Implementation of Machine Learning in Division for Health Services: Achievements in 2023 and Strategy Proposal for 2024

Utgitt av	Folkehelseinstituttet Område for helsetjenester
Tittel	Implementering av maskinlæring i område for helsetjenester: Resultater for 2023 og strategiforslag for 2024
English title	Implementation of Machine Learning in Division for Health Services: Achievements in 2023 and Strategy Proposal for 2024
Ansvarlig	Gun Peggy Strømstad Knudsen, konstituert direktør
Forfattere	Tiril C. Borge, Forsker Heather Melanie Ames, Seniorforsker Patricia Jacobsen Jardim, Seniorrådgiver Hans Bugge Bergsund, Forsker Martin Smådal Larsen, Rådgiver Christopher Rose, Forsker
ISBN	xxxxx
Publikasjonstype	Report (Rapport)
Antall sider	31
Oppdragsgiver	Folkehelseinstituttet
Emneord(MeSH)	biomedical; artificial intelligence, health; unsupervised machine learning; supervised machine learning; deep learning
Sitering	Borge TC, Ames HM, Jardim PJ, Bergsund HB, Larsen MS, Rose C, (2023). "Implementation of Machine Learning in Division for Health Services: Achievements in 2023 and Strategy Proposal for 2024 " [Implementering av maskinlæring I divisjon for helsetjenester: Resultater for 2023 og strategiforslag for 2024]. Oslo: Folkehelseinstituttet.

Contents

Key Messages

In 2020, the Norwegian Institute of Public Health (NIPH) established a machine learning (ML) team to align with the institute's automation and workflow innovation strategies. The ML team, has since become an international leader in integrating and implementing ML into evidence synthesis, achieving significant milestones, and securing official financing in November 2022, which contributed to much of the ML activities performed by the current iteration of the team.

This report is divided into two parts, covering the team's activities in 2023 and a strategy suggestion for 2024 based on the team's experiences from inception to late 2023.

In 2023, ML Team 3.0 accomplished a variety of project deliverables, including providing ML support to six teams, conducting teaching sessions, implementing an ML reporting template, and implementing e-learning course. Dissemination efforts included presentations, poster sessions, and publications, while evaluations encompassed various projects, including a pilot on interrater agreement using ChatGPT. Innovations comprised development of a scalable e-learning course, a survey on ML attitudes and barriers, and qualitative interviews.

In response to the evolving needs of our institute and in alignment with the strategic objectives of both the institute and the division, this proposal advocates for the elevation of the ML team to Division level in 2024. This will ensure long-term sustainability and decrease financial burdens for the cluster. In addition, we propose a restructured organizational framework with three teams: Innovation and Horizon Scanning, Evaluation and Evidence Building, and Implementation and Support, as well as a steering committee to coordinate activities and engage in external networking.

Title:
Implementation of Machine Learning in Division for Health Services: Achievements in 2023 and Strategy Proposal for 2024

Publisher:
The Cluster of Reviews and Health Technology Assessments, Division for Health Services, ML team 3.0

Type of publication:
Report

Executive summary (English)

Background

In early 2020, the Cluster for Reviews and Health Technology Assessments at the Norwegian Institute of Public Health (NIPH) established a dedicated machine learning (ML) team, aligning with NIPH's strategies for automation and workflow innovation. The Division for Health Services, driven by specific goals, sought to automate work processes and efficiently summarize evidence using ML. Since its inception in late 2020, the ML team has positioned NIPH as a leader in implementing ML into evidence synthesis, achieving significant milestones, and securing official financing in November 2022.

The report is divided into two parts. The first part outlines the team's activities in 2023, covering implementation, peer-to-peer support, dissemination, evaluations, innovation, horizon scanning, and external networking and collaborations. The second part offers suggestions for the 2024 strategy based on experiences from 2020 to late 2023.

Lessons learned in 2023 highlight external funding as a strength, collaboration with EPPI for addressing technical aspects, and the value of interdisciplinary collaboration. Weekly meetings, intrinsic motivation, and conference attendance have contributed positively. Challenges include the absence of protected time, difficulty keeping up with AI/ML developments, and the need for a rotating membership structure. Clear mandate from leadership is emphasized for improved team effectiveness.

ML team 3.0 deliverables 2023

ML Team 3.0 achieved a large range of project deliverables in 2023. The team provided ML support to 6 teams, conducted teaching sessions at the National Institute of Occupational Health in Norway (STAMI), implemented an ML reporting template, and developed an e-learning course on key ML functions. Dissemination efforts included ten domestic and international presentations, poster sessions, and publication of two protocols, one book chapter, and a paper for review. Evaluations were completed on the Cochrane RCT classifier, ML vs no ML use in reports, and a survey on ML use amongst the employees in the cluster. Ongoing evaluations involve a pilot project on interrater agreement using ChatGPT. Innovations included development of a scalable e-learning course and a survey assessing ML attitudes, barriers and perceptions, and qualitative interviews exploring the same concepts. While horizon scanning was limited due to resource constraints, the team has ongoing networking with eight external groups, fostering cross country collaboration on ML evaluations and strategy development.

Suggested ML strategy from 2024

In response to the evolving needs of our institute and in alignment with strategic objectives of both the institute and the division, this proposal advocates for the elevation of the ML team from cluster level to Division level. This transition is strategically positioned to ensure long-term sustainability of our ML efforts and decrease financial burdens for the cluster. The proposal aligns with both the institute and Division strategies, emphasizing the importance of evolving infrastructure, knowledge support, and expertise in cutting-edge methods like machine learning. Elevating the ML team is seen as imperative for innovation and collaboration in evidence synthesis and public health, preventing limitations in capacity and demotivation of team members.

Without this transition, there is a high risk of limited capacity, tasks, and time for the team, hindering its ability to keep up with advancements in the field. There is also the potential of becoming a "maintenance team," leading to demotivation and high turnover.

ML Team 4.0 organization

To address the challenges of rapid growth in ML and AI tools, a restructured organizational framework is proposed for the future, dividing the team into three distinct teams: Innovation and Horizon Scanning, Evaluation and Evidence Building, and Implementation and Support. Each team would have specific responsibilities, and a steering committee comprising the team leads would be established to coordinate activities and engage in external networking. The restructuring aims to enhance resource efficiency, introduce new competencies, increase the talent pool, and facilitate collaboration across departments. Clear criteria for team members are outlined based on their roles, emphasizing skills in ML and AI, communication, teaching, and experience in evaluation and implementation.

Suggested focus areas for 2024

The current ML team has put forward suggestions for key focus areas for the next iteration of the team, which align with the overarching goals of the Division and institute. These include seeking external funding for financial sustainability, improving machine readability through DOI numbers, leveraging OpenAlex for knowledge development, and fostering continued collaboration with external institutions, particularly EPPI. The team also aims to strengthen interdisciplinary collaboration within NIPH, build expertise in Generative AI, develop a comprehensive ML implementation package for training other institutions, explore and evaluate data extraction tools, seek cost funding for networking through ICASR, and investigate the application of ML/AI in the institute's registry work.

Preface

This report presents results for the current iteration of the ML team, "ML 3.0" and strategy suggestions for the next iteration of the machine learning team, "ML 4.0". The current team has crafted these suggestions based on our reflections of our successes, challenges, and learning during 2023.

Financing

Most of the work, particularly relating to innovation activities, was externally funded via DFØ. The remaining work was self-initiated and financed by the Cluster for Reviews and Health Technology Assessments, Division for Health Services at the NIPH.

With appreciation

The current team's learning and strategizing are due not only to the dedication of its members, past and present, but also to HTV leadership's investment and vocal support. There have also been numerous colleagues who have provided support, feedback, ideas, and opportunities including the team of librarians who have started evaluation work on OpenAlex. Outside of NIPH, James Thomas' mentoring and his team at EPPI Centre have continued to be instrumental to our understanding of ML and its potential to provide the most valuable evidence synthesis products to our commissioners.

Conflicts of interest

All authors declare they have no conflicts of interest.

Kåre Birger Hagen	Rigmor C Berg	Tiril Borge	Heather Ames
Research director	*Department director*	*Project leader*	*Implementation leader*

Background

Since early 2020, the Cluster for Reviews and Health Technology Assessments at the Norwegian Institute of Public Health (NIPH) recognized the potential benefits of employing machine learning (ML) in evidence syntheses. Consequently, a dedicated ML team was funded in late 2020, aligning with NIPH strategies for 2019-2024 focused on automation and workflow innovation.

Driven by division-specific goals, the Division for Health Services aimed to play an active role in automating work processes, utilizing ML to efficiently summarize evidence. Since its inception in late 2020, the ML team has positioned NIPH as a leader in integrating and implementing ML into evidence synthesis, strategically innovating to ensure a sustainable competitive advantage.

ML optimally utilizes scarce human resources by handling complex, repetitive tasks, allowing human involvement at critical points such as training, interpretation, and quality checks. The ML Team 2.0 secured official funding in November 2022, which has funded most of the ML related work in HTV during 2023.

This report is divided into two main sections: In the first part of this report, we will present our activities during 2023. These deliverables will be presented in the following categories: Implementation and peer-to-peer support, Dissemination, Evaluations, Innovation, Horizon scanning, capacity building and external networking and collaborations. In the second part of this report, we present our suggestions for further strategy for 2024, based on our experiences so far, from the team's inception in 2020 up until late 2023.

ML team 3.0 deliverables Q1 2023 – Q1 2024

Implementation and peer-to-peer support

Table 1 details team activities related to implementation and peer-to-peer support undertaken from January to December 2023. We have provided ML support to 6 teams across the institute, as well as holding ML teaching sessions at STAMI. We have also implemented the updated reporting template for ML use in our reports and an e-learning course to foster conceptual understanding of our most used ML functions.

Table 1: Implementation and peer-to-peer support activities in 2023

Date	Title	Type of deliverable
January	Implementing the new machine learning template for our systematic reviews and scoping reviews. These were developed, peer reviewed, and pilot tested before they were implemented	Implementation template
January	Help to report ML in « Triclosan coated sutures for prevention of surgical site infection: a health technology assessment»	Peer-to-peer support
March-May	Machine learning education for STAMI, and support to use EPPI. Also contributed to implementing ML in their systematic review protocol *"Title: Stable employment and mental health in the working age population, a protocol for a systematic review with meta-analyses of longitudinal studies. Authors: Fiona Aanesen; Rigmor C Berg; Patricia Sofia Jacobsen Jardim; Lillebeth Larun; Ingrid Løken Jørgensen; Benedicte Mohr; Karin Proper; Lars-Kristian Lunde"*	Machine learning teaching sessions at STAMI
March-May	Machine Learning support to "A knowledge resource for municipalities"	Peer-to-peer support

April-May	General EPPI Support to Norwegian Scientific Committee for Food and Environment (VKM), answering EPPI related questions and supporting the set-up of EPPI-VIS.	Peer-to-peer support
May-September	ML support for global health for a qualitative evidence synthesis entitled "Using evidence from civil society in health policy processes: a qualitative evidence synthesis"	Peer-to-peer support
October	Machine learning support for the project "CFS/ME IPD screening"	Peer-to-peer support
November	Implementing a conceptual e-learning with 4 modules for the most used machine learning algorithms/functions in EPPI Reviewer	E-learning
November-December	Machine learning support to "New national guidance on new recommendations for screening of resistant microbes", Division of Infection Control	Peer-to-peer support

In table 2 we provide a list of reports either published or completed in 2023 that have reported the use of at least one ML function. There might be other publications, like "forskningsomtaler", that have used ML but due to the individual publications word limitations, have not described their use of ML. ML support was provided to all but one of the reports presented in table 2.

Table 2: Internal reports using ML, published, or completed in 2023.

English Title	**Tool used**	**ML functions used**	**Provided ML support?**
Children and young people who perpetrate serious acts towards others: a rapid review	EPPI	OpenAlex, Priority screening	Yes*
Children and young people's opinions on topics in the proposal for a new Children's Act: a scoping review of Nordic qualitative studies	EPPI	Priority screening, Cochrane RCT classifier, Clustering	Yes*
The use of force and limit-setting for children and youth in residential childcare and foster care: systematic scoping review (update)	EPPI	OpenAlex, Priority screening	Yes*
Portable ECG equipment for the diagnosis of atrial fibrillation in the specialist health care: rapid HTA- scoping review	EPPI	Priority screening	No
Co-therapy and reflecting teams in couples- and family therapy: a mixed methods systematic review	EPPI	Priority screening	Yes*

Parental follow-up in family welfare services after child removal: a scoping review	EPPI	Priority screening, clustering	Yes*
What contributes to stable placements when children are placed in foster homes or institutions? Systematic literature search with sorting	EPPI	OpenAlex, Priority screening	Yes*
Individual placement and support for people with moderate to severe mental illnesses or substance abuse disorder: a systematic review	EPPI	Priority screening	Yes*
What are the characteristics of youth who are placed in care institutions in child welfare? A rapid review	EPPI	OpenAlex, Priority screening	Yes*
Consequences of the Covid-19 pandemic on children and youth's life and mental health: Second update of a rapid review	EPPI	OpenAlex, Priority screening, custom classifier	Yes*
Post COVID-19 condition after Omicron: a rapid review	EPPI	Priority screening	Yes*
Surgery for degenerative rotator cuff tears: a health technology assessment	EPPI	Priority screening, custom classifier	Yes
Transcutaneous non-invasive vagus nerve stimulation (gammaCore) for the treatment of cluster headache: A single technology assessment	EPPI	Open Alex, Priority screening, Cochrane RCT classifier	Yes
Triclosan coated sutures for prevention of surgical site infection: a health technology assessment	EPPI	Cochrane RCT classifier, Economic evaluation classifier, Priority screening	Yes
Coercion in mental health care and violence: systematic literature search with sorting	EPPI	Priority screening	Yes*
Language screening tools for children 0-5 years: a systematic scoping review	EPPI	Priority screening, OpenAlex, custom classifier, clustering	Yes*

**A member of the ML team was also a member of the project team*

Dissemination

The team has been very active in 2023 disseminating our work, mainly outside NIPH. We have held one presentation in-house and ten presentations outside of NIPH, where three were outside of Norway. Additionally, we have four presentations booked for January and February 2024. We have had three poster presentations, and one additional poster planned for a Norwegian implementation conference (NIMP) in January 2024. A summary of all our dissemination activities is found in Table 3.

Table 3: Dissemination activities Q1 2023 – Q1 2024

Date	English Title/Description	Type
March	Implementation and evaluation activities to build support for machine learning in evidence syntheses	Poster, NIPH Research and innovation day
May	Managing the information explosion: the usefulness of machine learning in SIA – the NIPH example	Presentation, INSIA annual meeting, Stockholm
August	Machine learning versus automation in evidence syntheses	Presentation, NORNESK webinar
September	Use of artificial intelligence and machine learning in evidence syntheses	Presentation, Division seminar
September	Can using the Cochrane RCT classifier help speed up study selection in qualitative evidence syntheses (QES)? A retrospective evaluation	Poster Cochrane Colloquium, London
September	Building acceptance for machine learning in study selection within a systematic review institution: Experiences from the Norwegian Institute of Public Health	Long Oral Presentation, Cochrane Colloquium, London
September	Connecting with other researchers who are working with ML or who want to work with ML. Networking for future collaborations.	Cochrane networking
September	NIPHs most frequently used machine learning functions in evidence syntheses	Presentation, NORNESK webinar
September	Information about ML team and our work published on the NIPH webpage about AI at FHI	NIPH AI webpage
September	Implementation and evaluation activities to build support for machine learning in evidence syntheses	Poster, Public Health Conference in Tromsø
September-October	To make the ML resources available on our SharePoint site more appealing and inviting, and more intuitive with regards to where you can find the different resources available.	Create new SharePoint site
October	How can machine learning be used to keep you updated on research areas?	Presentation, NORNESK webinar
October	Use of artificial intelligence and machine learning in evidence syntheses	Presentation, VKM
October	Building acceptance for machine learning in study selection within a systematic review institution: Experiences from the Norwegian Institute of Public Health	Presentation, Will Moy (Campbell Collaboration) & UK cabinet office
November	How to implement machine learning in evidence syntheses	Presentation, NORNESK conference, Bergen

November	How does the use of machine learning in evidence syntheses affect our work processes?	Presentation at NORNESK conference, Bergen
November	Experiences with the use of artificial intelligence and machine learning in evidence syntheses	Presentation, OsloMet seminar "Artificial intelligence in evidence syntheses - is OsloMet keeping up?"
January, 2024	How NIPH utilize AI tools in the evidence synthesis process	Presentation, University of Oslo
January 2024	Implementation guidance	Poster, Norwegian Network for Implementation Research (NIMP) Conference
February 2024	Tentative title: How can AI be used in evidence synthesis processes?	Presentation, Norwegian Poisons Information Centre
February 2024	Tentative title: Adopting and evaluating machine learning to accelerate evidence synthesis in public health and welfare (no formal invitation yet)	Presentation, National Academy of Sciences, Engineering and Medicine, Texas A&M Institute for Advancing Health Through Agriculture
February 2024	Tentative title: Building acceptance for machine learning in study selection within a systematic review institution: Experiences from the Norwegian Institute of Public Health	Presentation, UK Health Security Agency

Publications and preprints

In 2023 we have published two protocols and one book chapter, and one paper is in review per December 2023. Additionally, two papers and one report are in progress, aimed at publication during Q1 2024. Details on the publications are listed in table 4.

Table 4: List of publications and preprints published or planned Q1 2023 – Q1 2024

Date	Title	Type
January	**Muller, A. E., Berg, R. C., Meneses-Echavez, J. F., Ames, H. M. R., Borge, T. C., Jardim, P. S. J.**, Cooper, C., & **Rose, C. J**. (2023). The effect of machine learning tools for evidence synthesis on resource use and time-to-completion: protocol for a retrospective pilot study. Systematic reviews, 12(1), 7. https://doi.org/10.1186/s13643-023-02171-y	Protocol

August	**Rose, C**., Ringsten, M., Bidonde, J., Glanville, J., **Berg, R**., Cooper, C., Muller, A., **Bergsund, H**., **Meneses Echávez, J**., & Potrebny, T. (2023). Using a large language model (ChatGPT) to assess risk of bias in randomized controlled trials of medical interventions: protocol for a pilot study of interrater agreement with human reviewers. https://doi.org/10.21203/rs.3.rs-3288515/v1	Protocol
December	**Ames, H. N**., Noyes, J., and A. Booth (2023). Chapter 6: Selecting studies and sampling. Draft version. Cochrane-Campbell Handbook for Qualitative Evidence Synthesis, Version 1.0. Jane Noyes (Senior Editor) and Angela Harden (Senior Editor). London, Cochrane. https://training.cochrane.org/cochrane-campbell-handbook-qualitative-evidence-synthesis	Book chapter
In review	**Meneses-Echavez, J. F., Muller, A. E., Berg, R. C., Ames, H. M. R., Borge, T. C., Jardim, P. S. J**., Cooper, C., & **Rose, C. J**. The effect of machine learning tools for evidence synthesis on resource use and time-to-completion: a retrospective pilot study.	Paper
In progress	**Bergsund, H.B., Larun, L., Lidal, I., Poulsson, A., Borge, T., Jardim, P., Ames, H**. Tentative title: Developing a questionnaire to explore attitudes towards implementation of machine learning in a systematic review setting: a worked example.	Paper
In progress	Implementation guidance – Norwegian report	Report
In progress	Implementation guidance – English paper	Paper
In progress	Can using the Cochrane RCT classifier help speed up study selection in qualitative evidence syntheses? A retrospective evaluation	Paper

Evaluations

Complete

1. We evaluated whether the Cochrane RCT classifier could be used to identify qualitative studies. It was tested on 2828 included primary qualitative studies from a total of 102 QES's. The findings were presented as a poster at Cochrane Coloquium september 2023, and will also be published as a paper.
2. We have completed an evaluation comparing reviews that used versus did not use ML with respect to resource use and time-to-completion. The manuscript was submitted to Systematic Reviews on 6/10/2023 and is awaiting peer review. The protocol for the study is published (1).
3. In May we sent out a questionnaire to all employees in the cluster. The questionnaire was ment to assess the employees view on the ML functions they have used and the assistance they had received from the ML team and the ML resources. Some main findings summarized, based on open questions:
    a. **Challenges with the use of ML functions:**

Several respondents report limited knowledge and experience with ML functions, hindering independent setup and use. Some experience the complexity and limited application areas for certain functions such as clustering and classifiers. The time intervals between use create a need for repeated training. Changes over time in understanding and use of ML features, as well as challenges related to EPPI-Reviewer's usability, are also mentioned. Discussion of methodology and risk of bias is highlighted, including the need for harmonisation with risk of bias tools.

b. **Support needs and suggestions for improvements:**
Respondents value personalised support and guidance from ML team members. Availability and helpfulness of the ML team is positively rated. Materials and training resources work well, but some call for better identification of relevant materials. Discussions and conversations about ML functions should be increased. Resource persons with ML expertise, training and support for EPPI-Reviewer are emphasised as useful. A desire for more detailed guidance and early involvement of the ML team in projects is also mentioned.

c. **Improvement suggestions for easier use of ML functions:**
Respondents want more focus on experience sharing from others who have used ML functions. Increased user-friendliness and intuitiveness in the programmes and tools associated with ML are desired. Several suggest more training and information about new opportunities and functions, as well as clearer information about support contacts. The idea of super users and better understanding of areas of use for the municipal team (kommunelag) is emphasised, along with the need for regular use and practice with the tools to improve understanding and mastery.

Based on this feedback we identified areas where the team should put in efforts, which guided much of the work within the cluster the remaining year: Enhance and maintain knowledge, increase visibility and closer follow-up. To enhance and maintain knowledge, we have developed a new e-learning course. Additionally, a new ML learning week was arranged in November, spanning three full days, aiming to increase knowledge and familiarity with EPPI Reviewer and our most used ML functions. The e-learning modules were required pre-work before the ML learning week. Some sessions in the ML week were also open to all NIPH employees. To foster closer collaboration with EPPI, they held most of the sessions during the ML week.

To increase visibility and accessibility, we have fully renovated our ML and AI SharePoint site, with the aim of being a more intuitive and user-friendly "knowledge base". The SharePoint room houses relevant ML information, including reporting guidelines, help request forms, relevant literature as well as links to the e-learning course. This centralized resource aims to empower employees to find solutions independently, fostering a self-sufficient and informed workforce. To allow for closer follow-up, we have implemented more regular feedback rounds, both officially through surveys and unofficially through conversations with employees.

Ongoing

A pilot project to evaluate interrater agreement between human consensus risk of bias judgements and those made by a large language model (ChatGPT). A protocol has been

written and published as a preprint and submitted to BMC Medical Research Methodology for peer review. An international team has been formed, with researchers from Norway (NIPH, Western Norway University of Applied Sciences (HVL)), Sweden (University of Lund), and the United Kingdom (University of Bristol, York). The first phase of the work is underway, and a preprint of the protocol is published (2).

Innovation

All the innovation activities undertaken in 2023 were made possible due to external funding. The activities include development of a new and scalable e-learning course on our most used machine learning functions; development and implementation of a quantitative survey and conducting qualitative interviews, both assessing attitudes, barriers and experiences concerning the use of ML amongst employees. Additionally, we were able to perform capacity building activities both within the team and outside of the team.

Develop scalable e-learning.

Background: In March 2022, we piloted a one-week intensive training programme. It involved five members from the ML team and cluster management, and an average of 26 employees received 6.5 hours of training each. The pilot included digital workshops that built theoretical and practical knowledge. Despite the value of the training programme, several improvements were seen as needed:

a) increase transferability to our colleagues in other divisions and departments, where knowledge products are produced for different users and possibly with different processes.
b) compress the content so that it is feasible for new and existing employees in less time.
c) transform the way training is delivered so that it is scalable both in parts and as a whole, from live teaching to recorded videos, interactive training and tutorials, quizzes and independent tasks/exercises.

Creating an effective and engaging e-learning course involved careful consideration of various design aspects. Clear learning objectives were defined to ensure that the course content and activities aligned with employees' specific needs. Personalization features allow learners to showcase their skills and bypass content they already know, catering to novices by starting with foundational concepts and gradually progressing to advanced topics as confidence grows, while intermediate and advanced learners can concentrate on more advanced modules. Emphasis was placed on using plain and easy-to-understand language throughout the course. Assessments and feedback mechanisms are integrated into the course to help learners monitor their progress and pinpoint areas for improvement.

We developed e-learning modules on our most used ML functions: Ranking algorithms, classification algorithms and OpenAlex. The e-learning content goes through the conceptual aspects of these functions, to provide an understanding of what they are and how they work, as this is fundamental to understanding when it is appropriate to use them and how to use them correctly.

Development of survey

In May, four members of the team started developing a survey to assess attitudes, barriers and experiences concerning the use of ML. This would serve the dual purpose of both evaluating the results of capacity-building so far in the cluster, as well as informing the development of the implementation strategy. The process of developing the survey was carried out in four stages. First, an OpenAlex-based search of qualitative literature on ML experiences was carried out to inform the content of the questionnaire. Second, another OpenAlex-based search of studies describing relevant AI/ML attitude questionnaires was carried out, with the purpose of developing a pool of AI/ML-attitude questionnaire items. Third, items that corresponded to the most salient themes in the qualitative papers were drawn from the pool and used to develop our own AI/ML-attitude survey. Finally, a selection of our peers (members of the team, external evidence synthesis professionals and union representatives) were asked to review the survey and provide feedback (face validation). In response to the feedback, we created some additional items in areas that were not covered by the existing item pool and adjusted the way in which some of the items were phrased.

The survey has been distributed twice so far in the cluster, once before and once after the capacity building activities conducted outside the team in November.

Qualitative interviews by Comte Bureau

Comte Bureau was commissioned by NIPH to conduct 10 qualitative interviews with NIPH employees and managers to understand the attitudes, barriers, and experiences around the use of machine learning in the cluster. Prior to the interviews, three meetings were held in the autumn of 2023 to review the preparations for the interviews and lessons learnt from the parallel quantitative survey. The research questions from which the interview guide was developed were as follows:

1. What attitudes, barriers and experiences do FHl employees have to ML?
2. What personal barriers occur?
3. What structural barriers occur in relation to using ML in current projects in NIPH?
4. What does it take for NIPH employees to use ML in their everyday work?
5. What are NIPH employees' experiences of using ML with regard to
    a. How they have used ML so far
    b. How have they experienced work process changes associated with ML (if they have?)
    c. Perceived change in job satisfaction or engagement
6. How have employees experienced the capacity building activities that have occurred since September 2023 (e-learning course, ML week, change workshop by Mindshift at the cluster seminar in October)?
7. How have managers facilitated the use of ML?

The results from the interviews will be published in a report in December 2023.

Capacity building

Capacity building within the team

Scrum master course

In July, three members of the team participated in a Scrum master certification course. The aim with Scrum was originally to establish a more effective approach to software development, challenging traditional methods deemed dysfunctional, however the Scrum methodology can be applied across many different subject areas. Scrum is grounded in empirical process control theory, emphasizing a clear vision and learning through short iterations toward the goal. Work is carried out by small, self-organized teams, focusing on solving problems as they arise. The Scrum theory methodology emphasize transparency, inspection, and adaptation. The framework is minimalist, allowing Scrum teams to fill in details of their chosen methodology, fostering strong ownership and results. The three pillars of Scrum are transparency, inspection, and adaptation, ensuring visibility, frequent evaluation of artifacts and progress, and the ability to make necessary adjustments for goal attainment and improvement in both the product and the process. Overall, the Scrum methodology's emphasis on adaptability, collaboration, and continuous improvement makes it a valuable approach for teams working in dynamic and complex environments, where changes are frequent, and requirements evolve over time.

Digital learning institute – Digital design diploma

As the ML team's knowledge on e-learning development was very limited, three of the team members registered for A Professional Diploma in Digital Learning Design. The course content included learning about the foundations of instructional design, encompassing an understanding of learning theories, instructional design models, and the ability to analyse learner needs to formulate learning objectives. The program also explores digital learning technologies, including various tools and technologies utilized in online learning, and provides insights into Learning Management Systems (LMS) and content authoring tools. Content development is a key focus, involving the design and creation of engaging and interactive digital content, incorporating multimedia elements such as videos and interactive modules. There is focus on user-friendliness and accessibility, incorporating principles of design thinking into course development. Additionally, the program addresses adaptive learning and personalization techniques, including how to create personalized learning experiences and implement adaptive learning strategies based on learner needs. Effective communication strategies for online learning and fostering collaboration among learners in a digital environment are also integral components of the diploma program.

Capacity building outside of the team.

ML week 2023

In November three representatives from the EPPI centre held in-house training sessions on the EPPI tool and our most used ML functions. Forty-nine employees participated in an introductory workshop, where most participants (39 out of 49) were from other departments in FHI including: Norwegian Scientific Committee for Food and Environment (VKM), Global Health, Centre for Epidemic Interventions Research (CEIR), Chemical Toxicology, Air Quality and Noise, Norwegian Poison Information Centre, Library, Antibiotic Resistance and Infection

Prevention, Helsebiblioteket.no, Mental Health and Suicide, Food Safety, Childhood and Families and Physical Health and Ageing. A prerequisite to participate in the intermediate sessions on machine learning functions during ML week was to finish the e-learning course.

Change workshop by Mindshift at cluster seminar.

The use of ML and AI in our work processes inherently causes a change in how we work. Additionally, the cluster has been through two major reorganizations during 2022 and 2023 and we saw a need for increasing the employee's knowledge in handling change. Therefore, during the cluster seminar in October, we commissioned Mindshift to hold a workshop on change processes with a focus on digitalisation, where the main aim was to increase employees understanding of how to manage change as part of everyday working life. Part of the workshop was held separately for leaders and the remaining employees. The leaders' course covered topics such as the role of leaders in change processes, communication with employees, and guidance on preserving employee well-being. The employee course followed a similar structure, covering topics like self-management, handling changes in the work environment, and stress and stress management. Both courses were facilitated via practical tasks, plenary discussions, and feedback sessions.

Horizon scanning

Horizon scanning has not been prioritized this year due to limited resources. However, some activities have been carried out, and are listed in table 5.

Table 5: Horizon scanning and mapping activities during 2023.

Date	Description
February	Mapping machine learning functions in different screening tools (Cadima, EPPI Reviewer, Rayaan, Covidence, Distiller SR, Dextr). Published as an internal report in Norwegian.
March and September	Mapping and examine semi-automated data extraction tools (Dextr, RobotReviewer, Pitts)
Throughout the year	New tools with ML/AI functions that looks promising have been put on a list for future evaluations

External networking and collaboration

Below we list some of the groups and institutions with whom we presently engage in various capacities, ranging from networking to collaboration. Certain affiliations are characterized by networking connections and the exchange of expertise, while others involve more substantive collaborations, including contributing data for the advancement of machine learning functions, providing user input on tools and machine learning functions, and jointly undertaking initiatives such as strategy development and the planning of forthcoming evaluations.

Campbell Collaboration

Will Moy, CEO of the Campbell Collaboration, reached out to the ML team based on recommendations from James Thomas, as they, together with representatives from the UK Cabinet Office who work on evidence and evaluation, are exploring what it would take to support a step up in the use of ML in evidence synthesis. Based on this they wanted to have our expertise and experience on board. The Campbell Collaboration, together with EPPI centre and Future Evidence Foundation are putting together a proposal related to accelerating work on automation for Evidence Synthesis, based on the need for evidence synthesis needing to be faster, cheaper, and more widely available than it is now. This will include running evaluations that the group judges as contributing useful information regarding the performance of automation tools in evidence synthesis. The ML team has expressed interest in participating, but this project is still in its infancy, and the scope of the work which the ML team will contribute to is still unclear.

Cochrane Qualitative and Implementation Methods Group (QIMG)

HA sits as a co-convenor of the QIMG. As part of her role as co-convenor she has authored a chapter in the upcoming Cochrane-Campbell Handbook for Qualitative Evidence Synthesis on study selection and sampling where the use of machine learning is discussed. She is also an academic editor on the handbook. Together with Prof. James Thomas, she has responsibility for raising machine learning concepts with the QIMG when they are relevant or answering questions about potential use of ML in qualitative evidence synthesis.

EPPI Centre and National Institute for Health Care Excellence (NICE)

The study begun in late 2021 with NICE and EPPI Centre to improve the priority screening algorithms within the EPPI Reviewer software has been expanded to include experts from other European institutions. This collaborative study (k > 150 projects) is the largest simulation study of ML approaches with screening, and results will be used to suggest stopping criteria for screening, or when researchers can stop manual screening, as well as provide understandable metrics for researchers to evaluate algorithmic performance. Our role, and NICE's role, is to provide user input regarding the metrics and output of ML-assisted screening. Status: two algorithms have so far been analysed, with eight more remaining. We will provide EPPI with more datasets if needed for training and testing of the algorithms. Next steps are to maintain current collaboration, particularly with EPPI and in relation to the priority screening project and user test the stopping criteria once it is drafted.

INSIA Methods Working Group – AI Sub-group

NIPH is one of the member organizations in the International Network for Social Intervention Assessment (INSIA). In May, one of our team members (HB) was elected lead of an AI methods sub-group, which is currently made up of evidence synthesis professionals from the Swedish Agency for Health Technology Assessment and Assessment of Social Services (SBU), French National Authority for Health (HAS) and the Canadian National Institute of Excellence in Health and Social Services (INESSS). The group is working on a strategy for future projects, which will likely entail collaborations across the institutes on how to use AI in social intervention assessments.

International Collaboration for the Automation of Systematic Reviews (ICASR)

ICASR was launched with the aim of seamlessly integrating all components involved in automating the production of systematic reviews. The collaboration, including members from NIEHS (NIH), Cochrane, UCL, CREBP (Bond University), and others, focuses on principles such as efficiency improvement, automation across SR tasks, adherence to high standards, collaboration, open-source practices, and replicability. The first meeting in Vienna (October 2015) outlined these principles (3), laying the foundation for advancing automation in SR production.

The ML team held a presentation at the ICASR meeting June 2018 and has been involved with the group since then, However, owing to confusion surrounding the leadership transition within the ML team at the outset of the year, our involvement became dormant until September 2023. Subsequently, our participation was reinitiated following the Cochrane Colloquium. We have been invited to sit in a strategic planning group due to our knowledge and experience with implementation and have participated in to two meetings so far. We will continue our involvement to stay abreast of automation developments within the systematic review field and guide the direction for future meetings and conferences. Tasks might involve assisting with organizing and planning the future direction of ICASR and applying for EU funding for networking activities.

Julius Kühn-Institut (JKI)

We have established collaborations with JKI to share knowledge, resources and identify synergies. JKI has created their own systematic review software (CADIMA) and hired an AI researcher to further develop advanced, but user-friendly techniques, whereas NIPH relies on off-the-shelf products. We are both working towards the same goal, but from very different points of departure, and with different restraints and opportunities. JKI is continuously improving their software, and NIPH has provided them with data that is used as basis for development of semi-automated screening on both T/A and at full text level and using different classifiers for classification of references. They are also exploring possibilities for semi- automation of data extraction, and the ML team have provided input on our wishes for a data extraction function/tool to align with HTV needs for data extraction in our products, as well as providing data for them to test their developments on.

Robert Koch Institute

The Robert Koch Institute reached out to the ML team in May. They are the national public health institute in Germany, and they are exploring the possibilities of identifying existing machine learning tools for evidence synthesis and creating a workflow as part of a project they are starting on Public Health Impact Analyses. Their aim is to replicate a review that has already been done, with the help of the application of the tools, and to compare the results. In searching for institutes that have made similar efforts before, they came across the work conducted by the ML team and were very interested in exchanging ideas with us and in presenting their project. Currently we have had two meetings where they have shared their work and we have provided them with our experiences as well as input.

The Danish Center for Social Science Research (VIVE)

At the INSIA conference in May, one of the participating ML team members encountered a representative from VIVE, The Danish Center for Social Science Research. VIVE has a collaboration with Campbell, VIVE Campbell, which conducts systematic reviews and other high-quality reviews in the social sciences domain relevant to the Danish welfare system. The group have used some of the ML teams previous reports to inform their ML efforts within the evidence synthesis process. They have amongst other things an ongoing project where they are developing methods and programs for abstract screening using ChatGPT and intend to use the method and program in conjunction with the priority screening function in EPPI Reviewer for a complex screening process in an ongoing review project. The current aims for our collaboration are to exchange knowledge and experiences and to learn from each other, regarding ongoing projects and potential future project collaborations.

Lessons learned in 2023

What has worked well in 2023?
External funding has been a notable strength, providing crucial support for the team's initiatives. The collaboration with EPPI to address technical aspects of the software has been instrumental, offering valuable insights into implementation and a deeper technical understanding of the ML tools available. The external funding allowed team members to attend conferences which not only facilitated exposure to diverse perspectives but has also fostered networking opportunities, enabling the team to stay abreast of advancements in the field.

The interdisciplinary nature of the team has been a key success factor, with the inclusion of a librarian proving particularly valuable. Consistent weekly meetings have played a pivotal role in maintaining team cohesion, preventing potential drift. Furthermore, the team's intrinsic motivation for ML has been a driving force, propelling the group forward.

Challenges the team has faced in 2023.
On the flip side, certain challenges have emerged. The absence of protected time outside of DFØ has posed difficulties, hindering the team's ability to fully engage with ML advancements. Keeping up with the rapid developments in the AI/ML domain has proven challenging due to capacity constraints and we feel we have not been able to keep abreast of important developments in the field.

Ensuring a comprehensive understanding of ML usage across all teams and identifying areas requiring optimization and further learning has also been tasked prioritised due to time restrictions. The absence of a rotating membership structure has been felt as a limitation, as time constraints have impeded the inclusion of additional members in the team this year due to the amount of time needed to onboard new members unfamiliar with ML.

The presence of numerous uncertainties and unclear directives around the team's role and objectives has been a recurring issue, particularly in the initial months of the year when the team felt in a state of limbo or as if we were sitting on hold, negatively impacting both production and motivation. This underscores the importance of clearer expectations and guidance from leadership forming a clear mandate for the team’s work.

Suggested ML strategy from 2024

ML Team 4.0 as part of a division portfolio

In response to the evolving needs of our institute and in alignment with strategic objectives of both the institute and the division, this proposal advocates for the elevation of the ML team to the Division level. Such a move is strategically positioned to enhance collaboration and ensure sustainability of ML efforts in the evidence synthesis processes in the Division and the opportunity to expand to other research divisions and methods. Our rationale encompasses a multitude of factors that underscore the necessity of this organizational transition, which are presented below:

1. **Facilitating cross-departmental collaborations**: Elevating the ML team to Division level presents better opportunities for collaborations across departments and divisions, fostering a more integrated and interdisciplinary approach to our work.
2. **Ensuring team sustainability:** Operating at Division level secures the long-term sustainability of the ML team, by a) increasing the pool of potential team members, ensuring a diverse skill set and a broader range of expertise; and b) reducing vulnerability to employee turnover as there will be a larger pool of potential team members , safeguarding the continuity of core team members and ML/AI knowledge within the ML team. The team will also be more robust against future reorganizations.
3. **Financial benefits for the cluster:** Elevation decreases the financial burden for the cluster, as the cost for ML team members will be divided across departments involved with the ML team.
4. **Testing ML tools on primary research**: Being at Division level better enables the evaluation and testing of ML tools on primary research considering the majority of the primary research conducted in the Division is not based in HTV. Also, our existing ML expertise in evidence synthesis can help inform the work on primary research.
5. **Utilizing EPPI Reviewer across products and teams:** The transition facilitates increased use of the EPPI Reviewer across products and teams, improving the efficiency of evidence synthesis processes across the division.
6. **Capacity building and resource allocation:** Overall resources within the cluster will be less strained, as the pool of potential team members increases to all employees working within the division.
7. **Piloting interdisciplinary collaboration:** This move also provides an opportunity to pilot closer interdisciplinary collaboration ahead of potential new reorganizations.

8. **Global leadership position:** To maintain the ML teams position as a global leader in ML implementation, it is imperative to expand the pool of talent, ensuring the team's motivation is driven by innovation rather than mere fulfillment of requirements.
9. **Visibility and cross institutional collaboration opportunities**: Elevation enables easier implementation of ML across the institute and as well as externally through collaboration with other public health institutes.

Alignment with NIPH and Division Strategy

We have focused on aligning our strategy with the NIPH strategy (4). The NIPH strategy emphasizes the crucial role of evolving infrastructure in supporting knowledge and preparedness in the health sector, as well as societal and economic development, our institute recognizes the need for sustainable solutions to meet new opportunities and challenges. Additionally, the strategy emphasizes the ever-increasing demand for new knowledge as a foundational element for decision-making across all levels in health and care services.

Concurrently, within the Division strategy (5), four distinct focus areas are delineated: Knowledge support for municipalities, knowledge about interventions, future data basis for knowledge, and collaboration to strengthen public health institutions and systems in other countries. The Division strategy outlines specific efforts required to achieve these targets, underscoring the importance of autonomous employees who proactively contribute to the management of their divisions. This involves taking the initiative to explore new avenues, proposing innovative projects, and securing external funding. Furthermore, the strategy highlights the necessity for expertise in project management and cutting-edge methods, including machine learning and automation.

This proposal resonates with the institute's strategy of evolving infrastructure to meet new challenges and supports the increasing need for knowledge in health decision-making. It also aligns with the Division strategy, emphasizing autonomous employees, expertise in project management, and the integration of new methods such as machine learning.

Elevating the ML team to Division level is a strategic imperative that aligns seamlessly with both the institute and Division strategies. This move positions us to be at the forefront of innovation and collaboration in the rapidly evolving landscape of evidence synthesis, technology, and public health.

ML Team 4.0 organization

By elevating the ML team efforts to Division level, we not only ensure the sustainability of the team but also contribute significantly to the overarching goals of our division and institute. To further facilitate the divisions ML and AI work, we propose a restructured framework for Team 4.0 with a clearer division of tasks and responsibilities, driven by two primary factors. Firstly, the rapid growth and expansion of the field of ML and AI make it impractical for a single team to adequately perform horizon scanning, evaluation, and implementation. This has currently led to a strain on the team's resources, hindering the execution of activities at a satisfactory level. The second is the opportunity to bring in team members with specific competencies and/or particular interest within evaluation, implementation, or teaching.

These new team members will bring the skills and motivation we need to continue our work and receive support and onboarding from the ML team.

The proposed structure involves dividing the existing team into three distinct teams, each with a team lead, and establishing a steering committee to coordinate activities. The teams and their respective responsibilities are outlined as follows:

1. Innovation and horizon scanning
This team will consist of a team lead and 1-2 team members. They will have responsibility for horizon scanning for new tools within ML and AI, mapping these tools, doing a preliminary test, and then deciding which tools should be evaluated. They should also be responsible for mapping already published evaluations for new relevant tools. They will also have responsibility for the OpenAlex automated search for new research on ML and AI that is relevant to our evidence synthesis work.

Effectively conducting horizon scanning and innovation requires individuals to possess a combination of essential qualities. These include being organized, having good searching skills and a forward-thinking vision and curiosity to anticipate and explore emerging trends. Furthermore, they should have an understanding of ML and AI and potential ways it can be used in our work, and clear communication skills.

2. Evaluation and evidence building
This team will consist of a team lead and 3-4 team members (dependant on the number of ongoing evaluations). When the innovation and horizon scanning team identifies a new tool or function that they feel would be a good fit for our work, the evaluation and evidence building team will plan and conduct an evaluation of the new tool, if we haven't already identified an evaluation already conducted by another evidence synthesis group. This team will ensure that all new tools are beneficial to our work by saving time, increasing productivity, or helping with complex problems. They will also ensure that the new tool maintains or improves the quality of our products. This team will feed back to the innovation and horizon scanning team about any specific tools or needs they feel the team should be trying to identify.

If an evaluation shows that a tool should be implemented into our work processes, the evaluation lead will present the tool to the implementation and support team and have a discussion around implications for implementation and the way forward.

Team members need to have a good understanding of ML and AI functions and how they work as well as our internal work processes to plan and conduct evaluations. It is a clear advantage if some team members have a background in analysis/statistics and experience in planning and conducting evaluations.

3. Implementation and support
This team will consist of a team lead and 4-5 team members broken into two working groups. This team will feedback user needs to the two other teams. The first working group of 2-3 team members will have responsibility for implementation and peer-to-peer support. The

second working group of 2 people will have responsibility for addressing how using ML/AI changes our workflow processes and implications for project management.

The implementation and user support working groups will develop new teaching materials and implementation strategies in communication with the evaluation lead when a new tool is identified for widespread implementation. The team will create an implementation plan and identify what type of teaching materials and user support are needed. The team will also give peer-to-peer support to project teams on request.

The work processes evaluation and implementation work group will focus on evaluating changes related to workflow and assessing implications for implementation related to these changes. They will oversee implementing any work process changes that are identified as being beneficial through evaluations.

Team members need to have a good understanding of ML and AI functions and how they work as well as good teaching and communication skills. A further benefit would be experience with implementation and evaluation of implementation projects and/or qualitative research. Team members should have good communication skills, be patient and have an ability to speak with hesitant or sceptical colleagues without being judgemental.

Team lead steering committee

The three team leads will sit in a steering committee where they will coordinate activities and discuss tools that are ready of evaluation and implementation. This group will also have responsibility for networking with external actors and sitting in working groups linked to project partnerships in, for example, ICASR. A fourth member of the team lead steering committee will provide administrative coordination. This role will entail having an overview over what is happening, setting up meetings, coordinating funding applications, sending out the newsletter and keeping the website and SharePoint room up to date.

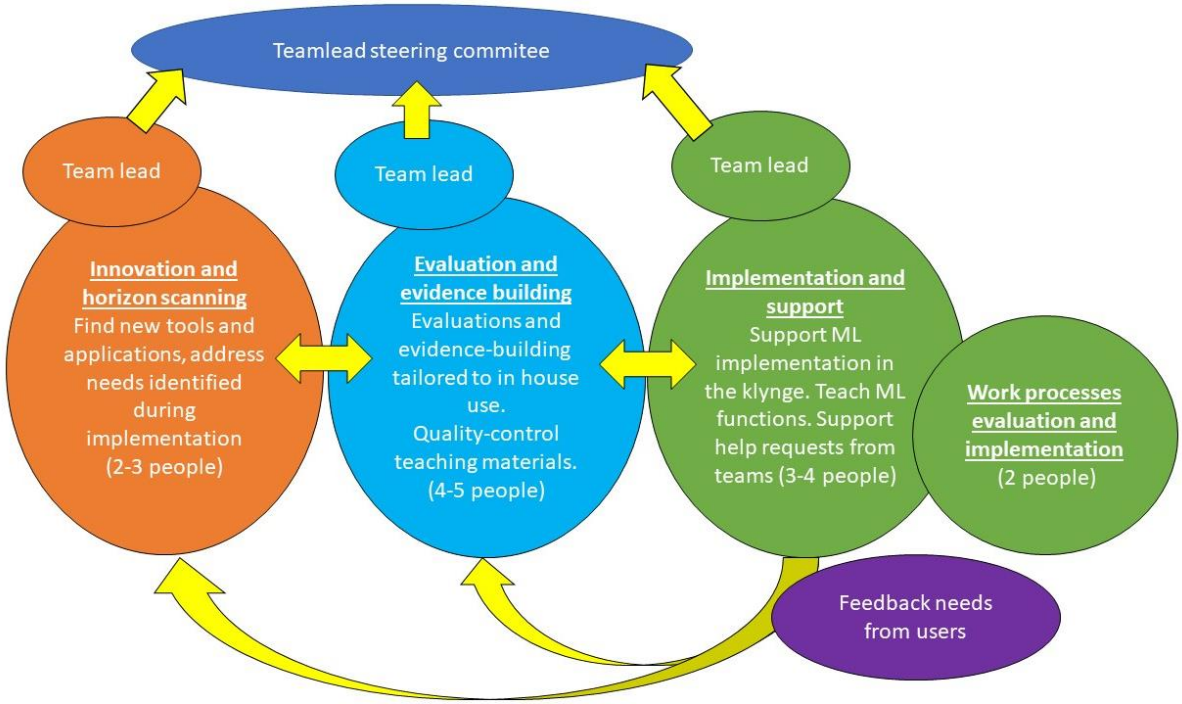

Figure 1: Proposed structure for team 4.0

Suggested focus areas for 2024

In the following section are some suggestions for possible focus areas for the ML team in 2024. These focus areas align well with both the Division and institute strategies, as well as supporting the team's overarching goals and contributing to its positioning as a leader in ML implementation within the evidence synthesis field.

1. **Seek funding from external and international sources:** This aligns with the institute strategy, emphasizing the need for sustainable solutions and investments. Seeking external funding contributes to the financial sustainability of the team and is in line with the overarching goals of the Division to collaborate and strengthen public health institutions globally. Potential relevant Norwegian funding sources will be the Norwegian Research Councils upcoming research efforts in artificial intelligence and digital technologies which will be increased by at least NOK 1 billion over the next five years, aimed to contribute to greater insight into the consequences of technological development for society. Possible international funding opportunities include the EU Networking grant (ICASR collaboration) and the UK better methods grant (Joint application with the EPPI Centre).
2. **Apply for DOI numbers for our reports:** This aligns with the divisions strategy for open science. Implementing the use of DOI numbers will allow machine learning based search and scrape to find the evidence we produce as it is not indexed in databases.
3. **Facilitate machine readability for open data sharing:** Engage with other systematic review groups such as EBM on FIHR (https://ebm.bmj.com/content/24/Suppl_1/A15.1) to begin working towards standardized formatting and language for systematic reviews to make open data sharing possible. Machine readability enhances the accessibility and

usability of the reports, contributing to the broader goal of knowledge support for municipalities and interventions.

4. **Increased use of OpenAlex for staying up to date in different topic areas**: Using OpenAlex to create libraries of relevant literature for different projects or topic areas. These libraries would be housed in EPPI Reviewer where research teams could code and categorise them for easy sifting. This aligns with the Division and institute strategies for knowledge development and staying current in relevant areas.
5. **Continued external collaboration with EPPI and other key institutions:** This aligns with the Division strategy's emphasis on collaboration to strengthen public health institutions and fosters interdisciplinary collaboration. Also, this will increase visibility of both the ML work being conducted in the Division and well as secure against unnecessary double work. Also, assisting EPPI with evaluating and testing various tasks puts us in a prime position of being a key user organization that can provide input to EPPI which will benefit our institution, with regards to improved interface and functionality of the EPPI tool. Evaluations that EPPI are interrested in involving us in are related to:
    a. Semi-automated data extraction and mapping of data
    b. New clustering functions with language models
    c. Stopping criteria for priority screening- testing not only the different criteria but testing the usability of the function.
    d. Vector database work
    e. Custom classifiers
6. **Strengthen interdisciplinary collaboration within NIPH.** During 2023 we have seen an increased interrest in using ML In evidence synthesis from groups at NIPH outside of our cluster. We have provided peer-to peer support for groups in the of Global Health cluster, the Norwegian Scientific Committee for Food and Environment and Division of Infection Control. Increased networking across NIPH divisions fosters interdisciplinary collaboration and aids in the dissemination of our work.
7. **Capacity building on Generative AI:** This aligns with the Division strategy's call for expertise in new methods, including machine learning. Capacity building on Generative AI ensures the team remains at the forefront of innovative methodologies and stays on top of the explosive development of Generative AI.
8. **Develop a complete ML implementation package for other institutions**: We have during 2023 made a scalable e-learning course on the conceptual parts of our most used machine learning functions. We want to continue developing the e-learning course as well as adding technical how-to components, to be able to provide a complete training package that can be delivered to other groups in-house as well as outside of FHI and to collaborating public health institutes globally. This aligns with the Division strategy's emphasis on collaboration to strengthen public health institutions by sharing expertise. This has also the potential of creating revenue.
9. **Further explore and evaluate tools related to data extraction**: This aligns with the Division and institute strategies for adopting new methods, contributing to efficiency and advancements in evidence synthesis, as well as staying at the forefront of innovation. Tools of particular relevance might be:
    a. **RobotReviewer**: This tool has been dormant for a while, but work will start up again. As this is a tool we have evaluated some parts of in previous work in 2022 (6), it will be very useful to continue evaluations of this tool

    b. **Generative AI tools**: Evaluations of data extraction tools using generative AI can be particularly relevant, as there are many tools available, and due to the nature of the tools and their AI functionality, many aspects of data extraction can be explored, e.g. PICO extraction, risk of bias evaluations for many different study designs as well as extraction of results.
10. **Explore possibilities of EU cost funding for networking through ICASR:** Seeking cost funding for networking aligns with the argument of seeking external funding and the divisions strategy for collaboration. This can enhance external collaboration efforts, contributing to shared goals and knowledge exchange.
11. **Explore possibilities with using ML/AI for the institutes registry work:** Registries will be a key focus Division for the institute the coming years due to the centralization of the health registries to NIPH.Exploring how AI/ML can be used in this work should be considered.

References

1. Muller AE, Berg RC, Meneses-Echavez JF, Ames HMR, Borge TC, Jardim PSJ, et al. The effect of machine learning tools for evidence synthesis on resource use and time-to-completion: protocol for a retrospective pilot study. Systematic Reviews 2023;12(1):7. DOI: 10.1186/s13643-023-02171-y
2. Rose C, Ringsten M, Bidonde J, Glanville J, Berg R, Cooper C, et al. Using a large language model (ChatGPT) to assess risk of bias in randomized controlled trials of medical interventions: protocol for a pilot study of interrater agreement with human reviewers 2023.
3. Beller E, Clark J, Tsafnat G, Adams C, Diehl H, Lund H, et al. Making progress with the automation of systematic reviews: principles of the International Collaboration for the Automation of Systematic Reviews (ICASR). Systematic Reviews 2018;7(1):77. DOI: 10.1186/s13643-018-0740-7
4. Folkehelseinstituttet. Strategi mot 2024 - Om Folkehelseinstituttets strategi – ti satsinger. 2023. Available from: https://www.fhi.no/om/folkehelseinstituttets-strategi/om-strategien/innledning-til-de-ti-satsingene/?term=
5. Område for Helsetjenester. Område for helsetjenester 1-5 år framover. Oslo: Folkehelseinstituttet; 2022.
6. Jardim PSJ, Rose CJ, Ames HM, Echavez JFM, Van de Velde S, Muller AE. Automating risk of bias assessment in systematic reviews: a real-time mixed methods comparison of human researchers to a machine learning system. BMC Med Res Methodol 2022;22(1):167. DOI: 10.1186/s12874-022-01649-y