



NorwAI

NORWEGIAN RESEARCH CENTER
FOR AI INNOVATION

ANNUAL REPORT 2023





Cover photo: AI student Kate Bendall peeking under the hood of the Kaia robot - understanding the black box of AI. Photo: Kai T. Dragland, NTNU

Contents

- 4 Language models are taking off
- 6 Fagkompetent generativ språkteknologi
- 10 Large Language Models at University of Oslo
- 13 The project "MIMIR" on copyrighted content
- 14 How can (Norw)AI protect personal data?
- 17 The path to NorLLM: The Norwegian Generative Languages Models
- 18 Four models built - three new ones in the pipeline
- 19 Demand for Norwegian models
- 20 Some observations from the language model work
- 21 Do we need Norwegian models?
- 23 Too little, but not too late
- 24 Newcomers with daunting tasks
- 27 Consortium
- 28 Organization
- 29 Center Management Team
- 30 The mysteries of Hybrid AI
- 35 NorwAI delivers
- 38 The Arendal experience Cooperation between partners
- 40 Members of the board
- 41 The Scientific Advisory Board
- 42 The Innovation Advisory Board
- 44 Research Activities Visions and Views
- 48 Research strategy
- 50 Publications in 2023
- 53 NorwAI objectives
- 56 Building Trust in AI-Enabled Systems through Assurance
- 62 Two new Program Managers at NorwAI
- 65 Another new Program Manager at NorwAI
- 66 The EU-Project TrustLLM as an answer to Generative AI hallucinations
- 68 The forums for shared insights
- 70 NorwAI Innovate '23A unique AI showroom
- 76 Best poster award at NorwAI Innovate 2023
- 78 RecruitmentPhd-students
- 81 Web and social media
- 82 Thework packages
- 89 Presentations and communication
- 94 NorwAI in the media in 2023
- 97 Internal seminars
- 98 Personnel
- 108 Accounts

Language models are taking off



Jon Atle Gulla
NorwAI Center Director
Trondheim 2024-03-11

The NorwAI language model activities received national attention in 2023. Together with Sven Størmer Thaulow from Schibsted and Yngvar Ugland from DNB NewTechLab the team from NTNU presented our largest language model, the NorGPT-23B generative model, on 17th of August at Arendalsuka. Present were also two representatives from the Ministry of Local Government and Regional Development, minister Sigbjørn Gjelsvik and state secretary Gunn Karin Gjøl. The discussion afterwards was very encouraging, and there was a clear understanding that Norwegianbased language models are needed.



Tore O. Sandvik from the Labour Party, now state secretary of the Ministry of Trade, Industry and Fisheries, visited us on 14th of September, before the new minister of Digitalization and Public Governance, Karianne Tung, came to Gløshaugen on 14th of November to further discuss the prospects of training large-scale Norwegian language models. These discussions have of course not been limited to government parties, and Venstre with Guri Melby and Grunde Almeland paid us a visit on 6th of December.

At the end of the year both NTNU and UiO were invited by the National Library to investigate the value of copyrighted material in large-scale language models. This was an initiative from the Ministry of Culture and Equality, and the hope is that this work will lead to larger and higher-quality training datasets for even better language models in the future. We are looking forward to an exciting year for Norwegian language models.

Collaboration across disciplines and sectors is important to NorwAI but can also be quite challenging and time-consuming. In 2022 we introduced the NorwAI forum to bring our partners closer together in deep scientific discussions. This was well received, and

consequently we expanded with both a forum on research and another on innovation last year. It seems that face-to-face meetings are still quite valuable for these forum events. It is interesting to note how different perspectives tend to appear when different companies with different backgrounds get together.

Whereas much of the early work in NorwAI concentrated on models and theories, the issues of trust and data sensitivity are now gradually complementing our research activities. Both DNV and Norsk Regnesentral have competence here that I think will be crucial to the success of many of NorwAI's activities.

We have also observed that companies are now gradually opening up and collaborating also on issues that are of great strategic importance to them. Our partners in the Hybrid AI work package are partly competitors but can still benefit from exchanging ideas and running common experiments. A good example here is the Valhall platform use case, in which Cognite provides real data from Valhall that is then used to configure and refine Kongsberg's simulator. An ambition for 2024 is to increase the level of collaboration and establish several new common projects among our partners.

Fagkompetent generativ språkteknologi



In this article in Norwegian, Åse Wetås, Director of the Language Council of Norway, discusses the importance of developing language models that can handle specialized terminology and language for professional use across various societal sectors. She emphasizes the need for relevant training data from all fields to ensure the models' accuracy and reliability. The article also highlights the challenges of creating Norwegian language models due to the limited amount of Norwegian text, especially in specialized fields, compared to English. She suggests that systematic efforts are needed to develop and make available texts that can improve language technology in Norwegian for different professional areas. The MIMIR research project by the National Library of Norway is mentioned, which aims to measure the impact of including copyrighted material in training generative AI models in Norwegian and to develop a compensation scheme for rights holders. The overarching goal for Norwegian language policy is to maintain Norwegian as a pillar of society and a comprehensive national language used in all aspects of society.



Norsk språk er et av verdens hundre største språk - men likevel et minispråk sammenliknet med engelsk.



Ase Wetås
Direktør, Norsk Språkråd

Nå som generative språkmodeller har vist seg å kunne skrive relativt flytende hverdagsnorsk, må vi tørre å sette forventningene til maskinene enda høyere. For at modellene trygt kan brukes til profesjonelle formål på alle samfunnsområder, må de også kunne håndtere og produsere fagspesifikt språk og spesialisert terminologi for disse samfunnsområdene på en korrekt, forutsigbar og pålitelig måte. I maskinlæringens tidsalder betyr det at det trengs relevant treningsdata fra alle fagfelt.

Med en gang man dykker dypere ned i leksikografien, ser man at grensen for hva som er «ordene i det norske språket» er relativt flytende. Nye ordsammensetninger lages hele tiden, og endringer i samfunn og teknologi bærer også med seg nye ord.

Mange av disse ordene er det vi kaller fagord eller terminologi, og disse er byggesteinene i fagspråket for alle fagfelt. Disse ordene er sjeldne, men samtidig svært viktige. Uten tilgang

til oppdatert fagspråk vil vi få språkmodeller som bruker upresise ord og tar feil avgjørelser.

Konsekvensene kan være at brukerne mister tillit og heller går over til engelskspråklige modeller, noe som igjen kan føre til at norskens posisjon svekkes i fagfeltene. Med det øker faren for at norsk blir redusert til et hverdagsspråk, og ikke lenger er et samfunnsbærende nasjonalspråk som kan brukes i alle deler av samfunnslivet.

Fagspråkets vei mot bruk i språkmodeller

Et fagords reise fra fødsel til bruk i språkteknologien kan være lang og strabasiøs. Først må det være enighet i et gitt fagfelt om hvilke norske ord som skal brukes om hvilke fenomener. Dette skjer gjennom en kombinasjon av naturlig spredning i språkfelleskapene og systematisk terminologiarbeid i virksomhetene.

Fordi dagens språkteknologi opererer med ordrepresentasjon basert på sammenheng med andre ord, må de som produserer tekst i et gitt fagfelt ta i bruk terminologien i riktig kontekst, altså i setninger. Deretter må disse tekstene struktureres i store nok datasett, slik at de kan tilgjengeliggjøres som treningsdata for språkmodeller.

Til slutt skjer selve treningen av språkmodellene.

Norsk og engelsk

Norsk språk er et av verdens hundre største språk - men likevel et minispråk sammenliknet med engelsk. Den totale mengden tekst som produseres på norsk er forsvinnende liten i forhold til storebror, av dette er det tilsvarende lite som er åpent tilgjengelig, av dette igjen er bare en brøkdel fagtekster fordelt på et stort antall ulike fagfelt, og bruken av konsistent fagterminologi er usikker.

Det gjør at vi må jobbe systematisk og smart for å utvikle språket vårt og for å tilgjengeliggjøre tekster som kan gjøre språkteknologi på norsk god og brukelig på ulike fagområder.

Fagkompetente modeller

Det er altså stort behov for mer fagtekst for trening av fagkompetent språkteknologi på norsk. Nøkkelen til tilgang på fagspesifikke tekster ligger delvis hos virksomhetene som opererer i et gitt fagfelt: Fagspråket må struktureres, innarbeides, brukes i tekst, og tilgjengeliggjøres.

Men den ligger også i den enorme digitaliserte samlingen til Nasjonalbiblioteket.

Gjennom forskningsprosjektet 'MIMIR' (se egen artikkel) skal Nasjonalbiblioteket i løpet av 2024 måle effekten det har å inkludere også opphavsrettsbeskyttet materiale i treningen av generativ KI på norsk.

Resultatene av dette prosjektet blir interessante for vurderingen av hvor sterke tiltak som bør igangsettes for å oppnå fagkompetente språkmodeller, og i oppdraget NB har fått ligger å utvikle forslag til en kompensasjonsordning til rettighetshaverne ved bruk av tekstene deres i trening av KI-modeller.

Hovedmålet i den norske språkpolitikken er at norsk skal være et samfunnsbærende språk - brukelig og i faktisk bruk på alle samfunnsområder. Da trenger vi å sette alle kluter til for å anskaffe fagtekstene som gjør de generative modellene språklig kompetente og faglig pålitelige på vårt eget språk!





Large Language Models at University of Oslo



Erik Velldal
Professor, University of Oslo

Many have declared 2023 as the year of Large Language Models (LLM), and it's hard to disagree. In the Language Technology Group (LTG) at the University of Oslo (UiO), developing language models for Norwegian has been an important priority for several years. While also a NorwAI partner, LTG has not been involved in the LLM efforts of the center. Nonetheless, language modeling has defined our activities in several other collaborations.

While basically all language models these days are based on the so-called Transformer architecture, they come in three main flavors: so-called encoder models, encoder-decoders, and decoders. Simplifying a bit, one can think of the difference as follows. Encoder models, with BERT being the most well-known example, are great for analyzing and forming representations of text. Decoders, like GPT-type models, are geared toward generating texts, while encoder-decoder models like T5 can do both.

Models of all flavors

At LTG, we have trained Norwegian models of all flavors, starting with our first NorBERT models back in 2020, followed by NorT5, and most recently the GPT-like NorMistral and NorBLOOM models that we trained in 2023. An important principle has been to make all models publicly available without any restrictions on use.

Our Norwegian versions of the popular Mistral and BLOOM models are the biggest we have trained so far, with 7B parameters. For training data, we relied exclusively on publicly available sources of Norwegian text, combining the Norwegian Colossal Corpus (NCC) made available by the National Library, combined with various sources of web-crawled text and programming code. While some of the models were trained from scratch for Norwegian, we also trained a version of NorMistral initialized from a pre-trained English model, but still using a custom tokenizer specifically for Norwegian. This “warm-started” model has proved to have the best performance in our evaluations so far.

Preliminary GPT-like models

Importantly, we consider these GPT-like models to be preliminary: So far, they are only pre-trained base models, and we will soon release



... language modeling has defined our activities in several other collaborations.



improved versions that have undergone additional instruction-tuning and preference optimization.

Both next steps in the training pipeline are crucial to arrive at the type of open-ended chat-based problem solving one has come to expect from models like ChatGPT. We have previously released an instruction-tuned version of our Norwegian T5-model, Chat-NorT5, based on machine translated data. It is clear, however, that creating high-quality specifically for Norwegian is needed to fully benefit from these additional fine-tuning steps.

As is so often the case, the bulk of the heavy lifting is carried out by PhD- and postdoctoral fellows, and these LLM efforts are no exception: LTG PhD-fellow David Samuel and postdoc Vladislav Mikhailov have been vital in carrying out both the training and evaluation of the models. Another vital factor has been the computing infrastructure.

LUMI, a supercomputer

In the age of LLMs and deep learning, access to high-performance computing (HPC) facilities with sufficient GPU capacity is crucial to stay competitive. In early 2023, LTG was invited to take part in the pilot testing of the GPU-partition of the new supercomputer LUMI. At the time of writing, LUMI is by far the fastest

supercomputer in Europe and ranked 5th globally. While physically located in Finland, the cluster is hosted by a consortium that includes Norway (represented by Sigma2) among nine other European countries.

In Norway we have generally had the luxury of being quite well off when it comes to national HPC facilities, and LTG has a long tradition of putting these to good use. Still, there is no denying that LUMI has opened entirely new possibilities. When we trained our first BERT-models on the Saga cluster back in 2020, it would take us two months to train a 100M parameter model. On LUMI, we can now train the same model in six hours. And it takes us two weeks to train our 7B parameter GPT-like models.

Benchmarking

An essential part of model development is benchmarking, i.e., being able to systematically evaluate and compare different models across different downstream tasks. While LTG has already created many resources for benchmarking Norwegian language models, like those collected in the NorBench suite we released in 2023, this is mostly geared towards encoder-based models, and we are currently focusing on adapting and extending this to generative models.

These efforts will be intensified in 2024 through the ongoing collaboration between NorAI and LTG in the “Mímir” project – a national project including Nasjonalbiblioteket - explained in the following article.

The project “MIMIR” on copyrighted content

At the end of 2023, an initiative emerged that brought the three most active environments in Norway with expertise in language models to collaborate more closely. The “Mimir” project united the National Library of Norway, the University of Oslo, and NorwAI in a joint effort.

On December 8, 2023, the National Library of Norway received the following assignment from the Ministry of Culture and Equality:

“We refer to the dialogue with the National Library about the development of generative artificial intelligence, and issues raised by the institution related to the use of copyright-protected material in the training of Norwegian language models. The Ministry hereby requests the National Library to initiate a coordinated research/development project to possibly investigate the value of copyright-protected material in the training of Norwegian generative language models. Relevant Norwegian research environments should be invited to participate in the project. Authors’ and publishers’ organizations are invited to follow the project. We ask the National Library, based on the results of the research project, to consider the basis for a possible compensation scheme for Norwegian rights holders, and possibly develop a proposal for such a scheme.”

On December 20, NorwAI had a meeting with the National Library, which briefed on the mandate from the department and aired the idea of a collaboration with NorwAI and the Language Technology Group (LTG) at the University of Oslo.

Into 2024

In a meeting on January 5, the National Library, NorwAI, and LTG agreed to run a joint project until the summer of 2024 to assess the value of copyright-protected content in large generative language models.

The compensation scheme is kept outside of this project. The National Library takes on the responsibility of creating one training dataset without copyrighted content and one with it. It is up to NorwAI and LTG to define which models need to be trained and which evaluations must be conducted to assess whether copyrighted content in the training data results in better language models.

As of writing (February), the project aims to report in the summer of 2024, and the intention is to use both NTNU’s infrastructure Idun and the joint European infrastructure LUMI in Finland for the training.



Jon Atle Gulla
Professor
NorwAI

How can (Norw)AI protect personal data?



Anders Løland
Research director
Norwegian Computing Center (NR)



Protecting personal information is challenging with complex AI models that are hungry for data. NorwAI's pledge to provide an individualized AI experience that provably respects privacy concerns is therefore more important than ever.

For traditional data, like tabular data, the privacy challenges are fairly well understood. With unstructured data like text, however, privacy is less well-defined. Take this sentence as an example:

"The applicant, Dr Royce Darnell, who was born in 1929, has been unemployed since the Trent Regional Health Authority ("the RHA") terminated his employment as a consultant microbiologist and Director of the Public Health Laboratory in Derby."

The name is a direct identifier, which is fairly easy to locate and possibly mask, while access to some of the phrases in italics can be seen as indirectly identifying information, depending on which external sources of information – like the Internet – the reader might have access to.

Without metrics or benchmarks to measure how well personal data in texts are protected it

is not possible to measure how a large language model deals with privacy or train/tune a large language model so that it does not remember too much personal information. And since privacy breaches can happen in many and sometimes mysterious ways, one single metric is not sufficient. At NR, also via NorwAI, we have therefore over several years been developing NLP (natural language processing) methods to:

- 1 detect entities (like names, addresses, etc.)
- 2 link entities (Dr Royce Darnell may be called just Royce in parts of the text, or even Roycy?)
- 3 detect confidential attributes (like sexual orientation)
- 4 measure how well a method is able to detect or mask personal information (that is evaluation metrics to protect personal data)

..we have constructed the TAB (Text Anonymization Benchmark)..



Such methods can for example be used when training or fine-tuning the NorGPT model to reduce the chances that the model contains or leaks personal data.

Open science is not possible without open data

Since personal data typically is not available for open science, we have constructed the TAB (Text Anonymization Benchmark); an open-source annotated corpus developed to address this shortage. The corpus comprises 1,268 English-language court cases from the European Court of Human Rights (ECHR) enriched with comprehensive annotations about the personal information appearing in each document, including their semantic category, identifier type, confidential attributes, and co-reference relations. The TAB corpus goes beyond traditional de-identification, and explicitly marks which text spans ought to be masked to conceal the identity of the person to be protected.

When GPT-4 was released in 2023, the TAB was utilised to test its anonymization performance (see the paper “Sparks of artificial general intelligence: Early experiments with GPT-4.”).

Is anonymization possible at all from a technical or legal viewpoint?

As mentioned, much of the legal and technical literature on data anonymization has focused on structured data such as tables. However, unstructured data such as text documents or images are far more common, and the legal requirements that must be fulfilled to properly anonymise such data formats remain unclear.

In the absence of a definition of the term ‘anonymous data’ in the GDPR (General Data Protection Regulation), we have – together with legal experts from the University of Oslo and the Oslo Metropolitan University – examined what conditions must be in place for the anonymization of unstructured data. We concluded that anonymization of unstructured data is virtually impossible as long as the original data continues to exist. Hence, for all practical purposes only a risk-based approach is realistic, meaning that we can estimate a certain privacy risk for a text corpus or a language model, but a strict guarantee is not possible.



Adjunct Associate Professor Jon Espen Ingvaldsen of NorwAI has played a key role in developing the NorLLM models introducing the thinking and architecture to numerous audiences in 2023. Photo: Ole Martin Wold

The path to NorLLM: The Norwegian Generative Languages Models

SFI NorwAI – the Norwegian Research Center for AI Innovation – is working intensively to develop generative language models that can benefit Norwegian society. Generative language models gained widespread attention in 2023 after the international breakthrough of the large language models from OpenAI. In this section, we summarize the current status of NorwAI’s work with NorLLM (Norwegian Large Language Models).

Our key messages are as follows:

- » Norway needs control over its own generative language models which are built on Norwegian data and values.
- » We have a well-functioning system for collecting and managing published content for use in large language models.
- » Lack of computational resources hinders both training and operation of large language models in Norway.
- » There is a need for structures and mechanisms to ensure that training data, fine-tuning data, and alignment methods align with Norwegian values and support open models.
- » NorwAI with its partners have the necessary expertise and experience and aims to develop Norwegian language models for the benefit of Norwegian society.

Data Scientist Lead Daniel Pleus at the Schibsted AI Enablement Team told the NorwAI Innovate audience of the contributions to NorLLM from the publisher giant at NorwAio Innovate 2023.



Four models built

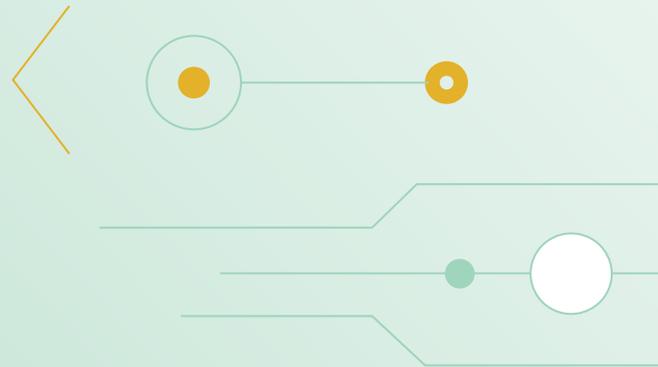
- three new ones in the pipeline



Jon Atle Gulla
NorwAI Center Director



Rolf Dyrnes Svendsen
Communication Manager, NorwAI



NorwAI has built four distinct Norwegian generative language models. During the winter of 2023/2024, an additional three models are being developed which will be made available in spring 2024. Collectively, these seven models represent steps toward NorwAI's ambition to build a comprehensive generative foundation model for general use, with approximately 40 billion parameters by the end of 2024.

Through proper training, the number of parameters captures subtle nuances in language, including orthography, syntax, semantics, and pragmatics. These parameters are learned during training, and more training data typically results in both more precise parameter values and more factual information from which to draw answers (and thus fewer hallucinations). The initial four models had increasing data availability and expanded functionality, as shown in the table below. These models handle both written forms of Norwegian, although they tend to default to Bokmål when lacking relevant responses in Nynorsk. The next three models will be trained on Norwegian data

without copyrighted content, using different architectures to uncover technological features and assess the practical suitability of each model architecture.

In 2024, NorwAI's language modeling efforts will diverge into two distinct tracks. There is a need to train a Norwegian generative model that is sufficiently large to support the Norwegian language at the same level as English and other major languages supported by international models. Simultaneously, there is a demand for smaller models that can be easily optimized for specific use cases and controlled locally using proprietary data and custom adaptations.



Cst. state secretary Tore O. Sandvik of the Trade and Industry Ministry has for a long time advocated Norwegian language models.



The parliamentarians Guri Melby and Grunde Almeland of Venstre at Stortinget visited NorwAI to learn more of the coming LLMs.



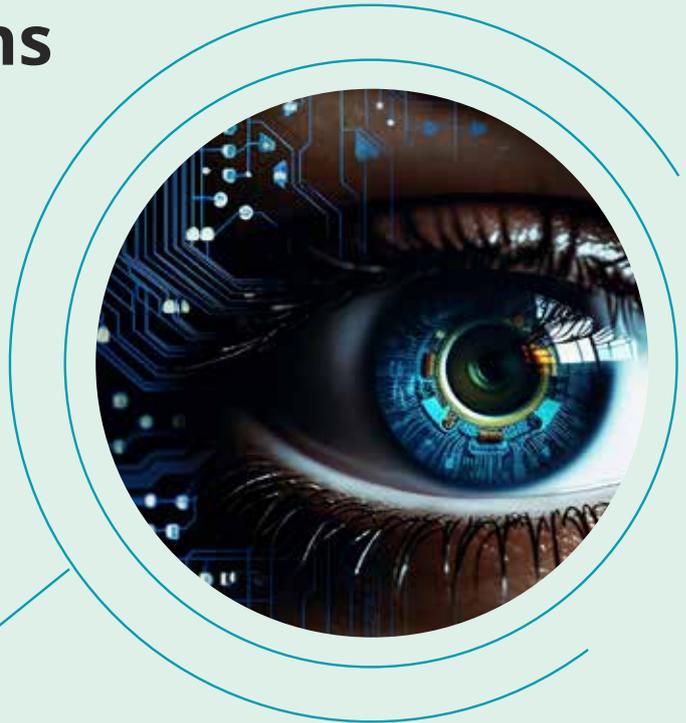
Karianne Tung, Minister for Digitalization spend a whole day at NTNU to update herself and give views on the frontrunning digital work of the university.

Demand for Norwegian models

NorwAI has been approached by several public organizations and private enterprises seeking an alternative to international models. These entities have primarily raised two concerns regarding existing commercial models: (i) handling sensitive and copyrighted data, and (ii) the lack of quality in Norwegian language generation. This suggests that Norwegian language models should be made available in various ways:

- » General, centralized Norwegian language model administered as part of a national infrastructure: Some desire a secure Norwegian model that can produce better and more reliable Norwegian texts. These models should be accessible via APIs and adhere to strict Norwegian standards.
- » Task- or domain-specific models (fine-tuning): Some inquire whether NorwAI can customize models for specific purposes or tasks. Others prefer to make these adaptations themselves and run the models locally without sharing their own data.
- » Models combined with proprietary data: The technique known as Retrieval Augmented Generation (RAG) enables the extraction of relevant data from an enterprise system and combines it with the language model to generate accurate responses. Local data is often sensitive, not used for model training, and typically more up-to-date and specific than the data in training sets.

Some observations from the language model work



TRANSPARENCY

International models often lack transparency. We don't have access to the details of their datasets, nor do we receive information about the methods they use for data cleaning, fine-tuning, model adaptation, or handling toxic or discriminatory language.

COPYRIGHTS

Most international models utilize copyrighted content without proper authorization. This includes for example Norwegian books available in online collections.

SUSTAINABILITY

Training large language models requires significant computational resources, substantial energy consumption, and involvement of many individuals. OpenAI has faced criticism for employing underpaid Kenyans to filter out toxic or biased language from their models. Training large language models can easily consume around 1,000 megawatt-hours of electricity.

VALUES AND NORMS

International models tend to reflect the norms and values of dominant languages in their training data. This manifests itself in the types of questions the models avoid answering and, more significantly, in the overall meaning and balance of the generated text.

LANGUAGE VARIANTS

Language models are surprisingly good at distinguishing between different languages. Abstractions from one language in the model appear to assist in handling related languages. Training texts in Bokmål (one of the Norwegian written forms) seem useful for generating responses in Nynorsk (another Norwegian written form), while Finnish training data can help address challenges in collecting sufficient textual data in Sami.



Do we need Norwegian models?

Generative artificial intelligence (AI) is advancing rapidly. The generic properties of language models suggest that they will be integrated into many digital services, contributing to efficiency and productivity across private and public sectors. Currently, foreign commercially driven companies offer these services, quickly establishing habits among large user groups. Consequently, Norway becomes nationally vulnerable to digital dependence on external entities.

Today, international actors control data, training, customization, operation, and distribution of large generative language models. These models mirror the values and norms they are trained upon. Although currently closed systems, they are poised to become part of future infrastructure. When assessing their significance, democratic control is a crucial perspective.

Web vs. Domain-Specific Language

International models are multilingual and primarily trained on English. However, they lack substantial data for smaller languages. Most of their training data is scraped from the internet, failing to capture the depth and breadth of knowledge and expertise found in Norwegian academic and fictional texts.

Norwegian remains a vibrant and essential language, used across all significant domains in Norwegian society. Norway is also a multilingual nation, with Sami and Norwegian as national languages, including multiple written forms. NorLLM's models have demonstrated

proficiency in both Bokmål and Nynorsk. Experiences gained from incorporating related languages into the training data enhance the possibility of building a multilingual, national model.

In models developed by international tech companies, Norwegian plays a subordinate role. These foreign companies have little incentive to accommodate the specific needs of such a small, multilingual society like Norway's. If this trend continues, the ability of the Norwegian languages to be fully functional across all important life domains will weaken.

NTNU and NorwAI operate within the Norwegian tradition, where assumptions and conditions are agreed upon, and the parties mutually exchange information and insights. A Norwegian language model must consider our society's principles and regulations. Material with artistic significance, domain-specific literature, and culturally valuable content are protected and regulated through copyright. Privacy holds a strong position in our country. While these values may not be automatically respected by international actors, they can be better safeguarded through a national solution for language models.

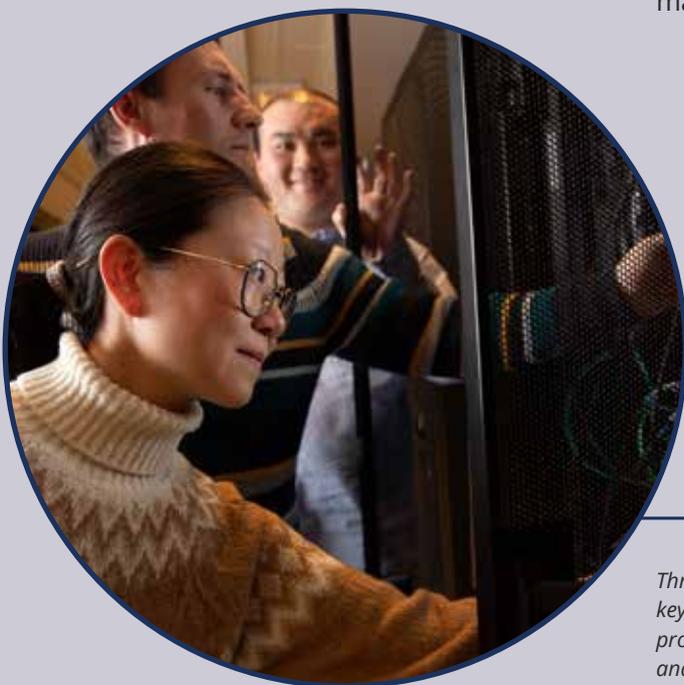


IDUN at work at night - NTNU has invested vastly in the build-up of the super computer cluster proving it's competence to build big, generative models.

Too little, but not too late

The training of language models in NorwAI is organized from the research center in Trondheim and is carried out by NTNU's scientific staff in collaboration with selected partners such as Schibsted. In addition, NTNU provides access to the supercomputer cluster Idun (<https://www.hpc.ntnu.no/idun/>) and substantial technical expertise. Idun, a collaboration between NTNU's faculties and the IT division, constitutes a professional infrastructure for high-performance computing that is well-suited for computationally demanding AI models.

However, available supercomputer resources in Norway are limited, and neither Idun nor Sigma2's clusters (<https://www.sigma2.no/>) are currently powerful enough for long-term training of language models. For instance, it took NorwAI 76 days to train NorGPT-23 on Idun in 2023. Some institutions, like the University of Oslo, utilize the Finnish LUMI supercomputer (<https://www.lumi-supercomputer.eu/>), but Norway has limited access to this infrastructure, which may also be less flexible for our language model work. Regardless, these clusters are primarily configured for research and cannot serve as a solution for model deployment and maintenance in the long run.



Three NorwAI post docs, here at work at IDUN, has played key roles in the NorLLM project: responsible for the training program Lemei Zhang (front), project leader Benjamin Kille and technical leader Peng Liu (back) has literally built the models from scratch.

Newcomers with daunting tasks



Vilhjalmur Arnar Vilhjalmsson

Research Assistant
NorwAI



Even Wienenga Lauvrak

Research Assistant
NorwAI

This summer, we started our chapter as fresh research assistants at NorwAI. Armed with nothing but our sprawling excitement and hunger for learning, we set out to aid the research of AI in Norway in any way we could.

What started as a daunting task turned into one of the best learning experiences any student could have asked for. We got to build close relationships with professors, post-docs, and the rest of the team at NorwAI, as well as work alongside them in their respective work. All this work has since accumulated in major contributions to the NorGLM project, even co-authoring papers.

My experience (Vilhjalmur):

Throughout the summer, I was tasked with handling the massive dataset used to train the NorGLM model; more precisely, making the NorGLM dataset easily accessible and transparent to the public. Handling a dataset of this scale requires careful management, and I devoted considerable time to optimizing every aspect of its processing.

Highlights

One of the highlights was an optimization I implemented in a processing pipeline, which reduced the processing time from over an hour to less than 30 seconds. I worked on the entire stack of the effort, developing the frontend, setting up the backend, and seamlessly integrating everything.

Taking a step back and reflecting on my time so far at NorwAI, the thing that stands out is the people I have the privilege of working alongside. Not only am I continuously impressed by the academic and intellectual prowess, but also by how kind and welcoming people are. One moment, we're diving deep into advanced AI concepts; the next, we're sharing laughs by the coffee machine.

My experience (Even):

Starting from the first week I was given the task of fine tuning and evaluating one of the NorGLM models, as I did not have the greatest expertise on the field of LLMs I were given the freedom to do my own research and learn with guidance from the team. After trial and error, I had learned a lot and my main task from there on was to fine tune and evaluate the different parameter models of NorGLM as well as other baseline models for comparisons.

This led to my contribution to a collaborative research paper that we submitted in December.

No wrong questions

I would say that my favourite part about NorwAI is how open everyone is and that there's no wrong questions. It may seem intimidating to ask professors or post-docs questions that you think you should know, but they are just happy to have a talk and help, and that what's make working as a research assistant at NorwAI so great.

*Even Wienenga Lauvrak and Vilhjalmur Arnar Vilhjalmsson are enjoying their work as research assistants.
Photo: Kai T. Dragland, NTNU*

Working as research assistant gives you a lot of opportunities that awaits to be taken. How much you want to do and at what level depends on your eagerness to look out for tasks, tasks that mainly comes from talking to the different professors and post-docs, showing interest, and asking to contribute. And the best part is that everyone is open, friendly, and ready for a chat. Some of the best memories we have made from the office is from sitting around the lunch table with all different kinds of people from the office and talking nonsense and just having a good time before we head back to exciting tasks.





A POWERHOUSE

*Some of the largest
and technologically most
ambitious companies
and research institutes in
Norway have joined
the consortium.*



Consortium

RESEARCH PARTNERS

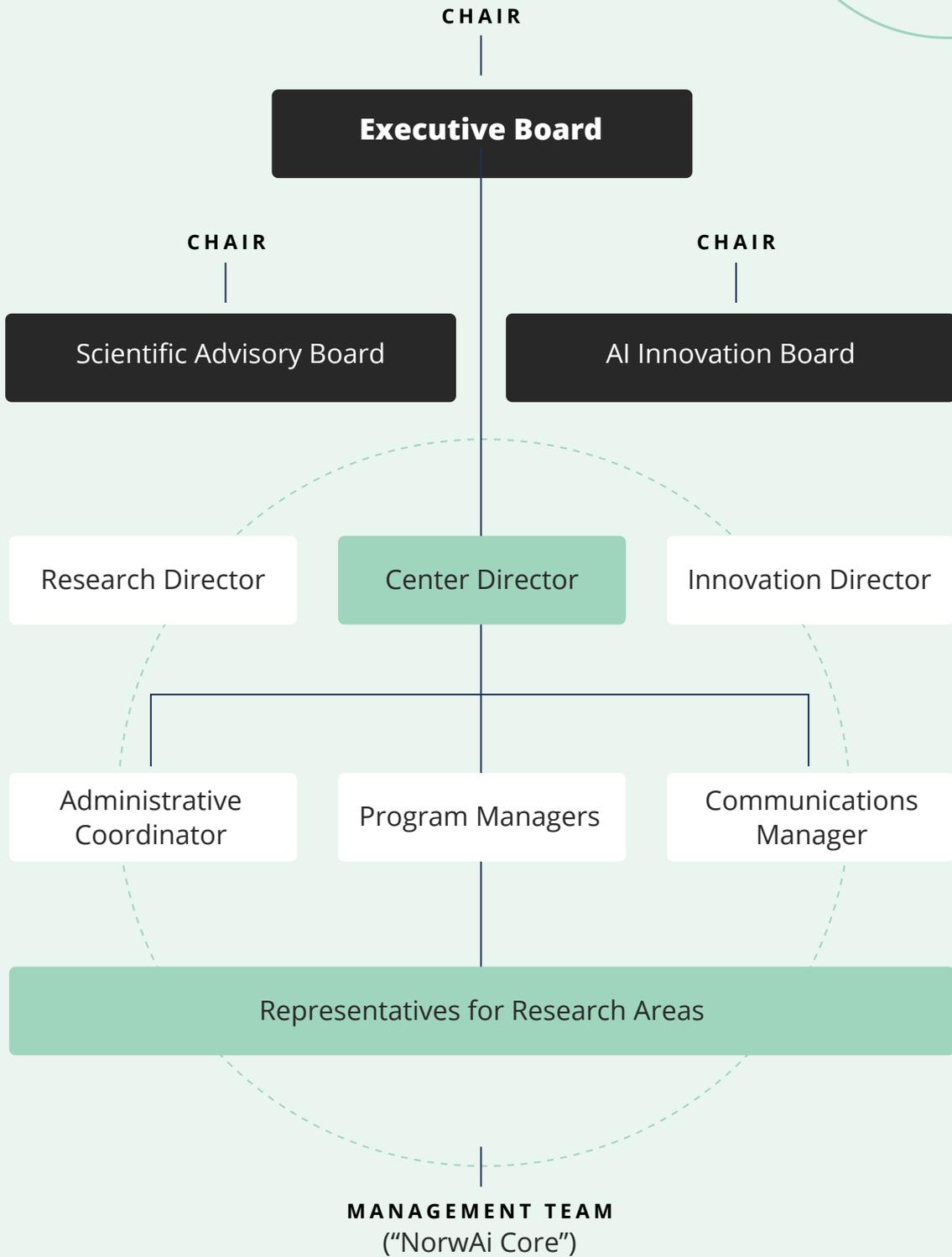
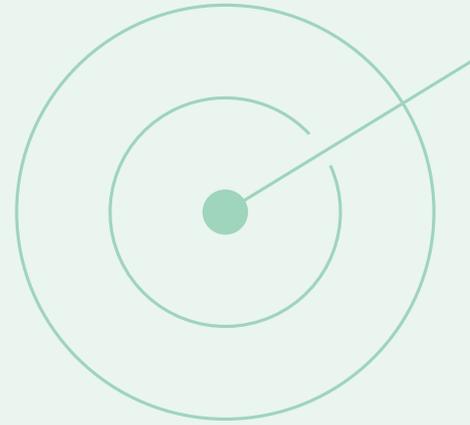
NTNU, the Norwegian University of Science and Technology, Department of Computer Science is host for the center, the other research partners are Norwegian Computing Center (NR), SINTEF, University of Oslo and University of Stavanger.

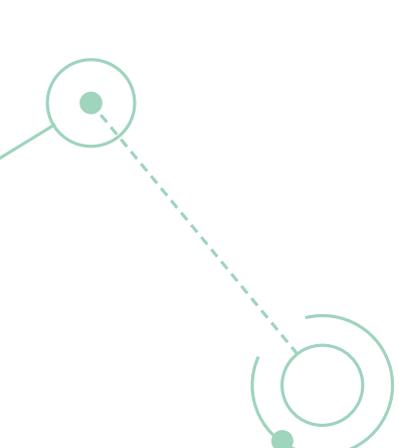
INDUSTRIAL PARTNERS

The group of industrial partners in Norway consisted in 2023 of ANEO, Cognite, Digital Norway, DNB, DNV, Kongsberg Digital, NRK, Retriever Norway, Schibsted, SpareBank 1 SMN, Statnett and Telenor.



Organization





Center Management Team

*The Center Management Team
is responsible for the day-to-day
operation of the center and
consists of:*



Jon Atle Gulla
Professor, NTNU, Center director



Benjamin Kille
Associate professor, NTNU,
Program Manager (from Nov -23)



Kjetil Nørvåg
Professor, NTNU,
Research director (until July -23)



Helge Langseth
Professor, NTNU, Research Area
Representative (until July -23)



Kerstin Bach
Professor, NTNU,
Research director (from Aug -23)



Signe Riemer-Sørensen
Research Manager, SINTEF,
Research Area Representative



Arne Jørgen Berre
Chief Scientist, SINTEF,
Innovation Director



Frank Alexander Kraemer
Associate Professor, NTNU,
Research Area Representative



Karolina Storesund
Senior Advisor, NTNU,
Administrative Coordinator



Jon Espen Ingvaldsen
Adjunct Associate Professor, NTNU,
Research Area Representative



Rolf Dyrnes Svendsen
Head of NxtMedia Lab,
Communications Manager



Terje Brasethvik
Adjunct Associate Professor, NTNU,
Research Area Representative



Özlem Özgöbek
Associate professor, NTNU,
Program Manager (from Nov -23)

The mysteries of Hybrid AI



The Valhall platform. Photo: Aker BP

...Cognite and Kongsberg compete daily in the Norwegian oil and gas market, but in NorwAI they have found a neutral arena...



Alexander Stasik
SINTEF Digital



Signe Riemer-Sørensen
SINTEF Digital



Andris Piebalgs
Cognite



Simone Casolo
Cognite



Eivind Roson Eide
Kongsberg Digital

2023 has seen heat records in Svalbard and extreme rainfall and floods in Norway (and other places). To prevent climate run-away, there is a pressing need to reduce the greenhouse gas emission from energy production.

Achieving this requires enhanced modeling of production plants, so they can be run optimally and serviced just-on-time to extend their lifetime without unnecessarily exchanging parts or downtime. This becomes crucial as renewable energy increases the dynamics of the energy system and will only be possible if we use AI in combination with existing physical knowledge in so-called hybrid AI .

The challenge of simulations

A digital twin is a digital model of a physical system which for a specified purpose, such as monitoring or predictions of what-if scenarios, serves as an indistinguishable digital counterpart of the physical system. For predictive tasks, digital twins often rely on heavy numerical modelling or simulations.

Simulations are based on first-principles models with model parameters that do not necessarily correspond to measurable properties of the

system e.g. the energy in each part of the system has to be interpreted through temperature and pressure at specific locations. The challenge lies in accurately determining model parameters that correspond to the actual system state.

The current workflow entails trial and error with reasonable guestimates until the simulation output matches the observed data. This matching effort scales exponentially with the number of input parameters, which is typically between tens to hundreds. Further, the system may change over time, so this matching must be carried out continuously.

The use case

The use case considered the first compressor stage on the Valhall platform with data provided by Cognite and simulator provided by Kongsberg through their k-spice framework. The compressor train refines the oil by removing water and gas.

It is a complicated process, sensitive to pressure, temperatures, composition and flow rates, and internal parameters of the specific equipment as well as time-varying degradation factors. Despite knowledge of the precise engineering layout, Kongsberg spend significant manpower setting up simulations to reflect the plant.

A series of methods

Simulation based inference is a series of methods where simulations are used to estimate unobservable parameters of a system. Each simulation is evaluated against observed data and its 'correctness' is quantified as 'likelihood' which is the probability of the observed data given the simulation input (similar to the loss function in a machine learning model).

The challenge is then to find the input parameters that maximize the likelihood without the likelihood function being available. Instead, we use machine learning techniques to estimate the unknown likelihood function from individual evaluations of the simulator and use this learned model to perform parameter

estimation. With only a few simulations, the estimate of the likelihood function will be crude, but each new simulation will improve it.

This approach drastically reduces simulation evaluations and therefore time, compared to brute force sampling of the simulator. Also, the probabilistic models provide confidence intervals for the parameters, highly relevant in safety-critical applications.

Toy example

A typical benchmark in process industry is the continuously stirred tank reactor where you have a tank with multiple chemical substances which are participating in an exothermic chemical reaction. For demonstration, one partner simulated a data set and gave it to SINTEF without the input parameters. A classic least squares optimization approach required **20 000 simulations** to identify the inputs. The simulation-based inference with a Gaussian processes model for the likelihood mapping required only **300 simulations**, showcasing a significant reduction in computational efforts.

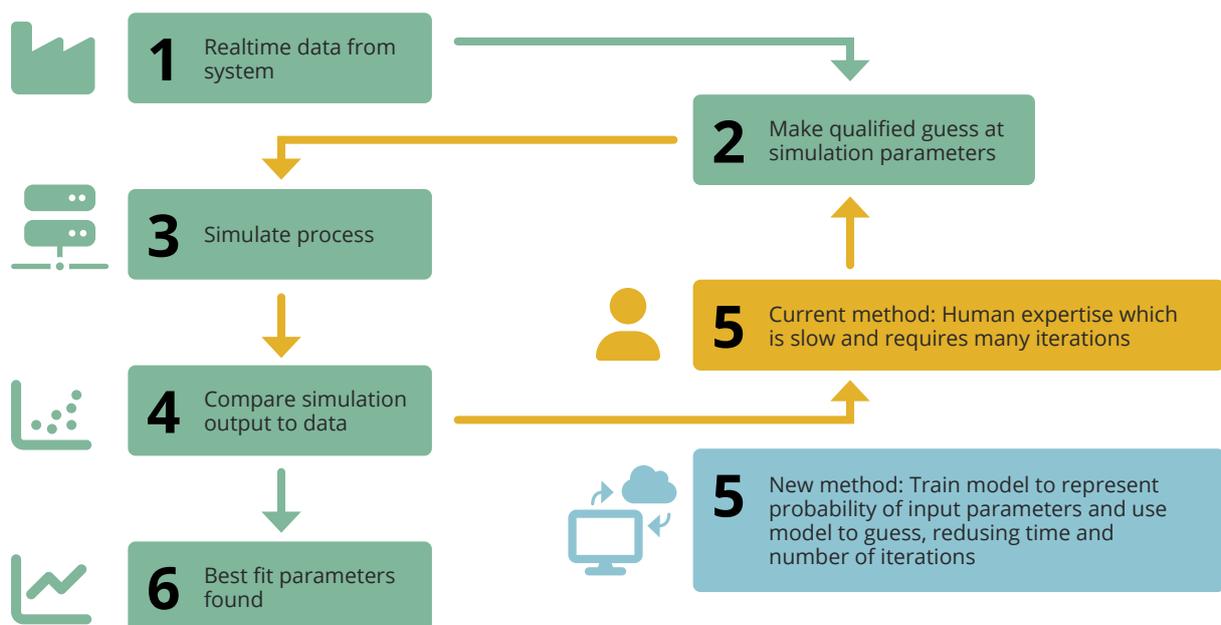


Illustration of the simulation tuning process. With the current manual method, the parameters are manually tuned for each iteration of the loop (2-4) which is not very cost efficient. The new method learns a probability distribution from the comparison to data and use the learned probability to guess the next trial parameters. This is very efficient and reduces the number of iterations significantly.

Real life application

The setup is currently being implemented by Kongsberg with data and pre-processing provided by Cognite. Cognite and Kongsberg compete daily in the Norwegian oil and gas market, but in NorwAI they have found a neutral arena where they can focus on challenging problems with SINTEF Digital as moderator and research partner.

Together, everyone has contributed to the development of the simulator-based interference model, Cognite with their colossal amounts of data and Kongsberg with good frameworks for simulation. Meeting places where industry get to voice their challenges and build relations from similarity in problems will become even more important with new regulations such as the EU AI Act which will require industry to document the safety of the algorithms used in critical infrastructure.

Upon successful implementation, the next steps involve applying the same setup to K-spice-based electric grid simulations in collaboration with Statnett, and Bladed wind turbine simulations provided by DNV with data from Aneo. Better tuning of the simulators can enable better condition monitoring and maintenance predictions, to reduce waste of resources from industry.



Signe Riemer-Sørensen
Photo Kai T. Dragland, NTNU



21.96

73.14

21.96

25.99

63.10

57.99

39.17

19.68

68.35

56.44

33.97

56.28

97.76

47.51

47.51

97.76

75.99

56.28

58.7

95.38

16.48

27.86

34

19.68

40.82

NorwAI delivers



➤ **Pål S. Malm**
Senior Project Manager
SINTEF Digital

In 2022, attempts were made to defraud DNB customers of NOK 1.2 billion. With the help of AI and other tools, the bank managed to save one billion of this. No wonder NorwAI's partners are impatiently waiting to implement the AI models they see prototyped in the project.

Artificial Intelligence is an enabling technology with short distance from cutting-edge research to innovations. The launch of ChatGPT in November 2022 has caused awareness of generative AI and large language models to skyrocket. The accessibility of generative AI has lowered the barrier to conduct internal experiments, and many of our partners' innovative initiatives are pertinent to our NorwAI.

Innovation happens through discussing new ideas, sharing new knowledge and collaborating on developing technologies that can be used for industry partners. The term "innovation" has different meanings in academia and industry. According to the EU Innovation Radar, industry innovation is usually focused on "business-ready" and "market-ready" technology, while

academia defines innovation as "exploring innovations" and "tech-ready" technology. Research and technology organizations (RTO) like SINTEF, is often bridging academia and industry. Use cases are useful to describe the "job that needs to be done". In NorwAI Annual Report 2021 this innovation approach was described as a way to start with the user, brand need or problem, as opposed to having data or technology as a starting point.

Innovation elements

The Research Council of Norway's definition of innovation describes most of the elements we associate with innovation in NorwAI:

"New or significantly improved products, processes, production or distribution processes, organizational or management forms, or

concepts introduced to enhance value creation or benefit to society.”

NorLLM is an innovation that has had a positive impact beyond its technical capabilities. Norwegian language model is on the political agenda, and there is public engagement for a national LLM for cultural preservation and representation for public services and education. Furthermore, we have gained an increased understanding and sense of the urgency of making national content and training data available.

“Using AI, DNB has saved itself several hundred million kroner in recent years, and customers one billion, in 2022 alone. The technology they use is cheaper and more accessible than ever.”

– Maria Ervik Løvold (COO DNB) to Dagens Næringsliv

DnB says, for example, that they have benefited greatly from NorwAI and NorLLM. Via NorwAI, they have had access to researchers with hands-on experience for DnB’s data science community to discuss to see how large language models in general can be tailored for internal use.

Already experimenting

DnB and Sparebank1 SMN are experimenting with AI and language models to detect fraud and financial crime. Attempts were made to defraud DnB customers of NOK 1.2 billion in 2022. With the help of AI and other tools, the bank managed to save one billion of this.

Criminals use AI to make more sophisticated fraud attempts and banks need to be more proactive in preventing financial crime. However, legal regulation restricts the use of transaction data internally and the sharing of data with other banks, although this could improve banks’ ability to detect fraudulent activity. Banking, finance and telecom have similar issues; They want to utilise internal data to improve services and customer dialogue and want to share data with other banks to prevent fraud and financial crime but find that sharing data is a demanding regulatory area.



EXAMPLES OF INNOVATIONS

DNV: AI-enabled systems Assurance describes a framework for securing systems with AI. It provides guidance on how to ensure systems with AI are reliable and managed responsibly throughout their lifecycle. It is part of DNV's digital recommended practice that provides guidance at all stages of the digitalisation journey.

1 Other examples: in the oil sector, one must have control of flows through the system. **Cognite and Kongsberg** collaborate to solve common challenges within Hybrid AI. Cognite with colossal sensor data volumes, Kongsberg with good frameworks for simulation models. Together with SINTEF Digital, they have developed a simulator-based interference model with the purpose of matching simulations to actual data. The model

serves as neutral ground for bringing the best of both worlds. The next step is to use the same framework with other partners.

- 2 Similarly, operation of wind turbines will be optimized. **Aneo** has wind power data and **DNV** has a simulation tool.
- 3 Also, the transmission capability of the power grid will be modelled. **Statnett** also do research within grid optimization.

All these three challenges require robust models of the systems, but data quality and simulation models are two basic challenges. Do computer-based models provide the right picture when copying critical infrastructure with digital twins? (to more details, have a look on the following article from Signe Riemer-Sørensen at SINTEF Digital).

Sparebank1 SMN are experimenting with AI to provide better savings advice to end users. They want to improve customer dialogue via their chatbot "Robot-Anne."

1 In a pilot called PT (personal trainer) customers receive simple advice on how to change their consumption patterns and save money. Kjersti Wold, head of Advanced Analysis, says that SMN has ambitions to expand the pilot on personal trainer on savings advice to include more areas for personalized advice, but need to do the regulatory clarifications before they develop it further.

2 Astrid Undheim, Executive Vice President of Technology and Development, states in SMN's own blog that they also use AI for background checks of new customers and as a tool in the fight against financial crime.

Schibsted: is testing available open-source Large Language Models to evaluate and deploy them into products.

- 1 "Briefly explained", an automatic summary of articles and transcribing from audio to text, a service that has saved the company for 16 000 man hours in 2023.
- 2 In the schibLM project a set of models are trained on Schibsted data to generate features in articles and ads based on other features.
- 3 Schibsted also instruct models to write front titles on a test basis.
- 4 In **FINN.no**, Schibsted is experimenting with being able to provide more relevant recommendations despite limited input data.

The Arendal experience

Cooperation between partners



Rolf Dyrnes Svendsen
Communication Manager
NorwAI

The line outside the seminar room was long. There was excitement in the air. Curiosity and expectations were high. It was premiere for the first public demonstration of NorwAI's 23 billion parameters generative language model, the NorGPT 23B, the largest in its kind in Norway in August 2023.

Partners DNB and Schibsted had joined hands in the making of the model, and now stood side by side for the presentation. Present was also state secretary Gunn Karin Gjøl who bluntly talked about her salvation to become a fan of Norwegian language models when asked. Her boss Sigbjørn Gjelsvik later joined in on the panel and agreed.

The cooperation between DNB and Schibsted proved the "NorwAI DNA" true. Building the consortium, companies that would invest personnel and know-how to accelerate AI was invited in to explore unknown technology territory. Both companies gave vital contributions to bring a Norwegian language model to life and furthermore laid the foundation to the NorLLM projects that now follow up the Arendal premiere.

Newcomer join in

Cross company cooperation and making use of mutual experiences was one of the reasons for Statnett to join the NorwAI consortium in 2023. The electricity grid giant has joined the energy cluster of big, industrial asset companies that constitute the industry group of NorwAI. They will bring new impulses to the established and unique cooperation between related and competing companies that has found common ground in the SFI.

Organizing the internal NorwAI Forum for partners twice a year has accelerated knowledge exchange between partners. The forums that emphasize innovation and research

Minister Sigbjørn Gjelsvik (right) of Local Government and Regional Development, joined in on the NorGPT premiere in Arendal. Also presenting Yngvar Ugland, Executive Vice President, DNB NewTechLab, Jon Atle Gulla, NorwAI director, Sven Størmer Thaulow, Chief Data and Technology Officer/ EVP at Schibsted and Rolf Dyrnes Svendsen, NorwAI Communication Manager. Photo: NorwAI





State secretary Gunn Karin Gjøl and NorwAI Executive Board chair Sven Størmer Thaulow agreed on the importance of Norwegian language models. Photo: NorwAI

respectively, also build personal relations between the many people involved in NorwAI, thus increase the opportunities to collaborate in new ways.

Workshops

The forums are supported with special workshops where companies dive deeper into more specific topics of common interest.

Furthermore, the NorwAI Innovate Conference has proved its success as Norway's most influential AI event by gathering a blend of business professionals, academics and AI student talents for the third time in a row.

The NorwAI consortium consisting of 11 companies and 5 research institutions represents considerable breadth with respect to business needs, experience with artificial intelligence and capabilities for commercializing technological research results to capture value. NorwAI's research strategy has a sensitivity towards this variety. Consequently, all NorwAI's

work packages always have more than one partner in every research area to promote cross industry and cross industry and academic cooperation.

Balanced leadership

NorwAI have strived to balance leadership in the work packages. Both experienced industry leaders and senior researchers are appointed leaders for the different groups. This applies for all three research areas such as Thrustworthy AI and AI in society, the language technologies as well as the work packages for asset industries.

Cooperation between research and business is key to a research-driven innovation center. To NorwAI, industry's own use cases identifying possibilities for innovation, is base for the work. Industry partners have listed their use cases, and efforts are invested in coordinating and establish a common ground for use case owners and the researchers designated to the processes.

Members of the board

CHAIR EXECUTIVE BOARD



Sven Størmer Thaulow

Chief Data and Technology Officer,
Schibsted Media Group

MEMBERS OF THE EXECUTIVE BOARD (2023)



Ingelin Steinsland

Vice Dean Research, Faculty of Information
Technology and Electrical Engineering, NTNU



John Markus Lervik

Chief Strategy and development
officer, Cognite



Sergei Savenko

CTO, Retriever



Liv Dingsør

CEO, Digital Norway



Frank Børre Pedersen

Vice President & Program Director
"Future of Digital Assurance", DNV



Stephan Oepen

Head of Department of Informatics,
University of Oslo



Trond Runar Hagen

Research Director, SINTEF



Birgitte Kvammen

IT Project and Portfolio Manager, NRK



Dagfinn Myhre

Senior Vice President, Telenor Research



Astrid Undheim

Executive Director, Sparebank1 SMN



Stein-Roar Skånhaug Bjørnstad

CTO, Kongsberg Digital



Tom Ryen

Head of Department of Electrical
Engineering and Computer Science,
University of Stavanger



Gøril Forbord

EVP Technology, Aneo



Anders Løland

Research Director, NR



Karl Aksel Festø

Head of CoE Advanced Analytics, DNB



Nenad Keseric

SVP Director Innovation and
Technology Development, Statnett

The Scientific Advisory Board

An Ambitious Board

The overall goal of the Scientific Advisory Board of NorwAI is to provide external scientific reviews of research activities, evaluate plans and progress, and contribute to shaping the center's research ambitions. The chairperson of the Scientific Advisory Board is professor Christian S. Jensen at Aalborg University.

CHAIR OF THE SCIENTIFIC ADVISORY BOARD



Christian S. Jensen
Professor, Aalborg University

MEMBERS OF THE SCIENTIFIC ADVISORY BOARD



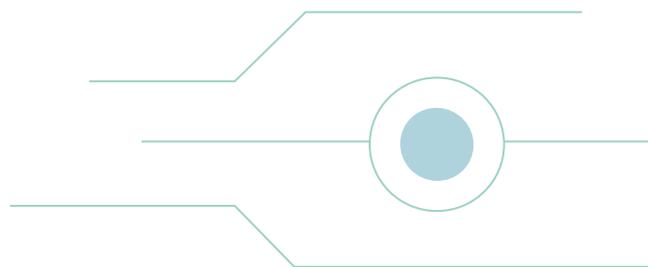
Concha Bielza
Professor, Technical University of Madrid



Maarten De Rijke
Professor, University of Amsterdam



Virginia Dignum
Professor, Umeå University



The Innovation Advisory Board

Exclusive group of innovators to advise NorwAI on innovation

The Innovation Advisory Board (IAB) will provide advice on how to create innovations from research for the NorwAI partners. The IAB will follow up on the innovation results on a regular basis and help to monitor the overall progress following SFI success criteria for innovation and commercialization. The Innovation Advisory Board is planned to meet with the Center director and the Innovation director twice a year.

The Innovation Advisory Board consists of a generous mix of internationally recognized experts from both abroad and Norway as its advisors. The Norwegians are well connected to the industry partners in NorwAI. The other two are prominent members of the AI community in both Europe and the US East Coast.

CHAIR OF THE INNOVATION ADVISORY BOARD



Ieva Martinkenaite
SVP Head of Research and Innovation, Telenor

MEMBERS OF THE INNOVATION ADVISORY BOARD



Sophie V. Vandebroek
Board Director, Trustee, Advisor



John Markus Lervik
Founder & Chief Strategy and development officer, Cognite



Saskia Steinacker
Global Head Strategy & Digital Transformation, Bayer



Øystein Larsen Indrevær
Senior Vice President Data-driven Sales, DNB



Research Activities

Visions and Views



I really like the excitement to explore the frontiers of our domain and our knowledge.

Kerstin Bach

Professor in AI
Research director at NorwAI

- We work on more than large language models at NorwAI. Our current research covers a variety of AI methods and applications through which we aim to create benefit of our partners and value for the Norwegian society.

Kerstin Bach, research director of NorwAI, emphasizes the diversity of the research activities when asked on her views of the SFI. She says NorwAI reflects the generic character of the AI methodologies:

- Our ongoing work spans over various domains, including telecommunications, finance, and energy, where we delve into analyzing streaming data. Specifically, our research initiatives encompass data platform development, anomaly detection, and hybrid AI. Additionally, we explore the development of reliable and ethical AI applications while examining their societal implications.

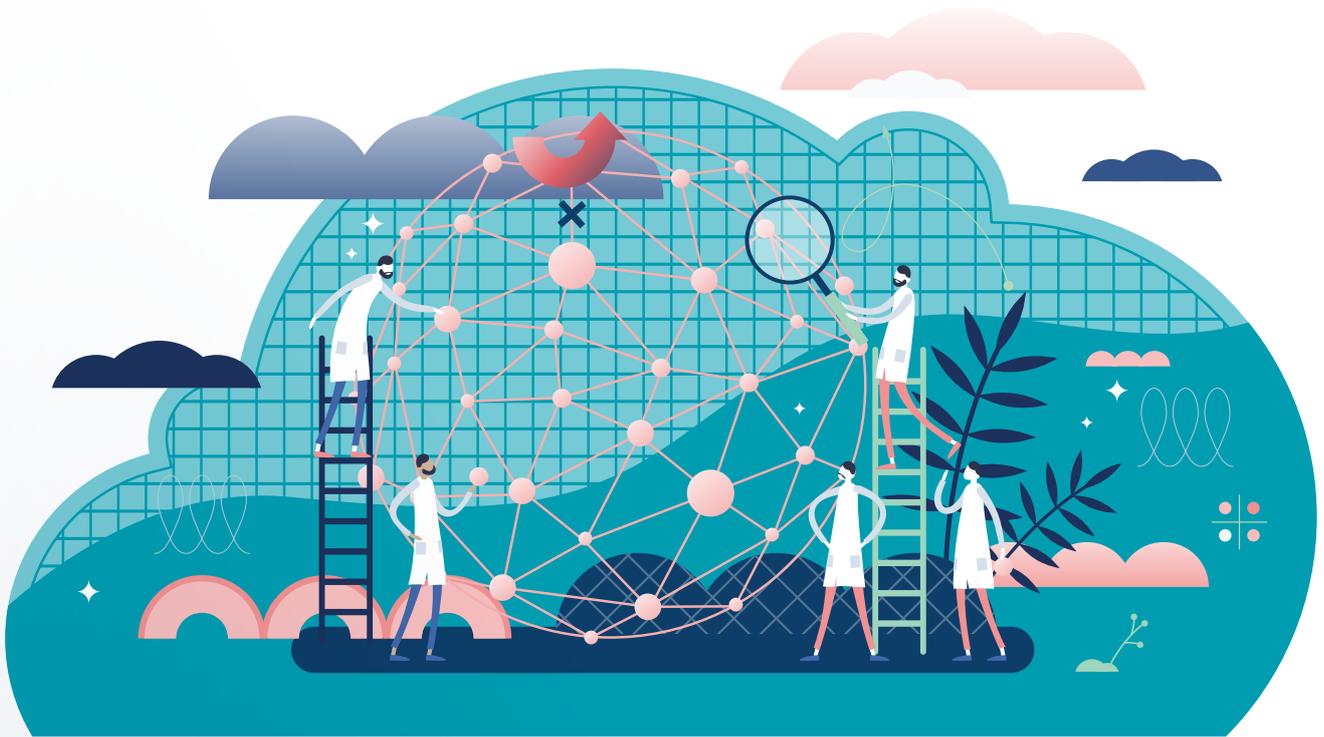
Talent pool

She is also enthusiastic seeing NorwAI's first generation of PhD's now approaching the end of their studies expecting interesting research

deriving from this talented pool of young researchers.

-The work of our PhD students and Postdoctoral fellows in the different work packages evolves. Our publications are accepted at international conferences and journals. In 2023 we have presented 13 conference papers and published 6 journal articles and expect an increase the coming years as our first generation of PhD students approach the final phase of their fellowships, she says, pointing at the long list of papers following this article in the Annual Report.

She enjoys discussing AI methodologies and applications with fellow researchers, but more importantly teaching it to students and bringing it into multidisciplinary projects.



- But even better is the dialogue with students, teaching and get challenged by their curiosity and eager to dig even deeper, be even more bold and see their eagerness to open new perspectives. I really like the excitement to explore the frontiers of our domain and our knowledge, says Kerstin Bach.

Kerstin Bach's own main research interest is methods for applied artificial intelligence.

- While the application domains differ, we are investigating how to use knowledge and experience in intelligent systems. Moreover, how to build systems that support complex, knowledge-intensive decisions using heterogeneous data sources, says Kerstin Bach.

About Kerstin

Kerstin has worked on developing AI methods and applications for more than 15 years. While the majority of her funded projects have been healthcare-related, her keen interest is how AI can support humans: such work can be decision support, explanation of AI models or understanding of large data. She has been the project manager of the selfBACK EU project

(2016-2021) that developed an e-Health application for low back pain patients and been involved in a number of EU-funded projects. With Norway being relatively small, it is important to work internationally and disseminate our work, but also build research networks and joining forces in solving problems through AI.

Kerstin Bach is a well-known and forward-leaning profile in the Department of Computer Science at NTNU, in which NorwAI and the largest academic AI research community in Norway is situated. Her core competence is Artificial Intelligence, in particular Intelligent Decision Support systems, and Machine Learning. In addition to her position at the NorwAI research center, she is the deputy head of the Data and Artificial Intelligence group and part of the Norwegian Open AI Lab.

Kerstin was awarded her PhD in Computer Science from the University of Hildesheim, Germany. Her main research interests are data-driven decision support systems as well as knowledge-intensive case-based reasoning. She is the chair of the German Society for Computer Science's Special Interest Group on Knowledge



Management and a board member of the Norwegian AI society.

NorwAI and NAIL

- You have positions in both NorwAI and the Norwegian Open AI Lab (NAIL), both situated at the same department at NTNU. People may wonder what the difference between the two is.

- NAIL is a hub for AI researchers to meet, facilitate research activities, participate in public debates, and collaboratively develop new projects. At NAIL, we as researchers work with students and partners from Norway. At NorwAI, financed by the Research Council and partners as a Research Center for AI innovation (an SFI), we work in close collaboration with industry partners on practical use cases derived from their needs and innovation programs for the defined period until 2028. And of course, we also educate many master students and PhDs, 500 and 20-40 respectively, in the 8-year lifespan of our center, says Kerstin Bach.

- Some say NorwAI is a closed community, not giving insights on what they are doing?

- Absolutely not. We publish research and results regularly, and our papers are accepted internationally. Our research fellows even win prizes. An open and across domain research is one of the trademarks of our work. We continuously give both academic and popular

keynote lectures. One of our major events in 2023 was the Arendalsuka where we publicly presented a preview of the Norwegian language model, NorGPT, which has the potential to become a national infrastructure in its field, says Kerstin Bach.

- Language models really caught the public eye in 2023. What is your assessment of its importance?

- Our Large Language Model effort is an extremely timely project. And with the EU AI Act to be implemented in Norway, our work on trustworthiness and AI's position in society are both interesting and important.

Furthermore, she says, the fast development of LLMs is an example where NorwAI need to work nationally and internationally:

- With our partners at NorwAI we have created a task force who drives the LLM development for the Norwegian languages. In parallel there are similar initiatives all over Europe and we had the opportunity to join an excellent consortium of European partners to create a platform for developing factful and trustworthy language models together. This is the TrustLLM EU project that commenced last November and one example of expanding the research beyond NorwAI, Kerstin Bach says.

Positions and projects

Kerstin was awarded her PhD in Computer Science from the University of Hildesheim, Germany. Her main research interests are data-driven decision support systems as well as knowledge-intensive case-based reasoning. She is the chair of the German Society for Computer Science's Special Interest Group on Knowledge Management and a board member of the Norwegian AI society.

PROJECTS:

[SmaRTWork](#) - a digital system for personalised return to work recommendations for sick-listed with musculoskeletal disorders (NFR, 2021-2027)

[LABDA](#) - A doctoral network for for Advanced Behavioural Data Analysis (EU Horizon Europe, 2023-2027)

[SupportPrim](#) - Using AI for treatment planning of musculoskeletal pain (2020-2024)

[Exaigon](#) - Explainable AI systems for gradual industry adoption (NFR, 2020-2024)

[AI4EU](#) - An AI-on-demand platform for Europe (EU H2020, 2019-2021)

[selfBACK](#) - A decision support system for self-management of low back pain (project manager, EU H2020, 2016-2021)

[Back-UP](#) - Personalised Prognostic Models to Improve Well-being and Return to Work After Neck and Low Back Pain (EU H2020, 2018-2021)

Research strategy

Our approach to research at NorwAI is fundamentally multi-disciplinary, consisting of both technical-oriented and socio-economic research. Current research on AI is diverse and reflect many disciplines or perspectives covering theoretical contributions from a wide range of disciplines together with more applied work from system designers and application developers working with data from many different domains.

Research in NorwAI focus on generic research areas within AI that can support the innovation activities in the center. The research areas will be evaluated every year in the center's lifetime, making sure new areas can be added based on

research needs from future innovation areas or in case of integrating new partners. Existing areas might also be merged or concluded and come to an end.

Our research is partner-driven and rooted in a strategic and systematic approach towards innovation to create lasting value and impact.

Publications in 2023



JOURNAL PAPERS

NAME	TITLE	TYPE/PLACE	DATE
Bernard, Nolwenn; Balog, Krisztian	A Systematic Review of Fairness, Accountability, Transparency and Ethics in Information Retrieval	ACM Computing Surveys	2023-12-15
Kille, Benjamin Uwe; Lommatzsch, Andreas; Özgöbek, Özlem; Liu, Peng; Eide, Simen; Zhang, Lemei	The Eleventh International Workshop on News Recommendation and Analytics (INRA'23)	RecSys '23: Proceedings of the 17th ACM Conference on Recommender Systems.	2023-09-24
Rønningstad, Egil	UIO at SemEval-2023 Task 12: Multilingual fine-tuning for sentiment classification in low-resource languages	Association for Computational Linguistics (ACL). Annual Meeting Conference Proceedings	2023-07-01
Zhang, Lemei; Liu, Peng; Gulla, Jon Atle	Recommending on graphs: a comprehensive review from a data perspective	User modeling and user-adapted interaction	2023-03-13

PUBLISHED CONFERENCE PAPERS

NAME	TITLE	TYPE/PLACE	DATE
Afzali, Jafar; Drzewiecki, Aleksander Mark; Balog, Krisztian; Zhang, Shuo	UserSimCRS: A User Simulation Toolkit for Evaluating Conversational Recommender Systems	WSDM '23: Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining	2023-02-27
Barnes, Jeremy Claude; Touileb, Samia; Mæhlum, Petter; Lison, Pierre	Identifying Token-Level Dialectal Features in Social Media	Proceedings of the 24th Nordic Conference on Computational Linguistics (NoDaLiDa)	2023-05-22
Bayrak, Betül; Bach, Kerstin	When to Explain? Model Agnostic Explanation Using a Case-based Approach and Counterfactuals	Proceedings of the 34th Norwegian ICT conference for research and education – NIKT 2022	2023-03-09
Bayrak, Betül; Marin-Veites Paola; Bach, Kerstin	Explaining your Neighbourhood: A CBR Approach for Explaining Black-Box Models.	CEUR workshop proceedings 2023 (ICCBR XAI'22: Explainable AI Challenge at ICCBR-2022, September, 2022, Nancy, France)	2023-05-11
Baumgartner, David; Langseth, Helge; Ramampiaro, Heri; Engø-Monsen, Kenth	mTADS: Multivariate Time Series Anomaly Detection Benchmark Suites	Proceedings 2023 IEEE International Conference on Big Data Dec 15 - Dec 18, 2023 • Sorrento, Italy	2023-12-18
Bernard, Nolwenn; Balog, Krisztian	MG-ShopDial: A Multi-Goal Conversational Dataset for e-Commerce	Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval	2023-03-19
Charpentier, Lucas Georges Gabriel ; Wold, Sondre; Samuel; David; Rønningstad, Egil	BRENT: Bidirectional Retrieval Enhanced Norwegian Transformer	NoDaLiDa 2023, main conference	2023-05-24
Doan, Tu My; Kille, Benjamin Uwe; Gulla, Jon Atle	SP-BERT: A Language Model for Political Text in Scandinavian Languages	NLDB 2023	2023-06-14
Lajewska, Weronika; Balog, Krisztian	From Baseline to Top Performer: A Reproducibility Study of Approaches at the TREC 2021 Conversational Assistance Track	Advances in Information Retrieval 45th European Conference on Information Retrieval, ECIR 2023, Dublin, Ireland, April 2–6, 2023, Proceedings, Part III	2023-03-16
Lajewska, Weronika; Balog, Krisztian	Towards Filling the Gap in Conversational Search: From Passage Retrieval to Conversational Response Generation	CIKM '23: Proceedings of the 32nd ACM International Conference on Information and Knowledge Management	2023-10-21
Olstad, Annika Willoch; Papadopoulou, Anthi; Lison, Pierre.	Generation of Replacement Options in Text Sanitization	Proceedings of the 24th Nordic Conference on Computational Linguistics (NoDaLiDa)	2023-05-22
Vassøy, Bjørnar; Langseth, Helge; Kille, Benjamin Uwe	Providing Previously Unseen Users Fair Recommendations Using Variational Autoencoders	ACM RecSys 2023	2023-09-23
Xing, Yujie; Gulla, Jon Atle	Improving Context-Awareness on Multi-Turn Dialogue Modeling with Extractive Summarization Techniques	NLDB2023	2023-06-14
Xing, Yujie; Liu, Peng	Prompt and instruction-based tuning for response generation in conversational question answering	NLDB2023	2023-06-22



NorwAI objectives

The objective of the NorwAI research center is to provide a strong and robust arena for industry, research and academic institutions to collaborate on the development of AI ideas and techniques, share results, and iteratively explore how technology can transform existing businesses and enable entirely new business avenues.

The primary objective – or mission – of NorwAI is to:

Accelerate the innovation of sustainable and trustworthy artificial intelligence solutions across Norwegian industries.

This primary objective is further broken down into objectives and secondary objectives, as shown on the next page.

Artificial intelligence constitutes a paradigm shift in computer science, enabling substantially shorter development cycles, extremely powerful solutions, and immediate transfer of technologies from one domain to another. The innovation cycles become very dynamic, representing both challenges and opportunities at the same time. New systems may completely transform existing practices, render old value chains worthless, or open for brand new business opportunities. NorwAI supports innovation as a platform for continuous

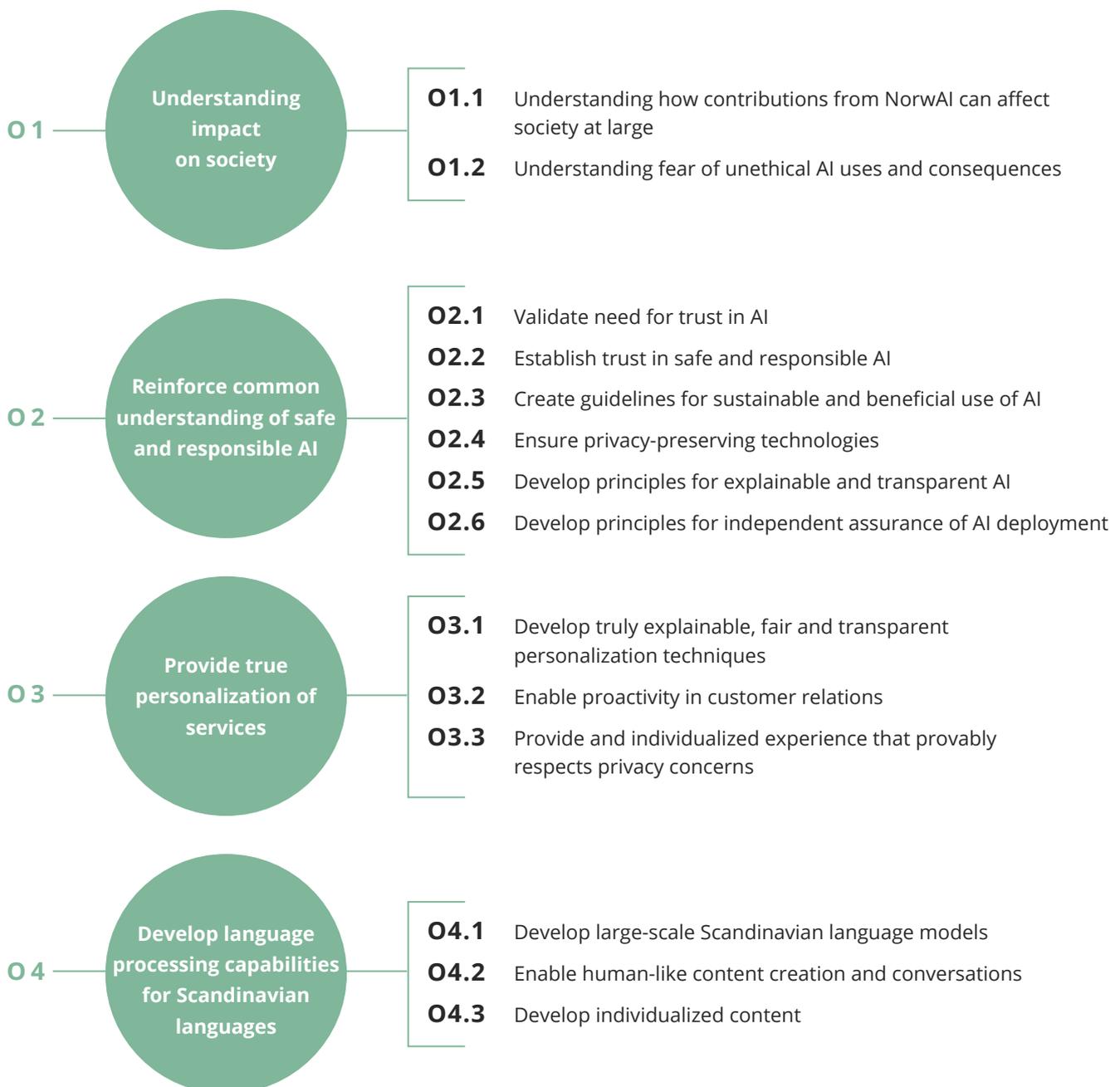
interaction between industry and academia, and will act as an ecosystem for creating alliances, joint venturing and building synergies among all partners. NorwAI will thereby enhance the ability of the business sector to innovate and create value through a greater focus on long-term research.

NorwAI acknowledges that the deep impact of AI makes it necessary to ensure efficient sharing of knowledge and enable businesses to adapt their innovation processes to this new situation and provides the skills for business transformation. At the same time, it is critical that the technology is applied with care and with respect for the needs of individuals and societies. The objectives of true personalization of services by providing data and platforms for AI innovations address the AI research needed to implement, deploy and evaluate companies' use case innovations. Finally, the overall quality and reputation of the research center will affect the center's ability to help companies develop and deploy research-based innovations and establish NorwAI as an international leading AI research and innovation center.

NorwAI objectives

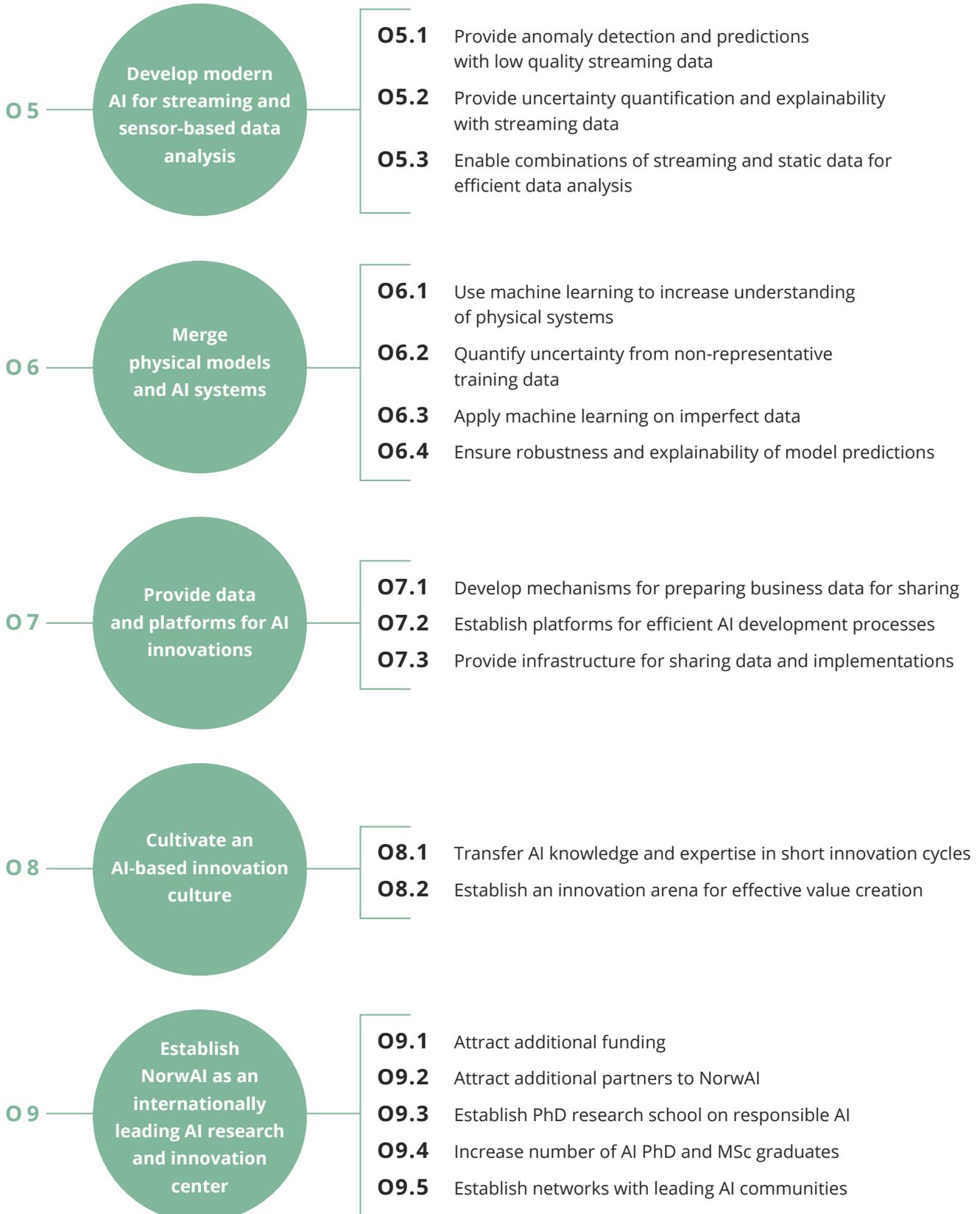
OBJECTIVES

SECONDARY OBJECTIVES



OBJECTIVES

SECONDARY OBJECTIVES



Building Trust in AI-Enabled Systems through Assurance



Andreas Hafver
Team leader - emerging technologies
DNV

AI technologies can enrich our lives and be harnessed to address societal challenges and advance industries towards a sustainable future. However, AI can also introduce new risks, vulnerabilities, and ethical concerns. For AI technologies to gain widespread acceptance and adoption, they must be trustworthy and managed in a responsible manner that protects the interests of all stakeholders.

One way to build trust in AI systems is through assurance, which involves establishing and verifying claims about the system's potential impacts. DNV has issued a new recommended practice, DNV-RP-0671, to guide actors across AI value chains to assure that AI-enabled systems are trustworthy and managed responsibly.

Like a lubricant

Trust is like a lubricant that reduces friction in society. We interact seamlessly with individuals, institutions, and technologies if we trust that they will safeguard our interests and not cause us harm. On the contrary, too much friction, in the form of mistrust, impedes business and hinders societal progress. Similarly, too little friction, in the form of misplaced or excessive trust, can also cause us to stumble.

The challenge lies in determining when trust is warranted, particularly when dealing with unfamiliar entities and technologies that we do not fully comprehend. This is when we need assurance.

What is assurance?

In essence, assurance of a system comprises two components: defining a set of requirements to a system and demonstrating that those requirements are met. Coming up with requirements entails describing and modelling the system and engaging with its stakeholders to identify risks and set risk acceptance criteria.



AI is deployed in more than

1/3

of public companies
worldwide



*...the unique challenges
posed by AI systems will
continue to grow.*



DNV CEO Remi Eriksen gave the main key note at NorwAi Innovate Conference 2023 focusing on the rapid growth of AI in industry and the need for trustworthy solutions. Remi Eriksen is also chairperson of NTNU.

In the context of AI-enabled systems, some requirements will pertain to the functions, performance, and other characteristics of the AI components within the system, imposed by the wider system and its stakeholders. In addition, the capabilities and limitations of AI components within a system may impose requirements on other parts of the system.

Demonstrating that requirements are met involves collecting evidence, which may include refined modelling and risk assessment. Through an iterative process, an argument can be built to substantiate claims about the AI-enabled system. One part of providing assurance is to verify the argument by scrutinizing the evidence and rationales. Another part of assurance is

to validate that the claims are relevant and adequately cover all legitimate interests of stakeholders. The result of this is warranted trust.

Depending on the application domain and applicable regulations, assurance may be provided by an independent third party, like DNV, or conducted by stakeholders themselves. Often, assurance is done to demonstrate compliance with regulations and standards, which is a ticket to trade in some industries. However, assurance can also be a way to foster trust with vendors, customers, and the wider society in wider ecosystems of trust.



The challenges

Assuring AI-enabled systems presents some unique challenges.

- » Firstly, AI-enabled systems tend to be complex, with emergent behaviours resulting from the interactions between AI components, other system components, humans, and the environment.
- » Secondly, AI-enabled systems tend to be dynamic, with some AI models capable of learning and adapting online as they accumulate data.
- » Thirdly, stakeholders' concerns relating to AI-enabled systems can vary widely depending on the application context, including impacts on safety, cybersecurity, and the environment; implications for privacy and intellectual property; as well as fairness, autonomy, equity, and other ethical considerations.

On top of this, different actors may be responsible for different parts of an AI system and may have diverging and sometimes conflicting interests.

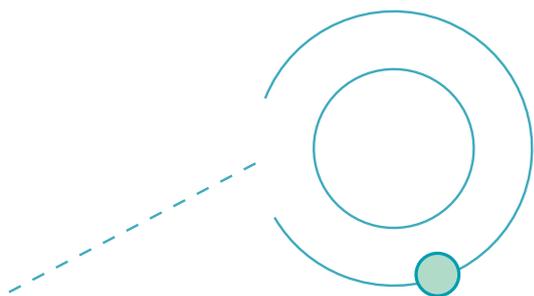
A new assurance approach

DNV-RP-0671 proposes a new assurance method to effectively address the mentioned challenges in assuring AI-enabled systems.

- » Firstly, it handles the complexity and emergent properties of AI systems by adopting a holistic systems approach that examines the interactions between AI components, other system components, humans, and the environment, rather than isolating individual parts.
- » Secondly, it accounts for the dynamic nature of AI-enabled systems by making the assurance process more agile and continuous.
- » Thirdly, it involves a diverse set of stakeholders in the assurance process to identify and address a wide range of concerns.
- » Lastly, it facilitates a modular assurance process where different stakeholders can take responsibility for their parts of the systems and still collaborate to build common trust.

The importance of assuring AI-enabled systems

Assurance plays a vital role in fostering warranted trust in AI-enabled systems. It not only helps in complying with regulations and standards but also promotes AI adoption and scaling by building trust among stakeholders. With the rapid advancement of AI technologies, the need for assurance methods that can effectively address the unique challenges posed by AI systems will continue to grow.



Expanding the AI talent pool

Expectations are high for the exchange of AI students between Norway and Canada. Associate Professor Özlem Özgöbek at NTNU, project coordinator, supports the statement from her colleague from across the Atlantic:

- Expanding the AI talent pool will yield positive impacts that help improve the quality-of-life for people around the world. This research exchange will prove to be invaluable for years to come, says Harold Godwin managing director of Waterloo. AI at the well respected University of Waterloo at Toronto, Canada.

The program kicked off in the winter of 2023 with Canadians arriving at NorwAI partners SINTEF, Cognite and NTNU. Vice versa Norwegians will take the reverse trip in 2024. The International Work-Integrated-Learning in

Artificial Intelligence (IWIL AI) project started by gathering partners from Canada and Norway digitally.

Strengthen relations

The IWIL AI project is led by The Department of Computer Science at NTNU, and is in close connection to SFI NorwAI and the Waterloo Artificial Intelligence Institute in Canada. External partners of the project are Cognite AS, SINTEF Digital and Aneo from Norway, and Manulife from Canada. The project is mainly run by two NorwAI staff: Associate professor Özlem Özgöbek as the project coordinator and Karolina Storesund as the administrative responsible.

The project aims to enhance the quality and relevancy of AI education through student exchange for bachelor, masters and PhD students in one of the partner organizations. In addition, it aims to strengthen the partnerships for research and innovation through international cooperation between the partners of this project.

*Stanley Maio at the poster session, NorwAI Innovate 2023.
(Photo: Alexander Stasik, SINTEF)*





First exchange students from the University of Waterloo at NorwAI with the project project management team. From left to right: Karolina Storesund, Elliot Song, Tony Li, Özlem Özgöbek.

Deliverables and contributions

During 2023 NTNU has received two students, NorwAI partners Cognite one and SINTEF have received one. All the students have successfully completed their four month stays with deliverables and contributions.

Stanley Miao, who stayed with SINTEF in Oslo, presented a poster about the work he has been involved in at the annual NorwAI Innovate Conference.

Elliot Song and Tony Li at NTNU also attended the conference and were responsible for a demo of the social robot Kaia.

In November 2023, the process for sending students from Norway started. The deadline for the second round of applications for a research stay in Waterloo will be in April 2024. The IWIL AI project is funded by the Norwegian Directorate for Higher Education and Skills (HK-dir) and will last until December 2026.

For more information, please check the web page: <https://www.ntnu.edu/iwilai>



Two new Program Managers at NorwAI

Associate Professor Özlem Özgöbek was appointed one of the two new Program Manager at NorwAI in 2023.

As a program manager, she is responsible of communication with the work package leaders and partners in the work packages Trustworthy AI and AI in Society. The packages respectively focus on reinforcing a common understanding of safe and responsible AI (TRUST) and examining societal aspects of artificial intelligence technologies as they are developed, debated, and implemented (SOC). Özlem will contribute to establishing better communication routines and new connections within NorwAI.

She is also coordinating the IWIL AI exchange program, connecting NorwAI to the University of Waterloo in Toronto, Canada, as well as lecturing and supervising master's and PhD students.

Özlem Özgöbek has a PhD degree from Ege University, Türkiye and previously worked as a post-doctoral researcher at NTNU. She is holding the associate professorship position since October 2019 at the Department of

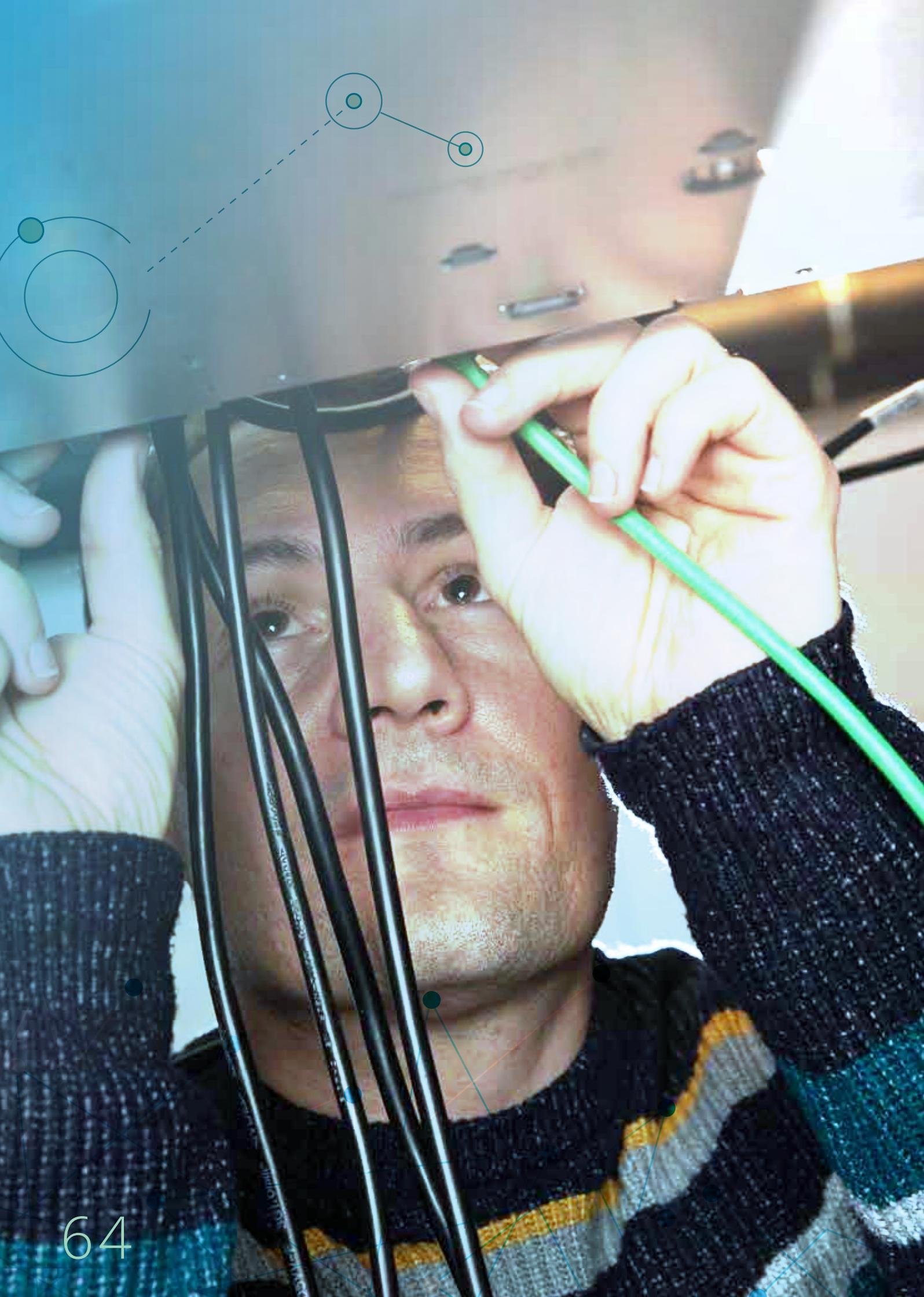
Computer Science, NTNU. Her research interests include recommender systems, disinformation detection systems and ethical aspects of artificial intelligence such as privacy and bias. She is involved in the organization committees of various events, scientific conferences and workshops, including events for gender balance in computer science.

- Why do you like working for NorwAI?

- *NorwAI brings many very important actors in artificial intelligence together for research-based innovation. It is exciting to be a part of this community and contribute to shaping the future of AI.*







Another new Program Manager at NorwAI

Associate Professor Benjamin Kille was appointed Program Manager at NorwAI in 2023.

He oversees the work packages Language and Personalization (LAP) and Data Platforms and Streaming Data (DATA). His task is to foster better communication between partners, creating connections and synergies between different research directions at NorwAI.

Benjamin holds a PhD from Technische Universität Berlin, Germany. He had worked as post-doctoral researcher at NTNU. He has been involved in activities around language models, personalization, and generative AI. Benjamin transitioned into the role as Associate Professor at NTNU in November 2023.

His current research interests include language technology, recommender systems, and generative AI. He co-organizes workshops and data-driven challenges, contributes to peer

review for various conferences and journals, and teaches the course on Recommender Systems at NTNU.

- Why do you like working for NorwAI?

- *Artificial Intelligence operates in an environment populated by academic, industrial, and public actors. NorwAI catalyzes innovation with the help of partners from different sectors. NTNU generally and NorwAI specifically have the resources to determine the use of AI in Norway.*



The EU-Project TrustLLM as an answer to Generative AI hallucinations

The last two years have seen the rise of Generative AI. Many models provide useful functions but tend to make up facts and respond overly confidently. How to mitigate that risk?

In November 2023, a consortium with partners from Norway, Germany, Sweden, Iceland, Denmark, and the Netherlands kicked off the Horizon Europe-funded project TrustLLM to develop open, trustworthy, and sustainable Large Language Models (LLMs).

Germanic language family

Initially, these models will focus on Germanic languages. NTNU, through SFI NorwAI, represents Norway in the consortium, and contributes with expertise on the development of LLMs and resources for Norwegian. Together with the partners, we hope to create a vivid eco-system fostering many innovations and more insights into the responsible use of generative AI.

The project faces several challenges. Building large-scale LLMs requires excessive amounts of texts and thousands of GPU-hours. Luckily, the consortium has experience in managing large data sets and applying computational resources to Europe's largest computer clusters.

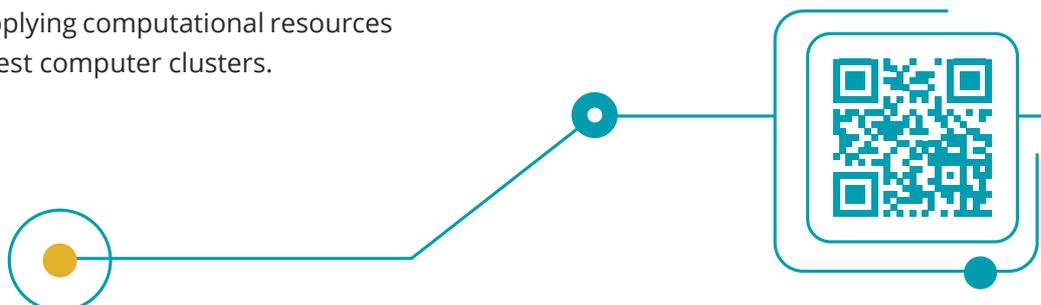
Evaluate

Evaluating generative AI represents a scientific challenge. Users perceive texts subjectively and annotators disagree. NorwAI is leading the task to define suited evaluation metrics to capture subtle concepts such as reliability, accuracy, fluency, factual consistency, or semantic coherence.

In the case of Norwegian, the models have to be trained and evaluated for both written forms: Bokmål and Nynorsk.

Finally, the consortium will apply the developed models to a number of use cases to show their utility and conformity with European values. NorwAI's member DNB and Schibsted have volunteered to examine the models within their organizations.

More information about TrustLLM can be found on: <https://trustllm.eu/>





PARTNERS:



The forums for shared insights

The NorwAI Forums have different focus. While the Oslo Forum headed by SINTEF look into innovation work, the Trondheim version headed by NTNU take a closer look into research.

Both Forums are cross company and cross institution gatherings with representatives from the NorwAI community attending and presenting use cases and insights. In 2023, two

Forums were organized in the two cities. In addition, SINTEF was also responsible for two workshops for the media/finance companies and the industrial cluster of NorwAI.

NORWAI FORUM: TRONDHEIM, MARCH 13TH

At the NorwAI HQ at NTNU, the Forum drew a number of participants from different partners.

Agenda

- ◆ On the role of Human Annotations in the era of LLM
- ◆ A Systematic Literature Review of User Trust in AI-Enabled Systems. An HCI-perspective.
- ◆ Topological Data Analysis of Industrial time-series
- ◆ Anomaly detection on graphs
- ◆ AI in society
- ◆ Creating Shared understanding about AI innovation for NorwAI



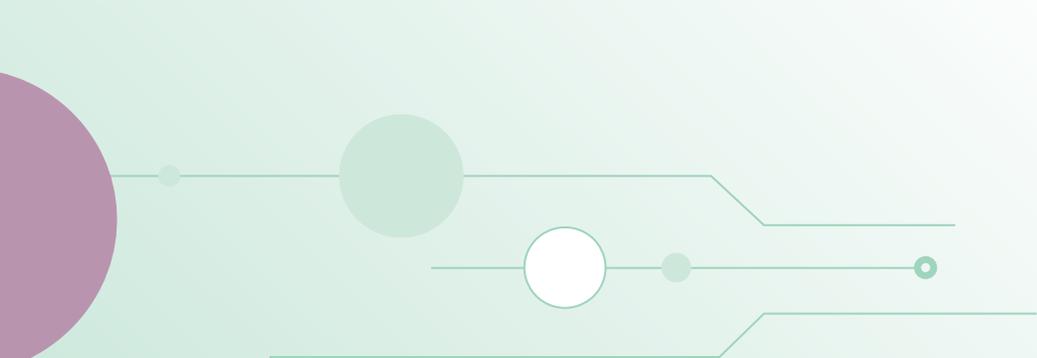
NORWAI FORUM: OSLO, AUGUST 30TH

At the DNB HQ in Oslo, 35 representatives gathered for lectures, insights and discussions.

Agenda

- ◆ Large language models development and application in industry
- ◆ Trustworthy AI and assurance
- ◆ Trustworthy AI and AI in Society
- ◆ Industrial challenges: Data sharing and anomaly/outlier detection
- ◆ Federated learning
- ◆ Synthetic data





Innovation workshops

NO1: AT DNV, IN MAY FOR THE ENERGY AND MANUFACTURING CLUSTER

Partners represented:

Kongsberg, DNV, SINTEF, NTNU, UiO, Statnett, Telenor, Aneo and Cognite

Main topic:

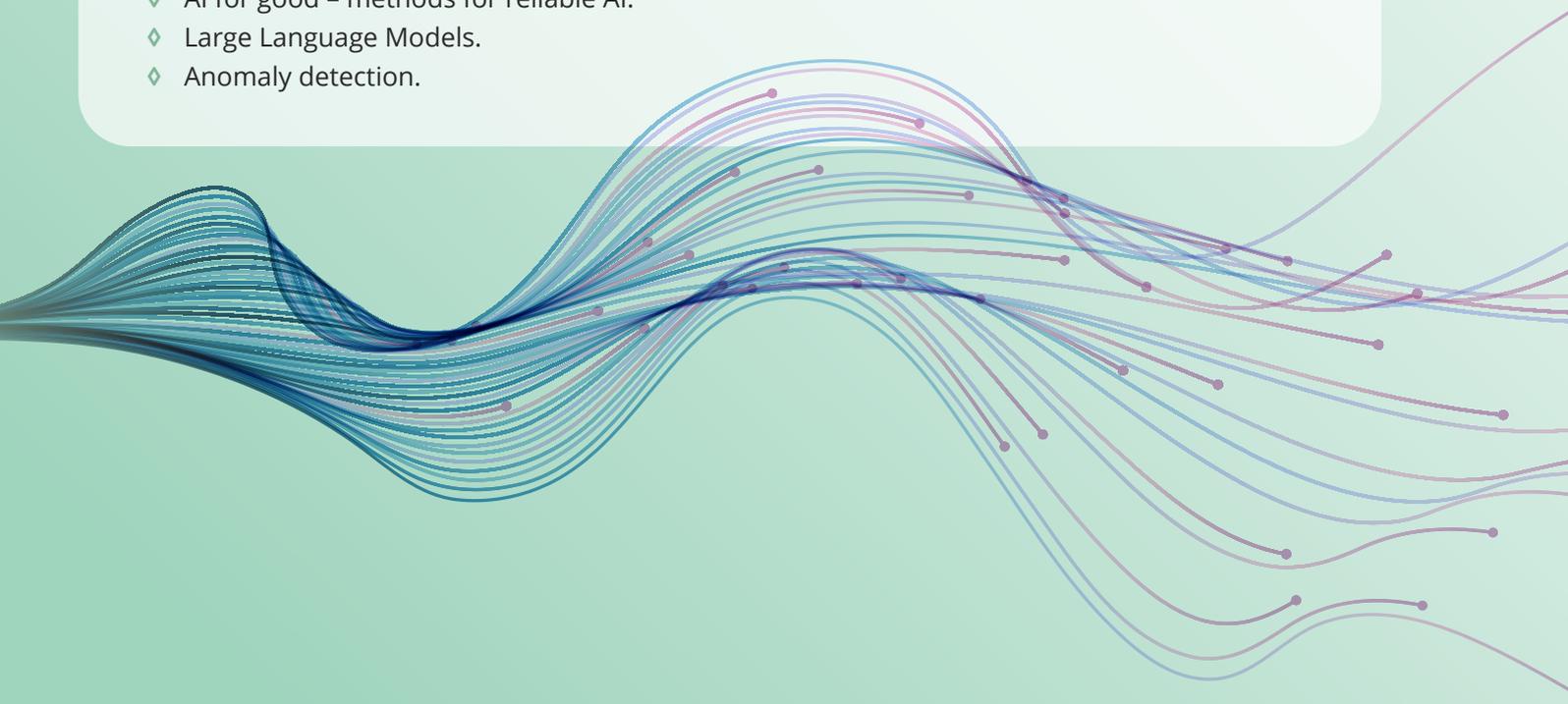
- ◆ Needs for competence and challenges for trusted AI.
- ◆ Needs for innovations and increased productivity by AI

NO2: AT SINTEF DIGITAL, IN JUNE FOR THE MEDIA AND FINANCE CLUSTER

Partners represented:

DNB, SpareBank1 SMN, Schibsted, DNV, UiO, SINTEF, NTNU

Main topics:

- ◆ AI for good – methods for reliable AI.
 - ◆ Large Language Models.
 - ◆ Anomaly detection.
- 



NorwAI Innovate '23

A unique AI showroom

Our ambition is to host the largest and most interesting AI tech conference in Norway for the Norwegian AI community. An extraordinary line-up of keynotes and a mixed academia and business audience provides a melting pot for cross professional collaboration, cross discipline exchange and cross generation meetings.

HOSTS



Professor Jon Atle Gulla
NorwAI Center Director



Rolf Dyrnes Svendsen
NorwAI Communication Manager



Karolina Storesund
NorwAI Administrative Coordinator



Political editor Siv Sandvik, Adresseavisen and cst. state secretary Tore O. Sandvik, Ministry of Trade and Industry at a panel discussion.

For the third year in a row, close to 200 people heeded the call to attend NorwAI Innovate. The mixed audience included company professionals, researchers and students.

NorwAI Innovate is a unique showroom for what to expect from the research center itself. The conference also examines specific AI verticals and invites national and international experts to speak about use cases and innovations in the ongoing AI transition in society. The 2023 version focused on trust, language models and

of course addressed the elephant in the room; i.e. AI regulations.

In the initial keynote, Group president and CEO of DNV Remi Eriksen, also chair person of NTNU, underlined the necessity of trust when AI technologies are introduced in industry.

Sidekicks like posters and prototype demos fueled NorwAI Innovate with the next generation AI talents. Above all, the lunch-to-lunch conference is a meet & greet event and a chance to get to know leading people in both the private and public sector concerned and engaged in AI developments.





Professor Fredrik Heintz
Linköping University
Coordinator TrustLLM project

KEY NOTE SPEAKERS



Alexander Stasik
SINTEF Digital



Jon Espen Ingvaldsen
NTNU/Kantega



Simone Casolo
Cognite



Christian Bendiksen
Bull & Co Advokatfirma



Keith Downing
NTNU



Tore O. Sandvik
Ministry of Trade, Industry and
Fisheries



Claes Lyth Walsø
Norwegian Police



Kristin Y. Pettersen
NTNU



Tore Tennøe
The Norwegian Board of
Technology



Daniel Pleus
Schibsted



Notto J. W. Thelle
Oslo Met



Viviana Gropengiesser
ETH AI Center



Grunde Almeland
Stortinget



Per Egil Kummervold
National Library of Norway



Yngvar Ugland
DNB



Johan Östman
AI Sweden



Remi Eriksen
DNV





NorwAI Innovate 2024

Trondheim
Sept 24th - 25th
Scandic Nidelven Hotel

SAVE THE DATE

Best poster award at NorwAI Innovate 2023

Xavier F.C. Sánchez-Díaz received this year's Best Poster Award worth NOK 10 000 at the NorwAI Innovate conference.

The award recognizes his work on an open-source library that provides reusable and modular components that can be used to customize problem solvers. The poster entitled "EvoLP.jl: Modular optimization in Julia using Evolutionary Computation", is part of his Ph.D. project under the supervision of professor Ole Jakob Mengshoel at the Norwegian Open AI Lab.

Research director, professor Kerstin Bach of NorwAI, handed him the award for an excellent poster presentation honored by the public vote and the award jury.



Xavier Sánchez-Díaz accepts the NorwAI Innovate 2023 Best Poster Award from Professor Kerstin Bach.

CONTRIBUTIONS

AUTHORS	TITLE
Agarwal, R.	Handling Dimension Varying Inputs in Online Learning. Poster at NorwAI Innovate Conference, Trondheim.
Baumgartner, D.	mTADS: Multivariate Time Series Anomaly Detection Benchmark Suites. Poster at NorwAI Innovate Conference, Trondheim.
Doan, T. M., Kille, B, Gulla, J. A.	SP-BERT: A Language Model for Political Texts in Scandinavian Languages. Poster at NorwAI Innovate Conference, Trondheim.
Eidnes, S., Stasik, A. Lye, K., Sterud, C., Bøhn, E. and Riemer-Sørensen, S.	Pseudo-Hamiltonian neural networks for learning ODEs and PDEs. Poster at NorwAI Innovate Conference, Trondheim.
Firouzjaei, H. A.	A deep learning-based approach for identifying unresolved questions on Stack Exchange Q&A communities through graph-based communication modelling. Poster at NorwAI Innovate Conference, Trondheim.
Fuchs, F. G., Stasik, A., Miao, S.	QuantumReservoirPy: A Software Package for Time Series Prediction. Poster at NorwAI Innovate Conference, Trondheim.
Jayawardene, I. Elvesæter, B.	Identifying Defects in Fishnets in Fish Farms. Poster at NorwAI Innovate Conference, Trondheim.
Khalitov, R. Yu, T., Cheng, L and Yang, Z.	ChordMixer: A Scalable Neural Attention Model for Sequences with Different Lengths. Poster at NorwAI Innovate Conference, Trondheim.
Knaus, T.	DECODING THE SOCIAL SIGNAL IN VOICE DATA – The emergence of computational methods to read human sociality. Poster at NorwAI Innovate Conference, Trondheim.
Løkkeborg, T.	Deep Reinforcement Learning for International Diplomacy: Learning to Play Map Variants. Poster at NorwAI Innovate Conference, Trondheim.
Michalowska, K., Goswami, S., Karniadakis, G. E., Riemer-Sørensen, S.	Multi-resolution learning with operator- and recurrent neural networks. Poster at NorwAI Innovate Conference, Trondheim.
Rønningstad, E.	NorwAI Innovate Poster Abstract: New Entity-centric Insights for Norwegian Sentiment Analysis. Poster at NorwAI Innovate Conference, Trondheim.
Sánchez-Díaz, X. F. C., Mengshoel, O. J.	EvoLP.jl: Modular Optimisation in Julia using Evolutionary Computation. Poster at NorwAI Innovate Conference, Trondheim.
Shamsaliei, S., Gundersen, O. E., Kjærnli, H., Langseth, H.	On Reliable Reporting of Conclusions in Model Comparison Studies of Deep Neural Networks.
Singh, A.	Deep Reinforcement Learning for Spatio-Temporal Wildlife Management. Poster at NorwAI Innovate Conference, Trondheim.
Song, E., Li, T	Kaia- the Social Robot. Demo at NorwAI Innovate Conference, Trondheim.
Treullier, C., Castagnos, S., Brun, A.	How a Multi-factorial Analysis of Polarization Paves the Way for Innovative Recommendation Strategies. Poster at NorwAI Innovate Conference, Trondheim.
Vassøy, B., Langseth, H. and Kille, B.	Providing Previously Unseen Users Fair Recommendations Using Variational Autoencoders. Poster at NorwAI Innovate Conference, Trondheim.

Recruitment

Phd-students

STARTED IN 2021



Katarzyna Michalowska

Nationality: Poland
Topic: Informed machine learning
Started: January 2021
Main Supervisor: Prof. Morten Hjort-Jensen, University of Oslo
NorwAI Work Package: HYB



Nikolay Nikolov

Nationality: Bulgaria
Topic: Flexible Deployment of Big Data Pipelines on the Cloud/Edge/Fog Continuum
Started: January 2021
Main Supervisor: Dr. Dumitru Roman (Senior Research Scientist at SINTEF Digital and Assoc. Professor at University of Oslo, Norway)
NorwAI Work Package: DATA



Bjørnar Vassøy

Nationality: Norway
Topic: Fairness, Accountability, Transparency and Privacy in Personalization/Recommender systems
Started: August 2021
Main supervisor: Prof. Helge Langseth, NTNU
NorwAI Work Package: LAP



David Baumgartner

Nationality: Austria
Topic: Data analysis with noisy and low-quality data streams
Started: September 2021
Main supervisor: Prof. Heri Ramampiaro, NTNU
NorwAI Work Package: DATA



Egil Rønningstad

Nationality: Norway
Topic: Norwegian Opinion summarization and Entity-level Sentiment Analysis
Started: October 2021
Main supervisor: Prof. Erik Velldal, University of Oslo
NorwAI Work Package: LAP

STARTED IN 2022



Nolwenn Bernard

Nationality: France
Topic: Study of fairness and transparency in conversational recommender systems
Started: February 2022
Main Supervisor: Prof. Krisztian Balog, University of Stavanger
NorwAI Work Package: LAP



Weronika Łajewska

Nationality: Poland
Topic: Personalizing Conversational Informational Access
Started: February 2022
Main Supervisor: Prof. Krisztian Balog, University of Stavanger
NorwAI Work Package: LAP



Tanja Knaus

Nationality: Austria
Topic: The materialization of the human voice into affective data - data practices of AI cloud services using vocal emotion recognition systems
Started: April 2022
Main Supervisor: Prof. Susanne Bauer, University of Oslo
NorwAI Work Package: SOC



Lena Jedamski

Nationality: Germany
 Topic: Trustworthy Artificial Intelligence
 Started: February 2023
 Main supervisor: Prof. Kerstin Bach, NTNU
 NorwAI Work Package: TRUST
 Also part of the PERSEUS Doctoral Programme



Vandana Yadav

Nationality: India
 Topic: Natural Language Processing for personalized content summarization
 Started: November 2023
 Main supervisor: Prof. Jon Atle Gulla, NTNU
 NorwAI Work Package: LAP
 Also part of the PERSEUS Doctoral Programme



Abdul-Kazeem Shamba

Nationality: Nigeria
 Topic: Artificial Intelligence for Anomaly Detection
 Started: September 2023
 Main Supervisor: Ass. Prof. Frank Kraemer, NTNU
 NorwAI Work Package: DATA



Serinha Murgorgo

Nationality: Zimbabwe
 Topic: Embracing Responsible AI: Integration of responsible AI principles in organisational practices
 Started: August 2023
 Main supervisor: Ass. Prof. Nhien Nguyen, NTNU
 NorwAI Work Package: INNOECO



Jessica Annalena Steppe

Nationality: Germany
 Topic: AI-Augmented Creativity in Innovation Teams
 Started: August 2023
 Main supervisor: Ass. Prof. Nhien Nguyen, NTNU
 NorwAI Work Package: INNOECO



Web and social media



INSTAGRAM
[@sfi_norwai](https://www.instagram.com/sfi_norwai)



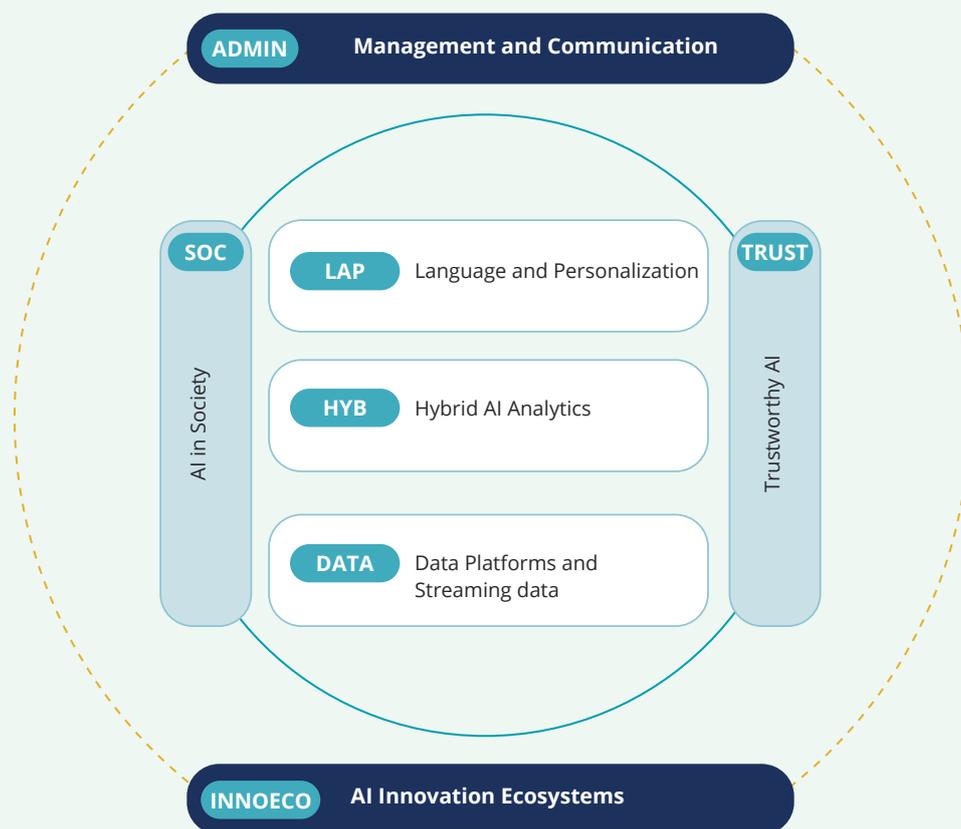
WEBPAGE
norwai.org



LINKEDIN
[www.linkedin.com/
company/norwai](https://www.linkedin.com/company/norwai)

The work packages

NorwAI consists of seven work packages, which have been organized as described in this figure:



ADMIN: MANAGEMENT AND COMMUNICATION

This work package concerns the day-to-day operations of the center as well as communication.

Work package leader: Jon Atle Gulla, Professor, NTNU



SOC: AI IN SOCIETY

This work package examines societal aspects of artificial intelligence technologies as they are developed, debated, and implemented. We focus on selected empirical cases, including

- Values in language models
- Societal aspects of voice and speech emotion recognition technologies
- Standardization of AI technologies and ethical, legal, and social aspects

Work package leader: Susanne Bauer, Professor, University of Oslo



TRUST: TRUSTWORTHY AI

The purpose of work package TRUST is to reinforce a common understanding of safe and responsible AI, specifically:

- Establish trust in safe and responsible AI
- Ensure privacy-preserving in AI technologies
- Create guidelines for sustainable and beneficial use of AI
- Develop principles for explainable and transparent AI
- Develop principles for independent assurance of AI deployment

Trust in AI is a necessary condition for the scalability and societal acceptance of these technologies. Without trust, innovation can be stalled. This research investigates, from an interdisciplinary perspective, the multiple dimensions of trust raised by the deployment of AI and builds tools, methods, and a framework for assuring the safe and responsible deployment of AI in industry and society. This work package aims to answer the question: How can such tools address the safety and needs of individuals, organizations and society at large, addressing both non-technical and technical issues? The research will address issues related to safety, explainability, transparency, bias, privacy and robustness, as well as human-machine interactions and co-behavior all in the context of industry regulations and societal expectations.

Work package leader: Andreas Hafver, Team Leader - Emerging Technologies, Group Research & Development



LAP: LANGUAGE AND PERSONALIZATION

The purpose for this work package is to develop personalization techniques and Scandinavian language processing capabilities to provide personalized content generation and:

- Develop truly explainable, fair and transparent personalization techniques
- Enable proactivity in customer relations
- Provide an individualized experience that provably respects privacy concerns
- Develop individualized content
- Develop large-scale Scandinavian language models
- Enable human-like content creation and conversations

Personalization and contextualization have been successfully employed in diverse applications over the past decade, and currently see an extended usage, for instance in proactive interaction with customers and individualization of news stories. LAP will contribute to developing such systems while ensuring that the system usage will be ethical and respecting users' requirements for privacy, fairness and accountability.

Building Scandinavian language models requires the compilation of large-scale reusable language resources, including general-purpose corpora from public sources (e.g., news and social media) as well as industry- and domain-specific text collections. We will address the scarcity of the latter by pre-training on the former and developing transfer learning methods. These large-scale language models will then be utilized in real-life scenarios by formulating a number of specific summarization, explanation, and conversational tasks based on our partners' use-cases. LAP will develop appropriate evaluation methodology with user-oriented evaluation measures and objectives. It will thus contribute to providing measurable quantification of the amount of domain-specific training material needed in order to provide a language service that is of sufficiently high quality.

Work package leader: Krisztian Balog, Professor, University of Stavanger



HYB: HYBRID AI ANALYTICS

The purpose of work package HYB is to:

- Develop robust, stable and explainable data-driven models for physical systems
- Constrain models to enforce meaningful predictions
- Transfer data-driven models from simulations to reality
- Characterize and quantify uncertainty of data-driven models

This work package will develop methods to predict and reduce the uncertainty of data-driven models. The models will be constrained by existing knowledge, allowing to interpret the model (explainable AI) and reducing the amount of required training data. Applying these methods on real world applications will allow the industry partners to better predict the behavior of their facilities and improve their simulations, e.g. for condition monitoring, predictive maintenance, optimal utilization.

Work package leader: Signe Riemer-Sørensen, Research Manager, SINTEF



WP DATA: DATA PLATFORMS AND STREAMING DATA

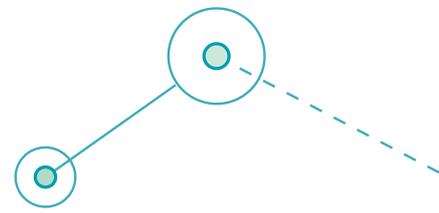
The purpose of this work package is twofold: 1) to develop modern AI for streaming and sensor-based data analysis and 2) to develop techniques and tools for the automatic creation and management of knowledge graphs.

The development of modern AI for streaming and sensor-based data will be done by

1. Providing anomaly detection and predictions with low quality streaming data
2. Providing uncertainty quantification and explainability with streaming data
3. Enabling combinations of streaming and static data for efficient data analysis

Streaming data can be used for automation, recommendations and decision making. Often this involves predictions and anomaly detection in multivariate time series, as well as providing explanations for the conclusions drawn. IoT sensors are increasingly instrumenting the physical world, and efforts have been made to use AI for solving these tasks also in low-quality data regimes. This research area will identify robust techniques for analysis of streaming data within several domains (including telco network, industrial IoT), with a particular focus on improving interpretability for cases with multivariate time series with low quality data.

Solving the research problems in STREAM is crucial to successfully innovate how IoT data can be fully used in anomaly detection and contribute to breakthrough in applying AI in predictive maintenance and operational availability.



Real impact of data-driven AI depends on the availability of live data of sufficient quality and quantity in an automatically discoverable format that both humans and machines can understand. DATA will investigate how the semantics of data, through automatic creation and mapping of suitable knowledge graphs, can be leveraged to scale AI models from one situation to all similar situations and how complex graph-based structures can be efficiently stored and processed.

Work package leader:

Andris Piebalgs, Data Science Lead, Cognite (until Aug -23);

Valeriya Naumova, Head of Data Science, Cognite (from Sep -23)



WP INNOECO: AI INNOVATION ECO SYSTEMS

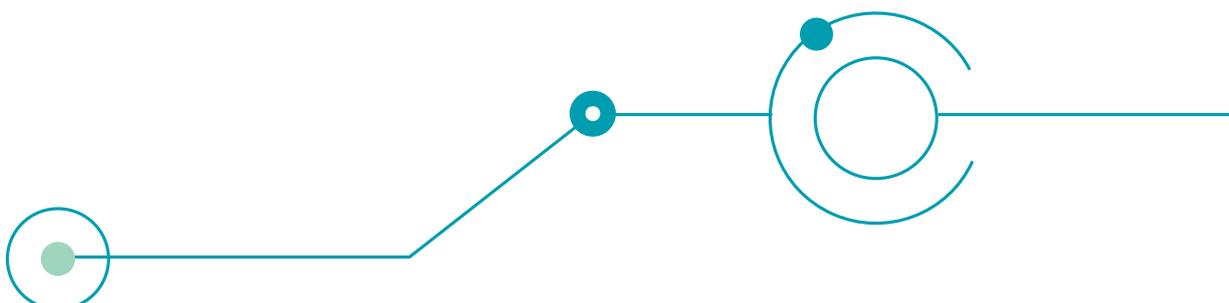
The purpose of this work package is to

- Create an Innovation Ecosystem among NorwAI partners for sharing both research findings, innovations, business solutions and change models within the AI domain.
- Share this knowledge with a broader audience, emphasizing small and medium sized businesses.

Artificial Intelligence still evokes more myths associated with threats than the possibilities technology brings to the table. AI experts are perceived speaking a foreign language when communicating to their organization, and most managers lack sufficient understanding of data driven business-models.

INNOECO will convey state of the art knowledge on AI-driven business models to a wide audience, including non-AI experts. This will be achieved by case analysis, conference presentations, and provision of a range of educational materials spanning from short tutorials to full university courses. The following topics will be covered: what AI is, its strategic impact on business model innovation, the business opportunities and limitations, implication on HR and teamwork, and innovation team dynamics, etc. To achieve these objectives, both AI and HR expertise will be convened.

Work package leader: Nhien Nguyen, Associate Professor, NTNU





Director of NorwAI, professor Jon Atle Gulla (right) and Associate Professor Benjamin Kille.



Presentations and communication

NAME	TITLE	TYPE/PLACE	DATE
Baumgartner, David; Storesund, Karolina;	Artificial intelligence - what is it, and how could it be useful in fire safety?	Seminar, Fire Research and Innovation Centre	2023-01-27
Kille, Benjamin	Large Language Models --- ChatGPT and its siblings	6th GAMI Media Lab Days (GAMI - Global Alliance for Media Innovation)	2023-01-27
Løland, Anders	Å forstå og bruke statistikk når man dekker forskning	Gjesteforelesning, Science and the media, Oslo Met	2023-02-01
Bach, Kerstin	Expertise in Action: Developing AI-Driven Solutions with Interdisciplinary Collaboration	Guest Lectures at US Naval Academy	2023-02-07
Gulla, Jon Atle	Hybrid Work with Large Language Models	IE-arrangement? Besøk USA	2023-02-10
Løland, Anders	Mer nyansert om algoritmetilsyn	Debattinnlegg: Morgenbladet	2023-02-10
Kille, Benjamin	Norwegian Language Models - Competing with Chat-GPT	Catch IDI	2023-02-24
Løland, Anders	Når virkeligheten kan hackes: hva er sant og hva kan vi stole på? Og hva gjør vi med det? (panelsamtale)	LYS OG MØRKE, TEKNOLOGI OG DEMOKRATI, EPISODE 2	2023-03-02
Kille, Benjamin	Building Large Generative Language Models at NorwAI	SFI CGF seminar	2023-03-16
Hammarstedt, Mark; Løland, Anders; Yazidi, Anis; Zouganeli, Evi; Ugland, Yngvar; Andersen, Eirik; Irgens, Morten	How can Norway become a Leading AI Nation, as Sweden & Finland - Panel discussion	Generative AI Meetup - Data & AI for Intelligent Innovation & Industries	2023-03-21
Løland, Anders	Panel: How Norway Can Become a Leading AI Nation, as Sweden & Finland	GENERATIVE AI WORKSHOP	2023-03-23
Bach, Kerstin	Master thesis @ NorwAI and the Norwegian Open AI Lab	BRAIN lunsjseminar	2023-03-24
Gulla, Jon Atle	Student information about NorwAI	BRAIN lunsjseminar	2023-03-24
Bach, Kerstin; Tessem, Bjørnar	Nordic Seminar on eXplainable AI (XAI)	2-day Workshop on eXplainable AI (XAI)	2023-03-29 - 2023-03-30
Eide, Simen	Relevance through diversity	Nordic Personalisation day, Oslo	2023-03-30

NAME	TITLE	TYPE/PLACE	DATE
Vassøy, Bjørnar	Consumer-side Fairness in Recommender Systems	Nordic Personalisation day, Oslo	2023-03-30
Zhang, Lemei	Pre-train, Prompt and Recommendation: Language Modeling Paradigm Adaptations for Recommender Systems	Nordic Personalisation day, Oslo	2023-03-30
Hafver, Andreas; Pedersen, Frank Børre	Beyond words? The possibilities, limitation, and risks of large language models	dnv.com	2023-04-12
Løland, Anders	Edrueilig og ansvarlig kunstig intelligens: Hvordan bør samfunnet håndtere de nye mulighetene?	Nasjonalt nettverk for kritisk medieforståelse	2023-04-18
Gulla, Jon Atle	En reise med språkmodeller	MUST AI Middag, Gyldendal	2023-05-02
Kille, Benjamin	Large language models at NorwAI	Friday Talk, Fakultet for informasjonsteknologi og elektroteknikk	2023-05-05
Løland, Anders	Hvorfor er kunstig intelligens så vanskelig å regulere?	Dagens næringsliv	2023-05-05
Gulla, Jon Atle	Kunstig intelligens i NorwAI	Gruppemøte for Ap i Stortinget	2023-05-08
Sven Størmer Thaulow, Schibsted	Kunstig intelligens: Dyrest å sitte stille	Medieleder 2023 Del 2	2023-05-10
Løland, Anders	Hva er kunstig intelligens, hva kan den gjøre og hva kan den ikke gjøre?	eForvaltningskonferansen 2023	2023-05-23
Gulla, Jon Atle	Norsk forskningscenter for AI-innovasjon	Presentasjon for SFI-Forum Forskningsrådet	2023-05-23
Håkon Høgset, Alexander Michael Ås, Marte Eggen	Demo Kaia	Regjeringens kick-off for ny nasjonal digitaliseringsteknologi	2023-06-06
Ingvaldsen, Jon Espen	[Presentasjon og deltakelse i panel]	Gyldendalkonferansen	2023-06-08
Michalowska, Katarzyna; Goswami, Somdatta; Karniadakis, George Em; Riemer-Sørensen, Signe	Neural operator learning with recurrent neural networks	Mathematical and Scientific Machine Learning (MSML 2023)	2023-06-08
Kille, Benjamin	Introduksjon til språkmodeller	AI-verksted for forlagsbransjen, Den Norske Forleggerforening, NAIL, NorwAI	2023-06-14
Bach, Kerstin	Trondheim Tech Port Innovation Breakfast: The Role of AI in Future Healthcare Services	Trondheim Tech Port Innovation Breakfast	2023-06-20
Bach, Kerstin	An Introduction to NorwAI	Innovation Summit, Pfizer, UK	2023-06-22
Gulla, Jon Atle	Norwegian Language Models	Møte Teknologirådet	2023-06-28

NAME	TITLE	TYPE/PLACE	DATE
Eidnes, Sølve; Bøhn, Eivind Egil; Lye, Kjetil Olsen; Myhr, Håkon Noren; Riemer-Sørensen, Signe; Stasik, Alexander Johannes; Sterud, Camilla	Pseudo-Hamiltonian neural networks	22nd ECMI Conference on Industrial and Applied Mathematics	2023-06-30
Michalowska, Katarzyna; Goswami, Somdatta; Karniadakis, George Em; Riemer-Sørensen, Signe	Multi-resolution learning with operator- and recurrent neural networks	CRUNCH seminar	2023-08-09
Løland, Anders	Når kommer neste nedtur for kunstig intelligens?	Dagens næringsliv	2023-08-10
Gulla, Jon Atle	Hvordan skape gode språkmodeller?	Presentasjon på Arendalsuka	2023-08-17
Thaulow, Sven Størmer	En revolusjon for norske medier--(og mange andre)	Arendalsuka	2023-08-17
Kille, Benjamin	AI and Language Models (Kunstig intelligens og språkmodeller)	NTL delegasjon besøk på AI-laben	2023-08-24
Gulla, Jon Atle	Hvordan kan en statlig norsk språkmodell se ut, og hva skal vi med den?	Nettmøte: Bør Norge bygge en utfordrer til ChatGPT?	2023-08-29
Andreas Hafver, Tita Alissa Bach, Asun Lera St.Clair	"Building trust in AI. Perspectives from different industries"	Building trust in AI - DNV	2023-08-30
Ingvaldsen, Jon Espen	Store språkmodeller	Frokostmøte om AI, DIGS	2023-08-30
Kille, Benjamin	Kunstig intelligens i oppvekstsektoren	KS Nordland, Bodø	2023-09-01
Kille, Benjamin	With the rise of GPT4, Bard, Llama 2 et al., is there still a need for localised (Norwegian) and specialised (medical) NLP models?	KIN nettverksmøte	2023-09-07
Svendsen, Rolf D.	NorwAI og norske språkmodeller	Kunstig Intelligens i Offentlig Sektor, KS	2023-09-13
Kille, Benjamin	Demystifying Large Language models ... in 15 minutes or so	DNB-styrets årlige seminar	2023-09-13
Gulla, Jon Atle	Språkmodeller i NorwAI	Techport TrondheimYes, møte om medieteknologi	2023-09-15
Bach, Kerstin	Artificial Intelligence Workshop	Workshop (artificial intelligence), WomENCourage	2023-09-20
Svendsen, Rolf D.	NorwAI og norske språkmodeller	Om norsk språkmodell, Innovarena, Rørvik	2023-09-20
Gulla, Jon Atle	Norske generative språkmodeller	Seminar hos Forskningsrådet	2023-09-21
Gulla, Jon Atle	Kunstig intelligens - en norsk modell / Sofaprat	Thamskonferansen	2023-09-21

NAME	TITLE	TYPE/PLACE	DATE
Terje Brasethvik , Rolf D. Svendsen	Norske språkmodeller	Styremøte i Det Norske Samlaget	2023-09-07
Løland, Anders; Aas, Kjersti	Generation of synthetic data: methods for, lessons from and challenges with tabular data	dScience Synthetic Data Generation Workshop	2023-09-25
Svendsen, Rolf D.	NorwAI og norske språkmodeller	Om språkmodeller på styremøte Kopinor	2023-09-25
Gulla, Jon Atle	Introduksjon til kunstig intelligens. Språkmodeller med dyp læring	Gjesteforelesning, NTNU Executive Education	2023-09-29
Nguyen, Nhien	Digital Strategy	Gjesteforelesning, NTNU Executive Education	2023-09-29
Ingvaldsen, Jon Espen	KI: Utvikling av store generative norske språkmodeller	Seminar, paneldebatt Norsk Oversetterforening	2023-09-30
Gulla, Jon Atle	NorGPT: Large Generative Language Models for Norwegian	Nordic Perspectives on Artificial Intelligence, Oslo	2023-10-13
Signe Riemer-Sørensen	Hybrid AI for industrial systems	Nordic Perspectives on Artificial Intelligence, Oslo	2023-10-13
Thaulow, Sven S.; Sahlgren, Magnus; Gulla, Jon Atle; Stenbom, Agnes	The Importance of Nordic Large Language Models; the status and way forward	Nordic Perspectives on Artificial Intelligence, Oslo	2023-10-13
Kille, Benjamin	NorwAI and research on Large Language Models	Visit of the French Ambassador to Norway	2023-10-18
Kille, Benjamin	Large Language Models-How are they created? How do they generate answers? And Panel debate	Noria	2023-10-19
Kille, Benjamin	NorGPT-Creating a Large Generative Language Model for Norwegian	IE Faculty Board meeting	2023-10-20
Zhang, Lemei; Liu, Peng	NorGLM: Norwegian generative language models	NTNU Tour under NOKIOS	2023-10-24
Gulla, Jon Atle	Industriell innovasjon i kunstig intelligens – klar til bruk?	Brohodekonferansen 2023, Frøya	2023-10-26
Ingvaldsen, Jon Espen	AI og store språkmodeller	Seminar UiT, ITA fagdager	2023-10-26
Ingvaldsen, Jon Espen	Generativ KI og viktigheten av store norske språkmodeller	Translatørdagen	2023-10-28
Ingvaldsen, Jon Espen	Responsible Norwegian Language Models	NorwAI Innovate Conference 2023	2023-10-31
Løland, Anders	Kunstig intelligens, hva er nå det egentlig?	Introduksjonsdagen for nyansatte i departementene og SMK	2023-10-31

NAME	TITLE	TYPE/PLACE	DATE
Kille, Benjamin	AI and Language Models for public news organizations	NRK	2023-11-01
Kille, Benjamin	NorGPT-en ny norsk språkmodell	Talk to a delegation of Nordea	2023-11-02
Nguyen, Nhien	Towards AI-driven business model: The case of NorGPT	Guest speakers, First Friday public event organized by EGGS Design, Trondheim	2023-11-03
Kille, Benjamin	NorGPT-arbeidet med en ny norsk språkmodell	Presentation to the Prins Daniel Fellowship	2023-11-08
Kille, Benjamin	Keynote under "Should Denmark invest in its own LLM?" and Panel debate	the Digital Tech Summit in Copenhagen	2023-11-09
Zhang, Lemei; Liu, Peng	NorGLM: Norwegian generative language models and an investigation on toxicity concerns	Research group for AI, ethics and philosophy - NTNU	2023-11-10
Gulla, Jon Atle	Den norske språkmodellen NorGPT	NTNU Kveld på Dokkhuset	2023-11-14
Ingvaldsen, Jon Espen	Large Language Models (LLMs): The Importance of Norwegian Models and how to Integrate LLMs with Business Data	DNB Data Science Meet-up	2023-11-20
Gulla, Jon Atle	Responsible language model with Norwegian content: The case of NorGPT (tentative)	IØT seminar "Leading Transitions"	2023-11-24
Özgöbek, Özlem	Fake news detection: From machine learning to LLMs	Artificial Intelligence Against Fake News: Workshop on Advances in Disinformation Detection (WADD 2023)	2023-11-30
Kille, Benjamin; Zhang, Lemei; Murad, Abdulmajid; Baumgartner, David; Xaviwer...			2023-12-08
Kille, Benjamin	Artificial Intelligence and Large Language Models—Developing Generative AI at NTNU	Studentkonferanse TDT70, NTNU	2023-12-08

NorwAI in the media in 2023

Trondheim's role in the GPT revolution

(The List Magazine , 2023-01-01)

Internasjonale akademikere kommer til Infomedia for å samarbeide

(uib.no, 2023-01-26)

Hvorfor er SFI-innovasjon så vanskelig?

(Khrono, 2023-01-28)

Microsoft med superutgave av ChatGPT - men også en ny advarsel

(Aftenposten, 2023-02-08)

Abels ekstra - Chat GPT tar feil!...sier Chat GPT - Regneoppgaven, siste kapittel

(Abels Tårn, NRK Radio, 2023-02-21)

-Vi tror våre norske modeller blir bedre

(Morgenbladet, 2023-03-31)

Språkmaskiner som kan sluke sivilisasjonen: Hva har vi sluppet løs?

(Morgenbladet, 2023-03-31)

KI lurte menneske trill rundt. Nå krever over 1000 eksperter forbud

(aftenposten.no, 2023-03-31)

Ikke bare klipp og lim: Nå skjønner ChatGPT sunn fornuft

(Aftenposten, 2023-04-17)

Inviterer til Kunstig intelligens-dugnad

(KOM24, 2023-05-10)

Kunstig intelligens: Dyrest å sitte stille, Sven Størmer Thaulow

(Medieleder 2023 Del 2, 2023-05-10)

Maskinene skal lære seg bedre norsk

(Finansfokus, 2023-05-23)

Blir eksamen rettferdig når ChatGPT er på banen?

(Gemini.no, 2023-05-24)

NorGPT kan bli redningen for det norske språket

(Universitetsavisa, 2023-06-06)

Vi har ingen relasjon til amerikanerne, men vi utnytter den samme teknologien

(Kapital, 2023-08-04)

Lenge til KI tar jobben din: - ChatGPT er en fantasifull kollega med ville ideer

(Universitetsavisa, 2023-08-10)

Arendalsuka: Første offentlige visning av NorGPT: - Regjeringen må kjenne sin besøkelsestid

(Universitetsavisa, 2023-08-17)

Arendalsuka: NorGPT er en modell som ennå ikke har lært seg folkeskikk

(Universitetsavisa, 2023-08-17)

Arendalsuka: Slik gikk det da ministeren spurte NorGPT om hvem som vinner valget

(Universitetsavisa, 2023-08-18)

Launch of Norwegian ChatGPT

(GCE NODE 2023-08-20)

Ambisiøs lansering av norsk «ChatGPT»

(Computerworld, 2023-08-22)

Mød Nor GPT: Norge får sin egen Chat GPT

(Børsen (Danmark), 2023-08-23)

Norske verdier har sjelden vært så truet

(Adressa, 2023-08-23)

Norge utvikler egne sprogmodel

(Børsen (Danmark), 2023-08-28)

Slik kan ChatGPT-juks avsløres

(Khrono, 2023-09-05)

Bruk KI-milliarden riktig. Sats på de åpne, globale språkmodellene

(digi.no, 2023-09-14)

Vi må skynde oss klokt

(dagsavisen.no, 2023-09-28)

Tusenvis av studenter bruker den nye tjenesten: - Positiv respons

(trd.by, 2023-10-05)

Brohodekonferansen 2023: Søkelys mot økt potensial i digitaliseringen innen havbruk

(ntnu.no, 2023-10-19)

Det er på tide å utforske andre sider av kunstig intelligens

(aftenposten-no, 2023-11-01)

På tide å utforske andre sider av KI

(Gemini.no, 2023-11-17)

Høyre foreslår å utvikle nordisk «ChatGPT»

(NRK.no, 2023-12-23)

Lanserer norsk alternativ til ChatGPT

(digi.no, 2023-12-29)



Internal seminars

NAME	TITLE	TYPE/PLACE	DATE
Knaus, Tanja (UiO)	WP2: AI in Society	NorwAI Forum, Trondheim	2023-03-13
Bach, Tita Alissa (DNV); (Khan, Amna; Hallock, Harry; Beltr�om, Gabriella; Sousa, Sonia)	A systematic literature review of user trust in AI-enabled systems: A human computer interaction (HCI) perspective	NorwAI Forum, Trondheim	2023-03-13
Balog, Krisztian; (Bernard, Nolwenn; Lajewska, Weronika) (UiS)	On the role of human annotation in the era of large language models	NorwAI Forum, Trondheim	2023-03-13
Casolo, Simone (Cognite)	Topological data analysis of industrial time-series - Topological data analysis for industrial data science	NorwAI Forum, Trondheim	2023-03-13
Shadrooh, Shiva	Graph-based anomaly detection	NorwAI Forum, Trondheim	2023-03-13
Zhang, Lemei; Liu, Peng (NTNU)	Empowering News Summerization with Pre-trained Language Models, Lemei Zhang & Peng Liu	NorwAI&NAIL AI Seminar	2023-05-26
Liu, Peng; Zhang, Lemei	Empowering News Summerization with Pre-trained Language Models	NorwAI&NAIL seminar series on AI Research and Innovation	2023-05-26
Murad, Abdulmajid (NTNU)	Efficient and Informative Data Acquisition: Leveraging Reinforced Active Learning with Realistic Uncertainty Estimates	NorwAI Partner Lunch Seminar	2023-08-23
Bauer, Susanne	Trustworthy AI and Assurance	NorwAI Forum Oslo	2023-08-30
Dautov, Rustem	Exploring Federated Learning	NorwAI Forum Oslo	2023-08-30
Hafver, Andreas	Assurance of AI-enabled systems	NorwAI Forum Oslo	2023-08-30
Hafver, Andreas; Bach, Tita; St.Clair, Asun Lera	Building Trust in AI, Perspectives from different industries	NorwAI Forum Oslo	2023-08-30
Ingvaldsen, Jon Espen	NorGPT	NorwAI Forum Oslo	2023-08-30
Ortiz, Pablo; Malacarne Sara	Domain adaptation for anomaly detection in the network	NorwAI Forum Oslo	2023-08-30
Treullier, C�lina (University of Lorraine, France)	Between understanding and modeling: the delicate quest of ethical and depolarizing news recommender systems	NorwAI Partner Lunch Seminar	2023-09-13
Ruocco, Massimiliano (SINTEF)	Irregular Time Series: Update and Current Work	NorwAI Partner Lunch Seminar	2023-09-27
Saeid Samizade, (Retriever)	Entity-based Sentiment model	NorwAI Partner Lunch Seminar	2023-10-11

Personnel

KEY RESEARCHERS

NAME	INSTITUTION	MAIN RESEARCH AREA
Andris Piebalgs	Cognite	Big Data, Information Retrieval, Machine Learning, Hybrid Analytics
Valeriya Naumova	Cognite	Big Data, Machine Learning
Andreas Hafver	DNV	Trustworthy AI
Anders Løland	Norsk Regnesentral	Machine Learning, Statistics
Ildiko Pilán	Norsk Regnesentral	NLP, Data Privacy, Clinical NLP, Computer-assisted Language Learning, Machine Learning
Mark Anderson	Norsk Regnesentral	NLP, Data Privacy, Syntax, Cyber Security, Explainable AI, Trustworthy AI, Machine Learning, Model Evaluation
Pierre Lison	Norsk Regnesentral	NLP, Privacy and Security, Machine Learning, Information Retrieval, Semantics, AI and Society, Big Data
Benjamin Kille	NTNU	NLP, Personalization, Machine Learning
Boye Høverstad	NTNU	Machine Learning
Frank Alexander Kraemer	NTNU	Machine Learning, Internet of Things
Helge Langseth	NTNU	Machine Learning, Probabilistic AI
Jon Atle Gulla	NTNU	NLP, Personalization
Jon Espen Ingvaldsen	NTNU	NLP, Innovation processes
Kerstin Bach	NTNU	AI and Society, Innovation Processes, Personalization, Machine Learning, Trustworthy AI
Kjetil Nørvåg	NTNU	NLP, Data Mining, Database Systems, Big Data, Information Retrieval
Lemei Zhang	NTNU	Personalization, Big Data, Information Retrieval, Machine Learning
Nhien Ngyuen	NTNU	Innovation processes
Özlem Özgöbek	NTNU	Personalization
Terje Brasethvik	NTNU	NLP, Innovation processes
Aleena Thomas	SINTEF Digital	NLP and analysis of fake news
Alexander Johannes Stasik	SINTEF Digital	Machine Learning, Hybrid Analytics

NAME	INSTITUTION	MAIN RESEARCH AREA
Arne Jørgen Berre	SINTEF Digital	AI Innovation, Semantics and Standardization for AI
Francisco Martin-Recuerda	SINTEF Digital	Large Language Models for entity matching for ontologies in industrial systems and digital twins.
Pål Malm	SINTEF Digital	Innovation processes
Signe Riemer -Sørensen	SINTEF Digital	Machine Learning, Hybrid Analytics
Pablo Ortiz	Telenor	NLP, Semantics, Machine Learning, Hybrid Analytics
Sara Malacarne	Telenor	Big Data, Machine Learning, Trustworthy AI, Time Series
Erik Vellidal	UiO	NLP, Machine Learning
Lilja Øvreid	UiO	NLP, Machine Learning
Susanne Bauer	UiO	AI in Society
Krisztian Balog	UiS	NLP, Semantics, Personalization, Information Retrieval, Machine Learning

VISITING RESEARCHERS

NAME	AFFILIATION	NATIONALITY	SEX	DURATION	TOPIC
Rohl, Andrew	Curtin University in Perth	Australia	M	2023-10-02 – 2023-10-04	Human activity recognition
Treuillier, Céline	University of Lorraine	Frankrike	F	2023-08-28 – 2023-12-22	News recommender systems
Wiratunga, Nirmalie	RGU Aberdeen	UK	F	2023-03-02 – 2023-03-03	Case-based reasoning
Mise, Pelin	MEF University Turkey	Turkiet	F	2023-07-20 – 2023-09-29	Fake News Detection

POSTDOCTORAL RESEARCHERS

Postdoctoral researchers with financial support from the center budget:

NAME	NATIONALITY	PERIOD	SEX	TOPIC
Liu, Peng	Kina	2020-12-07 – 2023-12-06	M	Language models for natural language processing, Conversational systems
Murad, Abdulmajid	Jemen	2023-03-23 – 2024-01-31	M	Deep Reinforcement Learning, Active Learning, and GFlowNet

Postdoctoral researchers working on projects in the center with financial support from other sources:

NAME	FUNDING	NATIONALITY	PERIOD	SEX	TOPIC
Kille, Benjamin	NTNU	Tyskland	2021-05-01 – 2023-04-30	M	Personalization, Natural Language Processing, and Machine Learning for Banking and Finance
Zhang, Lemei	NTNU	Kina	2022-04-04 – 2023-01-31; 2023-07-31 – 2025-10-03	F	Streaming Data Analytics
Ayetiran, Eniafe F.	ERCIM	Nigeria	2023-01-01 – 2023-12-22	M	Fake News Detection

PHD STUDENTS

PhD students with financial support from the center budget:

NAME	NATIONALITY	PERIOD	SEX	TOPIC
Baumgartner, David	Austria	2021-09-21— 2025-08-31	M	Data analysis with noisy and low-quality data streams
Bernard, Nolwenn	France	2022-02-01 – 2025-01-31	F	Study of fairness and transparency in conversational recommender systems
Jedamski, Lena	Germany	2023-02-06- 2026-02-05	F	Trustworthy Artificial Intelligence
Knaus, Tanja	Austria	2022-04-01 – 2025-03-31	F	Automation of affective data: infrastructure and data practices in voice recognition systems
Łajewska, Weronika	Poland	2022-02-01 – 2025-01-31	F	Personalizing Conversational Informational Access
Murgorgo , Serinha	Zimbabwe	2023-08-15- 2026-08-14	F	Embracing Responsible AI: Integration of responsible AI principles in organisational practices
Michałowska, Katarzyna	Poland	2021-01-01 – 2024-12-31	F	Informed machine learning
Nikolov, Nikolay	Bulgaria	2021-01-01 – 2024-12-31	M	Flexible Deployment of Big Data Pipelines on the Cloud/Edge/Fog Continuum
Rønningstad, Egil	Norway	2021-10-14 – 2024-10-13	M	Entity-level Sentiment Analysis
Shamba , Abdul-Kazeem	Nigeria	2023-09-01- 2027-08-31	M	Artificial Intelligence for Anomaly Detection
Steppe , Jessica Annalena	Germany	2023-08-15- 2026-08-14	F	AI-Augmented Creativity in Innovation Teams
Vassøy, Bjørnar	Norway	2021-08-02 – 2025-08-01	M	Fairness, Accountability, Transparency and Privacy in Personalization/ Recommender systems
Yadav, Vandana	India	2023-11-02 – 2027-11-01	F	Natural Language Processing for personalized content summarization

PHD STUDENTS

PhD students working on projects in the center with financial support from other sources:

NAME	FUNDING	NATIONALITY	PERIOD	SEX	TOPIC
Barlaug, Nils	RCN/Cognite	Norway	2019-08-01 – 2022-07-31	M	Data Integration
Bayrak, Betül	NFR/NTNU (Exaigon)	Turkiye	2022-01-19 – 2025-01-18	F	Explainable Case-Based Reasoning
Bekkemoen, Yanzhe	NTNU	Norway	2019-10-01 – 2024-09-30	M	Probabilistic approaches to explainable AI and reinforcement learning.
Bjørn, Anna Rodum	NTNU	Norway	2021-11-11 – 2025-11-10	F	Explainable deep bayesian learning
Doan, Tu My	NTNU	Vietnam	2020-01-06 – 2024-01-05	F	Political Text Mining
Firouzjaei, Hassan Abedi	NTNU	Iran	2019-06-20 – 2023-06-19	M	Querying and mining location-based social network data
Killingberg, Ludvig	NTNU	Norway	2019-10-01 – 2024-09-30	M	Bayesian deep learning and reinforcement learning
Logacjov, Aleksej	NTNU	Germany	2021-05-01 – 2024-04-30	M	Self-supervised learning for Human Activity Recognition
Maropaki, Stella	RCN	Greece	2016-05-01 – 2022-05-31	F	Database systems
Ottersen, Stuart	NTNU Helse: AI Ment	Norway	2023-02-01 – 2027-01-31	M	Using machine learning to predict response and dropout in RCT data from digital sleep therapy
Shadrooh, Shiva	DNB	Iran	2020-02-20 – 2024-09-30	F	Anomaly detection in streaming graphs
Stojanovic, Mateja	NTNU	Serbia	2020-12-01 – 2023-05-16	M	Recommender Systems for Enhancing Students' Learning in Higher Education
Sylvester, Sophia	NTNU Helse: AI Ment	Germany	2023-02-01 – 2027-01-31	F	Radar Data and Machine Learning in Psychiatric Care: Novel Approaches for Advancing Sleep/Wake Estimation and Aggression Prediction
Tiwari, Shweta	NTNU	India	2017-10-01 – 2022-01-11	F	Machine Learning Method for Outlier Detection in Energy Trading
Xing, Yujie	DNB	China	2021-01-30 – 2023-01-27	F	Generative Conversational Agents
Zhang, Lemei	RCN	China	2016-04-01 – 2021-03-01	F	User Modeling and Recommender System

MASTER'S DEGREES (2023)

Master students obtaining their degree on NorwAI topics in 2023:

NAME	SEX	PERIOD	TOPIC
Albizreh, Huda Mohamad Muataz	F	2022-08-01 – 2023-06-30	Reinforcement Learning for Optimizing Mobile Networks
Andersen, Ingunn	F	2022-08-01 – 2023-06-30	Tracing Truth: How Journalists in Faktisk Verifiserbar Used Digital Technologies to Verify Information from the Warzone in Ukraine - An Actor-Network Approach to Newsroom Practices in Norway
Barstad, Caroline	F	2022-08-01 – 2023-06-30	Privacy in Recommender Systems: Inferring User Personality Traits From Personalized Movie Recommendations
Barthold, Ingebjørg Dora Maxine	F	2022-08-01 – 2023-06-30	Privacy in Recommender Systems Inferring User Personality Traits From Personalized Movie Recommendations
Berg, Andreas Bjelland	F	2022-08-01 – 2023-06-30	Business models for small producers of large language models: The case of NorGLM
Bhaiji, Nora Andrea	F	2022-08-01 – 2023-06-30	Governing Speech in Digital Era: Norwegian Policy Responses to Disinformation and Digital Educational Tools for Increasing Critical Thinking Skills
Birkeland, August Asheim	M	2022-08-01 – 2023-06-30	Point Cloud Instance Segmentation Large Scale Process Plants
Bjørnland, Johan Fredrik Rønningen	M	2022-08-01 – 2023-06-30	Exploring Procedural Content Generation and Personalized Learning for Virtual Driving Education
Borgen, Haldis Kvalvåg	F	2022-08-01 – 2023-06-30	Detecting Private-Sensitive Content in Norwegian Social Media
Bru, Thea	F	2022-07-01 – 2023-05-31	From issue to action. The role of social and technical arrangements for enacting digital transformation in Norwegian higher education.
Ekornrød, Emil Nøding	M	2022-08-01 – 2023-12-31	AI-based Learning of Touch Sensing
Fossland, Tore Apeland	M	2022-08-01 – 2023-06-30	Semantic Similarity Search over Spatio-textual Data
Gedde, Yrjar Bringedal	M	2022-08-01 – 2023-06-30	Exploring Procedural Content Generation and Personalized Learning for Virtual Driving Education
Haidu, Victoria Heggland	F	2022-08-01 – 2023-06-30	Business models for small producers of large language models: The case of NorGLM
Hanssen, Jørgen	M	2022-08-01 – 2023-06-30	Expanding Our Knowledge of Maritime Trade with AIS and Explainable AI Systems
Hasund, Anne Marie Skaar	F	2022-08-01 – 2023-11-01	A Data-Driven CBR and Clustering Method for Identifying Physical Activity Phenotypes
Holmsen, Sigurd	M	2022-08-01 – 2023-06-30	Pseudo-Hamiltonian System Identification
Hossain, Md Amjad	M	2022-08-01 – 2023-06-30	Explain errors and improve timeseries forecasting models using XAI
Imran, Zaim Ul-Abrar	M	2022-08-01 – 2023-06-30	Exploring the Efficiency of Zero-Cost Proxies in NAS for Human Action Recognition
Johansen, Anne-Sofie	F	2022-08-01 – 2023-06-30	Nurse Scheduling and Rescheduling: Combining Optimization with Machine Learning-Driven Demand Predictions

MASTER DEGREES (2023)

Master students obtaining their degree on NorwAI topics in 2023:

NAME	SEX	PERIOD	TOPIC
Kohmann, Eivind	M	2022-08-01 – 2023-06-30	Investigating the Capability of Generative Adversarial Networks of Capturing Implicit Laws in Physical Systems
Langli, Karoline Lillevestre	F	2022-08-01 – 2023-06-30	Sentiment Analysis of Customer Emails Using BERT
Løkkeborg, Thomas Holene	M	2022-08-01 – 2023-06-30	Deep Reinforcement Learning for International Diplomacy: Learning to Play Map Variants
Matsen, Erik Aunvåg	M	2022-08-01 – 2023-06-30	Uncertainty-aware Deep Learning a.k.a. Deep Bayesian Learning
Merola, Marco Antonio	M	2022-08-01 – 2023-01-31	Artificial Intelligence for assessing road usage e-scooters
Myhre, Henrik Øberg	M	2022-08-01 – 2023-06-30	Uncertainty-aware Deep Learning a.k.a. Deep Bayesian Learning
Nag, Bendik	M	2022-08-01 – 2023-06-30	Nurse Scheduling and Rescheduling: Combining Optimization with Machine Learning-Driven Demand Predictions
Nilsen, Øystein Løndal	M	2022-08-01 – 2023-06-30	From Uni-Modal to Multi-Modal Fake News Detection: The Impact of Visual Cues on Detection Performa
Norum, Ole-Magnus Vian	M	2022-08-01 – 2023-06-30	Exploring Self-Supervised Learning for Bird Vocalization Classification
Nyheim, Bendik	M	2022-08-01 – 2023-06-30	A Novel Application of Machine Learning to Develop Pointing Models for Current and Future Radio/Sub-millimeter Telescopes
Nystad, Erik Johan	M	2022-01-01 – 2023-11-15	Large Language Models for entity matching and data integration
Orvik, Alexander	M	2022-08-01 – 2023-06-30	Network models of pandemic modeling
Riibe, Eskil Reitan	M	2022-08-01 – 2023-06-30	Self-Supervised Human Activity Recognition on Free-Living Sensor Data Streams
Salomonsen, Simen	M	2022-08-01 – 2023-06-30	Biologically-Plausible Neural Networks
Schau, Max Torre	M	2022-08-01 – 2023-06-30	Exploring the Efficiency of Zero-Cost Proxies in NAS for Human Action Recognition
Singh, Anmol	M	2022-08-01 – 2023-06-30	Deep Reinforcement Learning for Spatio-Temporal Wildlife Management
Solbjørg, Ingrid Amalie	F	2022-08-01 – 2023-06-30	Using Sentiment Analysis to Improve Course Recommendations for MOOCs
Stensrud, Christopher	M	2022-08-01 – 2023-06-30	Semantic Text Indexing
Torjusen, Hanna	F	2022-08-01 – 2023-06-30	Privacy in Recommender Systems Inferring User Personality Traits From Personalized Movie Recommendations
Tveit, Herborg Hermansen	F	2022-08-01 – 2023-06-30	Nurse Scheduling and Rescheduling: Combining Optimization with Machine Learning-Driven Demand Predictions
Tveiten, Lukas Nicholas	M	2022-08-01 – 2023-06-30	Frameworks for Distributed Processing of Temporal Graph Queries

MASTER DEGREES (2023)

Master students obtaining their degree on NorwAI topics in 2023:

NAME	SEX	PERIOD	TOPIC
Vu, Thanh-Hi Anthony	M	2022-08-01 – 2023-06-30	Ethical aspects of AI/recommender systems
Zachariassen, Oline Vikøren	F	2022-08-01 – 2023-06-30	Detecting Private-Sensitive Content in Norwegian Social Media
Øien, Erlend	M	2022-08-01 – 2023-06-30	Enhancing Online Lecture Recommendations Through Exploration of Student Behaviour Across Topics
Ågotnes, Henrik Ytrehus	M	2022-08-01 – 2023-06-30	Hydropower optimization using model-based Reinforcement Learning
Ås, Alexander Michael	M	2022-08-01 – 2023-06-30	A Norwegian Whisper Model for Automatic Speech Recognition

MASTER'S STUDENT ASSISTANTS

NAME	SEX	PERIOD	TOPIC
Adam Dybwad Sioud	M	2023-06-05 – 2023-09-01	Large-scale Language Model for Norwegian
Alexander Michael Ås	M	2022-08-01 – 2023-06-30	“Kaia” the social robot
Anna Østmo	F	2023-06-12 – 2023-07-07	Large-scale Language Model for Norwegian
Birk Søråsen	M	2023-08-01 – 2023-12-31	Annotation
Daniel Skinstad Drabitzius	M	2023-08-01 – 2023-12-31	Annotation
Eivind Kjosbakken	M	2023-08-21 – 2023-12-31	Large-scale Language Model for Norwegian
Elizabeth Ainan Pan	F	2022-08-01 – 2023-05-31	Communication
Ellen Margrethe Ulving	F	2023-08-01 – 2023-12-31	Annotation
Even W. Lauvrak	M	2023-06-01 – 2023-12-31	Large-scale Language Model for Norwegian
Håkon Bygdås Høgset	M	2022-08-01 – 2023-12-31	“Kaia” the social robot
Håkon Liltved Hyrve	M	2023-08-01 – 2023-12-31	Annotation
Henny Løfqvist	F	2023-08-21 – 2023-12-31	Large-scale Language Model for Norwegian

MASTER'S STUDENT ASSISTANTS

NAME	SEX	PERIOD	TOPIC
Ingrid Omdal	F	2023-08-21 – 2023-12-31	Large-scale Language Model for Norwegian
Isak Tønnesen	M	2023-06-26 – 2023-12-31	Large-scale Language Model for Norwegian
Iunia Melania Antal	F	2023-08-01 – 2023-12-31	Annotation
Jakob Tøssebro	M	2023-06-05 – 2023-12-31	Large-scale Language Model for Norwegian
Jonas Sundt	M	2023-06-12 – 2023-12-31	Large-scale Language Model for Norwegian
Marcus Tiedemann Økland Henriksbø	M	2023-08-22 – 2023-12-31	Large-scale Language Model for Norwegian (+ INNOECO)
Marte Eggen	F	2022-08-01 – 2023-12-31	“Kaia” the social robot
Selma Dale Ditlev-Simonsen	F	2023-09-04 – 2023-12-31	Large-scale Language Model for Norwegian
Terje Nissen Farup	M	2023-07-01 – 2023-12-31	Large-scale Language Model for Norwegian
Vilhjalmur Arnar Vilhjalmsson	M	2023-07-12 – 2023-12-31	Demonstrator for anomaly detection
Vegard Vaeng Bernhardsen	M	2023-10-08 – 2023-12-31	Large-scale Language Model for Norwegian

BACHELOR'S STUDENT ASSISTANTS

NAME	SEX	PERIOD	TOPIC
Elliot Song (Waterloo University, iWIL AI project)	M	2023-09-01 - 2023-12-31	Social robotics programming
Tony Li (Waterloo University, iWIL AI project)	M	2023-09-01 - 2023-12-31	Artificial intelligence specialist
Hargun Mujral (Waterloo University, iWIL AI project)	M	2023-09-01 - 2023-12-31	Software Engineering (Artificial Intelligence)
Stanley Miao (Waterloo University, iWIL AI project)	M	2023-09-01 - 2023-12-31	Programmer/software developer in artificial intelligence research



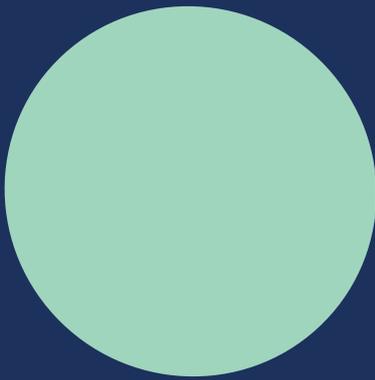
Accounts

2023	FUNDING	COSTS
The Research Council	11 701	
The Norwegian University of Science and Technology (NTNU)	10 253	15 169
Research Partners*)	2 208	8 993
Enterprise partners**)	10 278	10 278
Total	34 440	34 440

*) Norwegian Computing Center (NR), SINTEF, University of Oslo and University of Stavanger.

***) Aneo, Cognite, Digital Norway, DNB, DNV, Kongsberg Digital, NRK, Retriever Norway, Schibsted, Sparebank1 SMN, Statnett, Telenor.





NorwAI

Norwegian Research Center
for AI Innovation



 The Research
Council of Norway

 NTNU

Editor: Rolf Dyrnes Svendsen

Coordinator: Karolina Storesund

Center photos: Kai T. Dragland, NTNU - Ole Martin Wold

Graphic design: NTNU Grafisk senter