INSTITUTO SUPERIOR DE TECNOLOGIAS AVANÇADAS DE LISBOA

Master's in computer science

Thesis for master's degree in computer science

Adopting Copilot for Businesses

Helder Harshil Valabh

Thesis Advisor: Prof. Dr. Pedro Brandão

Lisbon

November 2025

INSTITUTO SUPERIOR DE TECNOLOGIAS AVANÇADAS DE LISBOA

Master's in computer science

Thesis for master's degree in computer science

Adopting Copilot for Businesses

Helder Harshil Valabh

Presidente: Professor Doutor Paulo André Reis Duarte Branco

Arguente: Professora Doutora Maria Inês Vasconcellos Furtado

Orientador: Professor Doutor Pedro Ramos dos Santos Brandão

Lisbon

November 2025

Acknowledgements

As I write this thesis, I can’t help but reflect on how the acknowledgements section is often overshadowed by the academic and technical content we present. Yet, behind every project like this lies countless hours of hard work, dedication, and support from others something that deserves recognition.

I want to extend my heartfelt thanks to everyone who stood by me during this long journey. To my girlfriend, thank you for your constant support and encouragement. And to my parents, thank you for always allowing me to pursue my passions freely, and for giving me the space to explore, even when things didn’t go as planned.

This thesis has been a personal ambition of mine, especially as we live in an era where technology and AI are rapidly transforming the world around us. AI feels like just the tip of the iceberg a gateway to many more innovations to come. As we move into the future, I truly believe it will continue to reshape how we live and work.

Abstract

Artificial intelligence (AI) has become a critical tool for transforming workplace productivity, and Copilot exemplifies this shift by automating repetitive tasks and improving workflow efficiency. This thesis explores the adoption and impact of Microsoft Copilot in enterprise environments, where the focus is on integration, challenges and difficulties.

Microsoft Copilot is an AI assistant that enhances productivity by leveraging advanced Large Language Models, including GPT-4, and integrating with Microsoft Graph to provide context-aware insights using organizational data. Copilot is designed to boost productivity, streamline workflows, and enable organizations to make better use of their existing data by providing AI-driven insights and automating tasks.

Using insights from official Microsoft documentation, this study analyzes how Copilot supports enterprise-specific needs while addressing challenges such as workforce training, implementation barriers, and data privacy concerns.

The objective of this thesis is to provide practical recommendations for companies looking to adopt AI tools like Microsoft Copilot, helping them overcome challenges and maximize their business value. Additionally, this research contributes both academically and to business practices by offering an analysis of the impact of these technologies. The study aims to support technicians and students in understanding the implications of AI and its role in the digital transformation of organizations.

Keywords: Copilot Microsoft 365, Microsoft Copilot, Generative AI, AI Training, LLM, Machine Learning

Resumo

A inteligência artificial (IA) tornou-se uma ferramenta crucial na transformação da produtividade no local de trabalho, sendo o Copilot um exemplo claro desta mudança ao automatizar tarefas repetitivas e melhorar a eficiência dos fluxos de trabalho. Esta dissertação explora a adoção e o impacto do Microsoft Copilot em ambientes empresariais, com foco na integração, nos desafios e nas dificuldades associados.

O Microsoft Copilot é um assistente baseado em IA que melhora a produtividade ao tirar partido de modelos avançados de linguagem natural, incluindo o GPT-4, e ao integrar-se com o Microsoft Graph para fornecer informações contextuais com base em dados organizacionais. O Copilot foi criado para aumentar a produtividade, simplificar os processos de trabalho e permitir que as organizações aproveitem melhor os dados que já possuem, através da automatização de tarefas e da geração de insights orientados por IA.

Com base na documentação oficial da Microsoft, este estudo analisa de que forma o Copilot responde às necessidades específicas das empresas, enquanto aborda desafios como a formação da força de trabalho, os obstáculos à implementação e as preocupações com a privacidade dos dados.

O objetivo desta dissertação é fornecer recomendações práticas bem como uma breve observação para empresas que pretendem adotar ferramentas de IA como o Microsoft Copilot, ajudando-as a ultrapassar desafios e a maximizar o seu valor para o negócio. Para além disso, esta investigação oferece um contributo relevante tanto para o meio académico como para a prática empresarial, ao apresentar uma análise do impacto destas tecnologias. O estudo pretende ainda apoiar técnicos e estudantes na compreensão das implicações da IA e do seu papel na transformação digital das organizações.

Palavras-chave: Copilot Microsoft 365, Microsoft Copilot, Generative AI, LLM, Machine Learning.

Table Of Contents

Contents

List of Tables

List of Figures

List of acronyms and abbreviations

AI – Artificial Intelligence

API – Application Programming Interface

ML – Machine Learning

LLM – Large Language Model

NLP – Natural Language Processing

Chapter 1: Introduction

1.1 Background and Motivation

The rise of Artificial Intelligence (AI) has been a defining force in reshaping the modern business landscape. Organizations across industries are increasingly adopting AI-powered tools to streamline operations, enable data-driven decision-making, and gain a competitive edge. Among these tools, Microsoft Copilot has gained significant attention for its seamless integration within the Microsoft 365 ecosystem. By automating repetitive tasks, enhancing collaboration, and providing intelligent insights, Copilot marks a substantial step forward in how enterprises can improve productivity and optimize workflows.

The motivation behind this research comes from the need to show more relevance for Copilot for businesses to adopt AI solutions that not only enhance efficiency but also address specific operational and organizational challenges. Despite its promising capabilities, the adoption of Microsoft Copilot is not without its hurdles. Many companies face difficulties in understanding its full potential, aligning it with existing workflows, and navigating challenges such as technical complexity, organizational change, and workforce adaptation.

Additionally, examining Copilot in relation to other AI tools such as Google Gemini and ChatGPT offers valuable context for understanding its unique strengths and areas for improvement within the competitive AI landscape.

This thesis explores how Microsoft Copilot is being used in real business scenarios. It focuses on real-world success stories and projects to show the practical impact Copilot has on productivity and workflows. The goal is also to give organizations some clear insights and strategies they can use to adopt Copilot effectively.

1.2 Objectives of the Study

The primary objective of this thesis is to examine the implementation of Microsoft Copilot within enterprises, focusing on its impact on productivity, challenges and benefits.

Specifically, this study will:

1. Investigate the role of AI in business, its benefits, and the challenges associated with its adoption.
2. Assess how Copilot enhances productivity in Microsoft 365 applications such as Word, Excel, and PowerPoint.
3. Analyse some of the benefits and obstacles enterprises face when integrating Copilot into their workflows.
4. Evaluate Copilot's scalability across industries and organizational sizes through success stories.
5. Provide insights strategies for implementing Copilot's long-term business impact.

These objectives guide the core research questions of this thesis, titled "Adopting Microsoft Copilot for Enterprises":

Why Copilot?

This question investigates what sets Microsoft Copilot apart from other AI solutions in the market, and why enterprises are choosing it as part of their digital transformation strategy.

What are the main challenges enterprises face when adopting AI, and how can organizations overcome these barriers?

This question explores obstacles such as employee resistance and technical limitations while identifying strategies to optimize AI adoption.

How does the implementation of Microsoft Copilot influence productivity in enterprise environments?

This question examines the impact of Copilot on employee efficiency, decision-making speed, and collaboration within daily workflows.

By addressing these questions, this study aims to bridge theoretical gaps in AI adoption literature while providing actionable insights for enterprises undergoing digital transformation.

1.3 Contributions and Impacts

The outcome of this research has the potential to provide meaningful contributions to different areas, including academically, IT professionals, enterprises.

From an academic perspective, it addresses research gaps in AI adoption, particularly the integration of tools like Microsoft Copilot within enterprise ecosystems. While AI implementation in businesses is a growing field of study, focused research on Copilot remains limited.

This thesis contributes to the field of education by exploring how tools like Copilot fit into everyday workflows. By examining its adoption, the research offers a nuance of the framework and practical guidelines that businesses can use laying the groundwork for future studies on other AI-driven productivity solutions.

For enterprises and IT professionals, this research provides practical recommendations for implementation. IT managers and decision-makers can leverage these findings to address key challenges, such as identifying the most impactful use cases for Copilot and overcoming employee resistance to AI adoption.

Moreover, by evaluating Copilot's impact on productivity, this study helps businesses understand the concrete benefits of AI, including efficiency improvements in data analysis, and task automation. Furthermore, successful AI adoption requires structured training and change management strategies, which this research highlights to ensure a smooth transition and minimize disruptions to existing workflows.

At the user experience level, this research explores how Copilot can improve work-life balance by automating repetitive and time-consuming tasks. By reducing manual workload, Copilot enables employees to focus on more strategic and creative work, addressing concerns about job displacement while highlighting its role in enhancing productivity.

Additionally, employees often struggle with information overload, and Copilot helps mitigate this challenge by offering AI-driven suggestions and automating routine processes, making work more manageable and efficient.

Beyond individual and enterprise applications, this research contributes to broader discussions on digital transformation and sustainability. By inspiring organizations to embrace AI tools, it accelerates innovation and the adoption of modern enterprise technologies. Furthermore, Copilot and similar AI tools contribute to more sustainable business operations by streamlining workflows, reducing inefficiencies, and optimizing overall performance.

Ultimately, this research not only explores the practical implications of Microsoft Copilot's adoption but also contributes to an understanding of AI's transformative role in modern workplaces. By addressing these critical topics, this study serves as a valuable resource for individuals, students and organizations looking to navigate the evolving landscape of AI-driven productivity tools. It highlights how far AI has come and demonstrates its potential for daily use, starting with Copilot as a key enabler in professional environments.

1.4 Structure of the Thesis

This thesis is structured with six chapters. The **first** chapter stablishes the context of the research by defining the scope of the study, which focuses on the integration of productivity-enhancing tools such as Microsoft Copilot in enterprise environments.

The research problem is mentioned, emphasizing the importance of assessing the practical impact of such tools on organizational efficiency. The chapter also outlines the specific objectives and research questions guiding the investigation, highlighting the study's significance to fields such as business administration, technology management, and organizational behavior. Finally, the structure of the thesis is presented, providing an overview of the chapters and their logical progression.

The **second** chapter reviews existing literature related to productivity-enhancing technologies, intelligent systems, and their application in business contexts. It begins with an exploration of what AI is and how it evolved to improve business operations and efficiency. The integration of intelligent systems into enterprise workflows is then discussed, with attention to factors influencing adoption and effectiveness. The chapter concludes with a detailed examination of Microsoft Copilot, its functionalities, and its integration with Microsoft 365, drawing from current research to present both expected and observed outcomes.

The **third** chapter, which is the methodology, outlines the research design, justifying the choice of a case study approach and the justification behind qualitative or mixed-methods research. It details the methods of data collection, including interviews, surveys, and document analysis, and explains the criteria for selecting case study enterprises. The analytical framework is also described, including techniques for qualitative content analysis and comparative analysis, with an explanation of how productivity will be measured.

The **fourth** chapter presents the success stories of enterprises that have implemented Microsoft Copilot, providing contextual information on each organization. A comparative analysis follows, examining the impact of Copilot on productivity across different business contexts and industries. The findings are analyzed to assess the effectiveness of Microsoft Copilot in enhancing organizational efficiency and productivity.

The **fifth** chapter examines the results obtained from surveys and interviews conducted to evaluate the adoption of Microsoft Copilot in enterprise environments. The analysis highlights key patterns, insights, and implications derived from the collected data, offering a thorough assessment of Copilot's role in business operations. This discussion seeks to connect theoretical approaches to adoption with the practical challenges organizations encounter during implementation.

Finally, the **sixth** chapter wraps up the main findings of the research and offers advice for businesses, including strategies for putting these tools into action.

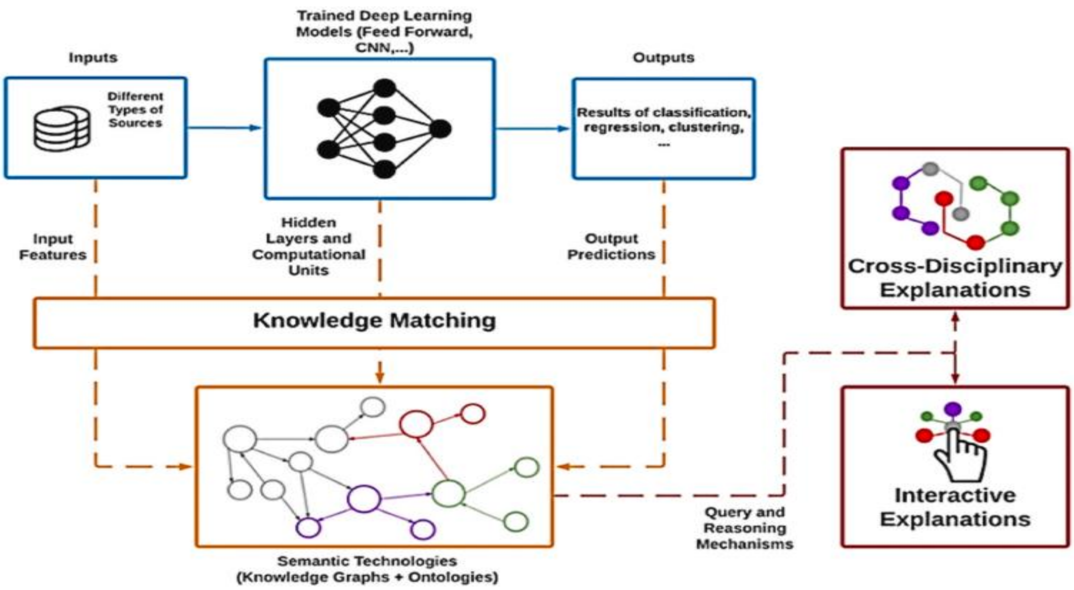
Chapter 2: State of Art

1.1 What is Artificial Intelligence?

AI refers to the ability of machines to exhibit intelligence like human cognitive functions, such as problem-solving and decision-making. Utilizing Machine Learning and Deep Learning techniques, AI systems process vast amounts of data to make autonomous decisions and adapt to new information. ***Figure 1*** will illustrate a basic AI model in action. These systems mimic human intelligence by analyzing data, identifying patterns, and interpreting natural language (Rashid & Kausik, 2024).

Figure 1

Working Principle of AI *(Rashid & Kausik, 2024)*



According to Rashid & Kausik the aim and objective of AI is to create intelligent machines called computed robots for various needs. AI is deeply integrated into our daily lives, often in ways people may not notice or fully understand. It plays a crucial role in various aspects of modern life, making it difficult to function without its assistance. Virtual assistants like Amazon's Alexa, Apple's Siri, Microsoft's Cortana, and Google Assistant rely on AI to interpret spoken commands and execute tasks (Rashid & Kausik, 2024) .

Streaming platforms such as Spotify and Netflix use AI to analyze user preferences and recommend personalized content. Similarly, companies like Google, Facebook, Amazon, and eBay track online behavior to deliver targeted advertisements. AI is present everywhere, shaping our experiences in daily life, entertainment, careers, and essential services (Rashid & Kausik, 2024)

It is also widely used in fields such as manufacturing, industrial automation, finance, data analysis, healthcare, agriculture, autonomous vehicles, workforce management, and decision-making, among many others (Rashid & Kausik, 2024).

1.2 Artificial Intelligence in Business

The integration of Artificial Intelligence (AI) into business operations has become a transformative force, reshaping how organizations function. AI technologies empower companies to automate tasks, extract valuable insights from data, and enhance decision-making speed and accuracy.

As Füller (2022) noted, AI is expected to have a profound impact across industries, comparable to the Internet's rise three decades ago or electricity's adoption a century ago, with the potential to contribute $13 trillion to global GDP by 2030.

AI has been particularly beneficial for small businesses, especially during crises such as the COVID-19 pandemic. In the insurance sector, AI-driven business models have diversified industry structures, allowing both traditional insurers and technology firms to redefine their roles within the value chain (Oluwatoyin Ajoke Farayola et al., 2023).

Despite its advantages, AI integration comes with challenges, including high technological investment, the need for specialized expertise, and ethical concerns related to data privacy. However, its ability to enhance decision-making, improve customer experiences, and create new business opportunities makes AI a strategic necessity in today's economy (Oluwatoyin Ajoke Farayola et al., 2023).

According to Rashid & Kausik the integration of artificial intelligence (AI) in agriculture is experiencing rapid growth, with a projected compound annual growth rate (CAGR) of approximately 25% from 2023 to 2031.

AI applications in precision farming have significantly improved efficiency, increasing crop yields by 20–30%, while AI-driven irrigation systems have reduced water consumption by up to 25%. Similarly, AI-enhanced fertilizer and pesticide application methods have optimized input usage and minimized waste.

The rising global population, which reached 7.9 billion by the end of 2022, has intensified the demand for food production. Traditional agricultural methods struggle to keep pace, leading to an increased reliance on chemical inputs such as pesticides and fertilizers.

However, excessive use of these substances depletes soil quality over time, rendering land unproductive and imposing substantial economic costs. AI technologies offer innovative solutions to these challenges by enhancing productivity, optimizing resource utilization, and mitigating the environmental impact of conventional farming practices.

Beyond agriculture, AI is also transforming the education sector. Personal computers, assistive technologies, and robotics are advancing student learning, particularly in early childhood education. Collaborative robots now assist teachers by helping young learners develop fundamental skills such as spelling and pronunciation while adapting to individual learning needs.

Additionally, digital education advancements have improved accessibility and adaptability, allowing students to engage with coursework both online and offline. Intelligent web-based frameworks now monitor teacher and student behavior, enabling real-time adjustments to enhance the learning experience(Rashid & Kausik, 2024).

In e-commerce, Amazon exemplifies AI integration through innovations such as voice shopping, product recommendations, facial recognition, home price estimation, visual search, and autonomous driving. Its AI-powered assistant, Alexa, facilitates voice-activated payments, while

Amazon Go, a checkout-free grocery store, leverages AI to track selected items and analyze shopper behavior in cities like Chicago, New York, San Francisco, and Seattle (Rashid & Kausik, 2024).

Despite the increasing demand for AI adoption in business, challenges persist, with approximately 85% of AI programs reportedly failing to achieve their objectives. Nevertheless, several e-commerce companies, including Clarifai, eBay, Anaplan, IBM, Emotive.io, inVia Robotics, and Alibaba, have successfully implemented AI to enhance various aspects of their operations.

Clarifai utilizes AI-driven image and video recognition for content moderation, while eBay employs AI to improve customer support, streamline shipping, and strengthen buyer-seller trust. IBM's Watson enables retailers to offer personalized shopping experiences using real-time consumer data, and Emotive.io enhances interactive conversational marketing through SMS-based customer engagement(Rashid & Kausik, 2024).

According to Rashid & Kausikby by the end of 2023, an estimated 73% of e-commerce transactions were projected to be completed via mobile devices, enhancing the mobile shopping experience. AI-driven applications such as Shopify, Square, WooCommerce, ShipStation, and eBay continue to shape consumer behavior. While AI offers significant advantages in e-commerce, overcoming implementation challenges will be essential to maximizing its impact on the sector.

The engineering industry is increasingly focused on artificial intelligence (AI), with ongoing discussions surrounding its benefits, challenges, risks, and long-term implications. While the precise trajectory of AI remains uncertain, there is a clear trend toward greater regulatory oversight. Lawmakers, particularly in the European Union, are expected to introduce legal measures to establish clearer governance frameworks for AI technologies (Rashid & Kausik, 2024).

Despite ongoing concerns, AI's transformative potential across industries is undeniable. Its rapid advancement is driving innovation in healthcare, finance, and transportation. AI is enhancing disease detection, personalized medicine, fraud prevention, market analysis, and autonomous vehicle technology. Additionally, it enables tailored healthcare interventions through genetic analysis. As AI continues to evolve, it is set to reshape industries, influence daily life, and significantly impact the global economy soon.

1.3 Challenges when Implementing AI

As artificial intelligence (AI) continues to evolve and integrate further into our daily lives, it is crucial to address the challenges and risks it presents. Ensuring that AI is developed and implemented responsibly, ethically, and for the benefit of all requires collective effort and collaboration (Rashid & Kausik, 2024)

Table 1 highlights the current challenges facing AI along with possible solutions. Addressing these challenges is essential to ensuring a promising future for AI. Key concerns include ethical issues, biases, and the need for continuous investment in AI development.

Table 1

Potential Challenges and Solutions for the Future of AI *(Rashid & Kausik, 2024)*

Challenges	**Potential Solutions**
Bias in AI models	Improved data collection and algorithm design
Lack of transparency	Increased transparency and explainability in AI
Job displacement	Investment in education and job training programs
Ethical concerns	Development of ethical frameworks and regulations
Security concerns	Development of more secure AI systems
Data privacy	Increased data privacy regulations and standards
Regulation and control	Development of regulatory frameworks for AI

While the rise in AI presents several obstacles and potential risks, it also offers significant benefits. By proactively tackling these challenges, we can help AI continue to drive innovation, advance various industries, and enhance lives globally (Rashid & Kausik, 2024).

Aside from these, according to (Rashid & Kausik, 2024) some other common challenges of AI and its future could be:

- **Data Privacy and Security:** A major challenge for AI is maintaining data privacy and security. Since AI systems rely on vast amounts of data to function, there is a risk of misuse or theft, which could result in privacy breaches or security threats. (Rashid & Kausik, 2024).
- **Bias, Fairness, and Ethics:** The concepts of bias, fairness, and ethics in AI technologies and systems remain insufficiently defined, and their functioning is not always well

understood. While technology itself may operate correctly and ethically, its fairness largely depends on the quality and impartiality of the data used for training.

If the input data contains biases, the AI system will also reflect these biases, leading to non-transparent decision-making (Rashid & Kausik, 2024) This can result in unfair outcomes, such as discrimination in hiring or loan approvals.

- **Lack of Transparency:** AI systems are often considered black boxes, the idea or logic of how they arrive at their decision or come to solve a problem is not clear yet to the researchers(Rashid & Kausik, 2024).
- **Lack of Skilled Professionals:** There is a lack of skilled professionals capable of designing and deploying AI systems in industrial environments. This talent gap could hinder AI development and delay its widespread adoption (Rashid & Kausik, 2024).

- **Existential Risk:** Some experts have expressed concerns about the possible existential threats posed by advanced AI, including the risk of an AI system becoming excessively powerful or beyond human control. (Rashid & Kausik, 2024).

Addressing these challenges demands a combined effort toward developing effective solutions. This can be achieved through increased investment in research and development, the enhancement of data privacy and security regulations, and the promotion of transparency within AI systems. These measures will contribute to the ethical and responsible advancement of AI technologies.

1.4 Threats of AI

As artificial intelligence (AI) progresses, it is essential to evaluate both its potential advantages and associated risks. Experts offer varying perspectives on the future implications of AI, with figures like Raymond Kurzweil from Google predicting that it will augment human intelligence, whereas others, such as Elon Musk, caution about the potential for disastrous outcomes.(Rashid & Kausik, 2024).

Addressing these issues effectively remains a challenge, primarily due to the rapid pace of AI development and the limited understanding of its complexities among policymakers, who ultimately influence AI regulation (Rashid & Kausik, 2024).

Table 2, as noted by Rashid & Kausik (2024), outlines the most common and recent AI threats and corresponding mitigation strategies. frameworks These challenges highlight the importance of robust research methodologies and collaboration to address AI's impact across different sectors. Although AI presents opportunities for innovation and growth, it also brings ethical, social, and economic challenges.

Table 2

Threats of AI and potential solutions *(Rashid & Kausik, 2024)*

Threat	**Potential Solutions**
Job Displacement	Education and training for new job skills
Malicious Use	Robust cybersecurity measures and ethical guidelines
Autonomous Weapons	International agreements and regulations banning their use
Superintelligence	Responsible development and regulation of AI technologies
Inequality and Bias	Diverse and representative data and ethical guidelines
Surveillance and Control	Strong privacy regulations and oversight mechanisms
Lack of Accountability	Transparent decision-making processes and human oversight

To optimize its advantages while minimizing potential risks, continuous dialogue and cooperation among researchers, politicians, and industry leaders are essential. Considering AI's profound effect on society, its development must be approached with responsibility and a strong ethical focus, ensuring it benefits humanity.

1.5 Ethical Implications

The implementation of artificial intelligence technologies raises several ethical and societal concerns, particularly related to algorithmic bias, data privacy, and the potential displacement of workers. Addressing these issues requires the development of robust governance frameworks for AI, the establishment of strict data privacy regulations, and the provision of retraining programs for employees impacted by technological changes. A thoughtful and responsible approach to the deployment of AI is essential to minimize its negative effects on society.

Ensuring the responsible development and application of AI demands addressing ethical concerns from the outset of AI creation. Developers and engineers must actively work to eliminate biases present in data and algorithms to prevent discriminatory practices.

Transparency in AI systems is essential, allowing users to understand the reasoning behind decisions made by AI agents. Furthermore, implementing mechanisms for error correction and providing recourse for users is vital for maintaining accountability. Protecting user privacy through robust data protection protocols and securing informed consent for data usage are fundamental ethical imperatives. Compliance with legal and ethical standards is critical for fostering public trust and reducing the potential harm caused by AI technologies. Ethical considerations play a crucial role in guiding the responsible integration of AI into society, ensuring fairness, transparency, accountability, and privacy(Hayawi & Shahriar, 2024).

The widespread integration of AI across various sectors has the potential to bring both positive and negative societal impact. On the positive side, AI technologies can contribute to environmental sustainability by optimizing energy consumption through smart grids and enhancing resource management, thereby reducing carbon emissions(Hayawi & Shahriar, 2024).

AI-powered predictive models also hold the promise of improving disaster preparedness and response, potentially mitigating the adverse effects of natural disasters, especially those influenced by climate change. However, the proliferation of AI also raises significant concerns, particularly regarding economic inequality, job displacement, and the equitable distribution of AI-driven benefits. These challenges underscore the need for comprehensive regulatory frameworks to ensure that AI benefits are shared fairly across all societal groups(Hayawi & Shahriar, 2024).

To ensure the ethical and fair use of AI, several measures must be adopted. Implementing techniques to mitigate bias within AI algorithms and datasets is crucial to achieving fairness. Transparency in AI decision-making should be prioritized to improve user understanding and allow for the explanation of AI-generated outcomes.

Strong data security protocols are necessary to protect user privacy, while accountability frameworks must hold developers and organizations responsible for the ethical deployment of AI systems. Regular assessments of the impact of AI technologies on different demographic groups should be conducted to identify and address disparities, ensuring equitable outcomes(Hayawi & Shahriar, 2024).

Furthermore, leveraging advanced deep learning architectures and large, diverse datasets can enhance the performance of AI agents, enabling them to effectively tackle a wide range of tasks while maintaining ethical standards in their operation(Hayawi & Shahriar, 2024).

The ethical and societal implications of AI require a proactive and structured approach to governance, regulation, and responsible deployment. Addressing biases, ensuring data privacy, and mitigating workforce disruptions are critical to fostering equitable AI integration.

Transparent decision-making, robust security measures, and ongoing impact assessments must be prioritized to uphold ethical standards. By implementing well-defined policies and fostering collaboration among policymakers, researchers, and industry leaders, AI can be harnessed to drive innovation while minimizing potential risks. A balanced approach that prioritizes both technological advancement and societal well-being is essential for the sustainable and ethical adoption of AI.

2.1 Machine Learning

Machine Learning (ML) is a subfield of computer science that sits at the intersection of mathematical and statistical techniques and computational algorithms. It utilizes algorithms rooted in Artificial Intelligence (AI) concepts, applying them to specific situations to identify patterns within sets of variables. The primary goal is often to predict a particular result of interest.(Miana et al., 2022).

2.1.1 Distinguishing ML from Traditional Systems

Most conventional techniques found in computer systems applied to medicine rely on rule-based algorithms, often referred to as “expert systems.” In these systems, developers manually encode existing medical knowledge regarding a subject using predefined rules. ML techniques, however, operate differently.(Miana et al., 2022) They are adept at handling many variables and seeking out new combinations that can reliably predict an outcome. This capability is particularly powerful when dealing with a high volume of data, such as in Big Data scenarios.

The concept of **Big Data** itself was defined by Doug Laney in 2001 using the “3 Vs” model: high volume, high velocity, and high variety of information. These characteristics necessitate new processing techniques to enable discoveries and optimize processes. The term Big Data can refer to either an enormous dataset that traditional data management tools cannot efficiently store or process, or it can describe the type of technology like storage facilities, tools, and processes used(Miana et al., 2022).

The Machine Learning Algorithm Development Process

According to Miana et al. the development of an ML algorithm is a structured process typically divided into three distinct phases: preprocessing, training, and model evaluation.

1. **Preprocessing**: This initial phase involves organizing the databank, clearly defining the research question that the ML model will address and partitioning the data into separate sets for training and testing the algorithm.
2. **Training**: During this phase, the algorithm learns from the data. The main difference between supervised and unsupervised learning models lies in this training algorithm.
3. **Model Evaluation**: In the final phase, the trained model's performance is assessed by comparing its outputs against the reserved test data, and the results are generated.

In **unsupervised learning**, the ML model extracts data characteristics and builds a representation without prior knowledge of the labels for each piece of data. It identifies information classification patterns heuristically. This lack of supervision can be advantageous, as it allows the algorithm to analyse and identify patterns that may not have been previously considered by human experts. In **supervised learning**, the ML model has access to data labels; the training samples are already correctly classified. The training is based on an iterative comparison between the results obtained from the model and these previously classified labels(Miana et al., 2022).

This process continues until a minimum error rate is achieved(Miana et al., 2022). ***Table 3*** summarizes the main characteristics of each type of learning model, as well as their advantages, disadvantages, and practical applicability.

Table 3

Comparison between supervised and unsupervised learning processes *(Miana et al., 2022)*

	Supervised learning	**Unsupervised learning**
Definition	Algorithms that learn relationships between input and output attributes based on a set of labelled examples.	Algorithms that attempt to find patterns in data clusters with similar characteristics, looking for unidentified or uninformed categories and outcomes.
Advantages	Analysis of multiple parameters; quick, automatic solution for large-scale questions and high accuracy.	Less human interference in data analysis; excellent for multimodal or multidimensional data sources; allows identification of new outcomes.
Disadvantages	Requires data to be labelled; may be impractical for large volumes of data. Tendency to overfit data.	High cost; complex techniques. It requires a large amount of data to elaborate the algorithm, and it can be challenging to interpret the results.
Main tasks	Regression, classification, prognostic model, and survival analysis.	Reducing dimensionality of the problem and grouping.
Examples of algorithms	Logistic regression, decision trees, random forests, and artificial neural networks.	Principal component analysis, hierarchical clustering, autoencoders, and linear discriminant analysis.

Essentially, ML algorithms learn through repeated observations. They work to establish a mapping pattern that allows them to label data and, critically, to create a model that generalizes the learned information. This generalization enables the model to accurately and reliably label new data that it has never encountered before(Miana et al., 2022).

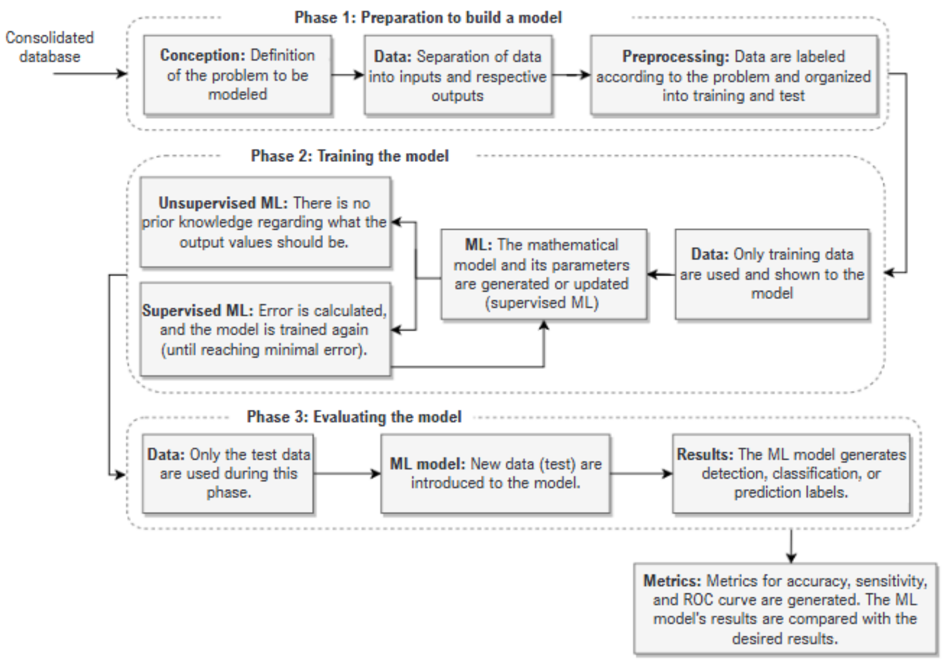
2.1.2 Common Machine Learning Techniques

Several Machine Learning (ML) techniques have found significant application as forms of computer-assisted diagnostic systems. These include prominent methods such as Artificial Neural Networks (ANNs), logistic regression, decision trees, random forests, Bayesian networks, deep learning, and Support Vector Machines (SVMs), among others. Some of these techniques utilize mathematical models by means of data for the purpose of learning and/or the organization of information. Others apply mathematical representations characterized by a high degree of abstraction, often involving complex mathematical models(Miana et al., 2022).

In such cases, it may not be possible to decipher or interpret the precise methods used by the algorithm to obtain its prediction, detection, or classification results; these ML models are, thus, commonly known as “black box” models(Miana et al., 2022).

Figure 2

Phases for developing machine learning algorithms *(Miana et al., 2022)*



Among these diverse approaches, **Artificial Neural Networks (ANNs)** represent a computational and mathematical model developed to function in a manner analogous to the human brain. An ANN is structured with several interconnecting elements, typically organized into a predictor layer, one or more hidden layers, and an output layer. The relationships and signal flow between these layers are inspired by the synaptic connections observed between biological neurons(Miana et al., 2022).

An ANN "learns" by means of these intricate connections between its layers, as well as through the adjustment of numerical weights associated with each connection. When input data is introduced in the predictor layer, it is propagated layer by layer. Mathematical processing takes place as data moves from one layer to the next, and the weights of these connections are dynamically updated according to the error identified in the result layer specifically, the discrepancy between the expected result and the one obtained by the network. This adaptive process is repeated iteratively until the error value is minimized to an acceptable level or until a pre-specified number of interactions has been completed(Miana et al., 2022).

Building upon the foundational architecture of ANNs, deep learning differs from more traditional ML techniques in that it processes more robust computational models equipped with multiple processing layers. Thus, the technique of deep learning works in accordance with ANN principles but is distinguished by a greater number of hidden layers and, consequently, a more extensive network of synaptic connections. Within a deep learning model, each layer reproduces a representation of data from the previous layer, allowing for the learning of increasingly complex features. The learning algorithm employed in deep learning can be either supervised or unsupervised(Miana et al., 2022).

A specialized application of neural network principles, particularly relevant given the high data volume and complexity involved in working with big data, is the autoencoder algorithm. This is a type of ANN designed specifically to reduce data dimensionality(Miana et al., 2022).

To achieve this, the autoencoder uses mathematical models with a high degree of abstraction to generate a new dataset that has reduced dimensionality while striving to maintain a representation as close as possible to the original input data. A fundamental difference between a standard ANN and an autoencoder is that the latter typically uses unlabelled data during its training phase(Miana et al., 2022).

Moving to other widely adopted paradigms, the decision tree algorithm is most often used when the dataset is relatively small. This algorithm is developed with a series of hierarchical yes/no questions designed to classify data into various categories. It utilizes a statistical model for data classification or prediction, employing the idea of nodes(Miana et al., 2022).

Each node represents a question, which then divides into possible outcomes these branches, in turn, lead to other possibilities, and this process is repeated until an outcome or classification is achieved. The primary advantages of this algorithm are its inherent simplicity and the intuitive nature of its interpretation, making it relatively easy to understand the decision-making process(Miana et al., 2022).

An extension of the decision tree approach is found in random forests, a technique widely used to solve both classification and regression problems. Random forests operate by combining multiple decision trees, where each individual tree is trained independently on different subsets of the data or features. Its main distinguishing features include a relatively simple underlying theory, the capacity for rapid data analysis, notable stability even in the presence of excessive noise within the data, and an automatic compensation mechanism that can help to mitigate the impact of biased data samples(Miana et al., 2022).

2.1.3 Limitations of Machine Learning

Machine learning (ML) models, while powerful, are not without their shortcomings, especially when applied to complex social systems. The author Malik proposes a "hierarchy of limitations" that arise from fundamental choices made when adopting an ML approach. Understanding these limitations is crucial for both developers and consumers of ML technologies to identify potential failure points and make informed decisions about their application.

1. The Constraints of Quantification

The initial decision to use quantitative analysis, the foundation of ML, introduces several limitations. Firstly, ML, as a quantitative method, inherently struggles to capture or quantify meaning making, which is often central to understanding human behaviour and social contexts. Qualitative approaches are typically better suited for exploring these nuanced aspects.

Secondly, ML models operate using proxy's measurable variables that stand in for the actual, often more complex, constructs of interest. For instance, a model might use healthcare costs as a proxy for illness. If these proxies are flawed or do not accurately map to the real-world construct, the model's outputs can be misleading, even if technically accurate on the proxy data. This can lead to issues like the "*quantitative fallacy,"* where reliance on what is measurable excludes other critical, unquantifiable factors(Malik, 2020).

Furthermore, an exclusive reliance on quantitative ML outputs risks delegitimizing lived experiences, particularly if model conclusions contradict the perspectives of individuals or marginalized communities who may lack the quantitative *"proof"* demanded by such systems. Finally, the concept of performativity highlights that ML models may not merely describe the world but can actively reshape it to fit their own logic. This is a significant limitation if the goal is objective understanding or if the induced changes are detrimental(Malik, 2020).

2. Assumptions in Probability-Based Modelling

Modern machine learning is predominantly statistical and probability-based, leading to further inherent limitations. It typically relies on structuring data into a data matrix. This framework assumes the world consists of *"fixed entities with properties,"* an assumption that can be at odds with sociological theories emphasizing the fluidity, context-dependency, and narrative nature of social phenomena(Malik, 2020).

Treating complex social identities like race or sex as mere variables in a matrix can oversimplify and misrepresent them. Moreover, ML's optimization processes are geared towards the central tendency. Predictions and classifications are therefore representative of group characteristics rather than truly individual outcomes. This approach is *"fundamentally unable to do anything with one observation of one entity"* in isolation and can disproportionately affect or misrepresent outliers(Malik, 2020).

3. Challenges of a Predictive Focus

The common emphasis in Machine Learning (ML) on achieving high predictive accuracy, often at the expense of providing causal explanations, introduces a distinct set of limitations. One primary issue is that prediction is not synonymous with explanation. ML models excel at identifying correlations within data to make predictions; however, they do not inherently uncover or detail the causal mechanisms driving these correlations. This fundamental gap limits their utility when the goal is to design interventions aimed at changing outcomes, as the *"levers"* of cause and effect remain obscure(Malik, 2020).

Furthermore, while techniques for *"interpretable"* or *"explainable AI"* are developing, it's crucial to understand the distinction between interpretability and true causal explanation. These methods can illuminate a model's internal workings or highlight the features it deems important for its predictions. However, this insight into model behaviour does not necessarily equate to an accurate explanation of real-world causal processes. Users might incorrectly infer causality from these model explanations, leading to flawed conclusions and misguided actions(Malik, 2020).

Another significant challenge is the fragility of correlations. Predictions based on non-causal correlations are inherently brittle. They might perform well in stable, static environments but can easily falter when the underlying system changes or when the model encounters *"distribution shifts* "new patterns of data do not present during its training phase(Malik, 2020).

The reliance on proxies also makes ML models vulnerable to gaming the proxies. If the specific variables (proxies) an ML model uses for prediction become known, and if there are significant stakes attached to the model's outcomes, individuals or systems may seek to manipulate these proxies to achieve favourable results. This phenomenon is well-described by Goodhart's Law: *"When a measure becomes a target, it ceases to be a good measure.* "Finally, by learning from historical data, ML systems risk optimizing for and perpetuating the status quo. If the data used for training reflects existing societal biases or inequalities, the model is likely to learn these patterns and may inadvertently reinforce or even amplify them in its predictions and applications, rather than challenging or helping to overcome these existing issues(Malik, 2020).

4. The Fallibility of Cross-Validation

The primary method for assessing ML model performance and generalizability within academic research is cross-validation.

However, its reliability is subject to several limitations:

- **Publication Bias:** The tendency in scientific publishing to favour positive results means that the reported successes of ML models might be an overestimation, as studies with negative or null findings from cross-validation are less likely to be published.
- **Overfitting to the Test Set:** Through repeated use of common benchmark datasets or in competitive environments, models can inadvertently become overfit to the specific test set. This makes the cross-validation results for that set overly optimistic regarding performance on genuinely new, unseen data.

- **Data Dependencies:** Standard cross-validation often assumes that data points are independent and identically distributed. However, real-world data frequently contains dependencies. If these dependencies are not properly handled in the data splitting process for cross-validation, the training and test sets will not be truly independent. This typically leads to the test error being an "overly optimistic" estimator of the model's true out-of-sample error.

The limitations outlined by Malik, from the foundational act of quantification to the specifics of model evaluation, underscore the necessity of a critical and informed approach to machine learning. Recognizing that ML models are not infallible tools but are embedded within specific conceptual frameworks and procedural choices is the first step.

By understanding these inherent constraints, ML practitioners can better mitigate potential failures, communicate results more honestly, and consider alternative or complementary methodologies to ensure that the application of machine learning is both effective and responsible.

2.2 Generative AI

According to Feuerriegel generative AI refers to computational techniques that are able of generating new content such as text, images, or audio from training data. The extensive flow of this technology with examples such as Dall-E 2, GPT-4, and Copilot is currently revolutionizing the way we work and communicate with each other(Feuerriegel et al., 2024). Generative AI systems can not only be used for artistic purposes to create new text mimicking writers or new images mimicking illustrators, but they can and will assist humans as intelligent question-answering systems (Feuerriegel et al., 2024)

As I mentioned before, generative AI has emerged as a transformative force across multiple industries, offering applications in content creation, business, education, healthcare, media, and gaming. However, its adoption also presents challenges related to misinformation, bias, ethical concerns, and security risks, necessitating governance and regulatory frameworks to ensure responsible usage(Fui-Hoon Nah et al., 2023).

Generative AI in our world plays a fundamental role in content creation, powering tools such as ChatGPT for text generation, Midjourney for image creation, and Deep Brain for video synthesis . In the business sector, AI enhances decision-making, automates repetitive tasks, and improves forecasting in industries such as marketing, healthcare, and finance.

The education sector benefits from AI-driven personalized learning, intelligent tutoring, and AI-assisted research. In healthcare, generative AI supports medical diagnosis, provides public health advisory services, and offers therapeutic recommendations(Fui-Hoon Nah et al., 2023).

Additionally, according to Fui-Hoon Nah AI is revolutionizing media, and journalism is revolutionizing by generating automated news articles, deepfake videos, and synthetic content. In the gaming industry, AI is used to create dynamic game plots, design characters, and enhance real-time 3D scene rendering.

2.2.1 Hallucination in Generative AI Models

Hallucinations refer to instances where AI systems generate information that is incorrect, misleading, or entirely fabricated. Despite appearing to be realistic statements, these outputs may not align with reality or verified facts. Various factors contribute to hallucinations, including biases in the training data, misinterpretation of context, or the inherent difficulties in modeling complex human language(Arsal et al., 2024). In essence, when an AI "hallucinates," it produces content that seems plausible or relevant but is not based on accurate information, raising concerns about the reliability and trustworthiness of AI-generated content(Arsal et al., 2024).

Ensuring the accuracy and dependability of information provided by chatbots is crucial, even beyond privacy considerations. These large language models (LLMs) are not entirely reliable, and users cannot fully trust the information they provide. Additionally, the misuse of AI-generated content can exacerbate privacy risks, highlighting the need for improved accuracy and integrity in AI systems(Arsal et al., 2024).

Hallucinations can lead to inaccurate data or fabricated content, which is particularly concerning as these tools become more integrated into various applications. For Google Gemini, which relies on pattern recognition and predictive analysis, combating hallucinations is essential for ensuring output accuracy, particularly in organizational security contexts. Likewise, Microsoft Copilot may generate faulty code suggestions that could introduce significant vulnerabilities. As the risk of misinformation grows with the widespread adoption of AI, improving the reliability of AI-generated content is crucial.

2.2.2 Challenges and Risks

While generative AI offers significant advancements across industries, its adoption is accompanied by a range of challenges and risks. One of the most critical concerns is the generation of misleading or entirely fabricated content, commonly referred to as AI hallucinations. This issue raises questions about the reliability and trustworthiness of AI-generated outputs, particularly in fields that require high levels of accuracy, such as healthcare, law, and journalism (Fui-Hoon Nah et al., 2023).

Generative AI models often inherit biases present in their training datasets, which can result in unfair or discriminatory outputs. These biases are especially problematic in decision-making processes, such as recruitment, financial assessments, and legal judgments, where impartiality is crucial(Fui-Hoon Nah et al., 2023).

Ethical considerations further complicate AI implementation, such as AI-generated content may include inappropriate material, privacy violations, or be leveraged for harmful activities such as deep-fake creation and misinformation dissemination(Fui-Hoon Nah et al., 2023).

Legal and intellectual property concerns also emerge as generative AI continues to produce content that may resemble copyrighted materials. The ambiguity surrounding ownership and copyright attribution presents challenges for businesses and content creators seeking to use AI-generated content without infringing on existing rights(Fui-Hoon Nah et al., 2023).

Furthermore, the widespread automation of tasks through AI raises concerns about labor market disruptions. While AI enhances efficiency, it also threatens job security, particularly for low-skilled workers whose roles are more susceptible to automation(Fui-Hoon Nah et al., 2023). This shift in employment dynamics necessitates workforce adaptation, emphasizing the need for reskilling and upskilling initiatives to ensure individuals remain competitive in an AI-driven economy.

Another pressing issue is data security and privacy because as generative AI interacts with users and processes vast amounts of data, there is an increased risk of sensitive information being exposed or misused. Ensuring data protection and implementing robust security measures is essential for maintaining trust in AI systems and safeguarding confidential business or personal information(Fui-Hoon Nah et al., 2023).

Considering these challenges, it is evident that the implementation of generative AI must be approached with caution. Addressing these risks requires a combination of regulatory oversight, ethical AI development, and enhanced transparency to maximize the benefits of AI while minimizing potential harm(Fui-Hoon Nah et al., 2023).

But even though all these challenges and risks of generative AI continue to evolve, its impact on industries and daily life will become more profound. I believe that ongoing learning and workforce adaptation will be essential to mitigate job displacement and ensure individuals remain competitive in an AI-driven economy. The era of generative AI is inevitable, and embracing this transformation is crucial for navigating the future of work and innovation.

2.3 LLMS Large Language Models

Large Language Models (LLMs) are advanced AI systems designed to predict the next word in a sequence based on patterns learned from vast corpora of human-written text. Unlike traditional rule-based systems, LLMs generate human-like responses by modeling the statistical likelihood of word combinations. While they can respond to complex queries, they do not possess true understanding or consciousness; rather, they generate outputs based on probabilistic word associations (Shanahan, 2024).

LLMs have evolved from simple statistical models to sophisticated transformer-based architecture such as GPT, BERT, and PaLM which leverage attention mechanisms to handle context over long sequences. These models have demonstrated significant capabilities in natural language processing tasks, including summarization, translation, question answering, and code generation.

Training LLMs involves gathering large-scale text datasets, followed by rigorous preprocessing to remove noise, ensure data quality, and protect user privacy (Usman Hadi et al., 2023.). The models are then trained using unsupervised learning, where they learn to predict masked or next-word tokens without explicit labels. This process demands extensive computational resources and careful parameter tuning.

2.3.1 Technical Foundations of Large Language Models

The operation of Large Language Models (LLMs) relies on several core technical components that collectively enable their performance and scalability. The process begins with **tokenization**, where input text is segmented into smaller units called tokens.

Algorithms such as Byte Pair Encoding (BPE), WordPiece, and UnigramLM allow models to handle large vocabularies and morphologically complex languages by breaking words into subword units or characters(Naveed et al., 2023.).

Since the transformer architecture processes tokens in parallel and lacks inherent awareness of sequence order, positional encoding is used to provide information about the position of tokens. While early models relied on absolute positional embeddings, modern LLMs increasingly use relative approaches. Techniques such as ALiBi (Attention with Linear Biases) and Rotary Positional Encoding (RoPE) are examples that help capture relative distances between tokens effectively(Naveed et al., 2023.).

At the heart of transformers is the self-attention mechanism, which allows the model to dynamically weigh the importance of different tokens when forming contextual representations. In encoder-decoder architectures, cross-attention enables the decoder to focus on relevant encoder outputs(Naveed et al., 2023.).

However, standard self-attention scales quadratically with sequence length, posing a challenge for long-context processing. To address this, efficient attention mechanisms have emerged, including sparse attention patterns, hardware-optimized methods like FlashAttention, and hybrid strategies such as LongT5's global-local attention and LongNet's dilated attention(Naveed et al., 2023.).

Further advancements like Position Interpolation (PI), Giraffe, YaRN, LM-Infinite, and Parallel Context Windows (PCW) enable models to handle longer sequences than seen during training without extensive retraining(Naveed et al., 2023.).

To introduce non-linearity, LLMs use activation functions such as ReLU and GeLU, with gated variants like ReGLU, GEGLU, and SwiGLU demonstrating improved performance in large models like PaLM. For stable training in deep networks, layer normalization methods like LayerNorm and RMSNorm are essential. Pre-LayerNorm is commonly used in LLMs, with alternatives like DeepNorm proposed for very deep architectures to address gradient instability(Naveed et al., 2023.).

Training modern LLMs requires distributed computing across multiple accelerators. Techniques include Data Parallelism, Tensor Parallelism, and Pipeline Parallelism—often combined into broader approaches like Model Parallelism or 3D Parallelism. Optimizer Parallelism, exemplified by ZeRO, reduces memory consumption by partitioning optimizer states and gradients(Naveed et al., 2023.).

A robust ecosystem supports the development and scaling of LLMs. Frameworks such as Hugging Face Transformers, DeepSpeed, Megatron-LM, JAX, PyTorch, TensorFlow, Colossal-AI, and BMTrain offer tools for model building, training, and optimization(Naveed et al., 2023.).

Effective data preprocessing is also critical. This involves quality filtering to remove low-quality or harmful content, deduplication to avoid overfitting, and privacy reduction techniques to eliminate personally identifiable information (PII) from datasets(Naveed et al., 2023.).

LLMs are trained using various pre-training objectives. The standard autoregressive approach, Full Language Modelling, predicts the next token based on previous ones. Other methods include Prefix Language Modelling, which allows bidirectional context in the prefix and unidirectional prediction for the suffix, and Masked Language Modelling, used by models like BERT and T5. Unified Language Modelling combines multiple objectives using flexible attention masking strategies(Naveed et al., 2023.).

Lastly, scaling laws offer insights into the relationship between model performance, size, dataset volume, and compute resources. Initial research showed that loss decreases predictably with larger models and datasets, but newer findings such as those from the Chinchilla paper suggest that optimal performance under fixed computational budgets is achieved by proportionally balancing model size and training data volume. These insights are vital for making cost-effective decisions in LLM development(Naveed et al., 2023.).

2.3.2 Key Pre-trained Models and Architectures

The development of LLMs has been marked by rapid progress, increasing scale, and architectural diversification. Early influential models established foundational paradigms, while subsequent work introduced variations, scaled existing architectures, and explored new training methodologies and data curation strategies(Naveed et al., 2023.).

Architecturally, LLMs largely follow variants of the transformer. Common types include:

- **Encoder-Decoder Models:** These models, such as T5 and its multilingual variant mT5, process input through an encoder stack and generate output using a decoder stack, leveraging cross-attention between the two. T5 pioneered the unified text-to-text framework. AlexaTM and UL2 are other examples.
- **Causal Decoder-Only Models:** This architecture, popularized by the GPT series, PaLM, BLOOM, LLaMA , Gopher, and others, consists solely of a decoder stack where attention is masked to prevent attending to future tokens. They excel at text generation and in-context learning.
- **Prefix Decoder (Non-Causal Decoder) Models:** These decoder-only models relax the strict causal masking, allowing bidirectional attention over prefix portions of the input, potentially improving understanding in certain tasks. GLM-130B uses an objective that enables bidirectional attention.
- **Mixture-of-Experts (MoE) Models:** MoE architectures, such as those used in GLaM , CPM-2, PanGu-Σ, Mixtral8x22b, Snowflake Arctic, Grok-1, and Gemini 1.5 Pro, incorporate multiple "expert" sub-networks (typically feed-forward layers) within each transformer layer. A routing mechanism directs each token to only a small subset of expert. This enables scaling model capacity significantly while keeping the computational cost per token relatively low, as only a fraction of parameters is active during inference. MoE models may also offer benefits for continual learning by potentially reducing catastrophic forgetting and allowing for the extraction of smaller, specialized sub-models(Naveed et al., 2023.).

Key milestones include T5's demonstration of transfer learning limits with its text-to-text approach ; GPT-3's highlighting of few-shot learning emergent from scale ; Chinchilla's formulation of compute-optimal scaling laws suggesting proportional scaling of model size and data ; the development of large open-source models like BLOOM and LLaMA which spurred wider research access; the push towards being multilingual ; and recent achievements in extremely long context windows (Gemini 1.5)(Naveed et al., 2023.).

Beyond general-purpose text models, specialized LLMs have been pre-trained for specific domains, leveraging curated datasets and tailored objectives(Naveed et al., 2023.).

Notable areas include:

- **Coding:** Models like Codex , AlphaCode , CodeGen, CodeT5+, and StarCoder are trained on vast amounts of source code alongside natural language, enabling code generation, completion, and understanding.
- **Scientific Knowledge:** Galactica was trained on a large corpus of scientific papers, textbooks, and databases to assist with scientific tasks.
- **Dialogue:** LaMDA was specifically pre-trained on dialogue data to improve conversational abilities, focusing on quality, safety, and being grounded.
- **Finance:** BloombergGPT was trained on a mix of general and extensive proprietary financial data, while Xuan Yuan 2.0 targeted Chinese financial contexts.

2.3.3 Fine-tuning and Alignment Strategies

Pre-trained large language models (LLMs) possess extensive world knowledge and linguistic fluency. However, they often struggle to consistently follow user instructions or comply with safety guidelines. Fine-tuning serves to adapt these models to specific tasks or desired behaviours. A specialized form of fine-tuning, known as alignment, focuses on making models helpful, honest, and harmless (HHH), ensuring they behave in accordance with human values and intentions.

Overall, fine-tuning enhances performance on downstream tasks and improves zero-shot generalization with only a modest computational cost compared to pre-training(Naveed et al., 2023.).

One prominent fine-tuning strategy is instruction tuning, where models are trained on datasets comprising instructions, optional inputs, and corresponding desired outputs. This technique enhances a model's ability to follow directions and generalize to novel tasks in a zero-shot manner. Instruction-tuning datasets can be manually curated by converting traditional NLP tasks into prompted formats or by collecting diverse human-written instructions (Naveed et al., 2023.).

Increasing the diversity and number of tasks and prompt formats generally leads to better performance, as demonstrated in models like OPT-IML and Flan. Alternatively, datasets can be automatically generated by powerful LLMs, a strategy used to train models such as Alpaca and Vicuna. More advanced approaches, like Evol-Instruct, prompt LLMs to create increasingly complex instructions, resulting in models like WizardLM and WizardCoder(Naveed et al., 2023).

Alignment with human preferences is critical for ensuring safety and usability. The dominant method is Reinforcement Learning from Human Feedback (RLHF), which involves a multi-step process:

1. **Supervised Fine-Tuning (SFT):** Models are initially fine-tuned using high-quality human demonstrations.
2. **Reward Model (RM) Training:** A separate model is trained to predict human preferences based on rankings of model-generated responses.
3. **Reinforcement Learning:** The SFT model is further fine-tuned using reinforcement learning to maximize the reward model's scores.

Prominent examples using RLHF include InstructGPT and LLaMA-2-Chat. Some variations incorporate separate reward models for helpfulness and safety or use rejection sampling instead of PPO(Naveed et al., 2023.).

According to Naveed et al due to the complexity and instability of the PPO stage in RLHF, several alternative alignment strategies have emerged:

- **Direct Alignment via SFT:** Techniques such as Direct Preference Optimization (DPO), Reward-Ranked Fine-Tuning (RAFT), Preference Ranking Optimization (PRO), and Rank Responses based on Human Feedback (RRHF) aim to align models directly using preference data, avoiding the need for separate reward models or reinforcement learning. "*Chain-of-Hindsight (CoH)"* replaces numeric rewards with textual language feedback.
- **Alignment with Supported Evidence:** To reduce hallucinations and enhance trust, models can be trained to provide cited or evidence-backed answers. This is often guided by reward models trained on human judgments of factual accuracy, as in GopherCite, WebGPT, and Sparrow.
- **Alignment with Synthetic Feedback:** To minimize the cost and scalability issues of human feedback, AI-generated feedback can be used. Methods like Constitutional AI or Reinforcement Learning from AI Feedback (RLAIF) involve LLMs critiquing outputs

based on predefined principles. Others simulate human preferences (e.g., AlpacaFarm) or use models to self-align with predefined ethical standards (e.g., Self-Align). Notably, Nemotron-4 340B achieved strong alignment performance using 98% synthetic data.

- **Prompt-based Alignment:** In some cases, LLMs can be steered toward safer or more desirable outputs through carefully designed prompts, without modifying the model's parameters. An example is self-correction prompting.

Red teaming further enhances safety by systematically probing models with adversarial prompts to elicit harmful or unintended responses. These failure cases are then used to fine-tune the model for greater robustness. Red teaming can be performed by human evaluators or by leveraging other LLMs to identify vulnerabilities(Naveed et al., 2023.).

Other fine-tuning strategies include Continued Pre-training, where fine-tuning data is mixed with a small portion of the original pre-training data to avoid catastrophic forgetting. This is particularly beneficial when working with small adaptation datasets(Naveed et al., 2023.).

Research on sample efficiency has shown that high performance can be achieved using relatively small but high-quality datasets for example, LIMA demonstrated impressive results using only 1,000 demonstrations highlighting that data quality may be more crucial than sheer quantity(Naveed et al., 2023.).

2.3.4 Augmenting LLM Capabilities

While large language models (LLMs) possess an extensive knowledge base acquired during pre-training and can adapt through in-context learning (ICL), their knowledge remains static and potentially outdated. Furthermore, these models are susceptible to hallucinations and lack access to real-time information or external functionalities. Augmented LLMs address these limitations by programmatically interfacing with external resources, such as knowledge bases and functional tools(Naveed et al., 2023.).

This augmentation allows the models to base their responses on current, factual data and to perform actions beyond mere text generation. Various architectures for external memory, including long-term, short-term, symbolic, and non-symbolic models, as well as different interaction protocols, have been explored to enhance this functionality(Naveed et al., 2023.).

A pivotal approach in this area is Retrieval-Augmented Generation (RAG), which enhances LLM output by incorporating information retrieved from external sources, such as documents, databases, or web searches, based on the input query. This retrieved context is then supplied to the LLM, enabling the generation of more accurate, timely, and verifiable responses. RAG has shown that even smaller models can be made competitive with much larger counterparts by augmenting them with relevant external information. The RAG pipeline consists of two primary components: the retriever and the language model, with the quality of the retriever playing a crucial role in overall performance(Naveed et al., 2023).

Zero-shot RAG utilizes pre-existing LLMs and retrievers, such as BM25 or frozen dense retrievers, without requiring retraining. Typically, relevant information is incorporated directly into the prompt, and techniques like iterative retrieval or dynamic retrieval triggers, such as FLARE, are employed to refine this process(Naveed et al., 2023).

For enhanced performance, RAG systems can be improved through the training of individual components. Pre-training the LLM with integrated retrieval (as in RETRO) can yield substantial improvements over RAG applied solely during fine-tuning. Additionally, focusing on the retriever, through methods like contrastive learning or using supervised signals from the LLM, can further optimize the retrieval of relevant context while keeping the LLM frozen. A more advanced approach involves jointly training both the retriever and the LLM, enabling end-to-end optimization of the entire system(Naveed et al., 2023.).

Efficiency improvements are also critical in RAG systems. Encoded context augmentation methods, such as Fusion-in-Decoder (FiD), use attention mechanisms to fuse retrieved information more effectively than simple concatenation, enabling the system to handle larger contexts efficiently. Web augmentation, which involves leveraging live web searches, ensures that the information used by the model remains up to date(Naveed et al., 2023).

In addition to knowledge augmentation, Tool-Augmented LLMs enhance the model's capabilities by enabling it to interact with external tools, such as APIs, calculators, or code interpreters. These systems rely on the LLM’s ability to reason and plan, decomposing tasks into manageable steps, selecting the appropriate tools, executing the tasks, observing the outcomes, and synthesizing a final response. This process is often iterative, allowing the model to refine its actions based on feedback(Naveed et al., 2023).

Zero-shot tool use enables LLMs to interact with external tools without the need for specific training, often guided by tool descriptions within the prompt, demonstrations (such as ART), or by interpreting API documentation. Tools like ToolkenGPT represent external tools as special tokens, facilitating integration with the model(Naveed et al., 2023).

Explicit training for tool uses further improves the proficiency of LLMs in manipulating these tools. Fine-tuning the models on datasets of tool interactions, such as through instruction tuning or self-play methods, helps models learn more effective strategies for tool interaction. In some cases, large-scale datasets of API calls or sophisticated data generation techniques, such as those used in ToolLLM, are employed to train models in tool usage(Naveed et al., 2023).

Multimodal tool augmentation extends this capability to scenarios involving multiple types of data, such as images, audio, or other modalities. In these cases, LLMs can orchestrate a variety of tools working across different input types, further broadening their applicability.

In sum, augmentation strategies significantly expand the practical utility and reliability of LLMs. By enabling these models to operate more effectively in complex, dynamic, and information-rich environments, augmentation transforms LLMs from static knowledge repositories into dynamic, context-aware systems capable of executing a wide range of tasks(Naveed et al., 2023).

2.3.5 Efficiency and Deployment Strategies

The substantial computational and memory demands associated with training and inference in large language models (LLMs) present a significant barrier to their widespread adoption. Consequently, enhancing the efficiency of LLMs, particularly in terms of reducing the resource requirements for both training and deployment, is a critical area of research(Naveed et al., 2023).

Parameter-efficient fine-tuning (PEFT) methods offer an alternative to the resource-intensive process of full model fine-tuning for adapting LLMs to downstream tasks. Rather than updating all parameters of the model, PEFT techniques focus on modifying only a small subset of parameters or adding a limited number of new, trainable parameters. This approach reduces both memory usage and training time while maintaining performance comparable to that of full fine-tuning, especially in low-resource settings(Naveed et al., 2023).

One such technique, adapter tuning, integrates small, trainable adapter modules, typically in the form of bottleneck layers, into the frozen transformer layers. Another widely used method, Low-Rank Adaptation (LoRA), updates weight matrices by learning low-rank adjustments, which can be merged with the frozen weights during inference, thereby introducing no additional latency(Naveed et al., 2023).

Prompt tuning and its variants, such as P-Tuning, optimize continuous vector embeddings (soft prompts) that are prepended to input sequences, enhancing model performance without requiring fine-tuning of the entire model. In prefix tuning, trainable prefix vectors are prepended to the attention mechanism's keys and values, with adaptive versions incorporating gating mechanisms. Bias tuning, or BitFit, fine-tunes only the bias parameters, providing an efficient approach that has proven effective in various contexts(Naveed et al., 2023).

Model compression techniques focus on reducing the size of trained models and lowering their inference costs. One such technique, quantization, reduces the precision of model weights and activations, transitioning from high-precision formats such as FP32 or FP16 to lower bit-width representations, such as INT8 or 4-bit. This reduction in precision can be challenging, particularly when large magnitude outliers emerge in activations during training.(Naveed et al., 2023)

Post-training quantization (PTQ) applies quantization after model training, utilizing methods like mixed-precision decomposition (e.g., LLM.int8()), activation smoothing (e.g., SmoothQuant), and weight optimization techniques (e.g., GPTQ and OWQ). Quantization-aware training (QAT) integrates quantization simulations into the fine-tuning process, enabling models to recover any accuracy loss that may result from the reduced precision.

Techniques such as AlphaTuning, PEQA, LLM-QAT, and QLoRA which combines 4-bit quantization with LoRA are prominent examples of this approach. Pruning, another form of model compression, involves the removal of redundant model parameters to reduce size and improve inference efficiency. Unstructured pruning removes individual weights based on importance metrics, such as the magnitude of the weight multiplied by the activation norm, as seen in methods like Wanda(Naveed et al., 2023).

This type of pruning does not typically require retraining. In contrast, structured pruning removes entire groups of parameters, such as attention heads, neurons, or entire layers. This approach results in improved hardware utilization and faster inference speeds. Examples of structured pruning methods include LLM-Pruner and SIMPLE, which utilizes learnable masks, as well as techniques that leverage matrix factorization(Naveed et al., 2023).

The efficiency techniques discussed are essential for enabling the deployment of powerful LLMs across a wider range of applications and hardware platforms. By reducing the resource demands of these models, these strategies make it possible to use LLMs in more diverse settings without compromising their performance.

2.3.6 LLMS Across Sectors

Large Language Models (LLMs) have emerged as transformative general-purpose tools, demonstrating utility across diverse fields due to their capacity for understanding and generating human-like text. Their adaptability allows them to perform tasks ranging from summarization, translation, and content creation to coding assistance, drafting communications, scheduling, automating customer service, and analysing large text datasets(Naveed et al., 2023).

The field is rapidly expanding into **Multimodal Large Language Models (MLLMs)**. These models integrate information from various modalities beyond text, such as images, video, and audio, aiming for a deeper contextual understanding akin to human perception. MLLMs like Flamingo, BLIP-2, MiniGPT-4, and Gemini are constructed through various strategies, including end-to-end pre-training with cross-modal fusion mechanisms, fine-tuning pre-trained LLMs with multimodal instruction datasets, or prompting techniques adapted for multimodal inputs. MLLMs enable applications like visual question answering, image captioning, and complex visual reasoning tasks(Naveed et al., 2023).

Furthermore, LLMs are increasingly functioning as the cognitive core or *"brain"* for LLM-Powered Agents. These are autonomous systems capable of planning, decision-making, and performing sequences of actions to achieve complex goals. LLMs enable agents to understand instructions, reason about tasks, generate plans, interact with tools, utilize memory, and adapt based on environmental feedback(Naveed et al., 2023).

Key components include sophisticated Planning and Reasoning, Feedback mechanisms for closed-loop control, Memory systems, and Multi-Agent Systems where multiple LLMs collaborate(Naveed et al., 2023).

Agents are being developed for web navigation, coding, and interaction with the physical environment, requiring grounding in real-world knowledge and affordances (e.g., SayCan, PaLM-E) for tasks like robotic manipulation and navigation(Naveed et al., 2023).

In specific sectors:

- **Medicine:** LLMs assist in clinical decision support, enhance patient interaction via chatbots, accelerate medical research by summarizing literature, aid medical education through simulations and feedback, and support public health initiatives.
- **Education:** They enable personalized learning experiences, support teachers with content generation and grading, facilitate language learning, and improve accessibility for students with disabilities.
- **Science:** LLMs expedite research by summarizing literature across domains, assisting hypothesis formulation, and improving scientific writing clarity.
- **Mathematics:** LLMs can provide step-by-step problem-solving assistance, help verify proofs and make complex concepts more accessible.
- **Law:** Applications include thematic analysis of legal documents, generating explanations of legal terms (potentially augmented with case law), legal reasoning, and question answering.
- **Finance:** Domain-specific models like BloombergGPT or open-source alternatives like FinGPT show strong performance on financial tasks, enabling applications like robo-advising and algorithmic trading. LLMs can also decompose complex financial tasks.
- **Robotics:** As mentioned with agents, LLMs aid in human-robot interaction, task planning, motion planning, navigation, and manipulation, enabling robots to understand environments and learn continuously.
- **Coding:** Beyond specialized models, general LLMs assist programmers with code generation, debugging, explanation, and translation between programming languages.

The expanding application scope underscores the versatility of LLMs, although responsible deployment necessitates careful consideration of their limitations and societal impacts(Naveed et al., 2023).

2.3.7 LLMS Challenges and Limitations

Despite their remarkable capabilities and rapid evolution, Large Language Models (LLMs) face a wide array of technical and ethical limitations that constrain their effective and responsible deployment. One of the foremost issues pertains to the immense computational resources required for training and inference, which not only result in substantial financial and environmental costs but also reinforce economic disparities by privileging organizations with extensive infrastructure(Naveed et al., 2023).

Furthermore, LLMs often inherit and potentially amplify the biases embedded within their training data, leading to outputs that may perpetuate stereotypes or result in discriminatory behaviour. This challenge underscores the pressing need for robust bias mitigation strategies.

In terms of reasoning, while LLMs exhibit emergent cognitive-like abilities, they frequently struggle with tasks that demand rigorous logical deduction, complex planning, or deep common-sense understanding, largely due to their reliance on statistical correlations rather than causal inference. Additionally, the phenomenon of *"hallucinations"* where models produce outputs that are factually incorrect, internally inconsistent, or contradictory to real-world knowledge poses significant risks, particularly in high-stakes contexts(Naveed et al., 2023).

This problem is exacerbated by the static nature of LLMs' knowledge, which is fixed at the time of training and becomes outdated, necessitating the use of augmentation techniques like Retrieval-Augmented Generation (RAG), which itself remains an area of ongoing refinement.

Safety and controllability remain central concerns, as LLMs can inadvertently or maliciously be prompted to generate toxic, harmful, or misleading content. Ensuring consistent adherence to desired behaviour and tone continues to be a difficult endeavour.

Compounding these issues are serious security and privacy vulnerabilities: LLMs are susceptible to a range of adversarial threats including prompt injection, jailbreaking, and training data poisoning, and may even memorize and regurgitate sensitive information. Their lack of robustness against subtle adversarial inputs further threatens their reliability, especially in safety-critical applications(Naveed et al., 2023).

Another critical limitation lies in the opacity of LLMs' internal workings. The high-dimensional, complex architectures render their decision-making processes largely inscrutable, thereby hindering trust, explainability, and accountability. Scalability for real-time or edge deployment presents yet another challenge, as inference latency and hardware constraints limit practical use, particularly for the largest models. Moreover, these models struggle with maintaining coherent responses over extended interactions, reflecting persistent difficulty in modelling long-range dependencies(Naveed et al., 2023).

The issue of catastrophic forgetting also emerges during fine-tuning, where adaptation to new tasks or domains may erode previously acquired knowledge, undermining multi-task learning and continuous adaptation.

Multimodal integration further compounds the complexity, as fusing disparate data types like text, image, audio, video introduces challenges in alignment, computational efficiency, and architectural compatibility. Additionally, the sensitivity of LLMs to prompt phrasing necessitates intricate and often non-trivial prompt engineering to elicit desired outputs(Naveed et al., 2023).

Finally, the swift explosion of LLM technologies requires the urgent development of comprehensive regulatory and ethical frameworks. These should address questions of accountability, auditability, societal impact, and equitable access, thus guiding the responsible and transparent integration of LLMs into real-world applications. Tackling these complex challenges remains essential to unlocking the full benefits of LLMs while mitigating the risks associated with their widespread use.

2.4 AI Agents

AI agents have become pivotal in modern business technology, functioning independently to accomplish specific objectives. These intelligent systems interpret their surroundings, analyze data, and make informed decisions. They gather information through various means, including sensor inputs, computational processing, and analytical techniques.

The evolution of AI in the workplace is characterized by a shift from passive assistants to proactive entities. As Hayawi & Shahriar (2024) distinguish, while *"Copilots"* function as supportive tools that require human initiation for tasks, the next generation of *"AI Coworkers"* possesses a higher degree of agency, capable of executing complex workflows autonomously.

This transition introduces critical challenges regarding trust and accountability. Unlike simple automation, autonomous agents operating as coworkers raise ethical questions about the *"agency"* of the software specifically, determining liability when an AI agent makes an incorrect decision without direct human oversight. Understanding this distinction is vital for analyzing the current implementation of Microsoft Copilot, which sits at the intersection of these two categories."

2.4.1 Types of AI Agents

According to Hayawi & Shahriar (2024) AI agents can be categorized into distinct types based on their level of complexity and decision-making capabilities. These categories range from simple reactive agents to advanced entities capable of learning and autonomous reasoning.

The categories of AI agents include:

- **Simple reflex agents** operate based on immediate environmental perceptions, responding to stimuli without retaining past information. Examples include barcode readers and automated vacuum cleaners(Hayawi & Shahriar, 2024).
- **Model-based reflex agents** extend this capability by incorporating an internal model of their environment, enabling them to interpret contextual changes. Self-driving cars exemplify this category(Hayawi & Shahriar, 2024).
- **Goal-based agents** function by making decisions that align with predefined objectives, such as automated shopping systems that follow a specified list(Hayawi & Shahriar, 2024).
- **Utility-based agents** further refine decision-making by evaluating multiple options to maximize efficiency, as seen in navigation systems optimizing routes or climate control systems maintaining optimal temperatures(Hayawi & Shahriar, 2024).
- **Learning-based agents** leverage experience and data to improve performance over time. These agents adapt through iterative learning, with applications such as spam filtering systems and autonomous robotic systems. Each type of AI agent plays a crucial role in advancing intelligent automation across various domains(Hayawi & Shahriar, 2024).

2.4.2 AI Agents and Their Roles

AI agents have significantly enhanced efficiency and safety across various domains, with their capabilities advancing considerably over the past decades. As a result, their role in different industries has expanded, enabling more sophisticated applications. One key application is their function as copilots, assisting in real-time data analysis and process optimization.

By automating routine tasks, AI copilots enable humans to focus on more complex problem-solving. Their applications include route planning, inventory management, supply chain forecasting, and logistics demand prediction.

Additionally, AI agents serve as coworkers, collaborating with humans to enhance productivity and facilitate teamwork. These agents contribute to tasks such as project management and data analysis while also offering personalized recommendations and assistance based on individual performance. Their integration into workflows demonstrates the growing synergy between AI and human expertise in modern work environments.

2.4.3 AI Agents Capabilities

AI agents exhibit a wide range of capabilities that enable them to operate independently and engage dynamically with their surroundings. These agents show a breakthrough in technological development, allowing for the execution of sophisticated tasks with minimal reliance on human oversight. Their functionalities extend beyond mere automation, playing a crucial role in enhancing efficiency, streamlining processes, and optimizing performance across various industries.

Key capabilities of modern AI agents include:

- **Perception:** AI agents can acquire and understand data from multiple sources, such as sensors, cameras, microphones, and text inputs. This capability relies on technologies like natural language processing (NLP), speech recognition, and computer vision to extract meaningful information from diverse inputs(Hayawi & Shahriar, 2024).
- **Reasoning:** AI agents analyse data, identify patterns, and derive conclusions using logical frameworks, probabilistic reasoning, and expert systems. These decision-making processes enable agents to assess situations and respond appropriately based on structured rules or learned patterns(Hayawi & Shahriar, 2024).
- **Learning:** Through machine learning techniques including supervised, unsupervised, and reinforcement learning AI agents continuously improve their performance. By detecting patterns and correlations within data, they refine their decision-making capabilities over time(Hayawi & Shahriar, 2024).

- **Adaptation:** AI agents are designed to adjust to dynamic environments, evolving requirements, and user preferences. Their flexibility allows them to optimize their actions, ensuring efficiency and effectiveness in changing circumstances(Hayawi & Shahriar, 2024).
- **Interaction:** AI agents can communicate with both humans and other agents using NLP, dialogue systems, and human-computer interfaces. These interaction mechanisms enhance collaboration in complex tasks, making AI systems more intuitive and user-friendly(Hayawi & Shahriar, 2024).
- **Problem-Solving:** A fundamental function of AI agents is problem-solving, which involves data analysis and the identification of optimal solutions. Problem-solving AI agents vary in complexity, from basic automata-based models that execute simple commands to proposition-based and predicate-based systems capable of solving intricate computational problems(Hayawi & Shahriar, 2024).
- **Predictive Analytics:** AI agents leverage large datasets for predictive analytics, using historical data to forecast trends and outcomes. Supervised learning is commonly applied in areas such as quality control, risk assessment, and security monitoring(Hayawi & Shahriar, 2024).
- **Automation:** AI agents streamline processes by automating repetitive and time-intensive tasks. They enhance productivity across various domains, including human resources, project management, finance, document processing, and supply chain operations. By reducing manual effort, these agents optimize business operations and decision-making(Hayawi & Shahriar, 2024).

AI agents are advanced assistants with diverse capabilities that allow them to perform a wide range of tasks across different sectors. In healthcare, they support doctors by analyzing patient images and symptoms to assist in diagnosing diseases. In finance, they monitor business operations and trigger alerts when discrepancies or issues are detected(Hayawi & Shahriar, 2024).

These agents can also engage in conversations with humans, providing quick responses to queries and assisting with online tasks, such as planning or making requests. In industrial settings, AI agents enable automated cleaning and maintenance of machinery through sensor data.

In agriculture, they assist farmers in optimizing production, while in autonomous vehicles, they play a key role in navigation. In educational environments, AI agents support teaching by delivering tailored lectures and helping clarify concepts for students(Hayawi & Shahriar, 2024).

2.4.4 AI Agents in Business

As previously discussed, AI agents are helpful in providing rapid responses and assisting with various tasks. However, it is crucial to further examine their role in the business environment, particularly in terms of their impact and potential applications.

Given that this thesis focuses on the influence of Copilot in business settings, it is important to recognize that Copilot itself functions as an AI agent. Its capabilities extend beyond simple task assistance to actively supporting professionals in their work processes.

By integrating AI agents like Copilot into business operations, organizations can enhance productivity, streamline workflows, and foster more efficient decision-making. Thus, the exploration of AI agents' influence on business practices is essential to understanding their broader implications for improving workplace performance and innovation.

One of the most transformative applications of AI in business is customer service automation. AI-powered solutions play a crucial role in Customer Relationship Management (CRM), automating interactions and improving customer experiences while reducing operational burdens.

Virtual assistants and chatbots enable businesses to handle customer inquiries efficiently, offering instant responses to frequently asked questions, resolving common issues, and providing real-time support. Companies like Amazon and Apple leverage AI chatbots to minimize wait times, enhance service efficiency, and deliver seamless customer experiences (Hayawi & Shahriar, 2024).

Beyond customer service, AI is revolutionizing market analysis and consumer behaviour predictions. AI-powered tools process vast amounts of data, including historical sales, social media activity, customer demographics, and online behaviour, to predict trends and personalize offerings(Hayawi & Shahriar, 2024).

E-commerce platforms such as Amazon and Alibaba use predictive analytics to recommend products based on consumer purchasing patterns, while streaming services like Netflix, YouTube, and Spotify tailor content suggestions based on viewing habits.

These advancements help businesses refine pricing strategies, manage inventory efficiently, and optimize marketing campaigns(Hayawi & Shahriar, 2024).

AI models also incorporate external factors such as economic trends, competitor activity, and social developments, providing organizations with a comprehensive view of market dynamics to minimize risks and maximize profitability (Hayawi & Shahriar, 2024).

Another major advantage of AI in business is automation of repetitive tasks and process optimization. AI-driven solutions streamline operations by handling tasks like document processing, data entry, and customer service inquiries, reducing manual workload and allowing employees to focus on more strategic initiatives. AI-powered analytics continuously assess data, offering recommendations for process enhancements and operational improvements(Hayawi & Shahriar, 2024).

This automation simplifies business procedures, lowers costs, and optimizes resource allocation. In marketing and sales, AI enhances targeted advertising efforts, enabling businesses to refine their strategies and increase profitability. Additionally, AI supports companies in staying competitive by adapting to market changes and fostering innovation in an increasingly dynamic business landscape (Hayawi & Shahriar, 2024).

The integration of AI agents like Microsoft Copilot into business operations has far-reaching implications. From enhancing customer interactions through automation to leveraging predictive analytics for smarter decision-making and optimizing workflows, AI is shaping the future of business strategy. Copilot exemplifies these advancements by supporting professionals in everyday tasks and facilitating digital transformation.

As AI adoption continues to grow, organizations must implement these tools strategically to maximize efficiency while ensuring they align with broader business objectives. Understanding the impact of AI agents is crucial for unlocking their full potential in fostering innovation, improving productivity, and driving long-term success.

2.4.5 Integration of AI Agents in the Workplace

According to the author Hayawi & Shahriar, over the past three decades, AI agents have fundamentally altered business operations by improving efficiency, automating repetitive tasks, and optimizing workflows.

AI technologies, including chatbots and virtual assistants, are now commonly deployed to manage customer service, providing immediate support, reducing response times, and alleviating the burden on human agents. Additionally, AI agents facilitate workflow management, optimize task allocation, and enhance team collaboration, thereby improving overall project management efficiency.

In human resources, AI can automate tasks such as candidate screening and employee training, resulting in significant time and cost savings. Overall, AI agents contribute to workplace productivity by allowing employees to focus on high-value tasks while automating routine functions (Hayawi & Shahriar, 2024).

The integration of AI agents into workplace workflows begins with identifying tasks that are repetitive and suitable for automation through AI technologies. Such automation can support business objectives, provide staff with clear guidance, and ensure data security and high-quality standards. AI-driven automation also enables continuous process monitoring, which optimizes operational efficiency across business functions(Hayawi & Shahriar, 2024).

Successful AI integration requires a collaborative approach between AI specialists and operational teams to develop tailored AI solutions that align with organizational goals. This approach involves clearly defining the roles of AI agents to ensure they complement human capabilities, rather than replacing them. Furthermore, it is essential to establish a culture of ongoing learning and adaptation, which includes providing training programs to equip employees with the skills necessary to effectively interact with AI systems(Hayawi & Shahriar, 2024).

Additionally, feedback mechanisms must be implemented to assess the impact of AI integration, enabling continuous adjustments to meet evolving business needs and technological advancements, this holistic approach ensures that AI agents contribute meaningfully to a more efficient and responsive business model.

2.5 Copilot

In the current thesis it was already what Generative AI, LLMS, AI Agents is, now we will do a deep dive on what Copilot is and what it can bring to the companies. Copilot is introduced as a daily artificial intelligence assistant designed to boost productivity and creativity across various Microsoft applications. It incorporates AI features into core Microsoft products such as GitHub, Microsoft 365, Bing, and Edge, with the goal of simplifying coding, improving workplace productivity, revolutionizing search experiences, and providing contextual support.

The Copilot experience integrates web context, work data, and real-time PC activities, while prioritizing privacy and security. It is built into Windows 11, Microsoft 365, and Edge, ensuring smooth access and continuous enhancement(Bano et al., 2025.)

According to Bano et al. (2025) research has shown mixed results regarding the potential of AI Copilots to improve productivity across various fields. In software development, GitHub Copilot has been linked to increased efficiency, with some studies suggesting it helps save time on coding tasks. In marketing analytics, AI Copilots have been utilized to support data processing and decision-making, potentially cutting down the time needed for tasks such as campaign optimization.

Knowledgeable workers, such as consultants and writers, have reported using AI Copilots to automate routine tasks, allowing them to focus more on higher-level work. According to Bano et al, (2025) research has shown that large language models, like ChatGPT 3.5 and GPT-4 have a positive impact on qualitative research by enhancing human capabilities instead of replacing them.

Their collaboration with human analysts results in improved outcomes, leveraging complementary strengths in classification and reasoning. However, these tools also raise concerns regarding data privacy, intellectual property, and the potential over-reliance on AI-generated content, which require further investigation(Bano et al., 2025.).

The existing literature on AI Copilots presents a nuanced view of their uses across various professional fields, showcasing both their advantages and challenges. In software development, GitHub Copilot is recognized for helping developers by generating code suggestions and addressing algorithmic problems.

While it boosts productivity, especially for experienced developers, it often produces faulty code and raises concerns about quality, security, and ownership (Bano et al., 2025.). Studies indicate that GitHub Copilot should be used responsibly and thoroughly tested to address these concerns, particularly the security risks(Bano et al., 2025.).

In fields like healthcare and education, AI Copilots such as GPT-4 and M365 Copilot have shown potential benefits, though the outcomes vary. In healthcare, ChatGPT-4 has effectively simplified medical jargon and answered clinical questions, but its susceptibility to adversarial prompts underscores the need for better security measures.

In education, these tools have enhanced student engagement and personalized learning experiences but concerns about their reliability and the lack of transparency in decision-making remain. Overall, the literature emphasizes the need for ongoing development and stronger safeguards to address the ethical and security concerns of using AI Copilots, especially in sensitive sectors like healthcare and cybersecurity (Bano et al., 2025.).

AI Copilots can enhance both productivity and efficiency by enabling users to complete tasks more quickly and accurately, reducing manual labor, and allowing individuals to focus on tasks requiring human creativity and decision-making (Bano et al., 2025.). In the context of AI Copilots, productivity is influenced by the specific needs and priorities of each domain, with its definition varying depending on the nature of tasks and desired results.

In software development, productivity is often centered around coding efficiency and accuracy, measured by metrics like lines of code written, debugging time, and feature implementation. AI Copilots such as GitHub Copilot greatly boost productivity by automating routine coding tasks, enabling developers to focus on complex problem-solving and design. In marketing, productivity is defined by the ability to optimize campaigns, generate creative content, and extract actionable insights from data, with tools like Google Bard and Microsoft Copilot enhancing these processes and improving engagement metrics (Bano et al., 2025.).

For knowledge work and research, productivity focuses on synthesizing information, generating reports, and making decisions efficiently, where tools like ChatGPT excel by reducing cognitive load, automating repetitive tasks, and offering insights that foster innovation (Bano et al., 2025.).

The domain-specific uses of AI Copilots highlight their versatility and effectiveness, underscoring the importance of developing customized metrics to accurately assess their impact across different professional settings.

While the productivity advantages of AI Copilots are evident, they also bring various risks and ethical concerns. Research indicates that AI Copilots, such as GitHub Copilot, can generate insecure or incorrect solutions, posing significant risks in areas where accuracy is vital(Pearce et al., 2021.). For example, a study found that about 40% of GitHub Copilot's generated code had security vulnerabilities, raising concerns about its reliability in sensitive software development tasks (Pearce et al., 2021). In addition, AI Copilots raise ethical issues regarding privacy, data governance, and intellectual property.

Since these tools often use large datasets that may include proprietary or sensitive information, their use can pose security and privacy risks (Bano et al., 2025). Moreover, the question of intellectual ownership of AI-generated content remains unclear, with concerns about potential copyright infringement (Bano et al., 2025).

Another ethical issue related to AI Copilots is the risk of de-skilling and automation bias (Bano et al., 2025).As AI Copilots take over more routine tasks, there is a concern that users may become too dependent on these systems, leading to decreased attentiveness and a decline in essential skills. This is especially concerning fields like software development, where a strong grasp of coding principles is necessary for ensuring quality control and fostering innovation (Bano et al., 2025).

Furthermore, dependence on AI-generated suggestions can result in automation bias, where users accept the AI's output without sufficient evaluation, which can amplify errors or biases within the AI models (Bano et al., 2025).

Despite the rapid development and integration of AI Copilots across various fields, there is a noticeable gap in research examining their potential to boost productivity and innovation in scientific research. From my perspective this gap in research presents an opportunity for the scientific community to explore how AI Copilots can specifically impact research productivity, efficiency, and scientific output quality.

2.6 Copilot Studio

Microsoft Copilot Studio is an AI-driven development platform that enables businesses and developers to create, customize, and manage their generative AI solutions. It seamlessly integrates with Microsoft's ecosystem, including Azure OpenAI Studio, Azure Bot Service, and Azure Cognitive Services(SEAN HOUGHTON, 2024.). Copilot Studio expands on Power Virtual Agents by incorporating advanced AI features, such as large language models and generative technologies. This advancement improves automation and decision-making while providing a low-code environment for building more sophisticated workflows.

Copilot Studio provides a comprehensive suite of built-in tools designed to assist businesses in tracking performance and identifying opportunities for refinement. By leveraging key strategies such as performance analytics, feedback loops, regular updates, and training enhancements, organizations can optimize their AI-driven solutions to better align with evolving business needs(SEAN HOUGHTON, 2024.).

One of the primary mechanisms for evaluating Copilot Studio's effectiveness is performance analytics. The platform offers built-in dashboards that enable businesses to monitor critical metrics, including response times, user engagement, and accuracy and these insights serve as a foundation for continuous improvement, allowing organizations to pinpoint areas requiring enhancement(SEAN HOUGHTON, 2024.).

2.6.1 Business Challenges Addressed by Copilot Studio

Copilot Studio is designed to address fundamental challenges encountered in business processes across a wide range of industries. One of the most pressing issues faced by organizations is operational inefficiency, often resulting from labor-intensive manual tasks that impede productivity. Copilot Studio mitigates this by enabling the creation of AI-driven copilots capable of generating automated responses, thereby allowing employees to concentrate on higher-value tasks(SEAN HOUGHTON, 2024.).

Another critical concern is the inconsistency in customer experience. Delivering high-quality interactions at scale remains a complex undertaking; however, through AI-powered automation, Copilot Studio contributes to improved service delivery by offering standardized responses, accelerating query resolution, and enabling more personalized customer engagement. Furthermore, while organizations gather extensive volumes of data, many lack the ability to convert this data into actionable insights. (SEAN HOUGHTON, 2024)

Copilot Studio addresses this limitation by transforming raw data into meaningful intelligence that supports informed decision-making. The platform also tackles resource constraints, which particularly affect organizations with limited technical capacity. Its low-code environment empowers non-technical users to design and implement tailored solutions with minimal dependence on IT resources (SEAN HOUGHTON, 2024.).

Finally, scalability remains a major challenge as businesses grow, and workflows become more complex. Copilot Studio, through its integration within the Microsoft ecosystem, offers a scalable solution that evolves with business needs without disrupting existing operations. By responding to these critical business challenges, Copilot Studio enhances organizational efficiency, supports better decision-making, and facilitates sustainable innovation in an increasingly dynamic marketplace (SEAN HOUGHTON, 2024.).

2.6.2 Copilot Studio Key Features

Copilot Studio integrates advanced functionalities that streamline the adoption of artificial intelligence and enhance business operations. By utilizing these capabilities, organizations are empowered to implement AI-driven solutions that significantly improve productivity, efficiency, and scalability.

One of the platform's central features is the ability to create customizable AI assistants. Through the Copilot Studio Agent Builder, organizations can develop tailored AI agents designed to address specific operational challenges and align with strategic goals. Furthermore, the platform offers a low-code development environment, which facilitates the creation of AI solutions with minimal technical expertise. This reduces reliance on specialized IT personnel and increases accessibility for a broader user base (SEAN HOUGHTON, 2024.).

Copilot Studio also offers seamless integration within the Microsoft ecosystem, including services such as Teams, Outlook, Power BI, and SharePoint. This ensures a unified and efficient workflow across various applications. In addition, the platform supports AI-driven automation, enabling businesses to streamline repetitive tasks and allowing employees to focus on higher-value responsibilities. Its deployment capabilities across Microsoft 365 channels further enhance accessibility and user engagement (SEAN HOUGHTON, 2024.).

Finally, Copilot Studio harnesses the power of advanced data analytics to deliver real-time insights. These insights support data-driven decision-making, thereby improving overall operational efficiency and strategic responsiveness (SEAN HOUGHTON, 2024.).

Copilot Studio incorporates advanced functionalities that streamline AI adoption and enhance business operations, using this type of feature empowers organizations to leverage AI-driven solutions for greater productivity, efficiency, and scalability.

2.7 Copilot for Microsoft 365

Microsoft introduced Copilot for Microsoft 365 in March 2023 as an addition to its suite of productivity apps, including Word, PowerPoint, and Teams. This AI functionality is powered by large language models from OpenAI, hosted in Microsoft's Azure data centers, and enhanced by data from users' documents, calendars, emails, and other information through Microsoft Graph, a unified API platform supporting Microsoft 365.

Copilot integrates with existing Microsoft 365 applications, enabling tasks such as drafting documents in Word based on PowerPoint presentations, generating meeting notes in Teams, and summarizing email threads in Outlook (Kytö et al, 2024.).

The value proposition of Copilot for Microsoft 365 centers on the potential time savings for knowledge workers in their everyday tasks. Initially, Microsoft surprised businesses with the sudden price of $30 per user per month for this add-on functionality.

Although this cost seemed high at first, many organizations quickly reassessed, realizing the tool's ability to save both time and money (Kytö et al, 2024.). In fact, it was deemed a *"no-brainer,"* as the average salary of knowledge workers in many Western countries exceeds the cost of the license.

This implies that, for companies to achieve a return on investment, each user would only need to save one hour per month using Copilot. Shortly after the general release of Copilot for Microsoft 365 to enterprise customers on November 1st, 2023, Microsoft Research published its initial studies on the tool's impact on information work (Kytö et al, 2024.).

The findings revealed a notable improvement in task completion speed across various scenarios, such as catching up on missed meetings and retrieving information (Kytö et al, 2024.). According to Kytö, (2024). users completed tasks in as little as 26.2% of the time it took the control group, who had no access to Copilot for Microsoft 365.

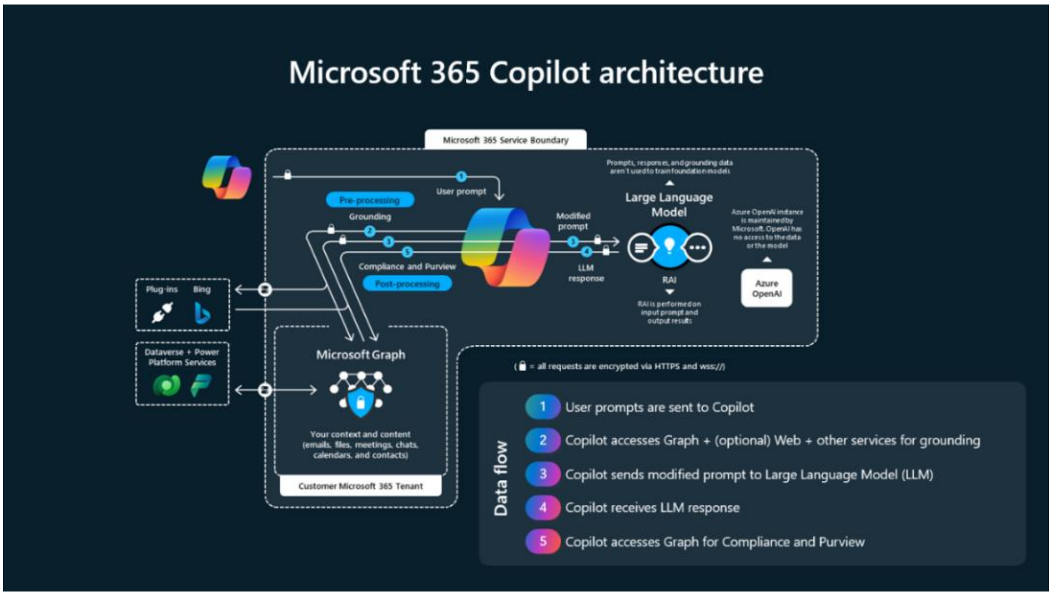
2.7.1 How does Microsoft 365 Copilot Work?

Microsoft 365 Copilot combines key components like LLMs, NLP, Microsoft Graph, semantic index, and Microsoft 365 Apps to help users utilize their existing data effectively. Users interact with Copilot through prompts via Microsoft 365 apps, with the service generating responses using data managed by Microsoft Graph from the organization's Microsoft 365 tenant (BJORDAL, M. H., et al 2025.).

Copilot acts as a bridge between user data and applications, facilitating seamless interaction as we can see from ***Figure 3*** below.

Figure 3

Copilot Architecture *(BJORDAL, M. H., et al 2025.)*



As mentioned before, Microsoft Graph serves as the central hub for Microsoft 365 Copilot, aggregating and searching content across the organization's tenant. It integrates data from services like SharePoint, Teams, and Outlook, enhancing prompts with contextual user information. Microsoft Graph ensures responses respect user permissions, data security, and compliance, allowing access to relevant data without switching between applications(BJORDAL, M. H. et al 2025.).

2.8 Functionalities in Copilot

Microsoft 365 Copilot is designed to improve several Microsoft applications by being integrated into them. This section will outline the key features of each selected application and provide examples of their most frequent use cases.

2.8.1 Microsoft 365 Copilot in Word

We use Word a lot in our daily work and Microsoft 365 Copilot in Microsoft Word is designed to assist in document creation and provide helpful suggestions and references to improve a user's work. It supports a range of writing tasks, from generating content from scratch to refining existing text and summarizing long articles (BJORDAL, M. H. et al, 2025).

When you open a new document or add a new line to your current one, a Microsoft 365 Copilot window will appear, allowing you to make requests. For example, you can ask it to write a paragraph on the latest trends in technology by providing general instructions. You can also provide more detailed input, such as notes, additional files, outlines, and your preferred state.

2.8.2 Microsoft 365 Copilot in PowerPoint

The purpose of Microsoft 365 Copilot in PowerPoint is to help users create visually engaging slides more quickly and generate content based on organizational data. Acting as a presentation assistant, Copilot transforms ideas into dynamic PowerPoint slides, including speaker notes and references. The tool creates presentations based on your outline or prompt and can even draw inspiration from linked documents, allowing you to convert a Word document into a PowerPoint presentation.

With specific prompts, you can modify layouts, modify text formatting, and synchronize animations to ensure your message is effectively communicated to the audience BJORDAL, M. H. et al, 2025)

2.8.3 Microsoft 365 Copilot in Excel

For regular Excel users, formatting data sheets to improve clarity and usability can be a time-consuming task. With Copilot, you can have a personal Excel assistant to help streamline these routine activities. Copilot can now analyze structured data, not just tables, and perform tasks like adding filters or splitting text. It can also assist in emphasizing key information through formula-based conditional formatting. For example, Copilot can simplify the task of highlighting critical data, such as identifying when spending exceeds revenue in a sales spreadsheet(Microsoft Community Hub, 2023).

2.8.4 Microsoft 365 Copilot in Teams

Microsoft 365 Copilot in Teams enhances collaboration by providing on-demand meeting insights. It uses data from multiple sources to offer real-time updates, facilitating data-driven decision-making and improving teamwork. During meetings, Copilot helps by answering questions, organizing key topics, and summarizing actions. It also allows users to request real-time updates without interrupting the flow, making it easier to join ongoing meetings(BJORDAL, M. H. et al, 2025).

Additionally, Copilot can extract meeting agendas, identify suitable participants for follow-up meetings, and assist in scheduling future meetings.

Microsoft 365 Copilot supports real-time meetings and chat threads by summarizing key points, identifying areas of disagreement, and generating notes. If someone misses a meeting, Copilot provides a summary of the outcome, lists important action items, and suggests follow-up questions for invites. In Teams chat threads, Copilot can summarize long conversations, allowing users to quickly catch up on key details, this is especially useful for teams in different time zones, enabling users to start their day with a short overview of overnight progress(BJORDAL, M. H. et al, 2025.).

2.9 Implementation of Copilot into businesses

We have previously discussed what Microsoft Copilot is and how it impacts the tools and tasks used in daily operations. However, we have not yet examined the practical aspects of its implementation, the challenges that may emerge, and the best practices to ensure successful deployment. Implementing artificial intelligence is not a simple "project" that can be delegated and expected to function autonomously.

Rather, it resembles assembling a complex puzzle requiring careful consideration of how each piece fits into the broader organizational context. This subchapter will explore the potential obstacles to implementation and outline best practices for overcoming them.

Before deploying Copilot, it is essential to understand both the purpose it serves and the method of implementation. Defining clear goals and objectives is a foundational step. This includes identifying key areas for improvement, setting measurable targets, and ensuring alignment with broader strategic initiatives.

Implementing Microsoft Copilot within an organization requires a structured and well-governed approach. Prior research emphasizes that successful adoption of AI tools depends heavily on organizational readiness, particularly in the areas of data hygiene, governance maturity, and workforce digital literacy (Füller et al., 2022). Before deployment, organizations must ensure that SharePoint and OneDrive environments are properly configured, permission structures are well defined, and documentation is accurate and searchable. Without this foundational work, Copilot's performance is limited, as the system relies extensively on organizational data to generate relevant outputs.

Organizations should also prepare employees for new ways of working that arise from AI-assisted tools. Studies show that the perceived usefulness and ease of use of AI systems strongly influence adoption success, and structured training significantly improves user confidence and outcomes (Hayawi & Shahriar, 2024). Ensuring users understand what Copilot can and cannot do reduces frustration and increases the likelihood of effective integration into daily workflows.

Once these steps are completed, gathering feedback from employees and users is crucial. Internal administrators and Microsoft product teams can use this feedback to optimize deployment, increase user engagement, and improve the overall Copilot experience.

2.9.1 Obstacles while Implementing Copilot

The implementation of Microsoft Copilot in organisational environments offers considerable potential for improving productivity, communication, and decision-making. However, research on enterprise AI adoption highlights several obstacles that organisations must address to ensure a secure, efficient, and responsible deployment. These challenges relate primarily to data governance, privacy protection, ethical considerations, and organisational readiness (Füller et al., 2022)(Hayawi & Shahriar, 2024).

A major obstacle concerns data security and privacy. Because Copilot operates on an organisation's internal content—including emails, documents, SharePoint sites, and Teams conversations—its performance depends heavily on the accuracy of access permissions and information governance structures. If organisational repositories contain improperly configured access levels or outdated permissions, Copilot may surface documents to users who would not normally have access to them. This risk does not result from Copilot itself but from the underlying data governance conditions within the organisation. Ensuring rigorous access control, sensitivity labels, and compliance policies is therefore essential before deployment.

Another significant challenge is content governance and information quality. Copilot generates responses based on existing organisational data. If documentation is incomplete, inconsistent, or poorly structured, the assistant may produce inaccurate or misleading outputs. Research shows that successful AI adoption requires mature knowledge management practices and well-maintained repositories to support reliable decision-making (Füller et al., 2022). Organisations must therefore invest in improving the clarity, ownership, and structure of their documentation before integrating Copilot into daily workflows.

The ethical and legal implications of AI-generated content present additional obstacles. As AI becomes increasingly integrated into business operations, organisations must consider issues related to transparency, authorship, copyright, and intellectual property. Hayawi & Shahriar, (2024)emphasise that organisations need clear guidelines regarding the acceptable use of AI-generated content, as well as procedures for reviewing and verifying outputs to avoid errors or regulatory non-compliance. These concerns are especially relevant in sectors with strict data protection requirements.

Furthermore, organisational readiness and workforce training remain recurring challenges. Employees must understand both the capabilities and limitations of Copilot, including when outputs require verification and how to provide effective prompts. Without adequate training, users may develop unrealistic expectations or rely on the system inappropriately. Studies show that structured training programmes significantly improve user confidence and enhance adoption outcomes (Hayawi & Shahriar, 2024)

Finally, cost and access management represent practical considerations during implementation. With a licence cost of €30 per user per month, organisations must evaluate which roles and teams will benefit most from Copilot. Not all employees require the same level of access, and adopting a targeted deployment strategy helps optimise costs while maintaining productivity gains. Aligning licence allocation with job functions, digital maturity, and information responsibilities ensures a more efficient and cost-effective implementation.

Overall, the obstacles associated with implementing Microsoft Copilot are not unique to this tool but reflect broader challenges inherent in integrating AI into enterprise environments. Through robust security policies, mature data governance, clear ethical guidelines, and structured training programmes, organisations can mitigate these challenges and unlock the full potential of Copilot.

2.9.2 Common Challenges and How to Avoid Them

Although Microsoft Copilot offers significant benefits, its implementation can present several challenges. Being aware of these potential pitfalls allows organizations to mitigate risks, saving time and resources while ensuring a smoother adoption process.

One common challenge is the lack of clear objectives. Establishing well-defined goals is essential for measuring success, as vague or undefined outcomes make it difficult to track progress and align Copilot's capabilities with business needs. Clearly defining objectives from the beginning ensures a structured and effective implementation (SEAN HOUGHTON et al, 2024.).

Another issue is limited customization, relying solely on out-of-the-box solutions may not fully accommodate an organization's unique workflows. Since every business has specific operational requirements, a generic AI solution may not be sufficient. Customizing Copilot to align with organizational needs enhances its performance and ensures seamless integration(SEAN HOUGHTON et al, 2024.).

Inadequate training is another barrier to successful adoption. Employees must be equipped with the necessary skills to use Copilot and Copilot Studio effectively. Without sufficient guidance, users may struggle to leverage Copilot's full potential. Allocating time and resources for comprehensive training ensures smoother adoption and maximizes its benefits(SEAN HOUGHTON et al, 2024.).

Additionally, the lack of performance monitoring can hinder Copilot's long-term success. Regularly tracking key metrics and gathering user feedback are crucial for maintaining efficiency. Failing to review performance data may lead to suboptimal results, whereas continuous assessment allows organizations to refine Copilot's functionality and optimize its impact*SEAN HOUGHTON et al, 2024)*.

Finally, scaling too quickly can introduce inefficiencies and operational disruptions. Expanding Copilot's use across multiple departments without adequate testing increases the risk of errors. A more effective approach is to begin with a controlled pilot project, evaluate performance, and gradually scale based on proven results(SEAN HOUGHTON et al, 2024.).

By proactively addressing these challenges, organizations can ensure a seamless and efficient integration of Microsoft Copilot, leading to optimized workflows, increased productivity, and long-term success.

3.0 Copilot Limitations

As good as tools are even in today's age, they have limitations. Limitations might not even mean a terrible thing as they create a door to the future to create a solution but even, so it is something that always needs to be talked about. A good example of limitations is Copilot on GitHub, we know that it offers substantial benefits to developers, but it also introduces potential security vulnerabilities.

The tool's reliance on extensive data input, which includes code and documentation, to generate relevant suggestions creates a risk of exposing confidential information. This exposure could lead to serious consequences for organizations, such as competitive disadvantages or legal issues. Therefore, the student concludes that while Copilot is a valuable tool, a careful assessment of its security implications is necessary(Vaswani et al., 2017a).

GitHub Copilot presents several challenges related to copyright, production code quality, and project integration. A key concern is that it generates code based on existing repositories, raising legal questions about ownership, especially when the original code belongs to a company. If copyright protections are insufficient, companies may stop sharing code, potentially weakening Copilot's training data.

Beyond legal concerns, Copilot performs well for small code snippets but struggles with large-scale production code, often producing inconsistent or incorrect outputs due to a lack of understanding of internal abstractions. Additionally, its limited project context leads to poor integration with other frameworks and components, requiring developers to spend extra time modifying the generated code. As a result, while Copilot enhances productivity, it cannot fully replace developers in complex software development tasks(Vaswani et al., 2017a).

In this case to ensure the quality of code generated by GitHub Copilot, several best practices should be followed. First, code review and manual checks are essential. While the generated code is mostly functional, it may include errors or irrelevant parts that need to be identified through careful review. This process not only ensures code quality but also helps developers understand the generated code's context (Vaswani et al., 2017a).

Another important practice is avoiding the inclusion of sensitive data in Copilot's training. Instead, developers should use dummy or testing data to prevent exposing customer privacy or production data. This approach is like using mock data during tests and ensures that sensitive information is protected(Vaswani et al., 2017a).

To detect hidden defects, developers should incorporate extensive integration and end-to-end testing. Copilot often struggles with integration, so these tests can uncover issues that might not be caught during code review. By adding test cases that cover edge cases and functionality, developers can ensure the quality of the generated code and avoid introducing bugs into the system (Vaswani et al., 2017b).

Finally, it's crucial to have a well-defined design document and architecture before using Copilot. GitHub Copilot should assist with repetitive tasks, not drive the entire project. A clear roadmap, established through team discussions, will guide how Copilot can help in implementing specific functionalities while maintaining control over the project's overall design.

A more recent article about GitHub Copilot made by Zhang indicates there is difficulty in integrating Copilot with Integrated Development Environments (IDEs) and other plug-ins. Some developers experienced conflicts between Copilot and existing shortcut settings, while others found that certain IDEs remain unsupported. Additionally, instability in Copilot's servers has led to accessibility issues.

The article also stated that the quality and diversity of Copilot's code suggestions were also criticized. Developers noted that Copilot sometimes provides a limited number of solutions, which may not be sufficient for complex tasks(Zhang et al., 2023).

Some reported that as code files grow, the relevance and accuracy of Copilot's suggestions deteriorate. Furthermore, concerns regarding code privacy were frequently mentioned, as developers worried about the unauthorized use of their code(Zhang et al., 2023).

While some users found Copilot superior to other AI-assisted coding tools, others reported a subpar user experience. These limitations underscore the need for continued improvements in Copilot's integration, stability, code generation capabilities, and data security measures(Zhang et al., 2023).

3.1 Copilot vs ChatGPT vs Gemini

Large language models (LLMs) such as Microsoft Copilot, OpenAI's ChatGPT, and Google's Gemini share a foundational transformer-based architecture (Vaswani et al., 2017a), but they differ significantly in purpose, training data, integration capabilities, and organisational applicability. Understanding these differences is essential for interpreting how users perceive Copilot relative to general-purpose AI systems.

ChatGPT functions as a broad conversational agent optimised for open-ended reasoning, creative ideation, and general problem solving (Shanahan et al, 2024). Its strengths lie in flexible linguistic generation, wide-ranging general knowledge, and adaptability across diverse domains. Gemini extends these capabilities by incorporating multimodal inputs text, images, audio, and structured data—and providing more advanced integrated reasoning across different information types (Alhur, 2024).

In contrast, Microsoft Copilot is designed specifically for enterprise productivity. Instead of operating as a standalone model, Copilot is embedded within the Microsoft 365 ecosystem and interacts with organisation-specific data through Microsoft Graph. This allows Copilot to summarise emails, generate documents, prepare presentations, analyse spreadsheets, and automate workflows using internal organisational content, while maintaining strict permission boundaries and compliance protections (Arsal et al., 2024). This enterprise grounding differentiates Copilot from ChatGPT and Gemini, which rely primarily on internet-scale public data.

These differences make ChatGPT and Gemini suited for broad, exploratory tasks, while Copilot is optimised for operational tasks within structured workplace environments. For example, ChatGPT and Gemini can generate creative narratives, brainstorm ideas, or answer general questions with extensive contextual flexibility. Copilot, by contrast, excels at tasks where organisational context is essential such as drafting an email consistent with a prior conversation thread or generating a summary based on an internal project document.

Security and compliance also represent key differentiators. Public LLMs like ChatGPT and Gemini operate outside enterprise boundaries, raising concerns about data exposure and regulatory compliance. Copilot, however, ensures that business data remains within the Microsoft 365 tenant and is not used to train public models, making it more suitable for organisations with stringent privacy and governance requirements.

Overall, while these three systems share the same underlying LLM foundation, they are optimised for distinctly different purposes. ChatGPT and Gemini prioritise general reasoning and multimodal intelligence, whereas Copilot is engineered as an enterprise assistant designed to augment workplace productivity through tight integration, contextual awareness, and privacy-by-design principles. These distinctions directly inform the user perceptions and adoption patterns explored in this study.

4.0 Empirical Synthesis

In this thesis, productivity is operationalised as the net value added to the workflow after accounting for verification and correction effort. To provide a proper context for the economic analysis of Microsoft Copilot, this thesis first redefines the concept of *"productivity"*. While traditional software metrics focus on efficiency, the integration of Generative AI requires a shift towards effectiveness. As highlighted in the optimization analysis by Rodrigues Pontes Reis Miranda, (2025). speed alone is insufficient the content generated is factually incorrect. Therefore, in this study, productivity is defined as the net value added to the workflow *after* accounting for the time users spend verifying the AI's suggestions.

This distinction explains the adoption barriers identified in recent qualitative research. According to Bano et al., (2025). users often experience *"unmet expectations"* regarding the tool's reasoning capabilities, leading to hesitation in integrating Copilot into critical tasks. Their study confirms that the primary barrier is not technical access, but the lack of organizational readiness specifically, the need for training to help employees manage trust and reliability issues(Bano et al., 2025.). These operational challenges directly influence the financial viability of the tool.

From a technology adoption perspective, these findings align with established acceptance models. In line with the Technology Acceptance Model (TAM), perceived usefulness of Copilot is high for routine tasks, but perceived ease of use is uneven due to the cognitive effort required for prompt formulation and output validation. The UTAUT framework further explains this pattern by emphasizing the role of facilitating conditions, such as structured training, governance, and organizational support, which strongly influence sustained usage. Where these conditions are absent, adoption remains cautious and fragmented despite perceived benefits.

Measuring the Return on Investment (ROI) for generative AI tools like Microsoft Copilot is a complex challenge that goes beyond simple productivity metrics. Traditionally, organizations estimate ROI by calculating the time saved per employee multiplied by their hourly wage. For example, if Copilot saves a user 10 minutes per day, the monthly savings might theoretically cover the license cost.

However, calculating the true ROI requires a more nuanced approach. As recent optimization studies suggest, organizations must account for significant *"hidden costs"* that offset these initial productivity gains (Rodrigues Pontes Reis Miranda, 2025). These include the "human-in-the-loop" verification time required to correct hallucinations, the resources allocated to mandatory training and prompt engineering workshops, and the extensive *'data hygiene'* projects needed to secure organizational repositories before deployment (Rodrigues Pontes Reis Miranda, 2025.).

Furthermore, the value of Copilot often manifests in qualitative rather than quantitative improvements, such as higher employee satisfaction or improved output quality, which are harder to monetize directly. Consequently, the net ROI is often lower or even negative in the initial months of adoption while the workforce climbs the learning curve, only becoming positive as organizational maturity with the tool increases (Rodrigues Pontes Reis Miranda, 2025.).

Finally, from a Diffusion of Innovation perspective, the empirical results suggest that most organizations are still in an early-adopter phase, characterized by experimentation rather than institutionalized use. In such stages, productivity gains are uneven, ROI is often delayed, and value materializes progressively as users gain experience and organizations mature in their use of the technology. This explains why economic impact tends to improve over time rather than appearing immediately after adoption.

Chapter 3: Methodology

3.1 Research Approach

For this thesis, I will adopt both qualitative and quantitative research methods. The goal is to address practical problems while simultaneously contributing to academic and technical knowledge. This approach will enable a comprehensive understanding of the topic from both theoretical and practical perspectives.

3.1.1 Qualitative Research

Qualitative research is a methodological approach focused on exploring and gaining deeper insights into real-world issues. Unlike quantitative research, which collects numerical data or tests specific interventions, qualitative research aims to understand participants' experiences, perceptions, and behaviors(Tenny et al., 2022). It answers questions about "*how*" and "*why*", "*how many*" or "*how much*".

This type of research can be conducted independently, relying solely on qualitative data, or as part of a mixed-methods approach, which integrates both qualitative and quantitative data to offer a more comprehensive understanding of a phenomenon(Tenny et al., 2022). Through qualitative inquiry, the study can assess how organizations interact with Copilot, what challenges they encounter, and how these insights can inform more effective AI integration strategies.

3.1.2 Quantitative Research

According to Sukamolson, 2007. quantitative research involves the collection and analysis of numerical data to identify patterns, determine averages, make predictions, test causal relationships, and generalize findings to larger populations. Unlike qualitative research, which focuses on non-numerical data (e.g., text, video, or audio), quantitative research is commonly used in fields such as biology, chemistry, psychology, economics, sociology, and marketing(Sukamolson, 2007.). It plays a key role in providing objective, measurable insights across various disciplines.

By combining qualitative and quantitative approaches, I aim to provide a detailed understanding of how Microsoft Copilot can transform enterprise workflows while offering actionable recommendations for successful adoption.

3.2 Data Collection

This research focuses on medium to large organizations that have either implemented or are planning to implement Microsoft Copilot within their environments. A sample of approximately one to three organizations will be selected based on their willingness to participate.

To ensure a comprehensive understanding of Copilot's adoption and impact, both primary and secondary data sources will be utilized. Primary data will be collected through surveys and semi-structured interviews. Surveys will be distributed publicly via LinkedIn, Discord, Reddit, and Microsoft Blogs to gather quantitative insights into Copilot's effect on productivity, decision-making, and overall efficiency.

Additionally, semi-structured interviews with IT technicians and users who have already had experience with Copilot will provide deeper insights into the challenges faced during implementation, strategies used and observed outcomes.

Secondary data will be gathered from Microsoft's published materials on Copilot adoption, such as their success stories. These sources will provide both qualitative insights, such as implementation strategies and organizational feedback, and quantitative data, including usage statistics and adoption reports.

To collect relevant information, various data collection instruments will be used. Surveys will focus on gathering user feedback regarding Copilot's impact on daily workflows, with example questions like: "What apps do you use with Copilot?" and "How has Microsoft Copilot affected your productivity?" Semi-structured interviews will allow participants to share their experiences through open-ended questions, such as: "What were the main challenges during the implementation of Copilot?" and "Would you recommend Microsoft Copilot to other professionals in your industry?" The full survey questionnaire is provided in Appendix D. Lastly, analysing Microsoft's success stories on Copilot adoption will provide valuable insights into best practices and common challenges.

3.3 Data Analysis

The survey data collected for this study underwent a light post-collection normalization process prior to analysis. This step was required due to the exploratory nature of the questionnaire and the presence of partially overlapping or ambiguously phrased responses, particularly in questions related to perceived productivity, efficiency, and challenges of Copilot usage.

Normalization consisted of grouping semantically equivalent responses under common analytical categories, without altering the original meaning or direction of participants' answers. No responses were removed, fabricated, or reweighted. This process aimed solely to improve interpretability and allow aggregated descriptive analysis of user perceptions.

Given this normalization step, the survey results are treated as indicative and exploratory rather than as a statistically rigorous quantitative dataset. Consequently, the findings are reported using descriptive statistics only and are interpreted cautiously, in line with best practices for early-stage technology adoption research.

3.4 Privacy and Ethical Considerations

All data will be anonymized to protect the identities of participants and organizations. Permission consent will be obtained from all participants before collecting data as they will be given the option to withdraw at any time if they wish to.

The study will follow academic and ethical guidelines and will undergo approval by my thesis advisor, and all collected data will be stored securely on password-protected systems, and access will be restricted to authorized persons.

3.5 Methodological Limitations

As an exploratory investigation carried out during the early stages of enterprise adoption of Microsoft Copilot, this study is characterised by several methodological limitations. Acknowledging these constraints is essential for contextualising the findings and guiding future research in this emerging field.

The first limitation concerns the sample size. With thirty survey respondents and two interview participants included in the final analysis, the dataset provides indicative insights rather than statistically generalisable conclusions. Although the mixed-methods approach enhances the interpretive depth of the study, the small sample size limits the ability to model variability across industries, organisational roles, and levels of AI proficiency. As a result, no claims of population-level inference can be made.

A second limitation relates to the reliance on self-reported measures, particularly for constructs such as productivity, efficiency, and perceived usefulness. Self-assessments are inherently subjective and influenced by user expectations, familiarity with technology, and individual interpretation of Copilot's capabilities. Without objective productivity metrics—such as time-on-task measurements, system telemetry, or automated usage logs—the findings reflect perceived impacts rather than quantifiable performance outcomes. In addition, the quantitative survey data underwent a light post-collection normalization process. This step was necessary due to the exploratory nature of the questionnaire and the presence of overlapping or ambiguously phrased responses. Normalization consisted of grouping semantically equivalent answers under common analytical categories to support coherent descriptive analysis. No responses were removed, fabricated, or reweighted, and the original intent of participants' answers was preserved. Nevertheless, this process further reinforces the exploratory character of the quantitative findings.

A third limitation stems from the semi-structured interview design. While this approach facilitated rich qualitative insights, it also introduced variability in the depth and focus of responses. The absence of a fully standardised interview script limited direct comparability across participants. Additionally, because only two interviews were deemed analytically suitable, the qualitative component represents a narrow but in-depth set of perspectives rather than a comprehensive range of organisational experiences.

Another limitation concerns organisational context and data quality. Participants' assessments of Copilot may reflect not only the tool's intrinsic capabilities but also the maturity of their organisation's data hygiene, governance practices, access control structures, and employee training. Consequently, perceptions of Copilot's effectiveness may be shaped by contextual factors that extend beyond the technology itself.

The temporal context of the study also introduces limitations. Copilot is a rapidly evolving technology, with frequent updates, new features, and expanded integrations. The data collected reflects user perceptions and system behaviour at a specific moment in time, and these findings may evolve as the technology matures and organisational practices adapt. Despite these constraints, the study offers valuable exploratory insights into early adoption dynamics and perceived impacts of Microsoft Copilot in enterprise environments. The findings should therefore be interpreted as an initial contribution that informs future longitudinal, experimental, and large-scale research on AI-assisted productivity tools.

Chapter 4: Success Stories and Analysis

This chapter examines the implementation of Copilot in enterprise environments, with a focus on its value, adoption, impact, and challenges. The selected success stories Kyndryl, Oxford University Hospitals NHS Foundation Trust (OUH), and IU International University of Applied Sciences represent different industries, including IT services, healthcare, and higher education. These success stories provide valuable insights into the role of AI-powered productivity tools in enhancing operational efficiency, decision-making, and personalized user experiences. A comparative analysis highlights key trends, challenges, and critical success factors for enterprises integrating generative AI solutions.

The organizations and companies selected for this success story and analysis are diverse in scale, ensuring a broad perspective. The primary criteria for selection include companies that have documented improvements in efficiency, security, and collaboration following the implementation of Copilot. Additionally, the analysis focuses on cases where Copilot has been deployed on scale, impacting a significant number of employees and business processes. These cases are drawn from Microsoft’s customer stories, which are publicly available online and on the respective company websites.

4.1 Success Stories

4.1.1 Kyndryl: AI-Driven IT Infrastructure Optimization

On October 2, 2024, Kyndryl, the world's largest IT infrastructure services provider, announced an expanded suite of services to accelerate the adoption of Copilot for Microsoft 365 in enterprise environments. Under the "Microsoft Gen AI for the Workplace by Kyndryl" initiative, these offerings focus on maximizing Copilot's impact by integrating advanced data services, optimizing user interfaces, and providing continuous performance tracking.(Kyndryl, 2023.).

Leveraging its expertise in enterprise AI deployment, Kyndryl helps organizations develop high value use cases for Copilot adoption. The services are designed to enhance decision-making, streamline operations, and drive digital transformation, ensuring a seamless and effective rollout of AI-powered workplace solutions (Kyndryl, 2023.).

Kyndryl's consultant approach includes tailored advisory, structured implementation, and managed services, enabling businesses to align Copilot deployment with their specific requirements. Through its Copilot Value Discovery Workshop, Kyndryl assesses an organization's technical needs, information governance framework, and AI readiness, ensuring a smooth and effective implementation(Kyndryl, 2023.).

Internally, Kyndryl has demonstrated measurable efficiency gains through its own Copilot deployment. A company survey found that 95% of users saved at least 20 minutes per day on knowledge discovery alone, highlighting the tool's potential to boost productivity and deliver a strong return on investment (Kyndryl, 2023.).

As a strategic global partner of Microsoft, Kyndryl continues to drive AI innovation in enterprise environments, expanding opportunities in data modernization, governance, and AI adoption. This partnership reinforces Kyndryl's commitment to providing scalable, secure AI solutions that enhance workplace efficiency and business performance(Kyndryl, 2023.).

4.1.2 Oxford Hospital University

Oxford University Hospitals NHS Foundation Trust (OUH) is among the largest teaching hospital trusts in the United Kingdom. Its IT department is responsible for supporting the delivery of acute healthcare services across Oxfordshire while also providing a technological infrastructure for other regional NHS organizations(Microsoft, 2025.).

With a user base exceeding 18,500 networked individuals, the OUH IT team identified a growing interest in the potential applications of generative AI to enhance operational efficiency and transform various aspects of professional workflows within the organization(Microsoft, 2025.).

OUH sought to enable its staff to explore and innovate with generative AI technologies within a secure, compliant, and controlled environment. To optimize the value of its existing technological infrastructure, the Trust implemented Microsoft 365 Copilot, ensuring seamless integration with its current Microsoft ecosystem. The IT service department at the University of Oxford conducted a week-long evaluation of Microsoft Copilot to assess its impact across departmental operations. Licenses were provided for IT staff, enabling them to explore the benefits of generative AI in tasks such as document summarization, content creation, meeting support, and task management(Saran, 2025).

Given the early stages of adoption, the department prioritized continuous monitoring and iterative assessment to determine the technology's overall value. While some employees were initially uncertain about AI integration, the department fostered a collaborative learning environment by organizing Microsoft briefing sessions, community engagement activities via Microsoft Teams, and an internal knowledge-sharing network(Saran, 2025.). Initiatives such as a "prompt of the week" challenge and drop-in help clinics further facilitated user engagement(Saran, 2025.).

Lee Massie Head of IT in Oxford University Hospitals NHS Foundation Trust has stated the following *"I've learnt that I can get 50% of what I need from some prompts, but by refining those prompts I can drive that quality up to 80 or 90%,"* Lee Massie also explained *"Once you start to learn how to drive Copilot properly, you get much better, more refined answers."*

Additionally, the university established an acceleration team to explore deeper integrations with Copilot Studio and optimize the tool's application across various workflows. Technology has proven particularly beneficial for smaller teams with diverse responsibilities, streamlining document creation and enhancing task management. Moreover, Copilot’s capabilities have supported neurodiverse users by improving workload prioritization and focus(Saran, 2025.).

Significant time reductions have been reported in administrative tasks. For instance, an information governance team member reduced complex formatting tasks from 15–20 minutes to mere seconds. A senior staff member estimated personal time savings of approximately one hour per day by leveraging Copilot for email organization and structured reporting(Microsoft, 2025).

Additionally, Copilot has substantially optimized meeting summaries, with one team leader reducing meeting documentation time from 90 minutes to 25 minutes, yielding a weekly saving of two to three hours(Microsoft, 2025).

Feedback from users highlights Copilot’s effectiveness in enhancing workflow efficiency. Teams have praised its ability to generate content, refine document formatting, summarize discussions, and improve real-time responsiveness in meetings. Executive leaders have noted "massive time savings" in taking notes and document summarization, while team leaders emphasize the benefits of automating routine document creation(Microsoft, 2025).

Oxford University is actively working to integrate generative AI into research, teaching, and administrative functions. Inspired by Copilot’s potential to enhance both current and future workplace operations, a delegation from the university visited Microsoft’s headquarters in Redmond to further explore strategic AI investment opportunities(Saran, 2025).

4.1.3 IU International University of Applied Sciences

IU International University of Applied Sciences (IU) is the largest private university in Germany, employing over 600 professors and 4,000 staff members to support a student body exceeding 140,000. The institution offers a diverse array of more than 200 Bachelor's, Master's, and MBA programs, equipping students with both current and future-oriented skills essential for emerging professions.

Notably, 94% of IU graduates secure employment within three months of completing their studies, highlighting the university’s strong emphasis on career readiness and industry relevance(Microsoft, 2025.).

Incorporating artificial intelligence (AI) into education is essential for creating personalized, scalable, and effective learning experiences. Preparing students for successful careers requires integrating academic knowledge with practical, real-world skills through tailored educational approaches that align with individual aspirations and professional goals(Microsoft, 2025.).

However, achieving personalized education presents challenges, including outdated technology, budgetary constraints, data privacy concerns, and faculty training needs. IU International University of Applied Sciences faced these challenges when seeking to implement technology-driven personalized learning paths(Microsoft, 2025).

Through its collaboration with Microsoft, IU integrated Microsoft 365 Copilot and other AI-powered tools into its educational framework. These innovations have significantly enhanced student engagement, accelerated learning progress, and better prepared graduates with in-demand skills. According to Dr. Marian Bodenstedt, COO Online DACH at IU Group, AI is particularly transformative in fields such as engineering, technology, business, and marketing, where efficiency gains from generative AI are already evident(Microsoft, 2025). AI enables personalized, cost-effective learning by adapting content to students' backgrounds, capacities, and interests, fostering a more individualized and engaging educational experience(Microsoft, 2025).

Microsoft 365 Copilot facilitates a range of academic and administrative tasks for both students and faculty, including automated meeting transcriptions, content summarization, flashcard generation, and writing assistance. Its integration within the learning environment enables the creation of personalized educational pathways, fostering a more interactive and adaptive learning experience.

Features such as deep dialogue learning provide additional support, particularly for remote learners, by offering tailored resources that enhance engagement and streamline the study process. Furthermore, Copilot assists students in managing their academic responsibilities by generating customized study plans that effectively balance educational, professional, and personal commitments, in that way optimizing their overall learning efficiency(Microsoft, 2025).

4.2 Comparative Analysis

These success stories illustrate both the flexibility and limitations of Microsoft Copilot across different sectors. Kyndryl's focus on operational efficiency contrasts with Oxford University Hospitals' (OUH) emphasis on administrative control and IU International University's drive for educational innovation. Despite these differences, all three organisations report trust in Microsoft

Copilot, face challenges related to user adoption and describe perceived productivity improvements.

Each organisation has approached Copilot implementation differently. Kyndryl has encouraged adoption through internal champions and enablement initiatives, while OUH has adopted a more structured, top-down governance approach to ensure regulatory compliance. In contrast, IU has prioritised a user-focused strategy, relying on workshops and training programmes to familiarise students and faculty with AI-driven tools. These differences highlight that successful adoption depends not only on the technology itself but also on organisational preparation, leadership alignment, and training efforts.

While all three organisations report efficiency gains, the extent to which Copilot drives meaningful transformation varies. At Kyndryl, AI adoption has contributed to shifts in employee roles and internal workflows, whereas IU has observed improvements in student learning experiences and academic support processes. OUH's careful, compliance-driven approach suggests that in highly regulated environments, AI technologies are more likely to introduce gradual, incremental improvements rather than rapid or disruptive change.

From a theoretical perspective, these cases align with several established frameworks. Kyndryl's experience reflects the resource-based view, in which AI is treated as a capability-enhancing asset that strengthens organisational performance. OUH's emphasis on compliance over innovation aligns with institutional theory, demonstrating how regulatory pressures shape the pace and scope of technology adoption. Meanwhile, IU's focus on AI as a tool for skill development and personalised learning supports human capital theory, which frames education and capability-building as sources of competitive advantage.

A key limitation across all three cases is the absence of clearly articulated return on investment (ROI) metrics, making it difficult to assess the long-term economic impact of Copilot adoption. This observation aligns with diffusion of innovation theory, which suggests that early adopters often prioritise experimentation and learning over formal evaluation during the initial phases of technology uptake.

At the same time, it is important to critically contextualise these success stories. All three cases are drawn from Microsoft-published customer narratives or partner communications and therefore primarily reflect positive implementation outcomes. As such, they are subject to potential selection and survivorship bias, as unsuccessful pilots, limited-impact deployments, or abandoned initiatives are unlikely to be publicly documented. Moreover, reported productivity gains are typically based on internal assessments or self-reported estimates rather than independently verified empirical measurements. Consequently, these cases should be interpreted as illustrative industry examples rather than as rigorous academic evidence.

Taken together, these findings reinforce this thesis's emphasis on the importance of strategic alignment when integrating AI into organisational processes. Rather than supporting the notion of a uniform technological disruption, the case studies reveal a spectrum of impact ranging from incremental efficiency gains to more transformative outcomes, depending on sectoral context, governance structures, and implementation strategy. In conclusion, the success stories of Kyndryl, Oxford University Hospitals (OUH), and IU International University illustrate the diverse applications of Microsoft 365 Copilot, demonstrating its potential to enhance IT operations, reduce administrative burdens in healthcare, and support personalised learning in education. At the same time, they highlight the inherent complexities of AI adoption, where factors such as governance, organisational readiness, peer influence, and contextual constraints play a decisive role in shaping outcomes.

Finally, it is critical to interpret these success stories with a degree of caution. Because Kyndryl, Oxford University Hospitals, and IU International University are featured in vendor-published literature, these narratives likely reflect 'survivorship bias,' highlighting best-case scenarios while potentially underreporting implementation friction. Unlike independent empirical studies, these vendor accounts may gloss over the significant integration challenges and employee resistance identified in broader literature. Therefore, these cases should be viewed as illustrative of Copilot's *potential* ceiling, rather than representative of the average enterprise experience, which—as our primary data will show—is often fraught with more practical hurdles.

Chapter 5: Results and Discussion

This chapter presents the findings derived from the surveys and interviews conducted to assess the adoption of Microsoft Copilot in enterprise environments. The analysis explores key trends, insights, and implications drawn from the collected data, providing a comprehensive evaluation of the role of Copilot in modern business settings. This discussion aims to bridge the gap between theoretical adoption strategies and practical implementation challenges faced by organizations.

5.1 Survey Results

The survey collected responses from 30 professionals across diverse industries, with the majority representing the technology sector. Given the limited sample size and the exploratory nature of the study, the survey does not aim to provide statistically generalisable findings. Instead, it offers an initial insight into how Microsoft Copilot is currently perceived and used in enterprise contexts, with the objective of identifying emerging patterns that may inform implementation strategies and future research. Survey results are discussed in this chapter; the survey instrument itself is available in Appendix D.

It is important to note that the survey results presented in this section should be interpreted as exploratory indicators of user perception rather than as statistically robust evidence. The findings are therefore reported descriptively and are used primarily to complement the qualitative interview analysis.

The first aspect examined concerns perceived productivity. As illustrated in ***Figure 4***, 35.7% of respondents reported that Copilot had "somewhat improved" their productivity, suggesting moderate perceived gains in workflow efficiency. A further 28.6% indicated that productivity was "greatly improved," reflecting more substantial perceived benefits in areas such as task initiation and decision support. However, 14.3% of respondents expressed uncertainty regarding Copilot's impact, which may be attributed to limited exposure, early-stage adoption, or inconclusive personal experience. The remaining respondents reported either minimal impact or no noticeable productivity change.

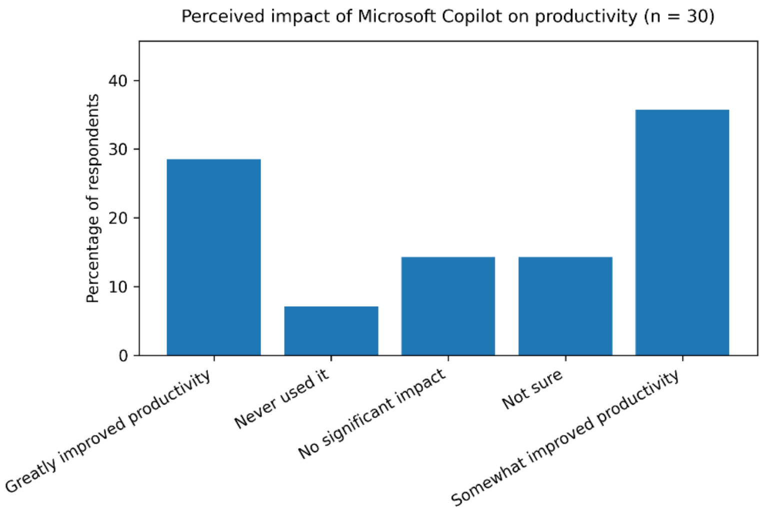


Figure 4 Perceived impact of Microsoft Copilot on productivity based on an exploratory survey of professionals (n = 30). The results reflect self-reported perceptions and indicate varying levels of perceived productivity change following Copilot adoption.

To further understand the nature of these perceived gains, ***Figure 5*** presents the specific benefits identified by respondents. The results show a strong emphasis on efficiency-related outcomes. Faster document creation was the most frequently cited benefit (60%), followed by a reduction in time spent on repetitive tasks (56%). These findings suggest that Copilot's primary perceived value lies in automating routine activities and accelerating the initial phases of content creation.

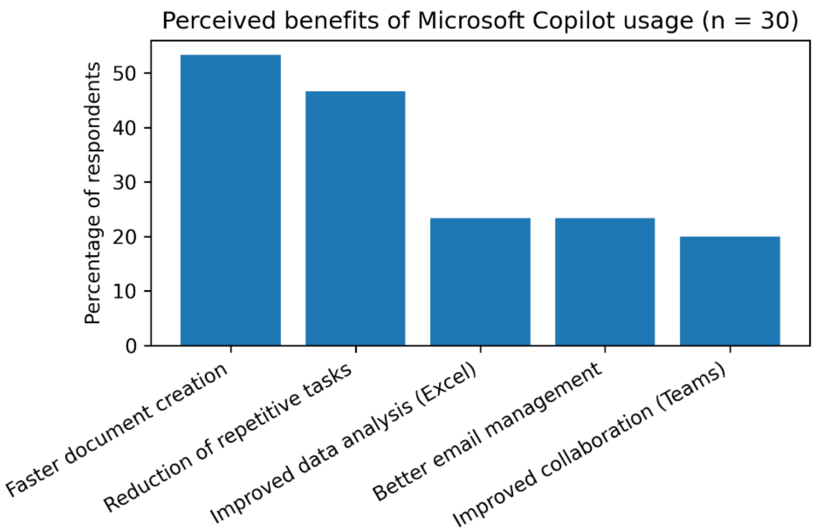


Figure 5 Perceived benefits of Microsoft Copilot usage as reported by survey respondents (n = 30). Multiple selections were allowed. The figure highlights efficiency-related benefits, particularly faster document creation and reduction of repetitive tasks.

Beyond these core benefits, respondents also highlighted improvements within specific Microsoft 365 applications. Improved data analysis in Excel and enhanced email management and summarisation was each reported by 28% of participants, indicating perceived value in both analytical and communication-related tasks. Additionally, 24% of respondents noted enhanced collaboration in Microsoft Teams, pointing to Copilot's role in supporting coordination and teamwork. A smaller but noteworthy proportion of respondents (16%) reported other benefits, including support for data gathering and code-writing shortcuts, suggesting that Copilot's utility extends beyond purely administrative tasks.

The survey also explored which applications respondents most frequently used Copilot with. As shown in ***Figure 6***, usage was concentrated within the Microsoft 365 ecosystem: 56% of respondents used Copilot with Word, 40% with Excel, 28% with PowerPoint, 28% with Teams, and 24% with Outlook. This pattern is consistent with Copilot's design as a productivity assistant embedded within Office 365 workflows and reinforces the observation that its perceived value is closely tied to document-centric and communication-heavy tasks.

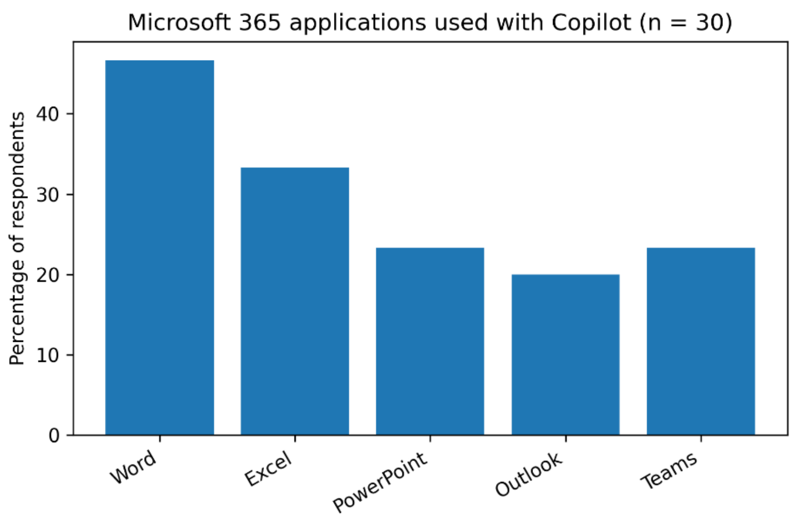


Figure 6 Microsoft 365 applications most frequently used with Microsoft Copilot according to survey respondents (n = 30). Multiple selections were allowed. Usage is concentrated within document- and communication-oriented applications.

At the same time, this concentration of usage within Office 365 applications highlights a potential limitation. While Copilot appears effective in enhancing productivity within Microsoft-native workflows, respondents implicitly suggested that broader integration with non-Microsoft tools and environments could further increase its usefulness and competitiveness relative to other AI agents. Participants were also asked to identify desired improvements to Copilot. Suggested enhancements included faster support for code development, full document translation, and improved output accuracy. One respondent summarised these expectations as follows:

"It is essential to enable more autonomous tasks to streamline operations and increase efficiency. By incorporating advanced automation, we can reduce the need for manual intervention, minimize errors, and enhance overall productivity."

To contextualise these expectations, ***Figure*** *7* presents the key challenges reported by respondents. Inaccurate or irrelevant suggestions were cited by 56% of participants, followed by lack of training or knowledge (32%), difficulty integrating Copilot into existing workflows (24%), and concerns related to data privacy and security (16%). This high rate of concern regarding accuracy is consistent with recent qualitative research by Bano et al. (2025.), who investigated user perceptions of generative AI in the public sector. Their study similarly identified reliability and the risk of hallucinations as primary barriers to trust, suggesting that the struggle with output accuracy is a systemic challenge across different organizational contexts, rather than an isolated issue within our sample. These combined findings indicate that, despite perceived productivity benefits, adoption barriers remain significant and are closely linked to organisational readiness, training, and governance practices

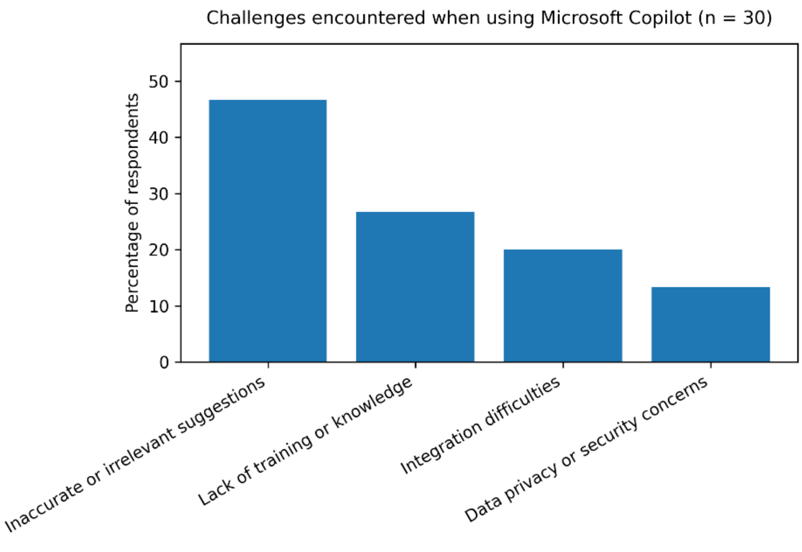


Figure 7 Challenges encountered by users when adopting Microsoft Copilot, as reported in an exploratory survey (n = 30). Multiple selections were allowed. The results indicate that accuracy, training, and integration remain key adoption barriers.

Finally, respondents were asked whether Copilot should be implemented within their organisation's workflows. As shown in ***Figure 8***, 67% supported adoption provided that clear guidelines and proper training were in place. This result reinforces the importance of structured onboarding, usage policies, and education in mitigating the challenges identified earlier.

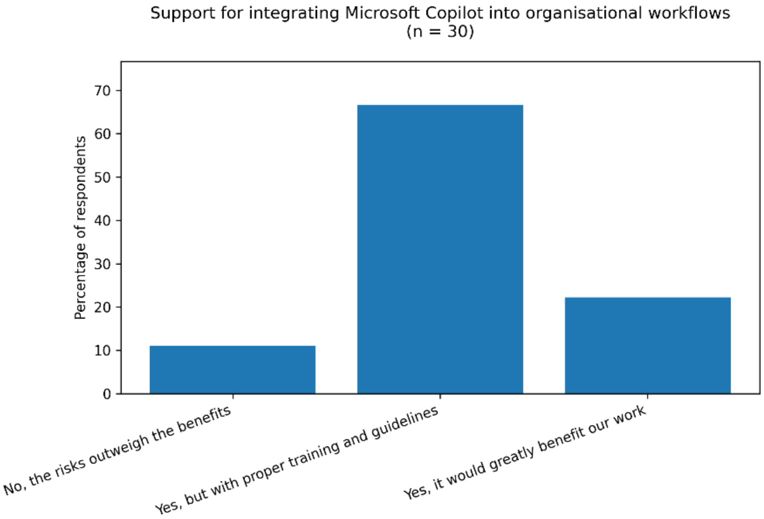


Figure 8 Support for integrating Microsoft Copilot into organizational workflows based on an exploratory survey of professionals (n = 30). Most respondents support adoption provided that clear guidelines and appropriate training are established.

Overall, the survey findings suggest that Microsoft Copilot is perceived as a productivity-enhancing tool, particularly for routine and document-oriented tasks, while also highlighting substantial limitations related to accuracy, training, and integration. Given the exploratory scope and limited sample size, these results should be interpreted as indicative rather than conclusive. They provide an initial empirical basis for understanding early adoption patterns and underscore the need for further, larger-scale studies to validate and extend these insights.

5.2 Qualitative Interview Analysis

The full interview transcripts are provided in Appendix B and Appendix C. To complement the survey data and provide a deeper understanding of how Microsoft Copilot is experienced in real organisational environments, two semi-structured interviews were conducted with individuals who interact with Copilot in distinct professional roles. One participant works as a Support Engineering Manager in a multinational corporation, representing a high-engagement end user. The other is an implementation specialist responsible for guiding enterprise clients through large-scale Copilot deployments. Given the limited number of interviews, the qualitative findings are not intended to be generalisable but to provide contextual and interpretative depth. Analysed together, these interviews offer a multifaceted perspective on Copilot's benefits, challenges, and broader implications for organisational productivity and readiness.

Although the participants operate in different contexts—one as a daily user and the other as an enterprise advisor—their experiences converge across several themes. These themes were identified through inductive thematic analysis and are presented below, supported by interpretative commentary. This approach allows for a cohesive synthesis rather than fragmented individual accounts, thus strengthening the analytical coherence of the findings.

A prominent theme emerging from the interviews is the perception of Copilot as an accelerator of initial task execution. Both participants highlighted the tool's capacity to reduce the cognitive load associated with starting a task. The Support Engineering Manager described how Copilot assists in drafting performance reviews, preparing technical documentation, composing emails, and structuring internal communications. By generating structured drafts or summarising long communication threads, Copilot "removes the blank page," enabling users to focus their cognitive effort on refinement and decision-making. Similarly, the implementation specialist emphasised Copilot's ability to rapidly summarise large documents and identify trends in organisational datasets, describing it as "an intern for every employee" that can substantially reduce perceived task effort. These reflections align with the survey findings, where respondents consistently reported improvements in routine or text-intensive tasks.

However, both interviewees underscored that Copilot's effectiveness is highly dependent on organisational data quality, particularly within SharePoint and OneDrive repositories. The implementation specialist stressed that organisations frequently overlook the need for *"data hygiene,"* yet Copilot's ability to generate useful and accurate content is shaped by the structure, clarity, and accessibility of internal information. Copilot's limitations thus become most apparent in environments where documentation is inconsistent, outdated, or poorly governed. The Support Engineering Manager echoed this point, noting that Copilot performs more reliably in enterprise-context queries—where the model is constrained to internal data—than in general web-based interactions.

A further theme relates to the continued need for human verification and critical oversight. Both participants observed that Copilot can occasionally produce outputs that are contextually plausible but technically incorrect, particularly when prompts are vague or when organisational documentation contains overlapping terminology. This verification burden supports the broader literature on generative AI, which emphasises the risk of overreliance on AI-generated content and the need for human judgement in high-stakes or technical environments. Rather than replacing human expertise, Copilot appears to redistribute cognitive effort toward evaluation, correction, and context-driven decision-making.

Training and AI literacy also emerged as critical determinants of Copilot's successful adoption. Both participants emphasised that Copilot "rewards" users who understand how to interact with it effectively. The implementation specialist stressed that enterprise-wide adoption requires structured onboarding, organisational champions, and clear use-case demonstrations. Without sufficient training, even highly capable users may struggle to extract value from Copilot, contributing to the neutrality and uncertainty observed in the survey results.

Another theme concerns security and trust, particularly around fears of data leakage or exposure. The interviews highlighted that secure adoption depends not only on built-in platform safeguards but also on appropriate access controls, governance policies, and organisational awareness. These findings help explain why Copilot is often preferred over general-purpose AI tools in enterprise environments.

Overall, the thematic analysis indicates that Copilot is experienced as a valuable but still maturing productivity tool. Its benefits are most visible in task initiation, summarisation, and information retrieval, while its limitations remain closely tied to data quality, user expertise, and the need for human oversight. These qualitative insights complement the quantitative survey results and are consistent with technology adoption frameworks that emphasise perceived usefulness, facilitating conditions, and organisational readiness.

Chapter 6: Conclusion and Recommendations

This dissertation examined the adoption of Microsoft Copilot in enterprise environments with the objective of understanding how it can be strategically integrated, which challenges organisations are likely to face, and why Copilot may be selected over other AI-based tools. By combining an exploratory survey with qualitative interviews and the analysis of documented implementation cases, the study provides an integrative perspective on early-stage enterprise adoption of AI copilots. The quantitative survey results offered an initial overview of user perceptions regarding productivity, benefits, and challenges associated with Copilot usage.

Given the limited sample size and the exploratory nature of the data, these findings should be interpreted as indicative rather than definitive. They highlight emerging patterns—particularly perceived efficiency gains in routine, document-oriented tasks—while also revealing persistent barriers related to accuracy, training, and workflow integration.

The qualitative interviews added depth and contextual understanding to these findings. They illustrate how Copilot is already supporting users in activities such as document drafting, information summarisation, and communication management. At the same time, the interviews underscore that Copilot's effectiveness is strongly influenced by organisational factors, including data quality, information governance, and user capability. Recurring concerns related to data privacy, hallucinations, and the need for human verification reinforce the conclusion that Copilot functions best as an assistive tool rather than as an autonomous decision-maker. Across both data sources, a consistent theme emerged regarding the importance of AI literacy and structured onboarding.

Many users lack a clear understanding of how to interact effectively with Copilot, and organisations often underestimate the effort required to support adoption through training, guidelines, and governance mechanisms. These findings align with established technology adoption theories, which emphasise the role of facilitating conditions, perceived ease of use, and organisational readiness in shaping successful uptake.

This study also presents several limitations. The exploratory design, small sample size, reliance on self-reported perceptions, and the rapidly evolving nature of Copilot constrain the generalisability of the findings. Furthermore, much of the available evidence on Copilot adoption originates from vendor-produced materials, which necessitates careful critical interpretation. As such, this dissertation does not aim to provide conclusive evidence of productivity impact or return on investment, but rather to identify key challenges, adoption conditions, and areas requiring further investigation. Despite these limitations, the dissertation contributes to both academic and practical discussions by synthesising current literature, user experiences, and implementation practices in a coherent analytical framework. Its primary contribution lies in clarifying the conditions under which AI copilots such as Microsoft Copilot can create value, as well as the risks and constraints that organisations must actively manage.

From a practical perspective, the findings suggest that organisations considering Copilot adoption should prioritise data governance, targeted deployment to suitable roles, structured training programmes, and clear usage policies. From a research perspective, future studies would benefit from larger and more diverse samples, objective productivity metrics, and longitudinal designs to assess how perceptions and impacts evolve over time.

In conclusion, Microsoft Copilot should not be viewed as a plug-and-play productivity solution, but as a socio-technical system whose value depends on strategic alignment, organisational readiness, and sustained human oversight. The true potential of AI in the workplace lies not in replacing human expertise, but in augmenting it responsibly, transparently, and ethically.

Bibliography

Alhur, A. (2024). *Redefining healthcare with artificial intelligence (AI): The contributions of ChatGPT, Gemini, and Copilot*. Cureus, 16(3), e57795. https://doi.org/10.7759/cureus.57795

Arsal, M., Saleem, B., Jalil, S., Ali, M., Zahra, M., Rehman, A. U., & Muhammad, Z. (2024). *Emerging cybersecurity and privacy threats of ChatGPT, Gemini, and Copilot: Current trends, challenges, and future directions* [Preprint]. https://doi.org/10.20944/preprints202410.1909.v1

Bano, M., Zowghi, D., Whittle, J., Zhu, L., Reeson, A., Martin, R., & Parsons, J. (2025). *Survey insights on M365 Copilot adoption* [Government report]. Australian Government – Digital Transformation Agency. [2412.16162] Survey Insights on M365 Copilot Adoption

Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). *Generative AI*. Business & Information Systems Engineering, 66(1), 111–126. [2309.07930] Generative AI

Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). *Generative AI and ChatGPT: Applications, challenges, and AI–human collaboration*. Journal of Information Technology Case and Application Research, 25(3), 277–304. (PDF) Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration

Füller, J., Hutter, K., Wahl, J., Bilgram, V., & Tekic, Z. (2022). *How AI revolutionizes innovation management – Perceptions and implementation preferences of AI-based innovators*. Technological Forecasting and Social Change, 178, 121598. https://doi.org/10.1016/J.TECHFORE.2022.121598

Hayawi, K., & Shahriar, S. (2024). *AI agents from copilots to coworkers: Historical context, challenges, limitations, implications, and practical guidelines* [Preprint]. https://doi.org/10.20944/preprints202404.0709.v1

Kyndryl. (2024). *Kyndryl launches services to speed adoption of AI-powered workplace solutions for Copilot for Microsoft 365*. https://www.kyndryl.com/us/en/about-us/news/2024/10/ai-workplace-services-Microsoft-365-copilot

Kytö, M. (2024). *Copilot for Microsoft 365: A comprehensive end-user training plans*. theseus.fi/bitstream/handle/10024/852578/Kyto_Miska.pdf?sequence=2&isAllowed=y

Malik, M. M. (2020). *A hierarchy of limitations in machine learning* [arXiv preprint]. http://arxiv.org/abs/2002.05193

Bjordal, M. H. (2025). *Unlocking organizational value through Microsoft 365 Copilot: A comprehensive analysis*

Miana, G., Paixão, M., Santos, B. C., Martins de Araujo, R., Ribeiro, M. H., Lopes de Moraes, J., & Ribeiro, A. L. (2022). *Machine learning in medicine: Review and applicability*. (PDF) Machine Learning na Medicina: Revisão e Aplicabilidade

Microsoft. (2025). *IU personalizes learning with Microsoft 365 Copilot*. Microsoft Customer Stories. https://www.Microsoft.com/en/customers/story/21865-iu-internationale-hochschule-Microsoft-365-copilot

Microsoft. (2025). *Oxford University Hospitals innovate with Microsoft Copilot to deliver efficiencies and improve workforce well-being*. Microsoft Customer Stories. https://www.Microsoft.com/en/customers/story/18859-oxford-university-hospitals-nhs-foundation-trust-Microsoft-365-copilot

Naveed, H., Ullah Khan, A., Qiu, S., Saqib, M., Anwar, S., Usman, M., Akhtar, N., Barnes, N., & Mian, A. (2023). *A comprehensive overview of large language models* [Preprint]. https://arxiv.org/abs/2307.06435

Farayola, O. A., Abdul, A. A., Irabor, B. O., & Okeleke, E. C. (2023). *Innovative business models driven by AI technologies: A review*. Computer Science & IT Research Journal, 4(2), 85–110. https://doi.org/10.51594/csitrj.v4i2.608

Pearce, H., Ahmad, B., Tan, B., Dolan-Gavitt, B., & Karri, R. (2021). *Asleep at the keyboard? Assessing the security of GitHub Copilot's code contributions*. Communications of the ACM. [2108.09293] Asleep at the Keyboard? Assessing the Security of GitHub Copilot's Code Contributions

Rashid, A. B., & Kausik, M. A. K. (2024). *AI revolutionizing industries worldwide: A comprehensive overview of its diverse applications*. Hybrid Advances, 7, 100277. (PDF) AI Revolutionizing Industries Worldwide: A Comprehensive Overview of Its Diverse Applications

Rodrigues Pontes Reis Miranda, R. (2025). *The impact and optimization of Microsoft Copilot in companies: Adoption, efficiency, ethical dilemmas, and user experience* (Master's thesis). NOVA IMS – Universidade Nova de Lisboa. run.unl.pt/bitstream/10362/190644/1/TGI4516.pdf

Saran, C. (2025). *Oxford University's IT department pilots Microsoft Copilot*. Computer Weekly. https://www.computerweekly.com/news/366617149/Oxford-universitys-IT-department-pilots-Microsoft-Copilot

Houghton, S. (2025). *What is Copilot Studio?* [Blog post]. https://www.aztechit.co.uk/blog/what-is-copilot-studio

Shanahan, M. (2024). *Talking about large language models*. Communications of the ACM, 67(2), 68–79. https://doi.org/10.1145/3624724

Sukamolson, S. (2007). *Fundamentals of quantitative research*. Fundamentals of quantitative research | Request PDF

Tenny, S., Brannan, J. M., & Brannan, G. D. (2022). *Qualitative study*. StatPearls. https://www.ncbi.nlm.nih.gov/books/NBK470395/

Microsoft. (2024.). *Unlock the power of Copilot in Excel, now generally available*. Microsoft Community Hub. https://techcommunity.Microsoft.com/blog/excelblog/unlock-the-power-of-copilot-in-excel-now-generally-available/4242810

Usman Hadi, M., Al-Tashi, Q., Qureshi, R., Shah, A., Muneer, A., Irfan, M., Zafar, A., Shaikh, M. B., Akhtar, N., Wu, J., Mirjalili, S., & Ali Al-Garadi, M. (2023). *Large language models: A comprehensive survey of applications, challenges, limitations, and future prospects* [TechRxiv preprint]. https://doi.org/10.36227/techrxiv.23589741.v4

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention is all you need*. Advances in Neural Information Processing Systems, 30, 5999–6009. [1706.03762] Attention Is All You Need

Zhang, B., Liang, P., Zhou, X., Ahmad, A., & Waseem, M. (2023). *Practices and challenges of using GitHub Copilot: An empirical study*. In Proceedings of the International Conference on Software Engineering and Knowledge Engineering (SEKE 2023) (pp. 124–129). https://doi.org/10.18293/SEKE2023-077

Appendices

Appendix A – Semi-Structured Interview Guide

Purpose of the Interview

The purpose of this semi-structured interview was to explore real-world experiences with Microsoft Copilot in enterprise environments. The interview aimed to understand perceived productivity impacts, adoption challenges, organisational readiness, and ethical or governance considerations related to the use of generative AI tools in professional settings.

The semi-structured format allowed flexibility to adapt follow-up questions based on the interviewee's role and experience, while ensuring consistency across interviews.

Interview Structure and Guiding Questions

Background and Role

1. Can you briefly describe your current role and professional background?
2. How are you involved with Microsoft Copilot in your day-to-day work?

Usage and Productivity

3. In which tasks or applications do you most frequently use Copilot?
4. How has Copilot affected your productivity or workflow?
5. In which situations do you find Copilot most useful, and where does it add less value?

Adoption and Organisational Readiness

6. What challenges have you observed when adopting Copilot in an organisational context?
7. How important is training and onboarding for successful adoption?

8. Do you think organisational readiness (data quality, processes, governance) influences Copilot's effectiveness? How?

Accuracy, Trust, and Human Oversight

9. Have you encountered inaccuracies or hallucinations when using Copilot?
10. How do you approach verification and responsibility for AI-generated outputs?
11. In your opinion, what level of human oversight is necessary when using Copilot in professional contexts?

Security, Ethics, and Governance

12. What concerns do organisations typically raise regarding data security and privacy?
13. How are these concerns addressed in practice?
14. Do you see ethical or accountability challenges emerging as Copilot becomes more capable?

Strategic and Future Outlook

15. How do you see the role of Copilot evolving in the future?
16. Would you recommend Copilot for all employees, or only for specific roles? Why?

Closing Question

17. Do you have any final advice or recommendations for organisations considering adopting Microsoft Copilot?

Notes

- All interviews were conducted with informed consent.
- Interviewees were free to skip any question.
- The guide was used flexibly, allowing follow-up questions where relevant.

Appendix B – Interview Transcript: Support Engineering Manager

The transcript was edited for clarity by removing filler words and repetitions. The semantic content was not altered.

Interview Type: Semi-structured interview

Participant Role: Support Engineering Manager (Multinational organization)

Interview Date: April 2025

Interview Duration: Approximately 43 minutes

Interview Mode: Microsoft Teams (recorded with consent)

Interview Transcript

Introduction

Interviewer:
Thank you for agreeing to participate in this interview. To begin, could you briefly describe your professional background and how you currently interact with Microsoft Copilot?

Participant:
I work as a Support Engineering Manager in a large multinational organisation. I use Microsoft Copilot primarily as an end user, not from a configuration or licensing perspective, but from a practical, day-to-day usage standpoint. My experience is focused on how Copilot supports my work in communication, documentation, and information retrieval within an enterprise environment.

Enterprise Context and Data Governance

Interviewer:
From an enterprise perspective, what are the key considerations when deploying Copilot?

Participant:
In a corporate environment, Copilot has access not only to web information but also to internal organisational data, such as documentation stored in SharePoint and OneDrive. This makes data governance critical. Organisations must carefully review access permissions and ensure that sensitive information is properly classified. If documents are incorrectly labelled or broadly accessible, Copilot may surface information that should not be visible. This is particularly important in regulated sectors such as banking, where strict compliance requirements apply.

Usage Scenarios and Productivity Support

Interviewer:
How do you personally use Copilot in your daily work?

Participant:
I primarily use Copilot as a writing and summarization assistant. For example, it is extremely useful for summarizing internal documentation, drafting structured texts, and supporting performance review. In my role, I regularly need to prepare performance evaluations, executive summaries, and internal reports. Copilot helps by generating structured drafts based on internal guidelines, which I then review and refine.

Email communication is another major use case. In customer-facing roles, communication must be clear and diplomatic. Copilot's review and rewrite suggestions help improve tone and clarity, especially in complex or sensitive situations.

Application Usage

Interviewer:
Which Microsoft applications do you most frequently use Copilot with?

Participant:
Outlook is probably the most frequent, due to constant communication with customers and internal teams. I also use Copilot with Word for documentation and with PowerPoint when preparing presentations.

For PowerPoint, Copilot is particularly helpful in the initial phase, where it can generate a first draft of slides based on a Word document or outline. This saves time during the setup phase, even though the content always requires review and refinement.

Accuracy, Hallucinations, and Verification

Interviewer:
Do you experience errors or hallucinations when using Copilot?

Participant:
Copilot has improved significantly over time, and hallucinations are less frequent than they were initially. However, they still occur, especially when Copilot is operating in a web context rather than using internal organisational data.

In enterprise scenarios, accuracy is generally higher because Copilot is constrained to internal documentation. Even so, verification is always necessary. For example, when dealing with technical error codes that exist across multiple systems, Copilot may confuse contexts unless prompts are very precise.

Trust and Human Oversight

Interviewer:
How do you approach trust and responsibility when using Copilot?

Participant:
Human oversight is essential. Copilot provides suggestions, but responsibility always remains with the user. I would never send content externally without reviewing it.
This is especially important in support environments, where incorrect information can have serious consequences. Copilot often provides sources, which helps, but users must still critically evaluate the output.

Training and AI Literacy

Interviewer:
How important is training in Copilot adoption?

Participant:
Training is critical. Copilot rewards users who understand how to interact with it properly. Prompt formulation is a skill, and organisations need to invest in training to help employees develop AI literacy.
Without training, users may either misuse the tool or avoid it altogether due to lack of confidence.

Outlook and Workforce Impact

Interviewer:
Do you see Copilot as a threat to jobs?

Participant:
I don't see it as a direct replacement for professionals, particularly in roles that require human interaction and contextual judgment. Copilot automates parts of tasks, not entire roles.
However, people who refuse to engage with AI tools may struggle in the future. AI literacy will

become an important professional skill, like how calculators became essential tools rather than threats.

Closing Remarks

Interviewer:
Do you have any final advice for organisations considering Copilot adoption?

Participant:
Organizations should invest in training, governance, and communication. Copilot has significant potential, but it must be implemented responsibly. It is not a magic solution, but a tool that can meaningfully augment human capabilities when used correctly.

Appendix C – Interview Transcript: Copilot Implementation Specialist

The transcript below was edited for clarity by removing filler words, repetitions, and transcription artefacts from the automatic recording. The semantic content of the interview was not altered. The participant requested that the discussion be treated as a professional perspective and not as an official company position.

Interview Type: Semi-structured interview

Participant Role: Copilot Implementation Specialist

Interview Date: February 2025

Interview Duration: Approximately 38 minutes

Interview Mode: Microsoft Teams (recorded with consent)

Interview Transcript

Introduction and Professional Context

Interviewer:
You work closely with customers implementing Copilot solutions. From your experience, what are the main challenges organisations face when adopting Microsoft Copilot?

Participant:
I usually start with a very simple question that often turns out to be quite difficult: *why* is the organisation purchasing Copilot, and what do they expect it to do once it is implemented? When organisations try to answer this, they often receive conflicting responses from different teams. This matters because Copilot depends heavily on organisational readiness, particularly data quality.

Data Hygiene and Organisational Readiness

Participant:
Copilot requires clean and well-managed data to work effectively. That includes SharePoint and OneDrive environments, which are often messy in many organisations.

Improving data hygiene is not something that can be done quickly. Cleaning up documents, access permissions, and metadata often requires weeks or months of coordinated effort. This is why we spend a significant amount of time helping organisations prepare their environments before or during Copilot deployment.

Training and Change Management

Interviewer:
How important is training when deploying Copilot?

Participant:
Training is essential. Organisations need to decide how they want employees to work with Copilot and then train them accordingly.

Without training, users may misuse the tool, misunderstand its limitations, or fail to extract meaningful value. These conversations alone—data readiness, application hygiene, and training—often require at least a month of preparatory work.

Security and Trust Considerations

Interviewer:
Many organisations are concerned about data security. How do you address these concerns?

Participant:
One of the main reasons organisations choose Microsoft Copilot is the built-in security model. Data used by Copilot is not sent to public models and is not used to train external large language models.
These guarantees are critical, especially for organisations that do not want to build their own AI infrastructure. Compared to developing custom solutions using platforms such as Azure OpenAI, Copilot offers a cost-effective and secure alternative.

Day-to-Day Impact and Productivity

Interviewer:
How has Copilot impacted your own daily work?

Participant:
Copilot has become a natural part of how I work. It removes the "blank page" problem by generating an initial draft, which allows me to focus on creativity and refinement.
It helps summarise large documents, identify trends, and proofread content. I often describe it as having an assistant or intern for every employee. In many cases, it reduces the effort required to complete tasks by 20–40%.

Copilot Studio and AI Agents

Interviewer:
Do you use Copilot Studio? How does it differ from standard Copilot usage?

Participant:
Yes, I use Copilot Studio. It allows organisations to extend Copilot's capabilities by connecting it to specific data sources, workflows, and automation tools.

With Copilot Studio, organisations can create customised agents focused on specific tasks, such as helpdesk support or human resources queries. These agents use curated datasets and can be shared across teams, rather than being limited to individual users.

Comparison with Public AI Models

Interviewer:
How does Copilot compare to tools like ChatGPT or Gemini?

Participant:
Public AI models tend to be less restricted. Microsoft Copilot includes a responsible AI layer that limits certain outputs to reduce organizational risk.

While this can restrict some responses, it provides a safer environment for enterprise use. Copilot is designed to prioritise governance, compliance, and transparency over unrestricted generation.

Future Evolution of Copilot

Interviewer:
What do you expect from Copilot in the future?

Participant:
Copilot is moving beyond static text generation toward actionable intelligence. It is beginning to take actions, such as scheduling meetings or updating documents automatically.
In the future, Copilot will likely have greater memory, improved contextual awareness, and the ability to perform recurring tasks autonomously. However, this also increases the importance of governance and oversight.

Licensing Strategy and Target Users

Interviewer:
Do you recommend Copilot for all employees?

Participant:
Not necessarily. Copilot is most valuable for users who create content—documents, code, reports, or analyses.

Because it has a cost, organisations should focus deployment on roles where it delivers the greatest value, such as knowledge workers, technical teams, and departments like HR or finance.

Closing Remarks

Interviewer:
Do you have any final recommendations for organisations considering Copilot adoption?

Participant:
Copilot offers clear advantages for organisations already using Microsoft 365, but it should be deployed strategically.

Successful adoption requires clear use cases, clean data, structured training, and careful governance. Copilot is not a universal solution, but when implemented thoughtfully, it can significantly improve productivity and decision-making.

Appendix D – Online Survey

This appendix presents the survey used in the exploratory survey on Microsoft Copilot adoption. The survey aimed to capture user perceptions regarding productivity, adoption challenges, and organisational readiness.

Q1. What is your job role? (Open-ended)

Q2. What industry does your company belong to? (Multiple choice)

Q3. How large is your company? (Multiple choice)

Q4. How frequently do you use Microsoft Office 365 in your work? (Multiple choice)

Q5. Are you familiar with Microsoft Copilot? (Yes / No)

Q6. If you have used Microsoft Copilot, which applications have you used it with? (Multiple choice – multiple answers allowed)

1. Word
2. Excel
3. PowerPoint
4. Outlook
5. Teams

Q7. How has Microsoft Copilot affected your productivity? (Multiple choice)

Q8. What are the main benefits of using Copilot in your work? (Multiple choice – multiple answers allowed)

Q9. What challenges have you encountered while using Microsoft Copilot? (Multiple choice – multiple answers allowed)

Q10. How concerned are you about data security and privacy when using AI tools like Microsoft Copilot? (Likert scale)

Q11. Does your company provide guidelines or policies on using Microsoft Copilot? (Yes / No / Not sure)

Q12. Do you believe Microsoft Copilot should be fully integrated into your company's workflows? (Yes / Yes, with guidelines / No)

Q13. What improvements would you like to see in Microsoft Copilot? (Open-ended)

Q14. Would you recommend Microsoft Copilot to other professionals in your industry? (Yes / No / Not sure)