RISE Computer Science & AI
open house
May 10, 08:30-13:00
Welcome to RISE Computer Science & AI open house on 10 May!

We will be showcasing cutting edge research, results, and initiatives in AI, Cybersecurity, Data Analysis, IoT, Data Centers and more, together with partners in industry and the public sector.

In this folder you can read more about our keynotes, RISE and partner presentations and the 77 parallel poster sessions

You can also find relevant links as well as links that will take you to each session or talk on May 10

Agenda:
8.30-9.30 introduction and invited keynotes on the next steps in industrial AI and data-driven innovation by technology leaders

Ali Ghodsi (CEO Databricks), Jim Dowling (CEO Hopsworks) and Staffan Truvé (CTO Recorded Future)

9.30-11.00 talks on key areas by RISE and partners

11.00-13.00 demos & posters session, where you can interact with 50+ experts on state-of-the-art technology topics
Keynotes
08.30-09.30 4x15 mins, one track session

Welcome and introduction
RISE Research Institutes of Sweden
Hanifeh Khayyeri & Sverker Janson

RISE Dept. of Computer Science welcomes you to this Open House event.

Ali Ghodsi, CEO Databricks
Fireside chat
Ali Ghodsi, Seif Haridi, Martin Körling

The journey from SICS/RISE to Berkeley and Silicon Valley, Databricks from Spark to full data platform, Cloud versus On-prem. Longer version will be available.

Jim Dowling, CEO Hopsworks
Fireside chat
Jim Dowling, Seif Haridi, Martin Körling

Hopsworks, the importance of Feature Stores, start-ups in Europe versus the US.

Staffan Truvé, CTO Recorded Future
Presentation
Staffan Truvé

Recorded Future is a leading provider of digital intelligence services. By combining persistent and pervasive automated data collection and analytics with human analysis, Recorded Future provides real-time visibility into the vast digital landscape and empowers clients to take proactive action to disrupt adversaries and keep their people, systems, and infrastructure safe.

Click here to join the keynote/presentation session on May 10 08.30-09:30
**Language Models for NLP – Is Bigger Always Better?**

Joakim Nivre

Recent advances in NLP, driven by large-scale pre-trained language models like BERT and GPT-3, suggest that the way forward is to build ever larger models on ever more data. Even if this is true, there will always be situations where it is hard or impractical to use the largest models. More generally, we should always strive to minimize the resources needed to achieve sufficient performance.

**Are we ready for the next cyberattack?**

Shahid Raza

Cybercriminals, we don’t like them, but we must give them credit for their spirit of innovation. The ever-increasing range and sophistication level of cyber-attacks makes cybersecurity a continuous process, like a game of cat and mouse, where the balance of power changes over time. This talk will highlight both short-term and strategic long-term research initiatives, with the potential to increase Swedish competitiveness in defending against cyber-attackers with heterogenous capabilities.

**Deep learning for real problems in the face of limited training data**

Olof Mogren

AI has become one of the most influential technologies, often powered by deep learning; a technology achieving astonishing results in a diverse set of tasks. However, limitations in training data may have severely detrimental effects. In this talk, we’ll see some examples of how RISE develops solutions to many challenging problems in the face of limitations in the training data.

**Why do we need data centers when we have the cloud?**

Tor Björn Minde

The cloud lives in datacenters, which have become the third main digital infrastructure of modern society, together with mobile networks and the core Internet. The ongoing discussion is now about how these can become even more efficient and even integrated parts of the energy system. The energy use of the facility envelope is still 10-20% of the data center total energy use. The remaining 80-90% come from the software applications running on the servers, which also need optimization. A holistic view on the energy efficiency of the total IT infrastructure is needed.
RISE & PARTNER PRESENTATIONS

AI & ECOSYSTEMS

Supporting SMEs in the Swedish and EU AI ecosystem?
Jeanette Nilsson

The Swedish and EU ecosystems for AI offer a wide range of initiatives and tools to support industry, SMEs, and the public sector in their digital transformation. SMEs seeking to explore the potential of data-driven innovation and AI can have difficulties navigating these offerings and getting help with their change journey. We discuss how RISE works with the AI ecosystems and supports partners.

AI PLATFORMS AND INFRASTRUCTURE

Is your multi-cloud strategy ready for the quantum age?
Martin Körling

Technical developments are going incredibly fast, both at the hardware level, with GPUs, HPC, and Quantum, and at the software level, with new cloud AI/ML tools and ML frameworks. Wouldn’t it be nice with a testbed, where you could evaluate full stack environments, plus take steps towards interoperability and portability to be able to handle all this?

INDUSTRIAL DATA ANALYSIS

Moving beyond prediction
Sepideh Pashami

Machine Learning has shown impressive results on predictive tasks such as image processing, pattern recognition, and learning by demonstration. However, most of today’s success stories focus on a particular technique, fine-tuned for an isolated problem. As we move towards more intelligent, advanced, and automated AI systems, we need to combine multiple capabilities (Hybrid AI), generalize better in new conditions (Transfer Learning), get closer to human-level intelligence (Causal Inference), and provide trustworthy and explainable solutions (XAI).

IoT

The battery-free Internet of Things
Carlos Penichet

Batteries have long been an obstacle to the full development of the Internet of Things due to their maintenance and deployment costs as well as their negative environmental impact. Energy harvesting is a promising alternative, but traditional radio transceivers consume too much for most harvesters. With our work we leverage backscatter communication techniques to enable IoT devices without batteries on harvested energy; thereby enabling a host of potential new applications in a sustainable manner.

Click here to join the keynote/presentation session on May 10 09.30-11:00
RISE & PARTNER PRESENTATIONS

SWEDISH NATIONAL SPACE AGENCY

Digital Earth Sweden, data-driven innovation on satellite data
Tobias Edman
The Swedish National Space Agency on the benefits of satellite data.

ERICSSON

AI by design, a business priority
Erik Sanders
We welcome our partner Ericsson to share their insights from working with AI.

VERTIV

Innovate for the new era of sustainable large-scale cloud, enterprise and edge digital infrastructures
Piotr Kowalski

Partner program
Cecilia Hyrén
A short summary of partner program
**AI@Edge – AI for network automation**

The introduction of 5G technologies is a paradigm shift: its high performance in terms of latency, bitrate, and reliability, call for a technological and business convergence between the cloud computing and the telecom worlds. Features like edge computing, network slicing, and better and more flexible radio connectivity can be used to support qualitatively different applications. Nevertheless, the challenges to be overcome in order to realize this connectivity/computing convergence are still notable. In particular, the increasing number of control and optimization dimensions of the end-to-end 5G infrastructure may result in an overly complex network that operators and vendors may find difficult to operate, manage, and evolve. The introduction of AI and Machine Learning (ML) technologies in the cloud-network convergence process will be crucial and help operators achieve a higher level of automation and increase network performance. The aim of the AI@EDGE project is to build a platform and the tools that enable secure and automated roll-out of large-scale edge and cloud compute infrastructures, with close to zero-touch of the underlying heterogeneous MEC resources (network, storage and compute resources).

**Jamming Detection with JamSense**

Low-power wireless networks transmit at low output power and are hence susceptible to cross-technology interference and jamming attacks. These may cause packet loss which may waste scarce energy resources by requiring the retransmission of packets. We present JamSense, a tool that is able to identify jamming attacks with high accuracy while not classifying Bluetooth or WiFi interference as jamming attacks.

**Links and contacts**

https://aiatedge.eu/
Contact:
Daniel Perez
Akhila Rao

**Click here to join poster on May 10 11:00-13:00**

**Links and contacts**

YouTube
Contact:
Thiemo Voigt
Data trace generation with Multi-Trace

Wireless, low-power, multi-hop networks are exposed to numerous attacks also due to their resource constraints. While there has been a lot of work on intrusion detection systems for such networks, most of these studies have considered only a few topologies, scenarios, and attacks. One of the reasons for this shortcoming is the lack of sufficient data traces required to train many machine learning algorithms. In contrast to other wireless networks, multi-hop networks do not contain one entity that can capture all the traffic, making it more challenging to acquire such traces. We present Multi-Trace, an extension of the Coojasimulator with multi-level tracing facilities that enable data logging at different levels while maintaining a global time.

ZeroIoT

ZeroloT enables the battery-free internet of things by advancing and combining backscatter communications and energy-driven intermittent computing. ZeroloT targets two application domains: In healthcare, implanted devices will improve the diagnosis of Sarcopenia and Osteopenia and the monitoring of prosthetic implants. In civil engineering, we will provide novel means to perform remote displacement monitoring for embankment dams.

Links and contacts

Contact: Niclas Finne

Click here to join poster on May 10 11:00-13:00
The Digital Futures Drone Arena

The Digital Futures Drone Arena is a platform where players in the digital transformation and society join in a conversation about the role of mobile robotics, autonomous systems, machine learning, and human-computer interaction. The platform takes the form of a novel aerial drone testbed, where drone competitions take place to explore the unfolding relationships between humans and drones. The inaugural competition is scheduled June 15-17, 2022 and asks participating teams to fly drones through an obstacle course in the shortest time.

Securing IoT Systems through Stateful Fuzz Testing

Constrained IoT operating systems comprise a plethora of protocol implementations that can be targeted by attackers. To detect security vulnerabilities, IoT developers employ advanced fuzz testing tools, but these tools fail to reach code that is dependent on protocol states. We propose a method to generate such states in a systematic way, and thereby improve the efficacy of fuzz testing.

Links and contacts


http://www.dronearena.info

Contact: Luca Mottola

https://assist-project.github.io/

Contact: Nicolas Tsiftes

Click here to join poster on May 10 11:00-13:00
Improved crypto and PKI support for Contiki-NG

MbedTLS is a lightweight cryptographic library designed for resource-constrained devices. In this project, we secure the popular IoT device management protocol LwM2M with DTLS from MbedTLS. The system is evaluated on the nrf52840 platform to gain insights on the overhead and possible optimisations of MbedTLS when integrated as a component into the Contiki-NG operating system.

AutoPKI: automated and scalable transfer of trust

IoT deployments grow in numbers and size and long-time support and maintainability become increasingly important. To prevent vendor lock-in, capabilities to transfer control of IoT devices between service providers must be offered. We propose a lightweight mechanism for transferring control using standard-based building blocks and show that it maintains desired security requirements.

Links and contacts
Contact: Jayendra Ellamathy
        Joel Höglund
        Nicolas Tsiftes

Links and contacts
Contact: Joel Höglund

Click here to join poster on May 10 11:00-13:00
Designing Touch Interaction

In designing for touch interactions, we move beyond the traditional graphical user interface, and make use of the intelligence that resides in the body. Through mid-sized haptic interactions directly touching the body, we create awareness and convey information. These interactions are showcased in two prototypes: the Pelvic Chair, aimed at pelvic floor health, and the Re-engage Seat in collaboration with Volvo Cars aimed at re- or disengage drivers in semi-autonomous cars.

Digital twin Kista (demo)

In this demo, we will show work on a digital twin to capture energy use and user behaviors in our office as well as a 4D map of the surrounding area in Kista. The twin will be shown at a fair later in May and has been developed in collaboration with among others Kista Limitless and Urban ICT arena.

Links and contacts

Contact: Anton Gustafsson

Click here to join poster on May 10 11:00-13:00
Radio Fingerprinting for IoT intrusion detection

An attacker can gain access to IoT networks by imitating the identity of an authentic IoT node. Radio fingerprinting distinguishes devices based on their unique hardware imperfections and makes it harder for an attacker node to impersonate another node. In this work, we use contrastive learning to distinguish between authentic nodes and intruder nodes. Furthermore, we investigate the impact of neural network quantization on our radio fingerprinting intrusion detection system.

Data programming and Streaming

There is an inherent need in our society for simpler and more effective ways to program data-driven systems. At one hand, high level declarative languages can be used to democratize data-driven applications, while at the same time automate code generation and optimisation for different target hardware. To that end, our team at RISE has led significant contributions in that area. This poster session highlights how we build on our prior work on Apache Flink’s state management to further introduce Arcon: a next generation streaming platform that allows analytical queries, ArcLang: a novel language that can be used to program reactive compute- and data-centric applications and related contributions in MLIR, a current standard in multi-level compilation for data analysis.
Graph ML Databases

Around 80% of the industry’s data query workloads by 2025 will be reducible to graph database queries according to Gartner. Graph databases currently take over the data management industry and RISE and KTH has a facilitating role in leading the creation of ORB: one of the most promising future graph database systems that will be able to support deep reasoning in polynomial time. Our architecture combines stream processing technology, error estimation and graph neural networks put together in a first-of-a-kind database system stack.

Serverless computing on the edge

Serverless is undoubtedly a key technology for the future of software, allowing high flexibility in utility computing without the need to declare and maintain virtual infrastructure such as servers and VMs. Despite its wide spectrum of applications, the serverless paradigm is currently inapplicable to edge computing due to the level of heterogeneity, asynchrony, uncertainty, and instability in such WAN-supported infrastructures. To that end, RISE is driving efforts to enable serverless computing across edge infrastructures. Our approach builds on decentralized resource management, synchronization primitives and principal application of transactional processing known in distributed databases in support of resilient service composition. Furthermore, we investigate faster ways to achieve state machine geo-replication while being robust to several types of network failures that can unexpectedly occur.
Privacy-preserving computing
Privacy-preserving computation allows multiple parties to securely share sensitive data and collaborate on this shared data. We are working on better ways to compose and integrate privacy preserving workflows with existing workflows; an ecosystem with privacy focus and native support for privacy primitives such as data ownership and secure multi-party computation. We hope to use these to support both GDPR requirements, and collaboration via advanced secure workflows, embedded natively within a widely adopted actor middleware software package.

Too slow and too expensive? Tools for software development
Do you need impartial help evaluating or engineering cutting-edge solutions for your software development process? Our researchers have decades of experience in adapting research results to solve industry problems. This session showcases some of our success stories, ranging from IDEs and compiler technology to automatic test case generation.

Links and contacts
Contact:
Jonas Spenger, Paris Carbone, Philipp Haller

Links and contacts
Contact:
Frej Drejhammar
Per Mildner
Peter A Jonsson
Mats Carlsson

Click here to join poster on May 10 11:00-13:00
Click here to join poster on May 10 11:00-13:00
AI & Cloud platforms/infrastructure, where are you running your AI, HPC, Quantum?

Currently, using AI/ML tools comes with several challenges. i) The evolution of new methods and software frameworks is very rapid, so there is a need to constantly be able to try and evaluate the latest. ii) There is a need to find and allocate infrastructure resources to run AI/ML training and serving workloads, often using several different providers, a multi-cloud setup (public cloud and in-house resources). iii) The ease-of-use for typical complete setups including infrastructure and platforms is not at a level to allow domain experts.

We present the status of ongoing work to address these challenges, based on recent developments in infrastructure-as-code tools. We look at multi-cloud aspects (portability and API harmonization across public cloud providers, HCPs), dynamic setup of test environments, and how all of this can be used to benchmark new AI/ML frameworks.

Resource-efficient machine learning

Training a deep neural network can be a time-consuming process using days of wall-clock execution time on a large compute cluster. For instance, training the GPT-3 natural language model used 355 GPU-years of computation. Models can also become exceptionally large with GPT-3 weighing in at some 700 GB and Google Brain being almost ten times that size. With those compute requirements come large power requirements; the compute resources used for training the GPT-3 model correspond to close to 1GWh of electricity which, if generated by a coal fired power plant leads to the emission of close to 1000 metric tons of CO2.

The large compute requirements are caused by a combination of the considerable number of parameters of the models, the large amount of data used for training and the iterative nature of the training process (many small updates to the model parameters to gradually improve the results on the training set).

When the model has been trained, it can be used for inference. For instance, a model trained to drive a vehicle using photos and videos of traffic situations is used to drive.
Resource-efficient machine learning (continued)

Here, the compute demands are smaller since just a single input at a time is used rather than a large training set and each input is used once rather than many times to gradually adjust parameters.

But resource demands may still be problematic since inference may occur in a more constrained environment, such as in a vehicle rather than in a data centre, and the acceptable latency may be a fraction of a second for each input rather than several days for the whole training process.

Thus, there is a dire need to find resource efficient implementation methods for training as well as inference.

Automating air traffic control

Much of our infrastructure is managed by humans in complex and advanced ways. One example is air traffic control (flygledning). It is a complex real-time management of distributed resources (airplanes and airspace) with operation procedures designed to work across jurisdictions, to conservatively handle both old and new airplanes, and to handle the unforeseen. This makes it an interesting and challenging area to automate with computer-controlled automation.

Links and contacts

Contact: Karl-Filip Faxén

Click here to join poster on May 10 11:00-13:00
Multilingual Image Search with M-CLIP

Recently, OpenAI released CLIP which caused a pivotal breakthrough in the long-standing AI endeavor of linking the textual and the visual domain. This model distinguishes how well an English text corresponds with a given image with unprecedented accuracy. This makes CLIP ideal for searching large untagged image databases using plain English. Capitalizing on the modularization of the CLIP architecture, we have successfully converted CLIP into various non-English languages. These multilingual CLIP models have been downloaded millions of times since their release and have enabled people worldwide to utilize the recent cutting-edge AI.

Big Data Analytics and AI for a Smarter Society

The rapid advancements in the field of AI and the abundance of data in a connected society are transforming the way we live and interact. In this poster we present a number of projects at RISE in which big data platforms and AI techniques are applied to improve our society in various ways including transport and automotive industry, e-health, and digitalization.
Towards biologically inspired meta-learning

The brain is a von Neumann computer. We are now beginning to understand its assembly language instructions. It is time to figure out its basic program and data structures. One of the most fundamental software components of the brain seems to be a bootstrapping recursive meta-learning module. Understanding this better may lead us to new ways of representation learning and handling heterogeneous data.

Links and contacts
www.drnil.com
Contact: Martin Nilsson

Click here to join poster on May 10 11:00-13:00

Jointly Learnable Self-Supervised Graph Representation Learning

Recently, several SSL methods for graph representation learning have achieved performance comparable to SOTA semi-supervised Graph Neural Networks (GNNs). One of the key challenges is data-augmentation, for which existing methods rely on heuristically crafted techniques. In this study, we propose a novel method for jointly learning both the augmentation and representation by leveraging the inherent signal encoded in the graph. Besides, to allow efficient use of resources we propose a new architecture that augments in the latent space as opposed to the input space.

Links and contacts
Contact: Zekarias Kefato, Sarunas Girdzijauskas

Click here to join poster on May 10 11:00-13:00
Decentralized Graph Neural Networks

Graph Neural Networks (GNNs) achieve state-of-the-art results in most graph representation learning benchmarks. However, compared to other deep learning models, their structure makes them hard to decentralize. Yet, decentralization is an important tool to achieve large-scale, data-private machine learning. In this work, we show how layer-wise, self-supervised learning may be used to train deep GNNs on a decentralized graph, where each node represents a separate computing device.

Explicit Spectral Graph Neural Network

GNN outperforms CNN in classification tasks since data relationships are explicitly accounted for during the training and testing. However, the relative relations between training and test data embeddings are not readily available for most datasets, preventing using GNN in most datasets. This work proposes to learn such a relation of training and test data embeddings into an affinity matrix using a spectral clustering setup. This affinity matrix and the eigenvector embeddings (from the spectral clustering) are the two components utilized for message passing in the GNN setup. We can observe an improvement in classification on popular datasets (i.e., MNIST and fashion MNIST) while utilizing a fraction of the labels from the training data.
Representation Learning on Heterogeneous Information Networks (HINs)

Graph representation learning (GRL) is a common approach to network mining and analysis. GRL is a set of techniques where the objective is to learn the graph/network structure and construct latent feature vector representations (embeddings) for the nodes and edges in the graph. Learning high-quality representations can be useful in automating prediction and other downstream tasks such as search and personalized recommendation. While most real-world networks are heterogeneous in nature, recent research efforts have focused on learning over homogeneous graphs. The typical complex rich structures of heterogeneous graphs likely cause the existing learning methods to fall inadequate when applied on them. In this poster we present a novel unsupervised model for learning node embeddings based on the structure of the graph. The new learning method is designed for heterogeneous graphs with learnable attention weights over the edge types in the vicinity of each node whose embedding is being learned.

MiniZinc for Problem Solving

Problems such as timetabling, scheduling, planning and rostering are central and often cost-critical components of many software solutions. Such problems are intrinsically hard to solve. RISE has a long track record in the theory, software and practice of solving those problems. A new interchange language, MiniZinc, for expressing such problems has recently emerged. This poster will give a brief introduction and some examples.
Rallying around AI – RISE Center for Applied AI

TBA

Links and contacts
Contact:
Magnus Hultell

Reinforcement Learning for radio resource management

Radio access networks beyond 3G divide and multiplex use of the radio resource into discrete channels and time frames. The resources are allocated to user devices and scheduled on antenna beams. Scheduling resources on beams that interfere, reduces efficiency of resource usage. Our result shows that it is possible train neural networks using reinforcement learning and advanced network simulator. This is Joint work with Diarmuid Corcoran at Ericsson AB, who also funded the work at RISE, 2018-2022.

Links and contacts
Contact:
Per Kreuger

Click here to join poster on May 10 11:00-13:00
Complex AI systems - The value of hybrids of AI techniques

For practical applications of AI based and autonomous systems, it is often not sufficient to use a single AI technique, but a system made up of combining several techniques may be required. Often this can be a combination of data-driven and domain knowledge-based techniques. Here I will describe a general methodology for such hybrid AI systems, and exemplify with recent projects.

Causal aware machine learning

Causal inference provides understanding about how the systems behave under changing, unseen situations. Knowledge about the causal relations allows us to answer “what if” questions, thus leading artificial systems towards higher awareness levels. True causal networks can be obtained using manipulations and interventions. Based purely on observational data, assuming faithfulness, one can, by running conditional independence tests, create a Markov equivalence class which is guaranteed to contain the true causal graph.

Links and contacts

Contact:
Anders Holst, Per Kreuger, Björn Burling

Contact:
Sepideh Pashami, Anders Holst
Natural Language Reasoning (NLR)

The goal of the NLR project is to develop machines that, based on information in text, are capable of reasoning on a human level. NLP applications, which use text as a data source, are lousy at reasoning, which is an obstacle to improvements to existing solutions and to the development of new applications.

IRRA Intent recognition for autonomous vehicles

Intention recognition is the task of inferring an agent’s intention based on its previous actions. It is crucial for human social intelligence which in turn enables understanding of, and the ability to predict, other humans’ behaviours, such as for example other drivers’ intent to overtake, stop, turn, or switch lanes. For making situation-based decisions, both autonomous and human drivers need to take the intentions of surrounding vehicles into account. This is especially true in a mix of autonomous and human drivers.

Existing algorithms and models for intention recognition need to be improved with respect to accuracy, transparency and scalability, to meet the requirements of the Swedish automotive industry and Trafikverket. The Vinnova FFI project IRRA (2019-2024) aims at addressing this knowledge gap and enable the creation of novel advanced and intelligent automotive services and products based on social intelligence and intention recognition. RISE contributes with novel logic-statistical hybrid algorithms for intention recognition and with novel algorithms for automated scene perception based on traffic data.
**Socio-economic effects of hard and soft digital infrastructure: from modelling to visualization**

In this poster we give an overview of our models to quantify and visualize the costs and benefits of digitalization, using two examples. The first is a model of the social economic impact of fiber broadband in Sweden using fixed-effect regression, controlling for municipality- and time-fixed effects over an 8-year perspective. The second is a dashboard-style visualization tool for 3 groups of effects of digitalizing the permit handling process in Sundsvall. A demo will be given to the interested visitor.

<table>
<thead>
<tr>
<th>Links and contacts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contact:</td>
</tr>
<tr>
<td>Jie Li, Marco Forzati</td>
</tr>
</tbody>
</table>

**Identifying causal relationships of cancer treatment and long-term health effects among 5-year survivors of childhood cancer in Southern Sweden**

Survivors of childhood cancer can develop late effects in adulthood. Knowledge about possible late effects can improve childhood cancer treatments and assist in follow-up. We developed a method to identify causative links between treatments and health outcomes. We applied it to a Swedish patient cohort and identified 98 causative links between treatments and outcomes, many of which are already known. Some, however, have not been previously described, including links between certain treatments and eye conditions or viral infections. We also confirm that childhood cancer survivors use more health care and have a higher mortality compared to the general population. This study helps to create a better understanding of the late effects of cancer treatment in children and may help to guide strategies to monitor and treat children to avoid these effects.

<table>
<thead>
<tr>
<th>Links and contacts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contact:</td>
</tr>
<tr>
<td>Helena Linge</td>
</tr>
<tr>
<td>Jan Ekman</td>
</tr>
<tr>
<td>Anders Holst</td>
</tr>
</tbody>
</table>
Privacy preservation by FHE enables data sharing in digital health management

RISE entry to the Vinnova innovation competition Vinter 2021-22
We present our experience at the Vinter competition 2021 organized by Vinnova. The Vinter competition invited solutions for digital health with a focus on interoperability. Our entry addresses privacy preservation by fully homomorphic encryption (FHE), and supports legal interoperability between e.g. users, health care and third party service providers. The work describes implementation of FHE in a digital health landscape. Its benefits, limitations and its potential for AI-based applications are discussed.

Swedish AI Society annual workshop
June 13-14 – hosted by RISE
Welcome to SAIS 2022: the 34th Swedish Artificial Intelligence Society (SAIS) annual workshop. The SAIS workshop has since its first edition been a forum for building the Swedish AI research community and nurturing networks across academia and industry. We invite researchers and practitioners in AI and related disciplines, in Sweden and the rest of the world, to join us in exchanging knowledge, news and results on AI related theory and applications.

SAIS 2022 is hosted by RISE and will take place in Stockholm, Sweden, on June 13-14, 2022. The event is planned to be hybrid combining in-person and virtual participations.
Machine Learning for Causal Inference in Observational Studies

To show a causal relation we need to observe the outcome of a treatment (factual), as well as the outcome in case the treatment had not been assigned (counterfactual). The fundamental problem of causal inference is that we can only observe one of the two potential outcomes. The gold standard solution is to run randomized controlled trials, which comes with its own limitations and pitfalls. Observational studies, including cohort studies, are an alternative, but they can normally only find associations and not causal relations. With the abundance of huge datasets, in healthcare or social studies, that are collected over many years, can machine learning models provide a solution for causal inference? What are the requirements, possibilities and limitations?

Learning Causal Predictors

Machine learning (ML) models have shown great success in the fields of image processing, language processing and stock market trading. One of the reasons behind its success is their great ability to learn patterns and correlations from big data. Nevertheless, these models still lack robustness when deployed in different environments, i.e., when exposed to unseen scenarios. One major drawback with current ML models is that they do not learn causation, instead weigh spurious features in predictions. A way to ease this problem is to add causal knowledge to ML models.
Pre-training Transformers for Molecular Property Prediction Using Reaction Prediction

Molecular property prediction is an important task within chemistry with application in drug discovery. However, the amount of available data on molecular properties is often very limited. Transfer learning has had a tremendous impact on fields like computer vision and Natural Language Processing and if it could be used within molecular property prediction it would have an impact of similar magnitude as in the aforementioned fields. We present a pre-training process for molecular representation learning using reaction data. The pre-training step is formulated as a reaction prediction task using a Transformer on SMILES and is evaluated on 12 molecular property prediction tasks from MoleculeNet within physical chemistry, biophysics and physiology. We can demonstrate a statistically significant positive effect on 9 of the 12 tasks compared to a nonpre-trained baseline model.

Machine learning for predicting water flow intensity based on physical characteristics of catchment area

In this pre-study we have investigated an ML model for water flow intensity prediction, which is given rain and temperature measurements and a top-view map of a catchment area. Given this input, the model outputs a water flow intensity map of the same spatial extent as the input map. Early results suggest that our ML model is able to predict water flow intensity more accurately than comparable baselines.
AI for drug influence detection using recorded eye scans

In this project, executed together with Eyescanner Technology AB, we are developing an AI system for determining whether a person is under the influence of drugs based on recorded eye scans. The two founders of Eyescanner have previously worked in the police, and their goal is to provide a technology which can improve upon the contemporary manual approach for deciding whether a subject is under influence.

Semi-supervised few-shot learning for sound event detection

In this project we introduce a semi-supervised method for the training of embedding functions to be used in downstream few-shot learning tasks. We include unlabeled data from the downstream task of interest as well as unlabeled data from the base data set and show how these can be included in an episodic training scheme by introducing a novel loss function for prototypical networks.
AI to learn about porcelain

This project is a collaboration with Rörstrands Museum, in which we develop a machine learning model to gain knowledge about porcelain from images. The model will build upon the digital archives of the museum, and the end goal is a service where people can upload images and learn about their own Röstrand porcelain.

Links and contacts
Contact: 
Ebba Ekblom
Olof Mogren

Click here to join poster on
May 10 11:00-13:00

Teaching a computer about beauty for image-based recommendations

The field of recommendation systems is dominated by collaborative filtering approaches that exploit vast datasets of user-item interactions to select the most relevant items for any given user. However, in social networks the popularity of items depends to a large degree on the author’s popularity, and the system can be exploited by user groups conspiring to push specific content regardless of its quality. An objective assessment of content quality independent of its social impact would be a valuable addition to many recommendation systems, but objective quality is often difficult to define. In this project we use a large image dataset derived from a social network for photographers, and manually labeled by experts, for a binary classification between photos that are visually outstanding and a random selection from all other photos. The network’s output scores let us create a ranking of the images that is in line with human appraisals, allowing for the preliminary conclusion that the network has learned to recognize beauty in images.

Links and contacts
Contact: 
Leon Sütfeld

Click here to join poster on
May 10 11:00-13:00
Machine learning and acoustic monitoring for wildlife conservation

In ecological research, years of audio data from natural habitats are being recorded using cheap acoustic sensors and machine learning is a promising way towards automated analysis. In this project we develop annotation efficient machine learning methods for soundscape analysis. We demonstrate the methods on recordings from wildlife conservation projects.

<table>
<thead>
<tr>
<th>Links and contacts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computational bioacoustics with deep learning: a review and roadmap</td>
</tr>
<tr>
<td>Perspectives in machine learning for wildlife conservation</td>
</tr>
<tr>
<td>Contact: John Martinsson</td>
</tr>
</tbody>
</table>

Click here to join poster on May 10 11:00-13:00

Distributed machine learning

The study of machine learning on decentralized data has been an interest for a long time, and has relations to many different research communities, such as cryptography, databases and data mining. The overarching goal being to efficiently learn from and analyze data distributed among several users, without the data leaving each user. Research in this field has become increasingly important in recent years. Smart and connected devices are continuously being adopted and more widespread, and the equipped sensors are improving and collecting a lot of data at a fast pace, which gives rise to a need for algorithms that can efficiently handle the large volumes of distributed data.

Federated learning (FL) is a distributed machine learning framework where a central server orchestrates the training of clients to learn a model on decentralized data. This entails several questions, and many challenges have emerged that are expected to grow in the near future. However, a central server might not always be available or even desirable, as it can be a large bottleneck or a possible source for malicious attacks. Decentralized machine learning is a research field addressing such problems. The aim of this PhD project is to study challenges and advance the research in both federated and decentralized learning. Open problems in these fields include developing algorithms for data heterogeneity, personalization, efficient communication, model compression and many more.

<table>
<thead>
<tr>
<th>Links and contacts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contact: Edvin Listo Zec</td>
</tr>
</tbody>
</table>

Click here to join poster on May 10 11:00-13:00
Resource-Efficient Multilingual NLP

How can we make resource-efficient NLP available in situations where we do not have access to super-computers? And how can we deliver high-quality NLP for languages with more limited resources than English and a few other languages, so that a larger part of the world's population can benefit from the technology? A promising approach to the second problem is to train multi-lingual models on data from multiple languages, hoping that the languages will mutually reinforce each other.

In this project, we explore modularization and model compression as key techniques for making multilingual NLP models more effective, by delivering higher quality for more languages, and more efficient, by requiring less resources to deliver these results.

Cross-lingual transfer of monolingual models

Recent studies in cross-lingual learning using multilingual models have cast doubt on the previous hypothesis that shared vocabulary and joint pre-training are the keys to cross-lingual generalization. We introduce a method for transferring monolingual models to other languages through continuous pre-training and study the effects of such transfer from four different languages to English. Our experimental results on GLUE show that the transferred models outperform an English model trained from scratch, independently of the source language. After probing the model representations, we find that model knowledge from the source language enhances the learning of syntactic and semantic knowledge in English.

Links and contacts
Contact: Joakim Nivre

Click here to join poster on
May 10 11:00-13:00
Semantic textual similarity for causality detection

Causality detection is the task of extracting information about causal relations from text. For Swedish, there is a lack of annotated data in this area. We investigated how a keyword filter and a Swedish semantic textual similarity model can be used to identify sentences containing causal relations and build a small demo to rank causal sentences in governmental reports relative to a specific query.

Reducing hazardous chemicals in society with AI and patent data

Patents contain information about the use of chemicals in products before they are introduced to the market. Patent data can therefore be utilised to proactively detect and avoid that hazardous substances end up in society and the environment. In a feasibility study, RISE, the Swedish Chemicals Agency and the Swedish Intellectual Property Office have demonstrated that natural language processing can be used for this purpose, enabling automatic, scalable and accurate detection of relevant patents beyond the reach of traditional keyword-based search.
ICE Test & Demo Datacenter

Sweden needs sovereignty and skills from the ground to the cloud. To support research and innovation projects for all of RISE in all areas from AI to data center heat reuse RISE has a test & demo facility called ICE in Luleå. World leading research in a number of areas is supported by two compute modules with thousands of CPU servers and 250 GPUs and a facility test lab containing compute modules, edge testbeds, liquid cooling testbed, wind tunnels, climate and heat chambers.

WeDistrict : Fuel cells and excess heat from datacenters put to use

Recovering heat from datacenters is a great opportunity, but also very challenging since the temperature of the heat often is low. In this project we are building a demonstrator to combine liquid cooled data center technology with solid oxide fuel cells, to maximize the system waste heat temperature, and in the same time produce green power for the servers running in the liquid cooled datacenter. The generated heat will be transferred to the local district heating network without using additional heat pumps, which normally is used when heat from datacenters is recovered and used for district heating purposes.

Links and contacts
ICE datacenter
Contact: Tor Björn Minde

Links and contacts
wedistrict.eu
Contact: Jonas Gustafsson

Click here to join poster on May 10 11:00-13:00
**E2P2 – Eco Edge Prime Power**

The E2P2 project will create a POC alternative prime power source that employs fuel cell technologies for on-site power generation of data centers. The concept of connecting fuel cells to gas networks to power resilient urban and edge data centers overcomes the need to have backup generation, thus reducing the emissions. The project will also create an open standard for fuel cell adaption to power data centers and analyse the combined social, environmental and commercial impact.

**Powering the edge**

To provide the brains for all our low latency computing needs, a new generation of edge data centers is making their way to our cities. However, at the same time the electrical grids in our cities are becoming increasingly congested and unstable. Within a growing portfolio of projects, we are looking for smart methods to design and operate these edge data centers so that the overall effect on the grid would be positive as e.g., virtual power plants or partaking in the frequency reserve.

**Links and contacts**

*E2p2.eu*

Contact: Jeanette Petersson

Click here to join poster on May 10 11:00-13:00

*Aniara*

Contact: Mikko Siltala

Click here to join poster on May 10 11:00-13:00
Environmental impact of data centres

In light of continuous digitalization, we as society are becoming more and more dependent on data centre services. Yet, little is known about the sustainability of data centres, the impact of operating them and the hotspots that are responsible for most of it. This poster presentation will explain how sustainability can be assessed and discuss the environmental performance of data centres.

Data center excess heat reuse, a matter of enhanced food self-sufficiency

During the worldwide pandemic and unstable world peace the import dependency from other countries has become obvious, whereas for Sweden most of the food has its origin in south of Europe. To enhance the food self-sufficiency focused work has been done at RISE ICE data center to evaluate the possibilities to utilize the data center excess heat for farming and food production. The north of Sweden which has the lowest levels of self-sufficiency it’s shown that 1 MW of data center power could increase the level with up to 8%.

Links and contacts
Stani Borisova presented Data Centers & LCA - Life Cycle Assessment

Contact:
Stani Borisova

Click here to join poster on May 10 11:00-13:00

Links and contacts
DC Farming - A guideline for implementing data center and greenhouse symbiosis

Contact:
Mattias Vesterlund

Click here to join poster on May 10 11:00-13:00
Homomorphic encryption for CPS

The design of intelligent control algorithms can be costly and is an investment that must be protected against reverse engineering. It can be safeguarded by running remotely from the cloud instead of locally on the equipment hardware, however such set-up requires careful consideration of customer data confidentiality. Fully Homomorphic Encryption is an emerging technology that permits computation on encrypted data and offers a solution that protects both customer data and manufacturer IP.

Links and contacts
Contact: Rickard Bränvall

Click here to join poster on May 10 11:00-13:00

Thermal energy storage for data centers

EDGE- solutions with small calculation clusters is a strong upcoming trend within the data center sector. With an optimized operation of the EDGE data center and a well dimensioned thermal energy storage (TES), the energy use for the datacenter (DC) can be minimized and with that also their climate footprint. In this poster session we will present one of our projects, DIR Borö, at RISE ICE datacenter, were we tested and evaluated the performance of a thermal energy storage tank at our EDGE test bed to see how long the tank can cool the datacenter without any additional energy. If you are interested to see the results and learn more, please visit our session!

Links and contacts
Contact: Tina Stark, Adrian Mellgren

Click here to join poster on May 10 11:00-13:00
Empowering a Pan-European Network to Counter Hybrid-Threats

In the EU-HYBNET project, a Horizon 2020 coordination and support action over 5 years, with 25 partners, we network with security practitioners, stakeholders, academics, industry players, and SMEs to find new and enhanced means to counter hybrid threats. We define common requirements, deal with performance needs, monitor research and innovation, deliver recommendations for uptake and industrialization of promising innovations and associated standardization.

Security Roadmap Towards 6G Systems

6G system is in an active race to be fully deployed by 2030. 6G system is expected to provide ultra-low latency, low power consumption, ultra-high capacity, seamless coverage, high localization precision, massive MIMO (small cell), millimeter-wave (mmWave), terahertz (THz) bands. Consequently, these high-performance specifications will enable new technologies within 6G systems such as the internet of everything (IoE), extended reality (XR), autonomous systems, Artificial intelligence (AI), machine learning (ML), heterogeneous wireless networks (HWN), intelligent and distributed environments, and visible light communications (VLC), wireless brain-computer interactions (BCI), etc. These new technologies will have significant impacts on the security and privacy of the upcoming 6G system. Therefore, novel security techniques (encryption, authentication, privacy-preserving, key management, access control) or some fundamental changes must be considered; for instance, distributed mutual authentication protocols are highly required for some 6G-based technologies (e.g., HWN, IoE), whereas end-to-end security and encryption protocols are necessary for some others emerging technologies. Hence, extensive research is supposed to be carried out to meet all these security requirements and ensure the reliability and functionality of the upcoming 6G system.
End-To-End Secure Group Communication for the IoT

The security protocol Group OSCORE protects end-to-end group communication for CoAP. Protection applies to a request addressing multiple servers and to its corresponding responses. Group members obtain the keying material from the responsible Group Manager when joining the group. The joining procedure uses the ACE framework, which ensures that nodes become group members only if authorized.

Innovation node in cybersecurity – Cybernode.se

RISE is running a national cybersecurity collaboration to accelerate research and innovation in the industry and public sector – a national node in cybersecurity. VINNOVA is responsible for the funding, while RISE’s cybersecurity unit coordinates the node’s work.
Intrusion detection in cloud containers

The popularity of containerization has extremely increased among organizations during last decade since container solutions offer multiple features consenting high time and cost savings and incomparable agility in application development and management processes. Nonetheless, cloud containers are an appealing target of malicious actors who look at virtualized applications as a new opportunity to exploit the huge amount of cloud resources and enforce their malicious activities. Cyber-attacks targeting virtualized applications with the aim of hiddenly stealing resources and private data have been widely demonstrated. When in-app protections fail to detect and prevent such attacks, it is crucial that the hosting operating system provides additional and innovative capabilities for monitoring container resources and detecting hidden malicious behaviours. Within the Cybersecurity Unit, we are currently working on AI solutions tailored to empower malware/intrusion detection in virtualized systems (e.g., cloud containers), with a focus on the analysis of data features observable by the cloud service provider at the

Key Update for OSCORE (KUDOS)

CoAP is a protocol similar to HTTP but designed for constrained devices and running over UDP. The OSCORE protocol provides end-to-end protection of CoAP messages. Peers using OSCORE should periodically renew their shared keying material, e.g. due to reaching limits for operations performed with the keys. Key Update for OSCORE (KUDOS) is a lightweight procedure for updating OSCORE keying material.
Vulnerability Analysis for Establishing More Secure Systems

Cyber-attacks are on the rise. For instance, most famed autonomous and/or electric cars are shown to be vulnerable as in the example of Tesla hacking showcases. Cybersecurity of any system should start from the attacker point of view, which might be referred to as “counter-intuitive” analysis. As such, Vulnerability Analysis helps researchers and especially security experts understanding limits of their cyber-defences as well open gaps (called vulnerabilities) that their defences fail to cover. Vulnerability Analysis of the systems and subsystems is utmost important for the cybersecurity experts in the intended domain while building/revisiting their cyber-security measures against adversaries. To ensure that organizations and their products are on the cyber-secure side of the spectrum while devising your next-generation products/services, it is central to ensure that these are cyber-secure, systems and subsystem components (along with their interactions with the surrounding peripherals, users, and the infrastructure) need to be analysed with agile Cyber-Security approaches such as Vulnerability Analysis and Attack Graphs.

Securing the Next Generation of Industrial Control Systems

Industrial Control Systems (ICS) are becoming more and more connected. While connecting systems increase flexibility productivity in ICS, it also introduces risks and security vulnerabilities. Media have reported several cyberattacks against ICS, and security is a top priority in the next generation of ICS. High availability requirements and severe consequences of cyber-attacks make securing ICS a challenging problem. In the next generation of industrial control systems, often called Industry 4.0, most parts are assumed to be connected. Securing Industry 4.0 requires research into secure communication, device management protocols and resilient software for ICS devices and Industrial IoT devices.
Federated Learning for IoT Security

With the introduction of new privacy laws such as GDPR in the EU encouraging (and in some cases mandates) sharing only minimal amounts of data, centralized machine learning trained with Big Data becomes infeasible. Federated Learning (FL) has emerged as a very promising paradigm for training distributed ML models. It allows to train a ML model collaboratively so that each participant in the learning task trains the model locally without sharing any private data. However, it brings new challenges such as communication overhead, statistical and systematical heterogeneity. These challenges have been proved that they degrade model’s performance and make learning process more difficult. Within Cybersecurity Unit, we are working on applications of FL to solve IoT security problems. Meanwhile, we address the challenges of FL, and strengthen FL in different aspects.

Formal verification of communication protocol

Our global ecosystem requires secure communications. To this end, protocols are specified and standardized. Formal verification is a useful tool that helps model and analyse these specifications to ensure they provide the required security properties. Strength:

- Strong guarantees w.r.t. the model
- Can be automated / tool-assisted
- Provides early guidance and feedback during design (before implementation)

Challenges (a.k.a. research opportunities):

- Adequation between model and reality
- Lack of intuitiveness (esp. for large scale adoption)
Cyber Range training and education

In order to raise cybersecurity competence among people working in Swedish organizations RISE is now offering Cyber Range courses that can provide hands on experience to better prepare people for potential threats. The courses target different segments within the organization to provide training tailored specifically to programmers, IT professionals and decision makers. Participants will get to experience and perform cyber-attacks in a virtualized environment to get a better understanding of how to protect their systems and properly respond in the case of an incident.

Links and contacts
Cyber Range
Contact: Gustav Midéus

Trusted Execution Environments for Resource-constrained IoT

Securing IoT devices is vital today as the security risks associated with these devices grow rapidly. The increase in Trusted Execution Environments (TEEs) in resource-constrained embedded device (i.e., TrustZone-M) in the critical infrastructure is a step towards isolation and secure execution of software. TEEs provide efficient mechanisms to isolate system resources and hence play a significant role in security-critical operations such as secure boot, crypto operations, software/firmware update and remote attestation. In the Cybersecurity Unit, we work on identifying and resolving challenges in TEEs like securing the inter-world communication, mitigating unauthorized activities within devices and remote attestation of software-state integrity.

Links and contacts
Contact: Anum Khurshid

Click here to join poster on May 10 11:00-13:00
5G Edge Testbed
Computing resources at the network edge, combined with low latency 5G networking is enabling a new class of real-time, compute intense applications. In a 5G testbed hosted by Luleå university of technology, RISE is hosting a compute-cluster at the network edge. On the combined 5G/Edge platform, experiments and projects are run at the infrastructure and application levels, accelerating adoption of edge computing and AI at the edge.

Digital Earth Sweden
Massive quantities of earth observation data are collected every day. Before they can be analyzed, many obstacles such as data storage and processing must be overcome. We lower the thresholds by offering standardized interfaces for accessing and processing the data. The goal is to increase productivity in research and production applications.
Data-science for Earth Observation (EO) Data

Via satellite-mounted imaging devices, earth observation gathers information used to assess the status, and changes in the environment. The Digital Earth Sweden hub is building EO competence through pilot projects. Some examples are coastline change detection or monitoring of grassland grazing. The methods investigated will be adapted to use on powerful computer clusters.

Click here to join poster on May 10 11:00-13:00

Links and contacts
Rymddatalabbet
AI för övervakning - en förstudie
Project Digital Earth Sweden
Contact:
Nuria Agues Paszkowsky
Erik Källman

AI-Ecosystems – ENCCS
ENCCS is the Swedish part in the project EuroCC with 33 members. The project aims to elevate the participating countries to a common high level in the fields of HPC, HPDA and AI. RISE is part of establishing ENCCS, we want to help SME, industry and public sector to start use to possibilities with HPC, HPDA, and AI. We provide activities and competencies to provide training and access to HPC.

Click here to join poster on May 10 11:00-13:00

Links and contacts
ENCCS
Contact:
Jeanette Nilsson
**ColonyOS – a meta operating system**

ColonyOS is a meta operating system for establishing trusted compute environments across decentralized, heterogeneous compute platforms. Each Colony is an abstract collection of compute nodes operating under a single identity, using a crypto protocol to enable secure and zero-trust process execution in arbitrary compute networks. This makes it possible to orchestrate complex machine learning workloads in compute continuums spanning devices, clouds and edge networks.

**Links and contacts**

- [https://github.com/colonyos](https://github.com/colonyos)
- [https://colonyos.io](https://colonyos.io)

Contact:
Henrik Forsgren
Thomas Ohlsson Timoudas

**Edge Computing, service continuity**

Edge computing is an opportunity for service providers to expand their services by offering computation resources at the edge of the network. Cellular networks provide mobility for basic services such as voice and data. Similar mobility support is needed to hand over from one edge host to another while maintaining service continuity. Edge services should enable live migration of application instances and context to the next optimal target Edge host while preserving the quality of service during the migration process.

**Links and contacts**

Contact:
Hamid Faragardi

Click here to join poster on May 10 11:00-13:00