Coco

Organizational Analysis in Public Administration: between traditional and digital tools for research on public managers
Giulia Cecchini

Multimodal Word–Context Analysis: Integrating Semantic, Emotional, and Visual Context in Political Discourse
Silvia Torta, Andrea Simonetti and Michele Tumminello

Automated Clinical Coding with Large Language Models
Irene Siragusa, Marco Speciale, Alessandro Albano, Roberto Pirrone and Antonella Plaia

C09 – Transformers, embeddings and LLM evaluation

Wednesday, 8 July 2026 **14:30 - 16:00** **Aula 12**
Chair: Rosario D'Agata

Decoding Self-Attention : How frequency matrices drive representation learning ?
Laurent Vanni, Dominique Longrée and Damon Mayaffre

A KAN-based non-linear factorization for Embedding Layer Compression
Oliver Giudice and Anna Russo Russo

Beyond ROUGE and BLEU: A Systematic Review of Evaluation Metrics for LLM-Generated Text
Nicolò Biasetton, Luigi Salmaso, Mahir Selek and Rosa Arboretti

Human perspectives in textual classification models
Emiliano del Gobbo, Alex Cucco, Roberta Siciliano and Lara Fontanella

Beyond frequency measures: Can contextual embeddings capture meaning change in scientific texts?
Jianying Liu, Kim Gerdes and Jean-Marc Deltorn

Analyzing the Impact of LLM-Based Augmentation on Document Classification
Vittorio Torri and Francesca Ieva

C14 – Political discourse and democratic communication

Wednesday, 8 July 2026 **14:30 - 16:00** **Aula 10**
Chair: Elio Mineo

Constructing Climate Change Discourse in the Digital Public Sphere: A Computational Analysis of Social Representations
Michelangelo Misuraca, Maria Spano and Luca D'Aniello

Temporal Dynamics of AI News Frames: A Multilayer Network Approach
Massimo Aria, Manuel J. Cobo, Luca D'Aniello, Michelangelo Misuraca and Maria Spano

Speeches in Austrian Parliament: Simpler and more active when party is in opposition
Michaela Nogolová, Michaela Koščová, Ján Mačutek and Radek Čech

Historical shifts in partisan language use in US presidential elections: A comparative analysis of televised debates from 2008 to 2024
Nozomu Hamada and Kaori Sugiyama

Argumentation Today: Opacity in CFA from a Brazilian Discourse Analysis Perspective
Mariana Alves

16:00 - 16:30 Coffee Break

16:30 - 18:00

S03 – Embedding-based Approaches to Text Categorization

Wednesday, 8 July 2026 **16:30 - 18:00** **Aula 9**
Chair: Domenica Fioredistella Iezzi

A Privacy-by-Design Framework for Chatbot Conversation Anonymisation with Hybrid NLP and Bootstrap-Based Risk Assessment
Samanta Pietropaoli

AI for Automatic Coding at Istat: From Semantic Models to Institutional Governance
Massimo Fedeli, Orietta Luzi, Mauro Bruno, Simona Pace and Sara Letardi

From Stylometric Patterns to Semantic Spaces: A Multi-Methodological Mapping of Global Corruption Discourse
Simona Mercurio and Giancarlo Carbone

Is there any AI-generated news in an online newspaper? An AI detection model for authorship attribution.
Andrea Sciandra

Power Dynamics in Relational Recovery from Severe Mental Illness: A Computational Analysis of Therapeutic Dyad Interviews
Safa Yassin, Matilde Trevisani, Arjuna Tuzzi, Elena Faccio, Ludovica Aquili, Michele Rocelli and Stefano Sbalchiero

C04 – Institutional, literary and conceptual corpora

Wednesday, 8 July 2026 **16:30 - 18:00** **Aula 12**
Chair: Céline Poudat

Datation de textes littéraires - Comparaison et interprétation de modèles de classification
Baptiste Bohet and Nicole Vincent

ADT vs Deep_Learning dans la recherche de l'intertexte : le cas de Jérôme de Stridon - Arianna Rosa
Arianna Rosa

CEReBRA : une approche guidée par embeddings contextuels et règles d'association pour l'analyse computationnelle de concepts philosophiques
Abdoulaye Ouedraogo, Francis Lareau, Ismail Biskri and Jean-Claude Simard

Analyse textuelle sur la diversité dans les entreprises canadiennes entre 2019 et 2022
Olivier Greusard, Marc Le Pouliquen and Nadia Tahiri

La "contractualisation" de l'université vue par ses acteurs : analyse diachronique d'un corpus de comptes rendus de conseils d'administration
Frédérique Sitri, Sascha Diwersy, Etienne Bordes, Hugo Dumoulin and Cyrielle Montrichard

C10 – Multimodal, web and user-generated content analytics

Wednesday, 8 July 2026 **16:30 - 18:00** **Aula 10**
Chair: Irene Siragusa

A Modular NLP Pipeline for Analyzing Spontaneous Multilingual Reviews in Urban Event Contexts
Sandro Stancampiano

From Narratives to Decisions: Fuzzy Text Categorization and Interpretable Market Intelligence from Consumer Review Corpora
Milena Nikolic, Teodor Kern, Marina Marjanovic and Srdjan Jovkovic

BRIDGING THE TOOL GAP IN VISUAL SOCIAL MEDIA RESEARCH: SMIQQDA, A FREE AND OPEN-SOURCE SOLUTION
Tiago Joseph, Michaël Stappers and Catherine Bouko

User-Generated Content analysis of cycling lanes: exploring spatial and temporal patterns in public acceptance
Alessandra D'Alessio, Dario D'Ingiullo, Lara Fontanella and Massimo Aria

20:00 Cocktail
Talea - Orto Botanico, Università degli Studi di Palermo

Thursday, 9 July 2026

09:00 - 10:30

S04 – Advanced Text Analytics for Official Statistics: Methods, Models, and Applications

Thursday, 9 July 2026 **09:00 - 10:30** **Aula 9**
Chair: Orietta Luzi

A Methodological Framework for Knowledge Graph Design Using Large Language Models
Adriana Fidanza, Raffaella Maria Aracri, Alessio Buonocore and Roberta Radini

Exploring User–Chatbot Communication: Methods, Insights, and Channel Differences
Samanta Pietropaoli, Salvatore Agrillo, Paola Bosso, Barbara D'Amen and Gabriella Fazzi

From textual agreements to structured data: using Large Language Models to enhance Official Statistics on anti-violence networks
Elena Catanese, Mauro Bruno, Paolo Pizzo, Giuseppina Muratore and Claudia Villante

Semantic Search Approach to Occupation Coding in Official Statistics
Giulio Massacci and Mauro Bruno

Using AI techniques to support cause of death coding for Official Statistics
Angela Pappagallo, Francesco Pugliese, Chiara Orsi, Tania Bracci, Francesco Grippo and Simone Navarra

C08 – Topic modeling, classification and interpretability

Thursday, 9 July 2026 **09:00 - 10:30** **Aula 12**
Chair: Laurent Vanni

Editorial Transitions and Thematic Change in the American Journal of Sociology: A Structural Topic Modeling Analysis
Stefano Sbalchiero, Valentina Rizzoli and Arjuna Tuzzi

Titles, Abstracts, or Full-Texts? A Comparative Study of LDA and BERTopic Performance across Text Levels
Francis Lareau and Christophe Malaterre

A Multi-Method Validation Framework for Large-Scale Multilingual Text Analytics
Stefano Blando and Domenica Fioredistella Iezzi

Transforming UN Internal Audit Reports into Oversight Knowledge: Lexicometric Perspective
Li Chen and Domenica Fioredistella Iezzi

Comparing textometric approaches: Reinert-style Descending Hierarchical Classification versus probabilistic and neural topic modeling on a web-sourced corpus of French rap lyrics
Robin Quillivic, Jules Dilé-Toustou and Pierre Ratinaud

C17 – Translation, stylistics and literary variation

Thursday, 9 July 2026 **09:00 - 10:30** **Aula 10**
Chair: Serge Heiden

Demonstratives in human and machine translation between English and French
Daniel Henkel

Beyond Questions: A Corpus-Based Analysis of Interrogatives and Their Responses in Spoken French
Yeo Jun Yun, Iris Eshkol-Taravella and Marianne Desmets

The French Drama Revolution: Political Economy and Literary Production (1700-1900)
Thiago Dumont Oliveira

From Lyrics to Values: A Computational–Sociological Analysis of Cultural Change in Italian Popular Music
Ciro Clemente De Falco and Emilia Romeo

A Computational Authorship Attribution in the Children's Tales of Oscar and Constance Wilde: The Case of "The Selfish Giant"
Liviu P. Dinu, Alina Iacob and Cosmin Ciotloș

10:30 - 11:00 Coffee Break

11:00 - 12:00 Keynote talk - Christof Schöch - *Open Science in Computational Literary Studies (Aula 9)*
Chair: Michelangelo Misuraca

12:00 - 13:30

S05 – Quantitative Text Analysis Methods for Scientific Literature

Thursday, 9 July 2026 **12:00 - 13:30** **Aula 9**
Chair: Massimo Aria

Assessing BERTopic stability on a specialized scientific corpus: seed sensitivity, outlier reduction, and temporal semantic drift
Carolina Vasconcelos, Sandro Mendonça and Bruno Damasio

Enhancing leader-based community detection in scientific collaboration using publications metadata
Geremia Sara and Domenico De Stefano

Mapping Communicative Purposes in Academic Discourse: designing an analytic linguistic pattern
Alessandra Belfiore, Walter Giordano, Massimo Aria and Corrado Cuccurullo

Mapping the Conceptual Landscape of Funded Computer Science Research in Spain
Manuel Jesus Cobo Martin, Verónica Duarte Martínez, Luis M. Castillo-Chávez and Jose A. Moral-Munoz

Modeling Emotion Constructs in Consumer Reviews: A Word2Vec–ELIta Hybrid Method Combined with PLS-SEM
Cosima Valentino

C16 – Andrea Camilleri and Sellerio corpora

Thursday, 9 July 2026 **12:00 - 13:30** **Aula 12**
Chair: Giovanni Boscaino

A New Investigation for Sellerio
Arjuna Tuzzi

Camilleri's Sicilian: From Stopwords To Topic Modelling
Vincenzo Giuseppe Genova and Andrea Sciandra

Classification of Andrea Camilleri's works using bottom-up and top-down approaches.
Andrea Sciandra and Michele A. Cortelazzo

Distinctive Lexical Traits in Andrea Camilleri: A Keyness Measures Analysis of a Corpus of Sellerio Crime Novels
Michele A. Cortelazzo, Matilde Trevisani and Arjuna Tuzzi

Assunti teorico-metodologici su linguaggio letterario esclusivo e dati testuali
Pietro Mazzarisi

C11 – Science, technology and knowledge networks

Thursday, 9 July 2026 **12:00 - 13:30** **Aula 10**
Chair: Gianluca Sottile

Language as a Predictor of Scholarly Citation in European Policy Documents
Francis Lareau, Jérémie Dion, Louis Renaud-Desjardins, Malvina Barra, Luisa Ripoll-Alberola and François Claveau

Mapping the languages of science: a comparative cross-media textual analysis of scientific expertise.
Gabriella Punziano, Rosanna Cataldo, Barbara Saracino and Ferdinando Iazzetta

Discourse structure across technological maturity levels: a textometric approach
Mélusine Caillard and Gaël Lejeune

Structuring sectorial knowledge: automated detection of lexical relations
Claudia Brunini, Eugenio Concetti and Giampiero D'Alessandro

A hybrid rule-based and data-driven pipeline for author name normalization in IRIS institutional repositories
Pasquale Pavone, Simone Righi and Sara Colombini

Identifying Knowledge Gaps in Textual Networks: A Structural Holes Approach to Literature-Based Discovery
Andrej Kastrin, Bojan Cestnik, Alja Nike Kastrin and Nada Lavrač

13:30 - 14:30 Lunch

14:30 - 16:00

S06 – Text Analysis in Economics

Thursday, 9 July 2026 **14:30 - 16:00** **Aula 9**
Chair: Davide Furceri

AI Meets Fiscal Policy: Mapping Government Spending Actions Across 64 Countries
Davide Furceri

BILATERAL CONFLICT RISK AND TRADE: MILITARY WARS, TRADE WARS, AND DIPLOMATIC NOISE
Jamel Saadaoui

From Fed Speeches to Signals Embedding Central Bank Communication in Macroeconomic Forecasting
Angelo Fama'

Global and local effects of Federal Reserve communication: uncertainty and real activity
Antonio Pietro Maria Morreale

C03 – Public, political and media discourse

Thursday, 9 July 2026 **14:30 - 16:00** **Aula 12**
Chair: Damon Mayaffre

Détection des isotopies du discours présidentiel d'Emmanuel Macron : de la cooccurrence à la self-attention
Damon Mayaffre and Laurent Vanni

Les formes du futur en planification urbaine. Analyse textométrique des temps verbaux en planification régionale (Île-de-France, 1919-2024)
Isabelle Chesneau

Investigation de l'évolution des discours médiatiques à propos des réseaux socionumériques à partir de la distance de Labbé et de sa version améliorée
Pierre Ratinaud and Brigitte Sebbah

Caractériser les commentaires en ligne : les émojis dans la loi « Duplomb »
Pascal Marchand and Pierre Ratinaud

Étudier les vidéos TikTok : comparaison lexicométrique des plus-values et des limites des retranscriptions automatiques versus les descriptions des vidéos.
Panos Tsimpoukis, Lucie Loubere, Nikos Smyrnaios, Pierre Ratinaud and Charlis Papaevangelou

C02 – French NLP, lexicon and contrastive neology

Thursday, 9 July 2026 **14:30 - 16:00** **Aula 10**
Chair: Jules Dilé-Toustou

Vers une correction automatique des formes inconnues : application du Masked Language Modeling de CamemBERT au corpus des Cahiers citoyens
Sami Guembour, Solomiia Korol and Catherine Dominguès

Séjour à l'étranger et développement des marqueurs discursifs en FLE : étude sur corpus chez les apprenants japonophones
Kaori Sugiyama

Le potentiel lexicographique des néologismes des TIC dans une perspective franco-slovaque : analyse par le biais de Sketch Engine
Lucia Rackova

Vers un chatbot basé sur la recherche sémantique et le raisonnement à base de règles pour l'optimisation des opérations de la Sécurité Sociale française
Wenxiang Zhou, Touria Aït El Mekki, Frédéric Lardeux, Marc Legeay and Jean-Olivier Boesmier

L'extraction semi-automatique de termes métaphoriques grâce à la spécificité sémantique : étude appliquée à un corpus français-italien de presse spécialisée dans le domaine de la bijouterie-joaillerie
Marius François and Maria Grazia Massimo

Utilisation des emprunts à l'anglais de type V-ING en espagnol, français et italien : une étude en diachronie courte basée sur les corpus du web du Wortschatz de Leipzig
François Maniez, Maria Belén Villar Diaz and Sidonie Larato

16:00 - 16:30 Coffee Break

16:30 - 18:00

C01 – Textometry, corpus methods and visualization

Thursday, 9 July 2026 **16:30 - 18:00** **Aula 9**
Chair: Bénédicte Pincemin

Articuler méthodes quantitatives et épistémologies critiques en linguistique : enjeux, tensions et perspectives méthodologiques
Tiago Joseph

A la croisée des disciplines informatique et psychanalytique : une application sémiotique de la méthode Reinert
Claire Tissot

De la mesure des paragraphes à la classification des visualisations : portée heuristique et démontrative
Claude Patricia Tardif

De l'analyse de données en cartes : Visualisations géographiques et chronologiques de corpus textuels
Jean-Marc Leblanc, Anne-Lise Humain-Lamoure and Marie Pérès-Leblanc

Etude des cooccurrences du verbe Dire dans un corpus d'entretiens à propos d'un événement traumatique
Anna Colli, Delphine Battistelli and Marie Chagnoux

Europarser: exploiter de larges corpus de presse
Marceau Hernandez, Gaël Lejeune, Félix Alié, Émile Rebours and Julien Bezançon

C05 – Italian institutional and cultural corpora

Thursday, 9 July 2026 **16:30 - 18:00** **Aula 12**
Chair: Andrea Sciandra

La sfida del lavoro da remoto. Analisi di focus group a responsabili delle risorse umane
Francesca della Ratta

Un'analisi testuale dei discorsi parlamentari italiani
Matteo Farnè and Marco Nannetti

A Multilevel Analysis of Italy's Three-Year School Planning Documents Using a Lexical and Distributional-Semantic Framework
Pasquale Pavone, Maria Teresa Sagri, Elettra Morini and Eniko Tolvay

The Lexical Construction of Women's 'Image': Party Specificity and Gender Asymmetry in Italy's Constituent Assembly
Amin Gino Fabbrucci Barbagli, Francesca Aulino, Rosario D'Agata and Domenico De Stefano

Perceptions of Emilia-Romagna: a cross-linguistic analysis of keywords in English and French travel literature
Rita Gramellini and Daniel Henkel

C13 – Health, clinical and accessibility text analytics

Thursday, 9 July 2026 **16:30 - 18:00** **Aula 10**
Chair: Chiara Di Maria

Living with osteoporosis across contexts: a lexicometric comparison of Italian and Spanish narratives
Chiara Tedesco, Ercole Vellone, Rosaria Alvaro, Gianluca Pucciarelli, Vicente Bernalte-Martì and Mariachiara Figura

Self-care behaviours in people with osteoporosis: a multidimensional analysis of textual data
Chiara Tedesco, Ercole Vellone, Rosaria Alvaro, Gianluca Pucciarelli, Vicente Bernalte-Martì and Mariachiara Figura

Text Mining Applications in the Evaluation of AIFA Drug Innovativeness Reports in Italy
Daniele Pirotta, Andrea Marletta, Carlotta Galeone and Paolo Mariani

Characterization of texts written by deaf people through syntactic markers
Mireille Esther Gettler Summa, Caroline Bogliotti and Anne Lacheret Dujour

19:45 Social Dinner

Friday, 10 July 2026

09:30 - 11:00

S07 – Finance and Textual Analysis

Friday, 10 July 2026 **09:30 - 11:00** **Aula 9**
Chair: Antonio Scalia

Web Automatic Investigator (WAI): an Agentic Framework for Regulatory Multimodal Analysis of Web Contents
Andrea Coletta, Miriam Corrado, Gabriele De Capoa, Oliver Giudice, Alessandro Maggi, Marco Mori, Anna Russo Russo and Valerio Signore

Interpretable Natural Language Concept Bottleneck Models for Default Risk estimation
Andrea Conti and Giacomo Morelli

Word2Prices: embedding central bank communications for inflation prediction
Douglas Araujo, Nikola Bokan, Fabio Alberto Comazzi and Michele Lenza

The cyber risk of non-financial firms
Francesco Columba, Manuel Cugliari, Marco Orlandi and Federica Vassalli

Processing instability and the value of information: from individual noise to common risk
Rudraksh Kansal and Taejin Park

C06 – Digital tools, annotation and corpus infrastructures

Friday, 10 July 2026 **09:30 - 11:00** **Aula 12**
Chair: Andrea Marletta

A Text analysis application: a library-led introductive course open to all University users
Deborah Grbac and Valentina Schiariti

JADT Anthology: decades of textual data analysis in one corpus.
Laurent Vanni

ActiveTigger: An open source collaborative text annotation software for computational social sciences
Emilien Schultz, Julien Boelaert, Axel Morin, Emma Bonutti D'Agostini, Annina Claesson, Étienne Ollion and Arnault Chatelain

ParaTaxe: a Suite of Tools for Processing and Publishing Multi-parallel Corpora
Olivier Kraif, Cyrille François, Elnaz Jalilian, Natacha Rimasson Fertin and Pascale Roux

Analyse de textualités multiples avec TXM : du texte brut aux éditions TEI, transcriptions synchronisées, traductions alignées et micro-textes
Serge Heiden

C07 – Statistical methods and pattern extraction

Friday, 10 July 2026 **09:30 - 11:00** **Aula 10**
Chair: Vito Rosario Muggeo

The UMCC: Usage and Measurement of Connectors in Context. A methodology and interface for measuring the use of connectors in different genres of corpora
Corinne Rossari, Claudia Ricci, Linda Sanvido, Jonathan Schaber and Gerold Schneider

The Motif Work Bench Pipeline: From Sequential Pattern Mining to Multidimensional Statistics
Hugo Dumoulin and Timothée Premat

The Moving Border Model: Applying Chronological Specificities to the Tuning of a Time Boundary
Bénédicte Pincemin, Céline Guillot-Barbance and Alexei Lavrentiev

Validation of Simultaneous Additive Trees in Textual Data Visualization
Ludovic Lebart

Pulling strings: scalable symbolic methods for Multi-Word Units detection in text corpora
Antonio Balvet

11:00 - 11:30 Coffee Break

11:30 - 12:30 Keynote talk - Guido Vetere - *Linguistic Competence and Large Language Models: Extracting Interpretable Lexical Knowledge (Aula 9)*
Chair: Roberto Pirrone

12:30 - 12:45 Award Ceremony: 2026 VADISTAT – Per Simona Balbi Young Award

12:45 - 13:00 Closing Ceremony

13:00 Lunch