# SYMPOSIUM ON AI, DATA AND DIGITALIZATION

SOGNDAL, NORWAY
9–10 MAY 2023

SAIDD 2023
AI. DATA. DIGITALIZATION

# Symposium on AI, Data and Digitalization (SAIDD 2023)

Sogndal – Norway | 9-10 May 2023

| | |
|---|---|
| Published by | Western Norway Research Institute / Vestlandsforsking |
| Address | PO Box 163, NO-6851 Sogndal, Norway |
| Title | Symposium on AI, Data and Digitalization |
| Funding | The Research Council of Norway |
| Year | May 2023 |
| ISBN | 978-82-428-0457-0 |



Edited by

Rajendra Akerkar
Big Data & Emerging Technologies Research Group
Western Norway Research Institute
Sogndal 6851, Norway

*E-mail*: rak@vestforsk.no
*Website*: bigdata.vestforsk.no

# Welcome to SAIDD 2023!

It is with great pleasure, I would like to welcome you to the Symposium on AI, Data and Digitalization (SAIDD 2023) taking place on May 9th-10th in Sogndal, Norway.

The symposium is supported by Western Norway Research Institute. Through knowledge-building, the institute seeks to contribute positively to innovation, policymaking, governance, and value creation. The institute aims to be at the forefront of national and international research, and take part in the development of Western Norway (Vestland), the region in which Sogndal is located. We have a cross-disciplinary research staff with competence in the academic fields of social studies, natural sciences, technology, and the humanities.

The digital revolution – which involves the use of digital Information and Communication Technologies (ICTs), the open data movement, novel enabling technologies, and the increase in the accessibility of big data, together with the advent of artificial intelligence and the Internet of Things – is now transforming society. The SAIDD 2023 symposium is intended to stimulate discussions on how data and artificial intelligence are setting the stage for the digital revolution and contributing to solving societal challenges. SAIDD 2023 provides a common space for dialogue between the academics, policymakers and professionals.

But first and foremost, I wish you an excellent stay in Sogndal, a beautiful place among the spectacular fjords in Western Norway. The Western Norwegian Fjords are listed by UNESCO on the World Heritage List and the Sognefjord region is prized by National Geographic Traveler for the un-spoiled fjord and the glaciers, mountains and cascading waterfalls, that offer excellent opportunities to ice-climbing, glacier walking, hiking and year-round skiing. This has made Sogndal a favorite destination for excitement- and activity-seeking visitors and residents. Enjoy the treasures of Norwegian hospitality, and also make time for some leisure in and around Sognefjord.

I look forward to your participation in the rich scientific program as well as enjoying the social gathering during the SAIDD 2023. I am delighted to see you in Sogndal!

**Anne Karin Hamre**
*Director*
*Western Norway Research Institute*

# Preface for the International Symposium on AI, Data and Digitalization (SAIDD 2023)

Rajendra Akerkar [1]

[1] *Western Norway Research Institute, Sogndal, Norway*

**Abstract**

This volume contains the papers presented at the Symposium on AI, Data and Digitalization (SAIDD 2023). The two-day symposium took place on 9 - 10 May 2023 in Sogndal, Norway. The symposium invited business as well as academic keynotes, panellists and papers exploring how data and artificial intelligence are setting the stage for the digital revolution and contributing to solving societal challenges. The symposium gathered the interested community and discussed the latest approaches for challenges both from the perspectives of academia and industry.

## Introduction

Today, more organizations are using and becoming familiar with the implementation of Big Data and Artificial Intelligence (AI) technologies in their digital transformation processes as it has proven to have great potential for growth. The dependence of digital transformation on AI is of utmost importance as it can help public and private sectors accelerate their digitalization processes. The greatest power of digitalization is to transform the physical world – improving productivity, innovation, and impact.

There are enormous AI opportunities for us across all areas of science, economy and government. Given the significant transformational scope of AI, it has the potential to have profound impacts on the economy and address regional, national and global challenges. This profound impact of AI on the economy has yet to be realised, and will rely on substantial research, development, innovation, and commercial efforts.

In order to realise the full benefits of Big Data and AI technologies, we will need to act to support the growth of our research and innovation capabilities, building on strong foundations that already exist. Furthermore, new challenges and prospects require new theory, methodology, best practice, systems and this should be developed, shared, and discussed by a wide range of stakeholders. Therefore, the objective of SAIDD 2023 was to bring prominent researchers, policy experts and practitioners together in order to foster a deeper understanding of how data and AI are setting the stage for the digital revolution and contributing to solving societal challenges.

## The SAIDD 2023 Symposium

The symposium was organized by our *Transnational Partnership for Excellent Research and Education in Disruptive Technologies for a Resilient Future* (INTPART DTRF) and *Big Data & Emerging Technologies Research Group* at the Western Norway Research Institute in Sogndal, Norway.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: rak@vestforsk.no
ORCHID: 0000-0002-4048-7226

This two-day symposium had a **three-fold perspective**, representing the views and activities of the **Policy, Industry** and **Research/Academia** sectors as follows:

- A **Science-meets-policy**, fostering exchange between policy experts, authorities, researchers and industrial partners dealing with AI, Data and Digitalization; followed by,
- A **Science-dedicated**, uncovering up-to-date contributions from the scientific community on the core issues.

This symposium featured six distinguished keynote addresses. Keynote speakers set the scene with important facts and prospects, and a number of selected participants provided key contributions in the form of position papers and 5-minute 'lightning talks' on defining and addressing the key challenges in this field, and a panel discussion focused on future directions. Position papers described open problems or neglected perspectives on the field, proposing ideas for bringing computational methods into a new application area, or summarizing the focus areas of a group working on AI to address societal challenges. Lightning talks presented bite-sized bits of science that showcase the essence of their research. These were opportunities for young researchers to introduce their work and to share areas of ongoing and emerging focus.

We trust that the scope of the SAIDD 2023 serves the interest of the scientific community, as well as the public and private sectors and the general public.

A total of 42 submissions were received, of which 12 position papers and 18 lightning talks were accepted for presentation at the symposium. To cope with timing constraints and allow for sufficient time for interactions during the conference, 12 papers were given a 15-minute presentation slot and 18 papers were given an opportunity for 5-minutes lightning talks.

The preliminary proceedings volume was published by Vestlandforsking (Western Norway Research Institute). Extended versions of accepted position papers and lightning talk papers will appear in the post-proceedings of SAIDD 2023 that will be published as a volume in Springer's CCIS series.

Each of the 42 submissions received exactly 2 reviews, a remarkable result, for which I am grateful to the members of the SAIDD 2023 Scientific Committee. In a few cases, we conducted an on-line discussion of the reviews and author responses, before reaching a final decision on the selected papers. The result was, I believe, a high-quality program that covers a wide spectrum of topics.

Constructive reviewer feedback is one of the keys to a high-quality symposium year after year, because it encourages each of us, as researchers, to make our work more complete, or to reach a little further. I would like to thank all the members of the SAIDD 2023 Scientific Committee, INTPART DTRF senior members and their sub-reviewers for their timely and insightful reviews which made this symposium possible.

Many colleagues and friends have contributed to SAIDD 2023. First, I thank the authors who submitted their work to SAIDD 2023 and who, by their contributions and participation, make this symposium a quality event. I am also grateful to the SAIDD Organising Committee for its help, to the keynote speakers for their inspiring talks, and, in particular, to the administration of my institute for all their support in organizing this event. Finally, I thank the EasyChair conference system, which made our job so much simpler, and the Research Council of Norway, INTPART DTRF project team and the "Technology and Society" group at Western Norway Research Institute for supporting the symposium.

May 2023 *Rajendra Akerkar*

## Symposium Organisation

### Symposium Convener

Rajendra Akerkar (Western Norway Research Institute, Sogndal, Norway)

### Organizing Committee

Reza Arghandeh (Western Norway University of Applied Sciences, Bergen, Norway)
Simen Eide (Schibsted, Oslo, Norway)
John Krogstie (NTNU, Trondheim, Norway)
Kenth Engø-Monsen (Smart Innovation Norway, Oslo, Norway)
Sachin Gaur (NORA, Oslo, Norway)
Hoang Long Nguyen (Western Norway Research Institute, Sogndal, Norway)
Vimala Nunavath (University of South-Eastern, Norway)
Endre Sundsdal (Norwegian Mapping Authority, Oslo, Norway)
Malin Waage (Western Norway Research Institute, Sogndal, Norway)
Bjørn Christian Weinbach (Western Norway Research Institute, Sogndal, Norway)
Sule Yildirim Yayilgan (NTNU, Gjøvik, Norway)
Svein Ølnes (Western Norway Research Institute, Sogndal, Norway)

## Scientific Committee

Nick Bassiliades (Aristotle University of Thessaloniki, Greece)
David Camacho (Universidad Politécnica de Madrid, Spain)
Jasmien César (Mastercard, Belgium)
Vadim Ermolayev (Ukrainian Catholic University, Ukraine)
Anna Fensel (Wageningen University & Research, The Netherlands)
Song Guo (Hong Kong Polytechnic University, Hong Kong)
Jason J. Jung (Chung-Ang University, Republic of Korea)
Yusheng Ji (National Institute of Informatics, Tokyo, Japan)
Pawan Lingras (Saint Mary's University, Halifax, Canada)
Pierre Maret (Université Jean Monnet, Lyon, France)
Andreas L Opdahl (University of Bergen, Norway)
Hemant Purohit (George Mason University, Farefax, USA)
Charles Robinson (Thales, France)
M. Sasikumar (Centre for Development of Advanced Computing, India)
André Skupin (San Diego State University, USA)
Xuan Song (Southern University of Science and Technology, Shenzhen, China)
Abhishek Srivastava (Indian Institute of Technology, Indore, India)
Fan Zipei (The University of Tokyo, Japan)

## Reviewers

Abhishek Srivastava
Andreas L Opdahl
Ankit Jain
Anna Fensel
David Camacho
Fan Zipei
Fedor Vitiugin
Hemant Purohit
Hoang Long Nguyen
Jason J. Jung
M. Sasikumar
Nick Bassiliades
Pawan Lingras
Pierre Maret
Reza Arghandeh
Rui Zhang
Rupendra Pratap Singh Hada
Song Guo
Sule Yildirim Yayilgan
Vadim Ermolayev
Vimala Nunavath
Xuan Song
Yasas Senarath
Yusheng Ji

## Symposium Supporters

## Table of Contents

**LIGHTNING TALKS**

# Keynotes

# Leveraging Cognitive Knowledge Graphs for Science and Economic Resilience

Sören Auer [1]

[1] *Director TIB, Head of research group Data Science and Digital Libraries, Leibniz Universität Hannover*

Following stations at the universities of Dresden, Ekaterinburg, Leipzig, Pennsylvania, Bonn and the Fraunhofer Society, Prof. Auer was appointed Professor of Data Science and Digital Libraries at Leibniz Universität Hannover and Director of the TIB in 2017. Prof. Auer has made important contributions to semantic technologies, knowledge engineering and information systems. He is the author (resp. co-author) of over 200 peer-reviewed scientific publications. He has received several awards, including an ERC Consolidator Grant from the European Research Council, a SWSA ten-year award, the ESWC 7-year Best Paper Award, and the OpenCourseware Innovation Award. He has led several large collaborative research projects, such as the EU H2020 flagship project BigDataEurope. He is co-founder of high potential research and community projects such as the Wikipedia semantification project DBpedia, the Open Research Knowledge Graph ORKG.org and the innovative technology start-up eccenca.com. Prof. Auer was founding director of the Big Data Value Association, led the semantic data representation in the Industrial/International Data Space, is an expert for industry, European Commission, W3C, the German National Research Data Infrastructure (NFDI) and the European Open Science Cloud (EOSC).

## Keynote

Knowledge graphs are meanwhile established representation forms for heterogeneous data and knowledge. However, knowledge graphs are still relatively simple structures with limited possibilities to represent and efficiently manage more complex structures, provenance, evolution or different levels of granularity. In this talk we will introduce the concept of cognitive knowledge graphs, which use richer atomic base entities - graphlets as constituents. With the Open Research Knowledge Graph (ORKG), we present an example of such a cognitive knowledge graph, which closely intertwines human and machine intelligence with the aim of semantically representing research contributions. As a result, the ORKG provides a digital research infrastructure enabling researchers and other stakeholders to obtain better oversight over the millions of annual research publications and better machine assistance for analysing the wealth of scholarly knowledge. As a second example of leveraging knowledge graphs, we will discuss the integration of data along supply chains to facilitate economic resilience in crisis situations.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: Auer@tib.eu


# Social Cybersecurity: Synthesizing Social Science Theory, AI & Network Science to Support Social Engagement

Kathleen M. Carley [1]

[1] *Professor, Engineering and Public Policy Department, Computer Science Department, and Social and Decision Sciences Department, Carnegie Mellon University. Director of the Center for Computational Analysis of Social and Organizational Systems (CASOS).*

Kathleen M. Carley is a professor in the Engineering and Public Policy Department, Computer Science Department, and Social and Decision Sciences Department at Carnegie Mellon University. She is the director of the Center for Computational Analysis of Social and Organizational Systems (CASOS), a university wide interdisciplinary center that brings together network analysis, computer science and organization science. Carley's research combines cognitive science, social networks, and computer science to address complex social and organizational problems. Her specific research areas are dynamic network analysis, computational social and organization theory, adaptation and evolution, text mining, the impact of telecommunication technologies and policy on communication, information diffusion, and disease contagion and response within and among groups particularly in disaster or crisis situations. Carley and her lab have developed infrastructure tools for analyzing large-scale dynamic networks and various multi-agent simulation systems.

## Keynote

Social-cybersecurity is a new emerging interdisciplinary computational social science aimed at characterizing, understanding, and forecasting cyber-mediated changes in human behavior, and the related social, cultural and political outcomes and developing and testing the technology needed in the cyber-infrastructure that is essential for the survival of civil interaction online. In this talk the nature of social-cybersecurity, the potential societal impact of the work in this field, the nature of related trans-disciplinary work, the criticality of translational research, and the role that AI, network science and social-science theory play in this field are presented. Key issues related to influence activities, disinformation, bots, and hate-speech are discussed. Empirical results from the application of new social-cyber technologies to areas such as the COVID-19 response, re-open America, elections in the Philippines, and other world events.

In presenting these results, the new BEND theoretical framework for assessing influence campaigns is described. The analytic pipeline needed to operationalize this – and the role of theory, AI, and network science in this pipeline – are explained. Key insights regarding the strengths and limitations of AI for social-cybersecurity are presented. The talk ends, with a description of areas where future research is needed to support social engagement and inhibit online harms.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: kathleen.carley@cs.cmu.edu

# AI Regulatory Developments: Deep Dive on the EU AI Act

Jasmien César [1]

[1] *Senior Counsel of Privacy and Data Protection for Artificial Intelligence at Mastercard*

Jasmien César is Senior Counsel of Privacy and Data Protection for Artificial Intelligence at Mastercard with global responsibility for privacy and data protection matters pertaining to Mastercard's AI solutions and uses. Jasmien leads development of Mastercard global strategy and policy for building privacy into the design of AI-powered products, ensuring they are trustworthy and human-centric, and oversees advocacy of Mastercard's privacy position on AI. Before joining Mastercard in 2019, Jasmien spent 5 years as an IT & Data Protection lawyer at the Brussels office of international law firm Bird & Bird. She holds a Degree in Law from KU Leuven (BE) and an LL.M. in EU Competition and Intellectual Property Law from the University of Liège (BE).

## Keynote

In addition to the myriad of existing legal frameworks (e.g., privacy and data protection, consumer protection, non-discrimination) already addressing Artificial Intelligence (AI) in a technology-agnostic way, we are witnessing a broad range of AI-specific regulatory initiatives around the globe. Many countries or regions are introducing principles-based AI laws and governance frameworks, focusing on principles such as AI risk or impact assessments, human oversight, data quality and accuracy, bias monitoring and mitigation, transparency and explainability, and cybersecurity and robustness. Everyone involved in the AI value chain - from developers to deployers, from research to industry, from public to private sector, from start-ups to multinationals - will need to navigate this increasingly complex regulatory landscape if they want to continue leveraging AI in a legally compliant way.

This keynote will provide an overview of the AI-specific regulatory landscape to date, delve into one of the most comprehensive legislative proposals addressing AI, the EU AI Act, and arm you with practical steps that you can take to prepare for AI regulation.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: jasmien.cesar@mastercard.com

# Responsible Artificial Intelligence: An Inclusive Road Ahead

Virginia Dignum [1]

[1] *Wallenberg Chair, Professor Responsible Artificial Intelligence, Program Director WASP-HS, Department of Computing Science, Umeå University*

Virginia Dignum is a professor in Responsible Artificial Intelligence and the scientific Director of WASP-HS (Humanities and Society). She is also associated with the Faculty Technology Policy and Management at the Delft University of Technology. Her research focuses on value-sensitive design of intelligent systems and multi-agent organisations, in particular on the ethical and societal impact of AI. Her work ranges from the engineering of practical applications and simulations to the development of formal theories that integrate agency and organization and includes a strong methodological design component. She is a Fellow of the European Artificial Intelligence Association (EURAI) and has been vice president of the BNVKI (Benelux AI Association). She was elected to the Swedish Royal Academy of Engineering Sciences (IVA) in 2020. She was member of the High-Level Expert Group on Artificial Intelligence from the European Commission, was member of the Executive Committee of the IEEE Initiative on Ethics of Autonomous Systems and sits in the scientific boards of the Delft Design for Values Institute, the AI4People - European Global Forum on AI, and the Responsible Robotics Foundation.

## Keynote

Artificial Intelligence (AI) has huge potential to bring accuracy, efficiency, cost savings and speed to a whole range of human activities and to provide entirely new insights into behaviour and cognition. However, the way AI is developed and deployed for a great part determines how AI will impact our lives and societies. For instance, automated classification systems can deliver prejudiced results and therefore raise questions about privacy and bias; and, the autonomy of intelligent systems, such as, e.g. self-driving vehicles, raises concerns about safety and responsibility. AI's impact concerns not only the research and development directions for AI, but also how these systems are introduced into society and used in everyday situations. There is a large debate concerning how the use of AI will influence labour, well-being, social interactions, health care, income distribution and other social areas. Dealing with these issues requires that ethical, legal, societal and economic implications are taken into account.

In this presentation, I will discuss how a responsible approach to the development and use of AI can be achieved, and how current approaches to ensure the ethical alignment of decisions made or supported by AI systems can benefit from the social perspective embedded in feminist and non-Western philosophies, in particular the Ubuntu philosophy.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: virginia@cs.umu.se


# eXplainable AI (XAI) a Basic Break Towards Synergistic Human-Machine Interaction and Collaboration

Fosca Giannotti [1]

[1] *Professor of Computer Science at Scuola Normale Superiore, Pisa & Associate at the Information Science and Technology Institute "A. Faedo" of CNR., Pisa, Italy*

Fosca Giannotti is Full Professor at Scuola Normale Superiore, Pisa, Italy. Fosca Giannotti is a pioneering scientist in mobility data mining, social network analysis and privacy-preserving data mining. Fosca leads the Pisa KDD Lab - Knowledge Discovery and Data Mining Laboratory, a joint research initiative of the University of Pisa and ISTI-CNR, founded in 1994 as one of the earliest research lab on data mining. Fosca's research focus is on social mining from big data: smart cities, human dynamics, social and economic networks, ethics and trust, diffusion of innovations. She is author of more than 300 papers. She has coordinated tens of European projects and industrial collaborations. Fosca is the former coordinator of SoBigData, the European research infrastructure on Big Data Analytics and Social Mining, an ecosystem of ten cutting edge European research centres providing an open platform for interdisciplinary data science and data-driven innovation. Recently she became the recipient of a prestigious ERC Advanced Grant entitled XAI – Science and technology for the explanation of AI decision making.

## Keynote

Black box AI systems for automated decision making, often based on machine learning over (big) data, map a user's features into a class or a score without exposing the reasons why. This is problematic not only for the lack of transparency, but also for possible biases inherited by the algorithms from human prejudices and collection artifacts hidden in the training data, which may lead to unfair or wrong decisions. The future of AI lies in enabling people to collaborate with machines to solve complex problems. Like any efficient collaboration, this requires good communication, trust, clarity and understanding. Explainable AI addresses such challenges and for years different AI communities have studied such topic, leading to different definitions, evaluation protocols, motivations, and results. This lecture provides a reasoned introduction to the work of Explainable AI (XAI) to date, and surveys the literature. A special focus will be on mechanisms to improve joint performance in high-stake decision-making such as methods aimed at engaging users with factual and counterfactual or other high-level explanations encoding domain knowledge and user background, methods focusing on conversational explainable AI and methods aimed at understanding the impact of explanation on expert users' information-seeking strategies, mental model updating, and trust calibration. Finally, we will motivate the need of paradigms that can promote collaboration and seamless interaction maintaining the human responsibility of the choice through a progressive disclosure to prevent cognitive overload.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: fosca.giannotti@isti.cnr.it


# The State of AI in the Nordic Region

Trym Holter [1]

[1] *Business Development Executive, Silo AI & Former Director, Norwegian Open AI Lab*

Trym Holter is a Business Development Executive at Silo AI. He holds a MSc (siv. ing.) from the Norwegian Institute of Technology (NTH) and a PhD (dr. ing.) from the Norwegian University of Science and Technology (NTNU) in 1998. For more than 25 years, Trym has been working in research and product development in areas such as machine learning, signal processing, acoustics, and electronics. Before joining Silo AI, he was the director of the Norwegian Open AI Lab, and prior to this he held the position as research director in SINTEF Digital. In addition to his background from research, he has held various positions in multinational companies (Honeywell and Motorola) as well as in a small start-up (Nacre AS). He is currently on the board of directors for start-up companies Minuendo and Nomono, as well as for OsloMet - Oslo Metropolitan University.

## Keynote

The values shared by the Nordic countries are often claimed to help make the region one of the most innovative and competitive in the world. While the Nordics may be forerunners in AI readiness, ethics, and trustworthiness, the region is still in the early stages of putting AI technologies to use in products and services. Clearer strategies and stronger collaborations are often being called for. The Nordic governments vary in their investments and strategies to accelerate the use of AI. Our universities rank well globally on AI research metrics, but we are falling behind on deploying solutions that capture value for businesses and users. Many also claim that general knowledge and education are still lacking.

In this talk we will discuss the current state of AI in the Nordic region. The focus will primarily be on the deployment of AI technologies in the private and public sectors, together with barriers and opportunities for continued development of AI based products and services.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: trym.holter@silo.ai

# Position Papers

# AI-based Strategic Foresight for Environment Protection

Jože M. Rožanec[1,2,*,†], Radu Prodan[3,†], Dumitru Roman[4], Gregor Leban[5] and Marko Grobelnik[1]

[1]Jožef Stefan Institute, Jamova cesta 39, 1000 Ljubljana, Slovenija
[2]Jožef Stefan International Postgraduate School, Jamova cesta 39, 1000 Ljubljana, Slovenija
[3]Alpen-Adria-Universität Klagenfurt, Universitätsstraße 65/67, 9020 Klagenfurt am Wörthersee, Austrija
[4]SINTEF, Strindvegen 4, NO-7465 Trondheim, Norway
[5]Event Registry d.o.o., Teslova ulica 30, Ljubljana, 1000 Ljubljana, Slovenia

**Abstract**

The Environment Protection Foresighter aims to deliver low-latency strategic foresight by analyzing massive amounts of data and leveraging Artificial Intelligence to extract knowledge, enrich existing data, and provide insights while involving experts in a feedback loop to learn from them and increase the foresight quality over time. In addition, it aims to identify how events relate to environment-related Sustainable Development Goals and perform horizon scanning for environment-related issues. It will be developed within the Graph-Massivizer project, funded by the Horizon Europe research and innovation program, leveraging the platform for high-performance and sustainable graph processing of data.

**Keywords**

Strategic Foresight, Environment Protection, Policy-making, Sustainable Development Goals, Artificial Intelligence, Graph Processing

## 1. Introduction

Strategic Foresight aims to provide a structured approach toward gathering information regarding plausible futures to prepare for change adequately. It provides expert insights regarding trends and emerging issues, which can be considered for strategic planning and policy-making. Strategic Foresight does not aim to predict the future but rather explore possible scenarios of interest and understand how specific actions could influence the future informing policy-making and decision-making. Foresight is considered increasingly important in TUNA (Turbulence,

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
*Corresponding author.
†These authors contributed equally.
EMAIL: joze.rozanec@ijs.si (J. M. Rožanec); radu.prodan@itec.aau.at (R. Prodan); Dumitru.Roman@sintef.no (D. Roman); gregor@eventregistry.org (G. Leban); marko.grobelnik@ijs.si (M. Grobelnik)
ORCHID: 0000-0002-3665-639X (J. M. Rožanec); 0000-0002-8247-5426 (R. Prodan); 0000-0001-6397-3705 (D. Roman); 0000-0003-1193-5955 (G. Leban); 0000-0001-7373-5591 (M. Grobelnik)


Unpredictability, Uncertainty, Novelty, and Ambiguity) conditions observed in a VUCA (Volatile, Uncertain, Complex, and Ambiguous) world order [1, 2].

There is a rising awareness of the need for strategic foresight for better policy-making worldwide. In 2018, the US Government Accountability Office created the Center for Strategic Foresight to provide the US Congress with reliable, fact-based information on emerging issues relevant to policymakers. Along the same line, the European Commission aims to cooperate with European institutions through an EU-wide Foresight Network to gain insights and develop future-oriented policies. Furthermore, it started publishing foresight reports to integrate strategic foresight into public administration [3]. However, in contrast to the EU and the US, little is known about strategic foresight, and its use in China [4].

Strategic Foresight has been recognized as a valuable tool to enable a more sustainable future [5]. In this line, the strategic foresight report published by the European Commission in 2022 outlined ten key action areas where synergies must be maximized toward a green and digital transition to be achieved by 2050.

A research study conducted in the US and published in 2018 has identified that the most frequently used methods in strategic foresight are horizon scanning and scenario planning [6]. Horizon scanning consists of systematically gathering and analyzing information on trends and developments relevant to the organization. Their relevance can be due to emerging threats, opportunities, or responsibilities. Four key stages can be identified: (i) discovery, (ii) evaluation, (iii) understanding, and (iv) decision-making. Conversely, scenario planning attempts to create descriptions of plausible future scenarios by harmonizing collected information into consistent descriptions regardless of the divergence observed across multiple trends. While Artificial Intelligence has not been widely adopted in strategic foresight, its value has been recognized [7]. Hybrid approaches have been developed, leveraging Artificial Intelligence to automate information scanning and data analysis and enabling the human professional to focus on more complex tasks [8, 9]. In particular, Geurts et al. [10] describe an approach that enables them to construct and refine knowledge graphs and use text mining techniques to perform topic modeling to analyze trends and identify weak signals. Relations to those weak signals are searched in the knowledge graph, and patterns are identified to identify plausible futures and assess their impact. Finally, the experts can determine appropriate strategies and policies according to the identified future scenarios.

We aim to develop a solution to ingest relevant data and semantic abstractions. We use artificial intelligence to track and forecast relevant trends. A graph is created and mined to find patterns and determine whether patterns observed in the past can emerge in the present and assess their potential impact. We aim to do so by leveraging the GraphMassivizer architecture [11]. In Section 2 we provide a detailed description of the use case, enabling technologies, objectives, and envisioned implementation.

## 2. Environment Protection Foresighter

### 2.1. Use Case

For successful environmental protection, there is a need to provide business and geopolitical foresight related to sustainable development goals (SDGs) regarding the environment (e.g.,

climate action, responsible consumption and production patterns, or clean and affordable energy, among others). While massive amounts of data can be collected, global media news provides a privileged view of the convergence of economy, politics, and science [12]. Semantic abstractions can be used to build a knowledge graph and better understand the patterns and dynamics that emerge and repeat over time. Furthermore, Artificial Intelligence can process the incoming data and the knowledge graph to find and extract causal templates. These are built by considering particular policy-making perspectives and provide insights into future scenarios by predicting possible events and consequences.

## 2.2. Objectives

Our research aims to leverage Artificial Intelligence to automate strategic foresight and speed up foresight creation between seven and thirty times, reporting with a latency below a day since the ingested event. In particular, we aim to address two specific foresight areas: (a) identify how events relate to environmental SDGs, and (b) identify emerging issues and opportunities through horizon scanning. We expect to develop an end-to-end system capable of analyzing incoming data, extracting relevant signals, identifying patterns, and alerting users on issues that are meaningful to their profile. Furthermore, we envision a collaboration between experts and such an Artificial Intelligence system to collect feedback and learn (a) what insights were correct and (b) what information displayed was useful or could be enhanced. The strategic foresight solution will focus on four environmental SDGs described by the United Nations: climate action, responsible production and consumption patterns, clean water and sanitation, and clean and affordable energy. Such a solution will be relevant to policymakers, governmental organizations, NGOs, and business strategists.

## 2.3. Enabling technologies

Many media event retrieval systems perform data gathering of media news events that provide access to crawled data as a service [13, 14]. The reported events can be enriched with semantic abstractions [15] and integrated into a knowledge graph, leveraging knowledge from external knowledge graphs too (e.g., Wikidata [16, 17, 18]). Among the abstractions, we also consider different conceptions of time [19], which lead to different representations based on time understanding (e.g., density, uncertainty, periodicity, and absolutist or relativist partitioning of time; and the clocks considered (e.g., global time, local time, logical clocks), among others). Further enrichment is possible by using machine learning classifiers to predict whether the events relate to a specific characteristic of interest [20, 21]. Regression models will forecast future values of relevant economic, political, and other indicators. Finally, graph processing techniques can unveil underlying causal templates and predict possible events and expected outcomes [22, 23].

## 2.4. Envisioned implementation

We aim to develop a tool based on the GraphMassivizer architecture [11]. We envision building a knowledge graph based on existing knowledge graphs (e.g., Wikidata) by aligning them with entity alignment techniques [24, 25]. When doing so, we understand the complexity of the task and the multiple open challenges regarding graph heterogeneity and conceptual

hierarchies that still exist. Such a graph will provide a basic understanding of the underlying world. Data regarding worldwide events will be continuously ingested from multiple sources. In particular, we have considered EventRegistry [14] and the GDELT project [26]. The events will be semantically enriched upon ingestion using a wikifier [27] and linked to the abovementioned graph. Furthermore, their content will undergo a text-mining process to extract relevant entities, their relationships, existing sentiments, and relevant insights (e.g., relationship to SDG goals). Graph processing techniques will be used to mine graph patterns and determine how patterns observed in the past are building in the present and may complete in the future. We consider this a critical capability that can be achieved with a rich semantic model, including a rich representation of time [19].

Two specific foresight areas will be addressed with the Environment Protection Foresighter. First, to identify how events relate to specific environmental SDGs, we will leverage a tool that extracts sentences for each event, tagging the subjects and causal relationships to a particular event of interest. Manual labeling is required to confirm whether the information extracted for each fragment is correct. Then few-shot learning models will be trained to learn how such fragments match specific SDGs and use them to classify existing and future data. Conversely, Horizon scanning will require topic modeling and monitoring of those topics and semantic abstractions to understand whether weak signals point to emerging issues of interest that can gain relevance over time. Furthermore, ingested data regarding economic, political, and other relevant indicators could be enriched with forecasts to understand likely future dynamics ahead of time. Finally, graph heuristics and graph neural networks could be used to contrast expected patterns against observed ones to identify changes and new, emerging trends that deviate from historical ones.

In Figure 1, we show an incomplete representation of events and abstractions regarding the Tunisian revolution between the end of 2010 and the beginning of 2011. The example does not consider the time variable and its possible ontological representations. In particular, it was observed that the uprising caused a broader instability in the MENA (Middle East North Africa) region, known as the Arab spring, which impacted energy prices worldwide [28]. Furthermore, it has been shown that oil price fluctuations affect the consumption of renewable energy sources [29, 30]. Identifying how latent conditions may trigger regional instability and affect renewable energy demand provides valuable insight that can be translated into concrete decision-making and policies before such events unfold.

## 3. Summary and Outlook

We consider technological research for the strategic foresight domain to be scarce, and more such use case adoption is needed. The development of the Environment Protection Foresighter will address current voids on how to ingest, enrich and analyze massive amounts of data for foresight purposes and create meaningful foresight insights. To achieve such a goal, we will exploit the GraphMassivizer software architecture. In addition, artificial Intelligence techniques will enable the automation of many tasks and analyses while involving foresight experts for feedback and outcomes curation.

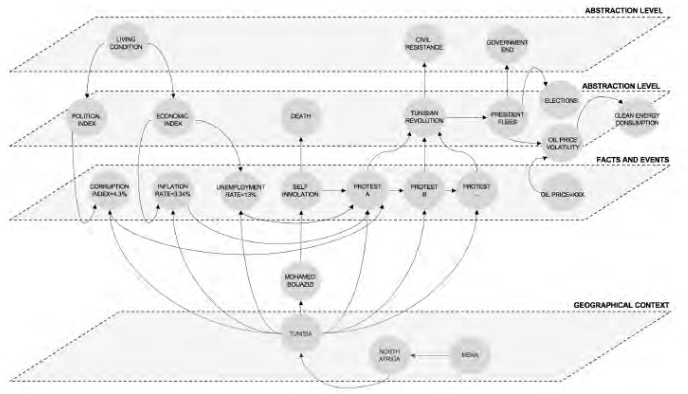



**Figure 1:** The figure represents events and abstractions related to the Arab spring. In particular, we consider some events related to the Tunisian revolution. By querying the highest abstraction levels in the graph, we could understand the catalysts of the revolution (poor living conditions and someone who self-immolated) along with the revolution outcome (the president fled, elections were held, and the crude oil prices increased worldwide, affecting the demand for renewable energy).

## Acknowledgments

The Graph-Massivizer project receives funding from the Horizon Europe research and innovation program of the European Union under grant agreement 101093202. The project started on January 1st 2023 and lasts three years.

## References

[1] A. Wilkinson, Strategic foresight primer, European Political Strategy Centre (2017).

[2] J. R. L. Kaivo-oja, I. T. Lauraeus, The vuca approach as a solution concept to corporate foresight challenges and global technological disruption, foresight (2018).

[3] A. B. Rosa, N. Gudowsky, P. Repo, Sensemaking and lens-shaping: Identifying citizen contributions to foresight through comparative topic modelling, Futures 129 (2021) 102733.

[4] P. Charon, In china (2021).

[5] T. Volkova, B. Dominiece-Diasa, Strategic foresight as an enabler of a sustainable future, International Multidisciplinary Scientific GeoConference: SGEM 19 (2019) 705–712.

[6] J. M. Greenblott, T. O'Farrell, R. Olson, B. Burchard, Strategic foresight in the federal government: A survey of methods, resources, and institutional arrangements, World futures review 11 (2019) 245–266.

[7] N. Reez, Foresight-based leadership. decision-making in a growing ai environment, in: International Security Management: New Solutions to Complexity, Springer, 2020, pp. 323–341.

[8] N. H. Parrish, A. L. Buczak, J. T. Zook, J. P. Howard, B. J. Ellison, B. D. Baugher, Crystal cube: Multidisciplinary approach to disruptive events prediction, in: Advances in Human Factors, Business Management and Society: Proceedings of the AHFE 2018 International Conference on Human Factors, Business Management and Society, July 21-25, 2018, Loews Sapphire Falls Resort at Universal Studios, Orlando, Florida, USA 9, Springer, 2019, pp. 571–581.

[9] P. Brandtner, M. Mates, Artificial intelligence in strategic foresight–current practices and future application potentials: Current practices and future application potentials, in: The 2021 12th International Conference on E-business, Management and Economics, 2021, pp. 75–81.

[10] A. Geurts, R. Gutknecht, P. Warnke, A. Goetheer, E. Schirrmeister, B. Bakker, S. Meissner, New perspectives for data-supported foresight: The hybrid ai-expert approach, Futures & Foresight Science 4 (2022) e99.

[11] R. Prodan, D. Kimovski, A. Bartolini, M. Cochez, A. Iosup, E. Kharlamov, J. Rožanec, L. Vasiliu, A. L. Vărbănescu, Towards extreme and sustainable graph processing for urgent societal challenges in europe, in: 2022 IEEE Cloud Summit, IEEE, 2022, pp. 23–30.

[12] V. Kayser, K. Blind, Extending the knowledge base of foresight: The contribution of text mining, Technological Forecasting and Social Change 116 (2017) 208–215.

[13] K. Leetaru, P. A. Schrodt, Gdelt: Global data on events, location, and tone, 1979–2012, in: ISA annual convention, volume 2, Citeseer, 2013, pp. 1–49.

[14] G. Leban, B. Fortuna, J. Brank, M. Grobelnik, Event registry: learning about world events from news, in: Proceedings of the 23rd International Conference on World Wide Web, 2014, pp. 107–110.

[15] M. Grobelnik, B. Fortuna, G. Leban, J. Rupnik, A. Muhič, A. Košmerlj, Xlike-cross-lingual knowledge extraction (fp7-ict-2011-7) (????).

[16] S. Malyshev, M. Krötzsch, L. González, J. Gonsior, A. Bielefeldt, Getting the most out of wikidata: semantic technology usage in wikipedia's knowledge graph, in: International Semantic Web Conference, Springer, 2018, pp. 376–394.

[17] C. Rudnik, T. Ehrhart, O. Ferret, D. Teyssou, R. Troncy, X. Tannier, Searching news articles using an event knowledge graph leveraged by wikidata, in: Companion proceedings of the 2019 world wide web conference, 2019, pp. 1232–1239.

[18] L. Wang, Y. Li, O. Aslan, O. Vinyals, Wikigraphs: A wikipedia text-knowledge graph paired dataset, arXiv preprint arXiv:2107.09556 (2021).

[19] V. Ermolayev, S. Batsakis, N. Keberle, O. Tatarintseva, G. Antoniou, Ontologies of time: Review and trends., International Journal of Computer Science & Applications 11 (2014).

[20] J. Galsurkar, M. Singh, L. Wu, A. Vempaty, M. Sushkov, D. Iyer, S. Kapto, K. Varshney, Assessing national development plans for alignment with sustainable development goals via semantic search, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 32, 2018.

[21] T. Matsui, K. Suzuki, K. Ando, Y. Kitai, C. Haga, N. Masuhara, S. Kawakubo, A natural language processing model for supporting sustainable development goals: translating semantics, visualizing nexus, and connecting stakeholders, Sustainability Science 17 (2022) 969–985.

[22] X. Ding, Z. Li, T. Liu, K. Liao, Elg: an event logic graph, arXiv preprint arXiv:1907.08015 (2019).

[23] J. Gao, X. Luo, H. Wang, Chinese causal event extraction using causality-associated graph neural network, Concurrency and Computation: Practice and Experience 34 (2022) e6572.

[24] K. Zeng, C. Li, L. Hou, J. Li, L. Feng, A comprehensive survey of entity alignment for knowledge graphs, AI Open 2 (2021) 1–13.

[25] Z. Sun, Q. Zhang, W. Hu, C. Wang, M. Chen, F. Akrami, C. Li, A benchmarking study of embedding-based entity alignment for knowledge graphs, arXiv preprint arXiv:2003.07743 (2020).

[26] E. Boudemagh, I. Moise, News media coverage of refugees in 2016: a gdelt case study, in: Proceedings of the International AAAI Conference on Web and Social Media, volume 11, 2017, pp. 743–750.

[27] J. Brank, G. Leban, M. Grobelnik, Annotating documents with relevant wikipedia concepts, Proceedings of SiKDD 472 (2017).

[28] B. Fattouh, H. Darbouche, The implications of the Arab uprisings for oil and gas markets, Oxford Institute for Energy Studies, 2011.

[29] N. Apergis, J. E. Payne, Renewable energy, output, carbon dioxide emissions, and oil prices: evidence from south america, Energy Sources, Part B: Economics, Planning, and Policy 10 (2015) 281–287.

[30] V. Troster, M. Shahbaz, G. S. Uddin, Renewable energy, oil prices, and economic activity: A granger-causality in quantiles analysis, Energy Economics 70 (2018) 440–452.

# Unpacking the tensions in the legal regulation of online hate-speech within the Norwegian context

Carol Azungi Dralega [1], Torborg Igland [2] and Gilda Seddighi [3]

[1] NLA University College, Bergtorasvei 120 Kristiansand, 4633, Norway
[2] NLA University College, Bergtorasvei 120 Kristiansand, 4633, Norway
[3] NORCE Research Institute, Mekjarvikveien 12, Randaberg 4070, Norway

**Abstract**

Recent research shows that online hate-speech is on the rise in western societies, including in Norway. This submission builds on a broader 4-year research initiative on: "Violence-inducing Behavior Prevention in Social-Cyber Space of Local Communities" funded by the Norwegian Research Council aimed develop radical technological solutions to support authorities in the fight against online-hate speech as well as generate new and deeper understandings of the complexities surrounding the vice on online hate. Our research contributes in two ways to the conference theme on how data and artificial intelligence are setting the stage for digital revolution and contributing to solving societal challenges. Firstly, our presentation contributes to strengthening our understanding of the positive strides, the grey areas, and the challenges from a legislative perspective. Here, we explore how the police identify, interpret, and apply the legislation on online hate speech, through data and artificial intelligence and the gaps thereof.

Secondly, this qualitative contribution feeds into the overall projects goal to develop a real time AI tool, to help authorities such as the police in their efforts to combat online hate speech.

Using human rights, freedom of expression and digitization as conceptual frames, the paper posits a social-economic, ethical, legislative and technological analysis of the legislative initiative on hate-speech (paragraph 185) including both positives as well as the tensions and consequences from a legislation context.

**Keywords**

Online hate-speech, legal grey areas, Racism, Police, Freedom of expression

## 1. Introduction and Background

Hate-speech, on or offline, poses a serious threat to human rights and democracy. Given the steady growth over the last years, of online hate-crimes (which is the focus of this paper), in Western societies, combating it has become crucial for governments, organizations, local communities and stakeholders (Zicarddi, 2020; Kalsnes, and Ihlebæk 2021, NOU 2022:9[i2]). The Council of Europe (SM/REC (2022)16), cognizant of the broader international and European human rights standards, recommends that all member states "protect human rights and fundamental freedoms in the digital environment, including by co-operating with internet intermediaries" (recommendation 3).

As in several countries in Europe, in Norway, the constitution and several legislations prohibit hate-based discrimination and abuse of individual and group rights. Specifically, the legal framework on hate speech appears in different chapters and paragraphs in the Norwegian legislation. One of the most central paragraphs is the so-called *paragraph on hate-speech,* §185. This paragraph is placed in Chapter

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: carol.dralega@nla.no (A. 1); torborg.igland@nla.no (A. 2); gsed@norceresearch.no (A. 3)

20 in the Norwegian legislation. Chapter 20 highlights the protection of society, public peace, order and security[ii] . There is an ongoing debate on the §185, whether the part of the law that protects (or not) individuals, should be moved to Chapter 24, which highlights the protection of personal freedom and peace. Both chapters are part of the criminal law in the legislation. Nevertheless, the §185 deals with racism and hateful utterances and is subject of this position paper on online hate-speech.

The ongoing debate and scholarship on online hate-speech in Norway gained momentum in 2020, after a legal precedence was set with its first guilty verdict on racial online hate crimes (Nguyen 2020). In August 2022, a special committee on freedom of expression released a new compendium that further defines and gives guidance on freedom of expression including online expressions (NOU 2022: 9). Despite the Norwegian Constitutional and legislative recourse, there are still lingering tensions especially when freedom of speech comes in collision with other rights and interests (Kierulf, 2021).

In the analysis we aim to unravel the grey areas within the legislation and punitive processes and how these present blurred lines that limit the effectiveness of legal recourse. We look at the role and importance of data and AI application in these processes. Of interest in this discussion is also on the degree to which such grey areas offer fertile grounds for a sense of entitlement in the public in general and with persons and groups that perpetuate online hate crimes as well as the consequences these have for those exposed to racism. We are not doing a legal analysis per se, but rather an analysis of how common perceptions of the concept racism corresponds to the experiences of racism, and how this again corresponds to the enforcement of the Law, particularly §185. Two cases of Sumaya Jirde Ali, with different outcomes, will serve as good examples in this matter and will be analyzed.

Our submission is aimed at sharing tentative research findings that are part of the larger project aimed at preventing violent behavior in social cyber spaces of local communities[iii]. The findings here are extracted from a study on several stakeholder approaches to the task of combating online hate-speech through data and/or AI. This presentation will focus on the work of the Police and relevant units. Using human rights as its overall frame, the main objective is to explore the legal, social, technological, and ethical parameters that defined and/or constrain the work of the Police and selected hate crime units. Specifically, the objectives are to:

- Explore how the Police and the different units identify and define parameters and frames of their work on online hate-speech, through data, AI and other means.
- Gain an understanding of the processes, actions and outcomes in fighting online hate-speech.
- Highlight tensions (or grey areas), especially in the legal framework and the implication on local communities, groups and individuals victimized by online hate speech.

In all this, we want to share the positive strides made by the Police as well as highlight the challenges they face in the fight against a novel, complex, multi-faceted and ever-changing societal problem. It is a goal of this research to contribute with recommendations, guidelines/solutions. This data will also feed into the development of real-time technical tools to support the police work.

## 2. Conceptual framework

Principles on fundamental Human rights such as: promoting human values and dignity; safety and security; fairness and non-discrimination; living in peace; accountability and freedom of expression and digitization – form the conceptual frames for the analysis (Kiritchenko, Nejadgholi and Fraser, 2021). We particularly revisit the debate on freedom of expression as a double-edged sword in the discussion on the its limits with regards to online hate speech. Our discussions unpack the tensions (and shortcoming especially within the existing legal frameworks) in a bid to fill gaps in the fight against online hate speech (Bennedek and Ketteman 20203 ; Calson, 2021; Kierulf 2021).

## 3. Methodology

Our discussion will share insights we obtained through qualitative methodologies including a workshop with stakeholders, interviews and focus group discussions with individuals within the Norwegian Police force between October 2022 and February 2023? We will triangulate the data with

literature on hate-speech and court documents from Norwegian Supreme Court on online hate speech, to find practice and identify grey zones and point out where the tensions might be.

## 4. Acknowledgement

This 4-year project was funded by the Norwegian Research Council, to whom we are grateful. We are also grateful to the informants who have been generous with information and insights. Appreciation also goes to the project manager Professor Rajendra Akerkar from Western Norway Research Institute as well as our colleagues from NTNU in Norway and George Mason University, USA for the critical discussions that have enriched this work.

## 5. References

Assimakopoulos, S., Baider, F. H., & Millar, S. Online hate speech in the European Union: a discourse-analytic perspective. Springer Nature. (2017)

Baider, F. Pragmatics lost? Overview, synthesis and proposition in defining online hate speech. Pragmatics and Society, (2020) 11(2), 196-218.

Banks, J. Regulating hate speech online. International Review of Law, Computers & Technology, (2010). 24(3), 233-239.

Brown, A. What is so special about online (as compared to offline) hate speech? Ethnicities, (2018) 18(3), 297-326.

Carlson, C. R. Hate speech. MIT Press. (2021)

Castaño-Pulgarín, S. A., Suárez-Betancur, N., Vega, L. M. T., & López, H. M. H. Internet, social media and online hate speech. Systematic review. Aggression and Violent Behavior, (2021). 58, 101608.

Corazza, M., Menini, S., Cabrio, E., Tonelli, S., & Villata, S. A multilingual evaluation for online hate speech detection. ACM Transactions on Internet Technology (TOIT), (2020) 20(2),

George, C. Hate-speech law and policy. The international encyclopedia of digital communication and society, (2015). 1-10.

Kalsnes, B., & Ihlebæk, K. A. Hiding hate speech: Political moderation on Facebook. Media, Culture & Society, (2021) 43(2), 326-342.

Kierulf, A. Hva er ytringsfrihet? Universitetsforlaget (2021).

Lovdata, Lov om straff, Lov om straff (straffeloven). Annen del. De straffbare handlingene. Kapittel 20. Vern av den offentlige ro, orden og sikkerhet, §185.

Mathew, B., Saha, P., Tharad, H., Rajgaria, S., Singhania, P., Maity, S. K., ... & Mukherjee, A. Thou shalt not hate: Countering online hate speech. In Proceedings of the international AAAI conference on web and social media (2019) (Vol. 13, pp. 369-380).1-22.

Nadim, M., & Fladmoe, A. Silencing women? Gender and online harassment. Social Science Computer Review, (2021) 39(2), 245-258.

Nadim, M. & Thorbjørnsrud, K. A study of participants in aggressive online debates.

«The struggle in the blurry areas of the debate: a study of participants in flaming and aggressive online debates». (2022)

Nguyên Duy, I. The Limits to Free Speech on Social Media: On Two Recent Decisions of the Supreme Court of Norway. Nordic Journal of Human Rights, (2020) 38(3), 237-245.

Ortiz, Stephanie M. Racists without racism? From colour-blind to entitlement racism online. Ethnic and Racial Studies. Volume 44, (2021) - Issue 14

Qian, J., Bethke, A., Liu, Y., Belding, E., & Wang, W. Y. A benchmark dataset for learning to intervene in online hate speech. (2019). arXiv:1909.04251.

Ullmann, S., & Tomalin, M. Quarantining online hate speech: technical and ethical perspectives. Ethics and Information Technology, (2020) 22, 69-80.

Ziccardi, G. Online political hate speech in Europe: The rise of new extremisms. Edward Elgar Publishing. (2020).

[i] NOU 2022: 9 - regjeringen.no

[ii] Høringsuttalelse til Ytringsfrihetskommisjonen fra Dommerforeningens menneskerettighetsutvalg.

[iii] https://www.vestforsk.no/en/project/violence-inducing-behaviour-prevention-social-cyber-space-local-communities

# Artificial Intelligence in the public sector in Norway: a hop-on-hop-off journey

Hilde G. Corneliussen [1], Gilda Seddighi [2], Aisha Iqbal [3] and Rudolf Andersen [3]

[1] *Western Norway Research Institute, Røyrgata 4, Sogndal, Norway*
[2] *NORCE Norwegian Research Centre AS, Nygårdsgaten 112, Bergen, Norway*
[3] *Rambøll Management Consulting, Harbitzalléen 5, 0275 Oslo, Norway*

**Abstract**

This paper discusses the use of artificial intelligence (AI) in the Norwegian public sector, particularly projects involving personal data. The study includes a survey with 200 public sector organizations and interviews with 19 of these using person data. The findings suggest that AI development in the public sector is still immature with few projects involving person data in production. Political pressure to use AI in the sector is significant, creating expectations about its use. Limited knowledge and focus on directing AI development among management has made individuals and units with the resources and interest in experimenting with AI an important driving forces. The study found that the journey from idea to production of AI in the public sector presents many challenges, which often leads to projects being temporarily halted or terminated. Furthermore, a wide definition of AI creates confusion about the use of different AI techniques and their role in decision-making. While AI can contribute to the streamlining and improvement of public services, it also has risks and challenges, including the risk of producing incorrect or discriminatory results. This, however, was not a significant concern or topic in the public sector AI projects. Instead, other concepts such as ethics, fairness, and transparency took precedence.

**Keywords**

Artificial intelligence, public sector, discrimination

## 1. Introduction

Artificial intelligence (AI) has made significant progress in recent years. It has been identified as one of the most important technologies of the 21st century [1], expected to have a major impact on solving small and large social challenges with an effect for private and public sectors and for individuals [2]. AI is considered an important tool for improving and making the Norwegian public sector more efficient. Adopting AI, however, also entails challenges and risks, including the risk of incorrect, unfair, or discriminatory results. This has concerned researchers [3-9] and national and international initiatives are on their way with suggestions for minimizing such risk [10-14]. This paper reports from a study of AI projects in the Norwegian public sector, mapping the use and plans for using AI, and the challenges experienced. A critical risk of AI producing discriminatory results for social groups and individuals is related to the use of personal data in an AI system. While this is a risk in private as well as public sectors, AI involves a particular set of challenges for the public sector. Public sector organizations are subject to strict laws and regulations for handling personal data, they are obligated to ensure equality for a diverse population in public services, and the "customers", the citizens, cannot choose another service provider. Thus, potential risks of using AI are critical to manage for the public sector, in order

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: hgc@vestforsk.no (A. 1); gsed@norceresearch.no (A. 2)
ORCID: 0000-0002-9661-9638 (A. 1); 0000-0003-1601-9286 (A. 2)

to avoid AI amplifying existing discrimination, creating new grounds for discrimination, or reducing citizens' trust in the public sector.

The public sector in Norway is working purposefully to contribute to the development and improvement of public services in an efficient and sustainable manner. Digitization and the use of new technology are important tools to achieve this, and here, AI can make a significant contribution. The use of new technologies helps to create public digital services that are tailored to users' needs [15]. Used at its best, AI can contribute to the streamlining and improvement of public services. There are good examples of the use of AI in Norway, for example in the healthcare sector. The desire for and the potential for further use are emphasized, for instance in the National Strategy for Artificial Intelligence [14] pointing out that Norway has many things in place for succeeding with AI. Norway is characterized by a high level of trust in the public sector, a high degree of digital competence in the population, a well-developed technological infrastructure, and a public sector that has come a long way in developing digital administration. Special advantages in the work with AI are good registry data and long time series, which may provide important access to data for developing AI [16]. The AI strategy also emphasizes that public organizations "have the capacity and competence to experiment with new technologies", which can be crucial for the public to be able to adopt new technology, such as AI.

## 2. Methods

A mixed methods strategy was used, starting with an online survey inviting 467 state and municipal organizations from various sectors including health care, education, labour and welfare administration, tax authorities and more to respond to questions about use and plans of using AI. 200 organizations responded to the survey, out of these, 59 had active projects or plans for using AI, and 39 involved person data. Since the study focused on risks of discrimination, the organizations that had AI projects and plans involving personal data were invited to in-depth interviews, resulting in a total of 19 interviews with governmental, municipal, and inter-municipal organizations. The informants included 18 men and 9 women. Some were managers, while the majority had a professional background in information technology and AI. Below we share some of the findings from the analysis, based on the in-depth study of those AI projects that involved person data.

## 3. Findings

The National AI strategy employs a wide definition of AI, from machine learning (ML) to automatised processes. We started the survey with a similar wide definition to let the respondents decide what they included in a definition of AI, thus, among the AI projects involved in this study were both ML and simple automated procedures.

Many public sector organizations are exploring how AI can be used to improve public services and make them more efficient. The AI projects we encountered had a variety of goals, from improving the quality of data, detecting suspicious patterns and errors in the data, predicting needs in the organization or users' behaviour, and more. Some of these projects had been initiated to explore the *possibilities* of AI for the services, or for testing AI models. Most of the AI projects involving person data were, however, still in an early stage, exploring and developing the possibility of using AI.

Our analysis showed that the journey from *idea* to *production* of AI in public sector presents many challenges that often lead to AI projects being temporarily halted or terminated, resulting in only a few AI projects operating on person data that had reached the production stage. The figure below illustrates identified challenges of AI projects as a series of elements that need to be in place to safely navigate from start and the design of the project to the end and being "in production".


**Figure 1**: A hop-on-hop-off journey of AI projects

Although the elements of the figure do not represent a perfect or necessary chronological development, the figure illustrates the development of an AI project with the different elements representing stops on a hop-on-hop-off journey: Not all the AI projects had started with an overall design or management strategy, and not all had reached the end stop of being in production. Most respondents to the survey agreed that there was a need for more knowledge about AI among leaders. The emphasis on exploring the technology also illustrated weak leader involvement in some of these projects. It is rather a matter of who wants to participate, one of the informants told us. Data involved many challenges, as found elsewhere [17], from access to GDPR and juridical frameworks that made it difficult for public sector to use their own data for developing AI systems, to known and unknown bias in data [18, 19]. Access to technical expertise varied with the size of the organization, and this also influenced whether the system was developed in-house or bought, sometimes specially designed, sometimes referring to shelf-ready AI algorithms. Some of the projects with external tech support had experienced that lack of domain competence and insights into the organization's requirements had introduced weaknesses and challenges to the AI project. Some of the projects had experienced that failing to involve juridical competence at an early stage had led to an abrupt stop, and some struggled with strict interpretations of GDPR and risk assessments that for some became barriers for exploring AI possibilities. Finally, few of the AI projects had engaged with questions of discrimination. "We haven't thought about that, thank you for reminding us", one informant said. We also found a tendency for other considerations and concepts such as AI ethics, fairness, and transparency, taking precedence while discrimination in line with the definition of the Norwegian Anti-Discrimination Act was only a topic in a handful of AI projects. This seems to reflect discussions of AI in general [20, 21].

## 4. Discussion and conclusion

Our study focused on the public sector organizations with ongoing AI activities with an emphasis on projects involving person data. The findings illustrate that the AI development is still quite immature in large parts of the public sector. The development seems to be relatively little leadership-driven and more driven by interest and willingness to experiment and test the technology. In addition, there is political pressure to use AI in the public sector, which was mentioned several times in the interviews. Thus, we can identify three levels of driving forces for the current AI development in the Norwegian public sector:

- The political level: strategies for digitalization in the public sector and for AI in Norway create expectations about the use of AI [14, 15].
- The management level: the management level in the public sector was considered to have limited knowledge of AI and a limited focus on establishing and directing the development of AI in the sector.
- The organizational level: individuals and units within each organization that expressed a "willingness to contribute" to AI development were important for initiating AI projects.

While the political level pushes AI into the sector, even creating a pressure for taking it up, the management level is less visible in many of the AI projects, while individuals and units with knowledge about and the resources to experiment with AI, seem to be the driving force in many organizations. This indicates that the public sector is still quite immature in the field of AI, also reflected in other studies [16, 22].

Another critical issue for engaging AI in the public sector was the confusion arising from the wide definition of AI in the National AI strategy, involving everything from simple automation to ML techniques. Only automated procedures can be used for decision making in public sector, while all kinds of AI techniques involving ML can only be used for supporting decisions, with humans having the final say. Putting these different technologies in the same pot creates a confusion both inside and outside the public sector, making it less clear when and how technology vs. humans is making the decisions.

This also highlights the importance of interdisciplinary competence for successful AI projects, as these projects do not operate in a digital vacuum but must interact with a range of different social, cultural, political, legal rules and regulations [10, 21].

Finally, the issue of discrimination is not high on the agenda of AI projects in the public sector. Other concepts such as AI ethics, fairness, and transparency appear to take precedence before discrimination, leaving a gap to be filled by future policy and practice.

## 5. Acknowledgements

This paper reports from the project "Use of artificial intelligence in the public sector and risk of discrimination", which was carried out on behalf of the Norwegian Directorate for Children, Youth and Family Affairs (Bufdir) by Rambøll Management Consulting and Vestlandsforsking (Western Norway Research Institute) in the period from November 2021 to December 2022.

We want to thank all contributors, from Norwegian and international AI experts participating in the reference group and in meetings to discuss the project. We also want to thank the contributors from the public sector responding to the survey and participating in interviews.

The paper is developed from the project report, available in Norwegian: https://www.vestforsk.no/en/project/artificial-intelligence-public-sector-and-risk-discrimination.

## 6. References

[1] Di Noia, T., et al., *Recommender systems under European AI regulations.* Communications of the ACM, 2022. **65**(4): p. 69-73.

[2] Sousa, W.G.d., et al., *How and where is artificial intelligence in the public sector going? A literature review and research agenda.* Government Information Quarterly, 2019. **36**(4): p. 101392.

[3] Belenguer, L., *AI bias: exploring discriminatory algorithmic decision-making models and the application of possible machine-centric solutions adapted from the pharmaceutical industry.* AI and Ethics, 2022.

[4] Barbieri, D., et al., *Artificial intelligence, platform work and gender equality*. 2022, European Institute for Gender Equality (EIGE).

[5] White, J.M. and R. Lidskog, *Ignorance and the regulation of artificial intelligence.* Journal of Risk Research, 2021: p. 1-13.

[6] Lepri, B., N. Oliver, and A. Pentland, *Ethical machines: The human-centric use of artificial intelligence.* iScience, 2021. **24**(3): p. 102249.

[7] Mannes, A., *Governance, Risk, and Artificial Intelligence.* Mannes, A.(2020). Governance, Risk, and Artificial Intelligence. AI Magazine, 2020. **41**(1): p. 61-69.

[8] Zuiderveen Borgesius, F., *Discrimination, artificial intelligence, and algorithmic decision-making.* 2018.

[9] Broomfield, H. and M.N. Lintvedt, *Is Norway Stumbling into an Algorithmic Welfare Dystopia?* Tidsskrift for velferdsforskning, 2022. **25**(3): p. 1-15.

[10] Bartoletti, I. and R. Xenidis, *Preliminary draft Council of Europe study on the impact of artificial intelligence, its potential for promoting equality, including gender equality, and the risks to non-discrimination*. 2022: The Gender Equality Commission (GEC) and the Steering Committee on Anti-Discrimination, Diversity and Inclusion (CDADI), The Council of Europe, https://rm.coe.int/gec-2022-9-study-on-ai-211022/1680a8ad89

[11] Xenidis, R. and L. Senden, *EU non-discrimination law in the era of artificial intelligence: Mapping the challenges of algorithmic discrimination*, in *General Principles of EU law and the EU Digital Order*, U. Bernitz, et al., Editors. 2020, Kluwer Law International. p. 151-182.

[12] UNESCO, *Artificial intelligence and gender equality: key findings of UNESCO's Global Dialogue*. 2020: Division for Gender Equality, UNESCO.

[13] Commission, E., *Laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts*. 2021, Office for Official Publications of the European Communities Luxembourg.

[14] Kommunal- og moderniseringsdepartementet, *Nasjonal strategi for kunstig intelligens*. 2020, https://www.regjeringen.no/no/dokumenter/nasjonal-strategi-for-kunstig-intelligens/id2685594/ (lastet 24. august 2022).

[15] Kommunal- og moderniseringsdepartementet, *Én digital offentlig sektor: Digitaliseringsstrategi for offentlig sektor 2019–2025*. 2019, https://www.regjeringen.no/no/dokumenter/en-digital-offentlig-sektor/id2653874/ (lastet 24. august 2022).
[16] Broomfield, H. and L.M. Reutter, *Towards a Data-Driven Public Administration: An Empirical Analysis of Nascent Phase Implementation.* Scandinavian Journal of Public Administration, 2021. **25**(2): p. 73-97.
[17] Gröger, C., *There Is No AI Without Data.* Communications of the ACM, 2021. **64**(11): p. 98-108.
[18] Friedman, B. and H. Nissenbaum, *Bias in computer systems.* ACM Transactions on information systems (TOIS), 1996. **14**(3): p. 330-347.
[19] Srinivasan, R. and A. Chander, *Biases in AI systems.* Communications of the ACM, 2021. **64**(8): p. 44-49.
[20] Gjerdsbakk, T.C.G., *Åpen og rettferdig kunstig intelligens.* Lov & Data, 2022. **150**(3).
[21] Gerards, J. and R. Xenidis, *Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law*. Gerards, J. and Xenidis, R. (2021) Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law. 2021: European Commission.
[22] Andréasson, U. and T. Stende. *Nordic municipalities' work with artificial intelligence*. 2019.

# Exploiting Privacy Inference for Data Auditing

Rui Zhang[1], Song Guo[1]

[1]The Hong Kong Polytechnic University, 11 Yuk Choi Road, Hung Hom, Kowloon, Hong Kong SAR

Abstract

The success of machine learning models is mainly attributed to the massive amount of training data. However, as the raw data is not available, one cannot tell whether private or biased data has been used or not. In this paper, we highlight the importance of data auditing for published models and propose to exploit privacy inference methods for judgment. We hope that this work can inspire subsequent studies.

**Keywords**

Machine learning, data auditing, data regulation, privacy inference

## 1. Introduction

Machine learning has made significant progress in recent years, with advancements in the areas of natural language processing, computer vision, and deep neural networks. These developments have led to the production of powerful models which are capable of performing incredible generative tasks, such as Stable Diffusion [1], DALL·E 2 [2] and ChatGPT [3]. These technological leaps are not only dependent on upgraded hardware computing resources, but also require massive amounts of data as fuel. For instance, ChatGPT is trained on 570GB of data obtained from the Internet, including websites, books and articles [4].

However, as the original training data is not available, people would like to know if there is inappropriate usage of private or disputed samples. As recently as February 2023, Getty Images has sued the creator of Stable Diffusion and claims that 12 million copyrighted images were 'unlawfully' copied [5]. Therefore, data audits of the published models are essential. On the one hand, it is possible to verify whether a given data is used by an enterprise to train product-level models without authorization. On the other hand, the users would like to know if the training data is disputed or discriminatory towards certain minority groups.

In this paper, we aim to highlight the importance of data auditing for published models and to give a feasible approach to auditing. The difficulty with auditing is how the auditor can make a judgment without access to the training data, just by the model parameters or the outputs returned by the API interface. We propose that auditors can fight fire with fire, in other words, to exploit existing privacy inference attacks to indirectly achieve the purpose of auditing. Specifically, member inference attacks (MIA) [6, 7, 8] can be applied to verify the misuse of sensitive data, while property inference attacks (PIA) [9, 10, 11] can be performed to inspect the characteristics or the distributions of the training data. We hope to raise awareness of the significance of auditing, and shed light on the following works in this area.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: csrzhang1@comp.polyu.edu.hk (R. Zhang); song.guo@polyu.edu.hk (S. Guo)
ORCHID: 8 0000-0002-4638-5132 (R. Zhang); 0000-0001-9831-2202 (S. Guo)


## 2. Privacy Inference for Auditing

In this section, we first introduce the two most prevalent attacks on privacy inference - membership inference attacks (MIA) and property inference attacks (PIA), and then describe how these approaches can be exploited for auditing the data. Finally, we formalize the challenge as a game of auditing and convert it to a classification or a regression problem.

### 2.1. Membership Inference Attacks

The membership inference attack (MIA) was initially proposed by Shokri et al. in 2017 [6]. As a privacy leakage approach, the goal of the attacker is to identify the membership of a specific sample, that is, to determine whether a given data has been used in the training process or not. Actually, MIA can be classified into two categories: white-box and black-box attacks, which reflects the difference in the capabilities of the attacker.

In the white-box attacks, the attacker has full access to the model and its parameters [12]. In contrast, in black-box attacks, the attacker only has access to the input-output behavior (e.g., predictions of the input data) of the model [6, 7]. Furthermore, in collaborative learning or federated learning, the updated gradients can also be exploited as white-box information for MIA [13, 14]. In general, white-box attacks are more powerful than black-box attacks, as the model parameters or gradients imply more information about the original data. Like the purpose of MIA, data auditing is intended to examine whether the training dataset includes a private sample when only the model or API access is available.

### 2.2. Property Inference Attacks

Property Inference Attack (PIA) is another type of privacy leakage that aims to infer sensitive global properties of the training dataset from a machine learning model. Different from MIA, PIA is more interested in extracting a certain attribute or distribution of the training dataset, rather than an individual instance. For example, in a dataset with ethnicity, the attacker may want to know which attribute value is dominant (e.g., if there are more whites) or what the precise ratio of the attribute value is (e.g., 60% whites vs. 35% whites) [11].

Similar to MIA, PIA can also be divided into white-box attacks [9, 10] and black-box attacks [11], corresponding to the access to the model parameters and the API interface respectively. For auditing applications, the usage of PIA as a tool can discriminate whether there is some preference, or even undesirable discrimination, in the training dataset. In Table 1, we summarize these approaches, the required knowledge and the objectives of auditing.

**Table 1**
Privacy Inference Attacks for Data Auditing

| Inference Approaches | Knowledge | Objectives of Auditing |
|---|---|---|
| Membership Inference (MIA) | Black-box / White-box | Existence of a target private sample |
| Property Inference (PIA) | Black-box / White-box | Preference of a biased distribution |

### 2.3. Problem Statement of Auditing

As privacy inference attacks are introduced as tools, we then formalize the problem of auditing. Inspired by previous works [15, 16], we consider a game of data auditing between an auditor $\mathcal{A}$ and a challenger $\mathcal{C}$. The goal of the auditor is to use privacy inference approaches on the challenger's model to find potential training samples (or properties) that are private (or biased). The challenger could be a company or any entity whose released model is suspected of unlawful data usage. With access to the model's parameters or the API's interface, the auditor has to estimate which feature $p$ from the prior knowledge space $\mathcal{P}$ is inherent in the training data $D_p$. By appropriately defining the space $\mathcal{P}$, this game can be generalized to infer the membership of a target sample or a property of the training dataset.

**Game 3.1** (Auditing Game). The game proceeds between an auditor $\mathcal{A}$ and a challenger $\mathcal{C}$. Both have access to a distribution $\mathbb{D}$, and know the prior space $\mathcal{P}$ and training algorithm $\mathcal{T}$.

(1) The challenger samples a dataset $D_p \leftarrow \mathbb{D}$ with a target feature $p$ from $\mathcal{P}$.
(2) The challenger trains a model $f_\theta \leftarrow \mathcal{T}(D_p)$ on the dataset $D_p$.
(3) The challenger gives the auditor query access to the model $f_\theta$.
(4) The auditor presents an estimation $\hat{p}$ about the target feature $p$.
(5) The auditor wins the game if $\hat{p} = p$ or $\|\hat{p} - p\| \leq \epsilon$ ($\epsilon$ is small).

The space $\mathcal{P}$ captures the auditor's prior knowledge about the possible values that the target features $p$ may take. In the membership inference game, for a particular target sample $x$ the feature is denoted as $\mathcal{P} = \{x, \perp\}$, where $\perp$ indicates the absence of the sample data. That is, the auditor needs to estimate whether the model $f_\theta$ is trained on $D_{p=x}(x \in D)$ or $D_{p=\perp}(x \notin D)$. While for the property inference game, the space $\mathcal{P}$ contains a set of target attribute values, i.e. $\mathcal{P} = \{a_1, a_2, \cdots, a_k\}$, wherein the set could be finite or infinite. The auditor provides an estimation $\hat{p}$ of the inherent property $p \in \mathcal{P}$ of the training dataset $D_p$. If the distance (i.e., $\|\hat{p} - p\|$) is bounded in a small range $\epsilon$, the auditor wins the game. Generally, the auditing of membership can be treated as a binary classification problem, while the game of property can be regarded as either a classification or a regression problem.

## 3. Conclusion

In this paper, we elaborate on the necessity of data auditing for published models, and propose a series of feasible approaches with reliable privacy inference attacks. We first introduce the membership inference attacks and property inference attacks, and explain why these methods can be exploited for examining the training data. Then, we formalize the challenge as a game of auditing, which can be converted to a classification or a regression problem. The practical implementation of solving such a problem is left to future research. We hope to attract the wide attention of data auditing and shed a light on the following works.

## References


[1] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, B. Ommer, High-resolution image synthesis with latent diffusion models, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 10684–10695.

[2] A. Ramesh, P. Dhariwal, A. Nichol, C. Chu, M. Chen, Hierarchical text-conditional image generation with clip latents, arXiv preprint arXiv:2204.06125 (2022).

[3] OpenAI, Chatgpt: Optimizing language models for dialogue, 2022. URL: https://openai.com/blog/chatgpt/.

[4] BBC Science Focus, Chatgpt: Everything you need to know about openai's gpt-3 tool, 2023. URL: https://www.sciencefocus.com/future-technology/gpt-3/.

[5] James Vincent, The Verge, Getty images sues ai art generator stable diffusion in the us for copyright infringement, 2023. URL: https://www.theverge.com/2023/2/6/23587393/ai-art-copyright-lawsuit-getty-images-stable-diffusion.

[6] R. Shokri, M. Stronati, C. Song, V. Shmatikov, Membership inference attacks against machine learning models, in: 2017 IEEE symposium on security and privacy (SP), IEEE, 2017, pp. 3–18.

[7] A. Salem, Y. Zhang, M. Humbert, M. Fritz, M. Backes, Ml-leaks: Model and data independent membership inference attacks and defenses on machine learning models, in: Network and Distributed Systems Security Symposium 2019, Internet Society, 2019.

[8] H. Liu, J. Jia, W. Qu, N. Z. Gong, Encodermi: Membership inference against pre-trained encoders in contrastive learning, in: Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security, 2021, pp. 2081–2095.

[9] G. Ateniese, L. V. Mancini, A. Spognardi, A. Villani, D. Vitali, G. Felici, Hacking smart machines with smarter ones: How to extract meaningful data from machine learning classifiers, International Journal of Security and Networks 10 (2015) 137–150.

[10] K. Ganju, Q. Wang, W. Yang, C. A. Gunter, N. Borisov, Property inference attacks on fully connected neural networks using permutation invariant representations, in: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, 2018.

[11] W. Zhang, S. Tople, O. Ohrimenko, Leakage of dataset properties in multi-party machine learning, in: 30th USENIX Security Symposium (USENIX Security 21), 2021, pp. 2687–2704.

[12] L. Melis, C. Song, E. De Cristofaro, V. Shmatikov, Exploiting unintended feature leakage in collaborative learning, in: 2019 IEEE symposium on security and privacy (SP), IEEE, 2019, pp. 691–706.

[13] M. Nasr, R. Shokri, A. Houmansadr, Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning, in: 2019 IEEE symposium on security and privacy (SP), IEEE, 2019, pp. 739–753.

[14] O. Zari, C. Xu, G. Neglia, Efficient passive membership inference attack in federated learning, arXiv preprint arXiv:2111.00430 (2021).

[15] S. Mahloujifar, E. Ghosh, M. Chase, Property inference from poisoning, in: 2022 IEEE Symposium on Security and Privacy (SP), IEEE, 2022, pp. 1120–1137.

[16] F. Tramèr, R. Shokri, A. San Joaquin, H. Le, M. Jagielski, S. Hong, N. Carlini, Truth serum: Poisoning machine learning models to reveal their secrets, in: Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, 2022, pp. 2779–2792.

# A Comprehensive Framework for Detecting Behavioural Anomalies in Elderly

Ankit Jain[1], Abhishek Srivastava[2]

*Indian Institute of Technology Indore, Madhya Pradesh, India*

Abstract

The world is seeing a rapid increase in the population of the aged. This combined with a shortage of affordable care-giving man-power leads to a dependence on automated systems for monitoring the well being of the elderly and detecting abnormalities. There exist techniques based on sensors of various types to detect and recognize daily activities of the elderly and detect anomalies. While such sensor based techniques are effective at detecting immediate exigencies, they are unable to comprehend gradual deterioration in the behavior of the elderly indicating conditions like dementia and alzheimer's, for example. This aspect is also not properly addressed in literature. This paper introduces a system for the comprehensive detection of anomalies in the activities of the elderly using a graph-based approach. This employs dynamic activity graphs where anomalies are detected using a dissimilarity score. The proposed approach is capable of detecting both short-term and long-term anomalies in the daily activities of the elderly.

**Keywords**

Behavioural Anomaly, Activity Graph, Graph Matching, Activity Detection and Recognition.

## 1. Introduction

The world continues to age rapidly, and the number of persons aged above 65 years is expected to reach 1.5 billion in 2050 [1]. Countries are therefore grappling with the issue of facilitating independent living for the elderly. With labour shortage being another issue, substantial investment is directed toward developing automated monitoring capabilities for the elderly. There are numerous approaches in literature for detecting daily activities of the elderly (and also in general) in indoor settings through the use of wearable sensors [2], ambient sensors [3], and vision sensors [4, 5].

Anomalies in the behaviour of the elderly may be classified as 'sudden' including phenomena like falls termed as *short term anomaly* and 'gradual' that encompasses variations in the daily routine of an individual over time termed as *long term anomaly*. Few examples of the long term anomalies in elderly are: elderly start spending more/less time in bed; started using toilets more than usual; not going out; missing other activities more often. Few of these anomalies may also be indicated by the short term anomaly detection systems but the nature of such anomalies will not be severe but if these patterns keep repeating for the longer duration than can be dangerous and termed as long term anomaly. Approaches in literature are mostly capable of recognizing short term anomalies and very few log and analyse changes in behaviour to detect long term anomalies. They miss out, therefore, on potentially detecting and comprehending

Proceedings of SAIDD 2023, May 09-10, 2023, Sogndal, Norway
EMAIL: phd1801101002@iiti.ac.in (A. Jain); asrivastava@iiti.ac.in (A. Srivastava)
ORCHID: 0000-0002-2119-222X (A. Jain); 0000-0002-6338-5476 (A. Srivastava)

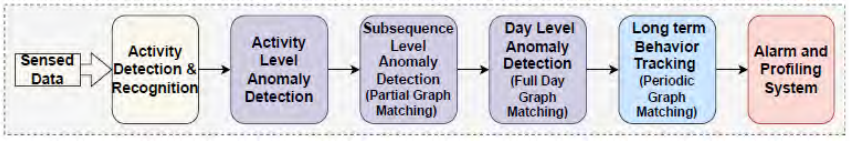




**Figure 1:** Workflow of the proposed Comprehensive Anomaly Detection System(CADS).

patterns in behavioural change that may not require immediate intervention but can indicate long-term pernicious health issues. The few endeavours in literature that do focus on long term behavioural anomaly detection take one of the following approaches: activity-wise anomaly detection [6, 7] that detects abnormalities in the execution of an activity, i.e., changes in the duration of the activity, or the location where the activity is performed; and sub-sequence wise anomaly detection [8, 9] that looks at anomalies in the ordering and sequencing of activities [10, 11].

These approaches classify behavioural anomalies either at a fine-grained level and look at individual activities separately, or a more coarse-grained level and focus on combinations of activities over the whole day. In addition to this, detected anomalies are handled similarly, irrespective of their level of severity. For example, sleeping more/less than normal is a less severe anomaly than missing medicine doses whereas falling down is the most severe anomaly and demands immediate intervention. There needs to be, therefore, a robust mechanism that immediately alerts the concerned caregiver for severe anomalies or keeps track of less severe anomalies that eventually over a period of time indicate a behavioural change and perhaps the build up of an adverse medical condition.

This paper aims at comprehensive activity monitoring and anomaly detection over the long term. In addition to this, a mechanism is proposed for triggering immediate intervention for severe anomalies; and for studying and analysing behaviour patterns over a period of time based on detecting and logging less severe anomalies.

## 2. Proposed Methodology

Existing approaches for activity monitoring and anomaly detection, as discussed earlier, look at limited portions of the activity network and are unable to comprehensively study behavioural patterns. Our objective is to develop a complete tool for monitoring the behavioural patterns of the elderly over short and long durations. The proposed system detects anomalies at the level of 1) individual activities, 2) within sub-sequences of activities, 3) within longer sequences of activities constituting the entire day, and 4) finally within activity networks spanning multiple days. Such comprehensive analysis and anomaly detection at varying levels of granularity leads to a good understanding of the behavioural patterns of the individual. This facilitates detection of anomalies in behaviour and drawing sound predictions on potentially adverse medical conditions.

The proposed system in Figure 1 comprises six modules starting from activity detection and extends up to the point that the system informs concerned personnel about the well being of the monitored individual. The activity detection and recognition module detects and classifies activities from streaming data emanating from various sensors (i.e., wearable, ambient, or vision). The proposed framework works at the activity level and utilizes existing activity detection and

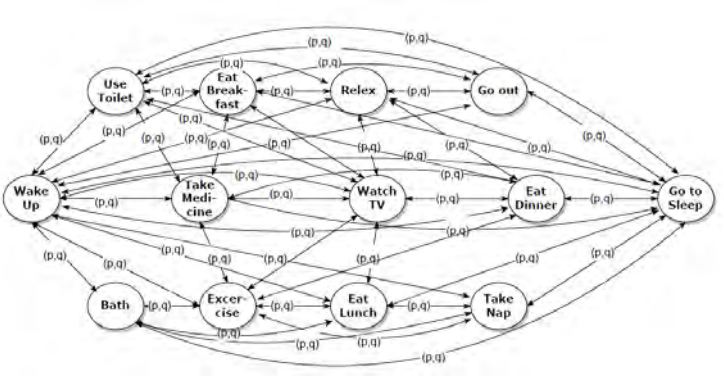



**Figure 2:** Daily Activity Graph for Elderly (Each edge A-B is represented by a tuple (p,q), where p is the transition probability from A to B and q is the transition probability from B to A.

recognition systems depending on the sensors involved. The daily activities of an individual are represented using a dynamic graph $G_t(V_t, E_t)$ [12] called the Daily Activity Graph (DAcG). A sample DAcG for an elderly person is shown in Figure 2. The nodes of the graph denote features describing an activity and the edges from a node denote the probabilities of occurrence of subsequent activities. The DAcG for several days together indicate the 'normal' behaviour of a monitored individual. We call this $G_{ref}$ the Reference Graph of daily behaviour.

Anomalies in the behaviour of a monitored individual are detected at three levels: activity level, sub-sequence level, and day level. An activity is classified as normal or anomalous based on features like activity start time, activity duration, location of activity. Based on these parameters a combined anomaly score is calculated and the nature of an anomaly is established (i.e., low, medium, high). A sub-sequence level anomaly is detected by matching[13, 14] a sub-graph $G_{t'}$ with the Reference Graph $G_{ref}$ with the latter being classified as anomalous or normal based on the matching scores. A *day level* anomaly is detected by matching the current DAcG $G_t$ with the Reference Graph $G_{ref}$.An anomaly and the nature of the anomaly is established by combining the matching scores and anomaly scores of individual activities.

The long term behaviour tracking module is responsible for tracking changes in the behaviour of the monitored elderly based on comprehensive assessment of his/her daily activities over a period of time (i.e., weeks, months, quarters, or years). This module comprises generating periodic graphs and comparing successive periodic graphs to identify abnormal patterns, if any.

The alarm and profiling module primarily works towards establishing a robust system for effective communication of an adverse medical condition to designated caregivers of the monitored individual.

## 3. Conclusion

In this paper, we proposed a comprehensive system for anomaly detection in the daily activities of the elderly. The proposed system is unique because it not only enables the detection of immediate anomalies but also enables analysis and comprehension of behavioural patterns of an individual over the long term. An understanding of the latter facilitates early detection and diagnosis of pernicious developments like dementia and alzhemer's. We expect the system to be of utility to the elderly and also assist caregivers in providing better care.

## References


[1] United Nations Department of Economic and Social Affairs, Population Division, World population ageing 2020 highlights, https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/undesa_pd-2020_world_population_ageing_highlights.pdf, 2020.

[2] W. Huang, L. Zhang, W. Gao, F. Min, J. He, Shallow convolutional neural networks for human activity recognition using wearable sensors, IEEE Transactions on Instrumentation and Measurement 70 (2021) 1–11.

[3] H. Y. Yatbaz, S. Eraslan, Y. Yesilada, E. Ever, Activity recognition using binary sensors for elderly people living alone: Scanpath trend analysis approach, IEEE Sensors Journal 19 (2019) 7575–7582.

[4] J. Zhang, F. Shen, X. Xu, H. T. Shen, Temporal reasoning graph for activity recognition, IEEE Transactions on Image Processing 29 (2020) 5491–5506.

[5] J. Yin, J. Han, R. Xie, C. Wang, X. Duan, Y. Rong, X. Zeng, J. Tao, Mc-lstm: Real-time 3d human action detection system for intelligent healthcare applications, IEEE Transactions on Biomedical Circuits and Systems 15 (2021) 259–269.

[6] P. Parvin, S. Chessa, M. Manca, F. Paterno', Real-time anomaly detection in elderly behavior with the support of task models, Proceedings of the ACM on human-computer interaction 2 (2018) 1–18.

[7] S. Saqaeeyan, H. Amirkhani, et al., Anomaly detection in smart homes using bayesian networks, KSII Transactions on Internet and Information Systems (TIIS) 14 (2020) 1796–1816.

[8] S.-C. Poh, Y.-F. Tan, X. Guo, S.-N. Cheong, C.-P. Ooi, W.-H. Tan, Lstm and hmm comparison for home activity anomaly detection, in: 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), IEEE, 2019, pp. 1564–1568.

[9] H. Gao, L. Zhou, J. Y. Kim, Y. Li, W. Huang, The behavior guidance and abnormality detection for a-mci patients under wireless sensor network, ACM Transactions on Sensor Networks (2021).

[10] L. Wang, Y. Zhou, R. Li, L. Ding, A fusion of a deep neural network and a hidden markov model to recognize the multiclass abnormal behavior of elderly people, Knowledge-Based Systems 252 (2022) 109351.

[11] C. Azefack, R. Phan, V. Augusto, G. Gardin, C. M. Coquard, R. Bouvier, X. Xie, An approach for behavioral drift detection in a smart home, in: 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE), IEEE, 2019, pp. 727–732.

[12] S. Khoshraftar, S. Mahdavi, A. An, Y. Hu, J. Liu, Dynamic graph embedding via lstm history tracking, in: 2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA), IEEE, 2019, pp. 119–127.

[13] M. Han, H. Kim, G. Gu, K. Park, W.-S. Han, Efficient subgraph matching: Harmonizing dynamic programming, adaptive matching order, and failing set together, in: Proceedings of the 2019 International Conference on Management of Data, 2019, pp. 1429–1446.

[14] V. Ravindra, G. Sanders, A. Grama, Identifying coherent subgraphs in dynamic brain networks, in: 2021 IEEE International Conference on Image Processing (ICIP), IEEE, 2021, pp. 121–125.

# I-KNOW-FOO: Interlinking and Creating KNOWledge Graphs for near-zero CO2 emission diets and sustainable FOOd production

Görkem Simsek-Senel [1], Hajo Rijgersberg [1], Bengü Öztürk [1], Jeroen Weits[2], and Anna Fensel[2,3]

[1] *Food Informatics, Wageningen Food & Biobased Research, Wageningen University & Research, Bornse Weilanden 9, Wageningen, 6708 WG, the Netherlands*
[2] *Consumption and Healthy Lifestyles, Wageningen University & Research, Hollandseweg 1, Wageningen, 6706 KN, the Netherlands*
[3] *Wageningen Data Competence Center, Wageningen University & Research, Droevendaalsesteeg 2, Wageningen, 6708 PB, the Netherlands*

**Abstract**

It is already known that the diet of the world's population has a massive impact on climate change. However, how climate change effects the growing conditions of ingredients for different foods and beverages, and emission rates due to, for example, production and logistics are still not known. In this work, different datasets have been explored to study the feasibility of interlinking datasets to automatically generate alternatives for the climate change–aware food items substitution. A core question to be answered is what the alternatives of the mostly consumed crops in current diets in the Netherlands in case of a climate change can be. The main crop attributes taken into account are nutritional composition and the growing conditions. The growing conditions of three most-consumed crops in the Netherlands have been linked manually to their nutritional composition data and a corresponding knowledge graph is created. This study shows that linking various data semantically promises to generate alternatives automatically.

**Keywords**

Climate change, food replacement, knowledge graph, reasoning, decision making

## 1. Introduction

It is already known that the diet of the world's population has a massive impact on climate change [1, 2]. However, still too little attention is being paid to the climate change's impact on the growing conditions of ingredients for different foods and beverages and to emission rates due to, for example, production and logistics. The provenance and climate change impact of various foods are often not clearly known or accessible, both for end consumers as well as for the whole supply chain elements.

To give an example, many food options are un-trivial and interdependent in terms of sustainability, for example, it may be not trivial to consumers that production of mineral water – due to the packaging materials used – may be more damaging the climate than the production of rice, and further aspects (e.g. logistics) become relevant. As in all information-intensive environments, food producers and consumers continuously face complex decisions on which ingredients or products to choose, in which amounts and how to process them or which alternatives to select for the products they consume regularly. To make decisions, they need access to data about these food items, for example their nutritional

[1]roceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: gorkem.simsek-senel@wur.nl (A.1); hajo.rijgersberg@wur.nl (A.2); bengu.ozturk@wur.nl (A.3); jeroen.weits@wur.nl (A. 4); anna.fensel@wur.nl (A. 5)
ORCID: 0000-0002-1391-7104 (A. 5)


value, taste, sustainability characteristics as well as the nutrients that they need to consume and logistics information. This information is still scattered, and the quality of the data varies. Meanwhile, data indicating climate change impact of different foods and beverages exists (or can be collected) as well as data on the supply chains. However, these data are still often not easily available and discoverable, and have no explicit connections between them.

In this work we aim to reach a clear understanding of the diets and how to make them equivalently nutritious, sustainable, but also evolve them accounting for the climate change or re-design the diets considering and adapting to climate change characteristics taking into account the growing conditions. The main question that is to be answered with this research is: "How can we interlink datasets so that an alternative to the current consumed products can be (automatically) found by taking into account the nutritional composition, growing conditions which will be effected by climate change and sustainability information?" and "To what extent can that process be automated?".

The objective of the research is to identify the relevant data and make them more accessible for discoveries and supporting (automatic) decision making in the food supply chain and for end consumers. Thus, the goal is to develop and generate knowledge graphs, benefiting from semantic technology which helps interlinking scattered information using standardized concepts. With employment of knowledge graphs and using them for interlinking, one will be able to create a web-like large-scale data infrastructure and tools to easily explore it for everyone, as well as to assist in making estimates of $CO_2$ footprints of various foods and beverages, and how to adapt the diets given the climate change.

The remainder of this document is organized as follows. Section 2 explains the methodology of the work. In Section 3, the results are presented, and Section 4 describes the conclusions.

## 2. Approach

As the datasets are scattered, the research started initially by making an inventory of the available datasets, ontologies and knowledge graphs on food products, and the impact of climate change on availability the food products availability. The datasets screened were SHARP Indicators Database, Food Consumption Impact datasets (Optimeal-Blonk Sustainability Datasets), RIVM Sustainability dataset, Pizza dataset, The Eaternity Database, Data Explorer: Environmental Impacts of Food, Dataset on potential environmental impacts of water deprivation and land use for food consumption in France and Tunisia, Ireland and UK database (Michael Clark), World Food LCA Database. Among those, a few are publicly available [3, 4, 5, 6, 7, 8, 9, 10, 11].

As the databases and datasets are from all over the world, the food products vary from one database to another and it is not straightforward to map them. Additionally, the existing food databases provide information about impact of food consumption on sustainability and they do not have a direct link to changing climate conditions which is required to determine alternatives for the original products that are currently part of the diets.

We then defined a use case that focuses on the most-imported crops in the Netherlands to connect the consumption to the changing climate. The imported crops are used to evaluate their important nutritional values and to find alternative crops in case the Netherlands may run out of the most-imported crops in a changing climate over years. For this goal we use the FAOSTAT Database [12]. The crop information was manually interlinked to growing conditions. The most useful information was considered to be found in the ECOlogical CROP Database (ECOCROP).

## 3. Results

The use case focuses on the most-imported crops in the Netherlands. We first have followed a manual process to run the use case. Our aim is to determine the most imported crops to the Netherlands in order to evaluate their important nutritional quantities and values and to find alternative crops to mostly imported crops in case Netherlands may run out of these in a changing climate over years. Top 10 commodities that were imported to the Netherlands within last 5 years (2016-2020) were screened using TRADE Datasets for Crops and livestock products in FAOSTAT database. Moreover, commodities supplied to the Netherlands were evaluated using Food Balance Datasets in terms of Domestic

Food Supply Quantity (1000 tonnes/yr) and Food Supply Quantity (kg/capita/year). These commodities are listed in terms of their import quantity, import values, supply quantities in descending order. We focused on three main commodities which are imported in high quantities and supplied to Dutch population, and selected *soybean, wheat* and *potato* as mostly imported and consumed food products. The next step was to find nutritionally similar alternative crops using the NEVO Dutch Food Composition Database. We have also searched for growing conditions of original and alternative crops, and developed knowledge graphs to link these data and re-use parts of the existing knowledge graphs. The alternatives are generated by manually processing the intersection of the different result sets.

**Manually generated alternatives.** To find alternatives to the three crops, we have focussed on parameters of climate resilience, nutrient-rich comparable crops and food products that have been screened using knowledge rules provided by a dietary expert using the NEVO Database. These possible alternative crops have then been evaluated in terms of their resistance to temperature increases in a changing climate using crop growth temperatures from the ECOCROP Database.

**Generating an ECOCROP ontology.** The ECOCROP database is transformed into a knowledge graph manually. First, the dataset has been cleaned. The measurementType 'optimalGrowthTemperature' has been subdivided into *maxGrowthCelsiusTemperature* and *minGrowthCelsiusTemperature* to distinguish between the two as well as add a unit into the predicate. The triples consist of the occurenceID as subject, measurementType as predicate and measurementValue as object. They have been transformed using OntoText's Refine tool and have been loaded into an RDF repository in RDF4J. OntoText Refine is a software tool that supports the transformation of string data into knowledge graphs [13].

**ECOCROP extension and interlinking to FIO and FoodOn.** We have extended ECOCROP manually by adding triples linking some of the *occurenceID*s in ECOCROP to the IDs of crops in FoodOn (including NCBITaxon [14]) and food items in FIO (Food Item Ontology [15]), based on the RIVM NEVO IDs. In FoodOn, we have chosen for instances of the organism class, because it represents the plants rather than the different foods that may originate from these plants. The plants are grown under (climate-changing or not) temperatures, namely, not particularly the foods. The relation used for linking the concepts is the *owl:sameAs* relation.

Subsequently, we have loaded the triples in the triple repository, where the information can be queried using SPARQL. In the future, this could be done by an automated tool. The query that we have formulated searches for crops that are more resilient to a warmer climate, being candidates to replace the current crop. So far in this exercise, we have only focused on the maximum growing temperature being one of the important factors in climate change on crop growth [16]. In our examples, the maximum optimal growing temperatures are 33°C for soybean, 23°C for wheat and 25°C for potato. Combining this information with nutritional values information, still leaves multiple options for food alternatives with similar nutrition characteristics. For example, for potatoes, possible alternatives are beans white/ brown dried, peas green dried, chestnuts raw, tapioca, cassava raw, taro raw, yam raw, tannia raw, beans black eyed dried, peas split yellow/green dried, tamarind, flour cassava.

The open access ECOCROP ontology and knowledge graphs created in our project are available at: https://git.wur.nl/FoodInformatics/i-know-foo.git

## 4. Conclusions

More sustainable food production, distribution and consumption options can be discovered by all stakeholders, eventually leading to near-zero CO2 emission diets and sustainable food production that will have a positive impact on climate change and will also be adaptive to it. Linking datasets and unchaining the information about crops and food products allow automatically finding nutritionally similar alternatives in a case of changing climate. This research demonstrates that automation is possible here. In this work, alternatives are generated manually for three most-imported crops in the Netherlands to showcase the feasibility of automatic generation. The growing conditions of the crops are defined in the created by us ECOCROP ontology that is based on open ECOCROP data. The linking between NEVO database and the ECOCROP ontology is done through the NEVO codes.

**Acknowledgements**. This work has been partially funded by WUR investment theme "Data Driven discoveries in a Changing Climate". The authors would like to thank Edith Feskens, Sander de Leeuw, and Karin Borgonjen for their contributions to this study.

## 5. References


[1] E. Stehfest, L. Bouwman, D.P. Van Vuuren, M.G. Den Elzen, B., Eickhout, B., P. Kabat (2009). Climate benefits of changing diet. Climatic change, 95(1), 83-102.

[2] B. Neha, T. Hills, D. Sgroi. *Climate Change and Diet*. No. 13426. Institute of Labor Economics (IZA), 2020.

[3] E. Mertens, G. Kaptijn, A. Kuijsten, H. van Zanten, J. M. Geleijnse, & P. van't Veer (2019). SHARP-Indicators Database towards a public database for environmental sustainability. *Data in brief*, *27*, 104617.

[4] Blonk Sustainability | Databases, https://blonksustainability.nl/tag/Databases

[5] RIVM Life Cycle Assessment (LCA) database, https://www.rivm.nl/life-cycle-assessment-lca

[6] A. Cortesi, C. Pénicaud, A. Saint-Eve, L.G. Soler & I. Souchon (2022). Life cycle inventory and assessment data for quantifying the environmental impacts of a wide range of food products belonging to the same food category: A case study of 80 pizzas representatives of the French retail market. *Data in brief*, *41*, 107950.

[7] Eaternity Database, https://eaternity.org/foodprint/database

[8] World Food LCA Database, https://ourworldindata.org/explorers/

[9] C. Sinfort, M. Perignon, S. Drogué, & M. J. Amiot (2019). Dataset on potential environmental impacts of water deprivation and land use for food consumption in France and Tunisia. Data in brief, 27, 104661.

[10] M. Clark, M. Springmann, M. Rayner, P. Scarborough, J. Hill, D.Tilman, D., ... & R. A. Harrington (2022). Estimating the environmental impacts of 57,000 food products. *Proceedings of the National Academy of Sciences*, *119*(33), e2120584119.

[11] B. Notarnicola, G. Tassielli, P.A. Renzulli, R. Di Capua, G. Saija, R. Salomone, ... & M. Mistretta (2022). Life cycle inventory data for the Italian agri-food sector: background, sources and methodological aspects. *The International Journal of Life Cycle Assessment*, 1-16.

[12] FAOSTAT, https://www.fao.org/faostat/en/#home

[13] Ontotext Refine tool, https://www.ontotext.com/products/ontotext-refine/

[14] NCBITaxon, http://purl.obolibrary.org/obo/NCBITaxon#

[15] Food Item Ontology, https://git.wur.nl/FoodInformatics/foodontology.git

[16] Hatfield, J. L., Boote, K. J., Kimball, B. A., Ziska, L. H., Izaurralde, R. C., Ort, D., and Wolfe, D. (2011). Climate impacts on agriculture: implications for crop production. Agronomy journal, 103(2), 351-370.

# A new approach to human cognitive modeling for the metaverse design system development

Yan Hong [1,2], Xingyuan Wang [1], Song Guo [2] and Xianyi Zeng [3]

[1] *Soochow University, 178 Ganjiang Road,215021, China*
[2] *The Hong Kong Polytechnic University, 11 Kowloon Hung Hom Yuk Choi Road , Hong Kong, 100872 ,China*
[3] *University of Lille, Métropole européenne de Lille , Roubaix, 59100, France*

**Abstract**

In this paper, an interactive design system facing the Metaverse is proposed, that is expected to be capable of integrating the design's physical content. For this purpose, the key to realize this system, the product design process and its related design knowledge, has been modelled. The designer's professional knowledge is extracted and simulated as several perceptual cognition models. A fuzzy transformer method is innovatively developed for the perceptual cognition model's computational modelling.

**Keywords**

Metaverse Design System, Design support system, Knowledge-based system, Human-centred design, Fuzzy transformer.

## 1. Introduction

The Metaverse is considered to be another revolution of the Internet [1], which can bring users an immersive experience and provide them with a channel connecting the real and virtual worlds [2,4,5,7-9]. In terms of design, the virtual environment provided by the Metaverse is completely a new environment [6]. The completion of design practice is based on a generalized design process and its related design knowledge [12].

The proposed system follows the general design process of Design-Display-Evaluation-Adjustment. The main sections of this work are as follows:

- An interactive knowledge-based design system facing the Metaverse is developed to enhance the user-Metaverse interaction.
- The general design process of Design-Display-Evaluation-Adjustment has been raised and applied to the proposed system to realize the interaction mechanism for the system.
- A fuzzy transformer method is innovatively developed for the perceptual cognition model's computational modelling.

## 2. The proposed system and its disciplinary

### 2.1. Principle of the proposed system

The structure of the proposed system is shown in Figure 1. The proposed system links the real world and the Metaverse. The general working principle follows the cycle of Generation-Display-Evaluation-Adjustment.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: yannic.hong@polyu.edu.hk (Y.Hong); wang@soochow.cn (X.Wang); song.guo@polyu.edu.hk (S.Guo); xianyi.zeng@ensit.fr (X.Zeng)

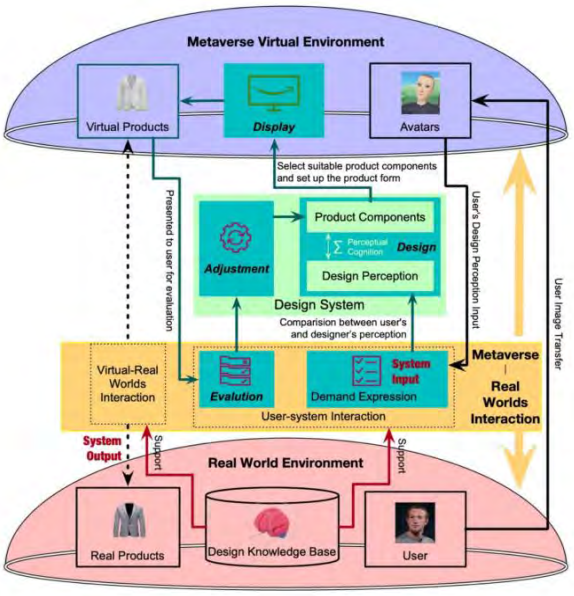



**Figure 1**: Working flowchart of the proposed automatic pattern generation system.

## 2.2. Principle of the proposed system

1. The basic process of design is "communication-conception-solution". The proposed cycle of Generation-Display-Evaluation-Adjustment interactive structure used in the proposed system follows the design process.

2. It is necessary to establish a Perceptual-Cognitive Descriptive Space (PCDS). The proposed Perceptual-Cognitive Descriptive Space (PCDS) consists of a set of paired adjectives.

3. The research on the matrix of product components is mainly based on ontology theory, which analyses the generation and physical structure of a product and its interrelationship.

4. Each product component corresponds to some perceptual image expression. The perceptual image of the Product Components Matrix (PCM) is expressed as the coordinate relationship of product components in the Perceptual-Cognitive Descriptive Space (PCDS).

5. As is shown in Figure 2, a fuzzy transformer method is proposed. Firstly, fuzzy logic, as a classical fuzzy modelling method is used to quantify the subjective evaluation results. In order to process the quantified data, a k-means-based transformer method is proposed to cluster the data and obtain the final result.

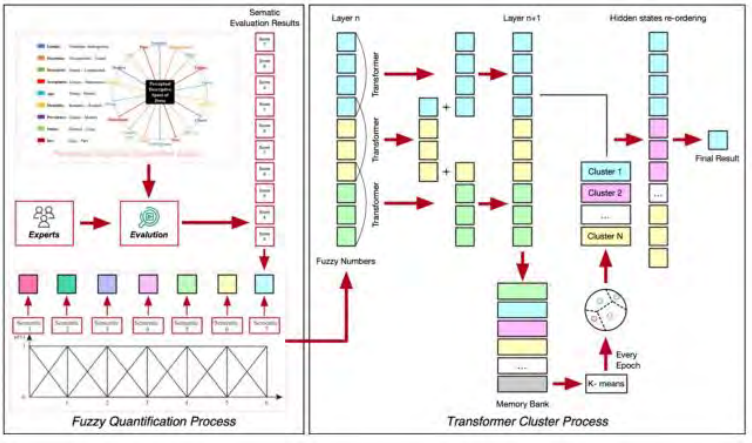



**Figure 2**: The working principle of the proposed Fuzzy Transformer method.

## 3. Case study and related experiments

### 3.1.1. Experiment I: Computational dress perceptual cognition modelling

Experiment I is designed to establish a Dress Design Knowledge Base (DDKB), namely the relationship between each dimension of the Dress Perceptual-Cognitive Descriptive Space (D-PCDS) and each alternative of the Dress Components Matrix (DCM).

### 3.1.2. Exploring the application of the computational dress perceptual cognition model: development of a personalized dress design system

Based on the computational dress perceptual cognition model established in this project, a personalized dress design system can be developed, as shown in Figure 3. The general working principle of the system is the "Design - Display - Evaluation - Adjustment" cycle. The cycle will be performed repeatedly by the user until a final satisfactory result is obtained.

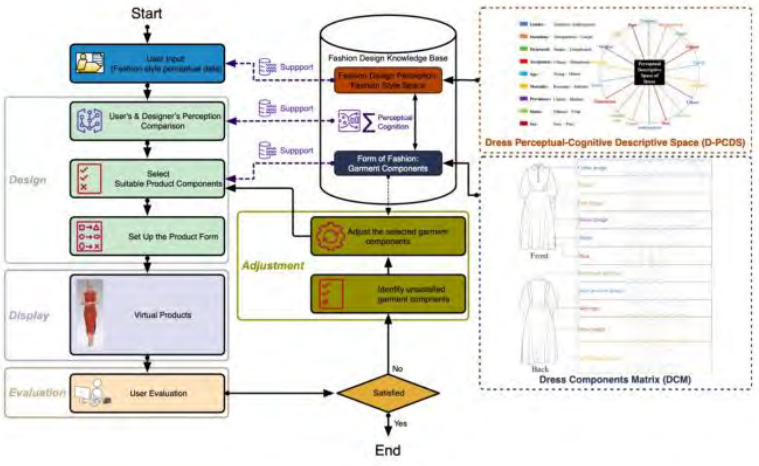



**Figure 3**: Working flowchart of the proposed personalized garment style design system based on the proposed perceptual cognition computational model (CMPCGSD).

### 3.1.3. System evaluation

In order to verify the effectiveness of the personalized dress design system proposed in this paper, a group of 60 female users aged 16-45 were involved in the validation experiment. The results show that the experimental data is effective.

## 4. Conclusion

This work aims at the extension of the application of the Metaverse to the design field. An interactive knowledge-based design system facing the Metaverse is developed. A fuzzy transformer method is innovatively developed to support this process. The proposed system greatly extends the application of the Metaverse and it can be applied to various product design system development.

## Acknowledgments

This research was funded by the National Natural Science Foundation of China (Grant Number: 61906129), the China Association for Science and Technology Y outh Support Talent Project (Grant Number: 2021-298), and the Hong Kong Polytechnic University GBA Startup Postdoc Programme 2022 (Grant Number: SP-22-13).

## References

[1] R. Cheng, N. Wu, S. Chen, and B. Han, “Will Metaverse be NextG Internet? Vision, Hype, and Reality,” IEEE Network, pp. 1-9, 2022, doi: 10.1109/MNET.117.2200055.
[2] M. Aloqaily, O. Bouachir, F. Karray, I. A. Ridhawi, and A. E. Saddik, “Integrating Digital Twin and Advanced Intelligent Technologies to Realize the Metaverse,” IEEE Consumer Electronics Magazine, pp. 18, 2022, doi: 10.1109/MCE.2022.3212570.
[3] B. Han, P. Pathak, S. Chen, and L. F. C. Y u, “CoMIC: A Collaborative Mobile Immersive Computing Infrastructure for Conducting Multi-user XR Research,” IEEE Network, pp. 1-9, 2022, doi: 10.1109/MNET.126.2200385.
[4] H. X. Qin, Y. Wang, and P. Hui, “Identity, Crimes, and Law Enforcement in the Metaverse,” arXiv preprint arXiv:2210.06134, 2210.06134, 2022.
[5] F. Y. Wang, R. Qin, X. Wang, and B. Hu, “MetaSocieties in Metaverse: MetaEconomics and MetaManagement for MetaEnterprises and MetaCities,” IEEE Transactions on Computational Social Systems, vol. 9, no. 1, pp. 2-7, 2022, doi: 10.1109/TCSS.2022.3145165.
[6] Y. Wang, J.-R. Chardonnet, F. Merienne, and J. Ovtcharova, “Using Fuzzy Logic to Involve Individual Differences for Predicting Cybersickness during VR Navigation,” in 2021 IEEE Virtual Reality and 3D User Interfaces (VR), 2021: IEEE, pp. 373-381.
[7] W. Yang et al, “Semantic Communications for Future Internet: Fundamentals, Applications, and Challenges,” IEEE Communications Surveys & Tutorials, pp. 1-1, 2022, doi: 10.1109/COMST.2022.3223224.
[8] Y. Wang, L.-H. Lee, T. Braud, and P. Hui, “Re-shaping Post-COVID19 Teaching and Learning: A Blueprint of Virtual-Physical Blended Classrooms in the Metaverse Era,” arXiv preprint arXiv:2203.09228, 2022.
[9] A. Musamih et al, “Metaverse in Healthcare:Applications, Challenges,and Future Directions,” IEEE Consumer Electronics Magazine, pp.1-13, 2022, doi: 10.1109/MCE.2022.3223522.
[10] X. Zhang, J. Wang, N. Cheng, and J. Xiao, “MetaSID: Singer Identification with Domain Adaptation for Metaverse,” in 2022 International Joint Conference on Neural Networks (IJCNN), 18-23 July 2022, pp. 1-7, doi: 10.1109/IJCNN55064.2022.9892793.
[11] P. Fettke, C. Houy, and P. Loos, “On the relevance of design knowledge for design-oriented business and information systems engineering,” Business & Information Systems Engineering, vol. 2, no. 6, pp. 347-358, 2010.
[12] “Leaf-nosed bat,” in Encyclopædia Britannica, ed: Encyclopædia Britannica Online, 2009.
[13] Y. Hong, X. Y. Zeng, Y. Y. Wang, P. Bruniaux, and Y.Chen, “CBCRS: An open case-based color recommendation system,” Knowledge-Based Systems, vol. 141, pp. Feb 2018, doi: 10.1016/j.knosys.2017.11.014.
[14] Y. Hong, X. Y. Zeng, P. Bruniaux, and K. X. Liu, “Interactive virtual tryon based three-dimensional garment block design for disabled people of scoliosis type,” Textile Research Journal, vol. 87, no. 10, pp. Jun 2017, doi: 10.1177/0040517516651105.
[15] L. Xue, Z. Y. Jin, H. Yan, and Z. J. Pan, “Development of novel fashion design knowledge base by integrating conflict rule processing mechanism and its application in personalized fashion recommendations,” Textile Research Journal, doi: 10.1177/00405175221129868.

# Data Marketplaces in the AI Economy

George Konstantinidis[1], Luis-Daniel Ibáñez[1] and Dumitru Roman[2]

[1]*University of Southampton, UK*
[2]*SINTEF, Norway*

Abstract

We envision a set of universal, trustworthy, transparent and user-friendly data market plugins (called UPCAST plugins) for the automation of data sharing and processing agreements between businesses, public administrations and citizens. Our foreseen plugins will enable actors in data spaces to design and deploy data exchange and trading operations guaranteeing (i) automatic negotiation of agreement terms, (ii) dynamic fair pricing, (ii) improved data-asset discovery, (iii) privacy, commercial and administrative confidentiality requirements, (iv) low environmental footprint, as well as ensuring compliance with (v) relevant legislation and (vi) ethical and responsibility guidelines. To achieve these we need to consolidate mature research in the areas of data management, privacy, monetisation, exchange and automated negotiation, considering efficiency for the environment as well as compliance with EU and international initiatives, AI regulations and ethical procedures.

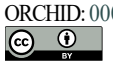
**Keywords**

Data Marketplaces, Data Privacy, Data Pricing, Data Sharing

## 1. Introduction

Digital technologies play a pivotal role in maximising the benefits of a data-driven society and a data-based economy but need to be solidly anchored to trustworthy, compliant, privacy-preserving and environmentally sustainable data sharing methods, architectures, and processes. These will enable citizens, businesses, and public administration/organisations to share and manipulate an ever-increasing amount of data, safely and efficiently. There is a worldwide strive to make our society empowered by data, where businesses and the public sector can make better and quicker decisions. Indeed, the European Strategy for Data[1] actively promotes the creation of common European data spaces providing a seamless common digital market of personal and commercial data to facilitate value creation and growth for businesses and organizations.

In this paper we advocate the need, and draw the map of steps, to design and deploy a set of universal plugins for data sharing, monetization and trading platforms that enable actors in common data spaces to collaboratively negotiate, improve and enforce data sharing contracts automatically (e.g. in the spirit of [1, 2, 3, 4] or [5, 6]), providing dynamic fair pricing mechanisms while implementing energy-efficient data exchange, ensuring privacy, confidentiality and legislation compliance and adhering to ethical and responsibility guidelines.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: g.konstantinidis@soton.ac.uk (G. Konstantinidis); l.d.ibanez@soton.ac.uk (L. Ibáñez); Dumitru.Roman@sintef.no (D. Roman)
ORCHID: 0000-0002-0877-7063 (G. Konstantinidis); 0000-0002-9421-8566 (L. Ibáñez); 0000-0001-7116-9338 (D. Roman)



[1] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066

To ensure businesses and administrations alike can maximise the returns from data driven innovation, we need to consider the following challenges:

- Data providers and consumers do not always understand the true value of data and are often unfamiliar with the requirements of data sharing and processing.
- Emerging data platforms and marketplaces lack the interoperability tools to exploit synergies and solve data tasks across users and platforms.
- There is an increasing complexity of drafting, negotiating and enforcing formal contractual agreements specific to privacy, regulation and other custom requirements of data sharing.

To overcome these, organisations need to be pro-active to ensure that the approach of valuing and promoting privacy, trust, and data sovereignty, as well as competition law, is promoted and facilitated. Data, tools and platforms that do not conform to these values are less likely to be used in business. Furthermore, different organisations may have different needs for each dimension of a data transaction, *e.g*, some organisations may need help pricing of their datasets, while others may need assistance discovering datasets relevant to their business processes. Our solution to these challenges is to develop a set of universal plugins for data marketplaces, each covering one core aspect of a data transaction. Plugins may be used independently, or in combination with one or more existing data marketplaces.

## 2. The UPCAST approach: Universal plugins for data markets

In the context of a data marketplace, we assume two types of organisations that enter data sharing or trading agreements: customers that want to create a data product or take advantage of a data driven opportunity that requires the design and execution of a data processing workflow, and providers that offer resources in the form of datasets or data operations, e.g., storage, cleaning, analytics, querying, integration, encapsulated as APIs or services [7]. In some data processing scenarios, an organisation can be both a customer and a provider. UPCAST's (see Fig. 1) aim is to provide end-to-end support to the following common scenarios:

**Scenario 1:** A customer organisation owns and controls some of the datasets that are required for the workflow it wants to implement. To complete the workflow, the customer needs to look

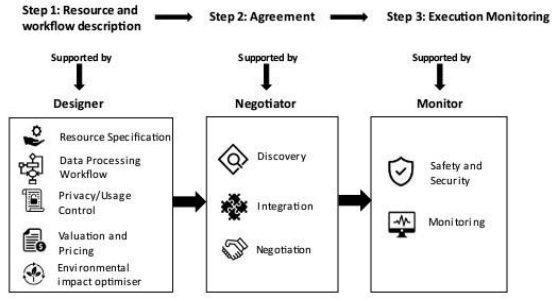



**Figure 1:** UPCAST Approach

in data marketplaces and find providers of further datasets or data operations. Providers on the other hand, enable the discovery of their resources within a marketplace. The customer discovers resources across marketplaces that match the operational, privacy, pricing, and environmental requirements of their desired workflow (see [5, 6, 8, 9] or [10, 1, 11, 12, 13, 14]). Once relevant resources are identified, customer and providers need to negotiate the usage and pricing conditions of each resource used in the workflow. All parties want to ensure that any privacy, legal, commercial, and administrative constraints are respected during the negotiation and subsequent execution of the workflow in a safe, secure and transparent way.

**Scenario 2:** A public administration entity owns or controls a set of datasets and would like to integrate them and enable their sharing or exchange in compliance with legal and administrative privacy and confidentiality constraints. They would also like to understand how much their data is worth in the context of a data marketplace, or, if they are to provide their datasets to an interested party, how much they are worth in the context of a particular workflow. The public administration entity is willing to exchange data for other data or services. All processes need to abide to legal, ethical, and environmental frameworks, and be auditable.

**Scenario 3:** A data marketplace provider wants to enable their data providers to trade data in compliance with legal and administrative privacy and confidentiality constraints. The marketplace also wants interoperability with other marketplaces to enable a wider data space. The marketplace would also like to understand synergies between datasets that they host order to suggest pricing functions and partnership opportunities to their clients.

With UPCAST plugins, customers and providers achieve their goals executing the following:

- Step 1 – Resource and Workflow Description: Providers and customers prepare descriptions of what they offer and what they need, respectively. Provider descriptions can be (i) Specification of technical characteristics including input/output description for data operations, dataset summaries and certificates of quality, environmental footprint or ethical data management, (ii) Usage conditions, that include constraints on privacy, confidentiality, data protection and pricing, among others. Customer descriptions sketch a high-level specification of the desired data processing workflow. They include the datasets and data operations they own, the requirements for datasets and data operations they need, and any additional conditions on the execution of the data processing workflow.
- Step 2 – Agreement: Based on a Data Processing Workflow specification, customer organisations indicate or discover relevant resources. Providers of these resources are invited to a negotiation where all parties reconcile any conflicting conditions to reach a contractual agreement. Resources are integrated and the entire workflow is optimised to satisfy privacy, pricing, environmental and other conditions.
- Step 3 – Monitoring execution: After an agreement is reached, the data processing workflow is executed in a decentralised fashion, in a secure, safe, transparent, and auditable way, via the processes in place in the underlying marketplaces.

UPCAST considers the role of third-party agents ('oracles'), such as human-based agents, to help meet the need for interfacing between legal language and code. The AI assessment enforces best-practices regarding the reliable, fair and transparent use and development of the automated techniques within the UPCAST framework.

## Acknowledgments

This work is partially funded by the projects UPCAST (HE 101093216) RAISE (HE 101058479), DataCloud (H2020 101016835), enRichMyData (HE 101070284), and Graph-Massivizer (HE 101093202).

## References

[1] R. Agrawal, J. Kiernan, R. Srikant, Y. Xu, Hippocratic databases, in: Proceedings of the 28th VLDB Conference, VLDB Endownment, United States, 2002, pp. 143–154.

[2] G. Konstantinidis, J. Holt, A. Chapman, Enabling personal consent in databases, Proc. VLDB Endow. 15 (2021) 375–387. URL: https://doi.org/10.14778/3489496.3489516. doi:10.14778/3489496.3489516.

[3] K. LeFevre, R. Agrawal, V. Ercegovac, R. Ramakrishnan, Y. Xu, D. DeWitt, Limiting disclosure in Hippocratic databases, in: Proceedings of the 30th International Conference on Very Large Databases, VLDB Endowment, 2004, pp. 108–119.

[4] R. Agrawal, P. Bird, T. Grandison, J. Kiernan, S. Logan, W. Rjaibi, Extending relational database systems to automatically enforce privacy policies, in: 21st International Confer- ence on Data Engineering (ICDE'05), IEEE, 2005, pp. 1013–1022.

[5] C. Debruyne, H. J. Pandit, D. Lewis, D. O'Sullivan, Towards generating policy-compliant datasets, in: 2019 IEEE 13th International Conference on Semantic Computing (ICSC), IEEE, 2019, pp. 199–203.

[6] C. Debruyne, H. J. Pandit, D. Lewis, D. O'Sullivan, "just-in-time" generation of datasets by considering structured representations of given consent for gdpr compliance, Knowledge and Information Systems 62 (2020) 3615–3640.

[7] A. Chapman, E. Simperl, L. Koesten, G. Konstantinidis, L.-D. Ibáñez, E. Kacprzak, P. Groth, Dataset search: a survey, The VLDB Journal 29 (2020) 251–272.

[8] Y. Wang, A. Kobsa, Respecting users' individual privacy constraints in web personalization, in: International Conference on User Modeling, Springer, 2007, pp. 157–166.

[9] P. Upadhyaya, M. Balazinska, D. Suciu, Automatic enforcement of data use policies with DataLawyer, in: Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, ACM, 2015, pp. 213–225.

[10] P. Ashley, C. Powers, M. Schunter, From privacy promises to privacy management: a new approach for enforcing privacy throughout an enterprise, in: Proceedings of the 2002 Workshop on New Security Paradigms, ACM, 2002, pp. 43–50.

[11] J. H. Kaufman, S. Edlund, D. A. Ford, C. Powers, The social contract core, in: Proceedings of the 11th International World Wide Web Conference (WWW), ACM, Honolulu, Hawaii, 2002, pp. 210–220.

[12] L. F. Cranor, Web privacy with P3P, O'Reilly Media, Inc., 2002.

[13] G. Karjoth, M. Schunter, M. Waidner, Platform for enterprise privacy practices: Privacy-enabled management of customer data, in: 2nd Workshop on Privacy-Enhancing Technologies. Lecture Notes in Computer Science, Springer, 2002, pp. 69–84.

[14] P. Ashley, S. Hada, G. Karjoth, M. Schunter, E-p3p privacy policies and privacy authorization, in: Proceedings of the 2002 ACM Workshop on Privacy in the Electronic Society, ACM, 2002, pp. 103–109.

# Understanding Sales Patterns using Unsupervised Machine Learning

Raavee Kadam [1], Pawan Lingras [1]

[1] *Saint Mary's University, 923 Robie St, Halifax, NS B3H 3C3, Canada*

**Abstract**

Point-of-sales data is very critical in retail to help businesses make evidence-driven decisions and formulate strategies to enhance their business operations and ensure customer satisfaction. Analysis of large and complex datasets gives better insights for decision-making, forecasting, and prediction in a retail environment. Timely identification of new trends helps a retail business align itself with the customers' expectations and changing market expectations. This study analyzes the sales patterns for different products within the portfolio of a large retail corporation and provides various insights for data-driven decision-making. Using sales data from 2015 to 2021, we use data summarization techniques and clustering algorithms to understand the sales patterns and analyze the impact of Covid-19 on sales patterns. The findings of this study are important for the organization to plan their product portfolio and offerings to keep up with the demands of the customers in the post-pandemic era.

**Keywords**

machine learning, sales forecasting, business intelligence, product planning

## 1. Introduction

In the retail industry, machine learning holds great promise for lowering product and service costs, lowering customer acquisition costs, enhancing business operations, and serving customers better given the large amount of point-of-sales data that is generated every single day. This study focuses on understanding the sales patterns of a large retail corporation in Canada. Given the amount of point-of-sales data that is generated daily from its stores, which are around 200, we are keen on understanding the performance of the company's stores and products. We use data summarization techniques and clustering algorithms to analyze around 7 years of data for the organization. Clustering is important in the retail business decision-making process since it enables making optimum retail decisions such as grouping together stores with similar customer bases, grouping stores based on both performance and non-performance parameters or choosing the best store locations with the goal of tailoring each location to better serve those customers. Clustering of stores will help the organization to understand the performance of various products within its portfolio and plan for effective demand and supply.

## 2. Literature Review

Unsupervised machine learning algorithms infer patterns from large and complex datasets when there are no output categories available on which the algorithm can try to model relationships. These algorithms try to use techniques on the input data to mine for rules, detect patterns, and summarize and group the data points, which help in deriving meaningful insights and describe the data better to the users [1]. Unsupervised learning is generally used to group patterns based on a similarity measure [2]. According to Ma and Sun [3], "in unsupervised learning tasks, the training dataset contains only

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: raavee.kadam@smu.ca (A. 1); pawan.lingras@smu.ca (A. 2)

the input variables, while the output variables are either undefined or unknown. The typical goal is to find hidden patterns in or extract information from the data" (p. 484). Clustering is a commonly used technique for this purpose. Clustering is an unsupervised data analysis technique used to identify hidden patterns in data, its properties and any outliers that might exist [4]. A conventional clustering technique uses the entire dataset and divides the dataset into meaningful groups. According to Wang et al. [5], "clustering is the process of grouping a collection of concrete or abstract object into multiple classes or clusters composed of similar objects" (p. 209). Popular unsupervised machine-learning algorithms include clustering, hierarchical clustering, k-means, principal component analysis, and association rules. K-means [2], a centroid-based clustering algorithm is frequently used for this purpose. The clusters identified by conventional techniques have crisp boundaries (hard clusters). However, in certain cases, more meaningful knowledge can be extracted by using overlapping boundaries (soft clusters). Rough clustering [6] and fuzzy clustering [7] have been proposed toward this end. Clustering that considers the time dimension of the data has also been used to identify dynamic changes to the groups [8]. Clustering that considers the time dimension of the data has also been used to identify dynamic changes to the groups [8]. The clustering of data streams provides some unique challenges related to speed, adaptability, scalability, and data variability [9]. Some of the clustering algorithms used for data streams are BIRCH [10], CluStream [11], D-Stream [12], DenStream [13], Online Divisive-Agglomerative Clustering (ODAC) [14] and DGClust [15]. Unsupervised machine learning techniques have been used within the retail environment to provide insights that help organizations serve their customers better. Yang and Nguyen [16] used point-of-sale (POS) data to explore sales performance using clustering, which enabled them to reveal the hidden structure of sales performance of retail stores. Similarly. Parikh and Abdelfattah [17] applied clustering algorithms on online transactions to provide strategies for customer purchasing behaviors and discover valuable customer groups.

## 3. Data

We used around 7 years of data from January 2015 to June 2022. Point-of-sales information was available for around 100 plus stores of the organization over the entire study period. This data consisted of stores that served around 50 of the forward sortation areas (FSA) within Canada. An FSA is a way to designate a geographical unit based on the first three characters in a Canadian postal code.

## 4. Analysis & Findings

We first commenced with data summarization to understand the sales information across all the stores of the organization. Table 1 below shows the annual sales information from 2015 to 2021. We have multiplied the actual sales figures with an arbitrary value of $x$ to denote the trend in sales and the total quantity of products sold.

**Table 1**
Annual Sales

| Year | Sales | Quantity |
|---|---|---|
| 2015 | \$30.25$x$ | 1.79$x$ |
| 2016 | \$30.51$x$ | 1.80$x$ |
| 2017 | \$30.97$x$ | 1.82$x$ |
| 2018 | \$31.64$x$ | 1.88$x$ |
| 2019 | \$32.92$x$ | 1.94$x$ |
| 2020 | \$36.29$x$ | 2.05$x$ |
| 2021 | \$36.69$x$ | 2.09$x$ |

We summarized data in several other ways as well to understand the sales for each day of the week, for each month of the year and so on. We then did a trend analysis to understand the trends in the sales of different products of the organization. There were 5 major categories of products within the company's portfolio, which we shall name Cat-A, Cat-B, Cat-C, Cat-D and Cat-E. Data was analyzed

to understand the sales trends for all five categories. The analysis revealed that Cat-A showed a sales pattern that was much different from the other products. We concluded that social distancing and a ban on social gatherings during the pandemic had a negative impact on the sales of Cat-A. We then began with clustering. First, we tried to find the exact number of clusters by plotting clustering errors against the number of clusters. We started with 5 clusters, which were not enough and then checked 10 clusters, which seemed too many. We settled at 7 clusters, which indicated an optimum value. Figure 1 below shows the cluster centroids for the five product categories over the 7 years.

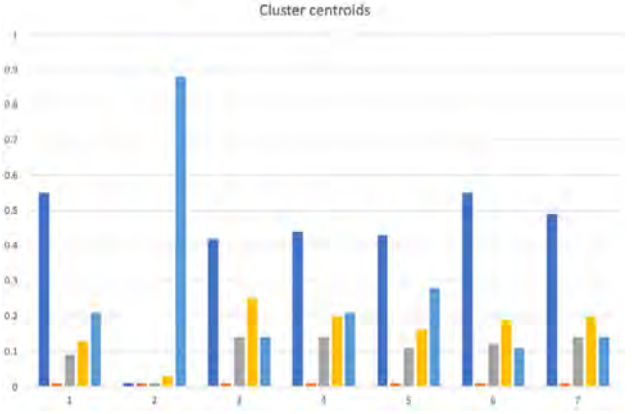



**Figure 1:** Cluster Centroids

We then identified the clustering of annual store patterns for the organization. The annual pattern for a store is treated as a separate object. Each store may have up to seven patterns for the years 2015-2021. Eight distinct groups were identified along with the transition of a store from cluster to cluster over seven years. Furthermore, we also identified the transition of an FSA from cluster to cluster and changes in the number of patterns in each cluster over the seven years. Clusters can shift; however, they continue to have the same profile. Product categories Cat-A, Cat-B, Cat-C, Cat-D and Cat-E have been used to label the clusters in order to identify the product profile of each cluster. Table 2 indicates the temporal shift in clusters i.e. the number of cluster memberships for stores in each year over the entire period.

**Table 2**
Cluster membership for stores

| Year | A41D22B20C17 | A42B30C15D12 | A44C24B17D14 | A46D21C19B12 | A46B21C17D15 | A42B30C15D12 | N | B60A19C08D06 | Grand Total |
|---|---|---|---|---|---|---|---|---|---|
| 2015 | | 18 | 24 | 2 | 31 | 28 | | 2 | 105 |
| 2016 | 2 | 18 | 24 | 4 | 30 | 26 | | 2 | 106 |
| 2017 | | 17 | 21 | 4 | 37 | 25 | | 1 | 105 |
| 2018 | 1 | 16 | 17 | 5 | 39 | 26 | | 2 | 106 |
| 2019 | 4 | 15 | 15 | 6 | 41 | 25 | 1 | 1 | 108 |
| 2020 | 44 | 9 | 4 | 40 | 8 | 1 | 1 | 1 | 108 |
| 2021 | 63 | 4 | | 34 | 4 | 1 | 1 | 1 | 108 |
| **Grand Total** | 114 | 97 | 105 | 95 | 190 | 132 | 3 | 10 | 746 |

## 5. Results & Discussion

On analyzing the company data for seven years, we identified seven optimal clusters as shown in Table 2. These clusters help us identify how the preference or affinity of customers for certain products of the organization has changed over the years, with the specific impact of the pandemic. From 2015, till we entered the pandemic in 2020, there was a significant change in the buying preferences of the customers, which can be seen by the shifting of clusters. Some clusters did not exist in 2015 but became dominant over the years, whereas some clusters that existed at the start of 2015 disappeared as we moved toward 2021. In 2015, there was nothing in the A41D22B20C17 cluster,

however, the same cluster became dominant in 2020 and 2021. Cluster A46D21C19B12 existed since 2015, however, it showed a substantial jump from 2020 onwards. On the other hand, cluster A44C24B17D14 existed from 2015 till 2019, however, it disappeared in 2020 and 2021. Similarly, clusters A42B30C15D12, A46B21C17D15 and A42B30C15D12 which existed consistently from 2015, reduced dramatically in 2020. So it can be observed that the shrinking clusters were A42B30C15D12, A A44C24B17D14, A46B21C17D15, and A42B30C15D12 whereas the expanding clusters were A41D22B20C17, A46D21C19B12. We had six clusters with relatively stable cluster sizes.

The Covid-19 pandemic transferred the retail environment with a huge shift in the buying preferences of consumers. Unsupervised machine learning algorithms are very useful in identifying such longitudinal changes in retail environments. One of the ways to approach clustering in the retail environment is to evaluate the sales performance by product categories of the organization. This approach helps us understand the variation in the demand for different products across the stores. Evaluating the sales performance by product categories can also help the organization understand who the target market for those sets of stores would be, and pricing and promotional strategies can be accordingly designed. In this research, clustering helped us understand the sales patterns of the company over the years and how the consumption of its different products changed during the Covid-19 pandemic. Such analysis can help the company incur several benefits in terms of inventory utilization, production and capacity planning, demand-supply management, etc. This in turn helps increase customer satisfaction due to the retailer's increased ability to provide a 'right mix' of products across its various stores. Clustering algorithms thus help the organization understand the demand for its various products and will help them make effective business decisions in the future as they move to a post-pandemic era.

## 6. References


[1] D. Fumo, "Types of Machine Learning Algorithms You Should Know," [Online]. Available: https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should-know-953a08248861.

[2] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A k-means clustering algorithm.," *Journal of the Royal Statistical Society. Series c (applied statistics),* vol. 28, no. 1, pp. 100-108.

[3] L. Ma and B. Sun, "Machine learning and AI in marketing – Connecting computing power to human insights," *International Journal of Research in Marketing,* vol. 37, no. 3, pp. 481-504.

[4] U. R. Hodeghatta and U. Nayak, Unsupervised Machine Learning. In: Business Analytics Using R - A Practical Approach, Berkeley, CA: Apress.

[5] H. Wang, J. Wang and Z. Zhong, "Research on Precision Marketing Strategy Based on Cluster Analysis Algorithm," 2020.

[6] P. Lingras and G. Peters, "Rough clustering," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, ,* vol. 1, no. 1, pp. 64-72, 2011.

[7] J. C. Bezdek, R. Ehrlich and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Computers & Geosciences,* vol. 10, no. 2-3, pp. 191-203, 1984.

[8] P. Lingras, M. Hugo, M. Snorek and C. West, "Temporal analysis of clusters of supermarket customers: conventional versus interval set approach," *Information Sciences,* vol. 172, no. 1-2, pp. 215-240, 2005.

[9] J. A. Silva, E. R. Faria, R. C. Barros, E. R. Hruschka, A. D. Carvalho and J. Gama, "Data stream clustering: A survey," *ACM Computing Surveys (CSUR),* vol. 46, no. 1, pp. 1-31, 2013.

[10] T. Zhang, R. Ramakrishnan and M. Livny, "BIRCH: A new data clustering algorithm and its applications," *Data Mining and Knowledge Discovery,* vol. 1, no. 2, pp. 141-182, 1997.

[11] C. C. Aggarwal, S. Y. Philip, J. Han and J. Wang, "A framework for clustering evolving data streams," in *Proceedings 2003 VLDB conference*, 2003.

[12] J. R. Chen, "Useful clustering outcomes from meaningful time series clustering," in *Proceedings of the sixth Australasian conference on Data mining and analytics - Volume 70*, 2007.

[13] F. Cao, M. Estert, W. Qian and A. Zhou, "Density-based clustering over an evolving data stream with noise," in *Proceedings of the 2006 SIAM International Conference on Data Mining*, 2006.

[14] P. P. Rodrigues, J. Gama and J. P. Pedroso, "ODAC: Hierarchical clustering of time series data streams," in *Proceedings of the 2006 SIAM International Conference on Data Mining*.

[15] J. Gama, P. P. Rodrigues and L. Lopes, "Clustering distributed sensor data streams using local processing and reduced communication," *Intelligent Data Analysis,* vol. 15, no. 1, pp. 3-28, 2011.

[16] C-L. Yang and T. P. Q. Nguyen, "Sequential Clustering and Classification Approach to Analyze Sales Performance of Retail Stores Based on Point-of-Sale Data", *International Journal of Information Technology & Decision Making,* vol. 21, no. 03, pp. 885-910, 2022.

[17] Y. Parikh and E. Abdelfattah, "Clustering Algorithms and RFM Analysis Performed on Retail Transactions," *2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, New York, NY, USA, 2020, pp. 0506-0511.

# Distributed Multi-agent Deep Reinforcement Learning for Traffic Engineering

Van An Le[1,*], Tien Thanh Le[2,1] and Yusheng Ji[1,2]

[1]National Institute of Informatics, Tokyo, Japan
[2]Department of Informatics, The Graduate University for Advanced Studies, SOKENDAI, Tokyo, Japan

Abstract
Recently, to address the explosion in network traffic and the dynamic of traffic behavior, reinforcement learning (RL) has become an essential technique in developing adaptive traffic engineering (TE) systems. In this paper, we propose a fully Distributed Multi-Agent deep reinforcement learning-based TE framework (DMATE). By carefully designing the local observation and reward function, we eliminate the necessity of centralized training, addressing the scalability issue and reducing communication overhead. We evaluate the proposed method on a real dataset and find that DMATE can achieve competitive performance to the centralized training approach while reducing communication overhead.

**Keywords**
Traffic engineering, multi-agent reinforcement learning, distributed network routing,

## 1. Introduction

Traffic engineering (TE) is concerned with reliable network operations and optimizing network resource utilization [1]. The explosion of traffic volume creates a strong urge to have efficient TE solutions that can adapt to the changes in network traffic. To this end, RL is believed the key technology for building autonomous network control systems [2]. Most of the proposed RL-based TE systems are centralized approaches (i.e., single-agent learning) which can take the advantage of global network view [3], [4], [5]. However, it faces several challenges: (1) the agent suffers from the scalability issue when the network size increases; (2) it requires high communication cost between network devices and the controller; (3) a centralized controller represents a single-point failure, which may degrade the network's performance.

In this paper, we propose DMATE, a fully Distributed and cooperative Multi-Agent reinforcement learning TE framework. In DMATE, each network node is an agent which independently chooses the routing path for the traffic flows. The agent makes decisions based on local information without sharing information with other agents. In our proposed approach, both the training and execution phases can be done in a distributed manner without causing any communication overhead. Each agent in DMATE will record the volumes of traffic flows that travel through the device (including the traffic flows sent from its own), which is called a partial traffic matrix and use it as the local observation for making the routing decisions. Our proposed

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
*Corresponding author.
EMAIL: anle@nii.ac.jp (V. A. Le); lethanh@nii.ac.jp (T. T. Le); kei@nii.ac.jp (Y. Ji)

method is trained using the MAPPO algorithm, which achieves strong performance in multiple popular multi-agent testbeds [6].

## 2. Network system

We target a well-known TE problem whose objective is to minimize the maximum link utilization. The network is represented by a directed graph $G = (V, E, c)$, where $V$ is the set of nodes ($|V| = n$), and $E$ is the set of the network's links. Each link $e \in E$ has capacity $c(e)$. A flow $d \in D$ is defined as the network traffic originating from a source node to a destination node in $G$. $m_d$ is the volume of traffic flow $d$, and $D$ is the set of all traffic flows. A path is defined as a sequence of nodes from the source to the destination. Let $P_d$ be a set of paths of flow $d$. These paths are pre-computed by the network operator and are the input for the TE optimization problem. In this paper, we consider a single-path routing problem in which there is one path in $P_d$ is chosen to route flow $d$. The objective of the problem is to minimize the maximum link utilization $u_{max}$, where $u_{max} = \max_{e \in E} u_e = \max_{e \in E} \frac{l(e)}{c(e)}$; $l(e)$ is the traffic load on link $e$.

## 3. The Distributed MARL-based TE: DMATE

DMATE includes multiple agents, in which each agent corresponds to one network node. Let $h$ denote the agent on node $h \in V$. Agent $h$ will decide the path for traffic flows in set $D_h \subset D$. $D_h$ is the set of all traffic flows originating from node $h$. The agent will periodically collect the local observation $\mathbf{s}_h^t$ and obtain the routing decisions $\mathbf{a}_h^t$. Then, the agent will receive the reward $r_h^t$ and collect the next observation $\mathbf{s}_h^{t+1}$. In the training mode, besides the actions, the agent will obtain the value of the current state ($v_h^t$). The information $(\mathbf{s}_h^t, \mathbf{a}_h^t, r_h^t, v_h^t, \mathbf{s}_h^{t+1})$ is stored in a local buffer and will be used for updating the agent's policy.

**Local observation**: We use both link utilization and traffic demands to be the local observation: $\mathbf{s}_h = [f_{d^*}, m_d, m_{d'}, u_e]$. We consider routing the top-k largest flows. $f_{d^*}$ are the destinations of the top-k largest flows. $m_d$ are the traffic demands of flows that originate or terminate at agent $h$. $m'_d$ are the traffic demand of the flows that have been routed through agent $h$ at time slot $t-1$.

**Agent's actions**: The action vector consists of the index of the pre-computed path for the top-k largest flows originating from the agent. Let $\mathbf{a}_h$ denote the action of agent $h$, then we have $\mathbf{a_h} = [a_d]_{d \in D_h^{topk}} \quad (a_d \in \{1, 2, ..., |P_d|\})$. $D_h^{topk}$ is the set of top-k largest flows originating from agent $h$. Other flows, which are not in $D_h^{topk}$, will be routed via the default paths.

**Reward function**: The reward function of each agent in DMATE is calculated using the maximum utilization among the link which are in the paths of flows $d \in D_h^{topk}$ of agent $h$: $r_h = -\max_{e \in p, p \in D_p} u_e$. Therefore, by maximizing the reward, the agents will minimize the maximum link utilization of the network.

## 4. Evaluation

We conduct an experiment to evaluate our approach with a real backbone network dataset (i.e., Geant network [7]). We use 200 traffic matrices for training and test with 500 traffic matrices. In this experiment, we compare the routing performance of our proposal, DMATE, with different TE algorithms by measuring the average $u_{max}$ of 500 traffic matrices in the test

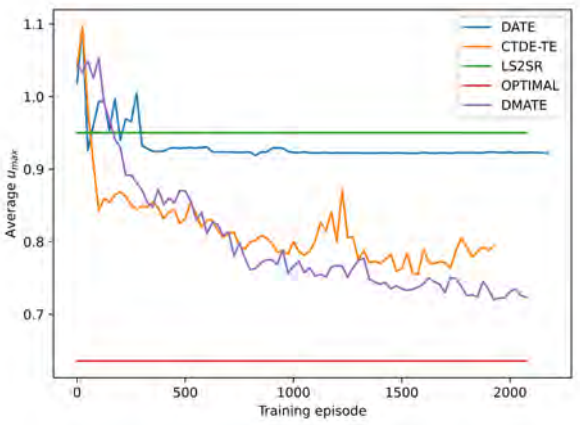



**Figure 1:** Performance comparison of different TE approaches.

set. The benchmarks include (1) **DATE**: the distributed and adaptive TE approach in [8]; (2) **CTDE-TE**: the RL-based TE approach with centralized training; (3) **LS2SR** [9]: this method obtains the routing rules by iteratively re-routing the large flows on high utilization links; (4) **OPTIMAL**: the routing decisions are obtained by solving the routing problem using a solver.

Fig.1 shows the average maximum link utilization of the network over the training phase. Our proposed method, DMATE, achieves better performance compared to DATE and LS2SR (reducing 21% and 23.6% in average $u_{max}$, respectively). Since DATE only uses link utilization as the local observation, it achieves faster training convergence but yields poor performance. LS2SR has a scalability issue when solving the routing problem of a large network under the time constraint. LS2SR solves the routing problem in 60 seconds while DMATE is less than 1 second). Compared to centralized approaches, DMATE has a similar performance to the CTDE-TE approach. With the global view, the agents in CTDE-TE seem to learn faster than DMATE in the first 500 episodes. However, CTDE-TE struggles in the rest of the training due to the large state space. In comparison with the optimal, DMATE achieves more than 80% of the optimal solution. Notes that the optimal method usually took a long time to find the solution, which is not practical in large networks with highly dynamic traffic changes [10]. Besides, DMATE does not cause communication overhead for sending the monitoring data to the centralized server.

## 5. Conclusion

This paper proposes a distributed traffic engineering framework, DMATE, that leverages the multi-agent deep reinforcement learning technique. We introduce the design of local observation, action, and reward function in order to achieve fully distributed system training without causing communication overhead. The experiment results show that DMATE can achieve better routing performance by minimizing the maximum link utilization of real network datasets. Next, we will improve the training convergence of the system by recovering the full traffic matrix based on the partial matrix. Therefore, the agent will have estimated traffic demand of flows that are not routed through it at the last time slot.

# References


[1] J. McManus, J. Malcolm, M. D. O'Dell, D. O. Awduche, J. Agogbua, Requirements for Traffic Engineering Over MPLS, RFC 2702, 1999. URL: https://rfc-editor.org/rfc/rfc2702.txt. doi:10.17487/RFC2702.

[2] T. Li, K. Zhu, N. C. Luong, D. Niyato, Q. Wu, Y. Zhang, B. Chen, Applications of multi-agent reinforcement learning in future internet: A comprehensive survey, IEEE Communications Surveys Tutorials 24 (2022) 1240–1279. doi:10.1109/COMST.2022.3160697.

[3] Z. Xu, J. Tang, J. Meng, W. Zhang, Y. Wang, C. H. Liu, D. Yang, Experience-driven networking: A deep reinforcement learning based approach, in: IEEE INFOCOM 2018 - IEEE Conference on Computer Communications, 2018, pp. 1871–1879. doi:10.1109/INFOCOM.2018.8485853.

[4] S. Troia, F. Sapienza, L. Varé, G. Maier, On deep reinforcement learning for traffic engineering in sd-wan, IEEE Journal on Selected Areas in Communications 39 (2021) 2198–2212. doi:10.1109/JSAC.2020.3041385.

[5] P. Almasan, S. Xiao, X. Cheng, X. Shi, P. Barlet-Ros, A. Cabellos-Aparicio, Enero: Efficient real-time wan routing optimization with deep reinforcement learning, Computer Networks 214 (2022) 109166. URL: https://www.sciencedirect.com/science/article/pii/S1389128622002717. doi:https://doi.org/10.1016/j.comnet.2022.109166.

[6] C. Yu, A. Velu, E. Vinitsky, Y. Wang, A. Bayen, Y. Wu, The surprising effectiveness of ppo in cooperative, multi-agent games, arXiv preprint arXiv:2103.01955 (2021).

[7] S. Orlowski, M. Pióro, A. Tomaszewski, R. Wessäly, SNDlib 1.0–Survivable Network Design Library, in: Proceedings of the 3rd International Network Optimization Conference (INOC 2007), Spa, Belgium, 2007. URL: http://www.zib.de/orlowski/Paper/OrlowskiPioroTomaszewskiWessaely2007-SNDlib-INOC.pdf.gz, http://sndlib.zib.de, extended version accepted in Networks, 2009.

[8] N. Geng, M. Xu, Y. Yang, C. Liu, J. Yang, Q. Li, S. Zhang, Distributed and adaptive traffic engineering with deep reinforcement learning, in: 2021 IEEE/ACM 29th International Symposium on Quality of Service (IWQOS), 2021, pp. 1–10. doi:10.1109/IWQOS52092.2021.9521303.

[9] V. A. Le, T. T. Le, P. L. Nguyen, H. T. T. Binh, Y. Ji, Multi-time-step segment routing based traffic engineering leveraging traffic prediction, in: 2021 IFIP/IEEE International Symposium on Integrated Network Management (IM), 2021, pp. 125–133.

[10] Z. Xu, F. Y. Yan, R. Singh, J. T. Chiu, A. M. Rush, M. Yu, Teal: Learning-accelerated optimization of traffic engineering, arXiv preprint arXiv:2210.13763 (2022).

# Machine learning for psychotherapy outcome predictions

Hans Jacob Westbye[1], MD, Andrew A. McAleavey[1,2,3], PhD and Christian Moltu[1,2], PhD

[1]District General Hospital of Førde, Førde, Norway
[2]Western Norway University of Applied Sciences, Bergen, Norway
[3]Weill Cornell Medicine, New York, United States

Abstract

Prioritizing the right patients and providing personalized treatment in a timely manner is crucial to improve access to healthcare. In psychotherapy at least 1 in 3 patients drop out of treatment with therapeutic alliance and patient motivation among the common predictors. Recommendations include strengthening the patient-therapist bond through developing common goals and checking in on progress and treatment path. Using a sample of 10,363 mental health from the USA, we used machine learning to develop a clinical feedback support tool to encourage patient-therapist goal alignment. A gradient-boosted decision tree was trained on pre-treatment patient-reported data to provide predictions of early treatment dropout, treatment duration, and symptom outcomes conditional on different treatment durations in out-of-sample patients. The models improved performance versus baseline predictions. The resulting decision support tool could assist in the collaborative selection of treatment goals, appropriate treatment intensity, and optimal allocation of resources. Results are discussed in the context of explainable AI and the ethical implications of predictive modeling in this context.

**Keywords**

Artificial Intelligence, Machine Learning, Psychotherapy, Outcome Prediction

## 1. Introduction

Psychological treatments almost always need to be personalized to individual patients, and therapists use their clinical experience to adjust their approach for a given patient. They may offer more intense, longer, or more emotionally challenging interventions depending on the patient's needs and characteristics. This leads to substantial variance in treatment approaches and outcomes, of which therapists are generally unaware: therapists regularly underestimate patients' negative experiences and fail to anticipate poor outcomes like dropout [1, 2]. This is unwanted on both an individual and systemic level. To better enable this treatment individualization, therapists increasingly use routine measurement using self-report questionnaires [3]. These self-report questionnaires are useful to track patient progress and have been shown to improve symptomatic outcomes of treatment when combined with feedback to therapists. Such patient reports generate large amounts of data that can be used for data-driven decision support tools, in order to empirically identify patients at risk of negative outcomes. In

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: hans.jacob.westbye@helse-forde.no (H. J. Westbye)


a study that compared 21 machine learning algorithms for predicting dropout from cognitive-behavioral therapy before the first session, the best-performing model had an AUC (Area Under the Receiver-Operator Characteristic Curve) of 0.6694 [4]. The modest accuracy of this prediction shows the difficulty of this problem. Nevertheless, even tools with limited accuracy may prove clinically valuable to therapists who can adjust treatment to reduce risk of a negative outcome. In the case of mental health, several predictions may be useful in clinical practice, including likelihood of dropping out of treatment, likelihood of completing a given treatment course, and symptomatic outcomes at the end of care. These are the focus of the present study.

# 2. Methods

## 2.1. Sample

Anonymized data was obtained from the digital measurement-based care instruments provider Mirah, Inc. The data was collected as part of routine practice in clinics using the Mirah software for clinical feedback and outcome monitoring in the United States. The sample included adult patients undergoing treatment between March 2016 and February 2022. Data were collected across multiple outpatient clinics and treatment providers resulting in a heterogeneous patient group seeking treatment for diverse mental health challenges.

## 2.2. Instruments

Norse Feedback (NF) is a clinical feedback system developed by the District General Hospital of Førde [5, 6, 7]. Items were generated from clinicians' and patients' needs and were tested and refined in clinical implementation studies [8]. NF consists of a maximum total of 88 items loading onto 18 dimensions. Patients respond to the items on a seven-point Likert scale. NF is a dynamic assessment that uses patient-adaptive computer logic to open and close dimensions depending on the patient reports on trigger items for each scale. Pragmatically, that means that some scales will not be present for some administrations, because they are determined to be irrelevant to the patient.

## 2.3. Description and preparation of data

The initial dataset consisted of $n$ = 11970 patients and initially $k$ = 2065 variables with high missingness. After dataset cleaning and variable selection, $k$ = 22 variables (18 NF dimensions, recent traumatic events (binary), in treatment at first assessment (binary), age and gender) for $n$ = 10363 patients were included for analyses. The dataset was randomly split by subject IDs into a training ($n$ = 8267) and test set ($n$ = 2096), withholding 20% of patients for the test dataset. The length of treatment was encoded as a continuous outcome using the number of unique assessments for each patient. The binary outcome "dropout" was encoded as 1 when the length of treatment was <=2 sessions. Training and test sets filtered for complete data were prepared to enable analysis with algorithms that do not handle missing data (train_nona: $n$ = 6937, test_nona: $n$ = 1752, $k$ = 20, two variables were excluded due to high missingness).

### 2.4. Machine learning algorithm

Variable selection reduced the overall missingness of data, but there were still many missing observations. Given this, we chose to use a machine learning algorithm that handled missing data directly. Gradient Boosted Decision Tree algorithms employ the gradient descent algorithm to minimize error in sequential models. The eXtreme Gradient Boosting (XGBoost) algorithm has proven to be a consistent top performer for many problems involving tabular data. Gradient boosted decision trees were among the best performing single models when comparing 21 machine learning models for dropout prediction [4].

### 2.5. Development and validation of prediction models

We defined three prediction tasks: 1) risk of dropout, 2) length of treatment (in number of sessions), and 3) outcomes given a completed treatment length. For task 1 we trained a generalized linear model (glm) on the train_nona dataset and two boosted decision tree models, one for the train_nona dataset (XGBoost_nona), and one for all observations (XGBoost_all). Results were compared to the baseline prediction of the probability of dropout using AUC. For task 2 we trained a glm model and two XGBoost models (glm, XGBoost_nona, and XGBoost_all) to predict the continuous outcome treatment length. Results were compared to the baseline prediction of mean treatment length using the root mean squared error (RMSE). For model 3 a total of 108 models were trained on all observations, one for each outcome (dimensions = 18) for each predicted treatment length (observations for treatment length 8/9 and 10/11 were binned for sufficient training data). For tasks 1 and 2 XGBoost hyperparameters were optimized through cross-validation using a grid search. For task 3 the same hyperparameters were used for all models. The performance of all models was tested in the out-of-sample test dataset. All data analyses were performed using R in the RStudio software for Windows.

## 3. Results

For predicting dropout the best performing model was XGBoost_all (AUC = 0.621) followed by XGBoost_nona (AUC = 0.601), and glm (AUC = 0.586) compared to baseline (AUC = 0.497). For predicting length of treatment the best performing model was XGBoost_all (RMSE = 5.910) followed by glm (RMSE = 6.169) and XGBoost_nona (RMSE = 6.211) compared to baseline (RMSE = 6.122). For predicting outcomes given a completed treatment length the XGBoost_all models performed significantly better than predicting the mean outcomes for a given treatment length (p-value = 0.001). The average RMSE of the 108 XGBoost_all models was 1.229 (95% CI 1.154-1.305) vs the average RMSE of the mean outcome prediction models being 1.456 (95% CI 1.343-1.568). An example of a clinical decision support tool using the results of this study has been developed and will be demonstrated at the SAIDD conference. Depending on the approval of the data owner this application can be made available online. The example decision support tool gives therapists an estimate of the risk of dropout, an estimated treatment length, and the corresponding estimated treatment outcome for each individual patient. This can facilitate in-treatment conversations to align treatment goals and expectations, and help trigger conversations about the motivation for treatment and risk factors for dropout.

## References


[1] J. Macdonald, J. Mellor-Clark, Correcting Psychotherapists' Blindsidedness: Formal Feedback as a Means of Overcoming the Natural Limitations of Therapists: Correcting Psychotherapists' Blindsidedness, Clinical Psychology & Psychotherapy 22 (2015) 249–257. URL: https://onlinelibrary.wiley.com/doi/10.1002/cpp.1887. doi:10.1002/cpp.1887.

[2] T. Probst, E. Humer, A. Jesser, C. Pieh, Attitudes of psychotherapists towards their own performance and the role of the social comparison group: The self-assessment bias in psychodynamic, humanistic, systemic, and behavioral therapists, Frontiers in Psychology 13 (2022) 966947. URL: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.966947/full. doi:10.3389/fpsyg.2022.966947.

[3] K. de Jong, J. M. Conijn, R. A. Gallagher, A. S. Reshetnikova, M. Heij, M. C. Lutz, Using progress feedback to improve outcomes and reduce drop-out, treatment duration, and deterioration: A multilevel meta-analysis, Clinical Psychology Review 85 (2021) 102002. URL: https://linkinghub.elsevier.com/retrieve/pii/S0272735821000453. doi:10.1016/j.cpr.2021.102002.

[4] B. Bennemann, B. Schwartz, J. Giesemann, W. Lutz, Predicting patients who will drop out of out-patient psychotherapy using machine learning algorithms, The British Journal of Psychiatry 220 (2022) 192–201. URL: https://www.cambridge.org/core/product/identifier/S0007125022000174/type/journal_article. doi:10.1192/bjp.2022.17.

[5] C. Moltu, M. Veseth, J. Stefansen, J. C. Nøtnes, Skjølberg, P.-E. Binder, L. G. Castonguay, S. S. Nordberg, This is what I need a clinical feedback system to do for me: A qualitative inquiry into therapists' and patients' perspectives, Psychotherapy Research 28 (2018) 250–263. URL: https://www.tandfonline.com/doi/full/10.1080/10503307.2016.1189619. doi:10.1080/10503307.2016.1189619.

[6] A. A. McAleavey, S. S. Nordberg, C. Moltu, Initial quantitative development of the Norse Feedback system: a novel clinical feedback system for routine mental healthcare, Qual Life Res (2021). URL: https://www.ncbi.nlm.nih.gov/pubmed/33851326https://link.springer.com/content/pdf/10.1007/s11136-021-02825-1.pdf. doi:10.1007/s11136-021-02825-1, edition: 2021/04/15.

[7] S. S. Nordberg, A. A. McAleavey, C. Moltu, Continuous quality improvement in measure development: Lessons from building a novel clinical feedback system, Qual Life Res (2021). URL: https://www.ncbi.nlm.nih.gov/pubmed/33591432https://link.springer.com/content/pdf/10.1007/s11136-021-02768-7.pdf. doi:10.1007/s11136-021-02768-7, edition: 2021/02/17.

[8] R. T. Hovland, C. Moltu, The challenges of making clinical feedback in psychotherapy benefit all users: A qualitative study, Nordic Psychology 72 (2020) 248–262. URL: https://www.tandfonline.com/doi/full/10.1080/19012276.2019.1684348. doi:10.1080/19012276.2019.1684348.

# Extreme and Sustainable Graph Processing for Green Finance Investment and Trading

Laurenţiu Vasiliu[1,*,†], Dumitru Roman[2,†] and Radu Prodan[3,†]

[1]*Peracton Ltd., Galway, Ireland*
[2]*SINTEF AS, Oslo, Norway*
[3]*University of Klagenfurt, Austria*

Abstract

The Graph-Massivizer project, funded by the Horizon Europe research and Innovation program, aims to create a high-performance and sustainable platform for extreme data processing. This paper focuses on one use case that addresses the limitations of financial market data. The project allows for the fast, semi-automated creation of realistic and affordable synthetic (extreme) financial datasets of any size for testing and improving AI-enhanced financial algorithms for green investment and trading. Synthetic data usage removes biases, ensures data affordability and completeness, consolidates financial algorithms and provides a statistically-relevant sample size for advanced back-testing.

**Keywords**

Graph processing, green finance, extreme data, synthetic financial data, data biases, serverless computing, sustainability

## 1. Introduction

The wide use, availability, accessible costs, interoperability, and analytical exploitation of financial data are essential for the European data strategy. For this, graphs are extreme data enablers that require further technological innovations to meet the needs of the European data economy. A study by IBM [1] revealed that the world generates nearly 2.5 quintillion bytes of financial data daily, posing extreme business analytics challenges. Graph-based technologies help pursue the United Nations Sustainable Development Goals by enabling better value chains, products, and services for green financial investments and deriving trustworthy insights for creating sustainable communities.

The improvement and optimization of green investments and trading face significant barriers. Historical securities' data, particularly on environmental, social, and governance data (starting from the early 2010s), is insufficient for in-depth testing, derisking financial algorithms, and

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
*Corresponding author.
†These authors contributed equally.
EMAIL: laurentiu.vasiliu@peracton.com (L. Vasiliu); dumitru.roman@sintef.no (D. Roman); radu.prodan@aau.at (R. Prodan)
ORCHID: 0000-0002-8247-5426 (R. Prodan)

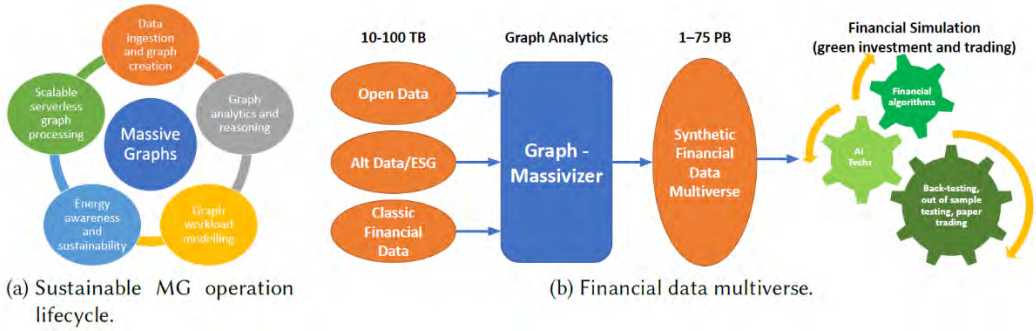



**Figure 1:** Green finance in Graph-Massivizer.

training AI models. Then, financial data is often difficult and expensive to access for training AI-driven financial algorithms. One historical record per security is commonly used to optimize a financial strategy, but this can lead to over-fitting and losses during live trading.

The *Graph-Massivizer project* [2] aims, among others, to remove the limitations of financial market data (limited volume, reduced accessibility, price barriers) by enabling fast, semi-automated creation of realistic and affordable synthetic extreme financial datasets, unlimited in size and accessibility. The extreme synthetic data goes one order of magnitude beyond the current big financial data features, aiming for PB in volume and affordable prices. The project researches and develops a high-performance, scalable, gender-neutral, secure, and sustainable platform based on the *massive graph (MG)* representation of extreme financial data. It delivers the *Graph-Massivizer toolkit* of five open-source software tools and FAIR graph datasets covering the sustainable lifecycle of processing extreme data as MG, displayed in Figure 1a.

## 2. Green Finance

Green finance targets financial products and services that direct investments into green-oriented enterprises. It aims for economic growth while reducing waste, pollution, and greenhouse gas emissions. Sustainable finance considers environmental, social, and governance factors for investment decisions for long-term sustainable economic activities.

**Financial Massive Graph (F-MG)** An F-MG is a hybrid, graph-based financial metadata structure (time series, values, boolean, monetary, securities taxonomies, statistical factors, rules). It helps research improved financial algorithms operating in five high-level steps: 1. historical financial data structure mapping into an F-MG; 2. synthetic data generation by preserving the original historic data statistical features; 3. missing data interpolation using ML inference and reasoning methods; 4. green financial investments and trading simulation, and 5. recommendation of the "greenest" investments and trading opportunities.

**Objectives** The green finance use case pursues two significant objectives on top of the Graph-Massivizer technology.

**Table 1**
Extreme financial data characteristics.

| *Data Characteristic* | *Big financial data state-of-the-art* | *Extreme financial data dimension* |
|---|---|---|
| *Volume* | Real financial data: 10 TB–100 TB | Synthetic streaming data: 1 PB–75 PB |
| *Value* | > € 200 000 per year | € 5000 – € 50 000 per year |
| *Veracity* | 20 %–50 % missing data | 100 % complete data |
| *Viridescence* | Unsustainable resource intensive analytics | Sustainable and energy-accountable graph analytics |

**Green financial data multiverse** Peracton Ltd. targets energy-efficient synthetic financial data generation (Figure1b) in a range size of 1 PB–75 PB, validated by standard green financial investment and trading algorithms. The developed technology promises a 90 % energy consumption accountability for extreme data creation streamed to end-users. Samples of the synthetic data will be available as open data for internal testing. The availability of significantly cheaper synthetic financial data for testing in extreme quantities allows more fintech companies, funds, and investors to test and derisk investment models.

**Greener financial algorithms and better investments** Peracton Ltd. aims to use the financial data multiverse for improved green AI-enhanced financial algorithms with reduced bias, risk, and higher performance while increasing the investment return by a realistic 2 %–4 %. It further targets an increase in excess return (alpha) by 1 %–2 % with a quick ratio higher than 1.5, reflecting healthy investments with lower risk and higher returns.

## 3. Summary and Outlook

Graph-Massivizer allows European green financial investors to avail of a massive financial data synthetic multiverse and a proven competitive, sustainable advantage. In comparison, other forms of analysis rely on present assumptions about *"what happened"* or *"what happens"*, correctly building and employing financial graphs to generate synthetic data can further reveal patterns suggesting what *"might happen"* with clear evidence for each connection or inference step. Graph processing facilitates problems solving driven by metrics related to costs or inefficiencies. The large-scale financial graph analytics market still traverses a developing phase, hampered by the lack of technology research and use case adoption. Graph-Massivizer provides for Europe these missing links.

## 4. Acknowledgments

The European Union funds this work through the Graph-Massivizer (Horizon Europe 101093202), enRichMyData (Horizon Europe 101070284), and DataCloud (Horizon 2020 101016835) projects.

## References

[1] A. Oriel, Accelerating the process of financial trading with big data analytics, https://www.analyticsinsight.net/accelerating-the-process-of-financial-trading-with-\big-data-analytics/, 2020.

[2] R. Prodan, *et al.*, Towards extreme and sustainable graph processing for urgent societal challenges in europe, in: 2022 IEEE Cloud Summit, IEEE, 2022, pp. 23–30.

# Lightning Talks

# Data Literacy for Citizenry

Inmaculada C. Haba-Ortuño [1], Fride H. Klykken [1] and Barbara Wasson [1]

[1] *Centre for the Science of Learning & Technology (SLATE), University of Bergen, Norway*

**Abstract**

We are in the era digitalization and datafication of society. The world is experiencing huge accelerating changes related to the abundance of data and appropriation of AI. However, is our current society been prepared for the arising possibilities and consequences of the digitalized and datafied life? In our international project we are focused on increasing the data and AI literacy for citizenry with an aim to empower adults to maneuver their way in this new reality.

**Keywords**

Data literacy, AI literacy, citizenry, playful approach

## 1. Introduction

Global society has been and is experiencing the digitalization. Digitalization leads to the generation of a lot of new types of data (e.g., video, sensor, voice), something which is often referred to as datafication of society [1]. Moreover, the implications of data and its use through algorithms in artificial intelligence (AI) are closely tied to data ethics and privacy rights, as concerns of surveillance capitalism ([2], [3]) are raised.

As a result of this status quo, researchers from European universities were funded through an Erasmus + partnership, Data Literacy for Citizens (DALI), which aims to empower adult citizens though fostering their competences related to data (data literacy).

## 2. DALI project: Data Literacy for Citizenship

DALI, with partners from Norway, Germany, United Kingdom and Spain, is researching how to provide different target groups of adults (from young adults until seniors) competences to be prepared for a datafied daily life. Adults are not only in charge of their own data but also of future generations, thus, to empower adults with literacy in personal data and AI has an enormous impact for citizenry.

The first steps were to map data and algorithmic literacy through a state of the field study using mapping review as methodology. The partnership also develop a data literacy framework for adults [4] through a Delphi study. Through 3 rounds and a half, 10 experts from 5 different institutions were involved in individual and synchronous collective phases.

Another output of the project is a repository of existing data literacy resources and activities [5], using Ronen and Shenkar's clustering [6] in order to attempt to search in equal conditions around the world and cultures. Following this we researched pedagogical strategies, combining different types of learning, such as Networked Learning [7], Playful Learning [8], [9] and Activity-Centred Analysis and Design [10], resulting in 10 Game-based Networked Learning Strategies [11]. The pedagogical approaches and co-creation strategies are the basis of 40 DALI games, which help citizens develop their data and AI literacy skills. The DALI playful Learning Toolkit, whose educational resources aims to transfer research results from academia to society in terms of data literacy.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: inmaculada.ortuno@uib.no (A.1); fride.klykken@uib.no (A.2); barbara.wasson@uib.no (A.3)
ORCID: 0000-0002-0712-8619 (A.1); 0000-0003-4070-3782 (A.2); 0000-0003-4897-1394 (A.3)


The growing impact of AI on all sectors and in daily live requires an upskilling of citizens so that they can address societal challenges, but also dealing with the societal challenges of data and AI. In this lightening talk the DALI framework and a selection of games will be presented.

## 3. Acknowledgements

This research has been supported by an ERASMUS+ Strategic Partnership (KA204-076492), funded by the European Union, specifically, the project "DALI - Data Literacy for Citizenship". Besides, thanks to all DALI partners who have been contributing to the project.

## 4. References

[1] K. Leurs, T. Shepherd, Datafication & Discrimination, in: M. T. Schäfer, K. van Es (Eds.), The Datafied society: Studying culture through data, 1st. ed, Amsterdam Press, Amsterdam, 2017, pp. 211-234.

[2] S. Zuboff, The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power, 1st. ed., PublicAffairs, New York, 2019.

[3] M. Landwehr, A. Borning, V. Wulf, The High Cost of Free Services: Problems with Surveillance Capitalism and Possible Alternatives for IT Infrastructure, in: Proceedings of the Fifth Workshop on Computing within Limits, LIMITS '19, Association for Computing Machinery, Lappeenranta, Finland, 2019, pp. 1-10, doi:10.1145/3338103.3338106

[4] V. Marín, I. Haba-Ortuño, G. Tur, D. Villar-Onrubia, Conceptualizando la alfabetización en datos: El marco DALI, in: O. L. Agudelo, B. de Benito, A. Darder, J. Moreno, J. Munar, F. Negre, A. Pérez, J. Salinas, G. Tur, S. Urbina (Eds.), Educación transformadora en un mundo digital: Conectando paisajes de aprendizaje, EDUTEC 2022, EDUTEC Asociación para el Desarrollo de la Tecnología Educativa, Palma de Mallorca, Spain, 2022, pp. 43-45.

[5] I. Haba-Ortuño, L. Castañeda, Actividades y recursos de enseñanza online para la alfabetización de datos, in: A. San Martín-Alonso, I. M. Gallardo-Fernández, J. Peirats-Chacón (Eds.), Libro de Actas. XXIX Jornadas Internacionales Universitarias de Tecnología Educativa: Aprender entre imágenes y pantallas, JUTE 2022, Red Universitaria de Tecnología Educativa, Valencia, Spain, 2022, pp. 817-824.

[6] S. Ronen, O. Shenkar, Mapping world cultures: Cluster formation, sources and implications, Journal of International Business Studies (2013) 867 – 897, https://doi.org/10.1057/jibs.2013.42

[7] Networked Learning Editorial Collective (NLEC), Networked Learning: Inviting Redefinition, Postdigital Science and Education (2020) 312-325. https://doi.org/10.1007/s42438-020-00167-8

[8] N. Whitton, Playful learning: Tools, techniques, and tactics, Research in Learning Technology (2018). doi: 10.25304/rlt.v26.2035

[9] S. Arnab, Game Science in Hybrid Learning Spaces, 1st. ed., Routledge, Oxfordshire, 2020.

[10] P. Goodyear, L. Carvalho, P. Yeoman, Activity-Centred Analysis and Design (ACAD): Core purposes, distinctive qualities and current developments, Educational Technology Research and Development (2021) 445–464. doi:10.1007/s11423-020-09926-7

[11] L. Castañeda, D. Villar-Onrubia, I. Haba-Ortuno, A. Y. Postigo-Fuentes, S. Arnab, Game-based Networked Learning, in: J. Jaldemark, M. H. Lindqvist, P. Mozelius, A. Öberg, M. De Laat, N. B. Dohn, T. Ryberg (Eds.), Proceedings for the Thirteenth International Conference on Networked Learning, NLC 2022, Networked Learning Consortium, Sundsvall, Sweden, 2022, pp. 273-277.

# Digital literacies of youth in NEET-situations: narratives from public and non-public service providers

Gilda Seddighi [1] and Hilde G. Corneliussen [2]

[1] NORCE Norwegian Research Centre AS, Nygårdsgaten 112, Bergen, Norway
[2] Western Norway Research Institute, Røyrgata 4, Sogndal, Norway

**Abstract**

Despite strong user-perspectives in digitalization of public welfare services, citizens who struggle with social, cognitive, or health-related challenges may be at risk of exclusion from these digital solutions. Youth, who are not in education, employment, or training (NEET) are particularly vulnerable to being excluded from digitalized public welfare services, in spite of the common assumption that young people, known as "digital natives," intuitively know how to use and develop digital skills.

This paper presents the results of the "Mapping the Need for Digital Literacies for Leaving NEET-Situations" project, which investigated how public and non-public service providers for NEET-youth perceive and work to improve digital literacies among these young people.

The results suggest that service providers do not have a routine in place to map and systematically improve the digital literacy of NEET youths. Instead, they tend to randomly map the digital literacies of these youths when they first have challenges accessing and benefiting from digitalized public services.

The study suggests that: 1) the concept digital literacy needs to be redefined in relation to NEET users and situations; 2) a tool is necessary to support service providers in mapping the digital literacies of NEET users; 3) a routine in the Norwegian Labour and Welfare Administration for mapping the digital literacies of NEETs is required; 4) an organization at the national level should support local service providers in improving the digital literacies of NEETs; and 5) an organization at the national level should facilitate local service providers' efforts to support NEETs' use of digitalized public services.

## References

1. Martin, A. and J. Grudziecki, *DigEuLit: Concepts and tools for digital literacy development. Innovation in Teaching and Learning in Information and Computer Sciences, 5 (4), 249–267*. 2006.
2. Helsper, E., *The digital disconnect: The social causes and consequences of digital inequalities.* The Digital Disconnect, 2021: p. 1-232.
3. Prensky, M., *Digital Natives, Digital Immigrants.* On the Horizon, 2001. **9**(5): p. 1-6.
4. Evans, C. and W. Robertson, *The four phases of the digital natives debate.* Human Behavior and Emerging Technologies, 2020. **2**(3): p. 269-277.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: gsed@norceresearch.no (A. 1); hgc@vestforsk.no (A. 2)
ORCID: https://orcid.org/0000-0003-1601-9286 (A. 1); https://orcid.org/0000-0002-9661-9638 (A. 2)

5. Kim, E. and S. Yang, *Internet literacy and digital natives' civic engagement: internet skill literacy or internet information literacy?* . Journal of Youth Studies, 2016. **19**(4): p. 438–456.
6. Balea, B., *Digital natives or not? How do Romanian adolescents cross the boundaries of internet common use.* Studia Universitatis Babes-Bolyai Sociologia 2016. **61**(1): p. 59–76.
7. Livingstone, S., et al., *Adolescents experiencing internetrelated mental health difficulties: the benefits and risks of digital skills*. 2022, KU Leuven ySKILLS GA 870612.
8. Buchert, U., et al., *Is digitalisation of public health and social welfare services reinforcing social exclusion? The case of Russian-speaking older migrants in Finland.* Critical Social Policy, 2022: p. 02610183221105035.
9. Midtgård, T.M., et al., *Digital ekskludering i NAV.* 2022.
10. Mascherini, M. and S. Ledermaier, *Exploring the diversity of NEETs.* 2016.
11. OECD, *Investing in Youth: Norway*. 2018.
12. Fyhn, T., R.L. Radlick, and V. Sveinsdottir, *Unge som står utenfor arbeid, opplæring og utdanning (NEET).* En analyse av unge i NEET-kategorien, 2021.

# Future Challenges of Question Answering

Pierre Maret[1,2,*], Dennis Diefenbach[2]

[1]Université Jean Monnet Saint-Etienne, CNRS (Laboratoire Hubert Curien UMR 5516) F-42023, SAINT-ETIENNE, France
[2]The QA Company, 20 rue Benoit Lauras, F-42000, SAINT-ETIENNE, France

Question Answering (QA) services consist of finding direct answers to questions sent by users against a corpus [1]. It generally uses diverse combinations of Deep Learning algorithms, NLP and Semantic Web techniques. The corpora that are targeted may be public or private, and consist of data (generally in the form of knowledge graphs [2]), documents [3], or they combine data and documents.

The challenges for the next generation of QA systems are heterogeneous. They have been impacted by the recently published ChatGPT and Bard systems and services. The challenges we have identified are listed hereafter, and one can imagine that they will be still evolving in the near future. They relate to engineering aspects and the methodologies, and to research tasks.

- Support of heterogeneous data formats
- Support of multiple information sources [4]
- Support of evolving information [5]
- Reduction of the cost of the information design
- Facility to produce the training data
- Reduction of the amount of training data required [6]
- Ability to explain of the returned answers [7]
- Conversational interface
- Multilingualism
- Sobriety, reduction of the environmental impact [8]
- Complementary and bridges between techniques.

Chatbots and Elasticsearch-based systems are also used to return information to users for answering to their information needs as QA systems do. However, the Chatbot technique requires the elicitation of all the intents (with their associated answers) the users could formulate. So far, this technique cannot be considered scaling. Elasticsearch-based systems use fast indexing techniques and can therefore scale with large corpora. But these systems don't consider relations between data in their implementation. Therefore, they don't benefit from the semantics incorporated into structured data or into the text sentences, to select the answers returned to users.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
*Corresponding author.
EMAIL: pierre.maret@univ-st-etienne.fr (P. Maret); dennis.diefenbach@the-qa-company.com (D. Diefenbach)
ORCHID: 8 0000-0001-5437-7725 (P. Maret); 0000-0002-0046-2219 (D. Diefenbach)


## References


[1] R. S. Roy, A. Anand, Question Answering for the Curated Web: Tasks and Methods in QA over Knowledge Bases and Text Collections, Synthesis Lectures on Information Concepts, Retrieval, and Services, Morgan & Claypool Publishers, 2021. URL: https://doi.org/10.2200/S0113ED1V01Y202109ICR076. doi:10.2200/S0113ED1V01Y202109ICR076.

[2] D. Diefenbach, A. Both, K. Singh, P. Maret, Towards a question answering system over the semantic web, Semantic Web 11 (2020) 421–439. URL: https://doi.org/10.3233/SW-190343. doi:10.3233/SW-190343.

[3] D. Chen, A. Fisch, J. Weston, A. Bordes, Reading Wikipedia to answer open-domain questions, in: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Vancouver, Canada, 2017, pp. 1870–1879. URL: https://aclanthology.org/P17-1171. doi:10.18653/v1/ P17-1171.

[4] S. Shekarpour, A.-C. Ngonga Ngomo, S. Auer, Question answering on interlinked data, in: Proceedings of the 22nd international conference on World Wide Web, 2013, pp. 1145–1156.

[5] P. Plessers, O. De Troyer, S. Casteleyn, Understanding ontology evolution: A change detection approach, Journal of Web Semantics 5 (2007) 39–49. URL: https://www.sciencedirect.com/science/article/pii/S1570826806000540. doi:https://doi.org/10.1016/j.websem.2006.11.001, selected Papers from the International Semantic Web Conference.

[6] Y. Wang, Q. Yao, J. T. Kwok, L. M. Ni, Generalizing from a few examples: A survey on few-shot learning, ACM Comput. Surv. 53 (2020). URL: https://doi.org/10.1145/3386252. doi:10.1145/3386252.

[7] B. H. van der Velden, H. J. Kuijf, K. G. Gilhuijs, M. A. Viergever, Explainable artificial intelligence (xai) in deep learning-based medical image analysis, Medical Image Analysis 79 (2022) 102470. URL: https://www.sciencedirect.com/science/article/pii/S1361841522001177. doi:https://doi.org/10.1016/j.media.2022.102470.

[8] S. Cerf, R. Bleuse, V. Reis, S. Perarnau, É. Rutten, Sustaining performance while reducing energy consumption: A control theory approach, in: L. Sousa, N. Roma, P. Tomás (Eds.), Euro-Par 2021: Parallel Processing - 27th International Conference on Parallel and Distributed Computing, Lisbon, Portugal, September 1-3, 2021, Proceedings, vol- ume 12820 of *Lecture Notes in Computer Science*, Springer, 2021, pp. 334–349. URL: https://doi.org/10.1007/978-3-030-85665-6_21. doi:10.1007/978-3-030-85665-6_21.

# Predictive Analytics for Product Consumption using Customer Demographics

Raavee Kadam [1], Pawan Lingras [1]

[1] *Saint Mary's University, 923 Robie St, Halifax, NS B3H 3C3, Canada*

**Abstract**

Demand and sales forecasting are of vital importance in a retail environment which enables organizations to envision their future sales revenue and allocate necessary resources according to the forecasts. With predictive analytics, organizations can use customer data generated from various sources to prospect for customer needs, find consumption patterns, understand customer behavior, etc. Industry trends can be analyzed to find actionable insights that can help the organization serve its customers better. Using predictive analytics, this study aims to predict the consumption patterns for various products of a large retail corporation using customer demographic data. The findings of this study are important for the organization to plan their product portfolio and offerings to keep up with the demands of the customers and serve their customers better.

**Keywords 1**

Machine learning, sales forecasting, demand & supply, predictive analytics

## 1. Introduction

In the retail industry, machine learning holds great promise for predicting the sales of products or services based on historical data. Machine learning algorithms can help analyze data from historical purchases, customer interactions, demographics, point of sales data, etc. to get an accurate sales forecast and predict which customer segment is most likely to buy your products or services. This can help the organization to plan product supply, provide targeted discounts, upsell or cross-sell and provide differentiated pricing to its customers. This study focuses on predicting the sales of various products of a large retail corporation in Canada. Given the amount of point-of-sales data that is generated daily from its stores, we are keen on understanding how the demography of the place will impact the sales of products of the organization. Using supervised machine learning and demographic information from the Canadian Census study, we predict the sales of various products, which we will name as Cat-A, Cat-B, Cat-C, Cat-D and Cat-E within the portfolio of the organization.

## 2. Literature Review

Supervised learning is generally used to classify an event, based on its features, into predetermined subsets [1]. Classification is the most widely used machine learning technique and finds application in areas such as spam filtering, loan approval and image analysis. The classification algorithm, called a classifier, is trained using historical data. This trained classifier can then be used to predict the behavior of a new instance. Decision tree [2], random forest [2], neural networks [3] and support vector machines (SVM) [4] are some of the established techniques used to develop a classifier. The conventional decision tree algorithm has been further enhanced to determine decision rules in cases of uncertain data, such as web usage ([5], [6]). Similarly, there have been efforts to improve the applicability of SVM to deal with datasets with soft boundaries and for multiclass classification [7]. Furthermore, traditional feature extraction uses hand-picked features that conform to physical models. These techniques are effective, but they require domain knowledge and usually only apply to one

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: raavee.kadam@smu.ca (A. 1); pawan.lingras@smu.ca (A. 2)

specific use-case. Unsupervised learning techniques, such as clustering, can be used to reveal latent relationships between objects within a dataset. Intuitively, membership to a cluster can be thought of as a feature of a data object, and clustering as a form of feature extraction. It is not surprising then that clustering has shown to be effective when used in conjunction with supervised learning techniques. Learning techniques when combined with socioeconomic variables can provided a variety of insights in the retail environment, and typically have a lesser acquisition cost as compared to psychological variables. Islam et al. [8] used socio-demographic variable to investigate the predictive accuracy of seven machine learning procedures for six grocery product categories. Their results indicated a roughly 20%–30% improvement in out-of-sample predictive performance. Similarly, Matuszelański and Kopczewska [9] used socio-geo-demographic data from the census to test a comprehensive customer churn model within the retail environment using machine learning methods. Their results listed the various factors such as demographic environment of the customer, customer location, etc. that influenced the customers' ability to churn.

## 3. Data

We used around 7 years of data from the company from January 2015 to June 2022. Point-of-sales information was available for around 100 stores of the organization over the entire study period. This data consisted of stores that served around 50 of the forward sortation areas (FSA) within Canada. An FSA is a way to designate a geographical unit based on the first three characters in a Canadian postal code. For customer demographic information, we used Statistics Canada data from the 2016 Census for region where the corporation is located. The census data contained different information for the population related to age, gender, income, education, employment, household data, ethnicity, etc. The census data was available FSA-wise, similar to the data from the corporation.

## 4. Analysis and Findings

Machine learning algorithms can help organizations forecast their sales for different products within their portfolio using large datasets. In sales forecasting, supervised learning algorithms can help businesses predict how consumers will behave based on their demographic information. This study used supervised learning to predict the sales of various products within the organization's portfolio using the demographic characteristics of the population. The products in the organization's portfolio included various products that we put in five categories, namely Cat-A, Cat-B, Cat-C, Cat-D and Cat-E. We aim to predict the consumption of two of its products, namely, Cat-A and Cat-B. We used 26 demographic variables to predict consumption and the quantity/percentage share. Since demographic data was available at the FSA level, predictions are done by FSA. Table 1 shows the percentile errors at 50% (median). In Table 1, 'percCatA' indicates the sales of Cat-A as a percentage of the total sales, whereas 'percCatB' indicates the sales of Cat-B as a percentage of the total sales. Table 1 shows high accuracy with median errors ranging from 1% to 4% for the percentage share. Predicting the proportional share of Cat-A was very accurate, however, the 15-19 dataset had the best accuracy, whereas the 20-21 dataset was a little better than the entire seven-year dataset.

**Table 1**
Percentile Errors

| Dataset | Variable | Median |
|---|---|---|
| 15to21 | Cat-A | 4.55 |
| 15to21 | percCatA | 3.73 |
| 15to21 | Cat-B | 3.77 |
| 15to21 | percCatB | 4.29 |
| 15to19 | Cat-A | 3.49 |
| 15to19 | percCatA | 1.33 |
| 15to19 | Cat-B | 2.21 |

| | | |
|---|---|---|
| 15to19 | percCatB | 2.62 |
| 20to21 | Cat-A | 6.66 |
| 20to21 | percCatA | 1.57 |
| 20to21 | Cat-B | 10.42 |
| 20to21 | percCatB | 3.17 |

## 5. Results & Discussion

We used Statistics Canada census data to get the demographics of the region in which the organization is retiling, and then predicted the consumption of the two prominent products of the organization based on that demographic data. Customer segmentation i.e. classifying consumers into manageable groups, based on demographic characteristics is a traditional retail activity used to identify the right mix of products, place, pricing and promotion for each segment. Demographics can help you identify and understand your target audience. This also helps the organization identify which customer segments it wants to cater to and focus on, with targeted sales and marketing strategies. Along similar lines, we wanted to understand how demographic characteristics are impacting the purchasing preferences of the customers of the retail organization. Table 2 indicates the detailed list of the demographic variables identified and used for prediction (ranked as per their importance).

**Table 2**
Key Demographic Variables

| Importance | Variable description |
|---|---|
| 205 | Average number of rooms per dwelling |
| 194 | Total - Tenant households in non-farm, non-reserve private dwellings - 25% sample data |
| 137 | Average household size |
| 132 | Total - Owner households in non-farm, non-reserve private dwellings - 25% sample data |
| 105 | Total private dwellings |
| 102 | Total - Immigrant status and period of immigration for the population in private households - 25% sample data |
| 100 | Total - Citizenship for the population in private households - 25% sample data |
| 96 | Total - Ethnic origin for the population in private households - 25% sample data |
| 86 | Number of persons in private households |
| 85 | Total - Age groups and average age of the population - 100% data; Both sexes |
| 81 | Total - Total income groups in 2015 for the population aged 15 years and over in private households - 100% data |
| 81 | Total - Income statistics in 2015 for the population aged 15 years and over in private households - 100% data |
| 81 | Population, 2016 |
| 80 | Total - Employment income groups in 2015 for the population aged 15 years and over in private households - 100% data |
| 73 | Total - Owner and tenant households with household total income greater than zero, in non-farm, non-reserve private dwellings by shelter-cost-to-income ratio - 25% sample data |
| 67 | Total - Occupied private dwellings by number of bedrooms - 25% sample data |
| 66 | Total - Private households by housing suitability - 25% sample data |
| 64 | Total - Household total income groups in 2015 for private households - 100% data |
| 64 | Total - Private households by tenure - 25% sample data |
| 63 | Total - Occupied private dwellings by number of rooms - 25% sample data |
| 63 | Total - Occupied private dwellings by period of construction - 25% sample data |
| 63 | Total - Occupied private dwellings by structural type of dwelling - 100% data |
| 63 | Total - Private households by number of persons per room - 25% sample data |
| 62 | Total - Private households by household size - 100% data |
| 60 | Total - Occupied private dwellings by condominium status - 25% sample data |
| 0 | Population percentage change, 2011 to 2016 |
| 0 | Total - Distribution (%) of the population by broad age groups - 100% data; Both sexes |

As we can see, the number of dwellings, proportion of tenant/owned households, income, population, number of people per dwelling, immigrant/citizens, ethnicity, and age were the prominent variables that had an impact on the consumption of the two products. Higher-income households with a higher social status can have a preference for certain kinds of products as opposed to lower-income products. The same applies to age groups where baby-boomers, millennials, Gen Z, etc. can have varied purchasing preferences. Moreover, households with different cultural backgrounds such as immigrants versus nonimmigrants, those from different ethnicities or countries can display affinity towards purchasing certain products that have significant relevance to their own culture. Identification of these variables provides the organization with a baseline for customer segmentation. Such customer insights can help the organization provide the 'right-mix' of products at various stores depending on the demographics of that particular location.

Societies today are increasingly diverse, and consumers have easy access to large amounts of information that influences their buying behaviors. Consumer demographic trends have been shaping the retail industry for a long time, and play a critical role in predicting sales, understanding customer preferences, deciding on store locations, managing the supply and demand of various products, etc. Organizations often research demographic data to understand how well a specific product or service is selling, who their customers are and where those customers are located. In this study, we used census data and tried to predict the consumption of two products of the retail organization. Census data consisted of a combination of various socioeconomic customer characteristics available FSA-wise that helped predict what would the customers buy in the future. Such insights based on demographic data help the organization to better understand its customers and design or increase product availability that is in line with customer preferences. The findings of this study would also help the organization manage the supply of its products in their different stores and design targeted pricing and promotional strategies to improve sales of its products.

## 6. References


[1] U. M. Fayyad, D. Haussler and P. E. Stolorz, "KDD for Science Data Analysis: Issues and Examples," *KDD,* pp. 50-56, 1996.

[2] L. Breiman, "Random forests," *Machine learning,* vol. 45, no. 1, pp. 5-32, 2001.

[3] P. J. Werbos, The roots of backpropagation: from ordered derivatives to neural networks and political forecasting (Vol. 1), John Wiley & Sons, 1994.

[4] B. E. Boser, I. M. Guyon and V. N. Vapnik, "A training algorithm for optimal margin classifiers," 1992.

[5] J. Bacardit, W. Browne, J. Drugowitsch, E. Bernadó-Mansilla and M. V. Butz, "Learning Classifier Systems," 2011.

[6] S. Trabelsi, Z. Elouedi and P. Lingras, "Classification systems based on rough sets under the belief function framework," *International Journal of Approximate Reasoning,* vol. 52, no. 9, pp. 1409-1432, 2011.

[7] P. Lingras and C. Butz, "Rough set based 1-v-1 and 1-vr approaches to support vector machine multi-classification," *Information Sciences,* vol. 177, no. 18, pp. 3782-3798, 2007.

[8] T. Islam, N. Meade, R. T. Carson, J.J. Louviere, and J. Wang, "The usefulness of socio-demographic variables in predicting purchase decisions: Evidence from machine learning procedures", *Journal of Business Research*, vol. 151, pp. 324-338, 2022

[9] K. Matuszelański and K. Kopczewska, "Customer Churn in Retail E-Commerce Business: Spatial and Machine Learning Approach", *Journal of Theoretical and Applied Electronic Commerce Research,* vol. 17, no. 1, pp. 165-198, 2022.

# The digital maturation of businesses in Sogn og Fjordane

Malin Waage [1], Bjørn Christian Weinbach [1] and Øyvind H. Larsen [1]

[1] *Vestlandsforsking, Sogndal, Vestland, Norway*

**Abstract**

Today, the amount of data collected by and available to businesses for value exploitation is increasing exponentially. However, we know less about businesses' willingness, knowledge, capability, and plans to exploit their data and take advantage of artificial intelligence and data for value gain. The small village of Sogndal, which hosts the location of the SAIDD symposium, is part of the larger region of Sogn og Fjordane. Here, many technology-, robotics-, and artificial intelligence-rooted companies and research groups have their primary location. Does this affect the general business life's ability and understanding of adopting digital technologies? What is the status quo of digital maturity among businesses here? To answer these questions, a questionnaire has been sent out to businesses in the region and the results will be presented at the symposium. The answers are important to better understand how data-, technology-, and robotics-science can contribute to strengthening the local business life.

**Keywords** Digital maturity, business life, Industry 4.0, Artificial Intelligence

The amount of data captured, collected, and used globally is increasing rapidly, and is forecasted to continue to increase into the future [1]. Exploiting such data, or information, is becoming an asset for governments and enterprises as we enter the fourth industrial revolution (Industry 4.0) [2], or "the digital shift", thus the possibilities in utilizing technology and data to perform better, more effectively, and sustainably are increasing rapidly across sectors. But how far are businesses in various sectors taking advantage of the digital possibilities? Does the size of the company play a role? What about the company's location in terms of population density, rural or urban?

The region of Sogn og Fjordane (SFJ) has a population of only approximately 109,000 [3]. Despite being rural, it is world-renowned for its robotics development through several companies and university groups, and many nationally and internationally known digital companies and companies known for being at the forefront of technology have their primary location there. How does this influence the general business life in the region? We have distributed a questionnaire to businesses in SFJ to gain insight into their willingness, ability, and plans to adopt digital technologies to create increased competitiveness and growth [3]. The first part of the questionnaire entails general information about the company and the work-position of the respondent. The main part presents many hypotheses essentially about the company's thoughts and plans on technology development, data utilization and artificial intelligence, while in the last voluntary part the respondent can explain with their own words their plans, challenges and competence needs regarding digitalization – the last part is an addition to the otherwise identical questionnaire sent out to businesses in other nearby regions [3]. The similarity of the questionnaires is important to enable a comparison and a merge of the answers when mapping the greater region of Norway (Vestlandet) in cooperation with Bergen Næringsråd and others. We plan to share and convey our questionary results with relevant business hubs and councils as well as policy makers. Here, we are interested in an exploratory discussion and exchange of insights on good practices for facilitating the adoption of digital methodologies and technology innovations in the business realm, with a view to strengthen the digital maturity status of companies. Key questions that may arise are:

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: mwa@vestforsk.no (A. 1); bcw@vestforsk.no (A. 2); ohl@vestforsk.no (A. 3)
ORCID: 0000-0002-6924-2273 (A. 1); 0000-0002-9812-100 (A. 2); 0000-0003-3515-6809 (A. 3)

What are recommended practices for conducting collaborative research between academia and industry, and how can such exemplars be most effectively disseminated to catalyze change?

## Acknowledgement

This work is funded by and a part of the 6-year long project Teknoløft Sogn og Fjordane, which aims to build capacity and expertise in digitalization and automation, robotics, and big data in the region of Sogn og Fjordane in close collaboration with local enterprises.

## References

[1] Taylor, P. (2022). "Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2020, with forecasts from 2021 to 2025". In Statista - The Statistics Portal. Retrieved February 09, 2023, URL: https://www.statista.com/statistics/871513/worldwide-data-created/.

[2] Hermann, M., Pentek, T., & Otto, B. (2016, January). "Design principles for industrie 4.0 scenarios." In 2016 49th Hawaii international conference on system sciences (HICSS) (pp. 3928-3937). IEEE.

[3] Warncke et al., (2022). "Digital Modenhet på Vestlandet Delrapport 1: Kunstig intelligens". Bergen Næringsråd and Norwegian Cognitive Centre (2022), online report. URL: https://d3gkcpa86cdznk.cloudfront.net/1669293750/kartlegging-digital-modenhet-delrapport-1.pdf

[4] Kyrkjebø (2018) «Teknoløft Sogn og Fjordane», project page at the HVL webpage. URL: https://www.hvl.no/teknoloft

# Automated Vegetation Monitoring using Satellite Images and Artificial Intelligence

Michele Gazzea[1,*], Reza Arghandeh[1]

[1]*Department of Computer science, Electrical engineering, and Mathematical sciences, Western Norway University of Applied Sciences, Bergen, Norway.*

Abstract

Vegetation Management is a significant preventive maintenance expense in many power transmission and distribution companies. Traditional Vegetation Management operational practices have proven ineffective and time-consuming. The rise of satellite imagery data and machine learning provides an opportunity to close the loop with continuous data-driven vegetation monitoring. In this work, we propose an automated framework for monitoring vegetation along power lines using high-resolution satellite imagery.

**Keywords**

Remote sensing, Power lines inspection, Artificial intelligence, Satellite imagery

Power transmission and distribution lines are one of the most critical infrastructures networks in our society, delivering electricity and power to all our activities and services. However, they are exposed to harsh weather conditions such as heavy snow, hail, and strong winds. Vegetation, in particular, combined with severe weather conditions, is the predominant reason for outages in power systems [1]. Trees can fall over the lines, causing electric disruptions, outages, and massive wildfires, especially in dry areas. Therefore, constant vegetation monitoring and management are crucial for all electric utilities.

Traditional monitoring procedures involve a time-based fixed-cycle approach. Operators usually send ground patrols to drive along the lines, visually inspecting and annotating risky areas and potential trees encroaching the ROWs. An alternative is flying over the lines with helicopters or drones for optical or laser scanning (LiDAR) surveys of the power lines. LiDAR data provide an accurate 3D representation of an environment. However, LiDAR data acquisition and processing are extremely pricey and time-consuming. Current procedures are generally time-consuming and extremely costly. The vast size of service territories further increases the cost. Power lines can easily span hundreds of kilometers, passing through inaccessible areas like mountains or fjords. Airborne sensors are limited by flight time, large area coverage and high operations costs. According to [2], LiDAR data is in the range of 62–240€ per km2, while aerial imagery is approximately 35–62€ per km2. Therefore, the typical vegetation monitoring cycle varies between one to ten years [3].

Nowadays, several commercial satellite providers offer easy-accessible images for any part of

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
*Corresponding author.
EMAIL: mgaz@hvl.no (M. Gazzea); arghandehr@gmail.com (R. Arghandeh)
ORCHID: 0000-0003-0759-4887 (M. Gazzea); 0000-0002-0691-5426 (R. Arghandeh)

the world with a high revisiting time (currently up to 0.25 - 0.5 meters/pixel resolution). Furthermore, the drop in launching costs and the growing number of satellites and mini-satellites in orbit combined with high-quality sensors have reduced the cost of satellite imagery to less than 15€ per km2 [2]. Satellite imagery is, therefore, currently the best trade-off between acquisition price, quality, and revisiting frequency [4]. The numerous sensors continuously monitoring the earth unfold the possibility of synergistic use of high-resolution satellite images. Common applications of remote sensing include disaster management (forest fires [5], hurricane impact [6], flood and drought monitoring [7]), agriculture (soil moisture monitoring [8], vegetation monitoring [9], [3]), deforestation ([10]) and urban planning [11], just to mention some. Consequently, satellite imagery brings the opportunity to combine scale, frequency, and cost efficiency to enhance situational awareness regarding vegetation encroachment in power lines' right-of-way using high-resolution satellite imagery. Therefore, vegetation management can be changed from traditional time-based monitoring to risk-based monitoring.

In this project, we propose a framework to monitor vegetation along power lines using high-resolution satellite images. We make use of the recent advances in computer vision. In particular, we use Convolutional Neural Networks (CNNs) to process the satellite images. CNNs have become the leading machine learning methodology due to their effectiveness at extracting feature representations from images for classification and segmentation purposes [12]. Trees are detected and classified from satellite imagery using tailored CNN-based models. For each location along power lines, we characterized trees based on the proximity and distribution with respect to power lines, number and density, and botanical species [9]. A risk analysis is then performed using a collision simulation in combination with weather data and forest dynamics, such as shielding and gap factors. The output is a risk map showing the locations with nearby trees more prone to failure due to wind. The proposed framework, integrated into the existing vegetation management pipeline, can help the electric utilities better understand the status of the vegetation along the grid. This would reduce the cost and time of power line monitoring by partially replacing ground patrols and helicopter or drone inspection with satellite data analytics and planning the vegetation management accordingly.

## Acknowledgments

This work is funded partly by the European Space Agency (ESA) under the Smart Grid Eye, from Space to Sky (GridEyeS) project [13].

## References

[1] H. Gugel, S. Ekisheva, M. Lauby, F. Tafreshi, Vegetation-related outages on transmission lines in north america, in: 2018 IEEE Power and Energy Society General Meeting (PESGM), 2018. doi:10.1109/PESGM.2018.8586462.

[2] H. O. Ørka, M. Hauglin, Use of remote sensing for mapping of non-native conifer species, https://www.miljodirektoratet.no/globalassets/publikasjoner/m490/m490.pdf, 2016. [Online].

[3] M. Gazzea, M. Pacevicius, D. O. Dammann, A. Sapronova, T. M. Lunde, R. Arghandeh, Automated power lines vegetation monitoring using high-resolution satellite imagery, IEEE Transactions on Power Delivery 37 (2022) 308–316. doi:10.1109/TPWRD.2021.3059307.

[4] L. Singh, W. Whittecar, M. DiPrinzio, J. Herman, M. Ferringer, P. Reed, Low cost satellite constellations for nearly continuous global coverage, Nature Communications 11 (2020). doi:10.1038/s41467-019-13865-0.

[5] L. B. Lentile, Z. A. Holden, A. M. S. Smith, M. J. Falkowski, A. T. Hudak, P. Morgan, S. A. Lewis, P. E. Gessler, N. C. Benson, Remote sensing techniques to assess active fire characteristics and post-fire effects, International Journal of Wildland Fire 15 (2006).

[6] M. Gazzea, A. Karaer, M. Ghorbanzadeh, N. Balafkan, T. Abichou, E. E. Ozguven, R. Arghandeh, Automated satellite-based assessment of hurricane impacts on roadways, IEEE Transactions on Industrial Informatics 18 (2022) 2110–2119. doi:10.1109/TII.2021.3082906.

[7] T. Lopez, A. Al Bitar, S. Biancamaria, A. Güntner, A. Jäggi, On the use of satellite remote sensing to detect floods and droughts at large scales, Surveys in Geophysics 41 (2020) 1461–1487. URL: https://doi.org/10.1007/s10712-020-09618-0. doi:10.1007/s10712-020-09618-0.

[8] E. Babaeian, M. Sadeghi, S. B. Jones, C. Montzka, H. Vereecken, M. Tuller, Ground, proximal, and satellite remote sensing of soil moisture, Reviews of Geophysics 57 (2019) 530–616. doi:https://doi.org/10.1029/2018RG000618.

[9] M. Gazzea, L. M. Kristensen, F. Pirotti, E. E. Ozguven, R. Arghandeh, Tree species classification using high-resolution satellite imagery and weakly-supervised learning, IEEE Transactions on Geoscience and Remote Sensing (2022) 1–1. doi:10.1109/TGRS.2022.3210275.

[10] Y. Gao, M. Skutsch, J. Paneque-Gálvez, A. Ghilardi, Remote sensing of forest degradation: a review, Environmental Research Letters 15 (2020) 103001. URL: https://doi.org/10.1088/1748-9326/abaad7. doi:10.1088/1748-9326/abaad7.

[11] N. Kadhim, M. Mourshed, M. Bray, Advances in remote sensing applications for urban sustainability, Euro-Mediterranean Journal for Environmental Integration 1 (2016) 7. URL: https://doi.org/10.1007/s41207-016-0007-4. doi:10.1007/s41207-016-0007-4.

[12] R. Kemker, C. Salvaggio, C. Kanan, Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning, Isprs Journal of Photogrammetry and Remote Sensing 145 (2018) 60–77.

[13] European Space Agency (ESA), Grideyes project, https://business.esa.int/projects/grideyes/, 2019. [Online].

# Violence-inducing Behaviour Prevention in Social-Cyber Space

Yasas Senarath [1], Hemant Purohit [1] and Rajendra Akerkar [2]

[1] George Mason University, Fairfax, USA
[2] Western Norway Research Institute, Sogndal, Norway

**Abstract**

Hate speech, radicalization and polarization in online social environments is one of the leading global societal challenges today. How to respond to online hate speech is a question troubling many democracies – including Norway. In this talk we will present research in progress to create the foundation of a transformative new technology for combating radicalization and hate speech in online social spaces.

**Keywords**

Big social data, resilience, malicious behaviour online, social media

In the era of digitalization, community resilience has increasingly become a priority for local to national governments. Available tools to support resilience initiatives of the local community services lack the ability to assess and mitigate the dynamic risks associated with the malicious public behaviour online such as spreading hate and misinformation on social media, which can significantly harm official response efforts during disasters in their communities. A variety of large-scale social media datasets, collaborative mapping tools, and data science approaches have emerged that can facilitate computational social science research to gain a better understanding of community resilience processes accounting for public behaviour and design actionable tools for the community services.

In this research, our primary objective is to improve community resilience using a principled approach of social cybersecurity by developing methods and tools to timely inform local Norwegian community services for proactive interventions at scale regarding violence-inducing social behaviours by individuals online.

In this talk, we will discuss our approach to analyse the big data sources at scale by taking in account all ethical, social and legal challenges and considerations, in contrast to only existing approaches of small-scale human observations, or survey-based analytical approaches.

## Acknowledgement

This project has received funding from the Research Council of Norway (RCN) as a Researcher Project for Technological Convergence related to Enabling Technologies.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: ywijesu@gmu.edu (A. 1); hpurohit@gmu.edu (A. 2); rak@vestforsk.no (A. 3)
ORCID: 0000-0002-8759-378X (A. 1); 0000-0002-4573-8450 (A. 2); 0000-0002-4048-7226 (A. 3)


# Assessing the Continuous Causal Responses of Typhoon Related Weather on Human Mobility: An Empirical Study in Japan

Hongjun Wang[1], Zipei Fan[2], Xuan Song[1] and Xuan Song[1]

[1] *Department of Computer Science and Engineering, Southern University of Science and Technology*
[2] *Center for Spatial Information Science, University of Tokyo*

## Abstract

Understanding the causal impact of natural disasters (e.g., typhoon, flood) on human mobility is of great significance, regarding the life quality and safety of urban residents. However, this topic is rarely investigated and similar works all struggle to accurately evaluate such effect due to the presence of other factors that may also influence mobility, such as time-periodic behaviors and the willingness of people. These factors are not within our research scope but will present non-negligible influences to experimental outcomes and thus their effects should be illuminated to achieve variable control for analysis. In the language of causality, these factors are termed as "confounders", which are usually unobservable and time-varying. To assess the pure causal impact of natural disasters, the key challenge actually lies in how to handle the confounders in problem modeling.

## Keywords

Causal Inference, Individual Treatment Effect, Smart City Development, Human Mobility

In this study, we utilize the weather data collected from two typhoon events in Japan (i.e., Typhoon Faxai and Hagibis), and the human mobility data of different Japanese counties over a long time period (including the time of two typhoons), to study the causal impacts of typhoons on human mobility. In other words, we aim to quantify to what extent that different kinds of typhoon-related weather will change human mobility. We address this problem by using a neural network-based framework that (1) is able to learn the representations of time-varying (unobserved) confounders from the observational data, and (2) quantifies the dose-response curve (i.e., causal effects) between typhoon related weather and human mobility via counterfactual deconfounding. Our framework has its root in a commonly used shared bottom architecture of causal reasoning but take a step forward to extract the information of unobservable confounders (a usual case in reality) from observations, so that we can perform more helpful analysis in realistic scenarios. Finally, we validate our design with real-world datasets and the results indicate the effectiveness of our proposed framework, where the assessment of causal effect is consistent with existing empirical researches on the control of natural disasters. We believe our research may shed lights on a promising direction towards a better and more intelligent people safety management against natural crisis.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: wanghj2020@mail.sustech.edu.cn (A.1); fanzipei@iis.u-tokyo.ac.jp (A.2); songx@sustech.edu.cn (A.3)


# Terminology Saturation Analysis for Machine Learning and Event Detection

Vadim Ermolayev [1] and Victoria Kosa [2]

[1] *Ukrainian Catholic Univ., 2a Kozelnytska st., Lviv, 79076, Ukraine*
[2] *Zaporizhzhia National Univ., 66 Zhukovskogo st., Zaporizhzhia, 69600, Ukraine*

**Abstract**

In this lightning talk, we inform about our recent research on terminological saturation analysis (TSA) in subject domain-bounded textual corpora. Further, we share our ideas on how TSA could be used for (i) constructing knowledge graphs that representatively describe a subject domain, or topic; (ii) extracting the smallest possible, yet representative, textual datasets from domain corpora for training machine learning models in natural language processing; and (iii) event detection from textual streams data.

**Keywords**

Terminological saturation analysis, knowledge graph, machine learning, event detection

Our recent research has demonstrated [1] that TSA could reveal patterns and trends in domain-bounded textual data – e.g. topical collections of scholarly publications[2]. We also discovered that this analytical approach could help detect trends of technology adoption in industry [3]. We found out that the major factors hampering terminological saturation were: (i) the immaturity of the domain implying that the domain-bounded corpus is too small; (ii) the heterogeneity within the domain – e.g. the fragmentation of the domain due to the competition among different R&D strands; or (iii) the volatility of the domain terminology over time. Based on these findings, it was remarkable to note that the existence of a terminologically saturated sub-collection in a corpus of texts – a terminological core collection – indicates the maturity and stability of the respective topic or domain. On the other hand, the absence of terminological saturation points out that an opportunity window is open for the further development of the focal domain, including the mergers of competing strands. Application wise, our research was aimed at ensuring the completeness of a text corpus in a domain for ontology learning from texts. However, the results seem to have a broader potential R&D impact.

One potentially good use is extracting the smallest possible, yet representatively complete datasets for training machine learning models for natural language processing tasks. Furthermore, a knowledge graph, built using the terminology extracted from such a terminological core dataset, could be used as a structured representation of the set of features in the domain. This might help make the outputs from the trained deep learning models better explainable.

Another potential use case could be in event detection and prediction using social media text streams. We hypothesise that a terminologically saturated topical stream of, say, tweets over a period of time might point out that either (i) the topic stream is dominated by one group using coherent terminology; or (ii) the majority of the community around the topic is focused on something very important, that already happened in the past or will happen soon. On the other hand, the lack of terminological saturation in a topical stream might point out that the situation around the topic is stable in the democratic sense, which is characterised by the plethora of different competitive opinions and judgements on the topic.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: victoriya1402.kosa@gmail.com (A. 1); ermolayev@ucu.edu.ua (A. 2)
ORCID: 0000-0002-7300-8818 (A. 1); 0000-0002-5159-254X (A. 2)



[2] Please see the review of the related work in our chapter [2].

## References

[1] Kosa, V., Ermolayev, V.: Terminology Saturation: Detection, Measurement, and Use. Cognitive Science and Technology, Springer Singapore (2022)
[2] Kosa, V., Ermolayev, V.: Related Work and Our Approach. In: Terminology Saturation: Detection, Measurement, and Use. Cognitive Science and Technology, Springer Singapore, p. 7--39 (2022)
[3] Kosa, V., Ermolayev, V.: Saturated Terminology Extraction and Analysis in Use. In: Terminology Saturation: Detection, Measurement, and Use. Cognitive Science and Technology, Springer Singapore, p. 155--170 (2022)

# Automatic Error Detection and Correction of Cadastral Boundary Information with AI

Simon Šanca[1], Sjur Kristoffer Dyrkolbotn[1], Leiv Bjarte Mjøs[1] and Reza Arghandeh[1]

[1]Western Norway University of Applied Sciences, Inndalsveien 28, Bergen, 5063, Norway

An effective land administration system promotes the underlying conditions for efficient registration of properties thus creating sustainable conditions for natural resource management [1]. Automating the mapping process for the natural and built environment is a crucial step towards sustainability and greater digitalization in the cadastre. Current methods for cadastral property mapping include the delineation of visible cadastral boundaries mostly using UAV images [2, 3] applying artificial intelligence.

There is a lack of systems that can efficiently recreate cadastral property boundaries from historical surveying records. In the past there have been some approaches to recreate historical city maps [4], to extract morphological features from historical maps [5] and a competition in Historical Map Segmentation (MapSeg) was organized [6]. Recreating cadastral property boundaries from historical records is an interesting challenge that needs to be solved. The Norwegian cadastre lacks improvement, especially the boundary determination system, which has uncertain quality [7]. In 2018 it was claimed that 50% of registered properties with boundaries have inadequate quality or completely lack boundary information [8]. Recreating boundary marks with standard land surveying techniques is a labour intensive and expensive job, therefore new techniques and methodologies need to be established to accelerate cadastral mapping from aerial images and historical records, including better methods for error detection and correction from existing data. The Norwegian cadastre could be improved by utilising AI-based methods for cadastral mapping, promoting greater digitalization and automation.

Our proposed research will focus on improving the cadastre by developing a document detection system applying natural language processing, computer vision and logical reasoning to automatically extract information from historical surveying records. In Norway the official documents describing property boundaries are the *målebrev* and the *skyldskifte*. Both documents contain a textual part and sometimes the map (sketch) part, graphically representing the measured property. We aim to develop our solution into a working QGIS-plugin, that can efficiently recreate properties using historical surveying documents as input. The output of our plugin will be the georeferenced property with exploited boundary information. The proposed idea is futher presented on figure 1.

An automated document detection system would help with error detection and correction, may help to solve potential boundary disputes, that arise between landowners. It would help land surveyors by processing large amounts of archival records of properties, requiring less manual extraction of information, as well as promoting digitalization in land surveying.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: simon.sanca@hvl.no (S. Šanca); sjur.kristoffer.dyrkolbotn@hvl.no (S. K. Dyrkolbotn); leiv.bjarte.mjos@hvl.no (L. B. Mjøs); reza.arghandeh@hvl.no (R. Arghandeh)

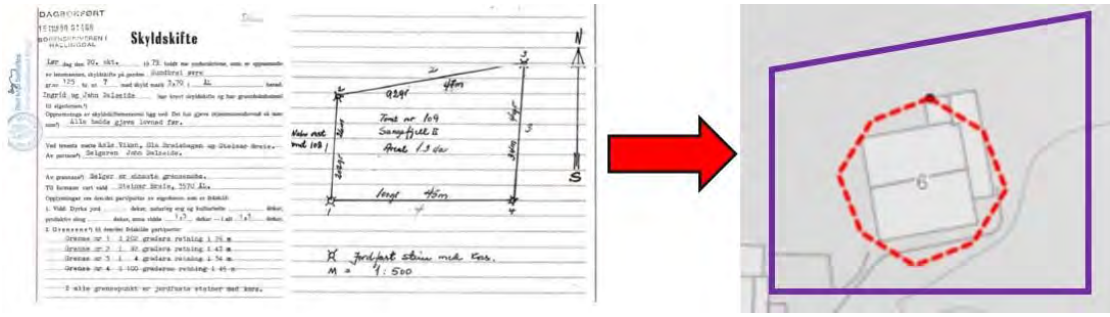



**Figure 1:** The idea is to detect the textual and the graphical part of the *målebrev* or skyldskifte and recreate the boundary of the measured property with AI, using natural language processing to read the textual part and computer vision to interpret the graphical part. The final result would be the georeferenced property boundary in marked with purple.

## References


[1] C. Persello, J. D. Wegner, R. Hänsch, D. Tuia, P. Ghamisi, M. Koeva, G. Camps-Valls, Deep learning and earth observation to support the sustainable development goals: Current approaches, open challenges, and future opportunities, IEEE Geoscience and Remote Sensing Magazine 10 (2022) 172–200. doi:10.1109/MGRS.2021.3136100.

[2] B. Fetai, K. Oštir, M. Kosmatin Fras, A. Lisec, Extraction of Visible Boundaries for Cadastral Mapping Based on UAV Imagery, Remote Sensing 11 (2019) 1510. doi:10.3390/rs11131510.

[3] S. Crommelinck, M. Koeva, M. Y. Yang, G. Vosselman, Application of Deep Learning for Delineation of Visible Cadastral Boundaries from Remote Sensing Imagery, Remote Sensing 11 (2019) 2505. doi:10.3390/rs11212505.

[4] A. Nobajas, F. Nadal, From historical map to online 3D recreation: the 1861 cadastral map of Horta (Barcelona), Cartography and Geographic Information Science 42 (2015) 211–223. doi:10.1080/15230406.2014.998285.

[5] Y. Chen, E. Carlinet, J. Chazalon, C. Mallet, B. Duménieu, J. Perret, Combining Deep Learning and Mathematical Morphology for Historical Map Segmentation, in: Discrete Geometry and Mathematical Morphology, Lecture Notes in Computer Science, Springer International Publishing, Cham, 2021, pp. 79–92. doi:10.1007/978-3-030-76657-3_5.

[6] J. Chazalon, E. Carlinet, Y. Chen, J. Perret, B. Duménieu, C. Mallet, T. Géraud, V. Nguyen, N. Nguyen, J. Baloun, L. Lenc, P. Král, ICDAR 2021 Competition on Historical Map Segmentation, in: Document Analysis and Recognition – ICDAR 2021, Lecture Notes in Computer Science, Springer International Publishing, Cham, 2021, pp. 693–707. doi:10.1007/978-3-030-86337-1_46.

[7] L. B. Mjøs, Matrikulær utvikling i Norge, Ph.D. thesis, 2016. doi:10.13140/RG.2.1.2856.2806.

[8] L. B. Mjøs, Cadastral development in Norway: the need for improvement, Survey Review 52 (2020) 473–484. doi:10.1080/00396265.2019.1637094.

# Causally guided Intelligent Transportation System

Jiyuan Chen[1], Jiayuan Zheng[2], Zipei Fan[3] and Xuan Song[1]

[1] *Department of Computer Science and Engineering, Southern University of Science and Technology*
[2] *School of Management, Zhejiang University*
[3] *Center for Spatial Information Science, University of Tokyo*

## Abstract

The past decade has witnessed Artificial Intelligence (AI) revolutionizing the transportation industry with advanced deep neural networks (DNNs). Compared with traditional models, deep learning methods offer superior predictive power in many practical application fields such as ride hailing [1], mobility prediction [2] and congestion forecasting [3]. However, despite the satisfying performances, DNNs preserve their black-box nature and make decisions depending heavily on statistical correlations. These drawbacks directly lead to DNN's lack of interpretability and robustness, which then become an important obstacle for their deployment into real-world scenarios. Recently, causal inference [4] has gain much attention in various research fields including computer vision [5] and graph learning [6]. Its combination with deep learning models, causally guided DNNs, is currently a hot topic because causality helps us to think beyond the black box of DNNs to offer interpretablity, and encourages the model to capture the true causal relations to offer robustness. To the best of our knowledge, the use of causal inference in current DNN-based intelligent transportation system is still in its very infancy. We believe this line of research can shed lights on a more powerful, transparent, and robust transportation system to provide more convenience and supports to residents and smart city development.

## Keywords

Deep Learning, Causal Inference, Smart City Development, Intelligent Transportation System

Lacking interpretability and robustness is a crucial problem for DNNs in transportation system. For example, without interpretablity, we will not be able to understand why model thinks there will be a traffic jam and thus cannot intervene to avoid it. And depending solely on statistical correlations will hurt model's robustness because of its inability to capture the true cause of an event. Considering a situation when we train a traffic prediction model, if in the training set the inflow of a road exhibit strong statistical correlation with that of another road ten miles away, the model will simply learn such spurious relation. However, it is not the true causality and is very likely not to present in the test set. This is a common issue called distribution shift [7] in machine learning, and can lead to huge performance degradation for models relying on statistical correlation. Causal inference, on the other hand, are by definition interpretable and robust. Causal links offer explanations of the model predictions, and underlying causal structure (relation) holds true for different environments to provide robustness. We believe current transportation system can greatly benefit from using causality to guide the learning of its DNNs. For instance, in traffic prediction, we can encourage the model to "think" why the inflow of the road will increase by requiring it to identify which adjacent neighbor contributes to the inflow. And we can also use counterfactual analysis [8] to interpret the model prediction by intervening the model input to observe the change of its outcome.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: 11811810@mail.sustech.edu.cn (A.1); zjyyy18318103682@163.com (A.2); fanzipei@iis.u-tokyo.ac.jp (A.3); songx@sustech.edu.cn (A.4)

## References


[1] J. Huang, L. Huang, M. Liu, H. Li, Q. Tan, X. Ma, J. Cui, D.-S. Huang, Deep reinforcement learning-based trajectory pricing on ride-hailing platforms, ACM Transactions on Intelligent Systems and Technology (TIST) 13 (2022) 1–19.

[2] M. Luca, G. Barlacchi, B. Lepri, L. Pappalardo, A survey on deep learning for human mobility, ACM Computing Surveys (CSUR) 55 (2021) 1–44.

[3] M. Chen, X. Yu, Y. Liu, Pcnn: Deep convolutional networks for short-term traffic congestion prediction, IEEE Transactions on Intelligent Transportation Systems 19 (2018) 3550–3559.

[4] J. Pearl, Causal inference, Causality: objectives and assessment (2010) 39–58.

[5] Y. Niu, K. Tang, H. Zhang, Z. Lu, X.-S. Hua, J.-R. Wen, Counterfactual vqa: A cause-effect look at language bias, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 12700–12710.

[6] Y. Sui, X. Wang, J. Wu, M. Lin, X. He, T.-S. Chua, Causal attention for interpretable and generalizable graph classification, in: Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2022, pp. 1696–1705.

[7] T. Fang, N. Lu, G. Niu, M. Sugiyama, Rethinking importance weighting for deep learning under distribution shift, Advances in neural information processing systems 33 (2020) 11996–12007.

[8] S. L. Morgan, C. Winship, Counterfactuals and causal inference, Cambridge University Press, 2015.

# Geolocation data as a research tool for the organization of the settlement system

## Case study of the spatial mobility model in Czechia

Václav Jaroš, Jaroslav Šimeček

*Ministry of the Interior, Czech Republic*

**Abstract**

Geolocation data is a widely used source of the spatial information about the population. Their great potential might be also used for population mobility research to identify spatial interactions forming the hierarchical structure of the settlement system. For this purpose, a model of data acquisition and their preliminary analysis was developed. This model represents an effective tool for mapping the mobility behaviour of the population. Using the example of Czechia, primary commuting links are identified, which are subsequently analysed in detail using GIS tools in both desktop and online environments. Therefore, important commuting centres of different hierarchical levels are defined by the volume and nature of spatial interactions. This approach is used as a source of important expertise for the proposals on subsequent Czech public administration reform. Nevertheless, the entire model is generally transferable, and the entire method of using the geolocation data for mapping the hierarchy within the settlement system can be replicated in other countries as well.

**Keywords**

Big data, mobility behaviour, data usage model

The whole approach is based on the presumption that mobile phones move together with their users for most of the day. Furthermore, the assumption of high penetration of the population by mobile phones is also crucial. In general, it can be concluded that in contemporary societies of developed countries, both assumptions are fulfilled.

This model is flexible in terms of the output databases produced. According to the primary setting, it produces a total of 15 attributes structured into 3 basic interconnected datasets: a) statistical data for individual municipalities, b) OD matrix showing commuting directions, and c) the average number of currently present population in every hour of the week (24/7) in each municipality. The output databases themselves do not indicate specific measured values for a certain day or period, but each attribute represents basically the number of people who reports a given type of travel (mobility) behaviour. This is no longer the geolocation data itself, but a summary of time-spatially aggregated statistics about geolocation data.

Primarily, the method is set to identify functional micro-regional commuting links. Microregions are territories in which a resident should be able to secure all his daily activities necessary and important for his everyday life. Their centres are primary commuting destinations for their surroundings and provide a sufficient range of job opportunities, primary and secondary education, health services, shops, etc.

This approach was used in the Czech Republic for a comprehensive revision of the spatial units of the public administration structure. The purpose of this activity was to harmonize the administrative units with natural commuting regions. Particularly, the aim was to ensure that public administration offices were located where people naturally concentrated. This leads to streamlining and deconcentration of the public administration and its adaptation to the needs of citizens. Based on this application example, it is also possible to conclude about the transferability of this approach and its applicability either in other territories (states) or in other scientific fields.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: vaclav.jaros@mvcr.cz (A. 1); jaroslav.simecek@mvcr.cz (A. 2)

## References


HALÁS, M., BLAŽEK, V., KLAPKA, P., KRAFT, S. (2021). Population movements based on mobile phone location data: the Czech Republic. Journal of Maps 17 (1), 116–122. ISSN 1744-5647, DOI: 10.1080/17445647.2021.1937730.

HAMPL, M. (2005): Geografická organizace společnosti v České republice: transformační procesy a jejich obecný kontext. Univerzita Karlova, Praha, 147 s.

HAMPL M., MARADA M. (2015): Sociogeografická regionalizace Česka. Geografie, vol. 120, no. 3, pp. 397–421. DOI:10.37040/geografie2015120030397

JAROŠ V. (2017). Social and transport exclusion. Geographia Polonica, vol. 90, no. 3 pp 247-263. DOI: 10.7163/GPol.0099.

MARADA M., FRÁNĚ L., JANOŠ V., JAROŠ V., KRAFT S., KŘÍŽ M., KOWALSKI M. (2016). Rychlá spojení metropolitních oblastí: Dopady (nové) dostupnosti na pracovní trh. Project TAČR TB0500MD005: Fast connections among metropolitan areas: impact of (new) accessibility on labour market.

MAZOUCH, P. a kol. (2017): Limity využití mobilních sítí ve statistických šetřeních ČSÚ. Project TAČR TD03000452. Limits of data from mobile sites in statistical surveys of Czech statistical office.

MV ČR (2020): Využití geolokačních dat mobilních operátorů pro potřeby veřejné správy. EAA grants, GG-PDP1-001, Improvement of preconditions for decentralisation and availability of public administration in the territory.

NOVÁK, J. (2010): Lokalizační data mobilních telefonů: Možnosti využití v geografickém výzkumu. Disertační práce. Katedra sociální geografie a regionálního rozvoje PřF UK, Praha, 175 s.

ŠVEDA, M. - BARLÍK, P. (2018). Daily commuting in the Bratislava metropolitan area: case study with mobile positioning data. In Papers in Applied Geography, 4(4): 409-423.


## Acknowledgement


Implemented as part of the project Improvement in Conditions for Decentralization and Accessibility of State Administration in the Territory financed from EEA and Norway Funds 2014 – 2021.

# Big Data and Machine Learning for Analysis of Numerical PDE

Roman Khotyachuk[1,*], Klaus Johannsen[2]

[1]University of Bergen, Allégaten 41, 5007 Bergen, Norway
[2]NORCE Norwegian Research Centre, Nygårdsgaten 112, 5008 Bergen, Norway

**Abstract**

Density-driven flows in porous media are frequently observed in natural and technical systems and have significant environmental, economic, and social impacts.

Numerical modeling of flows in porous media is used to investigate and predict their behavior in several important practical cases, such as risk assessment, water resource management, carbon sequestration, and geothermal reservoir study. Such modeling is challenging because of the high variability in flow processes, the non-linearity of governing PDEs, and the wide range of time and space domains where a problem is solved numerically.

In this work, we used Big Data and Machine Learning to analyze vast amounts of data produced by numerical solvers of PDEs. Our setup enabled mass parallel runs of the d$^3$f solver [1], efficient post-processing, and further analysis of large ensembles of numerical solutions of PDEs. Using both supervised and unsupervised machine learning techniques, we obtained new scientific results for the Elder problem [2], an example of flows in porous media.

**Keywords**

Scientific Big Data, Machine Learning, Numerical PDE, Elder problem

## 1. Project Description

We implemented the system of parallel runs of the d$^3$f PDE simulation software [1], which is integrated with the Hadoop/Spark cluster. Specifically, our Spark+d$^3$f setup is used to analyze the Elder problem [2, 3]. For this problem, we achieved the following results:

- Investigated the steady-state solutions of the Elder problem with regard to the Rayleigh numbers (Ra), perturbations, etc.;
- Analyzed the complexity of solutions regarding time and other factors using different data complexity metrics;
- Developed predictive models for the Elder problem based on ML techniques [5].

From the point of view of dynamical systems theory, this study is an attempt to investigate and understand a larger fraction of the entire phase space of the Elder problem. We are sampling the phase space with a larger number of samples (10000 or even more perturbed solutions) intending to understand the nonlinear dynamics of the Elder problem using data-driven approaches to complex nonlinear dynamical systems [4].

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
*Corresponding author.
EMAIL: roman.khotyachuk@uib.no (R. Khotyachuk); kljo@norceresearch.no (K. Johannsen)

## References


[1] Fein, E. $d^3f$ – ein Programmpaket zur Modellierung von dichtegetriebenen Strömungen. GRS, Braunschweig, GRS-139, **1998**, ISBN 3-923875-97-5.

[2] Elder, J.W. Transient convection in a porous medium. Journal of Fluid Mechanics, Volume 27, Issue 3, 24 February 1967, pp. 609 - 623. DOI: https://doi.org/10.1017/S0022112067000576.

[3] Simmons, C.T. and Elder, J.W. The Elder Problem. Groundwater, **2017**, Vol. 55, 926–930. https://doi.org/10.1111/gwat.12593.

[4] Kutz, J. N. Data-Driven Modeling & Scientific Computation: Methods for Complex Systems & Big Data. **2013**. Oxford University Press. ISBN 978-0-19-966034-6.

[5] Data Classification: Algorithms and Applications. Aggarwal, C. (Editor). Chapman & Hall/CRC, 2014. ISBN-10: 1466586745.

[6] Lorena, A., Garcia, L., Lehmann, J., Pereira de Souto, M.C., Ho, T. How Complex is your classification problem? A survey on measuring classification complexity. ACM Computing Surveys (CSUR), **2018**, vol. 52, pp. 1 -34. https://doi.org/10.1145/3347711.

# Returning Home Strategy Analysis using Mobile Sensing Data in Tohoku Earthquake*

Zhiwen Zhang[2], Zipei Fan[1,2,*] and Xuan Song[1,2]

[1] Southern University of Science and Technology, ShenZhen, China
[2] The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa-shi, Chiba, Japan

## Abstract

In recent decades, the frequency and intensity of natural disasters have increased significantly. Most serious disasters lead to massive population movements and evacuations. Analysis of these human activities is essential for planning effective humanitarian assistance, disaster control, and long-term social stability and reconstruction. Although there are a lot of related works for analyzing after a big earthquake [1, 2], few pieces of research consider the influence of personal factors on decision-making. In addition, the knowledge of what key factors impel a person to choose a returning home strategy is important for analyzing the human decision after a big earthquake. In addition, a large number of people remained in the companies or shelters due to the damage to the transport network. Therefore, understanding and predicting human behavior during a disaster will play a vital role in planning effective humanitarian relief, disaster management, and long-term societal reconstruction. However, such research is challenging due to the unavailability of reliable and large-scale human mobility data. Moreover, what key (sometimes hidden) factors do a person consider when making a particular decision for returning home after a big disaster. The relationship between a person making decisions after a big earthquake and the type and number of historical visits must be analyzed and understood. In this study, the smartphone is used to sense people's location. Then big and heterogeneous data were collected on the Tohoku earthquake in Japan in 2011 and discovered grid-based regions of different functions using both human mobility among areas and points of interest (POI) located in a region. This is a "9.0" earthquake that occurred in the Pacific Ocean 130 kilometers off Sendai City, Japan, on March 11, 2011, as a case study [3,4,5]. Then we jointly model the historical visit of the functional areas and decision-making after a big earthquake. Abundant work exploded the mobility pattern [6, 7, 8, 9] from historical visits collected by smartphones and employ them on analyzing people's behavior in disasters. For example, [10] developed a general probabilistic model to simulate population evacuation over complex geographic features in Japan in response to future disasters. Reference [6] found that the distribution of earthquake risk areas, people's emergency measures, and people's behavior can be tracked by using location data. But those work ignore the functional region [11] in practice, which may lack some necessary information for understanding human mobility. Finally, an empirical prediction is developed to estimate how people chose to return home after the Tohoku earthquake. An explainable analysis is conducted to explore the fundamental laws governing human mobility following disasters.

## Keywords

Emergency management, Big mobility data, Explainable knowledge, Decision-making strategy analysis

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: zhangzhiwen@csis.u-tokyo.ac.jp (Z. Zhang); fanzipei@iis.u-tokyo.ac.jp (Z. Fan); songx@sustech.edu.cn (X. Song)


# References


[1] Y. Pan, A. Darzi, A. Kabiri, G. Zhao, W. Luo, C. Xiong, L. Zhang, Quantifying human mobility behaviour changes during the covid-19 outbreak in the united states, Scientific Reports 10 (2020) 1–9.

[2] C. Xiong, S. Hu, M. Yang, H. Younes, W. Luo, S. Ghader, L. Zhang, Mobile device location data reveal human mobility response to state-level stay-at-home orders during the covid-19 pandemic in the usa, Journal of the Royal Society Interface 17 (2020) 20200344.

[3] N. Mori, T. Takahashi, T. Yasuda, H. Yanagisawa, Survey of 2011 tohoku earthquake tsunami inundation and run-up, Geophysical research letters 38 (2011).

[4] N. Mori, T. Takahashi, . T. E. T. J. S. Group, Nationwide post event survey and analysis of the 2011 tohoku earthquake tsunami, Coastal Engineering Journal 54 (2012) 1250001–1.

[5] Y. Y. Kagan, D. D. Jackson, Tohoku earthquake: A surprise?, Bulletin of the Seismological Society of America 103 (2013) 1181–1194.

[6] X. Chaoxu, N. Gaozhong, F. Xiwei, Z. Junxue, P. Xiaoke, Research on the application of mobile phone location signal data in earthquake emergency work: A case study of jiuzhaigou earthquake, PloS one 14 (2019) e0215361.

[7] J. Refonaa, M. Lakshmi, V. Vivek, Analysis and prediction of natural disaster using spatial data mining technique, in: 2015 International Conference on Circuits, Power and Computing Technologies [ICCPCT-2015], IEEE, 2015, pp. 1–6.

[8] M. I. Ramadhan, et al., An analysis of natural disaster data by using k-means and k-medoids algorithm of data mining techniques, in: 2017 15th International Conference on Quality in Research (QiR): International Symposium on Electrical and Computer Engineering, IEEE, 2017, pp. 221–225.

[9] N. Cárdenas-Benítez, R. Aquino-Santos, P. Magaña-Espinoza, J. Aguilar-Velazco, A. Edwards-Block, A. Medina Cass, Traffic congestion detection system through connected vehicles and big data, Sensors 16 (2016) 599.

[10] X. Song, Q. Zhang, Y. Sekimoto, T. Horanont, S. Ueyama, R. Shibasaki, Modeling and probabilistic reasoning of population evacuation during large-scale disaster, in: Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, 2013, pp. 1231–1239.

[11] J. Yuan, Y. Zheng, X. Xie, Discovering regions of different functions in a city using human mobility and pois, in: Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, 2012, pp. 186–194.

# Causal-based Spatio-Temporal Graph Neural Networks: Interpretable Deep Learning Framework for Multivariate Time Series Forecasting

Amir Miraki, Reza Arghandeh

*Department of Computer Science, Electrical Engineering and Mathematical Sciences, Western Norway University of Applied Sciences, Bergen, Norway*

## Abstract

The various types of data collected over time and space have led to the development of spatio-temporal data analysis as an emerging research field [1]. The ubiquity of spatio-temporal data today is unquestionable. For instance, GPS devices have become more widely available, mobile phones have sensors, and sensor technology has improved greatly over the past few years, providing more opportunity to collect data about individuals. Data collection has become easier as a result of all of these factors. Spatio-temporal data are collected in a wide variety of applications, such as smart grids, sensors, social, transportation, and electrical networks [2].

The use of Graph Neural Networks (GNNs) for forecasting spatio-temporal data has become increasingly popular in recent years [3, 4, 5]. GNN is a very suitable model for these types of data due to their irregular structure. Spatial dependence is a critical aspect of spatio-temporal analysis, since it represents the relationship between the variables by the graph, and the base of GNN is this graph. The graphs may be predefined or they can be created using classical methods such as the physical distance between sensors and correlations, for example. Despite their complexity, these networks remain difficult to interpret. Because of this, the models are unreliable and untrustworthy. In addition, recent research on interpretability aims to provide answers to questions related to causality, such as "Why does this model make these decisions?" or "Was it a specific feature that caused the decision made by the model?"

This research aims to improve the interpretability and explainability of GNNs by applying causality [6, 7]. Spatio-temporal data prediction methods based on GNN with an underlying causal graph will be examined in this research. The use of causal models can help us gain a better understanding of how variables are related to each other[8]. Therefore, we rely on DAG graphs and try to design the model for more accurate prediction while at the same time being able to interpret the model. In order to test the effectiveness of the proposed method,

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: amir@hvl.no (A. Miraki); reza.arghandeh@hvl.no (R. Arghandeh)

some benchmark data will be used to evaluate it. In summary, this work focuses on GNNs, explainability, and causal inference.

## References


[1] M. M. Alam, L. Torgo, A. Bifet, A survey on spatio-temporal data analytics systems, ACM Computing Surveys 54 (2022) 1–38.

[2] S. Wang, J. Cao, P. Yu, Deep learning for spatio-temporal data mining: A survey, IEEE transactions on knowledge and data engineering (2020).

[3] A. Kapoor, X. Ben, L. Liu, B. Perozzi, M. Barnes, M. Blais, S. O'Banion, Examining covid-19 forecasting using spatio-temporal graph neural networks, arXiv preprint arXiv:2007.03113 (2020).

[4] X. Wang, Y. Ma, Y. Wang, W. Jin, X. Wang, J. Tang, C. Jia, J. Yu, Traffic flow prediction via spatial temporal graph neural network, in: Proceedings of the web conference 2020, 2020, pp. 1082–1092.

[5] J. Simeunović, B. Schubnel, P.-J. Alet, R. E. Carrillo, Spatio-temporal graph neural networks for multi-site pv power forecasting, IEEE Transactions on Sustainable Energy 13 (2021) 1210–1220.

[6] J. Pearl, Causality, Cambridge university press, 2009.

[7] M. Bunge, Causality and modern science, Routledge, 2017.

[8] C. K. Assaad, E. Devijver, E. Gaussier, Survey and evaluation of causal discovery methods for time series, Journal of Artificial Intelligence Research 73 (2022) 767–819.

# Multisensor Fusion for Monitoring Unstable Rock Slopes - A Case Study from the Stampa Instability, Norway

Lukas Schild[1], Thomas Scheiber[1], Paula Snook[1], Stig Frode Samnøy[1], Alexander Maschler[1], Reza Arghandeh[1] and Lene Kristensen[2]

[1]*Western Norway University of Applied Sciences, Røyrgata 6, 6856 Sogndal, Norway*
[2]*Norwegian Water Resources and Energy Directorate, Abels Gate 9, 7030 Trondheim, Norway*

## Abstract

The unstable rock slope Stampa is located north-east of the touristic town of Flåm, Norway, above the Aurlandfjord and displays signs of post-glacial deformation over a large area encompassing a volume of several million $m^3$. Directly below the rock slope lies the European Road E16, a highly frequented connection between Bergen and Oslo. One identified unstable object, named Block 4a, of approximately $5,000m^3$ sits on a highly fractured base of approximately $40,000m^3$ and showed displacement rates exceeding $1cm$ per day in autmun 2022. Different failure scenarios and potential secondary events threaten the European Road at the foot of the slope and potentially also Flåm. An on-site sensor network with a range of in-situ instruments collects data since 2019. We combine this data with remote sensing data as well as weather station data from Stampa. Sensor Data Fusion is used to merge data of the different sensors and exploit the different sensor resolutions. This adds to the development of a data set with high spatiotemporal resolution capturing the physical properties of Block 4a, from which we derive 3-D displacement and the respective translation and rotation components. As the data set comprises complementary, redundant as well as cooperative data, Data Fusion results in a more accurate state estimation of Block 4a. Based on this data set we use explainable Artificial Intelligence to derive a model of the object, offering a better understanding of the mechanisms that drive the movement and failure of Block 4a. Moreover, the model can be used for short term predictions of object movement, incorporating weather forecast data for the area. Furthermore, a monitoring system will be developed using the accurate state information such as displacement, velocity and acceleration as well as inclination angle of the object, to support a human decision maker regarding evacuations for example. Combined with the short-term forecast, auto-adapting failure threshold values, actuators and near real-time communication, the system can further be extended towards an early-warning system and applied to similar unstable objects elsewhere. Hence, the automation of the data fusion process at low level can replace the common practice of manual information fusion on a higher level and contribute to decision making, supporting a transition towards safe, collaborative and intelligent early-warning systems.

## Keywords

Data Science, Sensor Fusion, Unstable Rock Slope, Monitoring, Early-Warning, Natural Hazards

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: lukas.schild@hvl.no

# Short-term Inflow Forecasting for a Use case in Western Norway Using Causal variational decomposition

Mojtaba Yousefi[1], Jinghao Wang[2], Jayaprakash Rajasekharan[2], Hossein Farahmand[2] and Reza Arghandeh[1]

[1]*Department of Computer Science, Electrical Engineering and Mathematical Sciences, Western Norway University of Applied Science Bergen, Norway*

[2]*Department of Electric Power Engineering, Norwegian University of Science and Technology Trondheim, Norway*

**Abstract**

Multi-step inflow forecasting is essential for hydropower producers to improve scheduling problems and meet the environmental requirement. In this paper, a new preprocessing framework using variational mode decomposition, causality feature selection and recursive forecasting techniques (CVD) is developed to forecast 24 hours ahead of inflow. Firstly, physically meaningful components (Modes) are provided using variational mode decomposition which decomposes time series. Then Causality feature selection identifies a subset of Modes with corresponding lag values which have significant contributions to the next state of inflow. Then, the selected subset Modes with corresponding lag values are used for training different machine learning algorithms to validate the performance of the developed CVD framework. The CVD framework is validated and tested to forecast 24 hourse ahead of inflow for a use case in western Norway.The simulation resultes shows that The CVD framework reducing Normalized root mean error of inflow forecasting by 25% when it is used with LSTM compared with a stand-alone LSTM when

**Keywords**

Deep learning, Causality, Data preprocessing, Spatiotemporal, Decomposition

## 1. Use case

The data is provided by Lyse company in Norway. The location of the use case is in Rogaland the southwest of Norway. The Lyse asset included Lyseboten I and II which are two power stations that are connected to three main reservoirs Breiavatnet (Bri), Lyngsvatnet (Lyn), and Strandvatnet (Str). In this use case river's stream is controlled by the bypass flow from the Bri reservoir to always meet minimum inflow at the NVE measurement station. The NVE measurement station is placed 20 Km away from Bri reservoir.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway
EMAIL: Myou@hvl.no (A.1)

## 2. Methodology

In this paper, a Causal variational Decomposition preprocessing architecture is developed. Variational mode decomposition technique is employed to remove noise and reduces the complexity of the time series and causality analysis is utilized to remove the irrelevant features of inflow by finding the most informative features and their corresponding lag values. CVD has three main modules. The first module decomposes the input data (meteorological and hydrological data related to the water inflow) into different Modes. The second module selects the most informative decomposed Modes related to the water inflow with proper time latency values. In the third module, only the selected latency values of decomposed Modes are used for training a multi-step time series forecasting algorithm. The contribution of this paper is presented below:

- A causal variational model (CVD) decomposition preprocessing model is developed to find the most informative variables among weather data, simulated and observed hydrological data for a use case in Western Norway to forecast multi-steps ahead of inflow.
- Developed CVD preprocessing module presents prominent results in improving short- term inflow forecast

## 3. Results

To evaluate the role of CVD as a pre-proccessing feature selection framework in improving the LSTM forecasting performance, comparision results have been presented in Table 1. The forecast performance, the Normalized Root Mean Square Error(NRMES) and computational time is compared with a stand-alone LSTM for four different scenarios.The first observation is that by comparing only LSTM with CVD-LSTM across four scenarios, the NRMSE error reduces significantly. For example, there is a 25% improvement inside scenario 4 when CVD is added to LSTM.

**Table 1**
Comparison of input data impact on LSTM and CVD-LSTM performance.

| Senarios | Data | Model | period | NRMSE | Computational time (s) |
|---|---|---|---|---|---|
| 1 | Historic inflow | LSTM | t+24 | 1.7 | 547 |
| 2 | Weather | LSTM | t+24 | 1.66 | 442 |
| | | CVD-LSTM | | 1.03 | 80 |
| 3 | Weather+ hydrological data | LSTM | t+24 | 1.06 | 629 |
| | | CVD-LSTM | | 0.8 | 96 |
| 4 | Weather+ hydrological+ HBV data | LSTM | t+24 | 0.68 | 900 |
| | | CVD-LSTM | | **0.51** | **76** |

# Collaboration approach for improving data readiness in SMEs*

B. C. Weinbach[1†], Malin Waage[1,*,†]

[1]Western Norway Research Institute, Norway

Abstract

One of our main goals is improving data economy readiness in regional companies in Norway. As part of our initiative, a collaboration was initiated with a local power company on inflow prediction. Hydrology is a complex topic and there is interest in investigating how accurately domain knowledge in hydrology and machine learning can be used to improve reporting of production capacity to power exchange markets. Our contribution is an approach on how researchers can help companies get started with machine learning. It borrows practices from agile software development methods.

**Keywords**

Machine Learning, Software Engineering, Power Production, Project Management

## 1. Background

We are collaborating with a local energy company, to conduct a case study on transitioning from a traditional, experience-based reporting of day-ahead power generation capability to a data-driven prediction based on past and future weather forecast and runoff data. Accurate predictions of power production are essential in addressing the current energy crisis. If successful, this case study will demonstrate that by planning the future power generation more accurately and efficiently. According to Beigaite et al. [1], one can increase the exploitation of the electric energy leading to increased profits and reduced risks. Here, we stress on the work method and software development.

The power plant in our test study is in a sub-polar region, and heavily dependent on snow melt, and hence, the runoff is affected by measures such as temperature, sunlight, wind, and rainfall in the entire runoff region. With so many factors at play, predicting the power plant's output for the following day based on a glance at the weather forecast is a difficult task, making it essential to find better forecasting methods, which is why we were contacted by the company.

Proceedings of SAIDD 2023, May 09–10, 2023, Sogndal, Norway

*Corresponding author.

†These authors contributed equally.

EMAIL: bcw@vestforsk.no (B. C. Weinbach); mwa@vestforsk.no (M. Waage)

ORCHID: 0000-0002-9812-1008 (B. C. Weinbach); 0000-0002-6924-2273 (M. Waage)

The local energy company, like many other SMEs [2], lacks required knowledge to enter the data economy. Having a few software engineers and few with a formal data science or machine learning background makes it challenging to get started. We propose an approach that is meant to ensure that data science researchers develop software on the companies' premises while also addressing researchers need for publications.

## 2. Software development methodology

According to Sommerville [3], software has in many cases been developed in a waterfall model. The waterfall model is very intuitive and compatible with other engineering disciplines. Software is developed in sequential steps until completion, then distributed to a customer. Requirements that is uncovered after release is reported in and developed in iteration 2.

The main criticism of this method is that one does not see the need for functionality until after the product has been delivered and entered production. For projects between researchers and software engineers the research environment produces an algorithm in a framework with good results but when released to software engineers the software is incompatible with their frameworks and needs. Our contribution is an approach that embraces agile methods. Our method can be described as follows:

1. **Less is more:** Prioritise development of small deliverable. Don't focus on the big picture and the advanced methods one wishes to create. This will come.
2. **Develop software on the user's premises:** Develop and deploy scientific software in the users framework. Use feedback on deliverables to inform decision about future deliverables. The less difference there is between the code that is written for the scientific publication and the code that is used in production, the bigger the impact of the code.
3. **Think quantitatively:** Define research questions while developing and see how deliverables and their effects can be quantified. Such that statistics and data can drive decision making. Additionally, collection of empirical data on the development process and the corresponding impacts make these experiences interesting for publication.

## 3. Benefits of the approach

One of the most important benefits of this approach is that it balances the different interests off the partners in research collaboration between researchers and companies that want to start using machine learning.

**The company's main interest is to get software to solve a particular problem which is economically viable**. This need has to be addressed by the researcher. Which is why there is emphasis on developing the software on the company's premises. This forces researchers to solve problems that are relevant in industry which may also raise new research questions and hypotheses.

**The researchers' main interest is developing a novel methodology that is publishable**. In machine learning research data is the driver of new methodologies. Researchers often have the necessary theoretical understanding to start development of new novel methods but lack

the access to data sources. Industry find themselves in the opposite situation, where data is in excess and lack necessary theoretical knowledge.

**The dynamic above is reinforced with SMEs**. Larger companies are able to limit the difference in theoretical knowledge and available data by employing machine learning engineers on demand and even having their own research and development department. For SMEs, which is more prominent in the region of Sogn og Fjordane than in the rest of Norway, this is not possible. The collaboration approach is therefore a necessity in ensuring SMEs capability of adjusting to the fourth industrial revolution.

**The collaboration approach stimulates knowledge transfer**. Norway experiences a shortage in ICT workers and many SMEs are not able to employ the necessary workforce in transitioning to the data economy. Workers that do not have formal a formal computer science background are set to solve tasks that is outside their comfort zone.

## 4. Conclusion

The collaboration methodology used should be further tested and improved on through feedback with companies. In the future, this approach might improve collaboration between researchers and regional companies for addressing data readiness in companies as well as giving researchers publishable results.

## References


[1] R. Beigaite, T. Krilavičius, K. L. Man, Electricity price forecasting for nord pool data, in: 2018 International Conference on Platform Technology and Service (PlatCon), 2018, pp. 1–6. doi:10.1109/PlatCon.2018.8472762.

[2] S. Mittal, M. A. Khan, J. K. Purohit, K. Menon, D. Romero, T. Wuest, A smart manufacturing adoption framework for smes, International Journal of Production Research 58 (2020) 1555–1573. doi:10.1080/00207543.2019.1661540.

[3] I. Sommerville, Software Engineering, 9/E, Pearson Education India, 2011.

ORGANIZED BY
THE INTPART DTRF PROJECT

SUPPORTED BY
WESTERN NORWAY RESEARCH INSTITUTE AND
THE RESEARCH COUNCILOF NORWAY

ISBN 978-82-428-0457-0