

Large Language Models and business applications in an R&D environment

Fredrik Forsén

Master Thesis

Industrial Management and Engineering

Vaasa 2024

DEGREE THESIS

Author: Fredrik Forsén

Degree Program and place of study: Industrial Management and Engineering, Vaasa

Supervisor: Mats Braskén

Title: Large Language Models and Business Applications in R&D environment

Date: 6.6.2024 Number of pages: 58 Appendices: 2

Abstract

The subject of this Thesis is usage of Generative AI and in particular Large Language Models (LLMs) within the R&D department of Mirka Power Tools. The research problem is to explore if R&D departments can improve productivity and innovation with the help of LLMs. The purpose of the Thesis is to find bottlenecks in the R&D work and make the identified bottleneck work processes more efficient, by speeding up time consuming work tasks or automated routine work. The methods used for this explorative research was conducted by 1) literature studies mentioned in the theory chapter and 2) a questionnaire consisting of eight AI and work-related questions that was sent out to the whole R&D department. The results showed that LLMs have potential to aid employees in various work tasks and make them work more efficiently, but it is important to be aware of the related risks (eventual false answers, biased answers or risks by using confidential data). The research gave a clearer fundamental understanding of the basics and limitations of LLMs and raised interest in further research. The survey gave valuable insight into today's work bottlenecks, and where the employees see the value of using LLMs for various applications. Respondents were 20 out of 64 people. The focus of LLM tool used in this survey is Microsoft Copilot. My recommendation is to use Copilot as an assistant and inspiration source in a wide spectra of R&D related process tasks and general office work - but think critically of what information you give and what you get.

Language: English

Key Words: Artificial Intelligence, Generative AI, Large Language Models, Copilot, Machine Learning, AI tools, LLM Business Applications, R&D, Neural Networks

EXAMENSARBETE

Författare: Fredrik Forsén

Utbildning och ort: Industrial Management and Engineering, Vasa

Handledare: Mats Braskén

Titel: Large Language Models and Business Applications in R&D environment

Datum: 6.6.2024

Sidantal: 58

Bilagor: 2

Abstrakt

Ämnet för denna avhandling handlar om användningen av generativ AI och särskilt stora språkmodeller (LLM) inom R&D-avdelningen vid Mirka Power Tools.

Forskningsproblemet är att utforska om R&D-avdelningar kan förbättra sin produktivitet och innovationsförmåga med hjälp av LLMs. Syftet med avhandlingen är att identifiera flaskhalsar i R&D-arbetet och göra de identifierade arbetsprocesserna mer effektiva genom att snabba upp tidskrävande arbetsuppgifter eller genom automation av rutinarbete. De metoder som användes för denna explorativa forskning inkluderade 1) litteraturstudier som nämns i teorikapitlet och 2) en enkät bestående av åtta frågor om AI och arbete som skickades ut till hela R&D-avdelningen. Resultaten visade att LLM har potential att hjälpa anställda med olika arbetsuppgifter och göra dem mer effektiva, men det är viktigt att vara medveten om de relaterade riskerna (eventuella felaktiga svar, snedvridna svar eller risker med användning av konfidentiell data). Forskningen gav en tydligare grundläggande förståelse för LLM:s grunder och begränsningar och väckte intresse för ytterligare forskning. Enkäten gav värdefulla insikter i några av arbetets flaskhalsar och var LLM kan användas för olika ändamål. Respondenterna var 20 av 64 personer. Huvudfokus för LLM-verktyget som användes i denna undersökning är Microsoft Copilot. Min rekommendation är att använda Copilot som en assistent och inspirationskälla i en bred skala av R&D-relaterade arbetsuppgifter och allmänt kontorsarbete - men tänk kritiskt på vilken information du ger och får från LLM-verktygen.

Språk: Svenska

Nyckelord: Artificiell intelligens, Generativ AI, Stora språkmodeller, Copilot, Maskininlärning, AI verktyg, språkmodell användningsområden, Forskning och Utveckling, Neurala nätverk

OPINNÄYTETYÖ

Tekijä: Fredrik Forsén

Koulutus ja paikkakunta: Industrial Management and Engineering, Vaasa

Ohjaaja: Mats Braskén

Nimike: Large Language Models and Business Applications in R&D environment

Päivämäärä: 6.6.2024

Sivumäärä: 58

Liitteet: 2

Tiivistelmä

Tämä tutkielma käsittelee generatiivisen tekoälyn ja erityisesti suurten kielimallien (LLM) hyödyntämistä Mirka Power Tools tutkimus- ja kehitysosastolla. Tutkimusongelmana on selvittää, voivatko tutkimus- ja kehitysosastot parantaa tuottavuutta ja innovaatiota LLM:ien avulla. Tämän tutkielman tavoitteena on tunnistaa pullonkauloja tutkimus- ja kehitystyössä ja tehostaa tunnistettuja työprosesseja nopeuttamalla aikaa vieviä työtehtäviä tai automatisoimalla rutiinivaihtelua. Tätä eksplanatiivista tutkimusta varten käytettiin metodeja: 1) teoriaosiossa mainitut kirjallisuustutkimukset ja 2) koko tutkimus- ja kehitysosastolle lähetetty kyselylomake, joka sisälsi kahdeksan tekoälyn ja työskentelyyn liittyvää kysymystä. Tulokset osoittivat, että LLM-malleilla on potentiaalia auttaa työntekijöitä erilaisissa työtehtävissä ja tekevät heidän työstään tehokkaampaa, mutta on tärkeää olla tietoinen niihin liittyvistä riskeistä (mahdolliset väärät vastaukset, puolueelliset vastaukset tai riskit luottamuksellisen tiedon käytöstä). Tutkimus antoi selkeämmän perustavan ymmärtämyksen LLM-mallien perusteista ja rajoituksista ja herätti kiinnostusta jatkotutkimuksiin. Kyselytutkimus antoi arvokasta tietoa joistain tämän päivän työn pullonkauloista ja siitä, missä vastaajat näkevät LLM-mallien käytön arvon eri sovelluksissa. Vastaajia oli 20 kaikkiaan 64 henkilöstä. Tässä kyselyssä käytetyn LLM-työkalun pääasiallinen kohde on Microsoft Copilot. Suositukseni on käyttää Copilot avustajana ja inspiraation lähteenä laajalla spektrillä tutkimus- ja kehitystyöhön liittyvissä prosessitehtävissä ja yleisissä toimistotoissa - mutta suhtautua kriittisesti siihen, mitä tietoa annat ja mitä saat.

Kieli: Suomi

Avainsanat: Tekoäly, Generatiivinen tekoäly, Suuret Kielimallit (LLM), Koneoppiminen (ML), AI työkalut, LLM liiketoiminta sovellukset, Tutkimus & kehitys

Table of Contents

1	INTRODUCTION	6
1.1	Background	7
1.2	Generative AI and Large Language Models	8
1.3	Research Problem	10
1.4	Research questions	11
1.5	Purpose of the Thesis	11
1.6	Limitations	12
2	COMPANY PRESENTATION- MIRKA POWER TOOLS R&D DEPARTMENT	13
2.1	R&D organization	13
2.2	R&D office work tasks	15
3	THEORY	17
3.1	What is artificial intelligence (AI)?	17
3.2	What is Machine Learning (ML)?	18
3.3	What are Artificial Neural Networks (ANN)?	20
3.4	What is Deep Learning (DL)?	21
3.5	What is Generative AI?	22
3.6	What are Large Language Models (LLM)?	23
3.7	How do LLMs work?	24
3.8	Fine tuning of LLMs	25
3.9	Examples of LLMs	26
3.10	Risks and limitations	27
3.11	What is prompt engineering?	29
4	RESEARCH METHODOLOGY	31
4.1	Research method	31
4.2	Data collecting method	32
4.2.1	Ethical considerations in data collection	32
4.3	Data analysis method	32
4.4	Result presentation method	32
4.5	Limitations in the research approach	32
5	LLM APPLICATIONS IN THE R&D PROCESS	34
6	AI SURVEY	37
6.1	Survey Questions	38
6.2	Respondents	38
7	RESULTS	39
7.1	Survey Results	39
8	Discussion of survey results	47

		5
9	CONCLUSION.....	50
10	DISCUSSION	51
11	FUTURE	54
12	BIBLIOGRAPHY	54
13	APPENDICES.....	59
13.1	Appendix 1-LLM training.....	59
13.2	Appendix 2-Transformers- a brief overview	60

1 INTRODUCTION



Fig.1 An AI generated image that symbolizes the orchestration of LLMs

“In the age of artificial intelligence, those who master the language of machines hold the key to unlocking unparalleled creativity, efficiency, and innovation. Just as a skilled conductor orchestrates an intricate symphony, learning to wield Large Language Models (LLMs) empowers individuals and companies to compose their own success. The lexicon of progress lies not only in words but in the algorithms that shape our future. Embrace the syntax, decode the possibilities, and watch your competitive edge crescendo.”

Source: (Microsoft, u.d.) when prompted to write a quotation about using LLMs in R&D work and creating a suitable image, which is shown in fig.1)

This thesis is written for and in co-operation with Mirka Power Tools Research and Development department.

The purpose of the thesis is to find out if and how the employees in the department can benefit in their work from the usage of Large Language Models (LLMs).

1.1 Background

The field of Artificial Intelligence (AI) has developed rapidly over the latest years driven by advancements in machine learning algorithms and the availability of massive amounts of data. Consequently, this has led to the development of a vast range of AI tools with unique capabilities and potential abilities across various sectors. Two main AI types are discriminative AI and generative AI. Discriminative AI *categorizes* (e.g., image recognition, narrow AI) and generative AI *creates* content (images, text, audio, code).

One major milestone in the AI development was the introduction of Transformer by Google in 2017, which is a neural network architecture that enabled the development of more capable Large Language Models (LLM). The LLM ChatGPT was released by OpenAI in November 2022 (OpenAI, u.d.), which quickly showed its potential in understanding and generating human like text, such as drafting stories, help in coding computer programs etc., while understanding context, irony, and humor. It proved clearly that machines indeed can “learn” the complexities of human language (Marr, 2023). Examples of what these tools can do include generating context aware text, automating routine tasks, generating creative content, and assisting in complex decision-making. AI tools are changing the way we work and live. As shown below in Fig.2, it quickly reached one hundred million users fast.

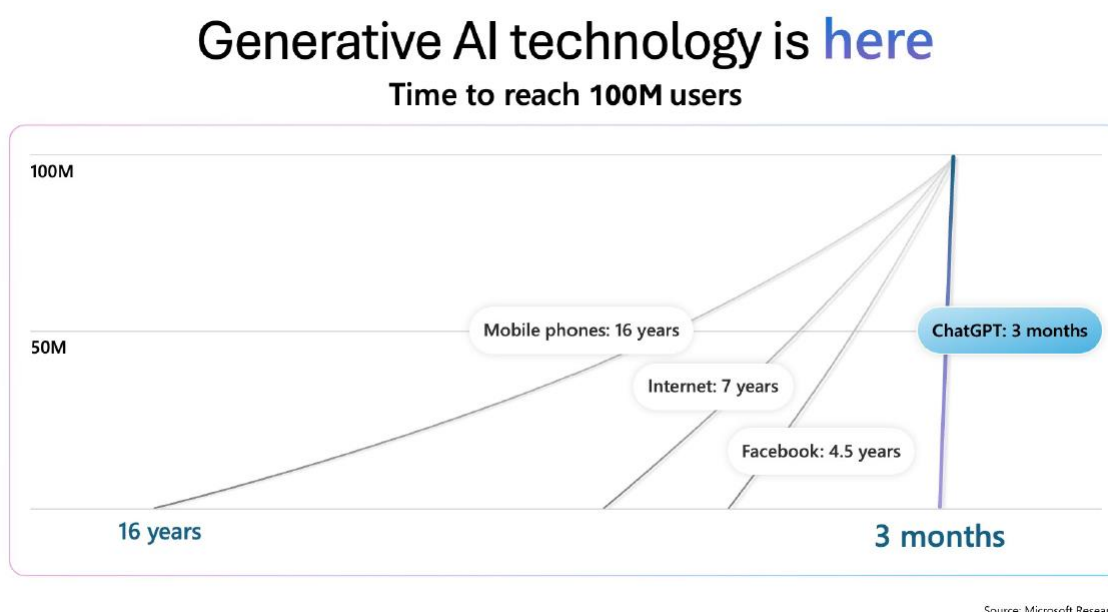


Fig.2 Generative AI and its rapid development hints about its significance.

“Change is the law of life. And those who look only to the past or present are certain to miss the future” -John F. Kennedy (Ekblom, 2015)

In a competitive business environment, organizations want to be more effective, creative, customer-focused, profitable, and successful in their work and performance. From an R&D department perspective, it is about making great innovations and products that create value for customers, stand out from competitors and drive the performance of the company. To achieve the desired goals and being competitive, continuous process improvements and new methods of work and thinking is essential. Since time, money and resources always are limited, it may be wise to reflect on what we do, why we do it and ask ourselves if there are more clever ways to organize the work towards desired goals such as customer satisfaction, great innovations, process optimization, economic KPIs and faster product development that more precisely meet market demands. And it is here where the exploration of AI tools in the context of R&D department work gets interesting.

From a leadership and management perspective, it is essential to understand the changing business environment and adapt to the new situation in a constructive way.

1.2 Generative AI and Large Language Models

While this thesis is primarily focused on one segment of the AI field, Large Language Models (LLM), it is still good to have some overview of the whole AI landscape.

Below follows a brief overview of some of the diverse AI applications that exist today (and that evolve by the day).

Tasks:

Natural Language Processing (NLP) tasks: analyze and understand human language, which includes semantics (meaning of the text) sentiment analysis (analyzing emotional tone, e.g. customer feedback), text summarization and machine translation (context-aware and not only word-to-word translation) (IBM, u.d.).

Computer Vision (CV) tasks: By recognizing patterns in the visual data and use those patterns to determine content in other images (Microsoft, u.d.).

- object recognition.
- image classification: E.g. (Google, u.d.).

- anomaly detection.

Tools:

Document Intelligence (DI) tools: Analyze several types of file documents such as pdfs, Word or Excel. For example, Copilot may summarize or extract key points out of pdfs. Or Copilot in Word that can do the same. Other examples are cloud-based AI services like Microsoft Azure where you can create intelligent document processing solutions of invoices, tax documents, insurance documents and so on (Microsoft, u.d.). SciSpace (Scispace, u.d.) is another tool that helps you understand research reports easier by simplifying academic complex text and extract relevant information from documents, or Perplexity.ai (Perplexity, u.d.) which is a powerful information search engine (powered by GPT-4 and Claude 2) with source citations, that also have a freeware version that can analyze pdfs, for example. These tools combine AI tools of NLP (find relevant information) and CV (understand pictures of e.g. graphs or mathematic formulas) to extract and understand information from the documents. Copilot also may work well if you add a picture of an mathematic formula, electrical diagram or similar in the prompt field and you want an explanation of what it is and how it works (utilizing Dall-E 3 feature) (Microsoft, u.d.). Another powerful tool is Wolfram Alpha that is a computational intelligence engine that can assist and compute answers within math, science, engineering, finance and so on (Wolfram Alpha, u.d.).

Automation tools: Automate repetitive tasks and processes, freeing up human resources for other activities (Microsoft, u.d.).

Predictive tools: Analyze data, find patterns, predict future trends to assist in decision making. Examples of tools for this is for example Microsoft Azure Machine Learning (Microsoft, u.d.).

Generative tools: A certain type of AI that uses Machine Learning to create several types of new content (Lawton, u.d.). Examples include:

- Text Generation: Large Language Models (LLM) such as ChatGPT, Google Gemini, Microsoft Copilot).
- Image Generation: Dall-E, Google Image FX or hotpot.ai (Google, u.d.).
- Sound generation: E.g. Suno (Suno, u.d.).
- Video generation: InVideo AI (make short videos from a text prompt).

Knowledge management tools: Utilize AI to organize and share insights from data within an organization (IBM, u.d.).

Chatbots and virtual assistants: Automating customer service interactions or versions trained on company data that may assist employees (Rebelo, 2023).

Personalization & recommendation tools: Tools that analyze user behavior and recommend relevant products, services, or content. Personalized marketing.

Classification Tools: AI tools that sort data into predefined categories, for example a spam filter.

While this diverse list of AI tasks and tools hints about exciting possibilities for almost any business field, this thesis will focus on the potential and applications of Large Language Models (LLMs) in R&D departments. The future is created now, and the big question is: “how can we utilize the new technology in the best possible way”? While lot of possibilities seem to appear, it is also essential to evaluate AI from a critical point of view and identify risks and limitations.

While AI development seems to develop exponentially by day, the learning curve of organizations tends to grow at a slower pace. To better recognize AI opportunities within the company, education about what AI tools can do and cannot is an important first step, to become more successful as individuals and as a company.

1.3 Research Problem

R&D work is a vague term, but in my definition includes many common offices work tasks such as email reading and writing, Excel related work, PowerPoint presentations, meetings, information search and so on. What is unique with R&D work compared to more traditional office work is the entire process from knowing your customers’ needs and behaviors and providing solutions and products that meets the requirements and specifications set from both the customer and the own organization. This often means idea generation, idea evaluation, meetings, testing, product development, define specifications, prototype making, problem solving, brainstorming, 3D modeling, data analysis, decision making, electronics and software development, simulations, supplier and customer contact, IP (Intellectual Property) related tasks and so on.

As an initial hypothesis and standing point regarding this explorative research work, I have a belief that LLMs can make R&D work more efficient (reduce work hours in order to achieve same amount of work) and making work of better quality (more input and perspectives from both colleagues and LLMs tend to give better ideas and results) in a lot of ways, if used properly and being aware of related risks.

The research problem for this Thesis can be summarized in the following question:

Can we improve productivity and innovation in the R&D department with the help of LLMs?

To find good answers to this rather wide question, we will break it down a bit, and ask us some more specific research questions, which will be presented in the next chapter.

1.4 Research questions

1. What are the current bottlenecks in office R&D work?
2. How can generative AI tools and LLMs be used to resolve these bottlenecks?
3. What LLM tools are currently available and what are their strengths, weaknesses, and limitations?

1.5 Purpose of the Thesis

The main purpose of the thesis is to make the identified bottleneck work processes more efficient, which means speeding up time consuming work tasks or automating routine work.

Other purposes are:

- Find out bottlenecks in today's way of work (find work processes to focus on and develop)
- Learn more about LLMs (=basic research in order to gain knowledge) and potential applications in R&D work where it could be useful (applied research in order to develop work processes) (Patel, 2011).
- Explore LLM applications (what is relevant to use).
- Give suggestions of how to work smarter and more efficiently with the help of LLMs.

1.6 Limitations

- Focus on R&D department tasks only
- Mainly from own work perspective
- LLM related AI solutions and ideas of how to make work more efficient and of better quality
- A basic overview of how the technology works but no exploration of every technical detail

2 COMPANY PRESENTATION- MIRKA POWER TOOLS R&D DEPARTMENT

Mirka Power Tools have since 2010 developed and manufactured advanced power tools for sanding and polishing in cooperation with external partners. The development has been rapid in terms of sales increase and personnel, as seen in Fig.3. The more the organization, product portfolio and sales grow, the need for good data structure and information sharing also increases. This puts bigger pressure on the organization to make sure that it has clear work and documentation processes and ways to organize, share and find relevant information.

The focus areas in product development have been on making excellent products that provide real customer value, for example the focus has been on dust free solutions, ergonomic and cordless tools.

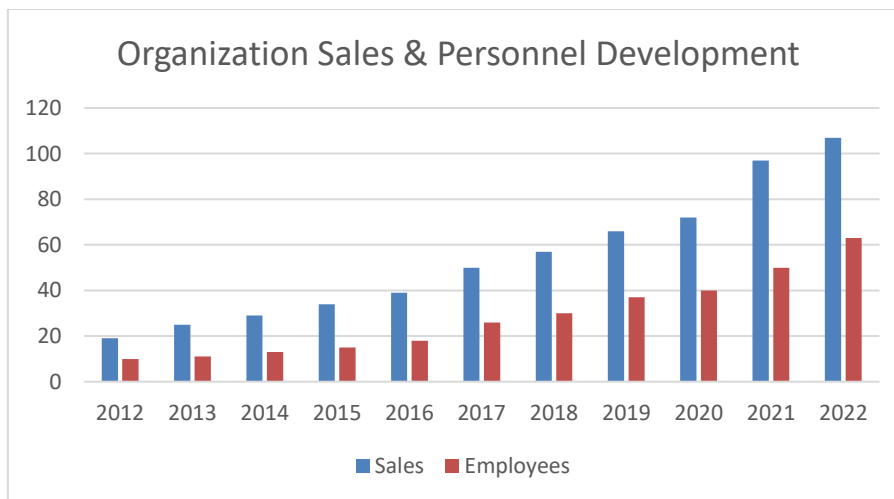


Fig.3 The sales and workforce has seen a steady growth in the past years

2.1 R&D organization

The R&D organization is divided into several sub departments that each handle various work tasks. The subdepartments are to various degrees dependent on each other and communicate and cooperate daily. All parts play a significant role in creating great products (that helps our customers solve their problems in the best way), improving products and processes, hence creating a solid foundation for the business to grow and prosper further. The R&D organization within the company today consist of following sub departments, as can be seen in Fig.4: R&D Product development (including Mechanical, Electrical and Software Engineers), Project Management, Compliance, Product Life Cycle, Test Department,

Quality and After Sales. The total workforce in the department is approximately 64 people, and approximately similar amount of personnel in production.

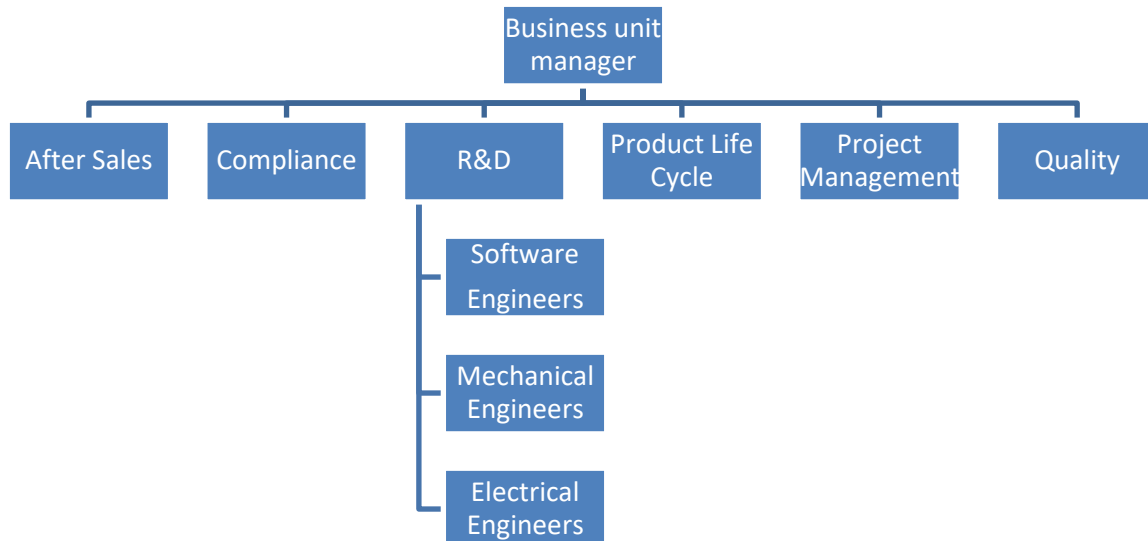


Fig.4 Organization overview

The R&D organization, or business unit, cooperates with other departments as well, that is not a formal part of the R&D organization and is not shown in the overview above. Some of these cooperation departments include Product Managers, Sector managers, Portfolio managers and Technical Specialists. While not part of the formal R&D organization, these mentioned roles also influence the R&D work. Business sector managers and product managers play a significant role in knowing the customers and passing the information forward to the R&D organization. It may often (but not always) be a Product or Portfolio Manager that request a certain product to be developed (after a business case has been found and has been approved on higher level) and speaks to the R&D Project Management Team, that initiate the project in cooperation with the R&D department.

Also important to mention is the company management on highest level strategic decisions that pour down through the whole Mirka organization and gives the vision of who we are and what direction and goals we want to aim for and what the focus areas in the R&D department are and sets limitations in budgets and workforce. Two trends and focus areas now are the green deal and the transition towards a more environmentally responsible future, and cordless tools.

2.2 R&D office work tasks

Knowledge workers (which R&D personnel can be seen as) typically combine multidisciplinary cognitive skills (analytic, reasoning, creative, communicative, and so on) with knowledge to make important decisions, develop new products and services, and find solutions to address business challenges. According to a survey from Forbes, the average knowledge worker spends 10 hours per week searching through tools and internal data systems for answers, which is about 25% of a typical workweek. Another challenge besides finding the correct information is to gain data insights from the often-unstructured data and communicate it to others. (Patenall, 2022)

According to a survey from Accenture, natural language tasks account for 62% of office workers time, and 65% of that could be made more productive by using LLMs to boost productivity (Accenture, 2023).

R&D office work consists of a diverse range of both general office work tasks and more specific R&D work tasks. In this thesis it is investigated how well LLMs can help with both these types of work task categories. Analyzing what these work tasks are in detail is not the main purpose of this thesis. However, general office work tasks such as email writing with suppliers and colleagues, email reading, Teams meetings, physical meetings, making & summarizing meeting notes, making to-do lists, reading, and understanding documents, information search, translations, documentation, PowerPoints, and Excel work is something that many face and do in their daily work. Some other R&D tasks depending on role may include brainstorming, problem solving, idea generation, idea evaluation, domain specific computer programs (3D modeling programs, electronics and software programs etc.), specification definition, prototype creation, IP related tasks, testing, data analysis, technology research, education, and finding parts and suppliers.

The goal for the R&D department is to create new products that solves customers problems (e.g. an efficient, ergonomic, dust free sander, that saves the customer time and money, preserves health and gives overall customer value) and fulfils the desired technical specifications (which sometimes are to be found out) and getting a product that hopefully exceeds the customers' expectations and is a gamechanger on the market and is in line with the company vision. In this case, the work tasks include more specific tasks, like finding out customer challenges and what is the best way to solve specific problems and how the tool could be designed in an optimal way, considering user needs, regulations, costs, competitor landscape and other aspects. Sometimes it is considered what engine or fan to use, what

material to use, what software algorithm to use, or thoughts about innovative new concepts and mixing different ideas to create something new, or designing ergonomic and esthetic looks, for example.

Below in Fig.5 is a graph with own-made examples of some of the work tasks related to “core R&D workers” (mechanical, electrical and software developers). This example is no absolute truth and may exclude some important work tasks, but it may be helpful as a start to a discussion regarding work tasks and what the R&D department generally puts most time into. Someone in the organization puts more time into meetings, and emails, for example. Analyzing actual time on different work tasks on whole R&D department level or personal level is left for someone else to analyze in more detail.

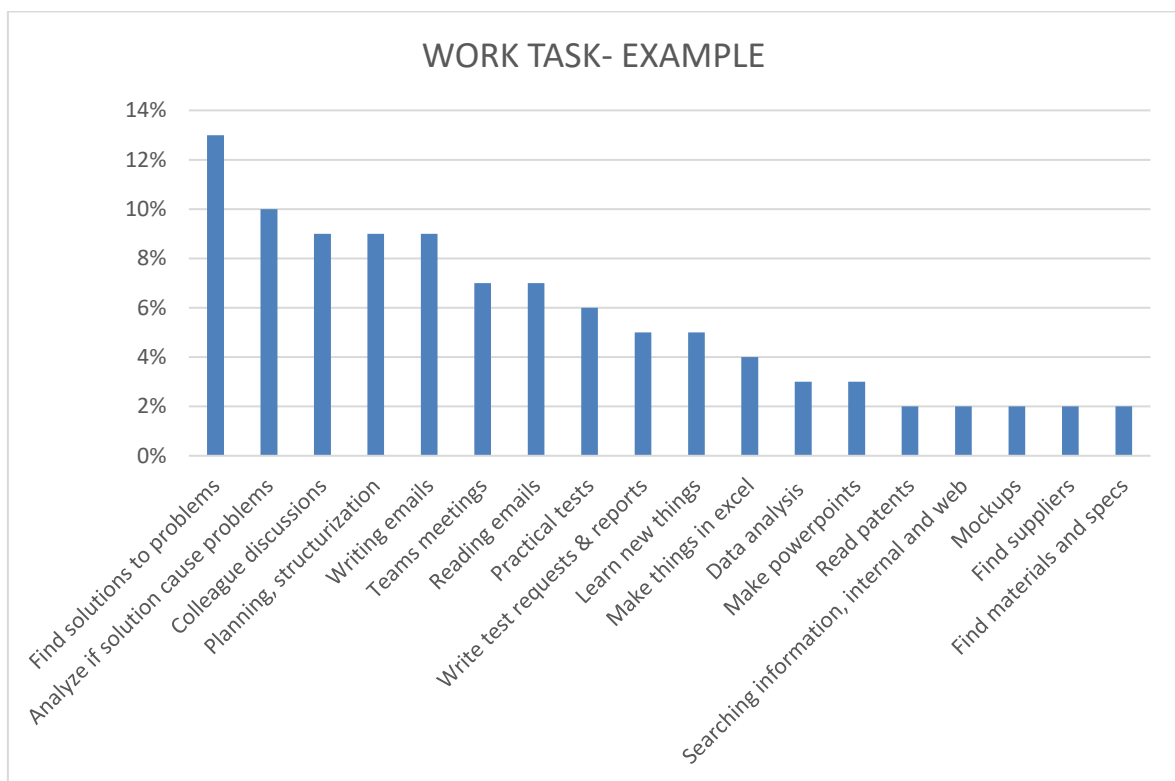


Fig.5 An imaginary own-made example of work tasks for R&D developers. This may be helpful when discussing work task topics.

3 THEORY

In the following chapter I will present some of the most relevant concepts regarding AI and LLMs. Knowing something about how the LLM technology works and its strengths, weaknesses, possibilities, and risks is essential when deciding how and what LLM tools to use and for what purpose. For more clarity, I decided to make a separate “Application Findings” chapter, but that also can be seen as a part of the theoretical framework.

3.1 What is artificial intelligence (AI)?

“Artificial intelligence, deep learning, machine learning—whatever you are doing if you don’t understand it—learn it. Because otherwise, you are going to be a dinosaur within 3 years.”- Mark Cuban

LLMs are AI models that are based on deep neural networks and are trained for the task of generating new text. Remembering Fig.6 may be helpful when considering how central AI terms relate to each other. I will later in the chapter present these terms in some more detail.

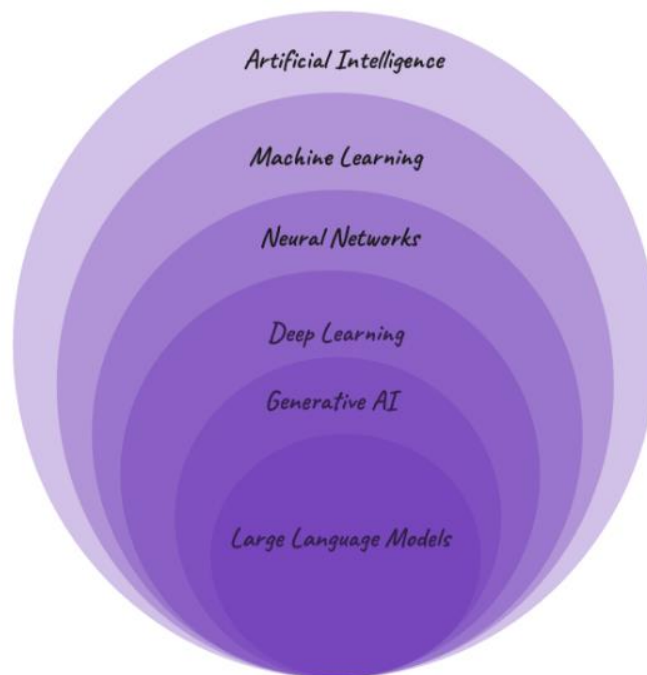


Fig.6 An illustration of how central AI terms relate to each other.

Defining Artificial Intelligence is not so simple. First, you should define what you mean with “Intelligence”. If we try to judge whether another human is intelligent, how do we do

it? For example, we might consider whether the person is good at solving problems, performs according to instructions, learns from past experiences, has lot of knowledge, thinks quickly, is insightful, is a good communicator and natural language capability, has good social skills, creates new ideas, judge his behavior, measure the score in an IQ test etc. In other words, “intelligence” is not an absolute term, and can mean a lot of things, depending on a person’s subjective thoughts and understanding of the term and the contextual situation. Considering the above-mentioned characteristics of intelligence is helpful when we argue whether a machine is “intelligent” or not (Rahman, 2020).

One definition of AI related to product development is the following: *Artificial intelligence is a disruptive technology that aims to reproduce human-like intelligence. Nowadays, companies utilize it as an assistant in each stage of the new product development to save time and resources* (Lemaire, 2023).

A quick introduction to central elements of AI is also shown in Fig.7. Usually, when we speak of AI today, we often refer to a type of machine learning (ML) called supervised learning using neural network models. Some more information of what that is, is described in the machine learning chapter. AI is the ability of computers to use ML to learn from data and use the learnt model to make human-like decisions.

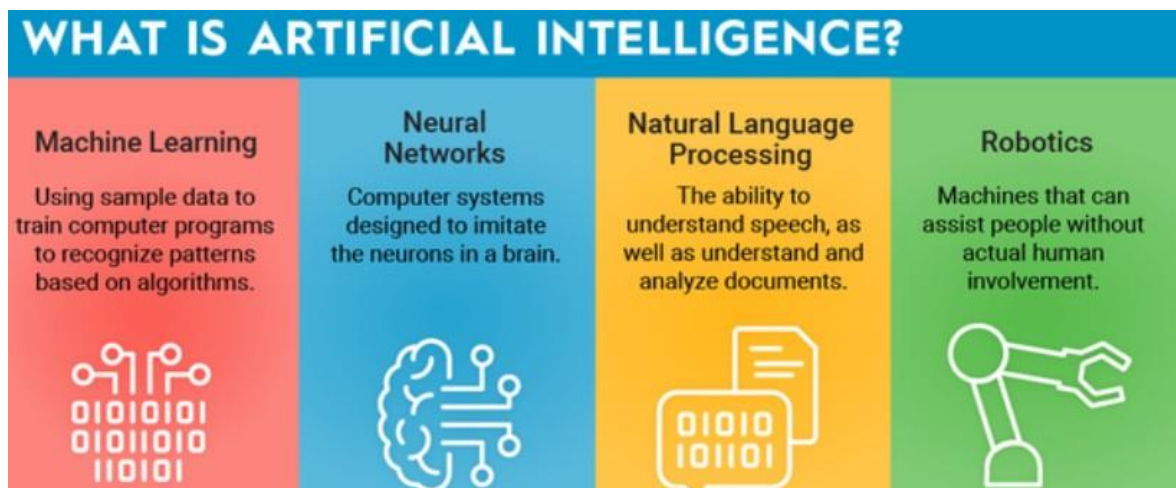


Fig.7 Central elements of AI (Motley Fool, 2024).

3.2 What is Machine Learning (ML)?

Machine Learning is a subset of AI that can solve tasks that are not possible to solve with traditional programming. It is a helpful way to understand data, by training on relevant data. As seen in fig.8 it can be used for several types of applications. It is used in a variety of tasks,

for example NLP, CV, speech recognition, email filtering, agriculture, and medicine. One example of machine learning is a filter that can recognize spam in your email box (by recognizing patterns of the language in spam emails). Image classification is another common machine learning example (identify and classify quality defects, categorization Cat VS Dog and so on). Another more R&D related example may be to use ML for analyzing for example motor sound on a tool and train the model to identify bearing failures or other quality defects, for example.

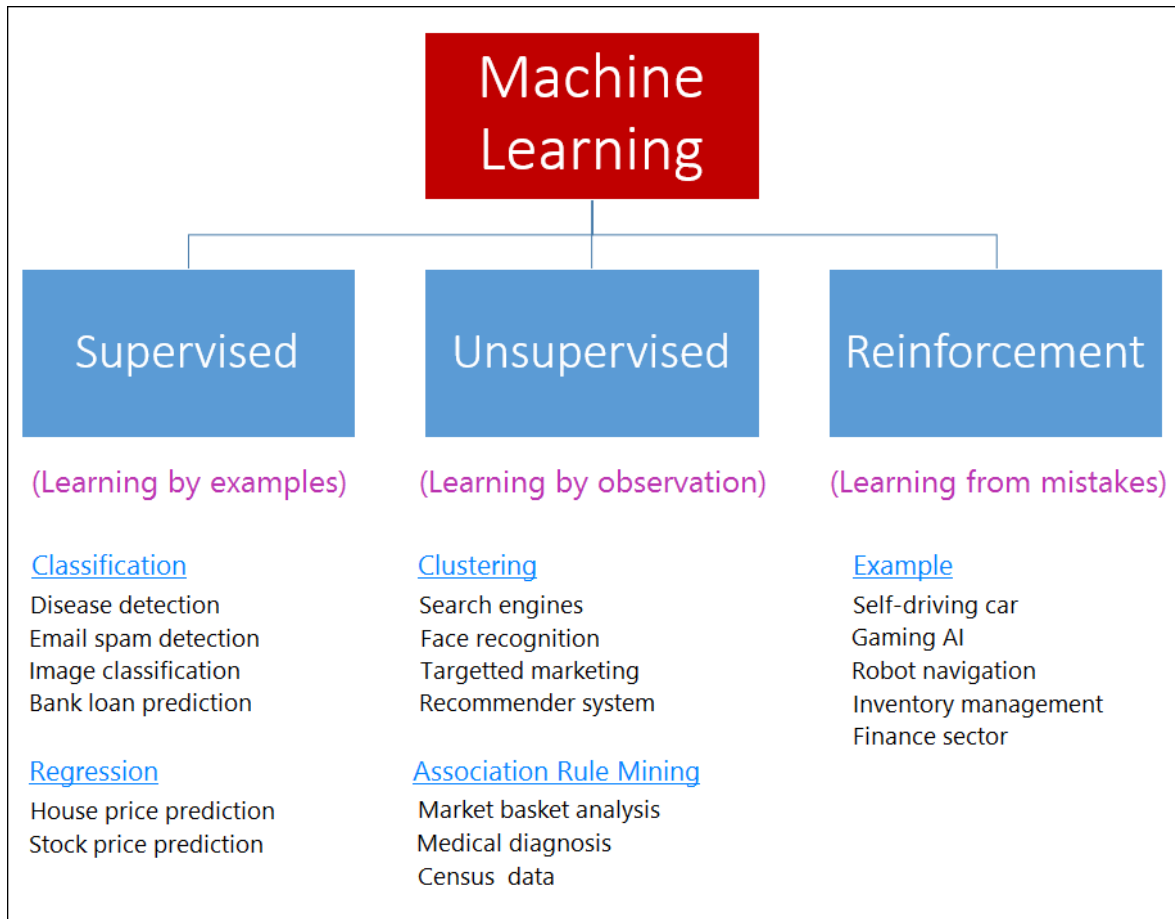


Fig.8 The three most common types of Machine learning (Techguruspeaks, u.d.).

Another example could be if we want to know how much an apartment is worth, we could make a database of other apartments and their market price, location, size, room numbers, age, and other relevant parameters. This is called labelled data and is a form of supervised learning, which is the most common machine learning method. Machine learning recognizes the pattern, or rules, in the training data, and can be based on that make a prediction for your apartment's market price.

Machine learning algorithms can be classified into a few categories (the two main categories are shown in fig.9):

- **Classifiers** (image, spam, fraud)
- **Regression** illustrating a linear relationship between variables (stock price, house price)
- **Clustering** (unsupervised classifiers)

(Campeato, 2020)

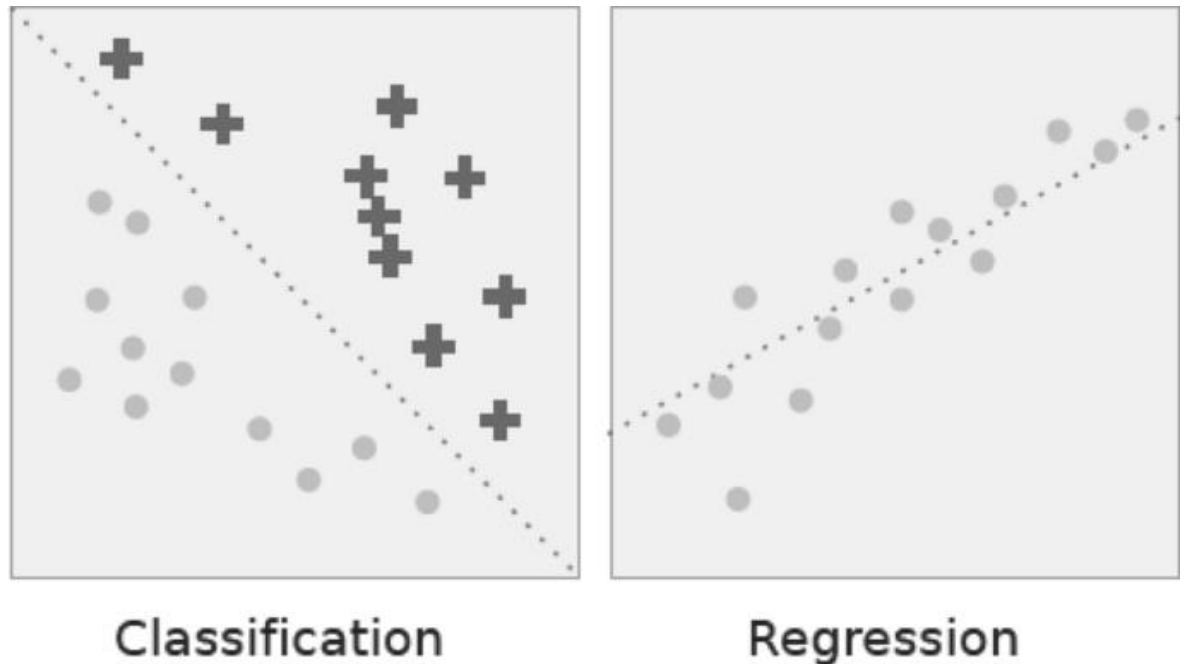


Fig.9 Classification VS regression. Classification categorizes, and regression is the linear relationship between two variables.

For more information regarding ML, check for example the following article (Sarker, 2021).

3.3 What are Artificial Neural Networks (ANN)?

Neural networks (NN) are a class of machine learning models inspired from the structure and function of the human brain (but still significantly differs from that). These models consist of connected nodes or “neurons”, that are organized as input, hidden and output layers. The nodes represent features of the data. For example, in an image recognition system, each node may represent a pixel value of the image. NNs are used for various tasks, including pattern recognition, speech recognition, natural language processing, image recognition and more (Oludare Isaac Abiodun, 2018). NNs are useful in many areas due to their ability to learn from data, and are often used for handling complex problems that do not have a known mathematic formula that describes the relation between input and output

variables. NNs are good at finding patterns in data and making predictions based on the training data. If we for example want to predict football results between two clubs, and we have a lot of different historical data that describes the [input data] (Home/away game, year, players who attend, win streak etc.) and [output data], (game result) we can build a NN model that predicts game results with very good accuracy-if the training data is of good quality and relevant. And that means that the data do need to have the actual relationships between variables, without a lot of excessive noise or faulty data, because we cannot expect to get out better data than we feed the NN with (Siljerud, 2023).

NNs learn and improve their accuracy by updating the network architecture and connection weights between nodes and layers, with help of backpropagation, so that the network can make better predictions, as illustrated in Fig. 10 (Minimize difference between prediction and the true value) (Mohaiminul Islam, 2019). This knowledge about how NNs learn to interpret data is fascinating and opens possibilities to automation of systems and to understand how complex systems that are dependent on many variables behave.

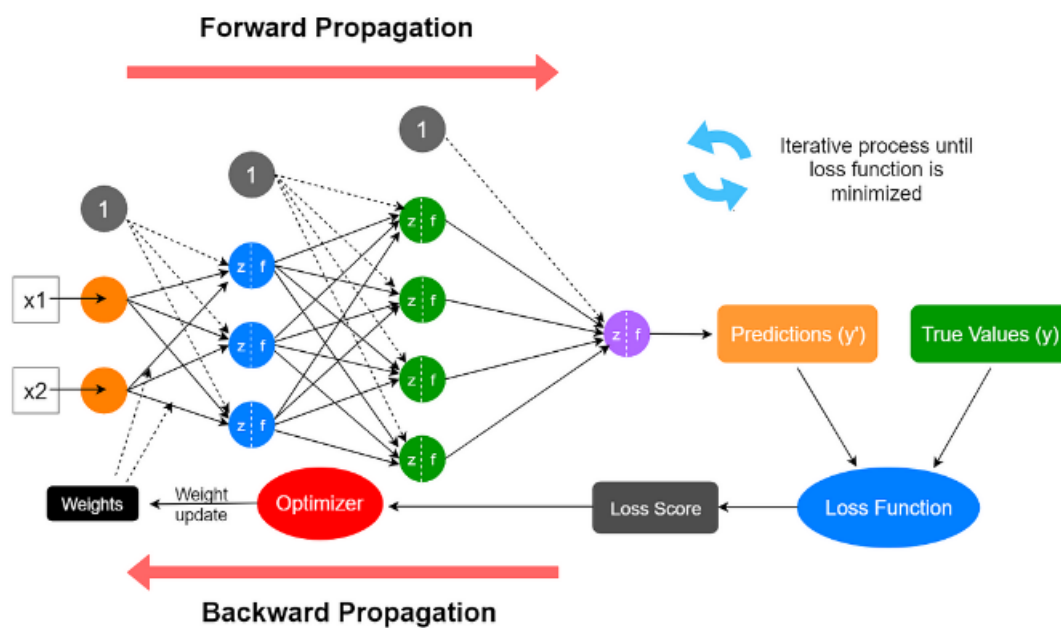


Fig.10 Example of a neural network

3.4 What is Deep Learning (DL)?

Deep Learning is a software technique and subset of Machine Learning that focuses on Neural Networks and algorithms for training Neural Networks. A Deep Learning model requires at least two hidden layers in a Neural Network. The more hidden layers a neural

network has, the *deeper* the learning model is. The layers capture features of the data. The neural network in Fig.11 has three hidden layers (Campeato, 2020). ChatGPT 3.5 has 96 layers and 175 billion parameters and is one of the largest deep learning models ever created (Knapp, 2023). It is not officially announced how many hidden layers ChatGPT 4 has. The term “deep learning” became popular back in 2010. By using big datasets and powerful computers running neural networks on Graphic Processing Units (GPUs), deep learning improved computers abilities to recognize images, process audio and play games (Tojin, Finkenstadt, & Folk, 2023).

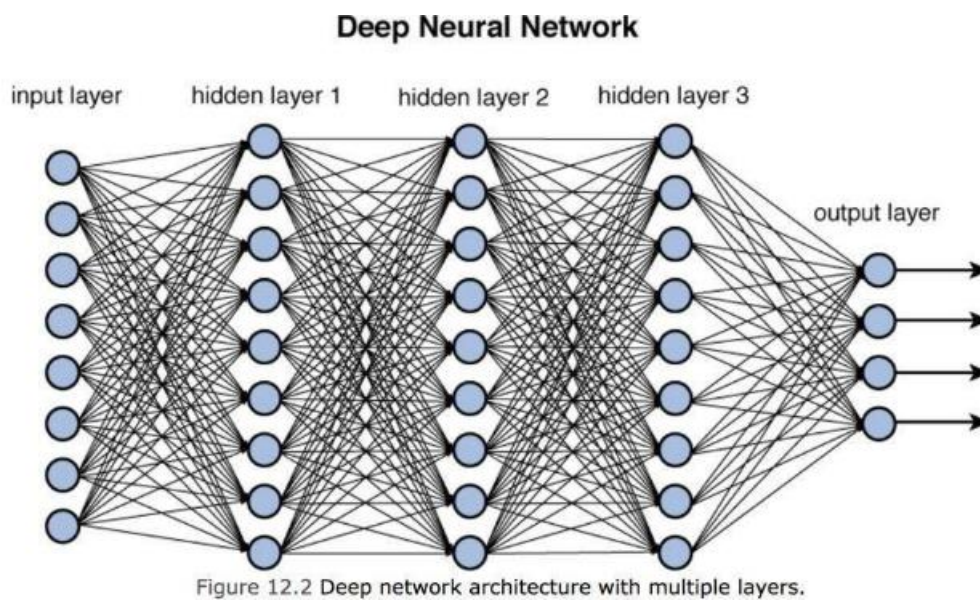


Fig.11 Illustration of a deep neural network with three hidden layers

3.5 What is Generative AI?

Generative AI is a pre-trained machine learning (ML) model that is used to generate new data, for example text, images, code, video, and sound, as seen in fig.12. One example of this is the Dall-E image generation tool that is trained on a lot of pictures. Large Language models such as ChatGPT also are included within the generative AI field.

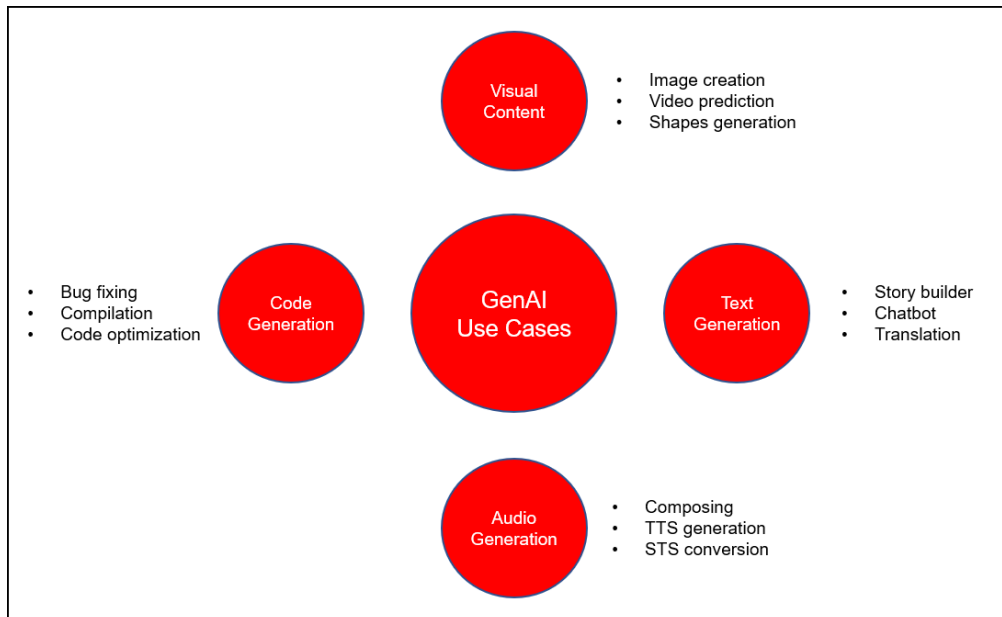


Fig.12 Generative AI use cases today. Source: (Lenovopress, u.d.)

Generative AI also includes an element of generating random answers that you do not have full control over. And even if you write the same prompt or instruction to ChatGPT twice, it is not likely you will get the exact same answer both times.

3.6 What are Large Language Models (LLM)?

A Large Language Model (LLM) is a probabilistic AI model of natural language. It can understand and generate human-like text. LLMs get these abilities by using massive amounts of training data such as books, articles, and web content to learn billions of parameters during training. During this training, the LLM learns patterns in the language and connections between words and general sentence structure. LLMs work by taking an input text and repeatedly predicts the next token or word (Wikipedia, u.d.). LLMs utilize deep learning, and teaches itself the “language rules and grammar” of the training data. No hard coding is needed of how a word should be spelled or how a sentence is structured in a specific context. The more training data of excellent quality, the better the quality of the text result will get. After the LLM model is trained (which requires considerable number of computational resources), they can be used for generation of new text content.

Tech companies like OpenAI, Microsoft, Alphabet, Amazon and Nvidia all have trained their own LLMs, and given them names like ChatGPT, Microsoft Copilot, Google Gemini, PaLM, Megatron, etc. The most famous LLM (or neural network) is ChatGPT that stands for Generative Pre-Trained Transformer. A transformer is a neural network architecture

that enabled the development of more capable Large Language Models (LLM) and was founded by google in 2017.

3.7 How do LLMs work?

LLMs might feel like magic when testing ChatGPT or Copilot for instance. But LLMs are experts in statistics and predictions. By training LLMs on massive amounts of text data, the model learns the grammar and associations between the words and sentences and fine tune the “weights” in the neural network. In Fig.13 below is an illustration of how LLMs work.

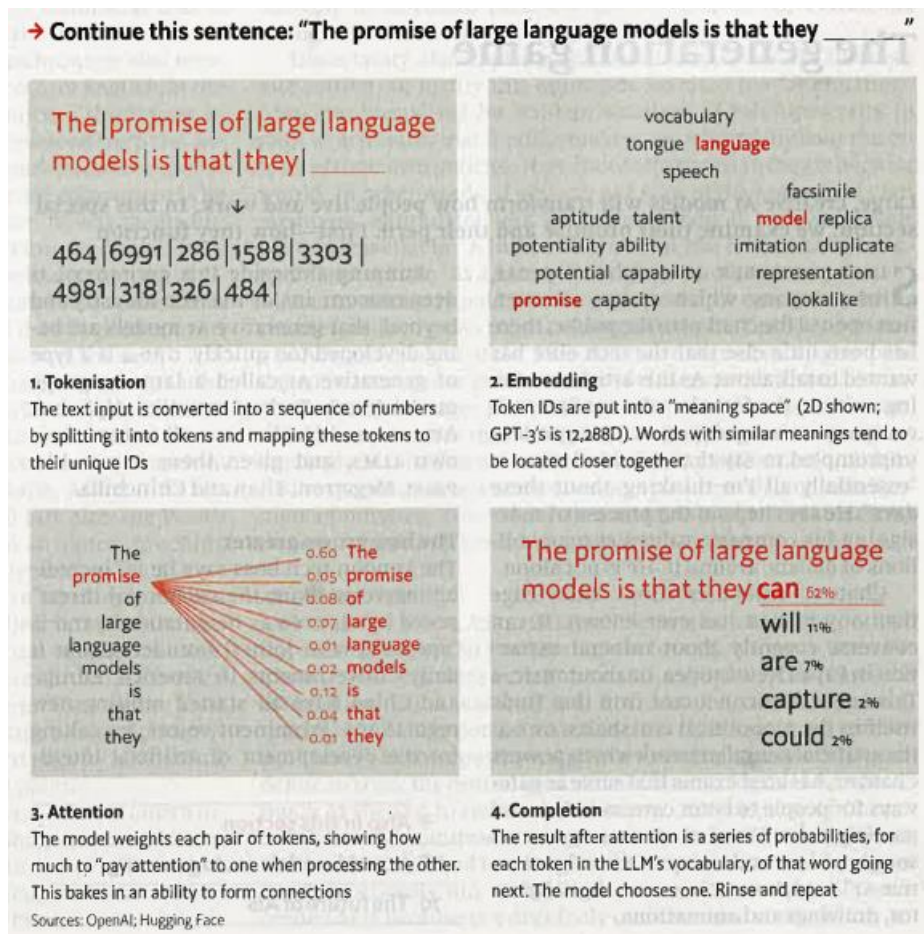


Fig13. Step-by-step explanation of how LLMs predict the next token or word. Source: (Tojin, Finkenstadt, & Folk, 2023).

First, the LLM Convert words to a format that neural networks can manage, which is a set of numbers. The text is divided into sets of characters called “tokens” that commonly occur together. GPT-4 can process a maximum of 32000 tokens, which is about the length of a novel. The longer the input-the higher the computing power is. The transformer architecture was introduced in June 2017 by Google and was a major milestone in the AI and LLM

development. A transformer is a neural network architecture designed for Natural language processing. (Hugging face, u.d.) for some more details, check appendix 2.

Traditional search engines like Google have a big drawback in the sense that they produce generic, non-personalized results which require the user to go through a lot of web pages before you eventually find the information you need. AI-based search engines, like Microsoft Copilot (previous Bing chat), Perplexity AI, etc. can understand the intent and context behind a search query and can provide more relevant, precise, concise and personalized results, based on user behavior and preferences with citations (Khan, 2023). For example Copilot (that is based on ChatGPT 4) goes on the internet and may find real life data, and can make for example charts of your chosen data. This is something that conversational agents like the free version of ChatGPT 3.5 cannot do.

3.8 Fine tuning of LLMs

LLMs like Copilot are pre-trained on diverse data, but in specific applications they lack knowledge. This is where fine-tuning of LLMs is a possibility. To utilize the general language understanding that the pre-trained model has, but to fine tune and tailor it for a specific task or problem, like becoming an expert on some company specific data. This may very well be a possibility in an R&D environment, for example domain-specific literature like compliance, material science, electronics, chemical engineering, patent analysis or other company data.

Training a LLM completely from scratch (not using ChatGPT or similar as base) to acquire the desired capabilities of language understanding and solving specific tasks, is in many cases not feasible, because it requires a lot of computational resources and is time consuming. Fine-tuned, tailor made LLMs created for specific tasks is something that is talked about more, to get better quality results, compared to a universal LLM model that knows “something about a lot of things, but does not always know the [important] details”. For more details on how LLMs are trained, see appendix 1.

From practical point of view: imagine using a pretrained AI model to answer questions about how to polish a car properly. But it does not give particularly good answers according to your needs. Then you can train and fine tune the model with relevant questions and answers that suit your needs to improve the performance of the model.

3.9 Examples of LLMs

The landscape of LLMs is big and evolves by the day, with new models introduced at a rapid pace. They all have their strengths and weaknesses. The “best” LLM depends on the needs of the user. For a comparison of some LLMs, and assistance in choosing the correct LLM for a specific application, see for example (Takyar, u.d.). Understanding the characteristics of different LLMs and architectures help to choose what type to use. Some examples of LLMs are:

- ChatGPT-3.5 GPT-4, and GPT-4o (Marr, 2023). GPT-4o was launched on May 13th, 2024, and has improved speed and accuracy over its predecessors. It has multimodal capabilities (Cotton, 2024).
- Microsoft Copilot (based on GPT-4). Emphasizes creativity and NLU (Microsoft, u.d.).
- Google Gemini (multimodal LLM) (Google, u.d.).
- Perplexity AI

Goes on internet to find real time data, focus on accuracy. Combines elements of traditional search engines and AI chatbots (Perplexity, u.d.).

- Open-source models (for example on Hugging face platform)

Open-source models offer transparency, customization, and cost-efficiency, while closed-source models may provide superior performance and access to advanced research.

I decided to focus primarily on Copilot because it is used in the department. ChatGPT and Gemini are also big players in the LLM arena and I wanted to research them a bit also. Perplexity was chosen as an example also because I have personally found it good at acting as a search engine and finding relevant sources and use it regularly.

3.10 Risks and limitations

There are several risks to consider when using AI and LLMs. Some of the risks are:

Sensitive Data Security

Following company guidelines of what data can be shared or not in, for example Microsoft Copilot or ChatGPT is a good starter. Generally, you should never write sensitive company data in these tools, especially freeware versions. The Enterprise version of Microsoft 365 Copilot could however be able to use internal information safely due to enterprise confidentiality agreements (Microsoft, u.d.). If you concern whether you could ask a specific thing, it is advisable to generally ask only general questions regarding information that already exist on internet. consult IT departments or those who make up the policy regarding AI tool usage. Also, an important reminder regarding protecting sensitive company data is to have cybersecurity thinking with passwords and phishing emails. Complex passwords, for example to Outlook and not sharing it with other sites is a good start, because easy passwords are easy to hack-and secret company data may be at risk.

Incorrect answers

LLMs (for example ChatGPT) may sometimes provide answers that at first sight may seem trustworthy but the case is not always so. If the training data is wrong, the result also may be wrong. LLMs also cannot fact-check their answers. So because of that it is important to fact-check the result from several other, more reliable sources, like official websites, scientific articles, study books, or fact-check websites like snopes.com, factcheck.org or politifact.com (Dackén, 2023). Another option could also be to ask the LLM for its sources regarding the asked questions, and it may be wise to ask oneself what it means and what potential complications it may get if the answer replied from the LLM turns out to be wrong. Critical thinking is always important, and it may be wise to have a skeptical but constructive attitude when using LLM tools (Sporrong, 2016). LLMs cannot replace humans, they are simply a tool and is best for helping and not “doing all the work” (Wästlund, 2023).

Biased answers

This depends on what data the model is trained on. If the model is trained on historical biased data, like old books with old beliefs that are not true, there is a danger of perpetuating and magnifying those biases, which eventually can lead to discrimination. One example of this could be if the training text data argues that men are better than women at math and

women are better at art, which is not true. If the text data it is trained on is based on human feelings and not truths, it may produce unfair results (Lemaire, 2023).

Unemployment and organizational changes

Automatization of tasks and work tasks that are changed due to new tools and methods of doing work may eventually lead to a need for a less work force and education for the personnel to adapt to the new work methods in the best possible way. An interesting report from Accenture also highlights some of the risks, fears and feelings related to LLMs and generative AI, from employee point of view and leadership point of view (Accenture, 2024).

Ethical and juridical responsibility issues

When analyzing data, questions regarding personal data and integrity are also important to consider. Everyone who manages personal data is obliged to follow the data protection regulation GDPR. It is advisable to be cautious and make sure that what data you collect, and gather is legal (Intersoft Consulting, u.d.).

Copyright and plagiarism issues is something that may arise and is important to consider as well, when using LLMs. (Wästlund, 2023) ChatGPT has also been trained on copyrighted material so it is yet to be seen whether this will be handled in a court case at some stage (Tojin, Finkenstadt, & Folk, 2023).

Making people lazy-If people completely stop to think on their own, people soon are helpless if the technology for some reason does not work. Developing your own problem-solving skills and encouraging creativity is still essential.

High cost of AI systems

Buying for example Copilot for Microsoft 365 costs \$30/user/month, and if a bigger company, the monthly sums quickly rise. So it is advisable to thoroughly think through if such systems are suitable for a specific company and gives more value in terms of improved productivity VS the added cost (Microsoft, u.d.).

Other concerns and limitations

Another concern is the costs and emissions related to LLM training and usage. An analysis has shown that training an LLM model with 200billion parameters produces approximately 75000kg of CO2 emissions, compared to only 900kg for a flight from New York to San

Francisco. (Hugging face, u.d.) This also leads to making pre-trained models more attractive. Pre-trained models, either open-source or commercial, only require fine tuning for the specific use case and can be built and used quickly (Lenovopress, u.d.).

AI is a tool in your toolbox and can be seen as a personal assistant. One limitation is that AI models do not have good ability to understand the difference between correlation and causation. Causation is the ability to recognize how change in one variable directly influences change in another variable. One example is that there is a correlation between people eating ice cream and drowning accidents. But it is not the ice cream that causes more drowning, instead it happens to be warmer days that you eat ice cream and go for a swim. So, it is no causation in that case (Dackén, 2023).

ChatGPT 3.5 is a type of Language Model, and was introduced in November 2022 by OpenAI, and is an application that can generate context-aware, human-like responses to your questions. The freeware LLM ChatGPT 3.5 is not always feasible, because it can't find online data, and is only trained on data until September 2021.

3.11 What is prompt engineering?

Prompt engineering is a term that is used to describe the (iterative) process to construct and adjust questions or instructions (“prompts”) that are given to an LLM-model to achieve the best results.

In practice it may involve experimentation with several instructions, that step by step may get more specific to get more focused and specific answers. Methodology may include detailed instruction prompts about answer length, format, tone, critical thinking, creative thinking, or something else. There are several methods how to construct the prompts, for example Zero-shot learning, Fine tuning, or Few-shot learning (Dackén, 2023).

In fig.14 there is a screenshot from when a prompt was written to Copilot, and where it is asked to answer on how LLMs can be used in A R&D department. The output gives some examples of how LLMs could eventually be used in work, for example by assisting in analyzing data and generating insights, simplify complex technical language, assist in translations with international team, and assistance with patent analysis and IP related tasks and more. This list was not the most comprehensive, and when the same question was asked in Google Gemini, my opinion was in that specific case that Gemini gave more and relevant answers. It is interesting to evaluate different LLMs and see how they may differ.

The screenshot shows the Microsoft Copilot interface. At the top, there are logos for Windows, MIRKA, COPILOT, and NOTEBOOK. The user's name 'fredrik.fo...' and a 'Protected' status are visible. The user asks: "how can R&D departments utilize AI and particularly LLMs in order to work smarter and more efficient?". Copilot responds with a detailed answer, starting with a security notice: "Your personal and company data are protected". The response states: "Certainly! Leveraging AI and LLMs (Large Language Models) can significantly enhance the efficiency and effectiveness of R&D departments. Here are some ways they can be utilized:" followed by five numbered points:

1. Accelerating Ideation:
 - o AI and LLMs can streamline the **ideation process** by analyzing vast amounts of data, identifying patterns, and generating innovative insights.
 - o By automating initial stages, they free up time for more strategic and creative aspects of innovation.
2. Enhancing Collaboration and Communication:
 - o LLMs simplify complex technical language, fostering understanding and engagement among cross-functional teams.
 - o They can bridge language barriers, promoting seamless collaboration among international teams.
3. Improving Decision-making and Risk Assessment:
 - o AI-powered analytics and LLMs help R&D teams understand the innovation landscape.
 - o By analyzing historical data and predicting trends, they provide insights for shaping R&D and innovation strategies.
 - o They identify potential risks associated with new projects, enabling data-driven decisions.
4. Streamlining Patent Analysis and IP Management:
 - o AI and LLMs transform intellectual property (IP) portfolio management.
 - o They assist in patent analysis, identifying relevant prior art and potential infringements.
5. Virtual Assistants and Knowledge Acquisition:
 - o LLMs can serve as virtual assistants, summarizing information, answering questions, and improving knowledge acquisition efficiency [1](#) [2](#).

In summary, integrating AI and LLMs into R&D processes can lead to smarter decision-making, faster innovation, and a competitive edge. 🚀

Below the response, there is a "Related to:" section with an advertisement for "Demo Gen AI with Now Assist | Generate Content Automatically". The ad text says: "Put AI to work to swiftly resolve issues with chatbots that comprehend human language. Not a ServiceNow® registered user yet? It only takes a few seconds to create an account." and includes the URL: <https://www.servicenow.com/generative-ai/demo>. At the bottom, there is a text input field with the placeholder "Ask me anything..." and a microphone icon.

Fig.14 Example of a prompted question in Copilot of how LLMs can be used in R&D departments.

4 RESEARCH METHODOLOGY

In the quest to find out answers to the previously mentioned research questions (1. What are the current bottlenecks in office R&D work? 2. How can generative AI tools and LLMs be used to resolve these bottlenecks? 3. What LLM tools are currently available and what are their strengths, weaknesses, and limitations?) I have chosen the method described below.

4.1 Research method

The research methods I have chosen is as follows:

- **Literature studies** about LLMs (books, articles, theses, relevant web sources) to learn and understand more about how LLMs work. This includes finding out strengths, weaknesses, possibilities, and risks related risks using LLMs, and how others have used LLMs in work.
- An **internal survey** on the client R&D department. I have made a questionnaire that has been sent out to sixty people with various LLM related questions, how they use ChatGPT/ LLMs, what risks and possibilities there are, what their bottlenecks in work are and so on.

The survey questions consist of both quantitative and qualitative questions. The investigation is mainly of an exploratory and inductive nature, to gain as much knowledge as possible about the subject from the respondents (Patel, 2011). Apart from these two methods to find out answers to the research problem, I also will do some personal reflections regarding my own work tasks, bottlenecks in my own work and some testing of LLM tools and my thoughts that will be mentioned in the discussion chapter.

- Evaluate LLM tools (gain knowledge in empirical way)
- Reflect how own work could be made more efficient with help of LLMs
- Ask Language models for suggestions and improvements

It would have been interesting to get other perspectives as well on this topic, for example combine the survey with interviews with persons on the R&D department, IT department, AI consults, or to make a survey sent out to other companies R&D departments also. Due to time limitations, the current method was chosen. Regarding the process of selecting

questions to include in the survey, they were to some extent driven by the curiosity to know how others perceive and use LLMs and learn something from each other.

4.2 Data collecting method

The data collecting method was to make the questionnaire in google forms and send it out to the respondents via email.

4.2.1 Ethical considerations in data collection

Ethical aspects were taken into consideration in the form of the fact that all responses were anonymous, and the ambition was to learn from each other, inspire and get inspired and bring benefits to the whole department.

4.3 Data analysis method

The data analysis method was then to read, summarize and visualize the responses in Excel. There have been no structured interviews with respondents.

4.4 Result presentation method

The ambition is to give a clear and quick insight into how the respondents think about the questions. The result from the survey is illustrated by graphs and tables that clearly define how many respondents have chosen a specific answer.

4.5 Limitations in the research approach

- The sample is not too big.
- Some questions in the survey with predefined answers may have relevant answers left out. For example, some applications where LLMs are used, that the survey maker did not think of.
- Some questions do not have standardized questions, which makes the summarization of these answers a bit more complicated.
- Trying to grasp challenges of whole R&D department with quite different work tasks is a challenge. Focusing solely on mechanical engineers, for example, would

have been easier to grasp, but then the sample would have been a lot smaller, and some valuable ideas may have been unnoticed.

A more detailed investigation into bottlenecks in work and reflection on work time segmentation at the R&D department would be interesting to research as well. Because if we had facts about how much time we put on different work tasks, it is easy to spot the bottlenecks in the work and easier to find focus areas that could be made more efficient. And reflect on what tasks could be made more efficient by the assistance of LLMs. This, however, would have made the scope too big for this thesis, and is left for someone else to do. The work field is wide and work processes can always be improved to achieve the desired goals faster and with better quality results.

5 LLM APPLICATIONS IN THE R&D PROCESS



Fig. 15 Engineers creating yellow Power Tools with the help of AI.

Source: Microsoft Copilot / Dall-E 3 (Microsoft, u.d.)

There are a lot of ways that LLMs can be used in R&D related work tasks. Here are some of them mentioned. An R&D process includes several key steps. LLMs can significantly help and be of assistance and improve efficiency in many of these steps.

Some of the R&D process steps include:

Idea generation and opportunity identification

Brainstorming and market research often plays an important part in generating new innovative concepts and exploring new opportunities. LLMs may here have the possibility to enhance the ideation process by acting as an assistant and suggesting interesting, inspiring ideas and help in identifying market opportunities and trends. According to Hwang & Won, participants contributed more ideas and ideas of higher quality, when they perceived their team working partner to be a bot. They concluded that AI-based systems could address the fear of negative evaluation and therefore allow for more idea-sharing (Lemaire, 2023). Visualizations in, for example DALL-E could also generate ideas.

Idea evaluation

Often there are a lot of ideas of new tools or features. But every idea of a new product or feature is not possible for several reasons. There is a need to find a clear focus on the most promising products or features that also are in line with the company strategy. The feasibility of an idea can be evaluated for example for its technical

design, how well it solves customer problems and fits the target market, its financial potential, price tag and considering the effort and resources needed.

Research planning

It is wise to think through the development process when developing a new product or tool (and adapting the development plan if needed). Defining (research) objectives with the development, finding out the suitable specifications, defining key customers and learning their daily challenges, and thinking about what methods or tests to do to find out these tasks. And make some timeline estimations and hopefully learn from past experiences and projects also. Here LLMs may have the possibility to help in making a research plan when designing a new tool with a little help of prompt engineering, for example.

Experimentation

Perform tests and simulations to gather data and evaluate ideas and concepts. Here LLMs may assist if you have questions regarding experimenting, or to summarize the results from the experiments.

Data Collection

Systematically collect data from experiments, tests, surveys, observations, or simulations. Here LLMs have immense potential to summarize and organize the data.

Analyze and interpret data

Analyzing and understanding raw data in, for example an Excel chart and finding patterns or correlations or analyzing customer feedback for sentiment analysis about a product are some examples of where LLMs can be of assistance.

Prototyping and development

LLMs cannot help in making the actual physical prototypes but may assist in coaching regarding features to include who the target users are, what typical user scenarios look like, idea refinement and communication and collaboration and documentation of ideas.

Testing and validation

Here the product is evaluated extensively to find flaws, or weaknesses or vulnerabilities in the design and find things to improve. And, to confirm the performance, functionality, safety, and durability of the tool. Here LLMs may assist in developing test methods for example.

Information finder and learning assistant

LLMs may here be helpful in searching for relevant parts to use, asking regarding technical specifications, summarizing technical documents, or asking them to extract specific data out of a document. Learning new things and asking for technical explanations, technical literature summaries, or questions for example regarding materials, physics, math, electricity or taking a photo of your problem and giving it to Copilot or any other preferably confidential multimodal LLM and ask how you could solve the problem are some interesting approaches also.

Iteration and refinement

To really get the data and feedback from test users and experts down to paper in an obvious way and taking the feedback into consideration when developing the concept further and discussing it with stakeholders. It is to be decided who to listen to because people tend to have different opinions. Is it the boss who has the final word, your gut feeling, key customers, data driven decisions from customers or what the majority in the R&D group thinks? LLMs may assist in organizing the feedback, or transcribing teams meetings to text where these topics and test results are discussed, or transcribing customer interviews to text for further (sentiment) analysis, summarizing the content and give action plans that are decided from the meeting (Reid, 2024).

Communication and documentation

Here LLMs are in a strong position and can assist in many ways. Email writing, presentation, translations, transcribing meetings, making to-do lists, summarizing meetings, writing test reports. Here the Microsoft 365 suite with integrated Copilot features in Word, Excel, PowerPoint, Teams, Outlook, OneNote etc. has exciting potential to unlock more fluent and efficient work methods and help in keeping the red thread in projects-having a clear picture of where in the development process we are, where we came from and where we are heading.

IP related tasks

Extract insights from patent databases and learn prior art, patent novelty evaluation, creating patent drafts and claims, identify patent trends, summarize patents, or rewrite the text in more easily understandable way (Saini, 2024).

6 AI SURVEY

As a reminder, the research problem of this thesis is to find out:

Can we improve productivity and innovation in the R&D department with the help of LLMs?

To find out an answer to this research problem, I formulated the research questions below, that I will try to answer, before it can be said whether LLMs can improve productivity and innovation in the R&D department. The research questions for the thesis were 1. What are the current bottlenecks in office R&D work? 2. How can generative AI tools and LLMs be used to resolve these bottlenecks? And 3. What LLM tools are currently available and what are their strengths, weaknesses and limitations? This survey aims to find answers to research questions 1 and 2. The first step was to learn more about LLMs and explore some potential applications, which has been presented in the theory chapter and the application study chapter. This knowledge of LLMs was helpful when defining relevant questions to the survey and was a steppingstone in finding out answers to research question 2. The second step in this research in order to find out answer to research question 1 was to explore the R&D departments challenges and bottlenecks at work and ask if they use LLMs, how they use LLMs and what tools or methods that possibly could make their work more efficient and of higher quality. To find out more about these topics, a survey with questions related to AI-use in work was sent out to the Power Tools department. 8 questions were asked, and answers between 4 to 20 responses were gathered to the different questions. The survey is of

explorative art and may hint about the employers' attitudes regarding AI in general and particularly LLMs and their usage. This knowledge may be of value both in terms of learning from each other and realizing the eventual possibilities by using LLMs but also when considering implementing eventual new AI initiatives and strategies from a management point of view.

6.1 Survey Questions

The following questions were asked:

Q1: *Do you use ChatGPT in your work?* (yes/ no question)

Q2: *f yes, for what do you use ChatGPT?* (Asked twice. First in freeform and after that as a multiple-choice question by ticking boxes of predefined user applications. The result presented here is the multiple-choice result)

Q3: *What risks do you see with using ChatGPT at work?* (freeform answer)

Q4: *What are the bottlenecks in your work today?* (freeform answer)

Q5: *If you could automate one work task-what would it be?* (freeform answer)

Q6: *What is the challenge of implementing AI tools at the company?* (multiple choice answer)

Q7: *Argument: AI means more possibilities than risks* (single choice answer 1...5)

Q8: *For what could an internal chatbot be usable?* (freeform answer)

6.2 Respondents

There were 20 respondents out of 64 employees in the department. 57 out of 64 are men, and most are educated engineers in ages ranging between 25 and 63, with a median age of 35 years.

Since the query was anonymous it is not known in exact detail who the respondents were.

However, the R&D department consists of:

- After Sales (service center organization, CRM systems, spare parts etc.)

- Compliance (works with regulations, standards, manuals etc.)
- Mechanical, Electrical and SW engineers
- Test engineers (performs testing and measurement on tools, parts, and writes test reports)
- Quality engineers (quality follow-up, statistics, risk analysis etc.)
- Project Management

7 RESULTS

Below are the results from the survey. For each of the eight questions, an appropriate chart and table is shown that includes the number of respondents and the task (s) or answer they have chosen.

7.1 Survey Results

Q1: Do you use ChatGPT in your work? (20 respondents)

Answer: 50% yes, 50% no.

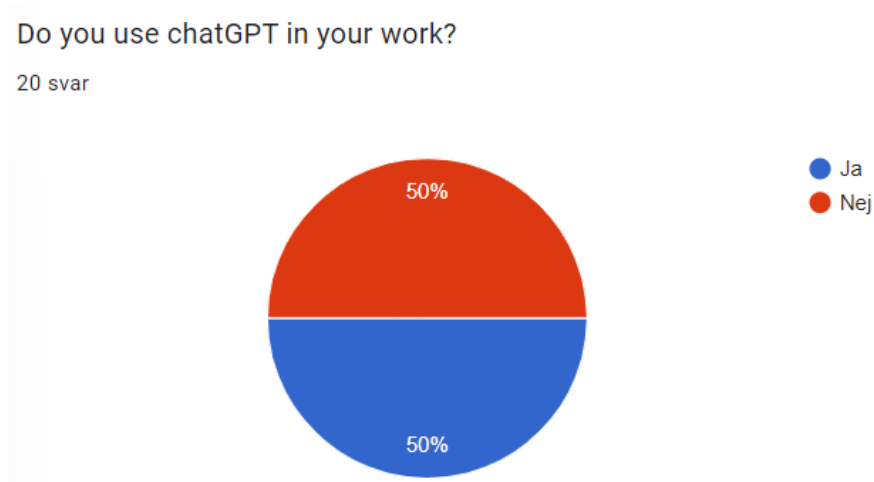


Fig.16 ChatGPT users in the department out of the 20 respondents.

Question respondents in total: 20

Answer	% that use	Respondents
Yes	50 %	10
No	50 %	10

Q2: For what do you use ChatGPT? (16 respondents)

Answer: Overall, the respondents seem to find value in ChatGPT for mainly technical explanations, information search and problem solving. A more detailed chart is seen below.

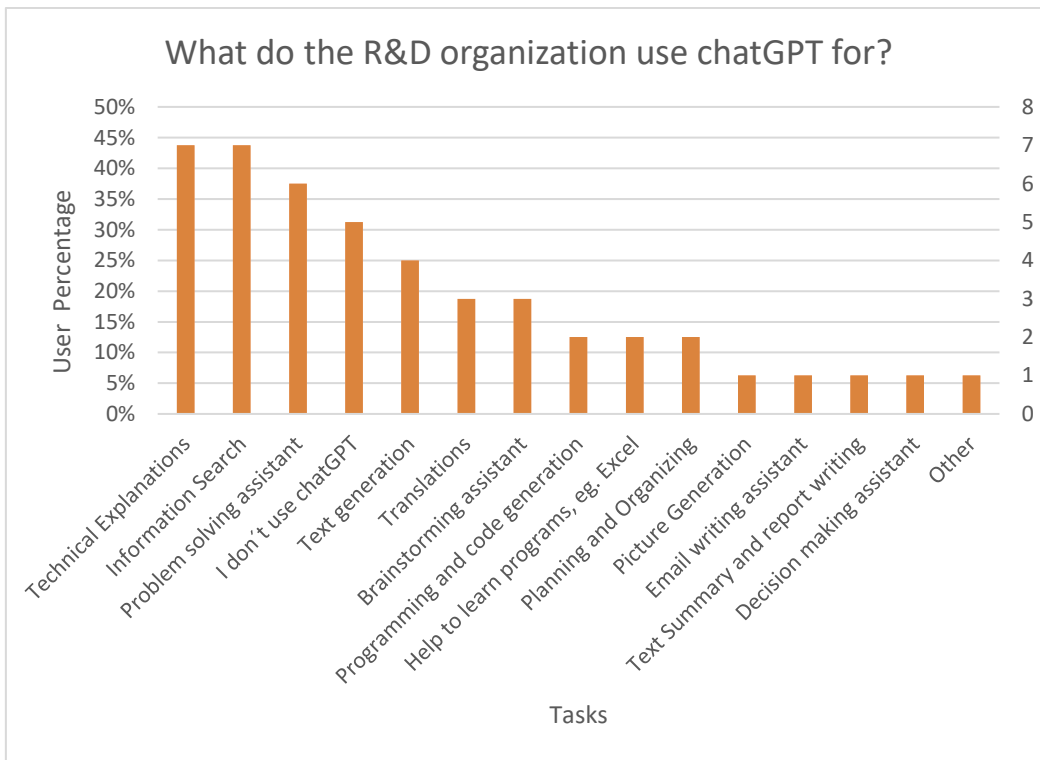


Fig.17 Some tasks the department use ChatGPT for.

Question respondents in total: 16

Task	% that use	Respondents
Technical Explanations	44 %	7
Information Search	44 %	7
Problem solving assistant	38 %	6
I don't use ChatGPT	31 %	5
Text generation	25 %	4
Translations	19 %	3
Brainstorming assistant	19 %	3
Programming and code generation	13 %	2

Help to learn programs, e.g., Excel	13 %	2
Planning and organizing	13 %	2
Picture Generation	6 %	1
Email writing assistant	6 %	1
Text Summary and report writing	6 %	1
Decision making assistant	6 %	1
Other	6 %	1

Q3: What risks do you see by using ChatGPT at work? (17 respondents)

Answer: Don't use confidential data is seen as the biggest risk.

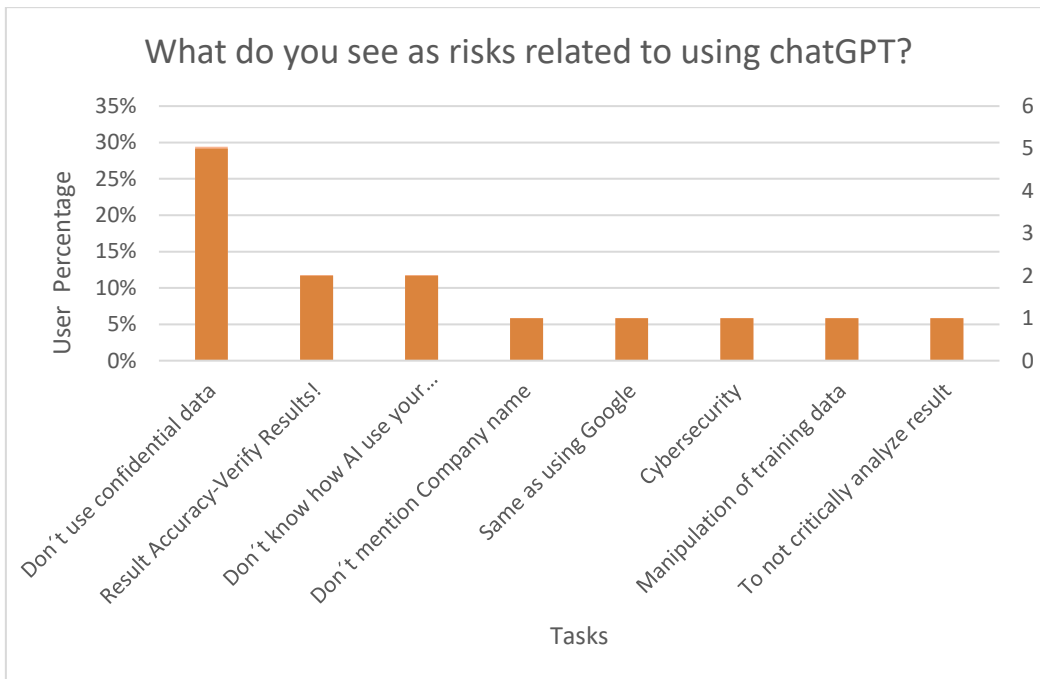


Fig.18 Risks related to using ChatGPT.

Question respondents in total: 17

Task	% that use	Respondents
Don't use confidential data	29 %	5
Result Accuracy-Verify Results!	12 %	2
Don't know how AI use your data	12 %	2
Don't mention Company name	6 %	1
Same as using Google	6 %	1
Cybersecurity	6 %	1
Manipulation of training data	6 %	1
To not critically analyze result	6 %	1

Q4: What are the bottlenecks in your work today? (12 respondents)

Answer: Information search and text reading are the biggest bottlenecks according to the survey.

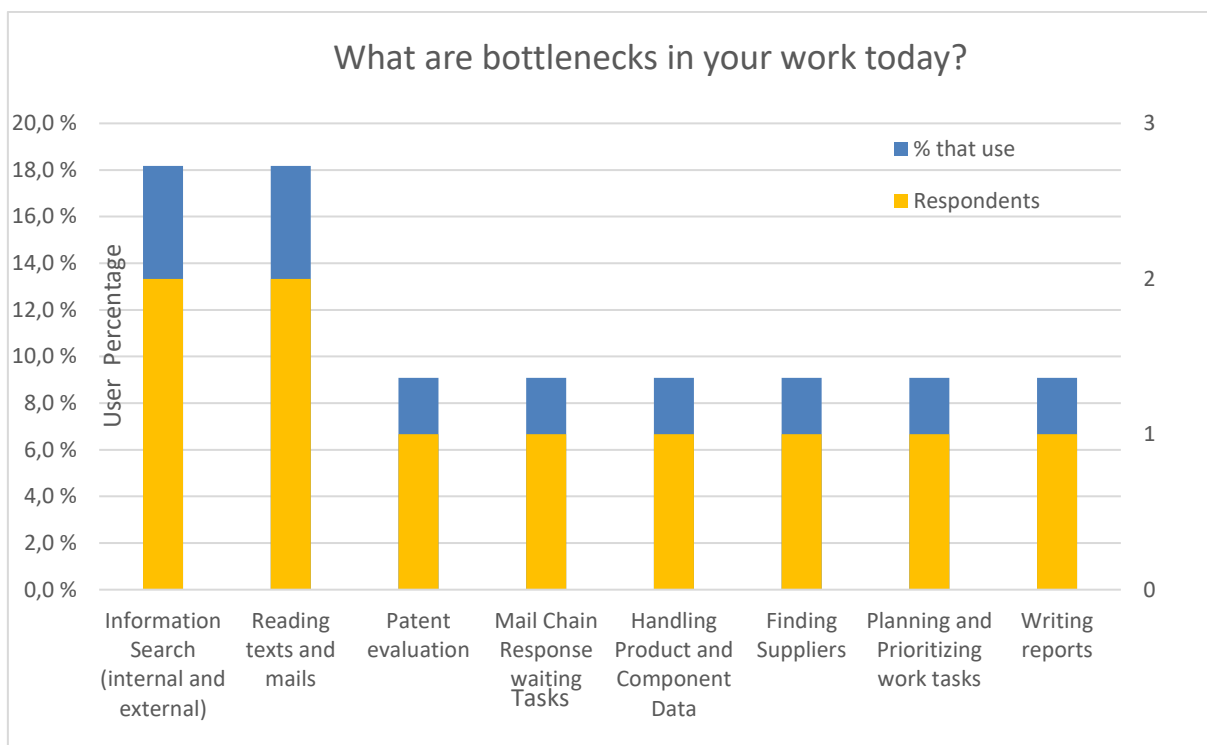


Fig.19 Bottlenecks in R&D work.

Question respondents in total: 11

Task	% that use	Respondents
Information Search (internal and external)	18,2 %	2
Reading texts and mails	18,2 %	2
Patent evaluation	9,1 %	1
Mail Chain Response waiting	9,1 %	1
Handling Product and Component Data	9,1 %	1
Finding Suppliers	9,1 %	1
Planning and Prioritizing work tasks	9,1 %	1
Writing reports	9,1 %	1

Q5: If you could automate one work task-what would it be? (13respondents)

Answer: Patent idea evaluation would be priority to automate according to the survey.

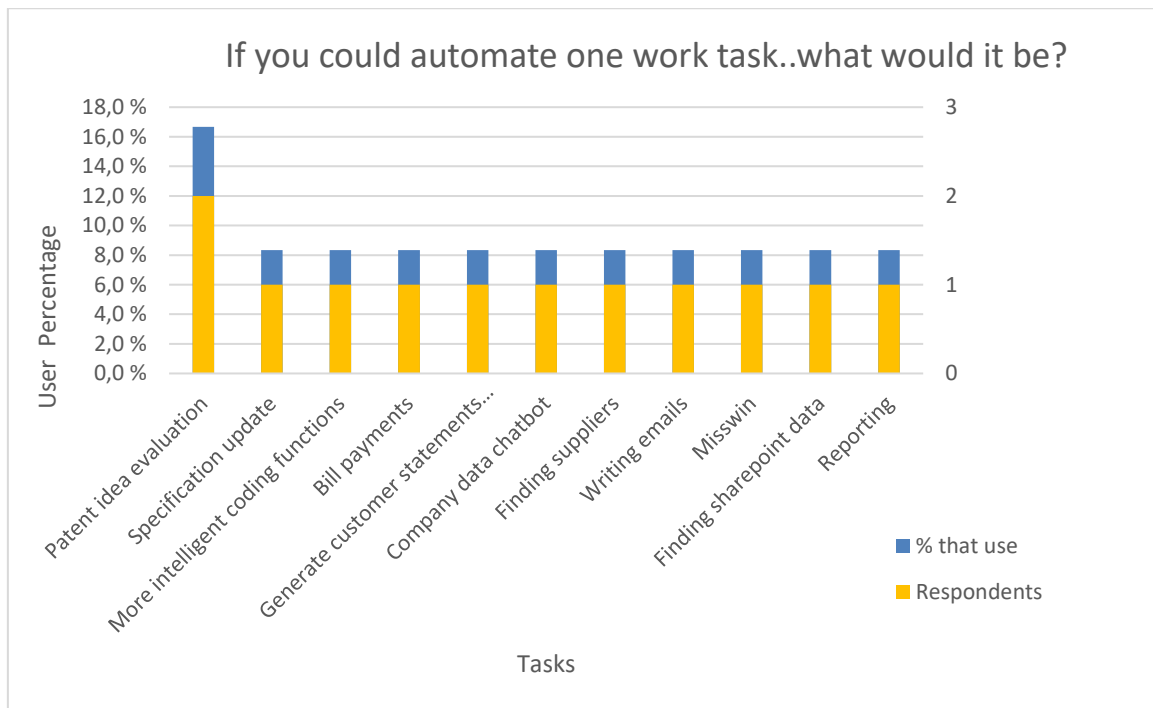


Fig.20 Automation of work task wishes.

Question respondents in total: 13

Task	% that use	Respondents
Patent idea evaluation	15,4 %	2
Specification update	7,7 %	1
More intelligent coding functions	7,7 %	1
Bill payments	7,7 %	1
Generate customer statements based on supplier data	7,7 %	1
Company data chatbot	7,7 %	1
Finding suppliers	7,7 %	1
Writing emails	7,7 %	1
ERP program tasks	7,7 %	1
Finding SharePoint data	7,7 %	1
Reporting	7,7 %	1

Q6: What are the challenges to implement AI tools at the company? (20 respondents)

Answer: Lack of knowledge, cybersecurity concerns and finding a clear focus are challenges emphasized in the result.

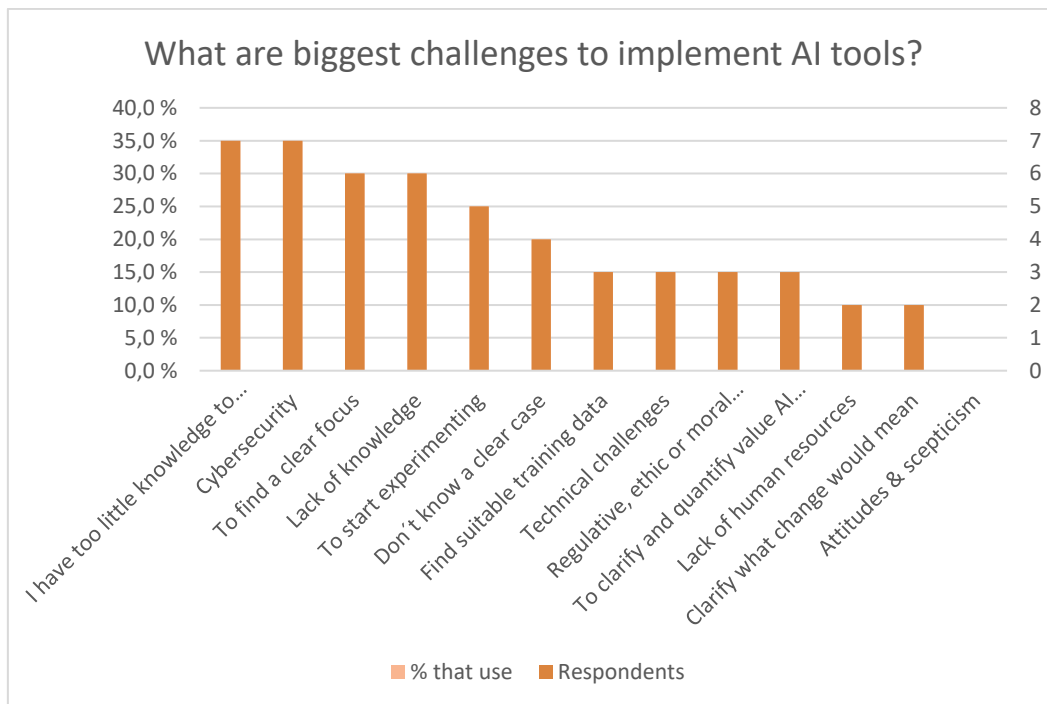


Fig.21 Organizational challenges related to AI implementation seen by employees.

Question respondents in total: 20

Task	% that use	Respondents
I have too little knowledge to answer	35,0 %	7
Cybersecurity	35,0 %	7
To find a clear focus	30,0 %	6
Lack of knowledge	30,0 %	6
To start experimenting	25,0 %	5
Don't know a clear case	20,0 %	4
Find suitable training data	15,0 %	3
Technical challenges	15,0 %	3
Regulative, ethic or moral concerns	15,0 %	3
To clarify and quantify value AI solution would give	15,0 %	3
Lack of human resources	10,0 %	2
Clarify what change would mean	10,0 %	2
Attitudes & skepticism	0,0 %	0

Q7: Statement: AI means more possibilities than risks (20 respondents)

1= I do not agree, 5 = Totally agree (20 respondents)

Answer: Average answer 3.85 (=” mainly agree”)

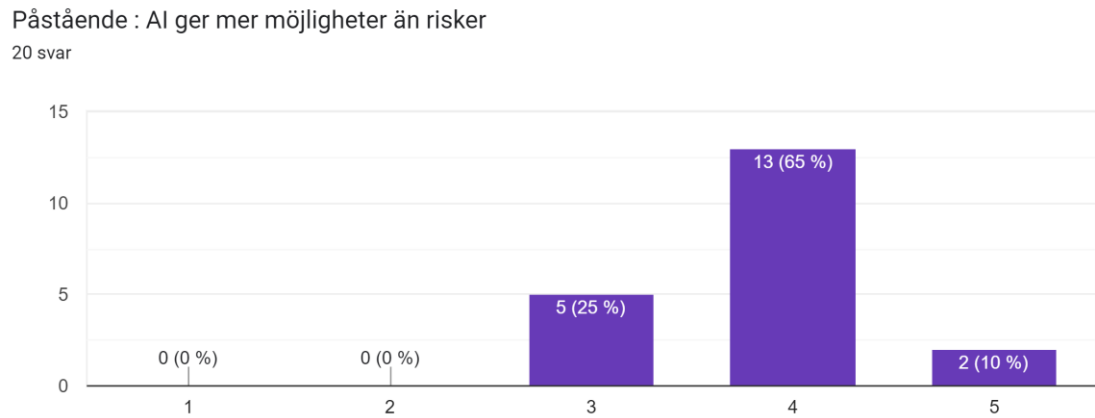


Fig.22 Organizational challenges related to AI implementation seen by employees.

Comment: The respondents mainly seem to have a positive bias towards AI.

Q8: For what could an internal chatbot be useable? (13respondents)

Answer: Lots of ideas summarized below.

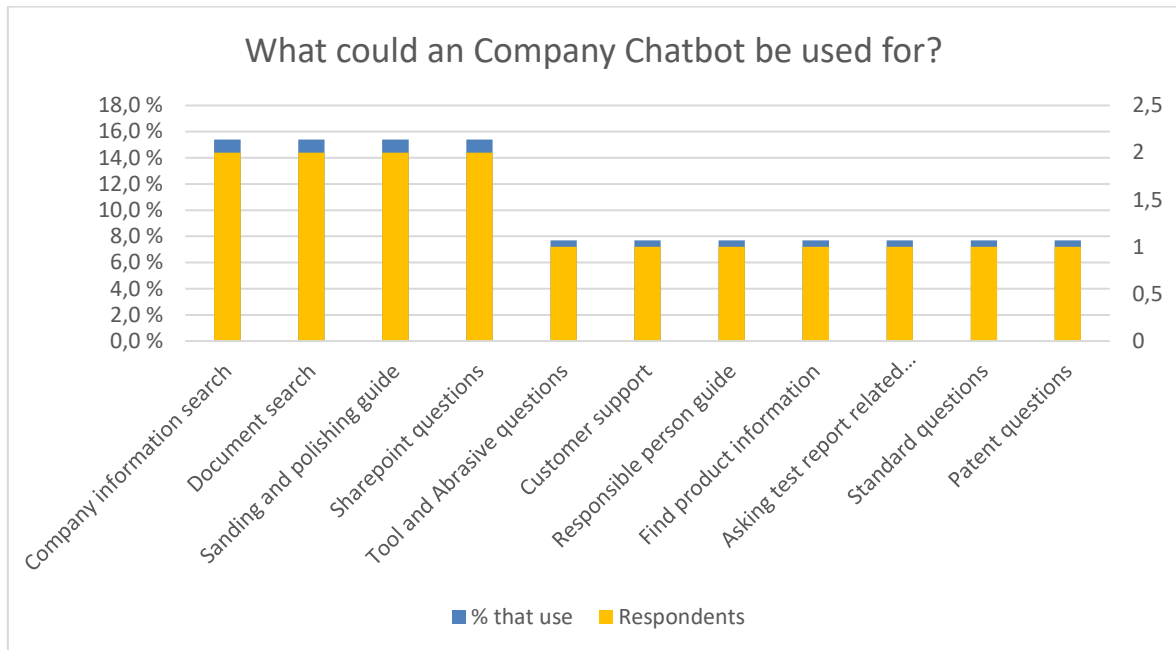


Fig.23 Some ideas what a company chat could be used for

Question respondents in total:

13

Task	% that use	Respondents
Company information search	15,4 %	2
Document search	15,4 %	2
Sanding and polishing guide	15,4 %	2
SharePoint questions	15,4 %	2
Tool and Abrasive questions	7,7 %	1
Customer support	7,7 %	1
Responsible person guide	7,7 %	1
Find product information	7,7 %	1
Asking test report related questions	7,7 %	1
Standard questions	7,7 %	1
Patent questions	7,7 %	1

8 Discussion of survey results

First, it can be mentioned that as the number of respondents varies for the questions, and in some questions are very few, this must be taken into consideration when analyzing the result. Additional to this is the fact that the respondents work with varying office work tasks, so the results are “the big picture” of how LLMs were used in this specific R&D department at the time when the survey was conducted in January 2024. The survey result shows some interesting results of how the R&D department see possibilities and risks related to using LLMs in their work. From the research problem point of view (that was: *Can we improve productivity and innovation in the R&D department with the help of LLMs?*) LLMs seem to be a versatile tool that can make R&D departments work more efficient in various ways and especially text related tasks by using tools like Copilot for example, if used properly and being aware of the related risks and knowing that LLMs basically are a statistical tool that predicts the next word or token in a certain context, and is limited to its training data. And like weather predictions, they are never 100% accurate.

Q1: The result shows that 50% of the respondents, or 10 out of 20 persons use Copilot/ChatGPT in work. It is unclear whether the other 40 people use it or not.

Q2: It is interesting to see how the department finds value in using ChatGPT and LLMs in several ways. The main applications according to the survey result are technical explanations (= learning), information search and using it as an assistant in problem solving. While many applications are listed, there probably are a lot of applications that still could be listed and relevant and useful. Also, it would add even more value and be easier to grasp if the segments were broken down into sub-segments and giving more detailed examples of “success stories.” For example, give concrete examples how it could assist in an innovation and conceptualization development process of a new tool, or extracting relevant knowledge from customers and the market. Besides all this I also look with excitement on the future and wonder if LLMs also at some point may revolutionize the IP process, as hinted about in chapter 5.

Q3: The main risks with LLMs in R&D work according to the result are seen to be the risk of writing confidential data, and the risk of getting faulty answers (that in worst case may lead to bad decisions). Always fact check the result from more trustworthy sources!

Q4: The main bottlenecks in the work according to survey are information search and reading text and emails. But only 12 respondents answered the question, and only two persons chose

these two as main bottlenecks, so perhaps not the most trustworthy result. But according to the mentioned theory, information search accounts for a big portion of knowledge worker's time. Also interesting is the spread and bottlenecks of contrasting characters. Other things I see as bottlenecks in my work as a mechanical engineer is to specify and create the final design of a tool. It is not a one-man show, and it depends on many factors where many need to give input before the decisions can be made. Also, a big bottleneck indeed is IP related topics such as evaluating ideas whether they are novel or not, analyzing the patent landscape, and reading and improving patent claims and text. Also, what takes time is meetings and drafting emails.

Q5: The most desired task to automate would be patent idea evaluation. I can only agree this would be helpful, and with more automation and assistance withing the whole Intellectual Property (IP) field in general. People don't really have time to sit for days or even weeks to read and fine tune patent text repeatedly, for something that eventually even never will be a valuable granted patent.

Q6: The main challenges to implement AI tools at the company are seen as lack of knowledge, cybersecurity concerns and finding a clear focus are challenges mainly emphasized in the result. Can companies trust that the big tech companies providing the LLMs manage their confidential information securely, or are open source, fine-tuned LLMs used on a local server something to consider for future applications? It was a broad and general question. No one has chosen that attitudes and skepticism would be a challenge to implement AI tools. Instead, it seems like more knowledge and exploration about possibilities where LLMs really could be effective could be a way forward.

Q7: The statement that AI means more possibilities, most of respondents seem to agree with and tend to have a mostly positive attitude towards AI. On a scale 1...5 (where 1 = Do Not Agree and 5 = Totally Agree) the average answer out of 20 respondents was 3.85, so people mostly tend to agree.

Q8: An internal AI chatbot could be useful for company information search, internal document search, answers to general company related questions among other proposals. To easily find answers to general company-related questions like what person to speak to when an IT problem occurs could be extra valuable for new employees, or where a certain policy document or document template can be found would save time and effort.

A potential bias in the result may also be the tendency to give answers related to the subject (AI and LLM related answers). For example, regarding bottlenecks in work, most of the answers are bottlenecks that LLMs eventually could assist with. But in terms of time spent on various tasks, meetings are one of the biggest time thieves for the average R&D person-but that bottleneck is not mentioned in the answers.

It is not possible by this survey to quantify how much time generally is saved for different use cases, and that is something that eventually could be investigated in a separate research case.

Also how do you measure the value of improving the language, clarity and structure in your test reports or emails? Or clarifying customer feedback, when prototype tools have been on test, and you want to have good quality data-driven decisions of how to improve the concept further? Or the value of getting context aware feedback on your questions?

There is a big likelihood that some key concepts are missing in the survey result of how LLMs can be utilized in R&D work. This is probably because people simply have not realized all the possibilities where LLMs can be of assistance (and neither have I), and the development in the field is so rapid.

I also think LLMs work best as an assistant, and not being “the master”. For example, seeing Copilot as a “colleague” to brainstorm ideas with, evaluate and develop ideas or concepts, search for information or explaining technical language and more. If cybersecurity and confidentiality issues are under control and the company is assured that confidential information is kept secret, then there is easier to think more of how this AI tool could be utilized for improved work efficiency, work quality and profession development.

If I had started the survey over from the beginning again, with the knowledge I have about LLMs and R&D work processes today, I would have formulated the survey and the questions differently. For example, questions related to Microsoft 365 version where Copilot are included not only in the edge browser but also in all the Microsoft programs such as Teams, Outlook, Word, Excel, OneNote and others, and what value people may find in the added features that give.

It would also have been nice to include some qualitative interviews with R&D people, experts in the field and the IT department to more thoroughly discuss both confidentiality issues and defining the “freedom to operate” with LLM tools and discussing the future, and

some discussion about the company AI strategy and investigate what AI strategy other companies has.

In R&D work much text is read, processed, analyzed, written and formulated all the time. People search for relevant information and want to learn new things. And if it is possible to find the right answers to our problems faster, one will become more efficient. Of course, as always when using LLMs like Copilot it is important to be critical about the result and fact-check the result from other reliable sources. The do's and don'ts when using LLMs in daily work crystallizes increasingly as the organization learn and makes policies and a strategy of how to use AI.

9 CONCLUSION

- This thesis contributes to shed light on some of the practical challenges and bottlenecks in work faced by professionals in R&D departments. It also contributes to a discussion about applications of LLMs in R&D departments.
- According to the survey, the theory and from own testing-LLMs shows potential to being an asset in R&D work and can save time and effort if used wisely. It has the potential for freeing up time for other activities and making the work more efficient. Generative AI and LLMs can be seen as a tool that strengthens our development and competence.

The research questions for this thesis were:

Research question 1: What are the current bottlenecks in office R&D work?

The main bottlenecks in the work according to the survey are information search and reading text and emails. This is in alignment with the theory that suggest that information search takes up lot of time for knowledge workers. Other bottlenecks were patent evaluation, waiting for mail response, handling product data, finding suppliers, work planning and report writing. 12 respondents answered the question.

Research question 2: How can generative AI tools and LLMs be used to resolve these bottlenecks?

Most of the bottlenecks found are of such a type where LLMs may be of help. Including for example Copilot for text processing tasks in the work processes and trying other LLMs such as Perplexity or Google Gemini for general context relevant information search could be a starting point. Another test could be to implement and use Copilot in Teams meetings,

to transcribe and summarize meetings automatically. It is then after that possible to generate to-do lists and summarize meeting reports for example.

Research question 3: What LLM tools are currently available and what are their strengths, weaknesses and limitations?

Currently at the R&D department Copilot is available in licensed and secure version in the web browser. The department does not have the Copilot app available in the Microsoft Office programs like Teams, Outlook, Word and Excel - yet. Some strengths of LLMs are to quickly and efficiently understand the meaning and context of what you search for or want to do. It is strong in creating new text, summarizing and explaining text, simplifying text, translating text and coming up with creative ideas. Some weaknesses are that they can sometimes hallucinate and give false answers, be manipulated and give biased answers depending on their training data, and that freeware versions of LLMs not should be used with confidential company or personal data, because these models may pass the information forward. Some other limitations are that LLMs sometimes are referred to as “black boxes” because it is difficult to explain their decision making, context window length and token limitations. The possibilities for more efficient, better-quality work and inspirational ideas for creative problem solutions seem endless, but being aware of risks and limitations with LLM usage is essential. Embrace the potential of LLMs as an assistant but don´t trust everything you read and always fact-check with other, more reliable sources. LLMs can leverage your professionalism in writing high quality text, learning new things faster and more. The art of prompt engineering (learn to ask the LLM for what you want in a precise manner) is vital in the quest of doing so. (Dackén, 2023)

“Language models are like skilled assistants: they excel at aiding us, suggesting solutions, and simplifying tasks. However, when it comes to groundbreaking research and innovation, they should remain our trusted companions, not the sole masters of our endeavors.” (Microsoft, u.d.)

10 DISCUSSION

The main purpose of the thesis was to make the identified bottleneck work processes more efficient. First it can be concluded that some bottlenecks were identified from the survey result in question 4. Information search and text reading were the two biggest bottlenecks according to the survey. It can be argued that it is likely that LLMs to some extent can make these two work tasks more efficient (among other NLP related tasks). Reflecting on this purpose, I personally think this is a process that takes time and people need time and

education and test themselves before this could happen. And the management also need input so they know what the bottlenecks are and what possible ways there are to address those bottlenecks. People also need the possibility to express their thoughts, experiences, and feelings regarding AI and LLMs and if they like they can test what makes their work processes more efficient and what may develop their professional skills. On a personal level LLMs have improved my efficiency at work. I have found and evaluated several new LLMs and tested to ask them various problems and often they help me forward in some way, even if there is a need to be skeptical to some answers, but if the reply suggest some (reliable) source, it is easier to determine what is true and what is not. This work has widened my views regarding LLMs and the wide AI field in general. This work still is only a scratch on the surface of the world of LLMs. A basic understanding of how LLMs work will be helpful when talking to IT architects and programmers about wishes from an R&D point of view and developing new tools and solutions. From curiosity perspective, one personal next step in the learning process regarding LLMs could be to test if we can use it in a more structured way in our work processes. For example, by transcribing meetings to text and process that data further, make to-do lists and action plans, summarize meeting reports, create timelines, and project plans and illustrate how the project process goes forward and making sure people focus on the right things. After a project is completed, perhaps we can summarize an extensive report out of all the gathered project data with the actual project timeline and challenges faced, so we can visualize the facts and be realistic and proactive for upcoming projects so we can avoid potential bottlenecks in the future.

Fine tuning of LLMs and the art of prompt engineering is likely something we will use more of in the future. As LLMs develop further, these techniques will likely play an important role in tailoring models to specific tasks like training on company data and giving fast answers to company related questions, summarizing customer feedback or extracting information from some other relevant knowledge base.

Regarding research question 1, the survey question had 13 respondents and it suggested that information search, reading and understanding [long & complex] texts as well as working with patents are seen as major bottlenecks. R&D work is multifaceted in its nature, and I can agree that all these sometimes may be bottlenecks. From a personal point of view, challenges and bottlenecks [in R&D work] may often come in phases. It may be so that you for a month or two sit and work with 3D modeling (drawing complex surfaces and getting it to work often requires patience and lots of iterations and waiting for 3D printing that often take 1week/iteration if bigger parts and you might lose focus as well. Then you work with

and test physical prototypes some months, Then you test at customers (and may get problems you not have noticed in test environment) and you have a lot of insights and opinions of how to develop further, and somehow need to rate all the input data and decide how to proceed. And you may face technical problems of various sorts that may require extensive work and testing before it is solved, or perhaps you need to make some changes so that the product will pass certification. And since projects may take between two to five years, it is a challenge to remember all that has been done in the project. If all knowledge that in a project has been gained- and remembered in a smart way that can easily be checked, there is less risk for double work and an immense potential for working more efficiently. To keep the red thread in the project and work in a logical way towards the set goals. Usually there are many things that needs to be done in a project, and unfortunately IP work-patent related tasks- are often not the highest priority, So any help with evaluating novelty of new ideas, help in showing the relevant patent landscape, writing patent drafts, suggestions of claims, and explanations of complicated patent texts would be very helpful and make the work more efficient. So, there may be lots of different bottlenecks in different stages in the work, but my top three bottlenecks from personal point of view is probably to imagine and create complex 3D models and defining its features and specifications, patent related work tasks and to find the best solution to a tricky problem. Regarding research question 2 (How can generative AI tools and LLMs be used to resolve these bottlenecks?) Summarized in the Survey Result chapter and Application Findings chapter, I found many things where LLMs could potentially be useful in R&D work and resolve bottlenecks. It was interesting to read from the survey what others use ChatGPT for. Many use cases were quite familiar from own knowledge and testing before, like summarizing text and using it as a context-aware search engine. It is always interesting to hear others' points of views. Some of my favorite things to prompt in Copilot may be to summarize and reflect on a text, Information search, assist in technical explanations of for example physical phenomena like electricity, gyro forces for better understanding, or how to do things in Excel or SolidWorks, or seeing it as a colleague to assist and brainstorm in problem-solving. From a practical perspective in real world application - one day we installed a Wi-Fi relay to actuate a light and could not get it to work. I did not find the solution with conventional googling or on the manufacturer's website, so I described the problem to Google Gemini, and it concluded I had old firmware in the app. After FW update in the app, it worked!

Personally, I rarely have used Copilot for email writing yet. I understand this may be of value when sending general mails to a broader public or rephrasing the tone of the language to fit

a certain category of receiver or improving the structure of the content for improved clarity. From personal point of view when mailing with suppliers with lot of project-specific details I prefer to keep the control and write by myself and make sure that no LLM rationalize away something important or write in a tone that “isn’t me”.

It was also interesting to read and learn from the relevant Thesis: *Application of the Latest Generation of AI in New Product Development Process* (Lemaire, 2023).

From a critical point of view, LLMs might not always be ideal where 100% correct answers are always needed. Also, in situations where a reliable source is needed to back up a given answer, LLMs might not always give source citations. There is also risk for potential bias in the answers the LLM gives, depending on what data the model is trained on. Open-source models generally are more open with how they are trained and could from that perspective be of interest (Canales, 2023).

11 FUTURE

The development of AI and LLMs is rapid. Who knows where we stand after a year or two? Internal company chatbots fine-tuned for specific tasks like internal information search or patent assistants and more Copilot licenses when people get more familiar with the tool and the department has a clear communicated AI strategy? I guess we have a lot of unseen yet. Hopefully, humankind uses these powerful tools in a responsible way for good and ethical purposes towards a brighter future and away from dystopic scenarios. Yesterday we walked, with today's LLMs we can cycle around in various directions, and tomorrow perhaps we drive an AI motorcycle that drives our efficiency and knowledge forward on a map with clear goals.

12 BIBLIOGRAPHY

Accenture. (2023). *A new era of AI for everyone*. Retrieved May 10, 2024, from <https://www.accenture.com/content/dam/accenture/final/accenture-com/document/Accenture-A-New-Era-of-Generative-AI-for-Everyone.pdf>

Accenture. (2024). *Work, workforce, workers -reinvented in the age of generative AI*. Retrieved May 20, 2024, from <https://www.accenture.com/content/dam/accenture/final/accenture-com/document-2/Accenture-Work-Can-Become-Era-Generative-AI.pdf>

Camposato, O. (2020). *Artificial Intelligence, Machine Learning, and Deep Learning*. Mercury Learning & Information.

- Canales, J. (2023, November). *Top open-source LLMs*. Retrieved April 20, 2024, from <https://www.datacamp.com/blog/top-open-source-llms>
- Corina. (2023, July 27). *Generative AI and LLMs in Business: Opportunities and Challenges*. Retrieved May 20, 2024, from <https://www.wingravity.com/blog/generative-ai-and-llms-in-business/>
- Cotton, R. (2024, May). *What is OpenAI's GPT-4o? Launch Date, How it Works, Use Cases & More*. Retrieved June 1, 2024, from <https://www.datacamp.com/blog/what-is-gpt-4o>
- Dackén, R. (2023). *AI Advantage-Mo Money, Mo sales*. Lund: Ekström & Garay.
- Eapen, T., Venkataswamy, L., Finkenstadt, D., & Folk, J. (2023, June). *How generative AI can augment human creativity*. Retrieved February 3, 2024, from <https://hbr.org/2023/07/how-generative-ai-can-augment-human-creativity>
- Ekblom, C. (2015). *Framgångsrika förändringar-Insikter, Strategier och Ledarskap*. Roos & Tegnér.
- Google. (n.d.). *Google Gemini Models*. Retrieved May 30, 2024, from <https://deepmind.google/technologies/gemini/>
- Google. (n.d.). *Image FX-google labs*. Retrieved March 10, 2024, from <https://aitestkitchen.withgoogle.com/tools/image-fx>
- Google. (n.d.). *Teachable Machine-Train a computer to recognize your own images, sounds, & poses*. Retrieved March 10, 2024, from <https://teachablemachine.withgoogle.com/>
- GrowthLoop. (2024, February 9). *Large language model (LLM)*. Retrieved March 25, 2024, from Growthloop: <https://www.growthloop.com/university/article/llm>
- Hugging face. (n.d.). *How do Transformers work?* Retrieved April 9, 2024, from <https://huggingface.co/learn/nlp-course/chapter1/4?fw=pt>
- IBM. (n.d.). *What is knowledge management?* Retrieved April 28, 2024, from IBM: <https://www.ibm.com/topics/knowledge-management>
- IBM. (n.d.). *What is NLP?* Retrieved April 29, 2024, from <https://www.ibm.com/topics/natural-language-processing>
- Intersoft Consulting. (n.d.). *General Data Protection Regulation*. Retrieved April 7, 2024, from <https://gdpr-info.eu/>
- Khan, U. (2023, August 9). *The Unstoppable March of Artificial Intelligence: The Dawn of Large Language Models*. Retrieved March 31, 2024, from [www.esignals.fi: http://urn.fi/URN:NBN:fi-fe2023080994491](http://urn.fi/URN:NBN:fi-fe2023080994491)
- Knapp, M. (2023, June 8). *How does chatGPT work anyway*. Retrieved May 23, 2024, from <https://www.currentdraft.com/p/how-does-chatgpt-work-anyway>
- Lawton, G. (n.d.). *What is generative AI? Everything you need to know*. Retrieved May 24, 2024, from <https://www.techtarget.com/searchenterpriseai/definition/generative-AI>

- Lemaire, J. (2023, December). *Application of the Latest Generation of AI in New Product Development Process*. Retrieved April 24, 2024, from Theseus: https://www.theseus.fi/bitstream/handle/10024/851525/Lemaire_Naudin.pdf?sequence=2&isAllowed=y
- Lenovopress. (n.d.). *Reference Architecture for Generative AI Based on Large Language Models (LLMs)*. Retrieved April 9, 2024, from <https://lenovopress.lenovo.com/lp1798-reference-architecture-for-generative-ai-based-on-large-language-models?orgRef=https%253A%252F%252Fcopilot.microsoft.com%252F#use-cases>
- Marr, B. (2023, May 19). *A Short History Of ChatGPT: How We Got To Where We Are Today*. Retrieved May 16, 2024, from Forbes: <https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/?sh=253f163e674f>
- Microsoft. (n.d.). *Azure AI Document Intelligence*. Retrieved March 10, 2024, from <https://azure.microsoft.com/en-us/products/ai-services/ai-document-intelligence>
- Microsoft. (n.d.). *Copilot for Microsoft 365*. Retrieved April 5, 2024, from [www.microsoft.com: https://www.microsoft.com/en-us/microsoft-365/enterprise/copilot-for-microsoft-365#Pricing](https://www.microsoft.com/en-us/microsoft-365/enterprise/copilot-for-microsoft-365#Pricing)
- Microsoft. (n.d.). *Copilot-your everyday companion*. Retrieved March 18, 2024, from <https://copilot.microsoft.com/>
- Microsoft. (n.d.). *Microsoft 365 copilot privacy*. Retrieved May 21, 2024, from <https://learn.microsoft.com/en-us/copilot/microsoft-365/microsoft-365-copilot-privacy>
- Microsoft. (n.d.). *Microsoft Azure Machine Learning*. Retrieved March 3, 2024, from https://azure.microsoft.com/en-us/free/machine-learning/search/?ef_id=_k_CjwKCAiA0bWvBhBjEiwAtEsoWwLUwze3LUGVFtpQNMmu_8Vv8EL9IVd00Y8fcDe-Xk6QtWDaHXuOKhoC3EwQAvD_BwE_k_&OCID=AIDcmmftanc7uz_SEM_k_CjwKCAiA0bWvBhBjEiwAtEsoWwLUwze3LUGVFtpQNMmu_8Vv8EL9IVd00Y8f
- Microsoft. (n.d.). *What is computer vision?* Retrieved April 28, 2024, from <https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-computer-vision#object-classification>
- Microsoft. (n.d.). *What is RPA? How robotic process automation streamlines business processes*. Retrieved April 28, 2024, from <https://powerautomate.microsoft.com/en-us/what-is-rpa/>
- Mohaiminul Islam, G. C. (2019, June 29). *American Journal of Neural Networks and Applications. Vol 5*. Retrieved 4 19, 2024, from https://d1wqtxts1xzle7.cloudfront.net/89062715/10.11648.j.ajjna.20190501.12-libre.pdf?1658993904=&response-content-disposition=inline%3B+filename%3DAn_Overview_of_Neural_Network.pdf&Expires=1717597790&Signature=HDIH4AxDQSFsjte-bv4EjE5SoG0cHLOq7B1FmRibnUXL

- Motley Fool. (2024, 1 16). *What is artificial intelligence?* Retrieved February 12, 2024, from <https://www.fool.com/terms/a/artificial-intelligence/>
- Oludare Isaac Abiodun, A. J. (2018, November 13). *State-of-the-art in artificial neural network applications: A survey*. Retrieved April 28, 2024, from [https://www.cell.com/heliyon/pdf/S2405-8440\(18\)33206-7.pdf](https://www.cell.com/heliyon/pdf/S2405-8440(18)33206-7.pdf)
- OpenAI. (n.d.). *OpenAI blog*. Retrieved January 12, 2024, from Open AI: <https://openai.com/blog>
- Ozdemir, S. (9-2023). *Quick Start Guide To Large Language Models-Strategies and Best Practices for using ChatGPT and other LLMs*. New Jersey: Addison-Wesley Professional.
- Patel, R. &. (2011). *Forskningsmetodikens grunder: Att planera, genomföra och rapportera en undersökning*. Lund: Studentlitteratur.
- Patenall, H. (2022, Jul 26). *Who are knowledge workers and how does AI technology speed up their work?* Retrieved 6 6, 2024, from <https://www.aiimi.com/insights/who-are-knowledge-workers-and-how-does-ai-technology-speed-up-their-work>
- Perplexity. (n.d.). *Perplexity AI*. Retrieved March 10, 2024, from <https://www.perplexity.ai/>
- Rahman, W. (2020). *AI and Machine Learning*. SAGE Publications India Pvt, Ltd.
- Rebelo, M. (2023, December 1). *The best AI chatbots in 2024*. Retrieved April 25, 2024, from <https://zapier.com/blog/best-ai-chatbot/>
- Reid, B. (2024, February 14). *Copilot 365 in Teams meetings- To Transcribe or not?* Retrieved May 20, 2024, from <https://c7solutions.com/2024/02/copilot-365-in-teams-meetings-to-transcribe-or-not>
- Runa Pathel, B. A. (2003). *Forskningsmetodikens grunder*. Lund: Studentlitteratur.
- Saini, D. (2024, January 4). *Navigating the Future of R&D: Integrating LLMs into Patent Processes*. Retrieved May 23, 2024, from XLScout: <https://xlscout.ai/navigating-the-future-of-rd-integrating-llms-into-patent-processes>
- Sarker, I. H. (2021). *Machine Learning: Algorithms, Real-World Applications and Research Directions*. Springer Link.
- Scispace. (n.d.). *SciSpace-Do hours worth of reading in minutes*. Retrieved March 10, 2024, from <https://typeset.io/>
- Siljerud, P. (2023). *AI för offentlig sektor-Insikter, inspiration och möjligheter*. Stockholm: Kajsa Lindroth.
- Sporrong, W. (2016). *Kritiskt tänkande i teori och praktik*. Lund: Studentlitteratur AB.
- Suno. (n.d.). *Make music with your mind*. Retrieved June 1, 2024, from <https://suno.com/>

- Takyar, A. (n.d.). *Comparison of LLMs*. Retrieved May 10, 2024, from <https://www.leewayhertz.com/comparison-of-llms/>
- Techguruspeaks. (n.d.). *Types of machine learning*. Retrieved May 31, 2024, from <https://www.techguruspeaks.com/types-of-machine-learning/>
- Tojin, E., Finkenstadt, D., & Folk, J. (2023). The generation game. *The Economist*.
- Wikipedia. (n.d.). *Large Language Models*. Retrieved April 1, 2024, from https://en.wikipedia.org/wiki/Large_language_model
- Wolfram Alpha*. (n.d.). Retrieved May 22, 2024, from <https://www.wolframalpha.com/>
- Wästlund, H. (2023). *AI för mindre företag*. Tipsförlaget.

13 APPENDICES

13.1 Appendix 1-LLM training

Training a LLM involves several steps, including:

1. **Identify goal or purpose for the LLM.**
2. **Data collection:** gather huge and diverse dataset of text.
3. **Data preprocessing:** cleaning of the data, removing “noise” and converting the text into suitable format for training (e.g., numerical representation)
4. **Architecture selection:** Choose appropriate neural network architecture, for example an transformer-based architecture like GPT.
5. **Model initialization:** Initialize the LLM with pre-trained weights.
6. **Fine-tuning:** Fine tune the LLM for a specific task, e.g., text generation, question answering or translation.
7. **Loss function and optimization:** Define a loss function that measures the difference between predicted and actual text. Optimize the model using gradient descent or other optimization algorithms.
8. **Hyperparameter tuning:** (Experiment with hyperparameters (e.g., learning rate, batch size and sequence length) to find optimal settings to improve the predictive accuracy.
9. **Training:** Train the model on the dataset. Give input text sequences to the model and adjust weights (in the neural network of the model) based on the loss (the difference between the true answer and the predicted one).
10. **Evaluation:** Evaluate the model’s performance and accuracy
11. **Deployment:** Once the model is trained, the LLM model can be deployed for a specific application (e.g., chatbots and content generation). (GrowthLoop, 2024)

13.2 Appendix 2-Transformers- a brief overview

A transformer is a fundamental architecture in natural language processing (NLP) and machine learning. It enables efficient training of large-scale language models. Here are the key points:

1. **Working Principle:**

- Transformers are based on self-attention mechanisms.
- They process input sequences in parallel and can find long-range dependencies. For example, when LLMs generate a sentence, it later in the sentence adjusts the words in such a way so that it grammatically and contextually fits to what has been written before.

2. **Architecture:**

- Transformers consist of an encoder-decoder structure or a single encoder for tasks like language modeling.
- Layers of attention heads learn contextual relationships between words.

3. **Attention Mechanism:**

- Attention computes weighted sums of input representations.
- It assigns higher weights to relevant words and lower weights to irrelevant ones.

4. **Pre-training and Fine-tuning:**

- Transformers are pre-trained on massive text corpora.
- Fine-tuning adapts them to specific tasks (e.g., sentiment analysis, translation).

5. **Application examples in R&D:**

- Automating literature reviews: Summarizing research papers.
- Idea generation: Brainstorming novel concepts.
- Predictive modeling: Forecasting material properties.
- Customized chatbots: Giving instant answers to researchers.

6. **Challenges:**

- Transformers require enormous amounts of data for effective fine-tuning.
- Hyperparameter tuning is crucial for best performance.

In summary, transformers empower R&D teams by efficiently managing language tasks, enabling innovation, and bridging the gap between research and practical applications.