Analytical Solutions and Reasoning
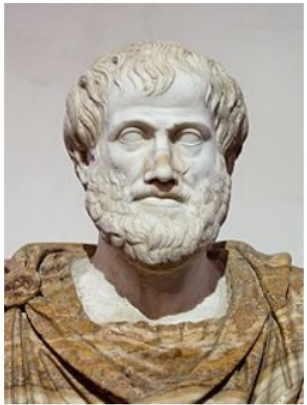
MSc topics 2023

Universitetet i Oslo

29 September 2023
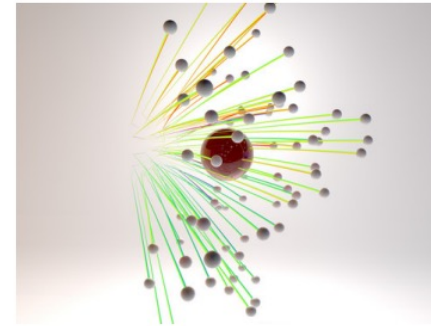
ASR – Analytical Solutions and Reasoning
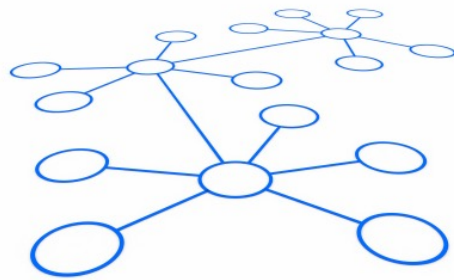

Modeling

Simulation

Logic, Reasoning

Semantic Web Technologies

Model-driven Systems

Databases

ASR **MSc topics**

ASR:

`https://www.mn.uio.no/ifi/studier/masteroppgaver/asr/`

Topic areas

- Machine Learning, Data Science
- Ontology-Based Information Systems
- Modeling and Analysis
- Data Streams and Applications
- Logic and Theoretical Computer Science

Digital Hospital Ward

Contact: Einar Broch Johnsen (einarj@ifi.uio.no) and
Lizeth Tapia Tarifa (sltarifa@ifi.uio.no)

Students will work on a digital twin of a single hospital ward, the location with be at OUS or Ahus.

Students will work towards efficiency of resources within healthcare.

Projects will focus on building the data infrastructure and framework needed for modelling to deal with predictive scenarios, for example, time in the ward, time to discharge, changes in personnel/staffing, and introduction of new technologies and innovations in treatment.

Marine Digital Twins

We are looking for students to work on topics related to biodiversity monitoring data in the Digital Twin of the Ocean. New technology (imaging, acoustics, DNA based, satellite) produce huge amounts of readings that require processing, aggregation and translation into useable data products such as species occurrences, densities, biovolumes, species migration routes, etc. In addition, data can come from sensors in sea monitoring areas, or attached to sea vessels. The masters student defines a project around incoming data streams connected to Digital Twin simulation frameworks.

- Integrate biodiversity data and translate to data products for marine digital twins
  - Projects are related to the EU Digital Twin of the Ocean or the Oslofjord project.
- Creating a digital replica of marine ecological processes

Contact: Einar Broch Johnsen (einarj@ifi.uio.no) and Lizeth Tapia Tarifa (sltarifa@ifi.uio.no)
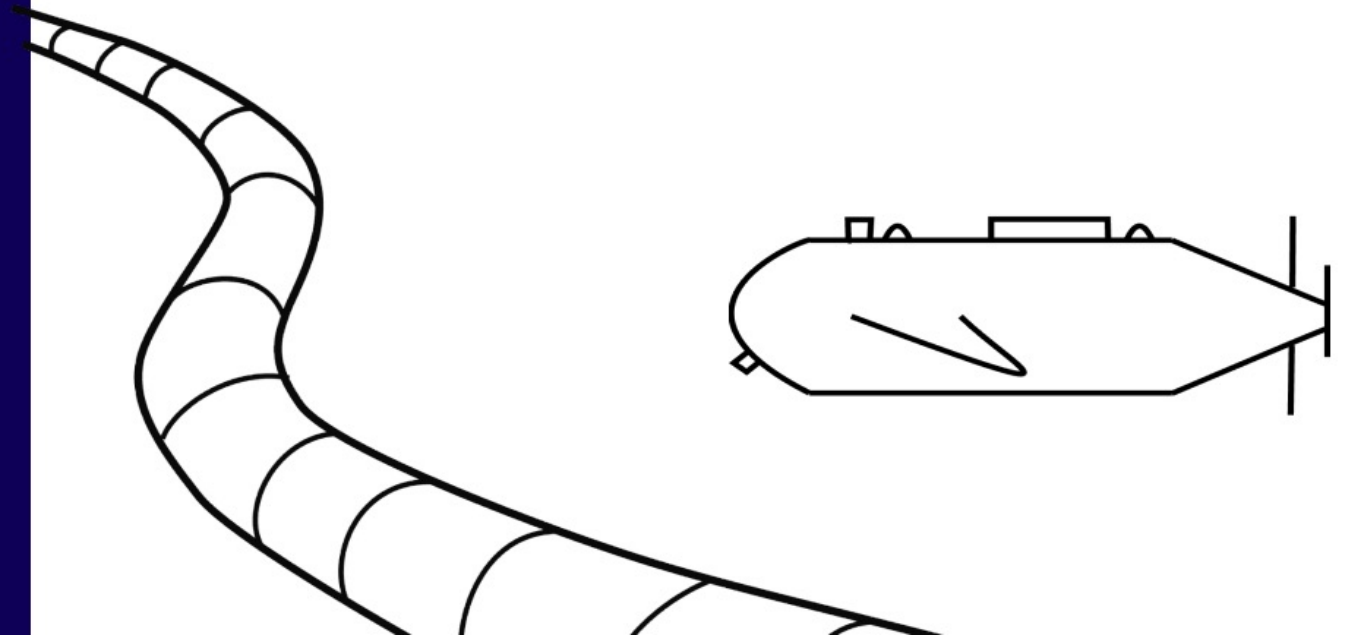
Modelling and Analysis of a Self-Adaptive Underwater Robot
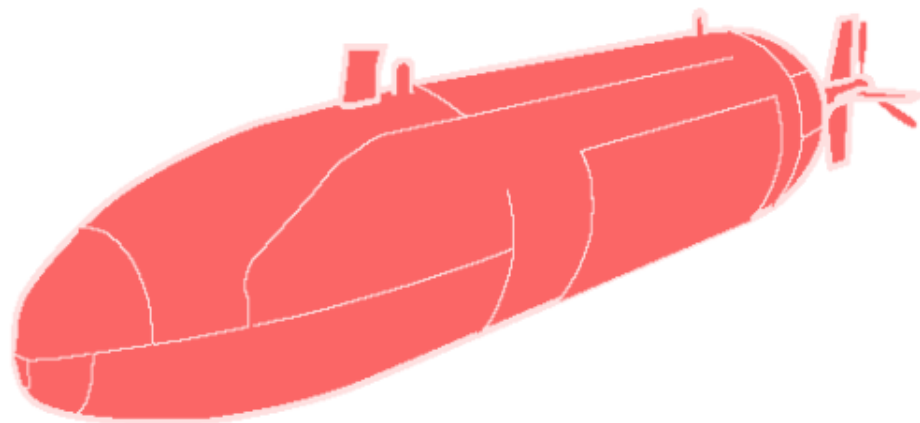
Juliane Päßler

julipas@uio.no

You will learn about

- self-adaptive underwater robots

- modelling self-adaptive systems

- analysing self-adative systems

Cognitive Robot Control

an AI system for underwater robots

Website for Topic

Tobias John

tobiajoh@ifi.uio.no

Involves:

- Design control program

- Design knowledge base

- Implementation and testing of robustness
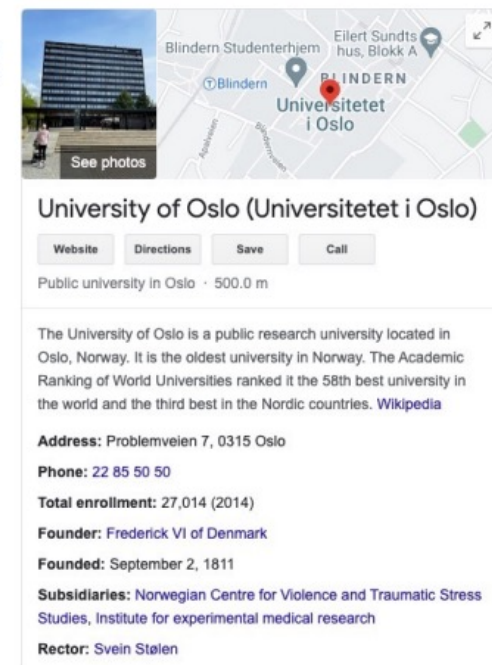
Ontology Summarization via Machine Learning Techniques

- **Problems**
  - Ontologies are getting bigger
  - Limited capacities of human user to understand contents
- **Goal**: extract top-k most relevant axioms that related to users' interests
  - Combine with machine learning techniques to extract relevant information
  - Apply tools on real-world ontologies/knowledge graphs
- **Relevant courses**:
  - IN3060/IN4060 Semantic Technologies
  - IN3050/IN4050 Artificial Intelligence and Machine Learning
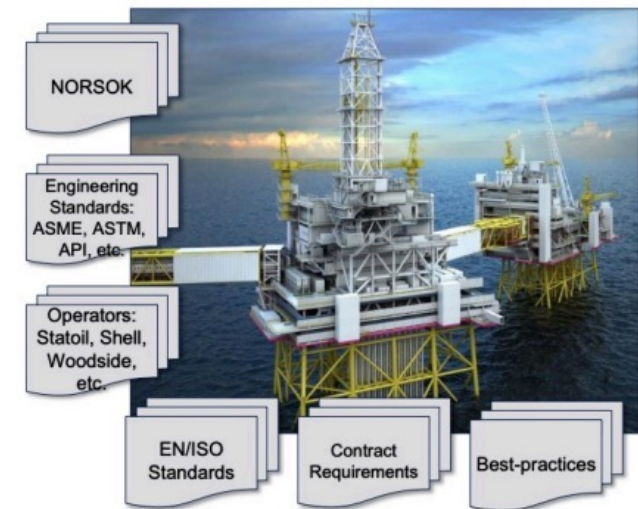- **Supervisors**: Jieying Chen (jieyingc@ifi.uio.no)

University of Oslo (Universitetet i Oslo)

Website   Directions   Save   Call

Public university in Oslo · 500.0 m

The University of Oslo is a public research university located in Oslo, Norway. It is the oldest university in Norway. The Academic Ranking of World Universities ranked it the 58th best university in the world and the third best in the Nordic countries. Wikipedia

Address: Problemveien 7, 0315 Oslo

Phone: 22 85 50 50

Total enrollment: 27,014 (2014)

Founder: Frederick VI of Denmark

Founded: September 2, 1811

Subsidaries: Norwegian Centre for Violence and Traumatic Stress Studies, Institute for experimental medical research

Rector: Svein Stølen

sfi Centre for Research-based Innovation
The Research Council of Norway

Knowledge Extraction from Large Scale Ontologies with Aibel

aibel®

- **Aibel**: leading service company within the oil, gas and offshore wind industries

- **MMD ontologies**: in OWL 2.0, > 1.8m axioms

- **Pains**
  - Poor reasoning performance
  - Difficult to maintain

- **Goal**: exact relevant knowledge but still maintain logical guarantee
  - Apply theorical results in industry
  - Design a framework for MMD
  - More research possibilities

  Supervisor: Jieying Chen (jieyingc@ifi.uio.no)

NORSOK

Engineering Standards: ASME, ASTM, API, etc.

Operators: Statoil, Shell, Woodside, etc.

EN/ISO Standards

Contract Requirements

Best-practices

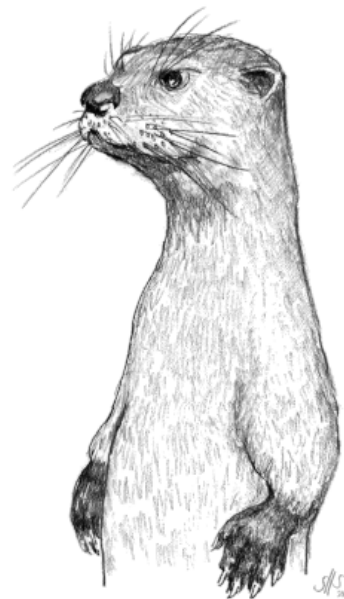# Ontology Diagnosis: Computing Justifications for OWL Ontologies

- **Problems**
  - Increasing number of large-scale ontologies are being developed
  - Difficult to debug ontologies and reuse their knowledge
- **Goal**: extract relevant ontology part to explain why certain conclusions follow from the ontology (axiom pinpointing / justification)
  - Extend JUST to deal with range restriction by using an extended calculus
  - Apply tools on real-world ontologies, e.g., MMD ontology from Aibel, SNOMED ontology
- **Relevant courses**:
  - IN3060/IN4060 Semantic Technologies
  - IN3070 /IN4060 Logikk
- **Supervisors**: Jieying Chen (jieyingc@ifi.uio.no)

    Martin Giese (martingi@ifi.uio.no)

Implement OTTR with sem.tech.-language

Supervisor: Martin Skjæveland/Leif Harald Karlsen – `martige@ifi.uio.no/leifhka@ifi.uio.no`

- ◆ OTTR templates are macros for RDF and OWL
- ◆ Currently implemented in Java
- ◆ Both templates, instances and expansion could live in same triplestore
- ◆ Lots of languages supported by triplestores (SPARQL, SHACL, rules...)
- ◆ <u>Thesis:</u> Investigate ways of implementing expansion of templates in these languages

```
:Pizza[owl:Class ?pizza, xsd:string ?label] :: {
    o-rdf:Type(?pizza, owl:Class),
    o-rdfs:SubClassOf(?pizza, :NamedPizza),
    o-rdfs:Label(?pizza, ?label)
} .
```

```
:Pizza(ex:margherita, "Margherita") .
:Pizza(ex:grandiosa, "Grandiosa") .
```

⇵

```
ex:margherita a owl:Class ;
    rdfs:subClassOf :NamedPizza ;
    rdfs:label "Margherita" .
ex:grandiosa a owl:Class ;
    rdfs:subClassOf :NamedPizza ;
    rdfs:label "Grandiosa" .
```

Angluin's exact learning framework: Repurposing

Horn Logic Algorithm

Queries

Answers

Learner

Teacher (Oracle)
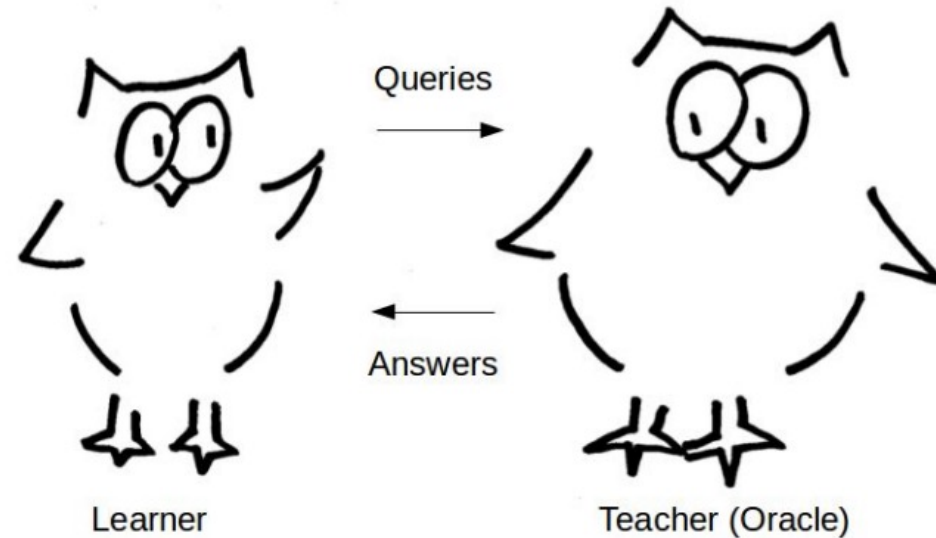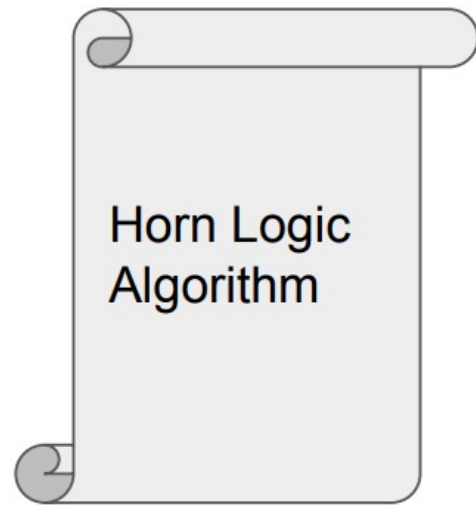
Detection of Harmful Biases

Language Model

[Learning Horn envelopes via queries from language models (IJAR 2023)](#)

S. Blum, R. Koudijs, A. Ozaki, S. Touileb

Supervisors:
Ana Ozaki
Ingrid Chieh Yu

Angluin's exact learning framework: Repurposing

Ontology Learning

Algorithm for EL Ontologies
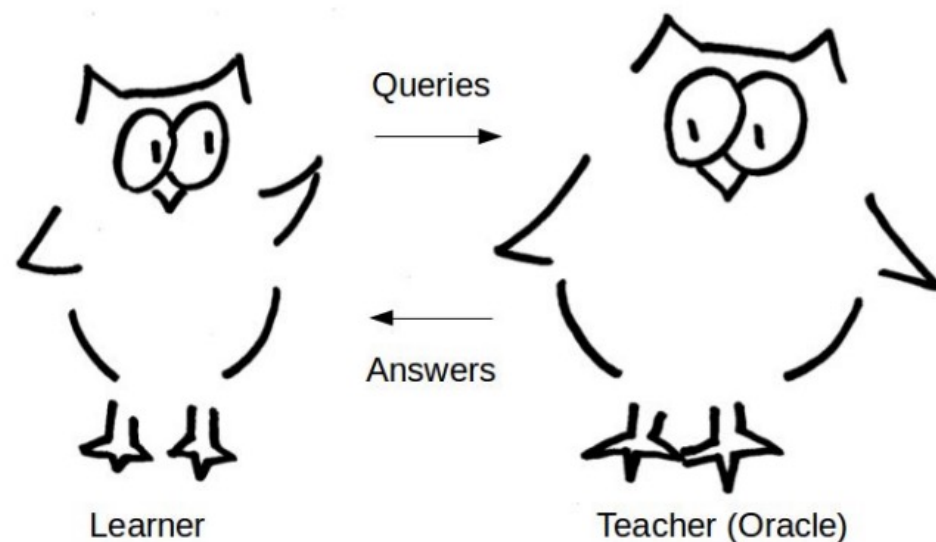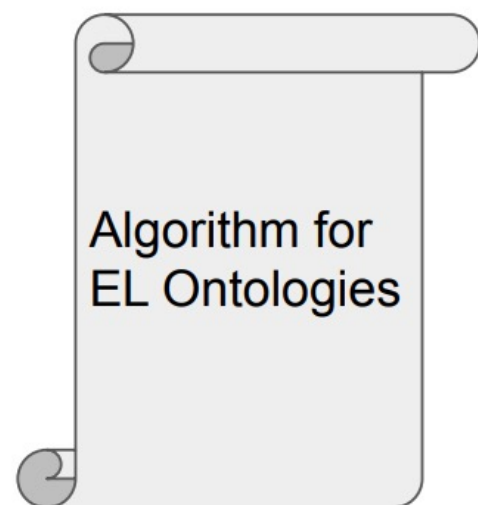

Queries →

← Answers

Learner   Teacher (Oracle)

Language Model

Supervisors:
Ana Ozaki
Ingrid Chieh Yu

Learning Horn envelopes via queries from language models (IJAR 2023)

S. Blum, R. Koudijs, A. Ozaki, S. Touileb

Learning / Handling Moral Preferences

Research Questions

- Can we learn (good) preference models?

- How should we best aggregate these preferences to a societal preference?

- ... [your own idea] ...
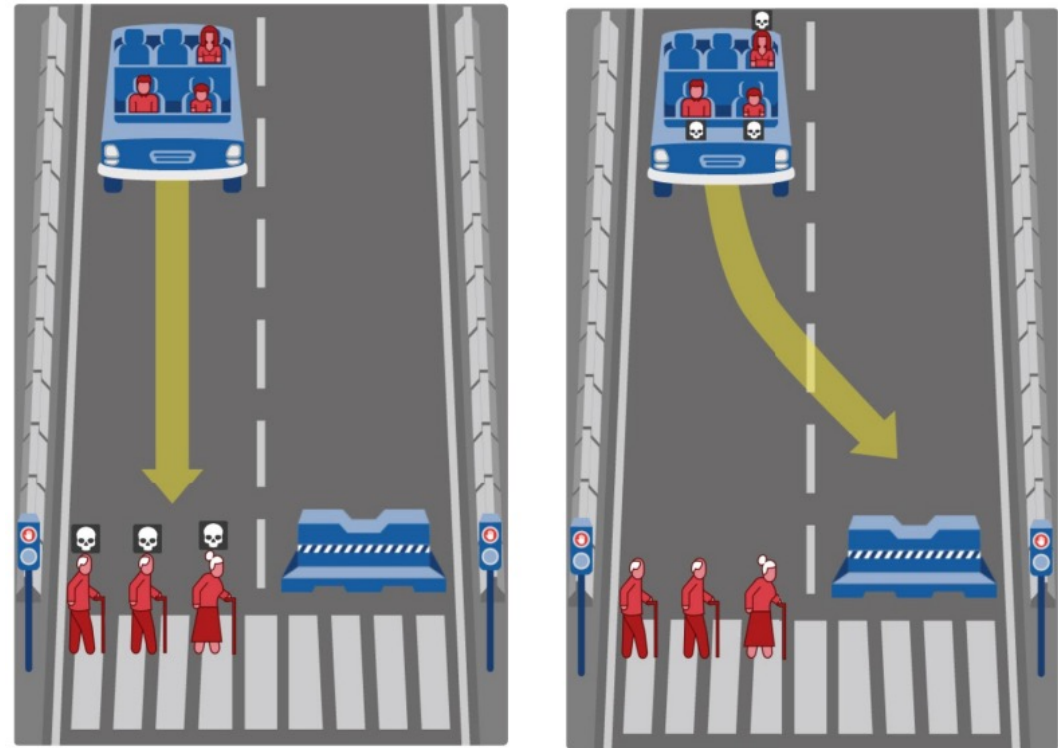
Interested? Write to:

Anne-Marie George
Postdoc
ASR Group, Ifi
annemage@uio.no

Ana Ozaki
Associate Prof.
ASR Group, Ifi
anaoz@uio.no

Who would you rather kill?

Source/ Good Read: The Moral Machine experiment (mit.edu)

Dynamic link prediction and time prediction on Temporal Knowledge Graphs

Standard (static) KG fact: (Obama, president, US)

Temporal KG fact: (Obama, president, US, [2009, 2017])

Tasks:

- **Dynamic link prediction**

  (?, president, US, [2009, 2017])

  (Obama, president, ?, [2009, 2017])

- **Time prediction**

  (Obama, president, US, ?)

(visit, 01-01-2014)
(investigate, 01-29-2014)
(visit, 10-03-2014)
(veto, 05-14-2014)
(sign formal agreement, 01-15-2014)
(sign formal agreement, 10-12-2014)
(consult, 02-16-2014)
(criticize, 03-01-2014)
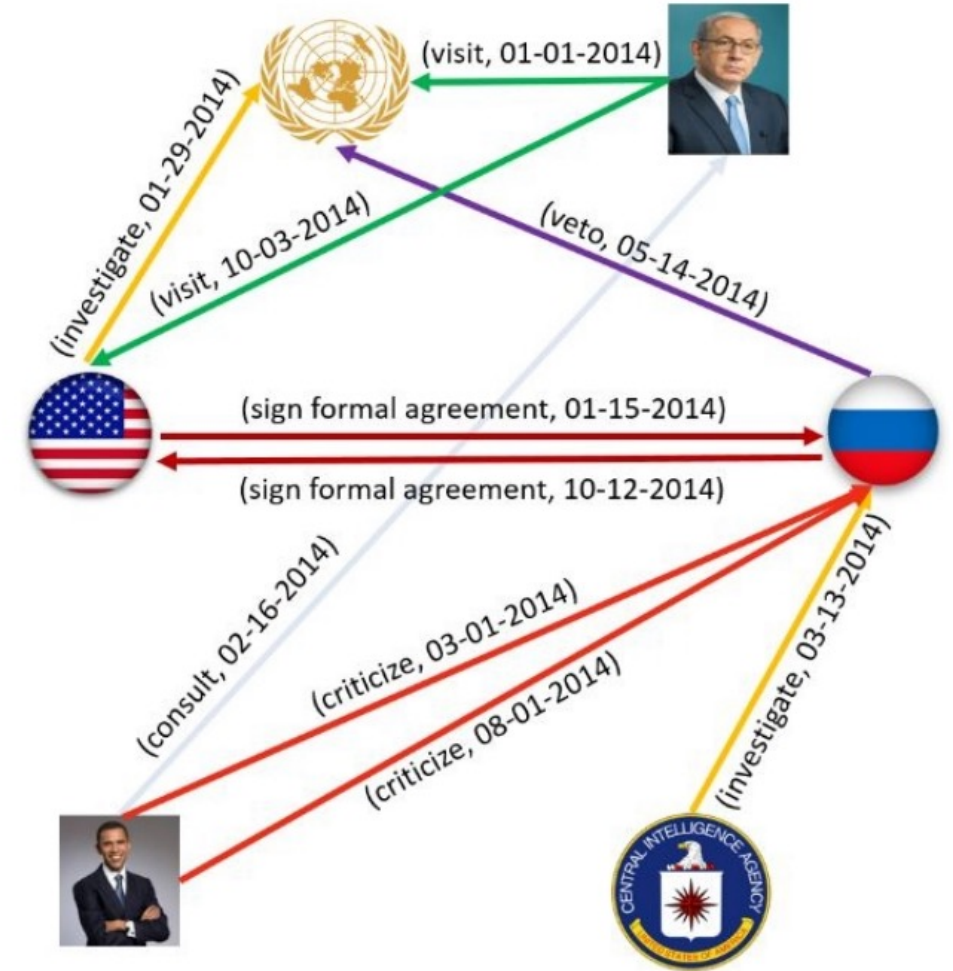(criticize, 08-01-2014)
(investigate, 03-13-2014)

Image taken from here

Supervisors: Egor V. Kostylev (egork@ifi.uio.no), Roxana Pop (roxanap@ifi.uio.no)

How trainable are GNNs?

(Comparing trainability of Graph Neural Networks for logic-expressible functions)

- Graph Neural Networks (GNNs):
  - a family of modern NN architectures
  - work with structured data (e.g., Knowledge Graphs)

- (Theoretical) expressivity of GNNs:
  - different GNN architectures express different sets of functions on graphs
  - well-understood by connecting to logics

- (Practical) trainability of GNNs:
  - "can express in theory" does not imply "can be trained"
  - Question:
    Which GNNs architectures can be trained for which logic-expressed functions?
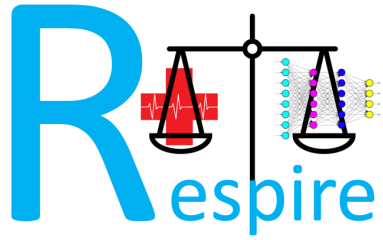
Supervisors: Egor V. Kostylev (egork@ifi.uio.no), Roxana Pop (roxanap@ifi.uio.no)
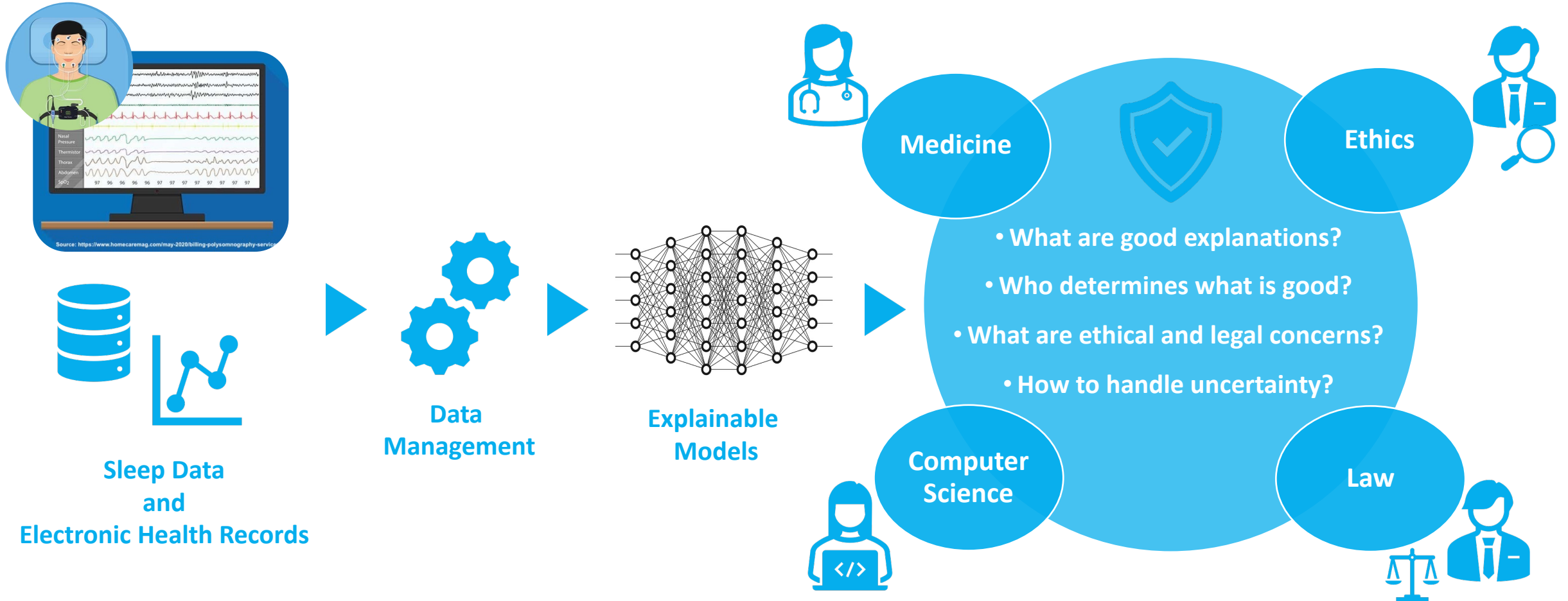
# Knowledge Graph for Sleep-related Respiratory Disorders

In the medical domain, Knowledge Graphs (KGs) have many applications such as structuring research knowledge and integrating it with study results to help in reliable diagnostics.

In the field of Sleep-related Respiratory Disorders, a KG would be highly beneficial for domain knowledge representation and integration.

Medical Knowledge Graph

https://doi.org/10.1016/j.bdr.2020.100174

Tasks involved in the development will include:

1. Existing medical ontologies will have to be analysed and validated with domain experts.

2. Medical knowledge needs preprocessing into a graph structure reflecting the ontology (NLP techniques might be necessary).

3. Appropriate databases and tools will be required to store and query the graph and other relevant data and metadata (Neo4j, PostgreSQL, Apache Airflow, …)

4. Accessing and visualising the graph will be handled through microservices and off-the-shelf applications like Apache Zeppelin.

https://www.mn.uio.no/ifi/studier/masteroppgaver/asr/knowledge-graph-for-sleep-related-respiratory-diso.html
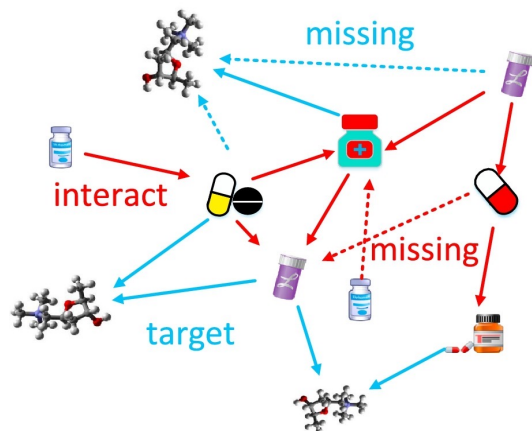
Identifying Respirations with Smart Sleeping Monitor and ML

- Human identification by measuring respiration patterns using Somnofy (non-contact smart sleep monitor)
- Analyze the respiration curves

Detect Obstructive Sleep Apnea (OSA) in infants

SOMNOFY

https://www.mn.uio.no/ifi/english/research/projects/respire/

Detecting Respiratory Events with Simulations and ML

- Apply Machine learning on the data stored by the mechanical ventilator to detect respiratory events from the simulation model

RespiSim® System

Vital Signs Monitor and Virtual Patient Chart (Learner View)

Instructor View

Ventilator

RespiPatient

Lungs (ASL 5000)

Mobile Cart

ASL 5000

Valve for Unintentional leak

Valve for intentional leak

Flowsensors

A

Ventilator

B

ASL 5000 Lung Model

Optional placement for flowsensor A

Optional restistance
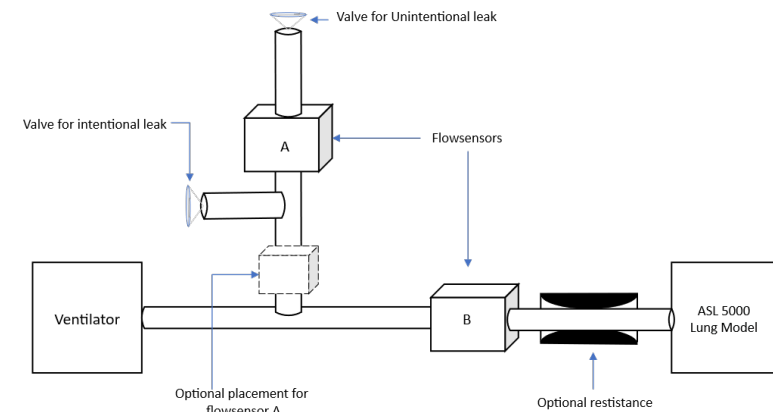
Parrot project: Privacy protection in data stream processing

Use Case

Marvin — GPS — entirely / area / during job / emergency → Family, Friends, Job, Health Service

System View

Data Producer
Data Consumer

Privacy Requirements (PR) | GPS Location | Data Utility needs | Select Marvin's location; | Location (exact, coarse, ...)

Stream Processing System (SPS)

Privacy Preserving Mechanisms (PPM)

Parrot

Data Provider → PR
Data Consumer → DU, Query Q

Parrot Magic?

PPM → SPS

```
if (PR ≤ DU){
    find PPM with PR ≤ PPM ≤ DU
    QPP := apply(Q, PPM)
}else{
    print("Sorry mate can't help you")
}
```
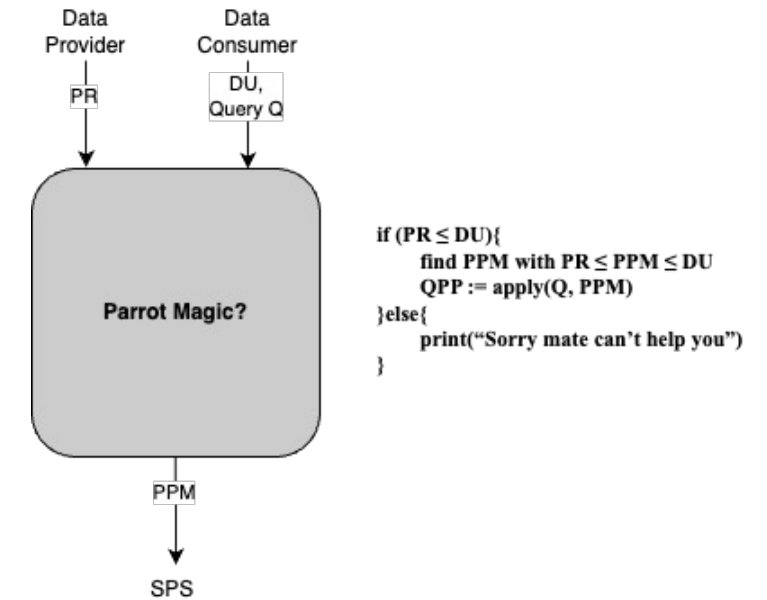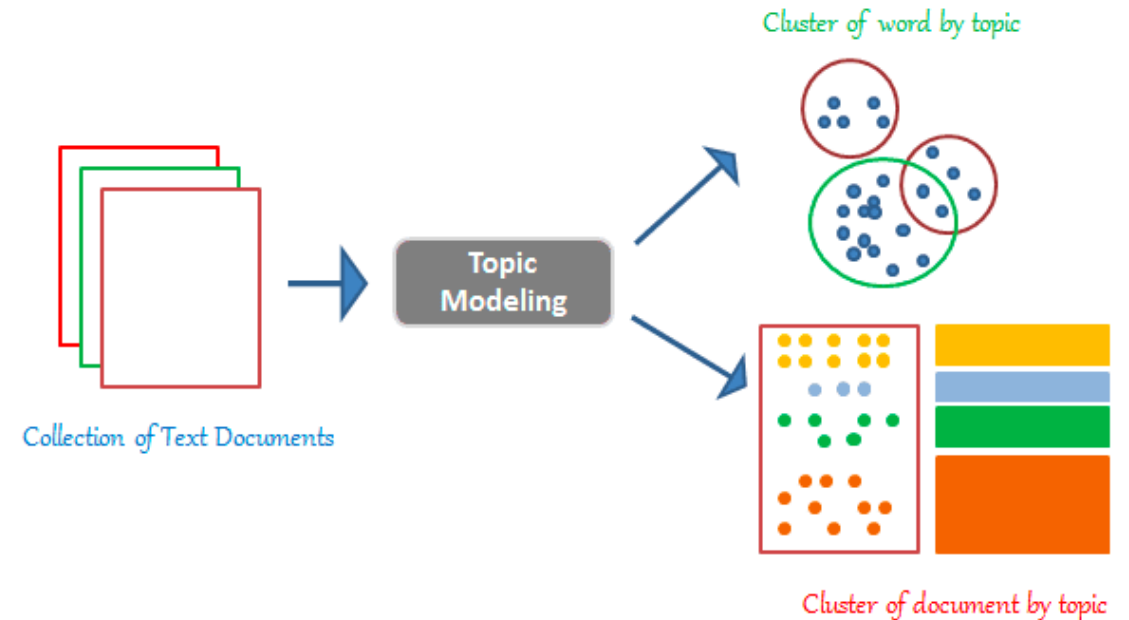
• Challenges: Understanding PR, Sensor information, PPMs and DU

What do we commit to? Using ML to understand Privacy Policies of Health Related IoT Devices

- Do we know what we commit to when signing off contracts for smart watches, cameras, wearable sensors, smart hubs etc.?

- Time to change that – lets understand what are we signing off to

- Use topic modelling to find topics within these privacy policies

- Use Pre-trained privacy policy classifiers to understand what is being discussed in these policies

- Can we use LLMs (Large Language Models) for summarizing these privacy policies? Making them more user friendly

- In short – perform a wide spectrum analysis of these privacy policies

Cluster of word by topic

Collection of Text Documents

Topic Modeling

Cluster of document by topic

Ontology of Information that can be Inferred from Sensors

Tasks:

1. You will use literature, source code repositories, etc., to determine which information can be derived from a sensor.

2. You will create an ontology from this information.

3. Finally, we will pick out some examples and try to prototypically implement them.

You should have some knowledge about ontologies, the Android platform and have some interest or experience with Sensors.

Thesis Announcement
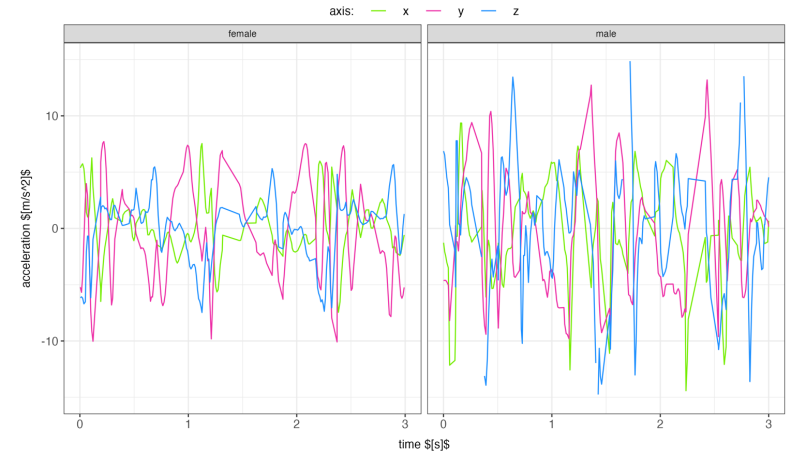
Analysing Data Utility needs for Apps on Health, Sport,…

Tasks:

1. Together with you, we will select for example 3 applications for the analysis.

2. You will determine the requirements of the selected applications, e.g. based on technical specifications and documentation.

3. You will create a testbed with which the privacy and utility of an application can be evaluated under different condition.

4. You will use the testbed to assess the selected apps and analyse the outcome.

You should have experience with the Android platform, have interest or experience with Sensors, in data privacy and PPMs.

Thesis Announcement

Ontology of Privacy-Preserving Mechanisms for Stream Processing

Tasks:

1. You will determine characteristics of PPMs, such as the supported data type, the achievable privacy level and the utility.

2. Based on these knowledge, you will analyse existing PPMs and create an ontology to represent this information.

You should have some knowledge about ontologies, have interest or experience in SPSs, data privacy and PPMs.

Thesis Announcement

Learning and Differential Privacy

How to infer parameters from privatised observations?

- It is known that local differential privacy results in slightly slower estimation rates
- However, optimal inference methods for such data have not yet been developed
- This thesis will focus on advanced statistical inference methods for efficient inference under privacy constraints

- Knowledge gained/required: statistics, optimisation, differential privacy

Supervisor: Christos Dimitrakakis <chridimi@ifi.uio.no>
Webpage: https://sites.google.com/site/christosdimitrakakis

About me (Knut Omang)

- Ph.D from Ifi - high performance networks, 'IIer' since then

- Scali (→ Platform Computing): "Affordable supercomputing"

- Fast (→ Microsoft)

  - Search engines: My part: search within documents = "teasers/summaries"

- Paradial (→ Logitech): Firewall traversal for realtime multimedia

- Oracle: Driver/kernel and Qemu (virtualization) development

- From Feb'20: Properate AS

  - Cloud based monitoring platform for commercial buildings

  - Monitors > 550 buildings, 1.6M m2 - all over Norway

  - Built on top of Cognite Data Fusion

PROPERATE 1

Thesis areas

- Challenges related to data gathering and presentation of data from millions of sensors
  - feed from many types of sensor hardware
  - data modeling, ontologies
  - creative new user interfaces/presentations
- Cloud/microservices related
  - containers, Kubernetes
  - Web services, REST APIs, protocols, functionality, security
  - Virtual network technologies
- Virtualization using QEMU
  - Innovations in QEMU, explore and understand features
  - Device emulation, passthrough/SR-iov, paravirtualization, nested virtualization
- Linux kernel development/test related topics
  - Kernel support for containers and VMs
  - Experiment with new APIs, options for building

https://properate.com

Data Science and Engineering Master Thesis With Equinor

equinor

Dirk - Your Main Contact

- Vice President Marketing and Supply Solutions, Equinor
  - M&S: Digital solutions for sales and distribution of energy, from algo trading to invoicing and ship chartering
- Associate Professor at UiO since 5 years
- Co-teaching IN-STK 5000/9000
- Past master thesis supervised/supervising
  - Machine learning for cyber intrusion detection (ongoing)
  - Distributed energy systems modelling (ongoing)
  - Deep Policy Gradient Methods in Commodity Markets
  - Gender prediction on Norwegian Twitter accounts
- Contact: dirkh@uio.no, dirh@equinor.com

Application of advanced anomaly detection in data monitoring

- Equinor is using advanced anomaly detection methods in a tool called <u>Omnia.Prevent</u> to detect potential early signs of failures in industrial equipment
  - Methods such as separation forests and autoencoders are used
  - The system is built for scalability. Today, Prevent is analysing data from more than 18,000 sensors spread across 30+ oil and gas installations offshore and onshore – with more to come.
- At the same time, Equinor is building a data platform that aims at making high-quality data available to many users across business units to become more data driven. This is essential for Equinor to achieve its digital ambitions.
  - To ensure the users can trust the data, its quality must be known.
  - Most data gets continuously updated (such as production data, commodity prices, etc.)
  - This requires continuous data monitoring
- The topic of this thesis would be to investigate if Omnia.Prevent and the methods used therein can be suitable to automatically monitor data at scale and alert data engineers and users of anomalies to be investigated.

Development of a Comprehensive Systematic Trading Tool for Gas Trading

- The issue
  - Gas market trading dynamics are becoming more complex, and traditional manual methods of trading are no longer sufficient.
  - Systematic trading, which relies on mathematical models, predefined rules, and algorithms, has gained significant attention in recent years for providing consistent and profitable trading opportunities.
  - This master's thesis aims to develop an advanced systematic trading tool for energy trading that can significantly support decision-making processes and risk management, ultimately increasing profitability and efficiency.
- The task
  - Conduct a thorough review of existing systematic trading techniques.
  - Identify the key features and requirements necessary for building an effective systematic trading tool for the gas market, considering data sources, analytics capabilities, and risk management functionalities.
  - Design and develop a comprehensive and user-friendly software tool that incorporates real-time data integration including weather pattern, Global supply demands, geopolitics events…etc., advanced financial technical analysis, and algorithmic trading strategies.
  - Provide financial insights and metrics to each trading strategy including risk management capabilities, and profitability through back-testing.

Unleash the power of our data platform using generative AI

- Generative AI methods, such as ChatGPT, are capturing people's attention and imagination, providing powerful tools for broad use across industries and fields.

- Equinor is building a data platform that aims at making high-quality data available to many users across business units to become more data driven. This is essential for Equinor to achieve its digital ambitions.
  - Often, however, specialist tools and knowledge is required to access this data and extract insight.
  - These tools range from Excel, to SQL and Python, as well as low-code tools such as Power BI

- The vision of this project would be to provide a chat-like user interface, potentially embedded in everyday collaboration tools such as MS Teams, to get insight into our vast collection of data on the spot.

- We would start with investigating the use of GenAI to identify the right dataset in our large and growing collection to answer a plain-text question and estimate the quality of the results.

- Once feasibility is established, one would investigate which possibilities exists to then query one or many datasets in an automated fashion, e.g. having ChatGPT create and QC SQL queries.

- At the same time methodology would have to be developed to give quantitative and qualitative estimates for the quality of the results and identify risks associated with using such a system..

Equinor can also cater for specific interests

- Equinor is an international energy company with 22,000 employees in 30 countries. Our ambition is to be a leading company in the energy transition, and we aim to create value through the opportunities the energy transition brings, breaking new industrial ground by building on our 50 years of experience.

- With our wide range of industrial and research activities, we can provide exciting master topics in many different fields. Take contact with Dirk if you are interested in writing your thesis in an industrial setting, but none of the topics seems to fit.