



15-16 SEPTEMBER 2021 FNP 2021

Welcome to the 3rd Financial Narrative Processing Workshop (FNP 2021). This year we are running a 2-day event sponsored and funded by the Data Science Institute (DSI) at Lancaster University and Yseop. This is an international gathering of researchers and keynote speakers working on Financial Narratives from computing, accounting & finance.

SHARED TASKS COMPETITIONS

The winning team from each shared task will receive a certificate of achievement and a prize voucher of US\$650.

- Financial Narrative Summarisation (FNS 2021)
- Causality Identification in Financial documents (FinCausal 2021)
- Financial Document Structure Extraction (FinTOC 2021)

Keynote Speakers:



Emmanuel Walckenaer
CEO of Yseop, France

Prof Jochen L Leidner
Sheffield University, UK



Dr Kim Trottier
HEC Montréal, Canada

Dr Claire Hardy

Lancaster University, UK



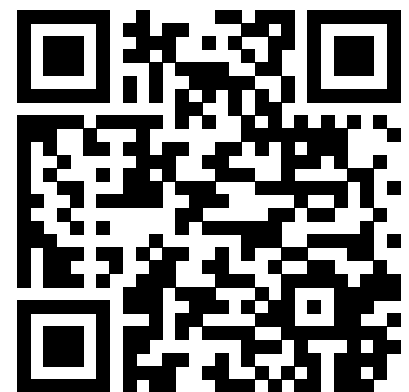
Prof Steven Young
Lancaster University, UK

Prof Antonio Moreno
UAM, Madrid, Spain



Dr Vasiliki Athanasakou
Saint Mary's University,
Canada

Dr Raymond Ng
Signal AI, London, UK



<http://wp.lancs.ac.uk/cfie>

Data Science
Institute

Lancaster
University



yseop



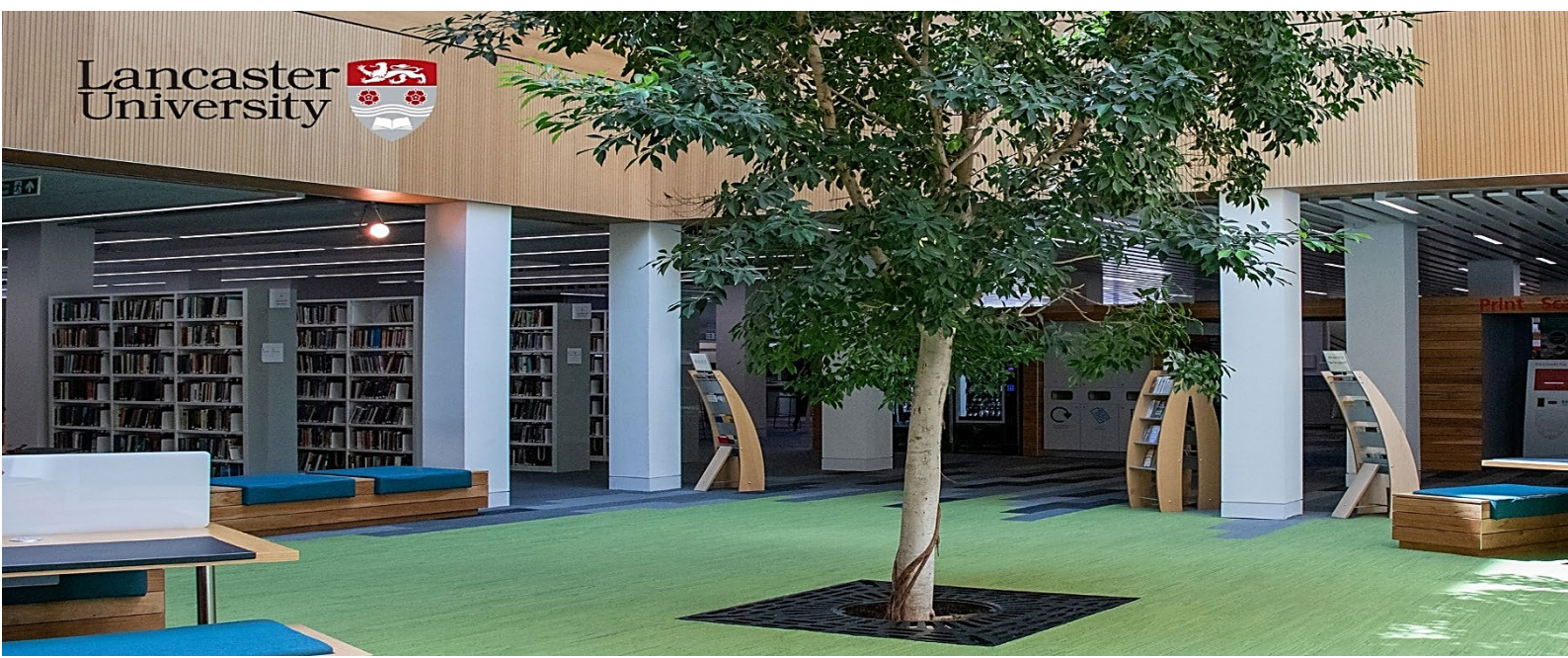
FNP 2021

15-16 SEPTEMBER 2021

Programme

Day 1 – Wednesday - 15/Sept/2021

09:45 – 10:00	Welcome and opening remarks – Dr Mahmoud El-Haj
10:00 – 10:40	Jochen L. Leidner – Sheffield University, UK
10:45 – 11:10	Coffee Break
11:10 – 11:50	Antonio Moreno – Autonomous University of Madrid, Spain
11:55 – 12:35	Vasiliki Athanasakou – Saint Mary's University, Canada
12:40 – 13:40	Lunch Break
13:40 – 14:20	Steven Young – Lancaster University, UK
14:25 – 15:05	Claire Hardy – Lancaster University, UK
15:10 – 15:50	Emmanuel Walckenaer – YSEOP, France
15:55 – 16:15	Coffee Break
16:15 – 16:55	Raymond Ng – Signal AI, UK
17:00 – 17:40	Kim Trottier - HEC Montreal, Canada
17:45 – 18:30	Panel Discussion



FNP 2021

15-16 SEPTEMBER 2021

Programme

Day 2 – Thursday - 16/Sept/2021

09:00 – 09:10	FinCausal Shared Task Overview
09:15 – 09:30	FinCausal Shared Task Winning Team
09:35 – 09:45	FinTOC Shared Task Overview
09:50 – 10:05	FinTOC Shared Task Winning Team
10:10 – 10:20	FNS Shared Task Overview
10:25 – 10:40	FNS Shared Task Winning Team
10:45 – 11:15	Coffee Break
11:15 – 11:30	An annotation model and a corpus for opinionated economy and finance narrative detection (Jiahui Hu, LISN, Natixis, France)
11:35 – 11:50	Data Driven Content Creation using Statistical and Natural Language Processing Techniques for Financial Domain (Sohom Ghosh, Fidelity Investments, India)
11:55 – 12:10	A sequence to sequence transformer data logic experiment (Daxin Cui, Université Sorbonne Nouvelle, France)
12:15 – 12:30	Economic Causal-Chain Search and Numerical Forecast using Textual Data (Kiyoshi Izumi, The University of Tokyo, Japan)
12:35 – 12:50	Preliminary experimentation with combinations and extensions of forward-looking sentence detection wordlists (Jan Štihec, University of Ljubljana, Slovenia)
12:55 – 13:10	On Estimating Sentence Complexity of Domain-Specific Single Sentences: Are Complexity Factors Relevant to Automated Agents Relevant to Human Agents? (Esme Manandise, Intuit Futures, USA)
13:15 – 14:30	Lunch Break
14:30 – 15:30	FinCausal Shared Task Poster Session (Gather Town)
15:30 – 16:30	FinTOC Shared Task Poster Session (Gather Town)
16:30 – 17:30	FNS Shared Task Poster Session (Gather Town)
17:30 – 17:40	Closing remarks

FNP 2021 Keynote Speakers

Day 1: 15/September/2021

Talk time: 10:00 – 10:40

All times in UK Summer Time (GMT +1)



Speaker: [Prof Jochen L Leidner](#) – [Coburg University of Applied Sciences](#), German | [University of Sheffield, UK](#)

Title: Financial Narratives, Information Asymmetry and the Hunt for Alpha

Abstract: In this talk, I present an overview of financial trading and the role of information and its asymmetry in it. In particular, I describe a few sources of information that financial professionals regularly consult, as well as associated work to provide computational support for analysing and exploiting these sources for financial gain, in particular focusing on the world of "unstructured" (textual) language.

Bio: Jochen is Professor of Explainable and Responsible Artificial Intelligence in Insurance at Coburg University of Applied Sciences, Germany and a Visiting Professor of Data Analytics, Department of Computer Science, University of Sheffield, UK. He is also founder and Director of Polygon Analytics and KnowledgeSpaces. He has over a decade of scientific and managerial experience in the legal and financial data industry in the US, the UK and Switzerland. He holds a MPhil from the University of Cambridge and a PhD from the University of Edinburgh, and is twice winner of the Thomson Reuters "Inventor of the Year" award.

Talk time: 11:10 – 11:50



Speaker: [Prof Antonio Moreno](#) – [UAM](#).Autonomous University of Madrid, Spain

Title: Annotating discourse markers and key financial terms in Spanish with transformers

Abstract: Semantically processing CEO letters to shareholders is a complex and difficult task, both for human experts and for Artificial Intelligence systems. The reasons are several: a) there are far fewer examples of loss-making companies than profit-making ones (distribution 15 / 85); the narrative of loss-making companies are very similar to those of profit-making companies, as managers know how to mask bad news so that it does not affect the credibility and solvency of the companies they run. This affects both opinion analysis and the automatic classification of profitable and loss-making companies from the LTS corpus. In this talk we explore the use of key terms and discourse markers to look for linguistic differences in each group. We have developed financial term and DM taggers, using supervised ML techniques. From the manual annotation by linguists of these units, we use transformer-based deep neural networks that give us preliminary results of over 90% accuracy.

Bio: Antonio is a Professor of Linguistics and Director of the Computational Linguistics Lab at University Autonomous of Madrid (UAM), and Director of the UAM-IIC Chair in Computational Linguistics. He is a former Fulbright postdoc scholar at the Computer Science Dept., New York University and a former DAAD scholar at Augsburg Universität.

Talk time: 11:55 – 12:35



Speaker: [Dr Vasiliki Athanasakou](#), [Saint Mary's University](#). Halifax, Canada

Title: The “Shakespeare” of Annual Reports: The Emergence and Role of Corporate Reporting Agencies

Abstract: We examine the role of corporate reporting agencies (CRAs) in shaping the content of annual report narratives in a sample of UK listed firms. CRAs offer services varying from printing and formatting for publication to high-level content-shaping services, such as advice on annual report structure, messaging consistency, and compliance with regulatory requirements and best practice. Our analysis shows that the use of CRAs is widespread among UK firms, and that CRAs offering content-shaping services have a meaningful share of the market. Use of content-shaping CRAs services increases with new reporting guidelines on annual report content and eligibility for annual reviews conducted by financial reporting regulators. After controlling for firm fixed effects, we document a distinct impact of CRA-hiring on annual report properties. Content-shaping CRAs are associated with more readable annual reports, and lengthier and more consistent discussions of the value creation process. Further analysis shows that content-shaping CRAs help firms reach what investors view as optimal disclosure levels, by reinforcing firm-level learning over time and reducing information overload. Taken together, our results suggest that CRAs act as a source of “external assurance” on annual report narratives for compliance with reporting guidelines and best practice. **Co-authors:** Yasmine Chahed (University of Manchester) Lisa Goh (Hang Seng University of Hong Kong)

Bio: Vasiliki is an associate professor in accounting at Saint Mary's University and has a track record of working with financial narratives, financial reporting quality and the impact of regulation on capital markets and organisational performance.

Talk time: 13:40 – 14:20



Speaker: [Prof Steven Young](#) – [Lancaster University](#). Lancaster, UK

Title: Talking different languages? Identifying and realizing the synergies between NLP and financial markets research

Abstract: Financial markets are a rich source of machine-readable discourse, much of which comes ready-labelled (e.g., disclosures by fraud firms and takeover targets, topic-defined commentaries in press releases, award-winning annual reports, etc.). Parallel streams of research examining financial market data are evolving in computational linguistics and finance but the level of cross-fertilization remains low. I explore two related themes with the aim of promoting greater levels of cross-fertilization. The first theme reviews the research questions and agendas in finance where input from NLP specialists is required. The second aspect explores the resources and insights available from finance that provide new research opportunities for NLP specialists. I also examine the barriers limiting interdisciplinarity and propose practical strategies for encouraging greater dialogue between researchers in finance and computational linguistics.

Bio: Steve is a Professor in accounting at Lancaster University and has a track record of working with financial narratives, financial reporting quality and has strong links with the UK's FRC, FCA, PLSA and IRS.

Talk time: 14:25 – 15:05



Speaker: **Dr Claire Hardy** – [Lancaster University](#). Lancaster, UK

Title: From Financial to Female Employee Reproductive Health – Exploring Applications of Narrative Processing Across Disciplines and Topics Areas

Abstract: This keynote presentation will share the opportunities for narrative processing beyond those in the financial sector. In particular, Dr Hardy will present the ideas and collaborative working that is currently underway with colleagues working in the financial narrative field and how multidisciplinary working in the fields of computing and organisational health and wellbeing can help address key issues around health and wellbeing in the context of work and beyond.

Bio: Dr Claire Hardy is a Senior Lecturer in Organisational Health and Wellbeing at Lancaster University's Centre for Organisational Health and Wellbeing within the Faculty of Health and Medicine. Her research areas range from understanding and aiding employees who work internationally, employee stress and burnout, to reproductive health issues at work, and developing and evaluating workplace health interventions.

Talk time: 15:10 – 15:50



Speaker: **Emmanuel Walckenaer**, CEO, [Yseop](#), Paris, France

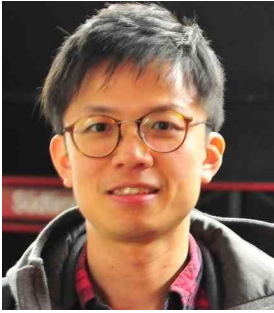
Title: Industry-scale Natural Language Generation: managing customer expectations and technical constraints

Abstract: Automatic generation of financial narratives is a challenging task: the system has to provide variability in the texts it produces while sticking to the customer's writing habits and being 100% faithful to the data it describes.

Yseop is a NLG specialist for more than 10 years, which focuses on the financial and pharmaceutical markets. We will present our new Augmented Analyst platform, an easy-to-use NLG platform for business experts that provides a no-code interface to a powerful, hybrid AI system. The system actually mixes a linguistically-sound deterministic approach of NLG with machine-learning components and domain-specific ontology to provide user and domain adaptation. Our platform is currently deployed at more than 30 Fortune 500 companies around the world.

Bio: Emmanuel joined Yseop as CEO in 2017, leading the company's growth and vision to bring the benefits of automation and NLG to enterprise companies globally. He brings over 25 years of international experience in high-tech service and business development. Emmanuel joined Yseop in 2017 from Sierra Wireless, where he was General Manager and Senior Vice President of the Cloud & Connectivity Services business unit. He led the growth of this unit across Europe and North America through strategic acquisitions and the development of a bespoke cloud and connectivity offering for the expanding Internet of Things (IOT) market. Before joining Sierra Wireless, Emmanuel was Senior Vice President at Gemalto where he was responsible for the development of global Telecom business services. He had previously held various positions with Esso across IT and business development. Emmanuel is a French national and holds an engineering degree from the Ecole Centrale Paris. Outside of the office he is a keen golfer, and enjoys getting plenty of fresh air while sailing.

Talk time: 16:15 – 16:55



Speaker: [Dr Raymond Ng](#) – [Signal AI](#), London, UK.

Title: Applying AI for decision augmentation

Abstract: In Signal AI, we use A.I. to monitor and analyse news and media to augment decision making -- our SaaS platform turns the world's data into knowledge and empower business leaders across a range of industries to make informed and confident decisions. For example, we allow PR professionals to understand the public perception about their brand to develop an effective communication strategy. Another example is monitoring and identifying risks and opportunities for investments. Quantity, efficiency, customisation and user interaction are fundamental requirements. In this talk, we will explain how A.I. techniques can be applied on vast amounts of text data to achieve this goal. We will look into particular examples of entity linking and sentiment analysis, and explain how the appropriate use of data and technology can help achieve robust performance in commercial standards, and discuss some core challenges in applying A.I. on a vast scale in a product development context.

Bio: Raymond is a Senior Data Scientist at Signal AI in London. His research areas cover speech technology and natural language processing. Before joining Signal, Ray was a postdoc researcher in the University of Sheffield. Ray was the PI/co-PI of six research projects (£380k funding) and he co-supervised various MSc/PhD projects on speech, language technology and machine learning topics. Ray is the recipient of Tech Nation UK "Exceptional Talent for Digital Technology" Tier 1 visa endorsement. He was the guest speaker of the Royal Institute Christmas Lectures 2017, the Intelligent Sensing Summer School QMUL 2019 and the London Text Analytics Meetup 2020.

17:00 – 17:40



Speaker: [Dr Kim Trottier](#) – [HEC Montréal](#), Montréal, Canada

Title: Canadian Annual Report Extraction (CARE) project

Abstract: Corporate Annual reports (AR) provide information about the performance and value of firms. Many stakeholders use this information: investors, researchers, regulators, policy makers, accounting professionals, governments, etc. Accessibility has come a long way since the days when AR were mailed to recipients. They are now digitized and available online. However, the PDF format continues to create operational challenges, therefore we propose to create the Canadian Annual Report Extraction (CARE) tool to facilitate analysis, particularly of the narrative. This project builds on the work of El-Haj, Alves, Rayson, Walker, and Young (2020) who created the CFIE tool for extraction of annual reports in the UK. The main differences between CFIE and CARE is the method of locating the sections within the AR and the text metrics provided as output. The CARE will be useful to NLP researchers who wish to explore finance and accounting questions in a Canadian setting, and in other jurisdictions as well after the second phase of our project.

Bio: Kim is an accounting associate professor at HEC Montréal Business School. With a background in capital markets, firm performance, and advanced analytics, she will provide cross-over expertise on financial reporting and artificial intelligence.