Large Language Models in Dutch Healthcare

Identifying the state of an early-stage Technological Innovation System

By Felix Filippini

Large Language Models in Dutch Healthcare

Identifying the state of an early-stage Technological Innovation System

MASTERS THESIS

BY

W.F.J. Filippini
5885671

Supervisor: Dr. Iryna Susha
Second Assessor: Dr. Giovanna Capponi

MSc. Innovation Sciences
Faculty of Geosciences
Utrecht University

Internship Supervisor: Drs. Anniek Stroomer

Capgemini Invent
Intelligent Health

Utrecht University

Capgemini invent

Contact details

Felix Filippini	(w.f.j.filippini@students.uu.nl)
Dr. Iryna Susha	(i.susha@uu.nl)
Dr. Giovanna Capponi	(g.capponi@uu.nl)
Drs. Anniek Stroomer	(anniek.stroomer@capgemini.com)

ABSTRACT

Technological innovation may offer a solution for the urgent problem of surging healthcare expenditures in Dutch healthcare. With the arrival of ChatGPT, the focus on artificial intelligence, particularly on the utilization of Large Language Models (LLMs) in healthcare, intensified. Nonetheless, the technological realities in the current state-of-the-art of LLMs in healthcare, in combination with the unique context of the Dutch healthcare sector, present considerable challenges for implementation.

In order to shed light on the transformative abilities of LLMs in healthcare in the coming years, this report uses the Technological Innovation System (TIS) analysis framework. The framework was operationalized to examine its systemic attributes and the potential emergence of systemic strengths and weaknesses, acknowledging the nascent nature of the technology. The Technological Innovation System of LLMs in healthcare is still in a developmental stage.

The theoretical aim was to extend the TIS framework, often used in sustainable technologies, to a case in the healthcare sector and to aggregate the current findings on early-stage TISs, applying them to the case context. The TIS framework provided multiple analytical data collection steps. First delineating the exact boundaries of the technological innovation system then, identifying the most important structural elements and finally the level of seven functional elements. Data was collected by a triangulation of desk research and two rounds of semi-structured interviews.

The results show how the Dutch Healthcare Large Language Model TIS (DH-LLMTIS) is dependent on the properties of the further developed Dutch Healthcare Artificial Intelligence TIS (DH-AITIS) in all but three functions, in a way that is complementary and not competitive. The global developments of Large Language Models in healthcare influenced the case context across four of the seven functional domains. Further analysis and comparison of the functions shows how the Dutch Healthcare Large Language Model TIS is characteristic of an early-stage formative TIS, moving towards a growing phase but held back mainly on three points. These points are: I. an absence of convergence of knowledge diffusion II. lack of unified regulations and industry standards III. Insufficient convergence of physical resource mobilization.

The results show how the TIS framework can be applied in an emerging healthcare technology context, emphasizing earlier findings on functions in early-stage TISs. The entire report provides a comprehensive overview of the elements decision-makers in the Dutch healthcare must consider to position themselves correctly to implement Large Language Models in the near future.

EXECUTIVE SUMMARY

Internship Context and Objective: This research was undertaken during a five-month internship within the 'Intelligent Health' Team of Capgemini Invent. The 'Intelligent Health' team is at the forefront of implementing digital healthcare innovations in the Dutch healthcare sector, both in the public and private sphere. With the emergence of generative AI and Large Language Models (LLMs) possibly significantly influencing the Dutch healthcare landscape, this research aims to provide the team with a comprehensive overview and deep insights into the latest technological capabilities of LLMs and the unique characteristics of the Dutch healthcare context.

Research Methodology: The study employs a modified Technological Innovation System (TIS) analysis framework to investigate systemic transitions and the emergence of technological innovations, specifically focusing on the Dutch Healthcare Large Language Model Technological Innovation System (DH-LLMTIS). This exploratory approach incorporates an analysis of academic literature, policy documents, yearly reports, web databases, news articles, and a bibliometric analysis. The research is further enhanced by two rounds of interviews with 21 key stakeholders, offering a validated perspective on the structural and functional aspects of the DH-LLMTIS.

Key Findings: A comprehensive assessment of the current state of LLMs in healthcare, outlining their capabilities and limitations is given. The crucial interdependencies between DH-LLMTIS, the broader AI Technological Innovation System (AITIS), and the Global LLM Technological Innovation System (GLLMTIS) are hightlighted. Notable observations include DH-LLMTIS's dependence on DH-AITIS in all areas except 'entrepreneurial activity', 'knowledge development', and 'legitimization', and its influence by G-LLMTIS in four functions; 'entrepreneurial activity', 'knowledge development and diffusion', 'guidance of search through regulations', and 'legitimization'. The study benchmarks the TIS's functional performance against expected developmental stages, identifying a transition from 'nascent' to 'formative' stages, with notable gaps hindering progression to the 'growth' stage. Key development challenges include the need for convergence in knowledge diffusion initiatives, streamlined communication of regulations and standards, and improved data mobilization as resource through policy and regulations.

Implications for the team: Incorporating the insights in this report will enable members of the ‘Intelligent Health’ team to purposively direct efforts into the right direction when being confronted with LLMs in their work, for which the comprehensive overview of the DH-LLMTIS can serve as reference work . The technological background and functional analysis underscore LLMs' potential for non-medical, intra-organizational, non-generative applications in the short term. Furthermore, the report shows the current status and maturity of attention given to AI and LLMs in actors and clients well known by the team. To conclude, the current expertise of the team on data availability, compliance with healthcare regulation and integration of European regulatory harmonization, was confirmed as being important for the development of a future DH-LLMTIS, parallel to its current importance for digital transformations that improve Dutch Healthcare.

List of abbreviations

AI	Artificial Intelligence
AP	Autoriteit Persoonsgegevens
DH- LLMTIS	Dutch Healthcare Large Language Model Technological Innovation System
DH-AITIS	Dutch Healthcare Artificial Intelligence Technological Innovation System
EHDS	European Health Data Space
EZK	Ministry of Economic Affairs
GDPR	General Data Protection Regulation
GenAI	Generative Artificial Intelligence
G-LLMTIS	Global Large Language Model Technological Innovation System
GPU	Graphic Processing Unit
IGJ	Health Inspection (Inspectie Gezondheidszorg en Jeugd)
IS	Innovation Systems
LLM	Large Language Model
MDR	Medical Device Regulation
MLP	Multi-Level-Perspective
NLP	Natural Language Processing
TIS	Technological Innovation System
TKZ	Top Clinical Hospital
UMC	University Medical Center
VWS	Ministry of Health

CONTENT

ABSTRACT 3

EXECUTIVE SUMMARY 4

LIST OF ABBREVIATIONS 5

CONTENT 6

1 INTRODUCTION 7

2 TECHNOLOGICAL BACKGROUND & CASE CONTEXT 11

2.1 ARTIFICIAL INTELLIGENCE AS EMERGING FIELD 11

2.2 NATURAL LANGUAGE PROCESSING THROUGH LLMS 12

2.3 LARGE LANGUAGE MODELS IN HEALTHCARE 13

2.4 A BRIEF OVERVIEW OF AI IN DUTCH HEALTHCARE 14

3 THEORY 16

3.1 INNOVATION SYSTEMS & TIS 16

3.2 AN INTEGRATED FRAMEWORK 18

4 METHODOLOGY 23

4.1 RESEARCH APPROACH 23

4.2 TIS ANALYSIS STEPS 23

4.3 DATA-COLLECTION 24

4.4 ANALYSIS 31

4.5 RESEARCH QUALITY INDICATORS 32

5 RESULTS 33

5.1 STRUCTURAL ANALYSIS 33

5.2 FUNCTIONAL ANALYSIS 38

6 DISCUSSION 64

6.1 DISCUSSION OF FINDINGS 64

6.2 THEORETICAL CONTRIBUTIONS & FUTURE RESEARCH 68

6.3 LIMITATIONS 71

6.4 PRACTICAL IMPLICATIONS 73

7 CONCLUSION 74

8 ACKNOWLEDGEMENTS 77

9 REFERENCES 78

APPENDIX A: INTERVIEW GUIDES 83

APPENDIX B: SEARCH QUERIES 86

APPENDIX C: CODING SCHEMES 89

1. INTRODUCTION

Demand for healthcare in the Netherlands, just as in Europe, is ever-rising. Contributing factors such as aging populations, unhealthy lifestyles, and expectations for ever-improving quality suggest that it is unlikely that this upward trend will soon reverse (EPRS, 2022; Rivm, 2018). Furthermore, as an unsurprising consequence of these increases, expenditures have risen substantially for decades (CBS, 2023; Eurostat, 2023). In the Netherlands, these costs are expected to rise gradually to €174 billion a year by 2040, doubling healthcare expenditures since 2015 and making the current trajectory unsustainable (Wouterse, 2017). Accordingly, cost containment has been one of the key priorities of the Dutch healthcare policies (Kroneman et al., 2016; Ministerie van Volksgezondheid Welzijn en Sport, 2022a; Rivm, 2018).

A way of coping with increasing demands and quality, without a cost increase, is through technological innovation (Ministerie van Volksgezondheid Welzijn en Sport, 2022a; Thimbleby, 2013). As we entered the age, some call the fourth industrial revolution (Lee et al., 2018), one of the fields of technological advancement that stands out, is that of artificial intelligence (AI). AI – a system's ability to understand and manipulate large data sets, mimicking human cognitive functions – is a concept that is both widely discussed and hard to grasp. While the term has been existing for decades, the sharp increase in applications and capabilities of AI-powered systems, brought to attention a 'spring' of AI that has already been in the making for years (Kaplan & Haenlein, 2019; McCarthy et al., 1955; The Economist, 2023a). Most notably is the introduction of 'ChatGPT' (a large language model (LLM) capable of impressive feats of natural language processing (NLP), more on that later) to the larger public. Models like ChatGPT belong to a class of AI able to generate humanlike content without predefined programming being thus called Generative AI or GenAI (Dietterich et al., 2022).

The promises that LLMs like ChatGPT hold, are thought to change virtually every sector of society, with healthcare often being named as top 3 of most affected (McKinsey, 2023; Rikap & Lundvall, 2021). LLMs are now named in an array of functionalities in medical-professional and patient facing applications using their NLP functions such as clinical documentation, summarizing research papers, creating discharge summaries and interpreting physicians notes (Meskó & Topol, 2023).

However potentially beneficial, the reality of putting these applications into practice in healthcare is subject to numerous obstacles, and it is not a given that it will succeed (He et al., 2019; Kelly et al., 2019; Economist 2023). Research has pointed out the tricky nature of technological innovation in hospital settings as it involves a large variety of stakeholders in a strongly regulated and institutionalized sector (Greenhalgh & Abimbola, 2019). For the roll-out of medical AI applications high-quality data availability is named as major hampering factor. Data sources are fragmented, unharmonized, and bound to multiple laws and regulations regarding patient privacy and data sharing (de Kok et al., 2023). The insufficient trust and legitimacy LLM applications enjoy is emphasized as well (Lambert et al., 2023). Unexplainable elements of algorithms or 'black boxes' are not tolerable for healthcare professionals who will therefore not alter their workflow because of them. One of the reasons for this, is the knowledge gap between the healthcare workforce and AI developers in general (Bienefeld et al., 2023). Already some of the expectations of medical AI applications are tempered such as that more data automatically equates to superior models because of lacking data quality, or that improved clinical performance leads to instant clinical confidence (Widner et al., 2023). Improving these processes demands time and scrutiny. Analyses of AI in the Dutch healthcare setting confirm similar barriers (Berenschot, 2022; Cap Gemini invent, 2020; KPMG, 2020; Wouda et al., 2019).

The problems in these articles point to the fact that technological advancements in themselves are not enough to foster technological change, but are dependent on multiple factors that make up the socio-technical 'system' in which the technology operates and that current literature on AI in healthcare does not take these factors into account often enough (Zahlan et al., 2023).

Because the complex organizational set-up, legal requirements and technological specification in the healthcare field, innovation in this sector has been the subject of several strands of theoretical frameworks. These frameworks take on different approaches and scopes (Moullin et al., 2015). For example, the widely used NASSS-framework[1] exists to study the implementation of healthcare technologies in a specific clinical setting within a healthcare organization including non-technology factors using seven 'domains'. However, while these frameworks examine relevant factors for a case study in a single organization, the analysis of the wider network of organizations and infrastructures is not the focus (Greenhalgh et al., 2017). Because the wide range of possible use cases of LLMs in healthcare, the strong interorganizational setting and its distinct global dependencies a more generic systemic transition framework can provide insights for the Dutch case (Rajpurkar et al., 2022).

Currently, several theories are dominantly in use to study technological transitions in systems and the adoption of technological innovations in organizations. Of these, the technological innovation system (TIS) analysis approach emerged as a prominent theory to capture the system's performance around one specific technological innovation (called the 'focal' technology) and is found to be mentioned in literature on AI innovations (Mariani et al., 2022). It can be defined as the set of actors, institutions, networks and physical infrastructure that form structural system elements that influence the speed and direction of technological change in a specific technological area (Wieczorek et al., 2013). Rooted in evolutionary economics, it highlights the importance of dynamic interplay between these 'structural' elements. Several key processes that take place between the structural elements were distilled, called 'functions', to measure the system's performance. In a TIS, all functions measured by performance indicators must show a minimal level of activity and quality for the system to perform well (Markard, 2020). The list of functions that are measured is distilled from various earlier innovation system literature and proved to be a strong basis for the assessment of one TIS or the comparison of different TISs (Bergek et al., 2008; Hekkert et al., 2007; Wieczorek et al., 2015).

As of now, limited research is done on the wider systemic aspects of LLM adoption in healthcare, using an overarching theory on technological transitions in systems. However, while applicable, TIS analyses are often focused on technologies with a sustainability aim, leaving other sectors underdeveloped (Boons & McMeekin, 2019). The TIS framework was never developed with a specific focus on sustainability, one of the reasons why a separate framework, the mission-oriented innovation systems (MIS) covering innovation steered by the achieving of societal goals, was recently conceptualised (Hekkert et al., 2020). This does not take away the usage of the TIS as a framework to study the state of emerging technologies not aimed towards a sustainability goal, but emerging nevertheless.

Therefore, in the fast-changing field of LLMs and their potential use in Dutch healthcare, it is relevant to take a closer look at a potentially emerging TIS. To the author's knowledge, no systemic analysis was performed for LLMs in healthcare in the Netherlands, which we will call the Dutch Healthcare Large Language Model Technological Innovation System or 'DH-LLMTIS'. Because the exploratory nature of the technology, analysing this TIS will require taking into account both the worldwide technological developments in the LLM field and the Dutch AI wide systems that are in place. This approach will try to get more grip on the state of an elusive field surrounded by high expectations and valid concerns (Sallam, 2023).

[1] The NASSS framework stands for non-adoption, abandonment, scale-up, spread, sustainability framework and was posed by Greenhalgh in 2017

While earlier work was done on the TIS of AI, it focused on the global system of innovation in the AI field itself that showed to be dominated by the US and China and their 'big tech' companies. It is through the global spread of digital innovations developed by these large incumbent companies that local transitions such as in the Netherlands are influenced as well, creating a distinct dynamic between local and global systems. In addition, a purely global view does not provide enough basis for understanding the dynamics of the implementation of LLM applications in a specific sectorial and geographical setting (Rikap & Lundvall, 2021).

To this date, one TIS analysis was performed on the use of healthcare AI from a life sciences perspective in the case of West-Sweden. This study, while paving the way for the analysis of a Dutch Healthcare Large Language Model Technological Innovation System (DH-LLMTIS) within the healthcare sector, was executed without fully integrating the TIS into the broader context of innovation theory (Apell & Eriksson, 2023). Furthermore, it underutilized the use of the different TIS phases and standard dynamics that literature describes or the comparison between TISs to interpret its findings (Bergek et al., 2008; Markard, 2020). The novel nature of LLM use in healthcare requires evaluation with earlier research on TISs in a formative and emerging situation (Makard 2020, Suurs 2010, Musiolik 2020). In addition, while the contextual delineation is an important part of a TIS analysis as signified by Markard et al. (2015), it is not included in the analysis of Apell & Eriksson (2023). Nonetheless, given that this research direction is relevant for the assessment of the Dutch LLMTIS, the research findings need to be integrated into new research to study the overlap with the Dutch context.

As a result, in this paper the technological innovation systems framework is adapted for an explorative interdisciplinary empirical analysis of the current developments of LLMs in healthcare in the Netherlands. LLMs are understood as a potential technology contributing to the quality and cost efficiency of the Dutch healthcare system through their NLP capabilities. The aim of the paper is to conceptualise and empirically describe the current state of its developments and based on that, describe the functioning of an early-stage TIS in the healthcare sector. The focus is on the Netherlands as a case study while the relevant worldwide developments of AI and specifically LLMs in healthcare are considered. This leads to the following research question:

RQ: What is the current state of the early-stage Technological Innovation System of Large Language Models in Dutch Healthcare?

To answer this research question, the following sub questions are answered:

- ***SQ1: What is the delineation and context of the Dutch Large Language Model Technological Innovation System in healthcare?***

- ***SQ2: What are the relevant structural elements in the Dutch Large Language Model Technological Innovation System regarding actors, networks, institutions and infrastructure?***

- ***SQ3: What is the performance of the different Technological Innovation System functions of the Dutch Large Language Model Technological Innovation System in healthcare?***

From a practical perspective, the research aims to assist policymakers in a healthcare setting in preparing for and managing the change processes linked to the emergence of LLMs in their field. If the actors of the Dutch LLMTIS are well informed about its state, it allows them to concentrate their efforts in the right place thereby improving the speed and direction of the implementation. When successfully implemented, these technological innovations promise better quality and efficiency in healthcare delivery, a positive prospect for both the healthcare sector that is under pressure, and society as a whole.

From a theoretical perspective, this research aims to add to the understanding of the dynamics in an early-state TIS and to test the validity of standard patterns of system functions, as described in earlier research. The case can be defined as a digital technological innovation that is dependent on multiple socio-technical systems that are in place in the Netherlands and worldwide. In doing so, the use of the TIS framework was extended and refined, bridging current gaps in the literature and guiding future research studying the systemic characteristics of emerging digital technology-related innovations in the medical field, with an emphasis on LLMs, inside and outside the Dutch context.

The remainder of the paper is structured as follows: Chapter 2 concerns the technological background and context of the focal TIS. Chapter 3 consists of the theoretical framework and proposed theoretical model. Chapter 4 addresses the methodology used for the data collection and analysis. Chapter 5 consists of the results of a 'structural' and 'functional' TIS analysis. Chapter 6 are the theoretical implications, discussion, limitations and policy implications. Chapter 7 consists of the conclusion and finally Chapter 8 consists of the references.

2 TECHNOLOGICAL BACKGROUND & CASE CONTEXT

2.1 ARTIFICIAL INTELLIGENCE AS EMERGING FIELD

Artificial intelligence can be defined as both a knowledge field and as a general-purpose technology with wide potential fields of application. While the field has a long and extensive history, modern-day artificial intelligence (AI) can be said to have started in 1956 with an academic summer project at Dartmouth College New Hampshire. Here, the foundations were laid for what the goals and directions should be of mimicking human cognition by computers. In these early days, they called out on the possibility of deconstructing language as a path towards creating AI as well as the use of imitations of human neural networks and the possibilities of 'self-improvement' by a machine. These are developments that shaped the AI models of today (McCarthy et al., 1955). Fast-forward to the 21st century, and the increased computational power of electronic hardware through the availability of graphics processing units (GPUs)[2], the increase of available data and new insights in model architecture, have come a long way in meeting those goals proposed in the summer of 1956.

Most current models are powered by machine learning, a subfield of AI that started in the 1990s, where predictive models automatically distill data-driven rules from large data sets without explicit specification from humans. From 2010 onwards, the re-introduction of the old concept of 'neural networks'[3] under the name of 'deep learning' gained progressively more traction showing spectacular results in a wide array of function and dominating the latest promising models (Bommasani et al., 2021). Neural networks are a model architecture that makes deep learning possible under the right conditions. In essence, it is a complex extension of the principles of linear regression, capturing the relationships between input variables and outcomes. They consist of an input layer, a hidden layer of 'neural nodes' and an output layer leading to outcomes that can be useful across a variety of tasks. Deep learning models consists of multiple hidden layers, making it possible to explore increasingly complex data relationships and indirect data patterns (Lecun et al., 2015). The capabilities of deep learning are only possible through a spectacular increase in computational power, mitigated by hardware innovations such as the 10-20 times increased capabilities of GPU's between 2018 and 2022 (Bommasani et al., 2021).

An essential development in the model architecture of neural networks, was the introduction of the transformer model in 2017 (Vaswani et al., 2017). It added a mechanism of continuous self-attention to capture the dynamics between data points of further removed position within the input data set. This led to a better representation of related datapoints in outputs, an ability that propelled the field of understanding language through natural language processing (NLP) and the field of computer vision forward (Wei et al., 2022). These new models are sometimes named foundational models, because they can have a relatively homogenous architecture (foundation) while being able to process a variety of input data (words, images, sounds, etc.) and using them for a wide range of downstream tasks (Moor et al., 2023).

---

[2] GPUs are used to perform large amounts of parallel computations also known as 'embarrassingly parallel' that are needed in training neural networks and LLMs

[3] Deep learning is understood as: models trained on raw inputs where higher-level capabilities emerge through training. Rather than having bespoke feature engineering pipelines for each application, the same deep neural network architecture can be used for many applications (Bommasani et al., 2021).

Furthermore, by the increased and progressive collection of data, data became more widely available. All these developments have led to two important concepts that took hold of the AI field, 'emergence' and 'homogenization'. Emergence is the behavior of a system that is implicitly induced instead of explicitly constructed, and homogenization refers to the unification of methodologies over a broad spectrum of applications, as seen in foundational models. Both concepts signal a gained momentum in AI research (Bommasani et al., 2021; Davenport & Kalakota, 2019; Esteva et al., 2019; Jiang et al., 2017). As of 2020, twenty-nine AI applications were approved for healthcare use by the FDA, mostly in the fields of radiology, oncology and cardiology. The first application was approved in 2016, showing its recent and fast development in the US market as market leader (Benjamens et al., 2020).

2.2 NATURAL LANGUAGE PROCESSING THROUGH LLMS

Within the wide array of applications that fall under AI, understanding human language has always been pivotal, reflecting its fundamental role in human intelligence (Davenport & Kalakota, 2019). NLP is the ability to interpret, manipulate and generate language in ways that resemble language generated by humans. Examples of these applications are: sentiment classification (predicting if a piece of text is positive or negative about a subject), sequence labelling (differentiating between a verb and noun or indicate all words that refer to subject), relation classification (determining the linkages between words) or generation tasks (producing new text on a specific subject). The advancements in fields such as neural networks, deep learning, transformer architecture described above, have made a profound impact on the field of NLP. Consequently, it marked the advent of the type of model coined large language models (LLMs)[4] (Wei et al., 2022). Given the complexity of human language and the interconnection of words, the models are trained on large amounts of data in a process that is called 'pre-training'. Pre-training is computational expensive but does not require the human labelling of data. Instead, the models derive patterns from the data without initial goals. In the process of 'fine-tuning', the model is adapted to perform specific tasks using a smaller body of labelled data in process that is less computationally expensive than pre-training (Wei et al., 2022). This fine-tuning used to be done solely using manually annotated data, which is laborious and carries significant costs. However the fine-tuning of the new models is done without human interference using unannotated data ('self-supervised') or with a combination of annotated and unannotated data ('semi-supervised'), drastically speeding up the process. One LLM can be fine-tuned for multiple tasks after pre-training. Furthermore, Brown et al. (2020) showed that when language models increase in parameter size, the ability to 'learn' from a decreasing amount of examples increased know as 'few-shot learning', adding to the emergent capabilities of the model.

As shown, the concepts of Natural Language Processing, Large Language Models, generative AI and foundational models are closely interlinked and sometimes used interchangeably. In this research, when is referred to Large Language Models, the more recent models after 2017 expressing high quality NLP capabilities are meant .

The concepts of 'generation', 'pre-training' and 'transformer' all come together in the 'GPT' model architecture. The first version of the GPT architecture was launched in 2018 and consisted of some 1.5 billion parameters. In 2020, a short time later, GPT3 was released including 175 billion parameters. The order of magnitude of these models to run within an acceptable timeframe was previously unthinkable, resulting in the prefix 'large' in LLMs. On November 30th 2022, a finetuned version of the GPT3 architecture, GPT-3.5, was released by the company OpenAI as 'ChatGPT', attracting over 100 million users within two months (Kung et al., 2023; Thirunavukarasu et al., 2023).

[4] LLMs are defined as a foundational models utilizing deep learning to perform NLP & Natural language generating (NLG) tasks

In the last year, ChatGPT has become the most well-known, but not the only, example of a LLM that has been developed. Multiple parties developed their own LLMs, all with parameters in the order of billions. However, size is not the only key to success. For example, some earlier version of the ChatGPT model by OpenAI outperformed models multiple times its size on standard operating tasks such as email writing (Thirunavukarasu et al., 2023). It is noteworthy that the infrastructure and energy costs associated with operating these models are substantial. The latest big models require state-of-the-art computational hardware and considerable electrical power (Peng et al., 2023; Thirunavukarasu et al., 2023). This means that hardware providers play an important role in the value chain of LLMs and steers the research on LLMs towards reducing the size and costs of newer models. **Figure 1** presented below provides a simplified depiction of the relevant concepts surrounding LLMs, three types of foundational models are used as examples namely: '*ChatGPT*', '*PaLM*', '*BERD*'.

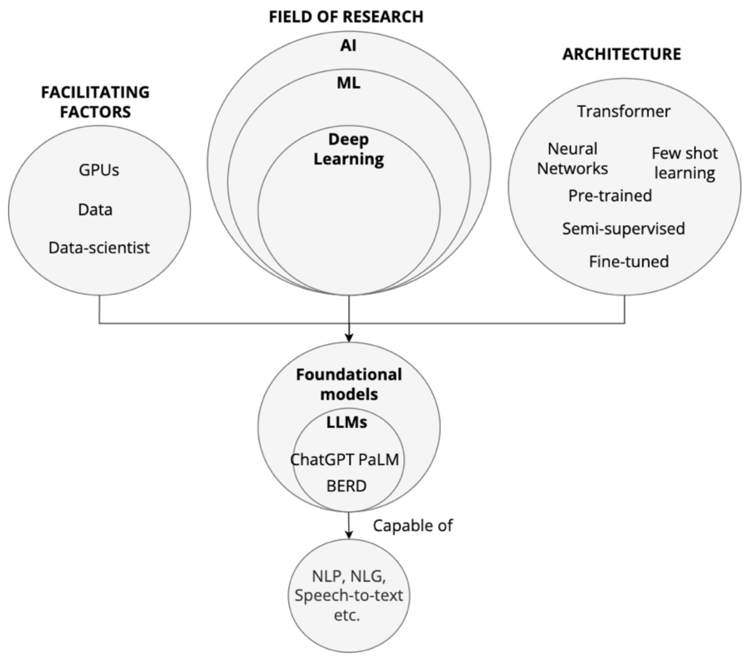


Figure 1*:* simplified representation of important concepts around LLMs in the research context

2.3 LARGE LANGUAGE MODELS IN HEALTHCARE

The most recent versions of ChatGPT and Med-PaLM2 achieved passing-level performance on the United States Medical Licensing Examinations, the standard test to evaluate prospective doctors in the United states. This sparked suggestions that LLM applications might be ready for deployment in clinical, educational and research settings within healthcare (Kung et al., 2023). A review of ChatGPT, revealed specific capabilities and recurring issues. Promising capabilities included streamlining the workflow, cost-saving, documentation, personalized medicine and improved health literacy. Commonly mentioned issues included ethical issues, copyright infringement, lack of transparency, legal challenges, bias, plagiarism and inaccuracy (Rathenau Instituut, 2023; Sallam, 2023; Thirunavukarasu et al., 2023) Significant excitement surrounds the potential applications, to the extent that some discussions suggest a paradigm shift or drawn parallels with the introduction of the web browser (Ali et al., 2023; Javaid et al., 2023; Moor et al., 2023; Patel & Lam, 2023). LLMs can have an impact on different aspects of the healthcare system. Patient facing applications include virtual assistance, responses to questions and help in education. Furthermore, its applications for the medical professional range from Electronic Health Record (EHS) summarization to clinical decision systems setting appointments and medication reminders (Javaid et al., 2023). The best performance of ChatGPT was attributed to its ability to summarize information, using technical language for communication among clinics as well as plain language for communication with patients and their families (Cascella et al., 2023). While these functionalities of LLMs show promise in healthcare, currently more research and development is required to fully realize this potential. General-purpose LLMs like ChatGPT are better tailored for casual chatbot conversations rather than healthcare contexts. This leads to a discrepancy between intended and actual use (Kung et al., 2023). Because of computing costs and performance quality, for clinical applications, domain specific LLMs are seen as an logical next step (Thirunavukarasu et al., 2023). However, given the challenges associated with LLMs around noisy data and the inherent risks of text generation errors, such as confabulation or hallucination, these factors pose extra risks in a sector dealing with human health. Wornow et al. (2023) examined the training data for 84 different LLMs in healthcare and concluded that a large number of the more recent and more domain specific LLMs are trained on just two publicly available data sets and publicly available articles on Pubmed, an academic literature database. This data contains a relatively limited 2 million clinical notes, 16 million abstracts and 5 million full-text publications. In this regard, a large part of high-quality data that is protected by paywalls, forms one of the barriers. The models using a greater percentage of its training data derived from electronic health records (EHS), showed better performance than those mainly trained on publicly available data-sets but at the same time expressed significant geographical and organizational dependences as codified medical knowledge is stored varyingly (Wornow et al., 2023). Again, this shows that quantity must not be mistaken for quality with the development of LLMs (Webster, 2023).

In order to mitigate the data availability problem, a recent study by (Peng et al., 2023) trained a new domain-specific medical LLMs called 'GatorTronGPT' that utilizes generated 'synthetic' data to improve its metrics and shows improved performance on clinical text generation compared to current models trained on 'real' data (Peng et al., 2023). To conclude, the use of LLMs in healthcare have recently gained the attention of the public and while in some metrics they already outperform humans, they are not without imperfections. A new avenue of research being the symbiosis of human-algorithm collaboration appears to express synergic effects, outperforming humans in and outside the LLMs research domain of AI (Park et al., 2023). In short term, general medical intelligence remains unattainable, but more locally trained, domain specific models are bound to be clinically relevant.

2.4 A BRIEF OVERVIEW OF AI IN DUTCH HEALTHCARE

Recently, the innovation landscape of AI in healthcare in the Netherlands on a systemic level has been the subject of 'grey literature'[5] reports commissioned by the Dutch Ministry of Health and its subsidiaries. Of the different types of AI applications, pattern recognition is used the most often (29%), followed by computer vision (24%) and NLP (16%). Thus, NLP being in the top 3 applications corresponding with international research on applications of AI in healthcare (Esteva et al., 2019; Wouda et al., 2019). In 2020,AI applications in the Dutch healthcare were most often used in diagnostics (35%) and intervention (16%) and less in prognosis and treatment, again being in line with the potential of the earlier defined areas of computer vision and NLP (KPMG, 2020). Within the healthcare sector, healthcare institutions, such as hospitals, are the most prevalent customer of AI applications (82%), with the most prevalent end-user being the medical specialists (63%)(Berenschot, 2022). Medical specialists often take on the role of the individual 'local champion' entrepreneur, which is mentioned as crucial for the start of the AI innovation trajectory (Berenschot, 2022; Strohm et al., 2020). The development of applications is often carried out by start-ups & scale-ups (50%), while healthcare institutions, knowledge institutions, and multinationals make up a smaller proportion (41%). In almost all applications, multiple different actors are involved, enabling successful implementation.

[5] Grey literature is defined as research published by sources outside the conventional academic sphere

There are multiple barriers to implementing AI that are mentioned for the Dutch healthcare case. For the cooperation of parties, the need for data and model sharing is both high and experienced as insufficient (Cap Gemini invent, 2020; M&I Partners, 2023). Data quantity is not seen as a problem, nor is the collection of increasing medical data. Data harmonization and making data findable and accessible is a main goal to improve sharing. The emphasis is on accessible decentralized data storage instead of an expanding centralized data source (Cap Gemini invent, 2020). Furthermore, healthcare institutions express that they are not adequately aware of the activities of different players, and sharing of running projects is minimal (KPMG, 2020). Cooperation of all the involved parties in the early phases is seen as important (KPMG, 2020). In a substantive amount of the applications, problems are experienced with the financial viability of the applications, insufficient financing, or no compliance with standing regulations. Additional factors that are hampering implementation that are mentioned are; acceptance among healthcare providers, upscaling of software and infrastructure, and the education of users (Berenschot, 2022; Cap Gemini invent, 2020; KPMG, 2020; Rivm, 2018; Wouda et al., 2019). The Dutch AI context is considered when analysing the DH-LLMTIS, a TIS that will be predominantly in an early-stage based on the technological developments described above.

3 THEORY

3.1 INNOVATION SYSTEMS & TIS

Since the notion by Schumpeter of ‘creative destruction’ as process of the ‘new’ replacing the ‘old’ in order to move forward economically, numerous schools of thought have tried to formulate theory that explains the dynamics of innovation and successful technological change (Schumpeter, 1937). From evolutionary economics, that is a critique on neoclassical economics and institutional theory, through seminal works by Freeman C, (1995), B.-Å. Lundvall (1992) and Nelson R (1993), the notion emerged that innovation and diffusion of technology is both an individual and collective act and takes part within a system.[6] This more systemic view was captured within the term ‘innovation systems’ (IS), first coined by Lundvall 1985. It can be defined as “all important economic, social, political, organizational and institutional, factors that influence the development, diffusion, and use of innovations” (Edquist & Charles Edquist, 2001; Hekkert et al., 2007; B.-A. Lundvall, 1985). It tries to explain why the rate and direction of technological change is not a simple competition between technologies but is formed by various existing and emerging innovation systems. The rigidity of these systems is often significant, leading to predetermined technological trajectories coined lock-in, that enables the study of expected system dynamics (David, 1985; Hekkert et al., 2007).

Over the years, taking different scopes as context, several strands of research concerned themselves with uncovering the dynamics of innovation systems including regional innovation systems (RIS) (Cooke et al., 1997), sectoral innovation systems (SIS) (Malerba, 2002) and the popular theory of national systems of innovation (NIS) (B.-Å. Lundvall, 1992). These approaches can be argued to complement each other rather than exclude each other (Edquist & Charles Edquist, 2001). Furthermore, the central idea of the rigidity of existing systems was also exemplified in the Multi-Level Model (MLP). Within this model, the pre-existing technology’s innovation system is designated as the “regime,” whereas the “niches” are the incubation rooms for emerging technologies or novelties. The fundamental concept posed by this model pertains to the circumstances under which a niche achieves such success that it integrates into the established regime[7] (Geels, 2002). However, to adopt a more specific framework to analyse the system around a specific technological innovation, the notion of a technological innovation system (TIS) developed parallel to the theories like the MLP. Technology being defined as the hardware, software and the technological knowledge embodied in both. First mentioned by Carlsson & Stankiewicz (1991) a technological innovation system can be defined as a set of networks, actors, institutions and physical infrastructure that jointly interact in a specific technological field and contribute to the generation, diffusion and utilization of variants of a technology, and/or product (Bergek et al., 2008; Markard & Truffer, 2008; Wieczorek et al., 2015). Compared to the previous innovation systems approaches, the number of relevant actors, networks, institutions and physical infrastructures of one technology are significantly reduced, improving the ability to study of systems specific dynamics. The article by Hekkert et al. (2007) conceptualised how to study the systems dynamics by putting forward a combined structural (static) and functional (dynamic) analysis of a TIS. The structural elements are the basis of the static analysis and consist of a characterization of the most important actors, networks, institutions and infrastructures present for the technology. It is simply not feasible to map all the dynamics that take place in a TIS. Therefore, Hekkert put forward the ‘functions’,

---

[6] Evolutionary Economics highlights the importance of innovators and rejects classic economic ideas of equilibrium. The Entrepreneur innovates in face of uncertainty creating a destructive force on current structures. The term ‘creative destruction’ coined by Schumpeter introduced extra uncertainty into economics. Evolutionary economics explicitly considers technological and social factors

[7] The MLP is said to give attention to competing technologies in their struggle to reach the regime. This is said to be under conceptualised in TIS. However, TIS approach is a fitting choice in technological innovation where the competitions are not primarily between different technologies or between the incumbent technology and the emerging one.

which are the most important activities that take place, as heuristic framework to analyse the TIS performance. These functions were distilled from earlier research and refined and updated over time and in different contexts, keeping the general structure the same.

The original seven that form the basis of a functional analysis are described as follows[8]: *F1 entrepreneurial activities*, the main source of uncertainty reduction in experimentation and essential for the functioning of a TIS. Entrepreneurs can be either new entrants or incumbents diversifying into new markets. *F2 knowledge development*, that is the creation of new knowledge, is at the heart of the innovative process. Taking the form of for example R&D and learning mechanisms. F2 has long been recognized as fundamental in the modern perspective of innovation. *F3 knowledge diffusion* through networks, being the exchange of information that is required to capitalize on new knowledge. However, functions 2 and 3 are often regarded as comparable and therefore aggregated (Apell & Eriksson, 2023; Bergek et al., 2008; Markard, 2020; Rikap & Lundvall, 2021). *F4 guidance of the search*, stemming from the fact that resources are inherently limited, is concerned with the way actors and institutions envision the trajectory of the technology. Active guidance of the search can steer targets towards one technology and away from another. *F5 market formation*, with a new technology having difficulty to compete with existing systems, artificial competitive advantage often needs to be created in order to make a new technology viable. Market formation can say something about the development stage the TIS is in. *F6 resource mobilization*, they can be present in the form of financial or human resources and form the 'fuel' for all the other functions. *F7 creation of legitimacy,* the legitimacy of new technology in the end determines its capabilities for 'creative destruction', for all technologies must deal with incumbent socio-technical systems (Bergek, 2019; Bergek et al., 2008; Hekkert et al., 2007).

The strength of this analytical framework provided by the functional approach to TIS is that it reduces the complexity of a wide array of technology dynamics to explainable dynamics (Bening et al., 2015). All different functions need to be accounted for in order for a technology to thrive. This has led to the detection of reoccurring functional performance patterns that can either improve or impede the formation of a TIS (Hekkert et al., 2007). Termed 'cumulative causation' it is the effect one changing factor has in a system of factors accelerating or decelerating the change process (Suurs & Hekkert, 2009). TIS framework does not focus on the potential power struggles between different technological systems, like other transition theories, such as the MLP. However, the function 'legitimacy' is understood as the alignment with current legislation, public opinion and the active defending of the technology by structural elements (Bergek et al., 2008; Hekkert et al., 2007; Markard et al., 2015; Markard & Truffer, 2008). After the technological review of the use of LLMs in a healthcare setting, it was found that the potential hampering factors or dynamics to its implementation in Dutch healthcare are not due to the power struggle between competing technologies.

After the initial formulation of the TIS framework, later research put the importance of contextual factors on a TIS forward in different ways. It being either the technologically or geographically position of a TIS within the context of other systems or TISs that have to be considered (Bergek, 2019; Bergek et al., 2015; König et al., 2018; Musiolik et al., 2012).

Previous research on a AI specific TIS by Rikap and Lundvall (2021) focused on global AI-TIS, led by by the US and China, and their major technology companies. However, such a global perspective is insufficient for understanding the specific sectoral and geographical dynamics of LLM applications in the Dutch context. For example, this research examines the use of AI associated terms in research published by 5 big American Tech companies. This analysis will therefore not be helpful to identify the most important players in the Dutch healthcare context.

[8] The numbers of this set of functions do not correspond with the numbers applied in this research

The only AI-TIS analysis in the healthcare context, performed by Apell & Eriksson (2023), while usefull for its methodological approach and insights in functional dynamics, presents shortcommings for the case delineation and context of this research. For example, the geographical delineation for West-Sweden as focus is not well explained. Additionaly, it does not mention the insights on the characterisation of various TIS phases and their functional dynamics as expressed in a significant body of research (Alkemade & Suurs, 2012; Bento & Wilson, 2016; Hanson, 2018; König et al., 2018; Markard, 2020; Markard et al., 2020).

3.2 AN INTEGRATED FRAMEWORK

The current use of LLMs in healthcare in the Netherlands is as of now limited as established in the technological context. However, give the potential that is attributed to LLMs in various areas of healthcare, it will be highly relevant to study what the characteristics of the Dutch Innovation System are for the implementation of LLMs in healthcare, as they will determine future successes. In this research, the theory on TISs was adapted to characterize these specific system dynamics.

Extension of the TIS framework

In practice, TIS analysis is predominantly applied to the formation of innovation systems in sustainable innovation, often focusing on sectors such as sustainable energy or electric vehicle implementation. This usage tends to blur the distinction between promoting sustainable technology and technological innovation at large (Bening et al., 2015; Bergek et al., 2015; Hekkert & Negro, 2009; Planko et al., 2017; Rikap & Lundvall, 2021; Suurs & Hekkert, 2009; Wieczorek et al., 2013, 2015). Although not initially designed with sustainability as its focal point, the TIS framework is recognized for its applicability across various sectors, including healthcare (Apell & Eriksson, 2023; Bergek, 2019). The healthcare sector, with its unique systemic requirements for successful digital implementation (Dorn, 2015), offers a relevant area to extend the TIS framework's usage, taking into account the sector-specific structural and functional differences compared to sustainability-oriented technologies. This approach was guided by the methodologies and findings of previous TIS analyses.

Structural and functional analytical steps

The structural elements in this research are defined, taking a socio-technical scope (including both the social and technical elements of the structure), to capture the distinct social elements of the implementation of a technology in a healthcare setting (Moullin et al., 2015). These definitions are adapted from work by Andersson et al. (2023), Bergek et al. (2008), Wieczorek et al. (2015) and Apell & Eriksson (2023) as follows:

- Actors: Actors, organizations or individuals contributing to the emerging technology in focus as a developer, adopter, regulator, financer, etc.
- Institutions: 'the rules of the game' such as laws, regulations, norms, search heuristics, etc.
- Networks: the configurations of actors and institutions in formal and informal ways such as joint projects, publications, partnerships and supply chain agreements.
- Technological structure: A notable addition to these three categories is the 'technological structure' which encompasses the upstream and downstream artefacts that accompany the technology in focus (Bergek et al., 2008; Suurs et al., 2010). In LLMs infrastructural artefacts are the computational hardware that is used to run the models or other cloud computing based solutions used.

In order to perform a functional analysis, the functions of multiple relevant TIS analyses were synthesized (Apell & Eriksson, 2023; Bergek et al., 2008; Hekkert et al., 2007; Jacobsson & Bergek, 2004; König et al., 2018; Markard, 2020; Rikap & Lundvall, 2021; Suurs et al., 2010; Wieczorek et al., 2015). **Table 1** illustrates how similar functions have been consistently identified in relevant literature on Technological Innovation Systems (TISs), with a focus on early-stage TISs and TISs in the field of AI in life sciences. It should be noted, as will be discussed later, that the precise delineation of these functions is not rigid but provides a comprehensive overview. The evolving dynamics that emerge from the data can alter the emphasis placed on specific functions. The final column of the table delineates the set of functions adopted as the analytical framework for this research.

The 7 functions used in this research were further defined and data was collected that is relevant and feasible for the TIS in focus. The resulting functions are presented in **Table 2**. This table includes the final 7 functions, the description used, the indicators used to measure them in this research and the expected dynamics in early-stage TIS phases that follow from literature. **Chapter 3** methodology further specifies how the data on the different functions was collected using a triangulation approach of combined desk research and semi-structured interviews.

Bergek 2003	Hekkert 2007	Bergek 2008	Suurs 2010	Wieczorek 2015	Köning 2018	Markard 2020	Lundvall 2021	Apell 2023	Dutch LLMTIS
F2 Knowledgde development & difussion	F1 Entrepreneurial activities	F2 Knowledgde development & difussion	F1 Entrepreneurial activities	F1 Entrepreneurial activities	F2 Knowledge difussion	F2 Knowledgde development & difussion	F2 Knowledgde development & difussion	F2 Knowledgde development & difussion	F1 Entrepreneurial activities
F3 Guidance of the direction of search	F2 Knowledge development	F3 Guidance of the direction of search	F2 Knowledge development	F2 Knowledge development	F6 Legitimicay creation	F3 Guidance of the direction of search	F3 Guidance of the direction of search	F6 Legitimicay creation	F2 Knowledgde development & difussion
F4 Market formation	F2 Knowledge difussion	F1 Entrepreneurial activities	F2 Knowledge difussion	F2 Knowledge difussion	F5 Resource mobilization	F1 Entrepreneurial activities	F1 Entrepreneurial activities	F5 Resource mobilization	F3 Guidance of the direction of search
F5 Resource mobilization	F3 Guidance of the direction of search	F4 Market formation	F3 Guidance of the direction of search	F3 Guidance of the direction of search	F3 Guidance of the direction of search	F4 Market formation	F4 Market expansion	F3 Guidance of the direction of search	F4 Market formation
F7 development of positive externalities	F4 Market formation	F6 Legitimicay creation	F4 Market formation	F4 Market formation	F4 Market formation	F6 Legitimicay creation	F6 Legitimicay creation	F1 Entrepreneurial activities	F5 Resource mobilization
	F5 Resource mobilization	F5 Resource mobilization	F5 Resource mobilization	F5 Resource mobilization	F7 development of positive externalities	F5 Resource mobilization	F5 Resource mobilization	F4 Market formation	F6 Legitimicay creation
	F6 Legitimicay creation	F7 development of positive externalities	FX Support from avocacy groups	F6 Legitimicay creation	F1 Entrepreneurial activities	F7 development of positive externalities		F7 development of positive externalities	F7 development of positive externalities

Table 1: The use of functions in earlier research

Characteristics of Early-stage TISs in literature

Significant literature addresses the categorization of TIS into various developmental stages, each characterized by distinct dynamics and patterns (Alkemade & Suurs, 2012; Bento & Wilson, 2016; Hanson, 2018; König et al., 2018; Markard, 2020; Markard et al., 2020). The earliest stage is described as the 'Nascent' stage followed by a 'formative' stage and then moving towards a 'growing' stage on its way to 'maturing', 'strengthening' (Alkemade & Suurs, 2012; Bento & Wilson, 2016; Bergek et al., 2008; Markard, 2020) or even declining (Markard et al., 2020).

This research does not aim to label these stages but rather to contextualize the focal TIS of LLMs within these established stages. As Bergek et al. (2008) mentioned, it would be unwise to judge the development of a TIS in a formative phase based on indicators of a further developed TIS. For instance, hoping to find strong synergies in a newly discovered technology would unjustly underestimate its potential for creative destruction. It can be safely said that a TIS of LLMs is still in an early phase, given the novelty of much of its technological breakthroughs and applications. Therefore, the TIS of LLMs in the Netherlands will be limited in specific structural and functional elements, it is important to note where and how this incomplete development could be improved upon in its current state.

Function	Description	Indicators	Examples of expected dynamics
F1. Entrepreneurial Activities	The presence and activity of entrepreneurs and the number of new applications	Number of different types of applications, number of start-ups and companies, assessabliity of entrepreneurship by experts	Low amount of entrepreneurs, new entrepreneurs moving in, dependent on entrepreneurs of developed TIS (Hekkert et. al 2007) First commercial deployment (Bento 2016) Actors provded by others TISs (Markard 2020)
F2. Knowledge Development & Diffusion	The scientific, technological and market knowledgbe base formation and knowledge exchange	Number of scientific publications, funded research projects, eduacational programs, assessements of function by experts	No patents (Markard 2020,) Knowledge aquired fast (no source) Large emphasis on R&D (Bergek 2008, Markard 2020) Product improvement not process improvement (Bento 2016) Convergence of networks, standards and regulation (Suurs 2012, Bergek 2008) Nascent phase, exclusively R&D Networks (Bento 2016)
F3. Guidance of the Search	The directionality given by actors and institutions through policy, regulations and their expectations on the subject and industry standards	Policy and strategic explicification promoting the technology. Regulations. End-users expectations and needs for the technology, assessements of function by experts	Importance of direction guidance through regulations (Suurs 2012) Alignment of regulations takes shape (Bergek 2008) Convergence of F2 strengenting F3 (Suurs 2009) Guidance through policy dependent on the contextual TIS (Markard 2020) Structuration becomes available, uncertainty exists in standards (Markard 2020) Decreasing number of institutions (Musiolik 2018) (Hekkert 2007)
F4. Market Formation	The market size, customers/ end-users, institutional stimulae for formation	Availability of innovations on the market. Artificial creation of niche markets , assessements of function by experts	A fraction of its potential market (Bergek 2008, Markard 2020) ten percent of its expected market potential reached (Bento 2016) Few and experimental customers (Bergek 2008) Formation of articial market opportunities (Hanson 2018) Fulfillment of other functions as driver for Market formation (Bento 2016, Suurs 2009) (Hekkert 2007)
F5. Resource Mobilisation	The availability of financial resources, human resources, infrastructure and the development hereof	Accessibility of data, financial and human capital. assessements of function by experts	Dependency on contextual TIS (Markard 2020)Clear start of mobilization of human an financial resources (Bento 2016) Financial government stimulation primarily in R&D (Markard 2020) Clear mobilization of physical resources markes the ending of the early-phases (Bento 2016, Markard 2020) Mobilization in all directions (Hekkert 2007)
F6. Legitimization	The social acceptance and compliance of the technology by actors	Sentiment analysis in media-outlets, assessements of function by experts	Alignment with the contextual TIS and societal problems (Markard 2020) More expert legtimacy compared to familiarization (Bergek 2008) Increase in alignment with industry values (Bento 2016) Competition between legitimacy of contextual TISs (Bergek 2019)
F7. Synergies & positive externalities	System specific intermediary goods are developing, pooled rescources, functional interactions, usable standards	Presence or establishment of standards and formal networks, assessements of function by experts	Synergies exist in the paved way of other TISs, the TIS itself is to emergent to express syngergies. Formation of intermediary services (Bergek et. al 2004, Hekkert et. all 2007) (Markard 2020)

Table 2*:* Classification and operationalization of the functional assessment

A relationship not uncommon for an early-state TIS is its dependence on the overlap with elements from existing more mature technologies that already developed a TIS, as showed in the analysis of the emerging sustainable energy system in Norway or the early-state TIS of ‘aquaponics; (Bergek et al., 2015; Hanson, 2018; König et al., 2018). In this research the more mature technology will be ‘AI’ in the broad sense of the technological term, as it has been around in healthcare for an more extended period of time and its development can be of a supportive role to the development of an DH-LLMTIS. Moreover, the fast worldwide developments in the LLM field described in the technological context are of influence on LLMs in the Dutch context as well.

Hanson showed that in this early-stage situation, firms engaging in exploration, new actors entering the field and the utilization of pre-existing networks are structural help in a one-sided relation termed ‘flow’. However, the temporal analysis done by Hanson suggests that these dependencies can decrease over time when a TIS moves from a formative/emerging stage towards a growing stages. Given examples like market developments, entrepreneurial activity and legitimacy building that can take place separately from the established TISs (Hanson, 2018). This context dependency and its importance is stressed by in terms of technological, sectorial, geographical and political factors by Bergek et al. (2015). They can be accounted for in an analysis as the interaction between a TIS structure and its associated TISs and is referred to as ‘structural coupling’ by Bergek et al. (2015) that is comparative to ‘flow’.

Furthermore, research concludes that a TIS in a formative stage is characterised by the fluidity of the emerging technology in its performance and direction and by a weakness, or even absence, of technological and institutional structures to support it. Additionally, it speaks of more actors being drawn in, networks being formed and institutions that are designed to make the technology fit better to its surrounding structures (Suurs 2010, Bergek 2004). Bento & Wilson (2016) posed specific indicators looking at multiple case studies, marking the beginning and ending of each ‘stage’ within the TIS, ranging from nascent, emerging, strengthened and mature. Important was the first commercial application he emphasized as starting point for a formative TIS. Additionally, Markard (2020) established different stages for TIS development combining TIS literature with industry life cycle literature. Here, the structural elements in the formative stage were characterized by an actors base set close to zero, little structure and a high degree of uncertainty. To add to that, the performance parameters of the technology in this situation are unclear and the TIS is dependent on the case context to adapt to it. In terms of the functionalities present in an early-stage TIS the importance of expectations in the guidance of the search function are especially stressed in work by Alkemade & Suurs (2012). This guidance can be expressed in policy measures, standards and regulation formation and the expectations regarding technological development and performance. They point to the dynamics of expectations that often take on the form of ‘hype-cycle characteristics’[9] (being sharp increases of expectations followed by sharp decreases).

Proposed framework TIS-TIS interactions

Considering the theoretical contributions in the field of emerging TISs, it can be stated that the literature is not completely uniform. Therefore, it is of added value to test the validity of multiple observations made in earlier research on early-state TISs to formulate more precisely what can be validated of TISs in early stages. Accordingly, **Figure 2** shows a proposed tentative visualization of how the early-state TIS can be visualized as positioned between the named other TISs, as a schematic adaptation on the position of an early-state TIS between other TISs as visualized by König et al. (2018). In the Figure, it is visible how the potential early-state DH-LLMTIS is situated as subdivision of the TIS of AI in healthcare in the Netherlands (DH-AITIS), LLMs being a part of AI technology. Furthermore, the global developments of LLMs (G-LLMTIS) influences parts of the Dutch system presented in the small overlap. As is known, the global development of LLMs is dominated by both American and Chinese research and companies in terms of publication, number of companies and funding. American companies have more overlap with the European than the Chinese and are therefore most important in the assessment of the G-LLMTIS (Rikap & Lundvall, 2021). In this proposed framework it is not fully clear in which areas this overlap exists in what way. This visualization explores the expectation that a picture emerges of where the Dutch LLMTIS falls short and is more dependent on the presence of the AI-TIS or Global LLMTIS. For example, a global research community can influence knowledge development and increase or decrease legitimacy of a technology. In contrast, market formation and guidance of the search in the form of formalized policy will show stronger local dependencies (Rikap &

[9] Gartner’s hype cycle is a popular framework to characterize the expectations of emerging technology (Linden 2003)

Lundvall, 2021). An adjusted visualization specifying the place and size of the overlap will aid in answering the delineation of research sub-question 1.

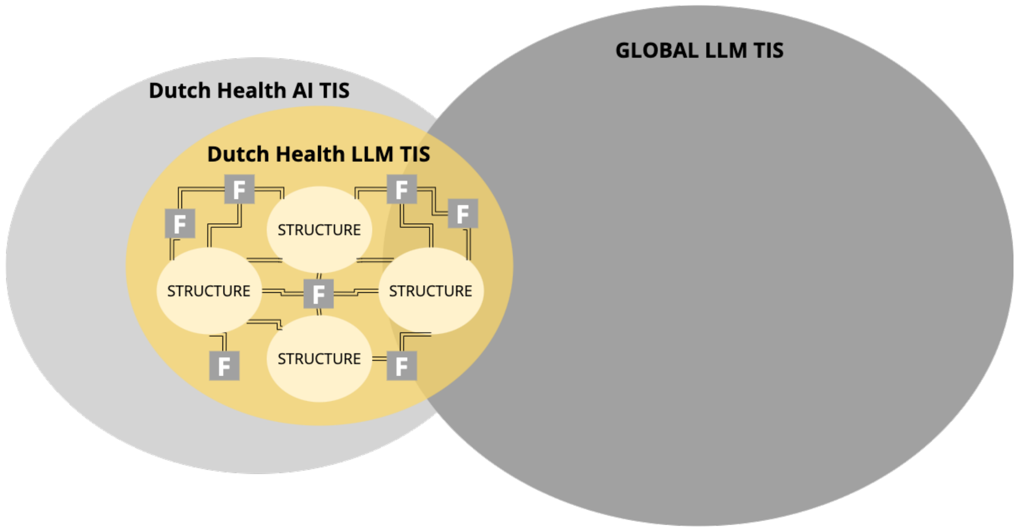


Figure 2*:* The Dutch Health LLMTIS positioned between the Dutch AI TIS and the global LLMTIS

Cumulative Causations

The set-up and development of different functions form standard patterns that are described in literature, taking on the form of cumulative causations or 'motors of innovation'. Suurs (2010) distinguished 4 different 'motors' of which the 'Science-Technology Push motor' and 'Entrepreneurial motor' were characteristic for the earlier stages of TIS development (Suurs & Hekkert, 2009). The Science-technology push motor of which a positive cycle starts with Guidance of the Search (F3) causing Resource Mobilization (F5) leading to Knowledge Development and Diffusion (F2), which lead to Guidance of the Search (F3) completing the cycle. Market Formation (F4) is absent in this motor. The Entrepreneurial Motor is the follow-up on the Science technology push and is characterized by a similar dynamic of function but involving significant entrepreneurial activity. This motor indicates the change from a more 'formative' phase to a 'growth' phase.[10] The findings on these standard patterns are compared to the case and add to the understanding of the functional performance and possible gaps of the focal TIS.

[10] The exact set of functions in the article by Suurs slightly differ from the final set of functions in this research, showing a strong overlap.

4 METHODOLOGY

4.1 RESEARCH APPROACH

In this research, a combined deductive and inductive approach was used in an explorative case study. In the first step, the existing theory of TIS analyses was taken as starting point and adapted to better fit the case of LLMs in healthcare. This resulted in an adapted structural and functional framework that directed the further gathering of data in a deductive way. It was designed on the notion that the performance of each function of the TIS can be determined based on the set of quantitative and qualitative assessments that collectively led to the performance of the entire TIS (Bergek et al., 2008; Hekkert et al., 2007). Furthermore, using this approach, it was assessed if the case expressed the characteristics that followed from the literature on Early-stage TISs. Moreover, new emerging dynamics or concepts leading to theoretic extension or alteration can inductively follow from the observations in the data.

Considering the theoretical aim of the research, a case study design was appropriate. Case studies are a suitable strategy when the research aims to deepen the understanding of a phenomenon in a real-life contextual setting and are often used in TIS analyses (Bergek et al., 2008, 2015; Bryman, 2012; Wieczorek et al., 2015) and allow for more than one unit of analysis as is needed in a TIS analysis. The data was gathered in mixed-method way and deemed suitable since the triangulation of multiple quantitative and qualitative data sources is expected to provide an in-depth analysis of the structural and functional elements of the TIS (Bergek et al., 2008; Corbin & Strauss, 2008; Markard, 2020; Zolfagharian et al., 2019). Data triangulation was achieved through the following data collection methods: (1) qualitative desk research for case exploration, case context and respondent search (2) qualitative and quantitative desk research combined with semi-structured interviews for a structural analysis (3) qualitative and quantitative desk research combined with a second round of semi-structured interviews for a functional analysis. Although the different steps of the TIS analysis are presented chronologically, they involved iterations between the steps to increase the relevance of the findings (Bergek et al., 2008).

4.2 TIS ANALYSIS STEPS

Figure 4 presents the steps of the TIS analysis that were performed as adapted from Bergek et al. (2008). 'I Delineation and context' established the technology in focus (focal TIS) and defined the context of the case through a qualitative document analysis and first informal talks with industry experts. 'II Structural analysis' was desk research, combined with a round of semi-structured interviews to determine the structure of the DH-LLMTIS. 'III Functional analysis' in this stage, another round of semi-structured interviews and desk research was performed to complete the functional analysis as adapted for this case. Finally, step IV and V present the synthesis of the analyses to pinpoint how the focal TIS performs and how the findings compare to previous literature. The next paragraphs specify the data collection and execution of the different steps of the methodology presented in **Figure 4**. Notable iterations are: the refinement of the delineation and context of step I during and after the interviews for the structural analysis, the refinement of the structural analysis during and after the desk research and interviews of the functional analysis, visible as the backward arrows in **Figure 3**.

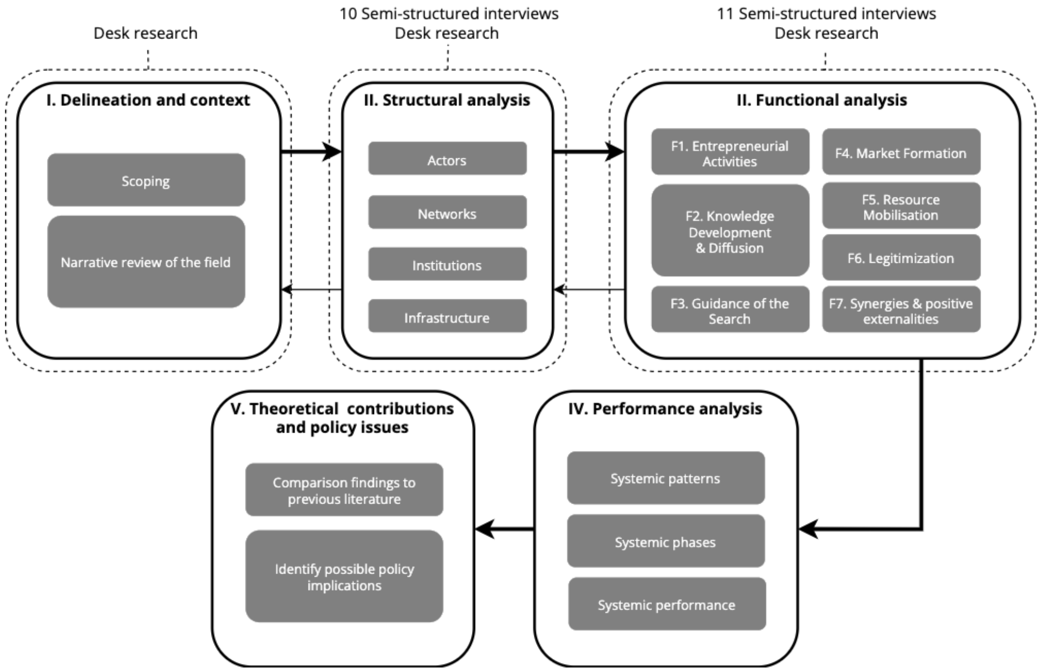


Figure 3: research steps of TIS analysis

4.3 DATA-COLLECTION

4.3.1 I CASE DELINEATION

During the initial phase of the case delineation and desk research, multiple choices are made in order to scope down to the focal TIS in terms of the technological field, the breadth and depth (being the range of studied applications and the sector of usage) of the focal technology and spatial delineation (Bergek, 2019; Bergek et al., 2008). This had as goal to answer sub question 1:

SQ1: What is de delineation and context of the Dutch Large Language Model Technological Innovation System in healthcare?

This was done using the following definition of technology, elaborated from the original definition of Carlsson & Stankiewicz (1991) as starting point: 'technology refers to material and immaterial objects – both hardware (e.g. products, tools and machines) and software (e.g. procedures/processes and digital protocols) – that can be used to solve real-world technical problems'(Andersson et al., 2023; Bergek et al., 2008). The real world technical problems are in this case defined as the pressure on a costly and labor-intensive healthcare system as stated in the introduction. The Dutch healthcare system is more expensive than average in the European Union, spending more than 11% of their GDP on healthcare (Eurostat, 2023). Rising costs have been structurally on the agenda of the Dutch government (Rivm, 2018).

For the document analysis, grey literature, press articles, policy/regulatory documents and academic literature were considered to provide context and aid the scoping process of the technology in focus. An initial search into the academic literature in the field of AI in healthcare led to the classification of specific subtypes of AI requiring their own technological development in the form of a TIS. Hereafter, the subject searched for was LLM technology in healthcare in The Netherlands. Additionally, official documents and strategy documents from government agencies and health consulting agencies provided insights into the general developments in the technological field. For scientific literature, articles that covered implementation or adoption of LLM applications in healthcare were sampled for. When LLM specific topics did not yield the necessary documents, documents regarding AI entire knowledge field were chosen and scanned for LLM specific content. Parallel to the desk research initial informal talks with experts in the field of AI and with an association to the healthcare sector were held with members of the 'intelligent health' team of Capgemini Invent and researchers at Utrecht University. This confirmed the scoping of AI as knowledge field to LLMs and their NLP applications. Further desk research into the applications and developments of LLMs and LLMs in healthcare provided the necessary technological background into the models and made the tentative positioning of the DH-AITIS developments and G-LLMTIS for the case in focus. The spatial scope was set on country-level, being the Netherlands, as healthcare systems express strong national characteristics (Kroneman et al., 2016). Chapter 2 provides the results of the scoping and case context that was done in step I of the research process. The final development of the integrated framework in the discussion chapter adds to this delineation in light of TIS-TIS interactions.

4.3.2 II STRUCTURAL ANALYSIS

This section was aimed at answering *Sub question 2:*

SQ2: What are the relevant structural elements in the Dutch Large Language Model Technological Innovation System regarding actors, networks, institutions and infrastructure?

The process of identifying the structural elements started out with desk research consisting of the same type of document sources as with the context delineation phase, dividing content into the four structural elements following the socio-technical scoping approach from Andersson et al. (2023). The goal of this search was to find data that directs to organizations and institutions with the highest volume and direction of technological activity and knowledge creation. This provided a starting point for the purposive sampling of respondents for a round of semi-structured interviews with industry experts. These interviews, served as validation and triangulation method (Bergek et al., 2008; Bryman, 2012). A general classification of the structural elements of a national healthcare TIS into three spheres was adopted from Larisch (2016) (governance, market and academic), guiding the process of data gathering and analysis to cover a full range of backgrounds. An overview of the guiding categories for the structural approach is given with three example actors per sphere in **Figure 4**. Actors types spoken to for the structural analysis are presented in bold.

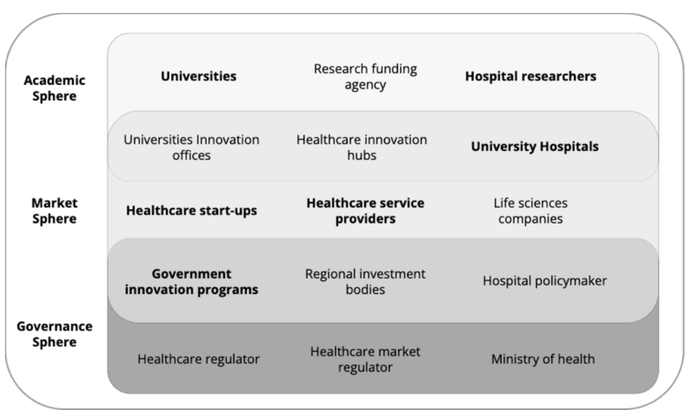


Figure 4: guiding categories for structural respondent analysis

Furthermore, although patent analysis can actively identify structural elements in a TIS framework, as suggested by Wieczorek et al. (2013) the nascent stage of LLM applications in healthcare rendered such analysis impossible. This determination emerged following initial consultations with industry experts, who highlighted the limited utility of patent analysis in this context due to the technology's novelty.

It was confirmed that in this field patents are lagging behind on recent developments and that no clearly defined patent codes can be used for a search. A general descriptive bibliometric analysis on the subject can also point to the knowledge institutions that are active as actor in the structure of the TIS (Bauer et al., 2017; Bergek et al., 2008). Therefore, a general descriptive bibliometric analysis to search for the main publishing institutions of literature on the topic of LLMs in healthcare in the Netherlands was performed. The databases of Web of Science, Pubmed and Scopus were searched. The search query was used to search for publications on LLMs and NLP in a healthcare setting in the geographical area of the Netherlands. Healthcare keywords were excluded in the PubMed search as it already only contains medical literature. Articles were included that consisted of at least one of the authors that was affiliated to a Dutch affiliation. To test the validity of this approach, the title, abstract and affiliations of 25 articles were checked manually per query to assure relevance. A cut-off date of 2017 was used, as it was the year the seminal article *Attention is all you need* was published, marking the acceleration of development towards the current LLMs. The complete search queries of the literature databases can be found in **Appendix B.**

Round 1 of explorative semi-structured interviews with experts

The purposive sampling of respondents was based on an iterative process of found structural elements in literature, the used guiding categories, and the snowballing of new respondents after initial contact, a response-rate of 91% was reached. The semi-structured interviews helped scrutinize the findings of the desk research and helped consolidate the structural analysis as performed in earlier TIS analyses (Apell & Eriksson, 2023; Hekkert & Negro, 2009; Suurs et al., 2010). Based on the theoretical considerations and the answering of sub question 1, an interview guide was formed (**Appendix A**). Questions were aimed at what respondents defined as the most important structural elements of the DH-LLMTIS applications. The answers also touched upon the structural elements that are in place overlapping with the DH-AITIS and G-LLMTIS without explicit steering away from the focal TIS. The interviews were held in either Dutch or English and all performed via Videocall. A larger part of the interviews were held in Dutch. The actors are abbreviated by their actor sphere of their official roles either being Market, Academic, Governance or multiple (A, M, G). Furthermore, if an observation was shared by all respondents when asked it is referred to Respondents Round 1 as (R1). An overview of the respondents is given in **Table 3**

Sphere	Function	Date and duration
Market	**M1** Consultant GenAI	18/9/23 (35 min)
Academic	**A1** Assistent Prof. Ai in Healthcare	21/9/23 (38 min)
Market	**M2** Consultant GenAI	25/09/23 (75 min)
Academic	**A2** PhD Ai in Healthcare	26/09/23 (44 min)
Academic	**A3** Prof LLMs in healthcare	27/09/23 (33 min)
Academic	**A4** PhD LLMS in healthcare	28/09/23 (62 min)
Market	**M3** Ceo Start-up LLM in healthcare	28/09/23 (85 min)
Academic/Market	**AM1** Researcher hospital LLMs in healthcare	10/10/23 (44 min)
Academic/Market	**M4** Ceo Start-up LLM in healthcare and investor	16/10/23 (38 min)
Academic/Governance	**AG 1** Prof and government innovation program	17/10/23 (78 min)
		Total: **10**

Table 3: overview of semi-structured interviews round 1

Respondents were always asked for consent on recording and are made aware of the content and goals of the research. In total 10 interviews were held with the aim of reaching data saturation. Respondent validation was performed, whereby the respondents are asked to review and validate their answers to the questions during and after the official part of the interview was finished. It ensured that there was good correspondence between the usage of the interview material by the researcher and the experiences and perspectives of the respondents. After the interviews, they were transcribed verbatim to avoid the loss of data. For anonymous data handling, interviewees are referred to as their actor sphere and a number, for example: 'A1'. The respondents were directly contacted by the researcher and asked to provide extra respondents through the snowballing method for both rounds of interviews (Bryman, 2012; Wieczorek et al., 2015).

4.3.3 III FUNCTIONAL ANALYSIS

This step was performed to answer *sub question 3:*

SQ3: What is the performance of the different Technological Innovation System functions of the Dutch Large Language Model Technological Innovation System in healthcare?

The development of each function was assessed through a combination of desk research, which is validated or elaborated upon with the findings from a second round on semi-structured interviews. Below an elaboration on the desk research of each function is given and the way findings were triangulated during the semi-structured interviews. The desk research began before the first interview of the second round of interviews and data collection continued during the first interviews.

Function 1: Entrepreneurial activity

This is measured by the presence and activity of entrepreneurs in the field in the form of companies, new start-ups and scale-ups as indicator. The amount, development over time, geographical spread, type of used technology and comparison to world-wide statistics was searched for. This was done using a web search that led to the identification of three start-up databases that were deemed suitable for the search.

Crunchbase.com, finder.techleap.nl and https://startups.eithealth.eu/ were identified as prominent databases to search for newly founded companies in and outside the Dutch healthcare setting. Searched was on healthcare related topics in the provided filters using categories such as: 'Biotechnology', 'biometrics', 'medical device', in combination with AI related keywords in the description such as: 'chatbot', 'natural language processing', 'generative AI ' and a geographical search for the Netherlands when necessary. There was no date range set as to include as many results as possible. Search results were explored manually on description to filter out companies that did not match the technology criteria of this research. After the initial search, the results of the three databases were combined and overlapping companies were filtered out. This yielded 137 results. The data on company name, founding date, city and description was included. Next, the website of all companies was searched manually to determine their current activities and to filter out the companies that make use of Large Language models or advanced NLP functionalities. This led to the identification of 7 companies using LLMs and or NLP and 1 directly providing services that enable LLMs and NLP in healthcare. A additional search on the biggest international database of the three databases, CrunchBase, provided insights in the order of magnitude of AI and LLM related companies in the US and countries in Europe comparable to the Netherlands. An overview of the used queries can be found in **Appendix C.** Respondents were asked to qualitatively assess the amount of entrepreneurial activity they encountered, its developed and if significant development were present since the cute-off date of 2017.

Function 2: Knowledge development and diffusion

A bibliometric analysis was performed as in Wieczorek et al. (2015) using the databases of web of science, Pubmed and Scopus. The first search query was used to search for publications on LLMs and NLP in a healthcare setting. Using the following keywords: large language model, generative AI, generative artificial intelligence, ChatGPT, ChatGPT, genai, natural language processing. Information was extracted on the development of publications over time, per geographical region, publishing institutions and funding agencies (of those databases were this information was available). Next, a query was made excluding 'ChatGPT' or 'ChatGPT' from the search, decreasing the search results. Furthermore, using the same terminology as with bibliometric analysis, the presence of educational programs were obtained by looking at the curriculum of all the Dutch Universities in a web search as was done in Wieczorek et al. (2013). Searched was for the existence of academic bachelor and master programs, courses and tracks on the subject of AI in healthcare or LLMs/NLP in healthcare. The complete queries can be found in **Appendix C.** The findings of the desk research were validated by the respondents

Function 3: Guidance of the search

The desk research for this function consisted of a document analysis of policy and strategy documents from the major actors identified in the structural analysis, with 2017 serving as the cut-off date. This approach focused on the most recent documents available on the topic of LLMs in healthcare. Given that policy documents are not published as frequently as academic literature, incorporating some temporal aspect was necessary.

The most recent annual reports from all Dutch Universities (from 2022), UMCs, and a selection of TKZs were examined for any formalized strategy or policy mentioning AI or LLMs/NLP. The keywords 'AI', 'Artificiële intelligentie', 'kunstmatige intelligentie', and 'artificial intelligence' were used for this purpose. Additionally, policy documents from government agencies were included in the analysis.

These documents were scrutinized for evidence of governmental commitment and expressed expectations, the presence and quality of regulatory regimes, and the policy instruments emerging from them. Since policy documents often address broader trends in innovation, a search was also conducted for general references to AI. It was particularly noted whether LLMs were mentioned separately within these documents. Interviews with respondents provided insight into their expectations and those of their professional environment regarding the focal technology. They were also asked about both the informal and formalized strategic steering toward the focal technology of which they were aware.

Moreover, the quarterly conference of the Dutch Healthcare Law Society was attended.[11] During this event, leading experts on healthcare law in the Netherlands gathered and featured a series of talks specifically about data and the current status of AI regulation in the healthcare sector. Notes from these talks were used to assess the position and perspectives of these key actors on the subject.

Function 4: Market formation

Due to the novelty of the TIS, it was impossible to study an existing market of LLMs that are in place. A document analysis was performed, looking if any artificial niche market is created to support LLMs in healthcare by means of regulations as well as grey literature on the expected market share of AI technologies in healthcare. Institutional stimulae are characteristic for an emerging market not able to survive on its own. Furthermore, the amount of commercially deployed applications now and in progress were assessed by the respondents as well as the viability of the landscape for technological application of LLMs in healthcare.

Function 5: Resource mobilization

Resource mobilization was categorized into financial, human, and infrastructural resources, as outlined by Bergek (2019). For the focal technology, infrastructural resources primarily encompassed computing power and data. The documents analysed included references to financial resources allocated to AI. Additionally, policy documents addressing the mobilization of data as a resource were examined.

To assess European funding for AI in healthcare, a search was conducted in the CORDIS database, following the methodology suggested by Wieczorek et al. (2015). The search criteria were: Domain of Application - Health, Programme - Horizon 2020 or Horizon Europe, Organisation Country - Netherlands, and the keyword ‘Artificial Intelligence’. This search yielded 83 results, which were manually scanned for relevant content. The funding from all Dutch agencies was subsequently compiled to present an aggregate number.

Moreover, respondents were inquired about the extent and sufficiency of the resources in the aforementioned categories that were allocated in their professional environment to the focal technology. This approach provided both quantitative and qualitative insights into the function 4 resource mobilization for AI, particularly focusing on financial, human, and infrastructural aspects.

Function 6: Legitimization

To gain a comprehensive view of the societal perspectives on LLM applications in healthcare in the Netherlands, a search was conducted for Dutch press articles using NexisUni. The timeframe for this search was set to include articles published after November 30, 2023, a date significant due to the release of the new ChatGPT. The Boolean query employed for this search was: ('ChatGPT' OR 'taalmodel' OR 'taal model' OR 'language model' OR 'llm' OR 'generatieve AI ' OR 'gen AI ' AND 'geneeskunde' OR 'gezondheidszorg' OR 'medisch' OR 'ziekenhuis' OR 'arts' OR ‘patient’), with the search limited to articles in Language Dutch. Two separate queries were executed, one inclusive of ChatGPT and another exclusive of it. City-specific newspapers were excluded to avoid the narrow focus of smaller news outlets.

A total of 146 documents were initially reviewed by their titles and abstracts to assess their relevance to the research. Subsequently, 55 articles were selected for inclusion in the study. After further review, 30 unique

[11]More information on the conference: https://www.verenigingezondheidsrecht.nl/index.php?mact=Vergaderingen,cntnt01,detail,0&cntnt01articleid=19&cntnt01pagelimit=10&cntnt01returnid=3

articles remained for coding. The articles were analysed using Nvivo software, utilizing a mix of open and closed coding techniques. The coding categorized articles as being positive, positively critical, neutral, critical, or negative towards the use of AI and LLMs/NLP in healthcare (closed coding). Additionally, the most frequently occurring arguments both in favor of and against the use of AI and LLMs/NLP were identified, aggregated, and incorporated into the analysis (open coding).

A second search was also conducted, which excluded articles mentioning ChatGPT. This search resulted in 2 additional articles.

Function 7: System-wide synergies

The presence of system-wide synergies is limited in early-stage TISs as indicated (Bach et al., 2020). However, insights from all functional assessments of namely the AITIS were taken into account in the performance analysis and synergic effects were elaborated upon.

Round 2 of semi-structured interviews

Respondents were gathered from a combination of snowballing from the first round of interviews and a purposive search for actors as identified through the structural analysis and the referral to relevant respondents through the personal contacts of the 'Intelligent Health' team at CapGemini Invent, within the Dutch healthcare sector. Respondents were either contacted via LinkedIn or Email. This process was performed until all the spheres of actors were sufficiently covered. The positive response rate for this round of interviews was 48%. **Table 4** gives an overview of the actors interviewed in round 2.

Sphere	Function	Date and duration
Market	**M5** Business developer NLP in healthcare company	9/11/23 (45 min)
Academic/Governance	**AG 2** Prof LLMs and government grant agency boardmember	9/11/23 (42 min)
Academic/Market	**AM 2**AI Innovation officer UMC	10/11/23 (53 min)
Market/Governance	**MG 1** AI Hub chairman	14/22/23 (50 min)
Governance	**G1** Health regulator AI specialist	14/11/23 (51 min)
Academic/Market	**AM 3**Local Champion UMC and applied data science lead	16/11/23 (43 min)
Market	**M6** EHS Business developer	17/11/23 (49 min)
Academic/Market/Governance	**AMG 1**UMC AI Lab, Doctor	23/11/23 (49 min)
Academic/Market	**AM 4**UMC AI Lab, Product developer NLP in healthcare	23/11/23 (61 min)
Governance	**G2** Health regulator AI lead	27/11/23 (42 min)
Market	**M7** CEO Start-up NLPs Healthcare	27/11/23 (53 min)
		Total: **11**

Table 4: Overview of semi-structured interviews round 2

The same procedures and quality steps were performed as described in the structural analysis. In **Figure 5** it is visible which extra actors from the different spheres that were spoken to in round 2 visible in bold italic.

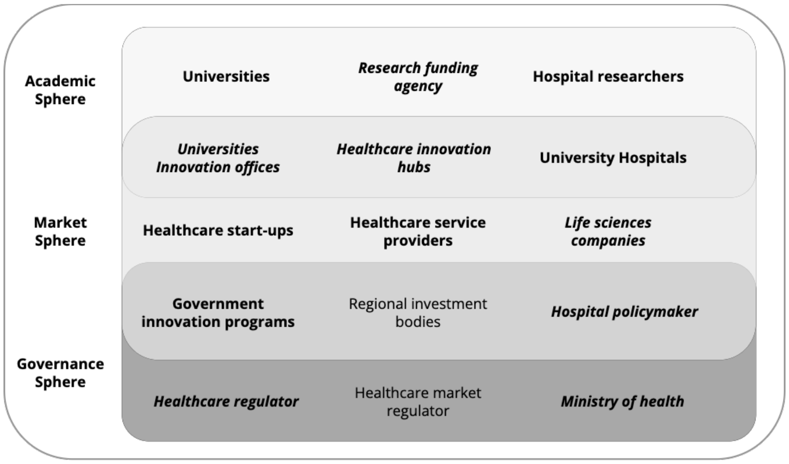


Figure 5: Overview of the spheres and spoken actors

The abbreviation for visions shared by the respondents from the second round is (R2). Based on the focus of the specific function, the interview guide was followed accordingly. This consisted of guiding diagnostic questions as presented in the interview guide in **Appendix B** and adapted from Bach et al. (2020) and Wieczorek et al. (2015). Questions were aimed at the thoughts and perceptions of valid stakeholders on the development and fulfillment of specific TIS functions. When question on the technology of LLMs did not yield a specific answer, actors steered towards AI as broad knowledge field or to the global LLM developments.

4.4 ANALYSIS

All interviews were initially transcribed using OpenAI's transcription software 'Whisper'. To ensure accuracy, the recordings were listened to afterwards, and the transcriptions were checked and corrected for any errors. Following this, the transcribed interviews were coded using the NVivo software. This coding was aligned with the TIS structures and functions as defined in the theoretical chapter and detailed in **Table 2**. Coding was conducted at the paragraph level rather than line-by-line to maintain sufficient contextual understanding.

The initial round of coding employed both closed and open coding methods. Closed coding was utilized to categorize responses into one of the predetermined structural or functional elements. In contrast, open coding allowed for the identification of concepts that enriched the case delineation, context, or structural analysis, thus enhancing the iterative nature of the analysis. A second round of more open coding was conducted, during which concepts and themes relevant to each category emerged without predefining the significance of subcategories. This was followed by a third round of axial coding, which involved synthesizing initial concepts into subcategories in line with the TIS framework and connecting them back to the theoretical structure. The goal of the final phase was to achieve theoretical saturation, meaning no additional data would significantly contribute to the existing coding (Bryman, 2012). The coding scheme and representative quotes are provided in **Appendix C**.

The structural analysis was composed based on the desk research and the coded first round of interviews. Subsequently, incorporating insights from further desk research and the second round of coded interviews, the results for each function were formulated. Where data indicated, the influence or presence of the DH-AITIS or G-LLMTIS was included and discussed. The functions were evaluated qualitatively, incorporating assessments by respondents, their dependence on existing structures, and their expected fulfillment as indicated by previous TIS literature on early-stage TISs, as summarized in **Table 2.**

4.5 RESEARCH QUALITY INDICATORS

To properly implement the research design, various quality indicators were considered. External reliability, which assesses the replicability of the study, was ensured through comprehensive documentation of all research methods and data collected. However, replicating the social context of the interviews in an exact manner is unfeasible. By extensively documenting the rationale for purposive sampling and transparently reporting the interview data, the transparency of this data collection step was enhanced.

Internal reliability, concerning the replicability of data interpretation, was a point of concern as data collection was conducted by a single researcher. Enhancing the reliability of the data involved consultations with the supervisor and member validation, allowing for the cross-verification of findings. This approach also increased internal validity, which evaluates the extent to which the importance of findings accurately reflects their significance within the case.

Assessing the external validity, or the generalizability of the findings beyond the scope of the case, is challenging to predict a priori. However, considering the globalized nature of both the healthcare and LLM fields, some degree of generalization beyond the Dutch context is anticipated. Conversely, the distinct national characteristics of healthcare systems mean that systemic issues particular to the Dutch context may not be universally applicable, posing a potential limitation to the generalizability of findings outside this Dutch case study (Bryman, 2012, part 1 ch 3).

5 RESULTS

5.1 STRUCTURAL ANALYSIS

5.1.1 ACTORS

As with all medical innovations, a large number of different actors are involved (Dorn, 2015). The general distinction between academic, market and governance actors as defined by Apell & Eriksson (2023) and Larisch et al. (2016) and confirmed for the Dutch case a government report, was kept as guiding framework to analyse the structure of actors.

The most prominent publishing affiliations of authors of scientific literature on the subject of LLMs in healthcare after the cut-off date of 2017, were identified with a descriptive bibliometric analysis of three academic literature repositories (as presented in the methodology). This yielded results presented in **Table 5.**

	Scopus		Web Of science		PubMed	
Position	**Affiliation**	**#**	**Affiliation**	**#**	**Affiliation**	**#**
1	Amsterdam University and UMC	*86*	Leiden University and UMC	*65*	Amsterdam University and UMC	*41*
2	Utrecht University and UMC	*74*	Utrecht University and UMC	*38*	Leiden University and UMC	*35*
3	Leiden University and UMC	*46*	Amsterdam University and UMC	*37*	Utrecht University and UMC	*29*
4	Groningen University and UMC	*43*	Rotterdam University and UMC	*27*	Groningen University and UMC	*21*
5	Maastricht University and UMC	*25*	Maastricht University and UMC	*17*	Maastricht University and UMC	*20*
6	Rotterdam University and UMC	*25*	Groningen University and UMC	*10*	Delft University of Technology	*15*
7	Nijmegen University and UMC	*18*	Delft University of Technology	*6*	Rotterdam University and UMC	*12*
8	Amsterdam Public Health	*16*	Nijmegen University and UMC	*2*	Nijmegen University and UMC	*10*
9	Delft University of Technology	*13*	-	-	University of Twente	*8*
10	Tilburg University	*8*	-	-	University of Eindhoven	*8*
	Topclinical	*13*	Topclinical	*5*	Topclinical	*5*
	Companies	*11*	Companies	*11*	Companies	*0*

Table 5: Dutch affiliations publishing on LLMs in healthcare between 2017 and 2023

The top 10 publishing affiliations signify the importance of the large universities as academic actors on the subject with an emphasis on universities affiliated with the 7 Dutch university medical centers (UMCs). The UMCs are both a knowledge hub and some of the biggest end-users, being the among the biggest healthcare providers in the Netherlands, therefore being both the biggest academic actors and an important market actor (A1, AG1, M3). Next to the UMCs, a group of 27 large and middle size Dutch hospitals not affiliated to a University called 'Top Klinische Ziekenhuizen' (TKZ), take a secondary role in terms of academic output. They make up a large part of the healthcare provision market in the Netherlands and can take a leading role in the implementation of technological innovations (AG1, AM1, M3, AMG1).

An important addition to these publishing affiliations is the role of American research on the subject of LLMs in healthcare which one acter mentioned as follows M1 "*In the Netherlands we are happy with 8 million for the development of new model. In the US, MIT invests 1 billion dollar into a new AI project. Some professors even avoid hot topics in LLM development because of the speed of American publishing.*" This confirms the dominance of the global, mainly American LLMTIS actors on the production of academic knowledge.

As they are the primary end-user (market actor) and an important developer of the Dutch healthcare sector, a brief overview of the structure of actors within Dutch hospitals is given. More elaborate dynamics are presented in the functional analysis. Currently, the type of actor and amount of actors differ per healthcare provider (R1, R2). Interest and movement of LLM usage can originate through individual medical specialist that show personal interest in the subject, defined in earlier research as 'local champions' (Strohm et al., 2020). The position of the medical specialist as important actor was emphasis several times while some medical departments show further knowledge into the field of AI than others (radiologist, cardiologist and oncologist are often more accustomed to using digital systems). As respondent AM2 indicated: *"Everyone has heard of it, some are exploring it a bit, but we are still at the beginning. A radiologist is a pleasant exception in this."* This is in accordance to literature on the maturity of AI applications in different hospital departments (Rajpurkar et al., 2022). In terms of policymakers or management within hospitals, the role of Chief Information Officers (CIOs) or Innovation officers was mentioned by several respondents (A2, AG1, AM1). Furthermore, in differing size and stage of maturity, data science teams (sometimes called Innovation teams, AI teams) have been set-up in the last decade, more on them in the functional analysis (A2, AM1, M3, M4). The IT-department, present in all hospitals, was determined as a guaranteed actor to deal with the practical side of any AI (and LLM) implementation but were not heavily involved in decision making (AM1, A2, M3, M4). Lastly, hospitals and healthcare providers have a small dedicated department for legal assurance of new used technologies, under which LLMs would also be assessed (A1, A3, AM1).

In terms of entrepreneurial market actors, several types were identified. A growing number of start-ups offering AI solutions with hospital wide applications exists in the Netherlands (AITIS). The entrepreneurial landscape is further analysed in the functional analysis of function 1 entrepreneurial activity. CEOs of start-ups stress the influence of large incumbent technology companies in hospitals in the Netherlands. M3 *"But it is a pity because it would be very nice if there was more competition between EHR suppliers and EHR suppliers that are easier to integrate for third parties where you can really bring innovation more easily that they become larger. But that is not the case because there is such a kind of monopoly of 3 major parties."* These incumbents include both companies providing medical technology services (Philips or Siemens were named) (M3), but mainly the actors in charge of the data provision of the Electronic Heath Records (EHR) (A2, M3, M4). It came forward that EHR market in the Netherlands is oligopolyzed with only three big players remaining, with the biggest commanding well over 2/3 of the markets share (Autoriteit Consument & Markt, 2021)

In the governance sphere, the following were identified: On the one hand, the Dutch government provides institutions handling national strategy and policy instruments on AI namely the Ministry of Healthcare (VWS)[12] and Ministry of Economic Affairs (EZK)[13]. On the other hand, the Dutch government provides the regulatory agencies namely, the Health Inspection (IGJ)[14] a subsidiary of the VWS, responsible for the supervision of good care and departments within the VWS like the department medical technologies, providing legal and practical guidelines for the IGJ to enforce upon.

5.1.2 NETWORKS

As of now, no designated LLM knowledge networks exist that is named as such (R1, R2). However all UMCs are involved in 'knowledge hubs' on the subject of AI (and subsequent LLMs) that were formed in approximately the last decade (R1, R2). Examples are the 'ICAI '[15] a grass-roots initiative to bundle research and development on AI under a common name, also in order to exchange knowledge, the 'CAIRELAB'[16] of the Leiden University Medical Center or the 'AI for health'[17] a collaboration between the Radboud University

[12] In the Netherlands the Ministry of Health is called 'Ministerie van Volksgezonheid, Welzijn en Sport' abbreviated as VWS
[13] In the Netherlands, the Ministry of Economic Affairs is called 'Ministerie van Economische Zaken en Klimaat' abbreviated as EZK
[14] Inspectie gezondheidszorg en Jeugd translates to: Inspection healthcare and youth. Moren information on: https://www.igj.nl/
[15] https://ic AI . AI /
[16] https://www.lumc.nl/over-het-lumc/maatschappelijke-rol/waarde--en-datagedreven-zorg/c AI relab- AI /
[17] https://www. AI -for-health.nl/

and RadboudUMC. These 'AI labs' are examples of the numerous knowledge 'hubs' that exist on the subject of AI with a primarily academic research perspective and partners. Examples of networks go by different names such as: knowledge hubs, initiatives, expertise centers, coalitions, working group, knowledge program, regional development platforms to name a few. A second class of initiatives has a broader scope, also including private, commercially oriented partners, NGO's of which the 'Nederlandse AI Coalitie'[18] is the most important. They are developed as a public private partnership and are often a central platform for multiple 'AI labs' connecting both hospitals, universities and companies. To add to that, inter-university collaboration projects such as the EWUU project that started in 2019 (between Utrecht, Eindhoven and Wageningen university) designating specific resources to AI labs stimulate the transfer of academic knowledge between actors. Networks, in the form of supply chain agreements exists in the relationships between healthcare providers and running agreements with external infrastructure parties that can provide future networks that stimulate the formation of the DH-LLMTIS. A further analysis of the characteristics of 'knowledge diffusion' landscape are given in Chapter 5.

5.1.3 INSTITUTIONS

The present institutional structures 'rules-of-the-game' in the form of laws, regulations was defined. The first point of interest in the institutions regarding LLM use in healthcare is the ambiguity, limited existence or absence of clear rules and guidelines. As is often with new technologies, guidelines are trailing behind on developments. However, LLMs in healthcare will have to adhere to existing formal institutions in the form of multiple existing regulating bodies and newly formed laws and regulations on AI, of which the most important ones that came forward are elaborated upon (R1, R2).

Medical Device Regulation (MDR)

LLM applications in healthcare can fall under the medical device regulation (A2, A3, A4, M1, M2, M3, M4). The medical device regulation, is a European wide harmonized act that provides the guidelines to which all medical devices must adhere that came into effect in 2021. Supervision and language specific guidelines are entrusted to national regulating agencies. In the Netherlands, this controlling role is entrusted to the IGJ. They assess the quality of new medical products. Medical devices can be categorized into 4 main categories. Class I: This category includes low-risk medical devices that are generally non-invasive and pose minimal risk to the patient or user. They are subject to the basic regulatory controls to ensure their safety and efficacy. Class IIa and IIb: Devices falling into these categories present a moderate to high risk to the patient, necessitating a higher level of scrutiny. Class IIa devices are associated with medium risk, while Class IIb devices pose a higher risk. They may be non-invasive or have limited invasiveness and require more extensive evidence of conformity to ensure their safety and performance. Class III: This classification encompasses high-risk medical devices that are invasive or implantable and may have a significant impact on the patient's health and safety. Due to the critical nature of these devices, they are subject to the most stringent regulatory requirements and rigorous assessment procedures to demonstrate their safety and clinical efficacy (European Parliament, 2017). Part of this regulation is the regulation specific on software as a medical device in which the conditions are stated as: "diagnosis, prevention, monitoring, prediction, prognosis, treatment or alleviation of disease, diagnosis, monitoring, treatment, alleviation of, or compensation for, an injury or disability, investigation, replacement or modification of the anatomy or of a physiological or pathological process or state, providing information by means of *in vitro* examination of specimens derived from the human body, including organ, blood and tissue donations".[19]

[18] https://nl AI c.com/
[19] Quoted from the European legal flowchart deciding if your software is going to be counted as medical device: https://health.ec.europa.eu/system/files/2021-03/md_mdcg_2021_mdsw_en_0.pdf

General Data Protection Regulation (GDPR)
As data is the fuel of LLMs, the gathering, sharing and transfer of data in all parts of the hospital or between hospitals, have to adhere to European and Dutch guidelines regarding protection of personal data use. The GDPR that came into effect in 2018, a similar European harmonization as the MDR, set up the principles and rules for correct data use by data governors. It defines individuals rights on data usage, the obligations of those processing data, measures for ensuring compliance and sanctions for breaches of the rules. These sanctions can be up to 20 million or 4% of global turnover.[20] However, rulings are often vague on their impact on AI -development. Study on the applicability of the GDPR of AI signifies that it is compliant with AI applications that 'balance' data protection but does provides limited guidance on how to achieve this goal (European Parliament Research Service, 2020). Enforcement of GDPR rules is entrusted to national agencies on data protection. The AVG[21] is the Dutch version of the Data protection law, based on and compliant with the European GDPR. The core principles are stated as follows: 1. Lawfulness, fairness, and transparency; 2. Purpose limitation; 3. Data minimization; 4. Accuracy; 5. Storage limitation; 6. Confidentiality and integrity. Following the technological qualities of medical LLMs mentioned in the technological context, these principles tend to create tension with the current limitations of current AI and LLM models.

Their compliance is enforced by the independent government organization Authority on Personal Data (AP)[22] who cooperates with the IGJ on healthcare cases. Since 2023 a new subsidiary of the AG is specifically assigned to monitor the correct usage of data in algorithms, including LLMs, called the '*directie coordinatie algoritmes*' (DCA), the organization is funded with 1 million euros per year increasing to 3.6 per year by 2026[23].

EU-AI act
The increase in AI use in general led the EU to formulating a separate EU-AI act, mentioned by all respondents from different backgrounds (R1, R2). The act, proposed in 2021 and accepted by parliament in 2023, consists of a regulatory framework that is the first attempt worldwide to regulate AI specifically. It strives to gain neutral definitions of AI and categorize AI using a 'risk-based approach'. The current information on the AI-act stems from its 2021 version with the processed alterations. However, this will give no certainty for the final version as the act is being adjusted as of now. The current version is debated about, a final version should be available by the end of 2023, coming into effect by the half of 2024 with an extended 2 year timeframe reserved for implementation[24].

The act categorizes AI systems into 4 different types of risk either being: (i) unacceptable risk, (ii) high risk, (iii) limited risk, and (iv) low or minimal risk. 'AI applications would be regulated only as strictly necessary to address specific levels.' Al medical AI systems would fall under the category iii high risk as stated: 'Systems used as a safety component of a product or falling under EU health and safety harmonization legislation (e.g. toys, aviation, cars, medical devices, lifts) (European Parliament, 2023) That is, if the usage can be categorized as a medical device. It is explicitly mentioned that providers of high-risk AI systems would be required to register their systems in an EU-wide database and that new products will fall under existing conformity networks such as the medical device regulation (European Parliament, 2023) However, 'AI systems presenting 'limited risk', such as systems that interacts with humans (i.e. chatbots) would only be subject to 'standard transparency' regulations. Enforcement is left to national market regulation committees and penalties of 30 million and 6% of global turnover are mentioned in case of non-compliance

[20] More information on the GDPR https://gdpr.eu/what-is-gdpr/
[21] AVG is the general regulation dataprotection (algemene verordering databescherming)
[22] AG' is the Dutch Authority Dataprotection (Authoriteit Gegevensbescherming)
[23] Translates to: direction coordination algorithms. More information on:
https://www.autoriteitpersoonsgegevens.nl/themas/algoritmes-ai/coordinatie-toezicht-algoritmes-ai/directie-coordinatie-algoritmes
dca#:~:text=De%20Autoriteit%20Persoonsgegevens%20(AP)%20is,directie%20Co%C3%B6rdinatie%20Algoritmes%20(DCA).
[24] https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

(European Parliament, 2023). Currently, negotiations are taking place on imposing specific obligations for 'general purpose AI models' and 'generative AI' models 'such as ChatGPT'. This will include a specific focus on LLMs. Currently it is stated that: foundational models (LLMs fall under this category as well) would be required to assess and mitigate the risks, 'comply with some design, information and environmental requirements' and 'register such models in an EU database'. Additionally, it is stated that 'all foundation models should provide all necessary information for downstream providers to be able to comply with their obligations under the AI act, disclose that the content was AI generated. On LLMs specifically it is said that: 'they have to publish information on the use of training data protected under copyright law'. Since then the compliance of the biggest publicly used LLMs with the statements in the EU-act was assessed by researchers at Stanford University. This paints a picture that sees some models scoring significantly lower than others influencing their entrance into the European market[25].

The uncommon positioning of AI applications, including LLMs, in-between data protection, Medical Device and European AI regulations, are a source of confusion for regulators (AG1, M3, M4). Legal research into the overlap of these different regulations concluded in 2021 that 'regarding the health care sector, in our view the current draft AI Act has no proper legal foundation seen from the proportionality and subsidiarity principles. This is not caused by the object of regulation, the use of AI systems, but by the choice to design the legislation'(Choi et al., 2022) . While the AI-act has been adjusted since, it indicates ambiguities on the acceptance of this new legislative package among lawyers.

5.1.4 INFRASTRUCTURE

The infrastructure that is of influence for LLMs in healthcare can be divided into two parts. The first part is the physical infrastructure that is the computing power and organization of data production and storage. The second part concerns the mobilization of data, this involves the extraction, harmonization, anonymization, pseudonymization and transfer of data. The mobilization of data will be elaborated upon in paragraph 5.2.5. Resource mobilization.

In terms of physical infrastructure, a two options are common. On the one hand, in 'off-premise' infrastructure actors mostly are limited in using infrastructure mainly provided by American large tech companies, to run or develop LLM models (AG1, M3). On the other hand, decentralized development of digital infrastructure 'on-premise' is improving (AM1, A4). It became clear that technological expectations are that smaller finetuned versions of LLMs can be developed, increasing the possibilities for more decentralized use while increasing performance of the models for that specific sector (AM1, AM3, A4).

The current dependency on US services was expressed in several parts of the infrastructure. Individual Dutch hospitals and universities do not have the capacity to run and develop entire LLMs in the order of magnitude of international renowned models like ChatGPT on their own. Therefore, often solutions are cloud-based. Thus hospitals are making use of agreements with American Cloud Providers, mainly Microsoft Azure and Amazon Web Services were mentioned (A3, AM3, AG1, M2, M3, M4). This will include a safeguarded area reserved for the actor in question to use for running their models. Development of models is both done as an adaptation of existing models developed by American Companies and by own making, using available information on earlier models. Models that are named to be used for adaptation include ChatGPT and MedPalm using tools such as Hugging Face (M2, AM1, A3). These models correspond with literature on medical LLMs that are developed.

When considering cloud options, the importance of anonymizing patient data, before it is offered to a cloud system, was stressed. However, ambiguity exists on good practice of anonymization and especially non-

[25] https://crfm-stanford-edu.proxy.library.uu.nl/2023/06/15/eu-ai-act.html

anonymized data or high-impact data makes the development of own infrastructure more needed. These insights resonate with the findings of the KPMG grey literature report in 2020. Here it is indicated that a scattered distribution of infrastructure solutions is present in Dutch healthcare, external and internal solutions where found to be almost 50/50 for both the data and model storage (KPMG, 2020).

Going forward, hospitals are considering their own independence regarding model development, in which the sensitive nature of medical data plays an important role (A3). Hospitals either build their own more powerful computing centers, capable of running smaller more specialized LLMs or exploring options of a more secured cloud option named a 'virtual private cloud' (A4, M4). The relation of hospitals to adjusting their existing infrastructure saw two distinct observations. First, it was indicated that incorporation into the existing IT is a long and laborious process. One actor mentioned seven months as minimum, another indicated that even this is fast (AM1, M4). This lengthy process decreases ease by which local LLM options could be incorporated into existing IT systems or form a barrier to scalability or transferability of models. Second, the relation of hospitals to their EHR providers was often mentioned (AG1, AM1, M2, M3, M4, AM2). EHR systems dictate for an important part the digital flexibility and possibilities of health data use and the implementation of new digital applications such as LLMs and are the main source of medical text data (Roest et al., 2019; Wornow et al., 2023). Actor M3 gave the following example: *'Sometimes a hospital can say, well, this is fantastic, this is truly beneficial. But an EHR provider can still refuse to create those spaces, and that is indeed a huge hurdle."* This relationship is further examined in the functional assessment.

5.2 FUNCTIONAL ANALYSIS

This section details the functional analysis results, merging desk research with insights from 11 in-depth interviews. These interviews collectively provided a qualitative review of the seven TIS functions defined in this research, distinguishing between DH-AITIS and DH-LLMTIS when necessary.

The findings consider the division between the different 'spheres'—governance, market, and academic—where relevant insights could be specifically attributed to one of these spheres. For instance, it was noted that the governance sphere does not significantly contribute to entrepreneurial activity within the TIS framework or that knowledge diffusion, by its nature, involves the transfer of information and insights between these different spheres, highlighting the interconnectedness of these domains within the TIS structure. Finally, at the end of each function, an synthesis of the results is given.

5.2.1 F1 ENTREPRENEURIAL ACTIVITIES

Entrepreneurial activities within the TIS are primarily centered around Market Sphere actors, many of whom are connected to the academic sphere through research. This section delineates these market actors, followed by observations on the global TIS.

Market Sphere (Start-ups)

A substantial number of entrepreneurial entities combining AI with healthcare were identified, totaling 137 companies (**Figure 6a**). This number was confirmed by respondents, who validated both the scale and specific company names, ensuring overlapping data (MG1, G1, AM3, M6, AMG1). Respondents, including various sector representatives like innovation managers and AI leads, reported a marked increase in AI application inquiries over the last 1.5 to 2 years. This surge, ranging from experimental ideas to commercial proposals, suggests growing entrepreneurial activity not yet reflected in new company formations (MG1, G1, AM3, M6), represented in the following quote by M6:

"I especially experience that as an enormous amount of email that has grown explosively. I had been on vacation for three weeks in the summer and I had not updated my mailbox. Then I came back to 4700 new emails, and I think about half of them were from interested parties who wanted to develop something with us or integrate something."

This increase in applications is also experiences by heads of departments of hospitals or the designated regulator who mention that they get notified of multiple serious AI applications per week where he could act upon (G1, AM3). A 2020 grey literature report analysed as much as 400 AI applications in the Netherlands, with only 18% demonstrating externally validated business cases. This highlights a disconnect between the siginificant amount of AI experimentation and viable commercialization in Dutch healthcare (KPMG, 2020).

The spread over the years, as shown in **Figure 6b,** paints a two-sided picture. On the one hand, it is clear that there has been an increase in start-ups founded in the Netherlands that use AI in healthcare over the last 10 years. In the year 2015 alone as much companies were founded as the total before 2013. On the other hand, the 'spring' of AI following the deep learning revival in 2017 is not clearly visible. To add to that, the number seems to be steady after 2017 and even decreasing after 2019. Respondents also indicate that they see an increase in research on both AI and LLMs / NLP models specifically, but that the amount of models that reach the ending stage of technological readiness levels is lagging behind (G2, AM4). Furthermore, with regard to start-ups and spin-ups from universities an innovation manager at an UMC indicates that this level has remained stable in the last years. *'Start-ups founded around 2015 – 2017 are just starting to make revenue and survive on their own'* (AM2). To add to that, an AI director at an UMC indicates that the people they come across in this market have stayed the same over the last 2 – 4 years (AM4). This can have multiple explanations as the number of entrepreneurs active in the field have been saturated and/or there is a delay before new companies show up in the statistics. This trend is also supported on a global scale by research from (Zahlan et al., 2023), indicating the same flattening of the trend around 2020, they pose that Covid-19, shifts in investors priorities and market consolidation. The saturation of the field is supported by a respondent who indicated that the AI in health sector is currently a small field in the Netherlands: *"Yes, we know the guys from [NLP Start-up], I know the people from [Incumbent LLM application] and I also know that our CCO is from [incumbent] and sold their voice division to [Incumbent LLM application] it's a pretty small world, I can tell you."* Either way, both the desk research data and interview data does not point to a surge in newly found companies.

The temporal distribution of start-ups (**Figure 6b**) reveals a growth in AI healthcare companies over the last decade, with a notable spike in 2015. However, the anticipated surge post-2017's deep learning advancements is not evident. The data shows stabilization post-2017 and a decline after 2019. Interviews suggest an increase in AI and LLM/NLP model research but a delay in achieving higher technological readiness levels (G2, AM4). Stability in university start-ups and spin-offs is noted, with those established around 2015 – 2017 beginning to generate revenue (AM2). Global trends show a similar plateau around 2020, attributed to COVID-19, shifts in investor priorities and market consolidation (Zahlan et al., 2023).

Regarding LLMs or NLP integration, only 8 out of 137 companies were identified. Post-2021, no new Dutch companies have emerged using LLMs or NLP, indicating a nascent stage in commercial deployment. These findings will be elaborated in Function 4: Market Formation.

The companies that were found to deploy LLMs/ NLP in healthcare in the Netherlands were all manually researched to find out their application area and maturity level. All of them were either in trial phase or just started out. Respondents were asked about these companies and had only heard of some of them, being in a trial phase as well (A2, M5, MG1, AMG1): *"For example, in [UMC] they do more NLP (Natural Language*

Processing) focused projects; there they do more with [application], but in my opinion, it is still in its infancy"(A2). Companies either used LLM /NLP for audio to text in a clinical setting or processing medical literature and documents. No company was found that generated new medical information or influence the decision process of doctors, this is further described in the 5.2.6. Legitimacy.

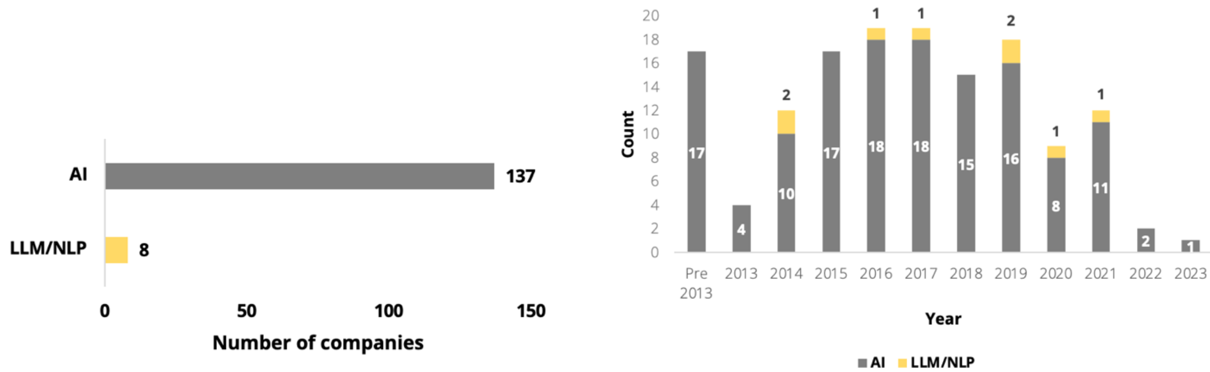
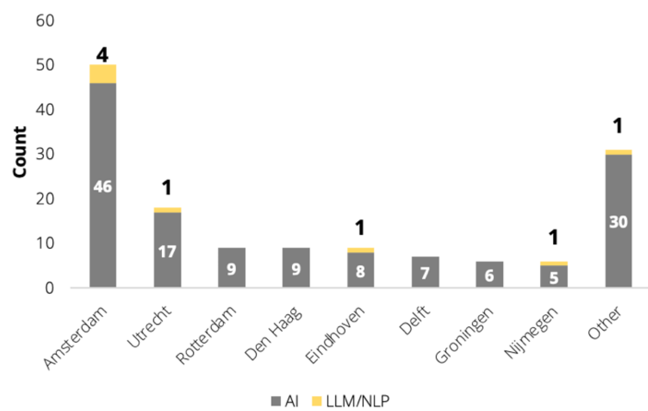


Figure 6: a. Dutch healthcare companies on AI and LLM/NLP in healthcare **b.** Companies using AI and LLM/NLP in healthcare founded in the Netherlands per year.

The geographical distribution shows Amsterdam as the frontrunner, followed by Utrecht (**Figure 7a**), with factors like UMC presence, technical workforce, funding, and talent influencing this trend (M7, AM2).



	AI	LLM/NLP
US	*2583*	*215*
UK	*371*	*25*
Germany	*180*	*7*
France	*145*	*7*
Switzerland	*95*	*9*
Netherlands	***74***	***3***
Sweden	*42*	*2*
Denmark	*34*	*0*
Belgium	*26*	*1*

Figure 7a: Companies on AI and LLM/NLP in healthcare per city in the Netherlands **7b.** Worldwide Crunchbase search on number of new companies per country

An additional category of start-ups was identified, focusing on creating prerequisites for AI healthcare implementation, such as data mobilization and legal compliance (M5, M6, M7, AM3). These complementary services are further discussed in Function 7: Synergies.

Market Sphere (Incumbents)

Beyond start-ups, incumbents in Dutch healthcare are also exploring AI and LLMs. EHR providers stand out as influential in entrepreneurial activity (M4, M5, M6, M7, AM1, AM3, MG1). Interviews revealed divergent

AI development strategies between the two largest providers. One provider fosters an environment conducive to using their text data for LLMs and NLP, promoting collaboration with UMCs, research projects, and start-ups, without intending to develop AI applications in-house or integrate them into their services. They specifically focus on LLM and NLP functionalities (M6).

However, the effectiveness of this collaborative approach varies. Some respondents recognize these efforts (M5, M7, AM2), but others, like two AI officers at UCMs (AM4, AMG1), haven't seen its practical application. In contrast, the other provider is less collaborative, not seeking partnerships and instead aiming to develop and launch their own AI applications (M5, M6, M7, MG1, AM4).

Market Sphere (Hospitals)

Hospitals, particularly University Medical Centers (UMCs), are crucial for entrepreneurial experimentation, often leading to AI application spinoffs with commercial potential (M5, AM2, AM3, AM4). Detailed discussion on this is in 5.2.2. Knowledge Creation and Development. In-house projects typically originate within UMC research groups and are enhanced by Innovation or Data Science teams. However, collaboration with external commercial entities is not standard, as many projects are designed for internal use (AM2, AM3, AMG1). One UMC reported 15 to 20 ongoing AI research projects, highlighting significant in-house entrepreneurial experimentation (AM2).

Regarding collaborations between hospitals and start-ups, perspectives vary. Some respondents note UMCs are well-positioned for such collaborations due to their advanced knowhow and infrastructure (AM4, M5), while others argue that 'Top Klinische Ziekenhuizen' (TKZs) are more agile due to their leaner organizational structure, enabling quicker movements (M5, M7).

G-LLMTIS

A Crunchbase comparison in (**Figure 7b**) positions the Netherlands alongside the US and European neighbors, indicating a competitive position in start-up founding, similar to to its peers as confirmed by Zahlan et al., 2023.

In short term, respondents view Dutch start-ups as viable against large, partly international incumbents, through leveraging the unique aspects of the Dutch language and healthcare system's data infrastructure (AM2, AM4, MG1). The Dutch market's lower priority for major international tech companies results in suboptimal performance in this region (M5, M6, M7, MG1, AM4). Key global influences on the Dutch LLMTIS include an international EHR provider and an American software giant owning a medical LLM/NLP company (M5, M7, MG1, AM4). While these global players could overshadow Dutch efforts in the future, their current performance doesn't surpass local initiatives (M5, M6, M7, MG1, AM4). Adding to this is the upcoming EU AI-act, with stringent LLM model design and training requirements, that might deter American market entry into Dutch healthcare in the near term.

F1 synthesis

In the 'formative' phase of entrepreneurial activity, typical characteristics include a low number of entrepreneurs and growth through the influx of new actors (Bergek, 2019; Bergek, Jacobsson, & Sandén, 2008). Analysis of the DH-LLMTIS reveal a recently established entrepreneurial community. However, there's a noticeable stagnation in new entrants, possibly due to market saturation or insufficient motivational drivers, impacting other functionalities.

A crucial milestone in transitioning from 'nascent' to formative phase is the first commercial deployment (Bento & Wilson, 2016). DH-AITIS has achieved this, with examples of commercial deployment reflecting dependencies on contextual TIS actors as posed by Hekkert et al. (2007). Conversely, DH-LLMTIS is nearing its first commercial deployment, evidenced by startups and pilot projects. Both Dutch and international incumbents are active in the focal TIS, with Dutch startups holding a competitive international position. These insights indicate that DH-LLMTIS's F1 is on track with early-stage development expectations, suggesting it's not yet poised for a growth transition.

5.2.2 F2 KNOWLEDGE DEVELOPMENT & DIFFUSION

The function of knowledge creation is divided into two segments: market and academic knowledge development. The market sphere focuses on the evolution of AI and LLM knowledge among market actors, particularly in hospitals, as entrepreneurs inherently possess knowledge on the subject. The progression of academic knowledge is scrutinized through specialized educational programs and publishing output over time. The function of Knowledge diffusion occurs across different spheres, and the impact of the G-LLMTIS is briefly considered.

5.2.2.1 F2 KNOWLEDGE DEVELOPMENT

Market Sphere

In hospitals, knowledge development primarily stems from University Medical Centers (UMCs) conducting AI and LLM/NLP research in collaboration with universities (M5, AM5, AMG1), aligning with the noted publication output. Recently, research is increasingly consolidated within AI labs, formalizing knowledge development. However, knowledge levels across hospitals are fragmented, varying by department, with radiology often leading (AM3) *"Well, I know that radiology has, of course, been familiar with these types of systems for a much longer time and is more familiar with them."* , as supported by (Strohm et al., 2020)(AM3, AMG1).

Additionally, on a more management level, the knowledge level differs greatly per hospital as exemplified by the account manager of an AI company in contact with multiple hospitals: *"It's so different in every hospital. It depends a lot on whether someone has a feeling for it. Frankly, even how old someone is. CMIOs who are 60 which we don't get much from compared to the example of a hospital pharmacist, he's 26, and he built a query in his system in half an hour and of course, there's a whole world in between."* Furthermore, the same dynamic between UMCs and TKZs becomes apparent. For example, TKZs do not have to take into account the opinion of data science teams and large research departments. Another reason for this dynamic, is the input of 'local champions' [26] (M5, M6, M7, AMG1).

However, hospital-wide the general knowledge level has increased with the introduction of ChatGPT, a dynamic shared by multiple respondents and elaborated upon in 5.2.6. Legitimacy (AG, AM2, MG1, G1, M6). This knowledge is far from in-depth as exemplified by the NLP professor: *"In a hospital setting, I have to talk about ChatGPT, and it starts off a bit informally, then I'm engaged in a conversation for about 15 minutes about the well-known things that I, of course, have already heard a thousand times."* Despite varying starting points, the integration of applications by incumbents can rapidly advance a hospital's knowledge position, even if it previously lacked in-house research capabilities (M5).

Academic Sphere

Educational programs on the subject of AI in healthcare were manually scanned at all Dutch universities. The results of which are visible in **Figure 8**. Of all the thirteen Dutch universities, eleven offered Bachelors and Masters programs in AI, all of which featured courses in NLP. eight offered specific courses combining AI with healthcare and four offered special tracks concerning AI in healthcare, two of them being at a technical university and two at a university connected to a UMC. One specifically mentioned using NLP in a

[26] Local champions are described in literature as individuals that can make a significant personal difference in the innovation process (Rogers 2003).

healthcare setting[27]. All of these specialized tracks and courses were formed in the last five years. This indicates both a growth in the knowledge development but also novelty and inconsistent standardization or formalization across Dutch universities.

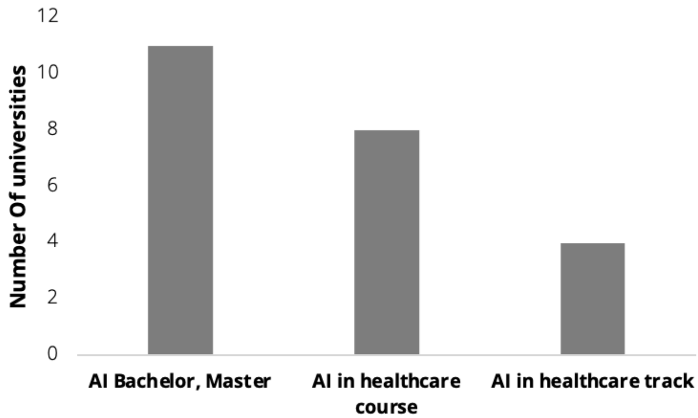


Figure 8: Academic educational programs on AI and AI in healthcare in the Netherlands (13 universities total)

AI's academic presence in the Netherlands has deep roots, dating back to the 1990s with the fusion of Data Science and Linguistics, forming a robust NLP research community (AG). This enduring interest in the AI field has fluctuated, mirroring global AI 'winters' and 'springs' (Bommasani et al., 2021) Since 2017, a notable rise in student applications for AI-related courses has been observed, yet university responses have been somewhat disjointed, with AI and NLP remaining scattered across academic departments (AG). This situation has led to universities often being reactive rather than proactive in addressing emerging AI trends. The historical depth and recent surge in student interest in AI and NLP are reflected in both academic output and the structure of educational programs, particularly those focused on healthcare applications. In that respect, Universities take on a role of laggard as said by the professor: *"And when such an emerging field has been waving red flags for five years, there's a chance, but just a chance, that a dean who is up-to-date with the times might take action on it, proactive action instead of reactive"..*

Considering the global transferability of academic knowledge, an increase in international publications on LLMs and NLP has also been noted. (**Figure 9a**) illustrates this global trend since 2017 across major academic literature databases. In the Dutch context, (**Figure 9b**) shows a similar, albeit smaller-scale, increase in publications, with a more than tenfold rise since 2017, indicating both the novelty and growing research interest in this area.

[27] https://vu.nl/en/education/master/artificial-intelligence

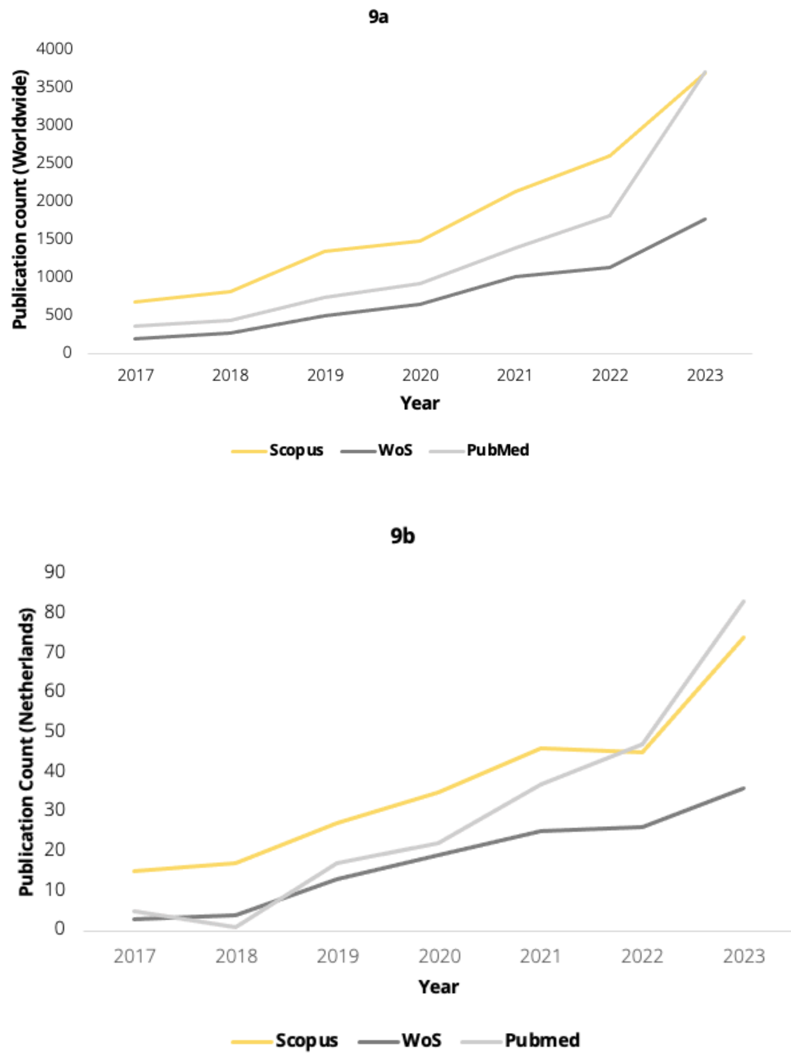


Figure 9: a. Worldwide publications on LLMs/NLP in healthcare per year **b.** Dutch publications on LLMs in healthcare per year

Furthermore, **Figure 10** shows the effect of removing articles mentioning ChatGPT since its launch in the end of 2022. It is seen that the percentage of academic literature mentioning ChatGPT is both significant but not overwhelming, indicating that knowledge development is not only focusing on this noteworthy application. More on this relation is said in 5.2.6. Legitimacy.

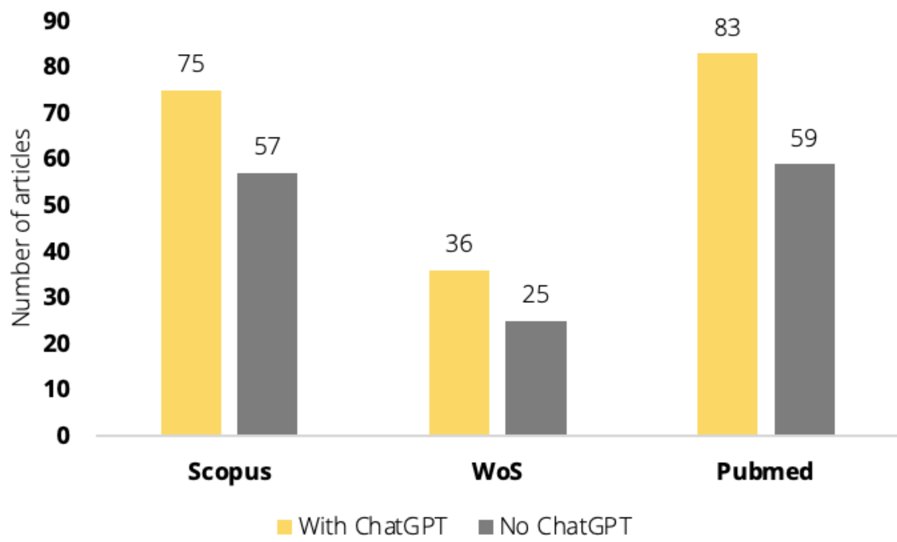


Figure 10: Dutch publications on LLMs/NLP in healthcare since 30th November 2022 with and without mentioning ChatGPT in title, abstract and keywords

G-LLMTIS

The comparison of Dutch academic output with other countries, detailed in (**Table 6**), positions the Netherlands at 11th and 12th places worldwide in terms of absolute numbers. This parallels the country comparison in the start-up landscape, affirming the Netherlands' robust international standing in academic research. While the United States emerges as the leader in sheer volume of publications, a per capita analysis offers a more nuanced perspective. The publication rate per capita in the Netherlands is comparable to that of the United States in Scopus and Web of Science (WoS) databases, and higher in PubMed, indicating a strong research output relative to population size.

	Scopus			Web of Science			Pubmed		
Position	**Country**	**# Pub.**	**Pub / Capita**	**Country**	**# Pub.**	**Pub / Capita**	**Country**	**# Pub.**	**Pub / Capita**
	World total	*12490*	1.5	World total	*5394*	0.7	World Total	*8418*	1.0
1	US	*4973*	14.8	US	*2735*	8.2	US	*3683*	11.0
2	China	*1423*	1.0	China	*549*	0.4	China	*925*	0.7
3	India	*1219*	0.9	United Kingdom	*431*	6.4	UK	*805*	12.0
4	United Kingdom	*932*	13.9	India	*294*	0.2	Canada	*405*	10.1
5	Canada	*526*	13.2	Canada	*272*	6.8	India	*404*	0.3
6	Germany	*505*	6.0	Australia	*250*	9.6	Germany	*343*	4.1
7	Spain	*473*	9.9	Spain	*201*	4.2	Australia	*302*	11.6
8	Australia	*450*	17.3	Germany	*182*	2.2	Italy	*268*	4.6
9	Italy	*393*	6.8	Italy	*165*	2.8	Spain	*258*	5.4
10	France	*373*	5.5	France	*140*	2.1	France	*248*	3.6
11	**Netherlands**	***260***	***15.3***	South Korea	*137*	2.7	Japan	*218*	1.7
12	-	-		**Netherlands**	***126***	***7.4***	**Netherlands**	***198***	***11.6***

Table 6: World ranking of number of publications on LLMs/NLP in healthcare since 2017, based on affiliation countries of all mentioned authors

F2 (development) synthesis
In early-stage TISs, patents are usually scarce, aligning with the decision to exlcude patent searches due to their expected minimal presence (Markard, 2020). This lack of patent activity is typical for the early phases of a TIS, characterized by intense R&D, technological uncertainty, and diverse experimentation (Bergek, Jacobsson, Carlsson, et al., 2008; Markard, 2020). This is particularly true for LLMs in healthcare, where technological performance remains uncertain (Drazen et al., 2023; Thirunavukarasu et al., 2023; Wornow et al., 2023). The absence of established key performance indicators, a focus on product rather than process innovation, and no dominant designs are indicative of a TIS's formative phase, with progress to later stages marked by the development of these elements (Bento & Wilson, 2016).

The increased activity of R&D in knowledge development within the DH-LLMTIS is evident. This is manifested in the growing number of academic publications, the establishment of dedicated AI research laboratories in academia and hospitals, and the integration of TIS-related content into educational programs. The modest volume of academic articles exclusively focused on ChatGPT suggests a broader scope of knowledge development, not solely driven by fluctuations in expectations (F3) and legitimacy (F6). A comparative analysis of academic knowledge development positions the Netherlands favorably on a global scale, indicating a robust contribution to this field. These observations imply that the F2 development function of the DH-LLMTIS is effectively advancing in its formative stage. The alignment of R&D focus, academic contributions, and emerging formal knowledge frameworks underscores the system's progression in its developmental phase.

5.2.2.2 F2 KNOWLEDGE DIFFUSION

Netherlands
The structural analysis reveals that knowledge diffusion for AI in healthcare in the Netherlands is active yet dispersed. There are no distinct actions dedicated to knowledge diffusion in LLMs and NLP (AM1, AG1). However, across the healthcare sector, a variety of knowledge diffusion networks exist, which can be categorized into five types.

Firstly, the Nederlandse AI Coalitie (Dutch AI Coalition) and its healthcare subsidiary form a central, government-funded network connecting academia, hospitals, companies, and regulators (AG1, AM3, M5, M7, AMG1, MG1, G2). An example of their knowledge diffusion is an introductory AI course for healthcare employees, reportedly completed over 10,000 times (MG1). However, its role as a central information hub is not universally recognized (AM2, M5), suggesting that diffusion may occur primarily at the management level.

Second, in the last four or five years decentralized or bottom-up knowledge networks have sprung up, mainly as diffusion tool between hospitals. Multiple examples came forward from the respondents; 1. The network AI network for medical specialists, founded in 2019, organizing information events for medical specialists on AI[28] (M7) 2. The informal network of data-science and AI teams of UMCs. Founded around 3 years ago, data-teams meet-up quarterly to openly share their thoughts and developments of AI in their corresponding hospitals. They are considering building a database of all the used models in their hospitals (AMG1, AM4). AM4 elaborates on the bottom-up character in the following way: *"It all started within the [UMC] with someone from the [UMC2]. Two managers from the then AI teams, I believe they were the first two AI teams at that time, who simply wanted to share their experiences.".* 3. as overarching bottom-up network to connect academia and market actors, the ICAI 'AI-labs' are an overarching 'network' of AI-labs in and outside the healthcare sector that was founded in 2018. It provides a framework for research to make a transfer to a commercial application however, active knowledge diffusion is not significant and is referred

[28] https://demedischscialist.nl/agenda/netwerk-artificial-intelligence

by two respondents connected to research groups at UMCs as a group of researchers operating under the same flag (AM2, AM4).

Third, the Expertise Centrum Zorg Algoritmen (Expertise center health algorithms)[29], founded in 2021, is a collaboration of 28 SAZ hospitals (a third category of general hospitals not belonging to the UMCs or TKZs) sharing knowledge and data on AI (MG1).

Fourth, direct C-suite meetings between CMIO's and other healthcare managers. It is indicated that technological developments on AI are shared directly between managers of hospitals and even search queries for an NLP program are sometimes transfered (M5). However, a general dynamic was touched upon that sees hospitals in the geographical south of the Netherlands sharing developments more openly than the UMCs of the biggest Universities in the West of the country (M5).

Fifth, a last category is the collaboration between universities such as the 'EWUU' project mentioned in the structural analysis. No large scale project of that kind was found on AI in healthcare specifically.

Despite these diffusion mechanisms, a KPMG report from 2020 indicated that hospitals often lack awareness of AI applications used in other hospitals (KPMG, 2020). This is exemplified by the diversity in networks and the misalignment of efforts within individual hospitals. For instance, different data science teams within a single UMC were found to be unaware of each other's work (M6): *"It turned out that in that room there were five different data science teams from one University Medical Center who did not know of each other's existence and were not informed of each other's initiatives."* indicating a lack of internal communication and oversight of AI models (M6,AM3, AMG1, G1).

G-LLMTIS

International knowledge diffusion in AI is primarily facilitated through a mix of academic publications, online resources, and healthcare conferences, both physical and digital. Notable American conferences like the HIMSS[30] and RSNA[31], are mentioned by an NLP healthcare company's account manager (M5), extensively feature AI, with RSNA explicitly focusing on NLP and LLMs. The upcoming US-based HLTH[32] conference in Amsterdam, highlighting "Beyond GPT: The new world of data and AI" as a major theme, exemplifies the growing attention to AI, including LLMs/NLP. These conferences, attended by diverse healthcare professionals, serve as platforms for knowledge diffusion and agenda setting in AI. These various knowledge diffusion networks and initiatives reflect not only large activity but also a significant need for knowledge exchange within the sector.

F2 (diffusion) synthesis

The formative phase of a TIS is marked by the emergence and subsequent convergence of networks, aligning with the institutional and structural context of the TIS (Alkemade & Suurs, 2012; Bergek, Jacobsson, Carlsson, et al., 2008; Suurs et al., 2010). Data clearly indicates network formation within the DH-AITIS, facilitating knowledge transfer to the DH-LLMTIS. International recognition of this knowledge diffusion is evident from the inclusion of AI topics at leading global healthcare conferences.

However, the current landscape remains fragmented at three levels. At the microscale (within organizations), grassroots knowledge-sharing initiatives exist, yet research groups within the TIS often operate in isolation. For example, hospitals are not fully aware of all operational AI models. At the mesoscale

[29] https://zorgalgoritmen.nl/
[30] https://www.himss.org/global-conference
[31] https://reg.meeting.rsna.org/flow/rsna/rsna23/RSNA2023/page/sessioncatalog
[32] https://europe.hlth.com/about-hlth-europe

(between organizations), numerous independent knowledge networks have formed, partly due to competitive dynamics in the Dutch healthcare sector. This competitive environment, while not the primary focus of this study, can pose challenges for TIS development. At the macroscale (national policy), the Dutch government has supported the creation of multiple, sometimes overlapping, knowledge networks.

The extent of knowledge diffusion activities aligns with the characteristics of a TIS in a formative phase, as opposed to a nascent phase which is typically dominated solely by R&D networks (Bento & Wilson, 2016). Initiatives like the AI coalition and the appointment of AI innovation officers in University Medical Centers (UMCs) underline this phase. These findings suggest that the F2 difusion function of the DH-LLMTIS is in line with expectations for a TIS at an formative stage. Nonetheless, key drivers for transitioning to a growth phase appear to be missing.

5.2.3 F3 GUIDANCE OF THE SEARCH

To assess the guidance of the search function, a combination of a document analysis combined with respondent validation giving a qualitative indication of the perceived guidance in policy, regulations and expectations was performed. Specific policy naming LLMs and NLP was sparsely available, the bulk of the found information concerned 'AI-wide' guidance of the search.

5.2.3.1 F3 POLICY

Governance Sphere

The Dutch government's formal policy on AI began in 2019 with the SAPAI (Strategisch Actie Plan Artificiële Intelligentie) program, initiated by the Ministry of Economic Affairs and Climate Policy. This policy direction started out by the 2018 'Vergroten, Versnellen, Verbinden' report from a collaboration of Dutch research institutes and commercial consultancy firms, urging action to maintain the Netherlands' international standing in AI research, development, and usage of the informal collaboration between several Dutch research institutes and commercial consultancy firms (Topteam ICT, 2018)[33]. The report highlighted the risk of falling behind as many OECD countries had already formalized their national AI strategies.

After this report was picked up, AI was seen as a *'key-enabling technology'*[34] by the Dutch government and from the SAPAI strategy, the Netherlands AI Coalition was formed and an investment plan was set up in the form of the AiNed Programme. Furthermore, through SAPAI the Netherlands Organization for Scientific Research (NWO) its subsidiary Taskforce for Applied Research (SIA), the Netherlands Enterprise Agency (RVO), applied research institutes (TO2) and a publicly funded private investment agency (InvestNL) were all involved in speeding up collaboration between government, industry, academia, and civil society (Claudio Lazo et al., 2023)[35]

Healthcare, significantly mentioned in SAPAI, received focused attention in AI policy thereafter. By the end of 2019, the 'waardevolle AI voor de gezondheidszorg'[36] 2-year program was launched by the Ministry of Health, aiming to "enhance value creation for healthcare providers, patients, and citizens through AI utilization". This initiative, involving multiple stakeholders including the healthcare department of the Dutch AI Coalition, aimed to provide clear guidelines for AI implementation in healthcare, categorized into strengthening shared ownership, stimulating sector-wide growth, and providing implementation guidelines.

[33] Translates to: Enlarge, Accelerate, Connect'
[34] https://www.nwo.nl/en/key-enabling-technologies
[35] This collaboration model is also known as the 'Quadruple helix model'
[36] Translates to: valuable AI for healthcare

The actions undertaken by the Ministry to provide guidelines took form in two policy documents: "Leidraad kwaliteit AI in de zorg" and "Nationale routekaart databeschikbaarheid AI"[37], a collaboration between multiple government organizations. Among them is Nictiz, a VWS funded agency founded in 2002 that has as goal to improve digital applications in Dutch healthcare[38] (Ministerie van Volksgezondheid Welzijn en Sport, 2022c).

The results of the 'waardevolle AI voor de gezondheidszorg' program are again explicitly mentioned in a letter of the minister of Health to parliament in 2022. The letter reiterates AI as a policy priority, with a subsequent Ministry report highlighting regulatory issues and the necessity for guidelines on appropriate AI use in healthcare (Ministerie van Volksgezondheid Welzijn en Sport, 2022b). Additionally, the 'integraal zorgakkoord' strategy alignment document, supported by various healthcare sector actors, outlined necessary digitalization steps for AI integration in healthcare, though not explicitly mentioning AI (van Volksgezondheid & en Sport, 2022b).

The Dutch AI Coalition's 2025 goals include implementing 21 AI applications nationwide, leading globally in at least one AI field, and developing an effective health AI ecosystem with over 300 participants (Nederlandse AI Coalitie, 2020). In summary, the Dutch government has played an active role in positively guiding AI development in healthcare over the past five years

Market Sphere

In order to assess the presence of AI and or LLMs/NLP in the formalized policy of healthcare providers, the most recent policy documents of all the 7 UMCs and two consortia of TKZ accounting for fifteen TKZ were manually scanned (10 documents). Of all the UMCs, 6 name AI as a strategic point, 1 just mentioning the term without further policy actions and 1 not naming AI at all. The reports of two consortia of TKZs, both name AI as strategic point. However, no extensive formalization for AI within the organization was further mentioned. The interviewed respondents confirmed that AI has been a policy subjects since the last years but that formalization in organizational structure is varying, this is confirmed by the consultancy report by M&I Partners (2023), finding AI policy formalized in 68% of the 42 researched hospitals. AI has been introduced as a subject, being in part of the conversation of the board (M6). Multiple respondents indicate that there has been a more formalized vision in the last couple of years (M6, M7, AM2, AM4). However, organization-wide alignment of an AI policy or vision is less common as exemplified by (AM3): "*I think every department has been dealing with its own AI piece. Probably not entirely uniform and probably not completely according to standard, but just out of interest. And in the long term, that is also not sustainable in any organization and also not for hospital operations, because it will require too much maintenance.*"

A significant insight from the interviews is the bottom-up nature of AI implementation in hospitals. Knowledge networks, departmental initiatives, and medical specialists independently experimenting with AI are key drivers of this bottom-up approach (AMG1, AM4, M6),. The accessibility of AI and LLM technologies has markedly increased, facilitated by open-source information and tools, as (AG) notes: with basic Python knowledge, anyone can access a wealth of NLP resources. This ease of access has broadened engagement with the technology, positively influencing the direction and adoption of AI (AM3). Consequently, the G-LLMTIS exerts a considerable influence on the Dutch market, shaping AI's trajectory and application within the healthcare sector.

Academic Sphere

Universities have been somewhat lagging in formulating clear, top-down policies on AI, as indicated by AG. The establishment of ICAI labs, initially as grassroots initiatives, exemplifies this lag, suggesting that

[37] Translates to: 'quality guideline AI in healthcare' and 'national roadmap data availability in healthcare'
[38] More information on Nictiz is available on their website: https://nictiz.nl/

university management was not initially attuned to AI trends. However, a review of the yearly reports from all thirteen Dutch universities reveals a different picture: AI is a featured topic in each report. Approximately half of these universities (seven) identify AI as a strategic theme or have significant initiatives related to it. Specifically, one university mentions AI in the context of healthcare. Nevertheless, none of these reports specifically mention LLM, NLP, or GenAI.

Although the NLP professor's assessment may reflect personal experiences and observations over an extended period, the current strategic documents from these universities in 2022 show a considerable level of interest in AI at the top management level. This suggests a growing acknowledgment and integration of AI themes within university strategies, aligning with the broader trends and developments in the field.

F3 (policy) synthesis

Alkemade & Suurs (2012) highlighted the importance of 'guidance of the search' in early TIS development through policy measures, industry standards, and regulations. Bergek et al. (2008) linked the onset of the formative stages to institutional alignment and network formation. Suurs & Hekkert (2009) described this phase's progression as a convergence of knowledge development and diffusion, emphasizing an interplay between functions F2 and F3.

Recent trends show a marked increase in policy measures in the focal TIS, particularly in Dutch AI healthcare policies since 2018, exemplified by initiatives like the SAPAI and AI policy implementations in hospitals and universities. This reflects the contextual dependence on DH-AITISs, as discussed by Markard (2020). These observations suggest that the F3 policy function of the DH-LLMTIS is rapidly progressing in its formative stage, especially in policy guidance, marked by a notable increase in policy direction and measures.

5.2.3.2 F3 REGULATIONS AND INDUSTRY STANDARDS

The structure of regulation of AI and LLMs in healthcare revolves around the three defined pillars as discussed in the structural analysis 1. Regulation of medical devices (MDR) 2. GDPR 3. Regulation of AI through the AI-act. Below, the most important observations that came forward from the qualitative assessment of the respondents on guidance via regulations are listed whereafter the events indicative of the Global LLMTIS guidance of the search are presented.

1. **Medical Device Regulations**

The Medical Device Regulations (MDR), monitored by the IGJ, frequently emerge as a crucial framework among stakeholders (R2 1-11). The MDR, predating the widespread adoption of AI, does not adequately address current technologies like AI and LLMs, lacking specific adaptations for these advancements (M7). This gap is highlighted by the designated regulator G1, who notes the MDR's insufficiency in guiding contemporary applications. Additionally, the distinction between simple search functions and the generation of medical information sets the boundary for classification under the Medical Device Software regulation (G1), aligning with recent discussions about regulatory oversight of LLMs in healthcare.

Meskó & Topol (2023) suggest differentiating LLMs trained for medical versus non-medical purposes. The iterative nature of LLMs and AI applications complicates certification processes, as it is impractical for developers to revalidate with each software update (M7, G1, G2). The FDA's method, which involves pre-agreeing to a performance bandwidth with the regulator, is recognized as a more fitting and developed approach (G1): "*The FDA in America introduced a framework a few years ago on how to deal with these types of issues using a method where you agree on a bandwidth in advance with the regulator. If my AI performs within that bandwidth then it's just okay.*"

Extensive literature is available on the technical difficulties of regulating technologies like LLMs where concepts like transparency, interpretability and responsibility requirements are often named (Meskó & Topol, 2023; Singhal et al., 2023).

Enforcement of the MDR in the Netherlands is characterized as stringent (MG1, AM2), prompting considerations for 'regulative oases' to foster innovation. The German DiGA system exemplifies a more flexible approach, allowing market entry under conditional standards, with the possibility of recall if necessary (MG1).[39] Because the regulatory challenges and strict enforcement in the Netherlands, actors tend to avoid MDR responsibilities, favoring in-house development within hospitals rather than assuming the role of a medical device producer (AM2, AM3, AM4, AMG1). Dedicated AI/Data science teams strive for centralized oversight of regulatory compliance (AM2, AM3, AM4, AMG1). When actors decide to enter the market, the IGJ is responsible for enforcement, necessitating up-to-date knowledge of individual applications and new market entries (G2): "*Yes, that would then be the IGJ's responsibility to check that subsequently. Being up-to-date about individual applications and about what new applications become available."* Manufacturers have a duty to seek review for their applications, and in ambiguous cases, regulators may consult with European agencies, initiating a lengthy and intricate process (G2).

2. General Data Protection Regulation

The interaction between AI systems and GDPR is evolving towards standardization, despite ongoing uncertatinties (M5, M7, AM2, AMG1). An NLP company's account manager notes a shift towards clearer data sharing standards: *"...various standard norms, registrations that also asked for... but at least you know what they are now, what is allowed and what is not, and still there is so much grey and black area."* (M5). This growing clarity reduces concerns about improper data usage, critical in healthcare (AM2, M7).

However, GDPR compliance remains complex, especially around anonymization standards, which are not clear to all (A3, M5). Expertise is often outsourced: "*Talk to the academic hospital; they have plenty of people who know about this stuff,"* (A1) and *"Data regulation? I prefer to stay out of that myself..."* (A3). LLM developers and guideline creators sometimes lack the necessary knowledge to ensure adherence and overview of the bigger picture (A1, A3).

On top of this unclarity, as with the MDR, it is expressed that the Netherlands is particularly strict regarding the enforcement of the GDPR, making no exceptions to the data sharing of patients that is not for research. This is officially restricted, even for inhouse use while it does happen in practice (M3, AM2). These factors have resulted in large differences between the perceptions of actors regarding internal or external data us. With internal data the only barrier found is the correct extraction of data from existing databases. However, the moment data has to be transferred outside the organization in any way, a cascade of 'difficulties' are perceived (M5 M7, AM2, AM3, AM4, AMG1).

As a result, in-house development is favored for its perceived safety and simplicity (AMG1, AM4). This cautious approach is exemplified by a hospital's reluctance to engage in a GDPR-involved project without collective responsibility and certainty of compliance (M3). The sector's stance on GDPR, similar to that on the MDR, is one of caution and a tendency towards internal development to manage regulatory challenges.

3. EU AI-act

As noted in the structural analysis, the AI-act will come into effect in 2024, with a two-year grace period for implementation. Stakeholders try to align with the AI-act in advance (R2 1-11). The Dutch healthcare sector's interests are indirectly represented in negotiations through the collaboration of IGJ and VWS with the Ministry of Economic Affairs at the forefront. These discussions are ongoing as of today (G1, G2).

[39] More information on: https://www.bfarm.de/EN/Medical-devices/Tasks/DiGA-and-DiPA/Digital-Health-Applications/_node.html

With the positioning and execution of the law on a national level, the government agencies IGJ and VWS both express their intentions to engage in knowledge provision and diffusion, communicating the implications of the AI-act with the healthcare sector. These diffusion networks were named: zorg voor innoveren[40], Dutch Ai Coalition, Government Entrepreneurship Agency[41]. To add to this, recently a platform project was started called 'Platform EU AI-act for healthcare'[42] by VWS to investigate the effects of the AI-act on Dutch healthcare. Precise execution of this knowledge diffusion and the allocation of responsibilities of regulations are currently unclear (G1, G2). However, in terms of law enforcements the IGJ sees itself as primary regulator of the AI-act in healthcare when it comes into effect (G1).

Market actors are maneuvering in response to the forthcoming regulations. Innovators cite ambiguities and slow progress around the AI-act's developments and information (G1, M7, AM2, AM3, AM4, AMG1). One actor talks of plans to develop bottom-up guidelines: *"The AI-act and all those things are too slow to deliver the deployment of LLMs in a responsible manner in the hospital... we actually have to define what is effective and what is safe for LLMs."* (M7). The act's requirements are too abstract for developers to operationalize (AG1, M7, AM2). For instance, the AI-act's mandate for "high-risk AI systems' data sets to be 'relevant, representative, free of errors, and complete'" is often impractical due to the inherent inaccuracies in healthcare data sets, which vary by provider, region, and disease. Therefore, a recent plead has been published to codify the transparency of data in a 5 layer system that is proposed, signifying the interaction between regulative choices and technical realities (Choi et al., 2022; Ferrari et al., 2023).

5.2.3.3 F3 EXPECTATIONS

In the Dutch governance context, the delegation of responsibility for new AI policy remains undetermined (G1, G2, AMG1, M3). Policy documents like "Leidraad kwaliteit AI in de zorg" and "Nationale routekaart databeschikbaarheid AI" were intended to delineate steps for AI in healthcare. From the reactions of respondents and the fact that even a commercial consultancy firm developed its own AI roadmap[43], it becomes clear that they did not provide sufficient centralized guidance to date (R2).

In pursuit of clarity regarding the regulatory direction in the market sphere, the Dutch healthcare law society's quarterly conference was attended. The event focused on the future of data regulation from various legal standpoints. Discussions highlighted the challenges in harmonizing health law across Europe, particularly in light of the subsidiarity principle, which underscores the significant role of national legislation. Concerns were also raised about compromising the doctor-patient relationship through data sharing, emphasizing patients' autonomy over their data[44]. The sector is noted for its deep fragmentation, with 45,000 healthcare providers, and faces multiple obstacles in adopting European data laws. Despite these challenges, the need is acknowledged to progress in anticipation of future hurdles and to undertake the necessary steps towards alignment. This sentiment is echoed in the Minister of Health's 2022 letter to parliament, which points out legal uncertainties from overlapping and vague regulations as a source of industry hesitancy. Such uncertainties lead to a trend of bespoke solutions to retain control, thereby causing significant fragmentation and affecting the scalability of innovations (Ministerie van Volksgezondheid Welzijn en Sport, 2022b). In summary, the existing regulatory, technical, and organizational constraints on AI (and LLMs) use, have set stakeholders' expectations towards more immediate, low-risk, non-medical, intra-organizational applications. Developers and users are calling for greater regulatory clarity—a resolution that appears distant at present (R2).

[40] More information on: https://www.zorgvoorinnoveren.nl/
[41] More information on: https://www.rvo.nl/
[42] More information on: https://www.linkedin.com/company/platform-eu-ai-act-voor-de-zorg/posts/
[43] https://www.ai-routekaart.nl/#start
[44] The subsidiary principle pertains to the scope of authority of European legislation in relation to the national legislation of the member states.

The complexity of compliance leads many to lack the means for full regulatory navigation. A 'checklist' has been proposed by several actors to streamline the process (M6, M7, AM2, AM3, AM4). The necessity of this is underscored by a statement highlighting the costly delays due to regulatory hurdles:

'I just literally see hundreds of entrepreneurs come by who just don't start right away with an Intended Purpose, don't start right away with Requirements, causing them to ultimately lose a year and a half of time. And that just says it all." (M7).

G-LLMTIS

Internationally, the past year has seen numerous appeals for regulation and oversight of AI, particularly Generative AI and Large Language Models (LLMs). Noteworthy events include an open letter from March 2023, signed by prominent figures including Elon Musk, urging a pause on AI advancements similar to ChatGPT-4 for at least six months[45]. Additionally, a global AI security summit was convened in the UK by Prime Minister Rishi Sunak in November 2023, advocating for "guardrails" in AI development[46]. That same month, France, Germany, and Italy agreed to endorse "mandatory self-regulation through codes of conduct" for foundational AI models, expressing reservations about the "un-tested norms" included in the AI-act[47]. These developments underscore the significant international interest in and momentum towards formulating regulatory frameworks for AI and LLMs, which will inevitably influence the regulatory landscape in the Netherlands.

F3 (Regulations and expectations) synthesis

In the formative stage of a TIS, emerging structures bring clarity to industry standards, regulations, and methods, alongside significant uncertainty (Markard, 2020). Hellsmark and Jacobsson (2009), and Musiolik et al. (2018) describe the post-formative stage as having higher institutional structuration, characterized by technology-specific entities like interoperability standards and safety regulations. Value chains solidify, and actors engage in formal and informal networks. Typically, the number of institutes decreases as convergence occurs.

However, such convergence is not yet apparent in the DH-LLMTIS, with ongoing structuration efforts facing alignment challenges, as indicated by data analysis and interviews. In the Netherlands, strict adherence to regulations like MDR and GDPR limits experimentation. Although European-wide harmonization suggests future structuration, it hasn't fully penetrated the field. Actors and regulators identify regulatory ambiguities, partly due to the nuances of DH-AITIS and DH-LLMTIS technologies. The AI-act adds uncertainty, with its implementation specifics still undefined and technological realities misaligned with policy proposals. Disagreements among European countries over the act's requirements, and the lack of a clear Dutch strategy for disseminating knowledge about the AI-act's effects, contribute to these uncertainties. Some actors are pursuing a bottom-up approach to clarify regulations, potentially leading to a dispersed knowledge diffusion pattern. However, as all main regulations (MDR, GDPR, AI-act) are in the implementation phase, structuration effects may lie ahead.

Leading healthcare law figures anticipate a challenging convergence path. Alkemade & Suurs (2012) associate 'hype-like' expectations with the formative phase. While some hype indications were found in knowledge formation (F2) and legitimacy, they were less evident in the assessment of (F3). Overstatements discussed in F4 market formation may also suggest hype-like expectations.

---

[45] https://futureoflife.org/open-letter/pause-giant-ai-experiments/
[46] https://www.aisafetysummit.gov.uk/
[47] https://www.reuters.com/technology/germany-france-italy-reach-agreement-future-ai-regulation-2023-11-18/

The DH-LLMTIS's current activities, aimed at reducing uncertainties through new regulations, indicate a transition towards a growth phase. However, the substantial lack of clarity aligns more with a nascent stage transitioning to a formative or emerging stage. These ambiguities underscore the complexity of the DH-LLMTIS's current developmental phase, balancing between emergent structuration and ongoing uncertainties.

5.2.4 F4 MARKET FORMATION

While the current market applications of AI in the Netherlands is small, business reports estimate an explosive growth in the worldwide market of AI in healthcare. Some see it growing from 11 billion in 2021 to 188 billion in 2030 (Statista, 2023) or 40 billion just for Generative AI (such as LLMs) in healthcare by 2031 is mentioned (Precedence Research, 2023). Though these figures are speculative, they indicate potential significant growth. Current market size in Dutch healthcare is unquantifiable in financial terms, but indicators suggest 111 AI applications are beyond the pilot phase, with 16% involving NLP (KPMG, 2020).

The entrepreneurial analysis confirms the presence of numerous early-stage start-ups in the Dutch AI landscape. However, they are generally not yet commercially sustainable and are in the initial stages of scaling up (G2, AM2, MG1, AM4). To date, there are no commercially deployed LLM applications in healthcare according to respondents (G1, G2, M5, M6, M7, AM2, AM3, AM4, MG1, AMG1), corresponding with international research that places LLMs mainly in the research phase (Yu et al., 2023). Nonetheless, trials using LLMs in hospitals have recently begun[48] (G2, M7, MG1, AM4), with the AI lead of an UMC reporting, *"I know of multiple hospitals participating in pilots... So it can be used with patients during real consultations."*

Identified barriers to market entry include prolonged development to implementation timelines and high initial costs, paralleling challenges in Healthcare Technology and Pharmaceutical sectors (DiMasi et al., 2016; Fogteloo, 2023). Regulatory compliance further complicates AI applications: *"The average time to register is 7 years... there is quite some grant money. And then there is nothing for a long time..."* (M7). UMCs' resource constraints also inhibit in-house business development: *"...an academic medical center does not yet have the technological expertise nor can bear the costs associated with bringing it to market..."* (AM2). Consequently, Dutch AI healthcare start-ups may even need to look beyond national borders for long-term viability, as the Dutch market alone may not offer sufficient cost savings potential (M7).

While UMCs are active in research, few projects evolve into commercial products. An accurate assessment of applications' maturity is necessary to measure commercial readiness. The perception of AI 'running' in hospitals often relates to research rather than clinical practice (AM4): *"Because I could also say that indeed we already use 15 AI models in practice, but those have been mainly tests with imaging tools . The definition of what is implemented can also differ per hospital...".*

Financial viability pivots on achieving significant cost savings (AM2, AM3, A4, M7), with the healthcare system's focus on cost reduction as a market entry precondition. Governmental policy instruments, like Germany's DiGa, could foster market entry for smaller entities by easing the process and creating shelter form full market pressure, yet no such mechanisms currently exist in the Netherlands, as interviews reveal (R2).

[48] More information on the news outing of one of the trials: https://www.skipr.nl/nieuws/etz-en-umcg-laten-ai-gestuurde-chatbot-medische-vragen-beantwoorden/

F4 synthesis

Bergek, Jacobsson, Carlsson, et al. (2008) and Markard (2020) note that in a formative phase, the established market is a small fraction of its potential future size. This is particularly true for the G-LLMTIS, where AI and LLMs in healthcare are expected to become a multi-billion dollar market by 2030, but the DH-AITIS currently has limited commercial applications.

Bento & Wilson (2016) propose that the end of the formative phase is indicated when about 10% of market potential is realized, usually determined ex post. While it's difficult to predict the total future market share, the 10% threshold has not beenr reached, given the nascent DH-LLMTIS market and modest commercial application levels in the DH-AITIS.

Bergek et al. (2008) observe that early-stage TIS customers are often scarce and experimental. They mention that the lack of a market at the start of the formative stage is common, and that institutional standards (F3) can catalyze market formation. Market actors note that unclear regulations and standards (F3) delay applications reaching the market, hindering entrepreneurial viability. This prolonged financial horizon reduces the attractiveness of investing in market formation, discussed in F5. Additionally, Hanson (2018) notes the absence of artificial market opportunities created by governments, typical in a transition from formative to growth phase.

These findings imply that F4 of the DH-AIMTIS, and by extension the DH-LLMTIS, is characteristic of an early-stage TIS not yet transitioning to a growth phase. Fulfillment of other functions is seen as key to further TIS development and improvement, as advocated by Bento & Wilson (2016) and Suurs & Hekkert (2009). The focus should therefore be on enhancing these other functional aspects to facilitate a future transition to a growth phase.

5.2.5 F5 RESOURCE MOBILIZATION

A document analysis was performed on the presence of funded projects on the this subject with a special interest in the type of actors who are contributing to these projects. Policy documents regarding mobilization of physical resources namely data mobilization, were analysed. Respondents were asked to reflect upon the presence and sufficiency of financial, human and physical resources. Again, separate resource mobilization regarding LLMs (and NLP) was sparsely found.

5.2.5.1 F5 HUMAN RESOURCES

Governance Sphere

In the governance of AI in healthcare, human resources are minimal and only recently developed. At VWS, one individual allocates a few hours weekly to AI in the medical device department. Another department focusing on data governance in healthcare dedicates one to three people to AI, while the digital transformation department doesn't have a dedicated AI team but *'includes AI in their work'* (G2). There is no organization-wide policy for AI knowledge diffusion, and employees involved in AI don't possess deep data science knowledge. The regulating agency IGJ has one person addressing software in medical devices, including AI, with a recent intent to slightly expand AI focus within the organization (G1).

Respondents didn't explicitly report a human resource shortage but noted challenges in keeping pace with new developments in AI (G1, G2). This view resonates with market actors: *"The Ministry of Health have been placing more people with substantive knowledge on the dossier, but these are people who have grown from a team that consisted of one person to a team that maybe consists of four or five people. But now they really have to think a lot about legal frameworks and legislation. And yes, if you try to legally cover everything with a team of five people following the wave, that's obviously a race you're never going to win."* (M6). G1 emphasized the need for a strategic approach in organizing AI expertise within the Dutch government, highlighting the scarcity of

AI experts: *"We are thinking about how to organize this in a future-proof way. Because if every regulator starts to recruit AI expertise in the market, then they are competing with each other. And there are simply not enough people to staff that. So maybe that needs to be done jointly."*

Market Sphere

In the market sphere, various roles and teams dedicated to AI have emerged in recent years (M5, M6, M7, AM4, MG1, AMG1). A notable development in scientific research departments within hospitals is the creation of the Chief Information Security Officer (CISO) role, specifically focusing on AI (M5). Respondents highlighted disparities in how organizations integrate and manage data science expertise. While there has been progress in hiring data scientists and improving data accessibility, there's still considerable inconsistency in the scale and centralization of these efforts across organizations (AM3, AM4, AMG1). *" No, the level of coordination changes, the level of knowledge varies. Yes, and maybe not much has changed or improved in the past year. You definitely see that more people are working with data, the liberation of data, and you also see more data scientists in hospitals, yes."* (M7). This shift from individual interest to formalized teams is also observed within EHR providers (M7). Predominantly, UMCs and larger TKZs have small but growing data science/AI teams (AMG1, AM4). One respondent noted a team's size change since 2017 due to fluctuating interest in AI, leading the hospital to prioritize other projects (AM4).

In conclusion, dedicated AI roles are becoming established in larger healthcare providers in the Netherlands. However, LLMs and NLP have not yet emerged as distinct functions or areas of focus.

Academic Sphere

Growth of human resources in Academia reserved for AI educational programs happened in a more organic way *"At some point, they really had to expand with more teachers and there was a huge student-staff ratio."* (AG1). The resource mobilization of Academia takes place mostly in the form of increased human resources.

5.2.5.2 F5 FINANCIAL RESOURCES

Governance Sphere

Since the SAPAI program several examples are present of financial resources being mobilized for the development of AI in all sectors. Below is an overview of key entities contributing to AI or AI in healthcare funding:

1. AINED Consortium: Originally allocated a sum of 1 billion euros, the AINED consortium's funding was adjusted in 2021 to 274 million euros over seven years from the Dutch National Growth Fund (Claudio Lazo et al., 2023). This funding supports various sectors represented in the Dutch AI Coalition, including healthcare.
2. ROBUST Program: In 2023, the ROBUST program, which focuses on AI solutions for societal issues, secured 87 million euros from the Ministry of Economic Affairs and the NWO (AG1, MG1).
3. NWO: As the Netherlands' leading academic research funder, the NWO doesn't have set quotas for AI and LLM research but is a prominent funder in this field, as indicated by Scopus web search results. Funding is provided based on 'organic research progress directions' rather than specific agenda setting, as noted by an NWO board member (AG1).
4. Health Holland: Recognized as one of the nine 'top sectors' by the Ministry of Economic Affairs and Climate, the Life Sciences & Health sector, under Health Holland, prioritizes funding Public-Private Partnerships (PPPs) and research. The PPS arrangement sees the Dutch government adding 30% to private investments, aligning with the Top Sector's agenda[49]. Digital transformation and AI were

[49] https://www.rvo.nl/subsidies-financiering/pps-toeslag-onderzoek-en-innovatie

key strategic themes for 2020-2023, with over 79 million euros in subsidies awarded in 2022 alone[50].

5. RVO and ROMs: The Dutch government's entrepreneurial agency RVO and regional development agencies (ROMs) are also noted as financial contributors. While their official strategy doesn't specifically target AI in healthcare, funding organically flows to AITIS and LLMTIS through various programs available to innovators (M3, MG1).
6. The number 1 financing affiliation of Dutch LLM and NLP healthcare articles on Scopus, was the European Horizon program[51]. A Cordis database search on AI in the domain of application health showed 84 multinational projects which were manually scanned on their association with AI and healthcare. The total funding to Dutch affiliations of these projects amounted to over 40 million euros since 2014.

Market Sphere

As highlighted in the discussion on Market Formation, the extended time-to-market presents significant financing challenges for AI start-ups in healthcare (M7, MG1, AM2). A 2021 study supports this, identifying room for improvement in the start-up climate for the Dutch healthcare sector. It notes that the complexity of the Dutch healthcare system and lengthy innovation cycles hinder the financial sustainability of start-ups (Techleap, 2021).

Hospitals allocate financial resources primarily through human capital but face constraints in setting aside specific budgets for AI projects. Additionally, hospitals are reluctant to bear the costs associated with medical technology entrepreneurship (AM2, M3). AI applications, in particular, are required to demonstrate cost savings to justify investments (M5, M6, AM2, AM3, AM4). This need for cost-efficiency echoes the broader issue of tight budgets in the healthcare sector, as mentioned in the introduction.

Academic Sphere

As the financial resources of the academic sphere overlap with the financial resources that are mobilized in the governance sphere, this category does not come forward as a significant functional dynamic.

F5 (human and financial) synthesis

The DH-LLMTIS's reliance on resource mobilization from the DH-AITIS is typical of a formative stage, as Markard (2020) notes. Bento & Wilson (2016) and Markard (2020) assert that significant mobilization of financial and human resources characterizes the transition from the formative to the growth phase. However, the literature provides limited specifics on the types of resources needed during the formative stage. Bergek, Jacobsson, Carlsson, et al. (2008) suggest that resource mobilization should stimulate not only the entry of firms but also the pursuit of varied ventures, implying a need for diverse resource types.

Financial resource mobilization in the DH-AITIS governance sphere has been notable, with initiatives like the AINED fund supporting knowledge development & diffusion (F3) and stimulating entrepreneurial activities (F1). This aligns with Markard's (2020) indication of financial stimulation in the formative phase. However, market actors in hospitals report a lack of significant financial resources for AI and LLMs in healthcare, with unsustainable cost increases being a critical issue. While academic knowledge development dynamics are consistent with earlier TIS phases, their additional resource mobilization is not a focus.

[50] https://www.health-holland.com/sites/default/files/downloads/Jaarraportage%202022%20LSHTKI-Definitief%20met%20persoonlijke%20handtekening%2027380989%202022.pdf

[51] Horizon 2020 (between 2014 – 2020) and Horizon Europe (between 2020 – 2027) are the most important research financing programs by the European Union

Human resource mobilization in governance appears insufficient. Instances of minimal staffing, like one person overseeing the AI strategy for the health regulator and one person part-time on AI application with MDR standards, suggest challenges in forming and communicating regulations and standards (F3). Despite this, there's been an increase in AI-dedicated personnel in hospitals and EHR providers forming AI teams. The ongoing clarity issues in F3 requires more dedicated personnel for aligning the TIS with hospital requirements.

The varied state of financial and human resource mobilization complicates pinpointing the exact TIS stage. Particularly, the government's human resource mobilization is underdeveloped. This indicates that comprehensive resource mobilization is lacking, suggesting that the formative phase is ongoing. The lack of mobilization in governmental human resources implies that the DH-LLMTIS needs further development and structuration for transitioning to the growth phase.

5.2.5.3 F5 PHYSICAL RESOURCES

The analysis of physical resources necessary for AI implementation in healthcare, including LLMs, reveals two key areas: computing infrastructure and data mobilization (R1, R2).

Computing infrastructure

The structural analysis indicated that hospitals are balancing their focus between on-premise computing solutions and cloud-based solutions, with no clear preference emerging (A3, AM2, AM3, AMG1). The current availability of computing power and data storage is not seen as a major barrier, aligning with Capgemini Invent's 2020 report on data availability in Dutch healthcare. Nonetheless, the development of a more centralized data infrastructure policy is considered beneficial for AI implementation in hospitals (AM2, AM3, AM4). Respondent AM3 highlighted the potential for synergistic effects from centralized data policies within hospitals, marking it as a key area for improvement. Overall, the data infrastructure was not singled out as a major obstacle in the interviews (R1, R2). Computing power and storage capacity are deemed sufficient in the short term to meet the needs of Dutch healthcare providers (AM2, AM3, AM4, M4).

Data Mobilization

Numerous interviews steered towards the key gateway facilitator for AI in general and LLMs/NLP that is data mobilization (R1, R2). Data mobilization, encompassing storage standardization, extraction, anonymization, and sharing, is essential for AI and LLM model development in healthcare. This complex process involves various projects and initiatives in the Netherlands and Europe, aimed at enhancing data availability in healthcare. Key examples are detailed below to highlight the activities and structures around data mobilization.

1. Government-Funded Agencies:
    a. Health-RI: A research foundation funded by the Dutch National Growth Fund, initiated in 2021 to improve the Dutch healthcare data infrastructure. It collaborates with VWS and EZK[52].
    b. Nictiz: Active since 2002 and commissioned by VWS, Nictiz focuses on enhancing digital technology usage in Dutch healthcare. It develops and manages standards for secure and efficient healthcare information exchange.
2. WEGIZ Law: The development and process of the WEGIZ law in the Netherlands, formally known as the 'Wet elektronische gegevensuitwisseling in de zorg', has as goal to enhance health data interoperability within the Dutch healthcare system. The law, which was introduced in 2021 and unanimously ratified by parliament in 2023, requires healthcare professionals to establish

[52] More information on Health-RI: https://www.health-ri.nl/

treatment standards that, once adopted by the a National Quality Register and gaining legal status, become enforceable harmonized norms. To ensure a more seamless exchange of health data, the WEGIZ law aims to bring healthcare organizations and professionals onto a common platform by adopting shared standards and data models. The impact of WEGIZ extends beyond the Netherlands as it aligns with the upcoming EU regulations on medical data interoperability, such as the European Health Data Space (EHDS), enabling data sharing across the European Union.

3 The European Health Data Space (EHDS): The EHDS is a regulatory framework aiming to enhance the access, exchange, and reuse of health data across the EU and was named by multiple respondents as influential on mobilization of data as resource (M3, M6, AG1, G1, G2) Key decisions include establishing interoperability of electronic health record systems, reinforcing individuals' control over personal health data, and creating platforms for cross-border data exchange. The European Council's position statement, released in December 2023, states the need to clarify the regulation's scope, aligns it with GDPR, and proposes governance structures with member-state oversight. Negotiations with the European Parliament currently take place about the exact implications of this position statement. However, the member-states did unanimously agree on the execution of the EDHS (European Council, 2023).

4 Cumuluz initiative: The Cumuluz project is an example of more practical execution of data mobilization. It is a data mobilization project coordinated by all Dutch UMCS, explicitly aligning with the WEGIZ law and striving to establish an overarching data hub. This is a target architecture for existing initiatives to converge towards, with privacy and security said to be effectively safeguarded. Implementation was estimated to take around 5 years at time of announcement in 2022[53]

These initiatives illustrate concerted efforts to enhance data mobilization in Dutch and European healthcare, vital for AI and LLMTISs development. However, actors encounter diverse data management standards, varying by country, region, or even hospital (AM3, M6), with EHR providers noting the personalisation options of their systems leading to different operational modes in hospitals (M6).

Actors need not only regulatory but also data standards (AM2, AMG1, M5, M6), as M6 emphasizes: *"So the hope is very real that from a European level, let's say, we are going to close the discussion once and for all, these will be the standards that we will adopt within Europe. Well then, we, like all other suppliers, will cheer, saying great, we can focus on that and then we can ensure that the data can be made available in the right way for each relevant purpose."*

In conclusion, while the list above captures major and current developments impacting the Netherlands, it's not exhaustive. Cap Gemini Invent's 2020 report identified 23 data mobilization projects. The multitude of projects, coupled with respondent insights, paints a picture of a fragmented landscape. Yet, initiatives like WEGIZ law and EHDS represent a move towards centralization, potentially harmonizing the field (M1, M2).

F5 (physical) synthesis

In the DH-LLMTIS context, resource mobilization for the DH-AITIS is crucial, as highlighted by Markard (2020a). Bento & Wilson (2016) and Markard (2020) suggest that distinct mobilization of physical complementary resources is a key indicator of the transition from the formative to the growth phase.

[53] More information on Cumuluz: https://www.cumuluz.org/ and https://www.nfu.nl/themas/data-infrastructuur/cumuluz-zorgplatform

The results indicate limited increases in computing power mobilization for AI and LLMs in healthcare. While centralizing computing capacity in hospitals is seen as beneficial for AI implementation, there's no strong indication that a lack of this infrastructure is seen as a major hampering factor moving forward.

In contrast, physical resource mobilization, especially data mobilization, shows dynamics akin to those in functions F2 and F3. The Dutch data mobilization landscape is active but fragmented. European harmonization efforts, like the EHDR and F3 regulations, are steering towards a more structured field. Initiatives such as the Dutch WEGIZ law, complementing the EHDS, and the UMCs' data sharing project Cumuluz, in line with WEGIZ, hint at early field convergence. In AI and LLM fields, data is a critical resource intertwined with policy regulations and industry standards, making its mobilization complex.

Considering the data standards F3, the movement towards convergence suggests an advanced formative stage moving towards growth, yet not fully achieved. Thus, physical resource mobilization aligns with a formative stage edging towards growth, contingent on successful data mobilization efforts. A clearer trajectory towards structured and cohesive resource utilization could potentially mitigate the system's transition into a growth phase.

5.2.6 F6 LEGITIMACY

In order to analyse the legitimacy of AI and more specifically LLMs and NLP use in Dutch healthcare, a news-outlet sentiment analysis was combined with qualitative assessment by respondents and insights of other function associated with legitimacy namely knowledge creation & diffusion and guidance of the search. First the findings from the sentiment analysis are presented whereafter the most important insights from the interviews are summed-up. Lastly, the differences between the pictures that emerge from both analyses are briefly discuss.

Sentiment Analysis

Of the 30 articles, 13 articles were generally positive, 13 articles were neutral with critical notes and 4 articles were generally negative regarding the use of AI in healthcare. Of the 30 articles, only 2 did not mention ChatGPT as seen in **Figure 11**. In contrast to the academic knowledge creation, ChatGPT is mentioned in almost all articles. This indicates how the reporting is narrowed around publicly known subjects and is not representable of all developments in the field.

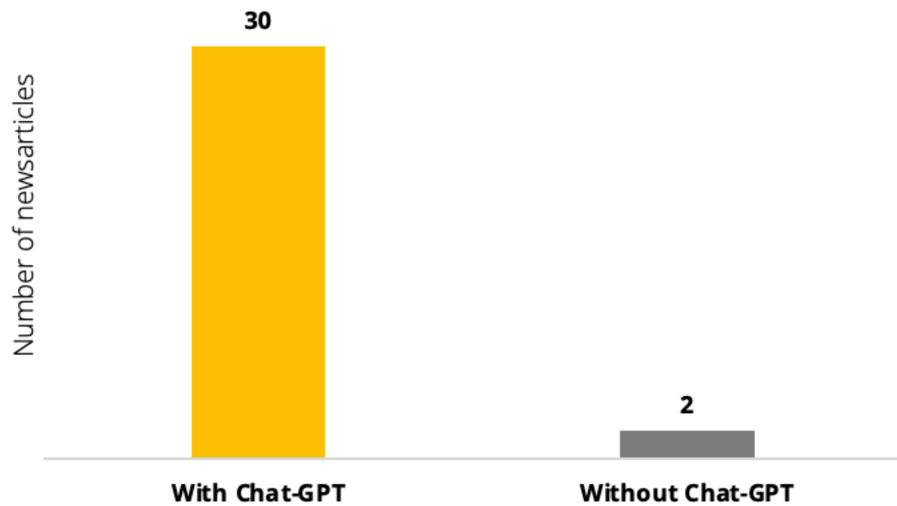


Figure 11: Dutch news articles LLM and NLP in healthcare since November 2022 with and without mentioning ChatGPT in title, abstract and keywords.

In general, the picture emerges that sees 'old' AI being a legitimate technology, while generative AI and subsequently LLMs have to be treated with caution and as a technology doctors would not want to use in the near future. Interestingly, when medical specialists are quoted, they come across as being more positive towards the use of AI than the author of the news article. Several reoccurring shortcomings of current LLMs and NLP that were named were: risk of bias, insufficient data availability and privacy, insufficient performance, missing regulation. These general conclusions are in line with the line of thought that has been mentioned throughout this report. The articles brought forward a division between the data scientist and doctor as a hard to bridge gap, indicating that doctors have insufficient knowledge on AI in general. The insufficiency of AI specific regulation is often mentioned, putting an emphasis on the fact that regulation should be strict and keep AI in check.

Interviews

The most important observations that emerged from the interviews regarding the perceived legitimacy, are presented below.

1. Impact of ChatGPT: ChatGPT has notably disrupted the general legitimacy of AI, making it a buzzword and sparking interest among various actors since late 2022 (AG, AM2, MG1, M6, AM). AG remarks, *"I think it just starts with a lot more awareness... there's a mix of high expectations and skepticism about privacy and data issues, and ChatGPT is hallucinating..."* Despite this increased awareness, ChatGPT and generative AI's readiness for clinical use remains contested (Meskó & Topol, 2023; Sallam, 2023; Wornow et al., 2023).
2. Regulatory Compliance: Uncertainty about LLMs/NLP's compliance with existing regulations is a current issue, causing hesitancy among stakeholders. The main concerns are data management within vs. outside organizations (with a preference for internal data use) and classification of applications as medical devices or otherwise. Non-medical device classification allows easier implementation but may lack long-term legitimacy (AM2, AM3, AM4, M5).
3. Reconfiguration vs. Generation: Systems that reconfigure existing knowledge rather than generating new medical insights are seen as more realistic and legitimate in the short term, aligning

with current literature that underscores the complexities of generative models in medicine (G2, M7, AM2, AM3).

4. Institutional Adoption: AI's widespread acknowledgment as a future healthcare technology is evident. The focus is shifting from whether to include AI in healthcare to how it should be integrated into policy (M7, G1).
5. Cost-Effectiveness: Financial viability is crucial, with AI applications expected to offer cost savings. This economic rationale boosts AI's legitimacy as a solution to pressing healthcare issues (M6, M7). The EHR manager notes, "*With the introduction of AI, there's a lot of demand due to its potential to improve efficiency, care quality, and address staff shortages.*"
6. Cyclical Interest: Characteristic for the AI field are several upswings and downfalls in interest that have taken place over the years that have raised or lowered the legitimacy of the field. The hype-like dynamics of the expectations of actors are congruent with this (M6, M7, AG, AM4).
7. ELSA Labs: The Netherlands has established ELSA labs to foster 'safe' AI development, contributing to the technology's legitimacy (MG1).

In summary, news articles were more critical than interview respondents, possibly reflecting the technical focus of those working in AI and LLMs in healthcare. While data mobilization issues, regulatory uncertainty, and ChatGPT's impact were common themes, respondents approached these topics from a practical standpoint, based on their professional experiences. News articles emphasized the divide between healthcare providers and data scientists more than the respondents did. In a sector that values proven results and seeks solutions to escalating costs, non-medical LLMs and AI applications in image recognition are currently viewed as more legitimate short-term options. ChatGPT's emergence has solidified AI's place in the healthcare landscape.

F6 synthesis

The process of aligning with contextual TISs and societal problems for legitimacy, as outlined by Markard (2020), is indicative of a formative stage. AI's advocacy in healthcare, addressing escalating costs, is a dynamic also applicable to LLMs, demonstrating their potential for societal legitimacy. Unlike other functions where the DH-LLMTIS is influenced by the broader DH-AITIS, the relationship here appears bidirectional, possibly due to the global legitimacy gained by LLMs, especially with successful G-LLMTIS applications like ChatGPT.

Bergek et al. (2008) describe legitimization in the formative stage as driven by 'expert' and 'rational' legitimacy, rather than 'familiarity' of usage, typical of later stages. However, the DH-LLMTIS seems to gain significant bottom-up legitimacy from personal application use like ChatGPT that can lead to familiarity of usage.

Bento & Wilson (2016) and Bergek et al. (2008) observe that increased legitimacy aligns with legislation and industry standards (F3), signaling the end of the formative phase and the onset of growth. However, such alignment in the DH-LLMTIS is ongoing, suggesting a delay in legitimization. Despite this, large incumbent institutions' engagement with the focal TIS could ensure future legitimacy through alignments.

Legitimacy in the DH-LLMTIS can be assessed along three axes: more cost savings implies greater legitimacy, increased intraorganizational data use correlates with legitimacy, and fewer medical device properties in the model enhances legitimacy.

In summary, the DH-LLMTIS's legitimacy showcases unique dynamics, deviating from traditional TIS analysis expectations of competitive dynamics between TISs. While this indicates a departure from the norm, the overall analysis confirms that the DH-LLMTIS is still in its formative stage. Advancing beyond this stage depends on aligning with regulatory frameworks, a step that is essential for transitioning to the growth phase.

5.2.7 F7 SYNGERIES

Currently, LLMs are in an experimental phase of development, which has not yet led to the establishment of specific synergies within the DH-LLMTIS. However, as previously detailed, various elements from the DH-AITIS are being leveraged to facilitate the development of DH-LLMTIS.

F7 synthesis

Bergek et al. (2008) highlight that the entry of new entrepreneurs into a TIS can improve functional dynamics and create positive externalities, serving as both a driver and an indicator of the strength of the other six TIS functions.

Markard (2020), however, notes that in the formative phase, specialized suppliers are typically absent, with intermediary actors and 'technology-specific' associations emerging as the TIS moves towards a growth phase. The data shows increased activity within the DH-LLMTIS and DH-AITIS, but a significant rise in new entrepreneurial actors is not evident, aligning with expectations of an early formative stage. Nonetheless, the development of intermediary actors within the DH-AITIS suggests a progression towards growth phase characteristics. For instance, companies facilitating AI implementation in healthcare, such as those involved in medical data extraction or regulatory compliance, indicate this shift. Currently, the presence of limited synergic effects supports the assessment of the DH-LLMTIS as being in an early formative stage, with certain elements signaling the beginning of a growth phase.

6 DISCUSSION

6.1 DISCUSSION OF FINDINGS

6.1.1 STRUCTURAL ANALYSIS

The division was kept between the 4 different types of structures that was sufficient for a comprehensive structural analysis. The structural analysis aids in answering RQ2, in which the division of three 'spheres' being academic, market and governance as posed by Larisch et al. (2016) was confirmed as a applicable framework to structure the search process in a sector that is known for its large amount of actors. This division combined with the socio-technical structural TIS approach as discussed by Andersson et al. (2023) did not result in significant new groups that could not be placed in one of the spheres. The DH-LLMTIS certainly involved technical elements in the form of the described infrastructure of computing and data management in hospitals, that rendered a pure social structural angle sub-optimal.

6.1.2 FUNCTIONAL ANALYSIS

The aggregation of the functions of multiple earlier TIS analyses proved a good starting point for analysing the TIS functions and analysis aided in answering RQ3. However during the practical executing, distinct intra-functional differences were found in dynamics of functions as they were chosen. This led to the segmentation of knowledge development and diffusion, which can be called 2A and 2B, a division that has been made often in TIS analyses and should be kept in the emergent context. The difference between creating knowledge through for example academic literature and diffusing it among actors showed unique characteristics. This was also the case for guidance through policy measures and guidance through regulations and different types of resource mobilization which we can call 3A and 3B. Prior literature does not make a clear difference between the mobilization of financial and human resource and physical resources. Perhaps this shows the particularity of the 'AI' context and the healthcare context. As described, data is a complex resource that is subject to regulations and development as well. Therefore this function can be divided into function 5A (financial and human resource mobilization) and 5B (Physical resource mobilization) Future analyses should consider these distinctions, especially in sectors where the complexity of resources is heightened by their regulatory and developmental contexts.

The synthesis of all the functional performances based on the synthesis of the functional analysis in the results section is presented in **Figure 12**. The figure shows how for most functionalities the LLMTIS expresses characteristics of a formative phase that is yet to transition into a growing phase. The insights of previous literature on early-stage TISs served as a yard-stick to determine the positioning of the different colors. The darkest color indicates the predominant overal position of functional development. The lighter colors present a gradual scale to which extent some aspects of the function express chacteristics. For example, knowledge difussion clearly has progressed through the nascent phase as significant difussion networks apart from R&D exists. However, no movement towards convergence of networks currently is observed, which is an essential marker for the transition to a growth phase. In the developments of syngeric effects, no clear influx of actors is observed, aligning with an early formative phase. However the existence of specialized intermediary actors shows some characteristics of a growing phase indicated by the light spread going into the growing phase in the figure.

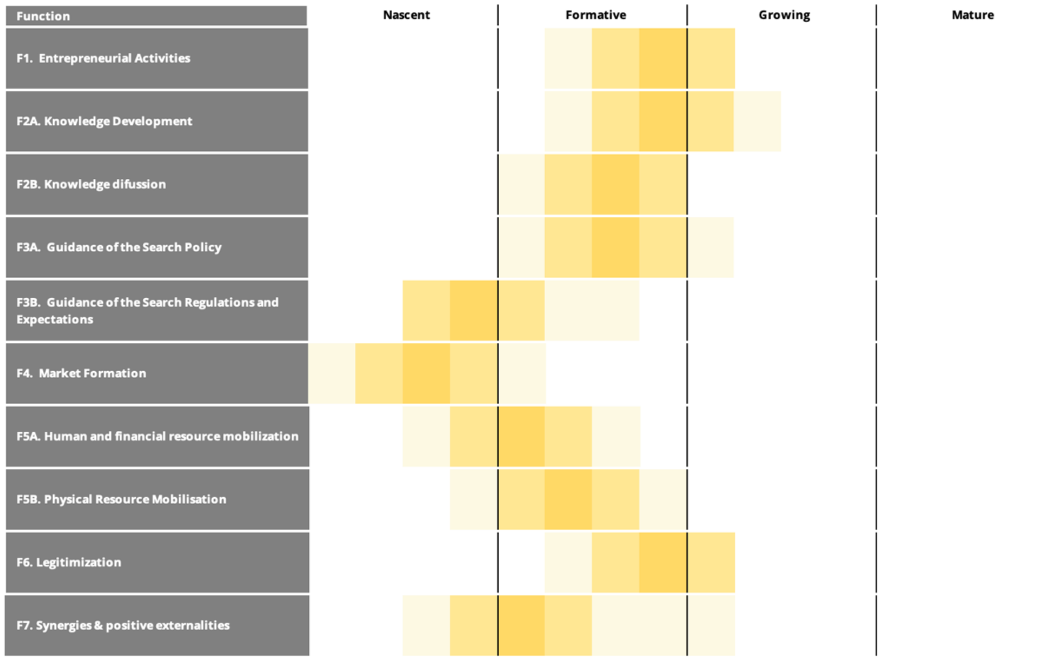


Figure 12: Indication of functional performances stages

6.1.3 TIS-TIS INTERDEPENDENCIES

In order to shed light upon the exact delineation of the DH-AITIS, G-LLMTIS and DH-LLMTIS the interaction between the different TIS are discussed, aiding the answering of RQ1.

The findings indicate that DH-Health LLMTIS cannot be distinctly classified as a separate TIS within DH-AITIS across all functionalities, as initially suggested in the old visualisation of **Figure 13a**. **Figure 13b** highlights functions in yellow (F1, F2, F6), where a distinct LLMTIS dynamic can be discerned within the Dutch AITIS framework.

In the realm of knowledge development (F2), a specific body of research on LLMs in healthcare has emerged, showing significant growth in recent years. In this function a distinction was between knowledge F2A and F2B. In terms of knowledge diffusion, no unique DH-LLMTIS dynamics were observed; instead, they are reliant on the diffusion networks of DH-AITIS. The functionalities of guiding the search (F3), forming the market (F4), and mobilizing resources (F5) exhibit complete reliance on DH-AITIS. Within the function of guiding the search, we observed notable variances between policy-driven guidance and regulatory and expectation-driven guidance, termed (F3A) and (F3B), respectively. Similarly, within resource mobilization, we distinguish between financial and human resources (F5A) and physical resource mobilization (F5B). The legitimacy function (F6) is notable, as DH-LLMTIS exists independently and interacts with DH-AITIS. No synergistic effects were identified within DH-LLMTIS; rather, they are entirely dependent on DH-AITIS (F7).

The Global LLMTIS overlaps with the Dutch context in four distinct functions, as illustrated in the figure. In (F1), global companies like Microsoft and Epic, while not dominating the Dutch market, influence the Dutch entrepreneurial landscape. In (F2A), international advancements in LLMs enhance Dutch knowledge development. Regarding knowledge diffusion, international healthcare conferences and online knowledge availability contribute to (F2B) in the Dutch context. The regulation and expectation-driven guidance (F3B) is significantly influenced by global calls to regulate AI. This is for example seen in the rapid development of the EU AI-act. Lastly, the legitimacy of the DH-LLMTIS is heavily influenced by developments in the G-

LLMTIS. News of breakthroughs in the application or technological abilities of healthcare LLMs in the global perspective will be able to directly sway the legitimacy of the DH-LLMTIS such as happened with the introduction of ChatGPT.

Figures 13a and b show the positioning of the early-stage DH-LLMTIS in the integrated framework and post-functional TIS analysis. In the revised framework, the distinct position of DH-LLMTIS shifts to a pronounced overlap with DH-AITIS. The functions highlighted in yellow are where DH-LLMTIS demonstrates unique dynamics. The overlap with G-LLMTIS is evident in functions 1, 2, 3, and 6. This analysis confirms that the overlap of DH-LLMTIS with DH-AITIS aligns with Hanson's (2018) contextual positioning.

It is clear that both contextual TISs serve as a breeding ground for the DH-LLMTIS and that ‘flow’ is present in four of the seven functions. Hanson differentiates the exploratory state from the formation phase, marked by the emergence of system-specific structures, including specialized knowledge and networks, entrepreneurship and legitimacy. While specialized entrepreneurship exists in a limited amount, the anticipated specialized networks are absent, suggesting a more nascent or exploratory phase. Bergek et al. (2015a) and subsequent studies by Bergek, (2019) and Markard (2020) on contextual TISs support this interpretation, noting in the 'early stages of development' a strong dependency on contextual TISs in similar technological or sectoral areas, referred to as 'structural couplings'. They pose that this balance shifts towards the focal TIS's independence in more advanced stages, moving from a formative to a growth phase. However, the data does not indicate such stages of independence.

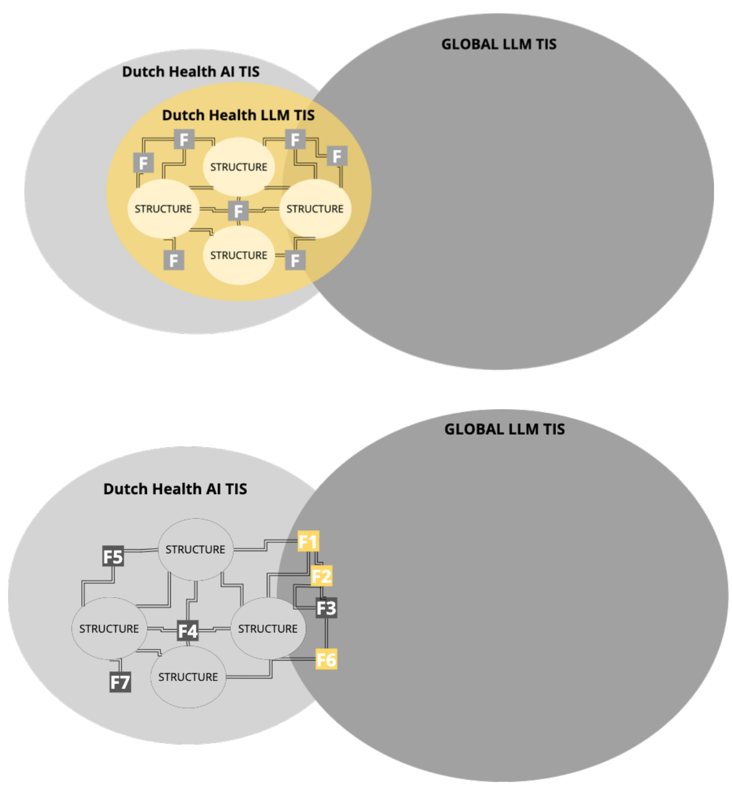


Figure 13a: Old integrated framework **12b**: new integrated framework

6.1.4 CUMULATIVE CAUSATIONS

In order to further analyse the processes behind the current state of the functional performances the cumulative causations (CC) are discussed below.

Bergek, Jacobsson, Carlsson, et al. (2008) refer to these CCs as 'inducement' and 'blocking' mechanisms, Apell & Eriksson (2023) as functional patterns, and Hekkert et al. (2007) and Hekkert & Negro (2009) describe them as virtuous and vicious cycles. Suurs & Hekkert (2009) provide an extensive description of recurring CCs, identifying two – the science technology push (STP) and entrepreneurial system motor (ES) – as characteristic of the early stages of a TIS. These 'motors of innovation' serve as analytical tools to examine the crucial interactions within DH-LLMTIS that could foster synergistic effects and potentially catalyze a transition to a more advanced 'growth' stage.

The first CC possibly present is the STP, which starts with positive research outcomes, leading to the establishment of government R&D programs (evidenced as guidance of the search through resource mobilization). This phenomenon is observed in the formation of a Dutch AI strategy (F3A) and the subsequent stimulation of R&D (F5A) in response to technological advancements in the field (F2AB). Such initiatives increase knowledge development and diffusion (F2AB), as seen in the increased knowledge development and the emergence of diffusion networks. This progression, in turn, leads to further guidance of the search (F3AB), evident in efforts towards greater regulatory alignment and policy formalization. A second cycle sees guidance (F3A) initiate resource mobilization (F5A), which then stimulates entrepreneurial activities (F1), as exemplified by the stimulation of academic and market collaboration through programs like AINED. The entrepreneurial activities are expected to enhance the guidance of the search (F3B), as observed through the calls from market actors for more structured and a clearer field of regulations.

In the STP, two potential areas for improvement in the DH-LLMTIS's 'motor' are identified. I. The first is the diffusion of knowledge through the convergence of existing diffusion networks. II. The second is the enhanced guidance of the search through more aligned and structured regulations. The interactions within this motor are depicted in **Figure 14**, highlighting the DH-LLMTIS functions that demonstrate individual development, marked in yellow.

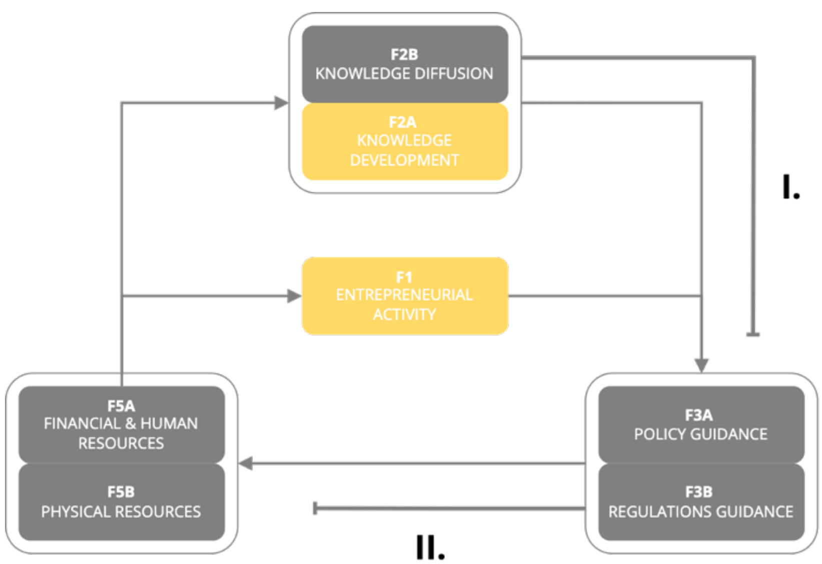


Figure 14: The STP motor of the DH-LLMTIS with possible improvements visible at I and II

Sequential to the STP motor, the entrepreneurial system motor (ES) often emerges. In the ES, entrepreneurs initiate innovative projects, a trend that is slowly becoming evident in the DH-LLMTIS (F1), driven by the perception of a future commercial viability shaped by the guidance of the search (F3AB). Given that most projects in this stage lack commercial viability, there is an increased need for resources in the ES. Suurs & Hekkert (2009) describe this phase as a period of intensified lobbying by entrepreneurial activities (F1) for increased governmental resources. This lobbying is expected to improve the initiation of more projects (F1), thereby increasing knowledge development and diffusion (F2AB), as well as influencing policy measures and expectations. Concerning this motor, one critical area for enhancement is identified: III. the current state of resource mobilization, particularly (F5B), which pertains to data mobilization. This aspect is perceived as insufficient and could benefit from more effective guidance of the search (F3). The present situation halts the commercial prospects, consequently leading to a decrease in entrepreneurial activities. The dynamics of this motor and its interactions are illustrated in **Figure 15**, highlighting the areas for potential improvement and development within the DH-LLMTIS.

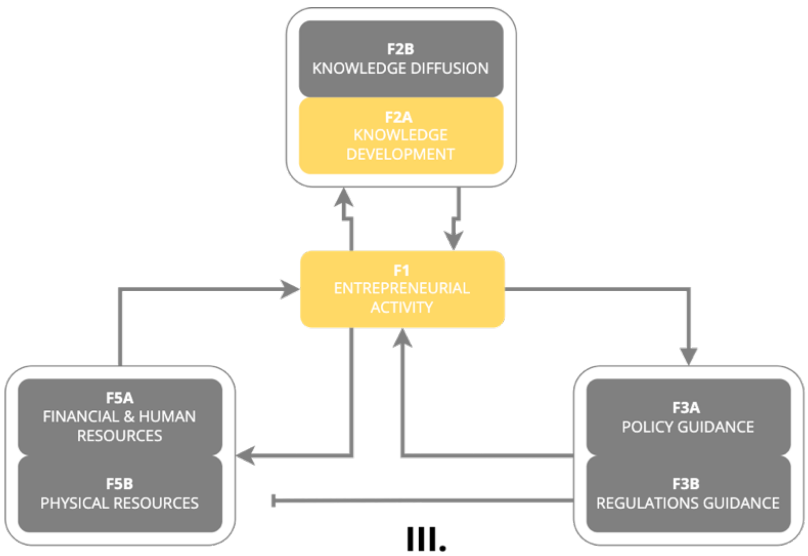


Figure 15: The ES motor of the DH-LLMTIS with possible improvement visible at III

Regarding this representation of the TIS dynamics, two observations can be made: 1. It is important to note that these motors do not engage with market formation processes until the later stages. This observation aligns with the findings of previous literature and suggests that the focus of TIS development should not be centered on market formation at this stage. The implication here is that unjustified emphasis on market formation could potentially misdirect resources and attention away from more crucial developmental aspects of the TIS at early stages. 2. These motors do not explicitly incorporate the influence of legitimization. This is because Suurs, in their framework, did not recognize legitimization as a distinct function, opting instead to focus on 'support by advocacy groups'. However, we can position the influence of legitimacy as a potentially positive factor impacting knowledge development (2A), knowledge diffusion (2B), and guidance of the search (F3AB). As highlighted by Bento & Wilson (2016), the roles of legitimization and market formation become increasingly significant towards the end of the formative stage and as the transition to a growth stage begins.

6.2 THEORETICAL CONTRIBUTIONS & FUTURE RESEARCH

The theoretical contributions of this research are founded upon a methodical reconsideration of the TIS framework's application within a healthcare context, concerning AI and LLMs. Revisiting the performed methodological steps and choices offers insights into the framework's utility. This in turn shows where theoretical contributions are made and avenues for future research could be pursued.

6.2.1 TIS ANALYTICAL FRAMEWORK

The first and important theoretical step is that of explicit delineation and context description of the focal TIS. In earlier TIS analyses, justification of the delineation is given variably, often no clear justification of the focal TIS is given, a common critique on TIS analyses (Markard et al., 2015). In recent work by Apell & Eriksson (2023) a TIS analysis was performed in the healthcare sector on the subject of AI. The societal implications of the usage of the Technology were not made clear, the geographical scope of West-Sweden was defined by 'researchers access to data' and the fact that 25% of the country lives in West-Sweden. Furthermore, no justification was given for the used division of technological application areas and maturity levels. In this research, a more in-depth justification is given for the usage of a TIS analysis in a national healthcare setting, opening up avenues for additional TIS analyses in this sector.

Additionally, Andersson et al. (2023) recently spoke of the deliberate choices that should be made when talking about a TIS, either being the technological innovation, a production system or both. This research reveals that an initial technological perspective on the application of LLMs in healthcare naturally widened during the analyis to incorporate production system factors, such as data quality and origin—elements that are integral to the technology's operational context.

6.2.2 BEYOND THE SUSTAINIBILITY CONTEXT

This research by no means overlooks the importance of the sustainability context. However, it also seeks to explore the broader applicability of the TIS framework, which extends beyond the realm of sustainability.

In several aspects the underlying assumptions focused on sustainability came forward. When guidance of the search policy was mentioned in literature, it emphasized examples that created an artificial competitive advantages for market actors compared to say 'incumbent energy suppliers', through instruments like subsidies. This could not be expected in case of the DH-LLMTIS and connects to the next point that is resource mobilization. In earlier TIS research, resource mobilization is often visualized as excess financial funds that can be used for fasten for example 'the energy transition'. In the DH-LLMTIS the healthcare sector had absolutely no excess funds to spend on technological innovations. On the contrary, the sector needs to reduce costs.

Additionally, legitimacy in TIS research often encompasses the competitive dynamics that arise when new technologies challenge established ones, a scenario common in energy markets. However, in the healthcare sector the dynamics may differ due to its inherent 'public good' nature. While there may be concerns about technology, such as AI, replacing certain tasks traditionally performed by humans, such as doctors, the overall sentiment leans towards pragmatic acceptance: if the technology proves to be effective and enhances healthcare delivery, it will be embraced.

Future TIS analyses within the healthcare sector should take these nuances into account, particularly the absence of excess funds for innovation and the unique aspects of legitimacy that differ from sectors like energy. The TIS framework's application should be sensitive to the sector-specific dynamics and the varying roles that different functions—such as policy guidance, resource mobilization, and legitimacy—play in the development of the system.

6.2.3. THE EARLY-STAGE TIS

In this research, various insights on the characteristics and dynamics of early stage TISs from previous literature were combined to categorize the developments in the focal TIS and better defined pressure points that ‘lag behind’ in stage development. The next question that emerges from this research is how to better categorize and delineate these stages?

Technological innovation is inherently gradual and does not conform neatly to discrete stages. Bento & Wilson (2016) attempted to establish quantitative benchmarks, such as achieving 10% of market potential, to delineate different stages of TIS development. However, the complexity and variability inherent across different functions and sectors render such one-size-fits-all measures insufficient for universal application such as this research shows.

Future case studies, especially those in sectors and technologies comparable to the ones examined in this research, can collectively highlight pressure points that signify transitions between stages, drawing on the conceptual framework of 'Motors of innovation' by Suurs & Hekkert (2009). Moreover, some TIS studies have employed ordinal scales to depict the functioning of a TIS system (Bach et al., 2020; Wieczorek et al., 2015). While this approach first was considered for this research, it was ultimately determined that an ordinal scale would not provide the depth of insight required. The challenge lies in quantifying the performances of functions within an early-stage TIS without a quantitative understanding of the expected performance benchmarks.

However, in future research, an ordinal assessment, particularly within the context of a comparative case study, may offer valuable insights. Such an approach could exmplify how different functions compare and contrast with one another within the same stage of development, or across different TISs. This could lead to a more nuanced understanding of how early stage TISs evolve, and the specific factors that influence their progression from one stage to the next.

6.2.4 ADRESSING CONTEXTUAL FACTORS

After an initial delineation was made, the next step is how to deal with the context of the TIS. The TIS was formulated as technology-centered approach to a system analysis, in order to simplify the amount of factors that are involved (Hekkert et al., 2007). However, a strategy how to sufficiently deal with ‘contextual’ factors deemed necessary over the years. Markard & Truffer (2008) show how to position the TIS within the more contextual ‘MLP’ framework, however this compromises the TISs descriptive capacities as heuristics framework, as Bergek et al. (2015a) argued. They elaborate on the interactions between a focal TIS and other TISs, other sectors, geographical context and political context. For the field of AI, the TIS-TIS interactions proved essential. As the technological delineation between LLMs, generative AI, foundational models and NLP models is one still under debate and current policy and conversations use these terms and the umbrella term AI interchangeably, the research led inevitable to the AI field as important contextual TIS for researching the LLM field.

6.2.5 FUNCTIONAL PERFORMANCE

In adapting this framework to the focal TIS, it was necessary to start with the standard dynamics to construct a visualization that reflects the unique interactions present. This led to the the identification of three hampering points in functional performance in the two ‘cycles’ of cumulative causation.

Although Apell & Eriksson (2023) did not employ the 'motors of innovation' concept, their findings on the ‘functional barriers’ resonate with the cycle patterns identified. However, in this research the relation between data and regulations is reversed saying that: *“increased access and structuring of data”*, *“could catalyse the alignment between regulatory requirements and technology developments”.* So while guidance and resource mobilization are identified the order is different. Regarding the large national and European efforts currently undertaken to increase the mobilization of healthcare data, this research would argue that new regulations will enable resource mobilization and not vice versa.

This perspective is especially important given the significant national and European initiatives underway to enhance the mobilization of healthcare data. These efforts support the argument that new regulations are likely to be the precursors that enable resource mobilization, which is a critical step in the maturation of the TIS, rather than an outcome of increased data availability and structuring. This viewpoint underscores the importance of regulatory frameworks in guiding the evolution and development of TIS, particularly in fields that are as data-driven and regulation-sensitive as healthcare AI and LLMs.

6.2.6 ADDITIONAL AVENUES FOR FUTURE RESEARCH

This study can give a theoretical guideline for further research to study the formation of a LLMTIS in a healthcare context. This leads to two main avenues for future research. The first avenue is within the application of TIS research. It is stimulated to conduct comparative case-studies between different national contexts. For example, in other European countries, it can be assessed how the European regulative aspects hold up, how and if national regulative exemptions are present, how EHR providers are involved in LLM development or how far the formation of dedicated AI-teams in hospitals has progressed. This would enable the research to challenge or validate the insights on the Early-stage TIS in this report. The second avenue is more practical and falls outside the realm of innovation theory but connects with the three distinct possible dynamics underdevelopments mentioned in the paragraph cumulative causation. These lead to follow three options for further research I. Apply theory on effective knowledge clusters and hub formation on the Dutch Healthcare AI case to identify possible routes for increased convergence in the diffusion landscape such as perform by the theoretical framework in Evers. (2008). II. Apply a technological alignment check of all the proposed and active LLMs in Dutch healthcare with the quality indicators in the relevant regulative bodies such as performed by Rishi Bommasani (2023). III. Apply a regulative alignment check of Dutch hospital with upcoming data sharing regulations such as the WEGIZ and EHDS to detect common shortcomings. This comparative research would challenge or validate the insights on Early-stage TIS and enhance the understanding of national differences in the adoption and integration of LLMs in healthcare. These practical research options would provide insights into the effective application of LLMs in healthcare, focusing on knowledge diffusion, technological alignment and regulatory compliance with standards and the regulatory alignment of data mobilization.

6.3 LIMITATIONS

6.3.1 DATA COLLECTION

The data collection for Technology Innovation System (TIS) analysis is comprehensive, requiring the researcher to make multiple decisions based on the applicability, feasibility, and availability of data related to the phenomena under study. The collected data during desk research was not exhaustive for the researched structures or functions. For instance, compiling all projects allocating financial resources to LLM-related projects was not feasible. Considering these constraints, several observations can be made about the data collection process. Start-up databases have been tracking companies formed over the years, but their historical data collection practices might lead to an underestimation of companies founded before 2000. However, the data from the last five years is considered most accurate, aligning with the focus of the analysis. In the bibliometric analysis, triangulating three databases improved the validity of the results, revealing increasing publication trends. However, a direct correlation between the number of publications and the quality of knowledge development is not always straightforward, nor is the assumption that more companies always equate to more mature development. The data collection for guidance of the search involved analysing numerous policy documents and grey literature, where the varying definitions of what should be included in annual reports might lead to inconsistencies. However, if AI is a significant part of an institution's policy, it is unlikely to be entirely excluded from the reports. Resource mobilization data collection faced the challenge of quantifying human, financial, and data resources mobilized for the focal TIS. Lastly, in legitimization, public news sources were compared with interview findings, but quantifying the overlap between public opinion and interviewees' perspectives was not possible. These limitations were partly addressed through triangulation with semi-structured interviews and by focusing on the most important observations within each function.

The interviews themselves express other limitations that are characteristic for using interviews for data collection. The interviews actors are but a fraction of the total amount of actors present in the DH-LLMTIS. Their perspectives, while often overlapping, are still individual perspectives. Whenever a qualitative assessement was given, it was strived for multiple actors to overlap in order to include the findings in the results. More descriptive explanations of processes did not require the validation by multiple respondents but strived was for at least one desk research source. Desk research was executed prior to the second round of interviews, therefore, the findings could be directly validated when the conversation steer towards them.

6.3.2 TECHNOLOGICAL NOVELTY

The technological novelty of LLMs presented significant challenges. The academic field itself struggles with defining concepts such as foundational models, generative AI, and the boundary between medical and non-medical applications. The novelty of the field restricted data collection in some areas, such as the unavailability of patents or the inability to conduct extensive temporal analysis of historical events. Over time, it is expected that definitions will become more aligned, and more data will become available. However, delineation between the DH-LLMTIS and the DH-AITIS was often ambiguous, as respondents sometimes used AI and LLMs interchangbly. Nevertheless, the two TISs were found to be complementary, and the ambiguities did not significantly hinder the identification of three specific functionalities of the DH-LLMTIS. Future research could benefit from a more stringent definition of terms in interviews, to clarify delineation.

6.3.3 UNDERANALYSING CONTEXT

Under-analysing the contextual aspects of a TIS is a known critique of the framework (Markard et al, 2015). The development of LLMs is influenced by global dynamics, particularly by major players in the US and China, which this research could not detail extensively due to the scope of the focal TIS. The influence of global developments exerting a significant influence on the Dutch context were asked for in the interviews, however following the definition of different types of contextual factors by Bergek et al. (2015). This limitation could be improved upon by explicitly further study the political and sectoral interactions named as influential context factors of the focal TIS.

6.3.4 PREVIOUS LITERATURE

Previous TIS analyses provided varying degrees of applicable material. For instance, the decline of nuclear energy or the implementation of maritime battery-energy solutions illustrate a preference for using TIS in sustainable energy cases. Their methodological steps sometimes required additional validation for the focal TIS used in this research. An example is the use of vocational courses for windmill installation in Wieczorek et al. (2015) which was not relevant in an AI and LLM context that does not involve significant installation for iplementation. The exact overlap between methods used in these prior studies and the current research could be seen as a limitation, requiring further validation in future research.

The limitations described above highlight the complexity of TIS analysis and underscore the importance of careful data selection, clear delineation of TIS boundaries, comprehensive consideration of contextual factors, and adaptable methodologies to suit the specificities of the technology and sector under investigation.

6.4 PRACTICAL IMPLICATIONS

The dynamics described earlier align with the recommendations found in grey literature reports on AI in healthcare. Key conclusions from these reports include insights from KPMG (2020), Berenschot (2022), Cap Gemini invent (2020), and M&I Partners (2023). These reports highlight challenges such as limited awareness among healthcare institutions about AI applications developed or implemented elsewhere, the need for better actor connectivity, the significance of a centralized data access point, data accessibility issues, and the confinement of AI experiments within scientific spheres. Building upon these recommendations, shortcomings that become appearant in the discussion of the cumulative casusations in the DH-LLMTIS can be extended as follows:

First, the need for convergence in knowledge diffusion networks is underscored by the emergence of dispersed grassroots knowledge networks. A potential strategy could involve assessing existing knowledge networks within organizations to identify and reduce overlapping information, thereby streamlining efforts. National diffusion networks might explore achieving a more centralized role in knowledge dissemination.

Second, aligning the policy and regulatory landscape with technological needs is a complex task. Current efforts include harmonizing data storage and sharing regulations, medical device regulations, and the regulation of AI through the European AI-Act, all of which require time to become fully effective. Success in implementing these regulations is crucial. However, in the Dutch context, there is no centralized information hub for regulation standards and effective communication to market actors, while there is a strict interpretation of laws that limits experimentation. Learning from international examples and employing more dedicated personnel to oversee approval of applications would be beneficial.

Third, the ongoing data interoperability projects, such as the WEGIZ law and the EHDS, are crucial but in not fully implemented yet. Until these initiatives mature, individual market actors could proactively adopt emerging standards and practices to facilitate smoother implementation. For developers, gaining experience with non-medical, intra-organizational, Dutch-language LLMs could offer larger short-term achievability.

In order to put these practitcal implications to use, the effects on the different spheres needs to be considered. A market actor in a hospital can consider the regulation of data-flow that enters the organization in the form of the mentioned difussion networks or actively position itsself to comply to all the upcoming harmonized regulative changes. Individual actors in the governance sphere cannot change government policy. However, high level policy changes in the governance sphere could improve the dispersed field of knowledge difussion networks and the unclarity of communcation of regulations. An academic actor could consider the axis along which LLM enjoy the highest legitimation in the short term to either improve legitimization among the far-ends of the axes through their research or focus research on applications that have the highest chance of short-term implementation.

To end the practical implications, a concluding note must be given on the maturity of LLMs in healthcare. This research extensively shows the activities that are undertaken on in the field of AI in Dutch healthcare. However, activity must not be mistaken for progress. A formative TIS can easily be halted in its developments only to be dismantled and be forgotten. Yes, LLMs provide exiting new capabilities and yes the societal pressure to decrease costs is high, but after its release just over a year ago, ChatGPT and its likings have not yet transformed healthcare in ways some news articles predicted. A realistic view of the technological opportunities within the regulative and data frameworks of LLMs in Dutch healthcare will be necessary to fine-tune the DH-LLMTIS in the right direction.

7 CONCLUSION

In order to mitigate the problem of a unsustainable cost-increase in Dutch healthcare, well implemented technological innovations can make a positive contribution. Therefore, the potential implementation of the promising technological field of Large Language Models and their Natural Language Processing capabilities is of interest. To comprehensively understand the factors influencing the implementation of this technology, the study employed the heuristic framework of TIS analysis, rooted in evolutionary economics. This approach facilitated an exploration of the current state and functionality of the TIS for LLMs in Dutch Healthcare, leading to the primary research question:

RQ: What is the current state of the early-state Technological Innovation System of Large Language Models in Healthcare in the Netherlands?

In order to answer this research question, the first sub question was:

- ***SQ1: What is de delineation and context of the Dutch Large Language Model Technological Innovation System in healthcare?***

The technological state-of-the-art for LLMs in healthcare show a dispersed image. On the one hand, the technology is not without flaws and from a performance perspective multiple, for healthcare significant, barriers still exist. On the other hand, the attention given to the field, is not without a cause and developments following each other up in rapid succession, especially since 2017. Both the academic field and the general public are astounded by the successful combination of LLM models with a human moderator in a healthcare setting. In the Netherlands, AI in healthcare has put a foothold in healthcare providers all-around, however current literature questions the viability of projects to transfer from research to clinical use. For this research the focal TIS was delineated as a technological application field of LLMs with their NLP functionalities. The geographical scope was set on the Netherlands with its distinct healthcare system. The temporal scope was set on the current state, taking a recent temporal perspective when necessary for creating a comprehensive analysis.

After the structural and functional analysis were performed, a more detailed picture of the delineation and positioning of the DH-LLMTIS compared to the DH-AITIS and G-LLMTIS emerged. It showed that the DH-LLMTIS largely depends on the DH-AITIS, except for the functions of entrepreneurial activity, knowledge creation, and legitimacy. These same functions are also influenced but not dependent on the G-LLMTIS. A prime example is the influence worldwide developments of LLMs have on the legitimization of the technology.

Against this contextual background and delineation, the second sub question regarding the relevant structural elements is answered.

- ***SQ2: What are the relevant structural elements in the Dutch Large Language Model Technological Innovation System regarding actors, networks, institutions and infrastructure?***

The most important actors were divided into three spheres namely the governance, market and academic sphere. In the governance sphere the leading actor is the Dutch Ministry of Health and its subsidiaries of the department of medical devices and the IGJ as health regulator. The ministry of economic affairs plays an important role in providing financial resources and representing the Dutch healthcare sectors interest in the European regulative negotiations. In the market sphere, the leading actors are UMCS and TKZs as end-users and developers and start-ups and EHR providers as entrepreneurial developers. In the academic sphere the Universities with a direct contact to UMCs were identified as the main actors. The structural networks are an multitude of different formal or informal networks used for functionalities such as knowledge diffusion. They have been founded in recent years regarding the topic of AI or AI in healthcare of which the Nederlandse AI Coalitie is a central hub. The most important formal institutions that are of influence of AI and therefore LLMs have been identified as the MDR the GDPR and the upcoming EU AI-act. The computing infrastructure for LLMs in healthcare currently exists in the form of centralized computing centers on- premise at market actors or as secured cloud options off premise, mainly provided by American tech-companies.

This analysis sets the stage for examining the seven TIS functions, addressing the third sub question:

- ***SQ3: What is the performance of the different TIS functions of the Dutch Large Language Model Technological Innovation System in healthcare?***

After analysis and comparison with previous literature, the performance of the different functions for early stage TISs was assessed. Key underdeveloped areas compared to expectations for an early-stage TIS include knowledge diffusion, regulatory guidance, and data mobilization. The Dutch regulatory landscape, characterized by numerous, unclear, and unaligned regulations, hinders AI/LLM development. Data mobilization as a physical resource is crucial for enhancing technological performance and legitimacy. Function 2B Knowledge diffusion while active, is not providing the right spread of harmonized knowledge on AI and therefore LLMs in Dutch healthcare among relevant actors. This is in two-way negative relationship with the underdeveloped function 3B guidance of the search through regulations and expectations. Current regulations and expectations in the Netherlands are numerous, unclear unaligned and do not provide any sort of artificial breathing space for AI/LLMs applications to take their first steps. Regulations are both trailing behind recent developments and are not coordinated and communicated in a centralized way from the governmental sphere. Furthermore, the mobilization of data as a physical resource analysed in F5B is of essential value to increase the technological performance and thereby the legitimacy of LLMs in healthcare, again leading to more downstream effects that can project the TIS development forward. As of now, the DH-LLMTIS is to underdeveloped to express its own internal synergies, however the DH-AITIS has started to show the first signs of synergies most notably by the appearance of intermediate suppliers.

These functional performances and developmental levels enabled the answering of the main research question:

RQ: What is the current state of the early-stage Technological Innovation System of Large Language Models in Dutch Healthcare?

The state of the early-stage DH-LLMTIS is characterized by several key observations: from a TIS perspective the individual DH-LLMTIS is an a nascent underdeveloped state. However, incorporated in the broader context of the DH-AITIS a more developed scenario emerges.

Some functional aspects are as expected in an early-stage TIS. Such as the low amount of commercially deployed applications or the small current market formation. It would be incorrect to judge these functionalities as incorrectly developing just because the promising capabilities of LLMs are recent. However, other functionalities can benefit from a specific directionality. The state of the knowledge diffusion landscape is active but dispersed and could benefit from centralized convergence of efforts. The guidance through regulations is currently unclear for actors while at the same time significant work is done on the formation of new policy for AI including LLMs, specifically via the EU AI-act. These new regulation will exists on top of standing regulation regarding data-use and medical devices, all which are recently harmonized on European Scale. Their implementation is not yet familiarized by market actors and it uncertain what the effects of extra regulation would be on their entrepreneurial and experimental activities.

The generative, ever changing technical character and hard-to-define medical/non-medical boundaries of LLMs make it also unlikely that on short term an easy to follow roadmap will be available for market actors. The role of the governance sphere here is to guide the structuration process and provide centralized communication as soon as possible. Again, well-coordinated knowledge diffusion would benefit from this and vice versa. Mobilization of data as a resource stands apart from this two-way dynamic, but follows a similar pattern. Data mobilization efforts are currently also dispersed in the Netherlands. In this function, significant activity is ongoing to centralize and harmonize these activities on a national and European level of which the EHDS is a prime example. In the current technological and systemic state, market actors have a tendency towards non-generative, non-medical, Dutch LLMs that use data from inside the organization.

If the systemic functions described above are improved successfully, the upcoming years promise the opportunity for the DH-LLMTIS to draw from their strengths, using their synergic effects. This would enable a better technological performance, more entrepreneurial activity and finally the formation of a market that sets the TIS up for a transition from a formative phase to a growth phase.

Theory on innovation should inherently enable us to be better positioned for the future. This can be done by analysing the past, however the field becomes most relevant on the intersection between insights in historic developments and insights in current developments that can improve the trajectory of promising technologies going into the future. The upside to this vision is the relevance of findings for actors developing their innovation in the present, the downsides are the uncertainties of any analysis and the sparce availability of data. In that sense, hindsight is more comfortable. In this research, theory on technological innovation was used as heuristic framework to allow for sensemaking of a technology and an innovation, that is exactly on that intersection, promising a bright future for healthcare innovation. Until that time we have to stay in the waiting room to see a digital doctor.

8 ACKNOWLEDGEMENTS

The successful completion of this master’s thesis is owed to the contributions of several key individuals whose support and guidance have been instrumental.

Foremost, I extend my deepest appreciation to my academic supervisor, Dr. Iryna Susha, for her mentorship. Her willingness to start supervising me during an unconventional period, consistent availability to schedule our meetings and her ability to make time to discuss all the steps along the way, have enabled this entire research process.

I am also grateful to my internship supervisor, Dhr. Anniek Stroomer, whose mentorship has helped further scoping my research in a relevant direction, and whos weekly meetings have kept me on track while also providing me with a special eye for textual details.

I must express my sincere thanks to Celine Nagtegaal for her warm introduction to the professional setting and environment of the team, being the ideal 'Invent buddy'. Furthermore, the entire Health Team at CapGemini Invent deserves special mention for their hospitable and stimulating work environment in the last months.

I am indebted to all individuals who have engaged with me as sparring partners. Their willingness to discuss and challenge ideas has been essential in ordering my thoughts and bringing clarity to complex topics.

Lastly, I must extend a special word of thanks to all the individuals who generously agreed to participate in interviews for this research. Their willingness to share insights, experiences, and time has been a cornerstone of this thesis.

9 References

Ali, S. R., Dobbs, T. D., Hutchings, H. A., & Whitaker, I. S. (2023). Using ChatGPT to write patient clinic letters. In *The Lancet Digital Health* (Vol. 5, Issue 4, pp. e179–e181). Elsevier Ltd. https://doi.org/10.1016/S2589-7500(23)00048-1

Alkemade, F., & Suurs, R. A. A. (2012). Patterns of expectations for emerging sustainable technologies. *Technological Forecasting and Social Change*, *79*(3), 448–456. https://doi.org/10.1016/j.techfore.2011.08.014

Andersson, J., Hojcková, K., & Sandén, B. A. (2023). On the functional and structural scope of technological innovation systems – A literature review with conceptual suggestions. In *Environmental Innovation and Societal Transitions* (Vol. 49). Elsevier B.V. https://doi.org/10.1016/j.eist.2023.100786

Apell, P., & Eriksson, H. (2023). Artificial intelligence (AI) healthcare technology innovations: the current state and challenges from a life science industry perspective. *Technology Analysis and Strategic Management*, *35*(2), 179–193. https://doi.org/10.1080/09537325.2021.1971188

Autoriteit Consument & Markt. (2021). *ZIS/EPD-systemen: marktproblemen en oplossingsrichtingen*. https://www.acm.nl/nl/publicaties/zis-epd-systemen-marktproblemen-en-oplossingsrichtingen-een-tussenstand#:~:text=Printen-,ZIS%2FEPD%2Dsystemen%3A%20Marktproblemen%20en%20oplossingsrichtingen%2C%20een%20tussenstand,die%20voortkomen%20uit%20de%20mededingingsregels.

Bach, H., Bergek, A., Bjørgum, Ø., Hansen, T., Kenzhegaliyeva, A., & Steen, M. (2020). Implementing maritime battery-electric and hydrogen solutions: A technological innovation systems analysis. *Transportation Research Part D: Transport and Environment*, *87*. https://doi.org/10.1016/j.trd.2020.102492

Bauer, F., Coenen, L., Hansen, T., McCormick, K., & Palgan, Y. V. (2017). Technological innovation systems for biorefineries: a review of the literature. In *Biofuels, Bioproducts and Biorefining* (Vol. 11, Issue 3, pp. 534–548). John Wiley and Sons Ltd. https://doi.org/10.1002/bbb.1767

Bening, C. R., Blum, N. U., & Schmidt, T. S. (2015). The need to increase the policy relevance of the functional approach to Technological Innovation Systems (TIS). *Environmental Innovation and Societal Transitions*, *16*, 73–75. https://doi.org/10.1016/j.eist.2015.07.007

Benjamens, S., Dhunnoo, P., & Meskó, B. (2020). The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *Npj Digital Medicine*, *3*(1). https://doi.org/10.1038/s41746-020-00324-0

Bento, N., & Wilson, C. (2016). Measuring the duration of formative phases for energy technologies. *Environmental Innovation and Societal Transitions*, *21*, 95–112. https://doi.org/10.1016/j.eist.2016.04.004

Berenschot. (2022). *Artificiële intelligentie en passende zorg*. https://www.zorginstituutnederland.nl/publicaties/rapport/2022/09/29/onderzoeksrapport-artificiele-intelligentie-en-passende-zorg

Bergek, A. (2019). Technological innovation systems: a review of recent findings and suggestions for future research. In *Handbook of Sustainable Innovation* (pp. 200–218). Edward Elgar Publishing Ltd. https://doi.org/10.4337/9781788112574.00019

Bergek, A., Hekkert, M., Jacobsson, S., Markard, J., Sandén, B., & Truffer, B. (2015). Technological innovation systems in contexts: Conceptualizing contextual structures and interaction dynamics. *Environmental Innovation and Societal Transitions*, *16*, 51–64. https://doi.org/10.1016/j.eist.2015.07.003

Bergek, A., Jacobsson, S., Carlsson, B., Lindmark, S., & Rickne, A. (2008). Analyzing the functional dynamics of technological innovation systems: A scheme of analysis. *Research Policy*, *37*(3), 407–429. https://doi.org/10.1016/j.respol.2007.12.003

Bienefeld, N., Boss, J. M., Lüthy, R., Brodbeck, D., Azzati, J., Blaser, M., Willms, J., & Keller, E. (2023). Solving the explainable AI conundrum by bridging clinicians' needs and developers' goals. *Npj Digital Medicine*, *6*(1). https://doi.org/10.1038/s41746-023-00837-4

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., … Liang, P. (2021). *On the Opportunities and Risks of Foundation Models*. http://arxiv.org/abs/2108.07258

Boons, F., & McMeekin, A. (2019). *Handbook of sustainable innovation*. https://econpapers.repec.org/bookchap/elgeebook/17966.htm

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). *Language Models are Few-Shot Learners*. http://arxiv.org/abs/2005.14165

Bryman, A. , 2012. S. R. Methods. 4th ed. (2012). *Social Research Methods 4th edition*. Oxford ; New York: Oxford University Press. .

Cap Gemini invent. (2020). *Nulmeting databeschikbaarheid in gezondheid en zorg*. https://pht.health-ri.nl/nl

Carlsson, B., & Stankiewicz, R. (1991). Evolutionary Economics On the nature, function and composition of technological systems. In *J Evol Econ* (Vol. 1). https://link.springer.com/article/10.1007/BF01224915

Cascella, M., Montomoli, J., Bellini, V., Ottaiano, A., Santorsola, M., Perri, F., Sabbatino, F., Vittori, A., & Bignami, E. G. (2023). Writing the paper "Unveiling artificial intelligence: an insight into ethics and applications in anesthesia" implementing the large language model ChatGPT: a qualitative study. *Journal of Medical Artificial Intelligence*, *6*. https://doi.org/10.21037/jmai-23-13

Choi, W., Van Eck, M., Van Der Heijden, C., Hooghiemstra, T., & Vollebregt, E. (2022). *Legal analysis European legislative proposal draft AI act and MDR/IVDR*. https://hooghiemstra-en-partners.nl/report-ai-act-in-relation-to-mdr-and-ivdr/?lang=en

Claudio Lazo, A., Bodea, G., van Ette, F., Ailisto, H., Stenberg, S., & Burden, H. (2023). *Comparison of AI Strategies in Finland, Sweden and The Netherlands-Case the Netherlands*. https://nlaic.com/bouwstenen/research-en-innovatie/ai-in-beeld-de-strategieen-van-nederland-finland-en-zweden-onder-de-loep/
Cooke, P., Gomez Uranga, M., & Etxebarria, G. (1997). Regional innovation systems: Institutional and organisational dimensions. In *research policy ELSEVIER Research Policy* (Vol. 26). https://doi.org/https://doi.org/10.1016/S0048-7333(97)00025-5
Corbin, J., & Strauss, A. (2008). *Basics of Qualitative Research (3rd ed.): Techniques and Procedures for Developing Grounded Theory*. SAGE Publications, Inc. https://doi.org/10.4135/9781452230153
Davenport, T., & Kalakota, R. (2019). DIGITAL TECHNOLOGY The potential for artificial intelligence in healthcare. In *Future Healthcare Journal* (Vol. 6, Issue 2). https://doi.org/10.7861/futurehosp.6-2-94
David, P. A. (1985). *Clio and the Economics of QWERTY* (Vol. 75, Issue 2). American Economic Association. https://www.researchgate.net/publication/305389640_Clio_and_the_economics_of_QWERTY
de Kok, J. W. T. M., de la Hoz, M. Á. A., de Jong, Y., Brokke, V., Elbers, P. W. G., Thoral, P., Castillejo, A., Trenor, T., Castellano, J. M., Bronchalo, A. E., Merz, T. M., Faltys, M., van der Horst, I. C. C., Xu, M., Celi, L. A., van Bussel, B. C. T., & Borrat, X. (2023). A guide to sharing open healthcare data under the General Data Protection Regulation. *Scientific Data*, *10*(1), 404. https://doi.org/10.1038/s41597-023-02256-2
Dietterich, T., Bishop, C., Heckerman, D., Jordan, M., & Kearns, M. (2022). *Adaptive Computation and Machine Learning*. https://lccn.loc.gov/2021027430
DiMasi, J. A., Grabowski, H. G., & Hansen, R. W. (2016). Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of Health Economics*, *47*, 20–33. https://doi.org/10.1016/j.jhealeco.2016.01.012
Dorn, S. D. (2015). Digital Health: Hope, Hype, and Amara's Law. In *Gastroenterology* (Vol. 149, Issue 3, pp. 516–520). W.B. Saunders. https://doi.org/10.1053/j.gastro.2015.07.024
Edquist, C., & Charles Edquist, B. (2001). *The Systems of Innovation Approach and Innovation Policy: An Account of the State of the Art*. https://www.researchgate.net/publication/228823918_The_Systems_of_Innovation_Approach_and_Innovation_Policy_An_Account_of_the_State_of_the_Art
EPRS, E. P. R. S. (2022). *Artificial intelligence in healthcare*. https://doi.org/10.2861/568473
Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S., & Dean, J. (2019). A guide to deep learning in healthcare. In *Nature Medicine* (Vol. 25, Issue 1, pp. 24–29). Nature Publishing Group. https://doi.org/10.1038/s41591-018-0316-z
European Council. (2023). *European Council*. https://health.ec.europa.eu/publications/proposal-regulation-european-health-data-space_en
European Parliament. (2017). *Regulation on medical devices*. https://eur-lex.europa.eu/eli/reg/2017/745/oj?eliuri=eli:reg:2017:745:oj&print=true1/206ELI:http://data.europa.eu/eli/reg/2017/745/oj
European Parliament. (2023). *EU-AI act*. https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence
European Parliament Research Service. (2020). *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*. https://doi.org/10.2861/293
Eurostat. (2023). *Government expenditures on health*. Https://Ec.Europa.Eu/Eurostat/Statistics-Explained/Index.Php?Title=Government_expenditure_on_health.
Evers, H.-D. (2008). Knowledge Hubs and Knowledge Clusters: Designing a Knowledge Architecture for Development. *ZEF Working Paper Series 27 Center for Development Research Department of Political and Cultural Change*. https://www.researchgate.net/publication/254659273_Knowledge_Hubs_and_Knowledge_Clusters_A_Knowledge_Architecture_for_Development
Ferrari, F., van Dijck, J., & van den Bosch, A. (2023). Foundation models and the privatization of public knowledge. In *Nature Machine Intelligence* (Vol. 5, Issue 8, pp. 818–820). Nature Research. https://doi.org/10.1038/s42256-023-00695-5
Fogteloo. (2023, February). Waarom we ons moeten ontworstelen aan de wurggreep van Big Pharma. *De Groene Amsterdammer*. https://www.groene.nl/artikel/je-moet-de-regie-terugpakken
Freeman C. (1995). The "National System of Innovation" in historical perspective. In *Cambridge Journal of Economics* (Vol. 19, Issue 1). https://about.jstor.org/terms
Geels, F. W. (2002). Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study. In *Research Policy* (Vol. 31). https://doi.org/https://doi.org/10.1016/S0048-7333(02)00062-8
Greenhalgh, T., & Abimbola, S. (2019). The NASSS Framework A Synthesis of Multiple Theories of Technology Implementation. *Studies in Health Technology and Informatics*, *263*, 193–204. https://doi.org/10.3233/SHTI190123
Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., Hughes, G., A'Court, C., Hinder, S., Fahy, N., Procter, R., & Shaw, S. (2017). Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *Journal of Medical Internet Research*, *19*(11). https://doi.org/10.2196/jmir.8775
Hanson, J. (2018). Established industries as foundations for emerging technological innovation systems: The case of solar photovoltaics in Norway. *Environmental Innovation and Societal Transitions*, *26*, 64–77. https://doi.org/10.1016/j.eist.2017.06.001
He, J., Baxter, S. L., Xu, J., Xu, J., Zhou, X., & Zhang, K. (2019). The practical implementation of artificial intelligence technologies in medicine. In *Nature Medicine* (Vol. 25, Issue 1, pp. 30–36). Nature Publishing Group. https://doi.org/10.1038/s41591-018-0307-0
Hekkert, M. P., Janssen, M. J., Wesseling, J. H., & Negro, S. O. (2020). Mission-oriented innovation systems. *Environmental Innovation and Societal Transitions*, *34*, 76–79. https://doi.org/10.1016/j.eist.2019.11.011

Hekkert, M. P., & Negro, S. O. (2009). Functions of innovation systems as a framework to understand sustainable technological change: Empirical evidence for earlier claims. *Technological Forecasting and Social Change*, *76*(4), 584–594. https://doi.org/10.1016/j.techfore.2008.04.013

Hekkert, M. P., Suurs, R. A. A., Negro, S. O., Kuhlmann, S., & Smits, R. E. H. M. (2007). Functions of innovation systems: A new approach for analysing technological change. *Technological Forecasting and Social Change*, 74(4), 413–432. https://doi.org/10.1016/j.techfore.2006.03.002

Jacobsson, S., & Bergek, A. (2004). Transforming the energy sector: The evolution of technological systems in renewable energy technology. *Industrial and Corporate Change*, *13*(5), 815–849. https://doi.org/10.1093/icc/dth032

Javaid, M., Haleem, A., & Singh, R. P. (2023). ChatGPT for healthcare services: An emerging stage for an innovative perspective. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, *3*(1), 100105. https://doi.org/10.1016/j.tbench.2023.100105

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. In *Stroke and Vascular Neurology* (Vol. 2, Issue 4, pp. 230–243). BMJ Publishing Group. https://doi.org/10.1136/svn-2017-000101

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. In *Business Horizons* (Vol. 62, Issue 1, pp. 15–25). Elsevier Ltd. https://doi.org/10.1016/j.bushor.2018.08.004

Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. In *BMC Medicine* (Vol. 17, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s12916-019-1426-2

König, B., Janker, J., Reinhardt, T., Villarroel, M., & Junge, R. (2018). Analysis of aquaponics as an emerging technological innovation system. *Journal of Cleaner Production*, *180*, 232–243. https://doi.org/10.1016/j.jclepro.2018.01.037

KPMG. (2020). *Inventarisatie AI-toepassingen in gezondheid en zorg in Nederland*. https://www.rijksoverheid.nl/documenten/rapporten/2020/10/05/inventarisatie-ai-toepassingen-in-gezondheid-en-zorg-in-nederland

Kroneman, M., Boerma, W., Van Den Berg, M., Groenewegen, P., De, J., & Ewout Van Ginneken, J. •. (2016). Health Systems in Transition, Netherlands Vol.18 No.2 2016. In *Netherlands Health system review* (Vol. 18, Issue 2). https://iris.who.int/bitstream/handle/10665/330244/HiT-18-2-2016-eng.pdf?isAllowed=y&sequence=5

Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health*, *2*(2), e0000198. https://doi.org/10.1371/journal.pdig.0000198

Lambert, S. I., Madi, M., Sopka, S., Lenes, A., Stange, H., Buszello, C. P., & Stephan, A. (2023). An integrative review on the acceptance of artificial intelligence among healthcare professionals in hospitals. In *npj Digital Medicine* (Vol. 6, Issue 1). Nature Research. https://doi.org/10.1038/s41746-023-00852-5

Larisch, L. M., Amer-Wåhlin, I., & Hidefjäll, P. (2016). Understanding healthcare innovation systems: the Stockholm region case. *Journal of Health Organization and Management*, *30*(8), 1221–1241. https://doi.org/10.1108/JHOM-04-2016-0061

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. In *Nature* (Vol. 521, Issue 7553, pp. 436–444). Nature Publishing Group. https://doi.org/10.1038/nature14539

Lee, J., Davari, H., Singh, J., & Pandhare, V. (2018). Industrial Artificial Intelligence for industry 4.0-based manufacturing systems. *Manufacturing Letters*, *18*, 20–23. https://doi.org/10.1016/j.mfglet.2018.09.002

Lundvall, B.-A. (1985). *Product innovation and user-producer interaction*. Aalborg University Press. https://www.researchgate.net/publication/251542478_Product_Innovation_and_User-Producer_Interaction

Lundvall, B.-Å. (1992). *National systems of innovation: towards a theory of innovation and interactive learning*. Pinter Publishers. https://www.jstor.org/stable/j.ctt1gxp7cs

Malerba, F. (2002). Sectoral systems of innovation and production. In *Research Policy* (Vol. 31). https://doi.org/https://doi.org/10.1016/S0048-7333(01)00139-1

Mariani, M. M., Machado, I., Magrelli, V., & Dwivedi, Y. K. (2022). Artificial intelligence in innovation research: A systematic review, conceptual framework, and future research directions. *Technovation*. https://doi.org/10.1016/j.technovation.2022.102623

Markard, J. (2020). The life cycle of technological innovation systems. *Technological Forecasting and Social Change*, *153*. https://doi.org/10.1016/j.techfore.2018.07.045

Markard, J., Bento, N., Kittner, N., & Nuñez-Jimenez, A. (2020). Destined for decline? Examining nuclear energy from a technological innovation systems perspective. *Energy Research and Social Science*, *67*. https://doi.org/10.1016/j.erss.2020.101512

Markard, J., Hekkert, M., & Jacobsson, S. (2015). The technological innovation systems framework: Response to six criticisms. *Environmental Innovation and Societal Transitions*, *16*, 76–86. https://doi.org/10.1016/j.eist.2015.07.006

Markard, J., & Truffer, B. (2008). Technological innovation systems and the multi-level perspective: Towards an integrated framework. *Research Policy*, *37*(4), 596–615. https://doi.org/10.1016/j.respol.2008.01.004

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. https://doi.org/https://doi.org/10.1609/aimag.v27i4.1904

McKinsey. (2023). *McKinsey Technology Trends Outlook 2023*. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-top-trends-in-tech#/

Meskó, B., & Topol, E. J. (2023). The imperative for regulatory oversight of large language models (or generative AI) in healthcare. *Npj Digital Medicine*, *6*(1). https://doi.org/10.1038/s41746-023-00873-0

M&I Partners. (2023). *AI Monitor Ziekenhuizen 2023*. www.mxi.nl

Ministerie van Volksgezondheid Welzijn en Sport. (2022a). *Integraal Zorg Akkoord - Samen werken aan gezonde zorg*. https://www.rijksoverheid.nl/documenten/rapporten/2022/09/16/integraal-zorgakkoord-samen-werken-aan-gezonde-zorg
Ministerie van Volksgezondheid Welzijn en Sport. (2022b). *Kamerbrief Waardevolle AI voor Gezondheid*. https://www.rijksoverheid.nl/documenten/kamerstukken/2022/05/09/kamerbrief-over-waardevolle-ai-voor-gezondheid
Ministerie van Volksgezondheid Welzijn en Sport. (2022c). *Programmatoelichting Waardevolle AI voor Gezondheid*. https://www.datavoorgezondheid.nl/documenten/publicaties/2022/03/30/programmatoelichting-waardevolle-ai-voor-gezondheid
Moor, M., Banerjee, O., Abad, Z. S. H., Krumholz, H. M., Leskovec, J., Topol, E. J., & Rajpurkar, P. (2023). Foundation models for generalist medical artificial intelligence. *Nature*, *616*(7956), 259–265. https://doi.org/10.1038/s41586-023-05881-4
Moullin, J. C., Sabater-Hernández, D., Fernandez-Llimos, F., & Benrimoj, S. I. (2015). A systematic review of implementation frameworks of innovations in healthcare and resulting generic implementation framework. *Health Research Policy and Systems*, *13*(1). https://doi.org/10.1186/s12961-015-0005-z
Musiolik, J., Markard, J., & Hekkert, M. (2012). Networks and network resources in technological innovation systems: Towards a conceptual framework for system building. *Technological Forecasting and Social Change*, *79*(6), 1032–1048. https://doi.org/10.1016/j.techfore.2012.01.003
Nederlandse AI Coalitie. (2020). *Manifest werkgroep Gezondheid en Zorg*. https://nlaic.com/toepassingsgebied/gezondheid-en-zorg/
Nelson R. (1993). *National Innovation Systems: A Comparative Analysis*. Columbia University - School of International & Public Affairs (SIPA). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1496195
Park, H. J., Kim, S. H., Choi, J. Y., & Cha, D. (2023). Human–machine cooperation meta-model for clinical diagnosis by adaptation to human expert's diagnostic characteristics. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-43291-8
Patel, S. B., & Lam, K. (2023). ChatGPT: the future of discharge summaries? In *The Lancet Digital Health* (Vol. 5, Issue 3, pp. e107–e108). Elsevier Ltd. https://doi.org/10.1016/S2589-7500(23)00021-3
Peng, C., Yang, X., Chen, A., Smith, K. E., PourNejatian, N., Costa, A. B., Martin, C., Flores, M. G., Zhang, Y., Magoc, T., Lipori, G., Mitchell, D. A., Ospina, N. S., Ahmed, M. M., Hogan, W. R., Shenkman, E. A., Guo, Y., Bian, J., & Wu, Y. (2023). A study of generative large language model for medical research and healthcare. *Npj Digital Medicine*, *6*(1). https://doi.org/10.1038/s41746-023-00958-w
Planko, J., Cramer, J., Hekkert, M. P., & Chappin, M. M. H. (2017). Combining the technological innovation systems framework with the entrepreneurs' perspective on innovation. *Technology Analysis and Strategic Management*, *29*(6), 614–625. https://doi.org/10.1080/09537325.2016.1220515
Precedence Research. (2023). NLP in healthcare and life sciences market. *Https://Www.Precedenceresearch.Com/Nlp-in-Healthcare-and-Life-Sciences-Market*.
Rajpurkar, P., Chen, E., Banerjee, O., & Topol, E. J. (2022). AI in health and medicine. In *Nature Medicine* (Vol. 28, Issue 1, pp. 31–38). Nature Research. https://doi.org/10.1038/s41591-021-01614-0
Rathenau Instituut. (2023). *Generatieve AI*. https://www.rathenau.nl/nl/digitalisering/generatieve-ai
Rikap, C., & Lundvall, B.-Å. (2021). *The Digital Innovation Race Conceptualizing the Emerging New World Order*. https://doi.org/10.1007/978-3-030-89443-6
Rishi Bommasani, K. K. D. Z. P. L. (2023). *Do Foundation Model Providers Comply with the Draft EU AI Act?* https://crfm-stanford-edu.proxy.library.uu.nl/2023/06/15/eu-ai-act.html
Rivm. (2018). *Volksgezondheid Toekomst Verkenning 2018 Een gezond vooruitzicht Synthese*. www.vtv2018.nl
Roest, A. A. W., Adrichem, R., Le Cessie, S., Hazekamp, M. G., Van Dam, N. A. M., Blom, N. A., Rammeloo, L. A. J., Filippini, L. H. P. M., Kuipers, I. M., & Ten Harkel, A. D. J. (2019). Risk of Clinically Relevant Pericardial Effusion After Pediatric Cardiac Surgery. *Pediatric Cardiology*, *40*(3), 585–594. https://doi.org/10.1007/s00246-018-2031-4
Sallam, M. (2023). ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. In *Healthcare (Switzerland)* (Vol. 11, Issue 6). MDPI. https://doi.org/10.3390/healthcare11060887
Schumpeter. (1937). Preface to japanese edition of "Theorie der wirtschaflichen entwicklung." *Journal of American Statistical Association*.
Singhal, K., Tu, T., Gottweis, J., Sayres, R., Wulczyn, E., Hou, L., Clark, K., Pfohl, S., Cole-Lewis, H., Neal, D., Schaekermann, M., Wang, A., Amin, M., Lachgar, S., Mansfield, P., Prakash, S., Green, B., Dominowska, E., Arcas, B. A. y, … Natarajan, V. (2023). *Towards Expert-Level Medical Question Answering with Large Language Models*. http://arxiv.org/abs/2305.09617
Statista. (2023). *AI in healthcare market size worldwide*. https://www.statista.com/statistics/1334826/ai-in-healthcare-market-size-worldwide/
Strohm, L., Hehakaya, C., Ranschaert, E. R., Boon, W. P. C., & Moors, E. H. M. (2020). Implementation of artificial intelligence (AI) applications in radiology: hindering and facilitating factors. *European Radiology*, *30*(10), 5525–5532. https://doi.org/10.1007/s00330-020-06946-y
Suurs, R. A. A., & Hekkert, M. P. (2009). *Motors of sustainable innovation*. https://www.researchgate.net/publication/255587265_Motors_of_sustainable_innovation
Suurs, R. A. A., Hekkert, M. P., Kieboom, S., & Smits, R. E. H. M. (2010). Understanding the formative stage of technological innovation system development: The case of natural gas as an automotive fuel. *Energy Policy*, *38*(1), 419–431. https://doi.org/10.1016/j.enpol.2009.09.032
Techleap. (2021). *Dutch healthtech report 2021*. https://techleap.nl/report/dutch-healthtech-2021-report-unlocking-its-untapped-potential/

The Economist. (2023a, April 20). *How AI could change computing, culture and the course of history*. https://www.economist.com/essay/2023/04/20/how-ai-could-change-computing-culture-and-the-course-of-history

The Economist. (2023b, June 30). *The widespread adoption of AI by companies will take a while*. https://www.economist.com/leaders/2023/06/29/the-widespread-adoption-of-ai-by-companies-will-take-a-while

Thimbleby, H. (2013). Technology and the future of healthcare. In *Journal of Public Health Research* (Vol. 2). https://doi.org/doi: 10.4081/jphr.2013.e28.

Thirunavukarasu, A. J., Ting, D. S. J., Elangovan, K., Gutierrez, L., Tan, T. F., & Ting, D. S. W. (2023). Large language models in medicine. In *Nature Medicine* (Vol. 29, Issue 8, pp. 1930–1940). Nature Research. https://doi.org/10.1038/s41591-023-02448-8

Topteam ICT, V.-N. I. T. B. D. (2018). *AI voor Nederland vergroten versnellen en verbinden*. https://www.vno-ncw.nl/sites/default/files/aivnl_20181106_0.pdf

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need*. http://arxiv.org/abs/1706.03762

Webster, P. (2023). Medical AI chatbots: are they safe to talk to patients? In *Nature Medicine*. Nature Research. https://doi.org/10.1038/s41591-023-02535-w

Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., Metzler, D., Chi, E. H., Hashimoto, T., Vinyals, O., Liang, P., Dean, J., & Fedus, W. (2022). *Emergent Abilities of Large Language Models*. http://arxiv.org/abs/2206.07682

Widner, K., Virmani, S., Krause, J., Nayar, J., Tiwari, R., Pedersen, E. R., Jeji, D., Hammel, N., Matias, Y., Corrado, G. S., Liu, Y., Peng, L., & Webster, D. R. (2023). Lessons learned from translating AI from development to deployment in healthcare. In *Nature Medicine*. Nature Research. https://doi.org/10.1038/s41591-023-02293-9

Wieczorek, A. J., Hekkert, M. P., Coenen, L., & Harmsen, R. (2015). Broadening the national focus in technological innovation system analysis: The case of offshore wind. *Environmental Innovation and Societal Transitions*, *14*, 128–148. https://doi.org/10.1016/j.eist.2014.09.001

Wieczorek, A. J., Negro, S. O., Harmsen, R., Heimeriks, G. J., Luo, L., & Hekkert, M. P. (2013). A review of the European offshore wind innovation system. In *Renewable and Sustainable Energy Reviews* (Vol. 26, pp. 294–306). https://doi.org/10.1016/j.rser.2013.05.045

Wornow, M., Xu, Y., Thapa, R., Patel, B., Steinberg, E., Fleming, S., Pfeffer, M. A., Fries, J., & Shah, N. H. (2023). The shaky foundations of large language models and foundation models for electronic health records. In *npj Digital Medicine* (Vol. 6, Issue 1). Nature Research. https://doi.org/10.1038/s41746-023-00879-8

Wouda, F., Hutink, H., & Nell, J. (2019). *Artificial Intelligence in de zorg*. https://nictiz.nl/publicaties/artificial-intelligence-in-de-zorg/

Wouterse. (2017). How to Finance the Rising Costs of Long-Term Care: Four Alternatives for the Netherlands. *Fiscal Studies, Vol. 38, No. 3, Special Issue on the Challenges of Public Financing and Organisation of Long-Term Care* , 369–391. https://www.jstor.org/stable/26605627

Zahlan, A., Ranjan, R. P., & Hayes, D. (2023). Artificial intelligence innovation in healthcare: Literature review, exploratory analysis, and future research. *Technology in Society*, *74*. https://doi.org/10.1016/j.techsoc.2023.102321

Zolfagharian, M., Walrave, B., Raven, R., & Romme, A. G. L. (2019). Studying transitions: Past, present, and future. *Research Policy*, *48*(9). https://doi.org/10.1016/j.respol.2019.04.012

APPENDIX A: INTERVIEW GUIDES

Semi-structured interviews #1

1. **Pre-interview**
   Send informed consent form

2. **Interview Questions Guide**

Introduction and confirmation of consent

Preliminary scoping questions for the structural elements
Part I: Actors in the field op LLM implementations in healthcare in the Netherlands?
Who are the main actors in the field? Prompts to give: 'these may include universities and research institutes, public bodies, influential interest organizations, venture capitalists, organizations deciding on standards, etc. And how have they developed in the last years?

Part II: Institutions
What are the main institutions present in the field" Prompts to give: 'These can include norms, laws, regulations, routines, culture, that are of interest in the field or the lack hereof.' And how have they developed in the last years?

Part III: Networks
What are the main networks in the field? Prompts to give: 'these could include the formal and informal networks that are used most often such as; standardization networks, technology platform consortia, public–private partnerships, supplier groups, important buyer–seller relationships, university-industry links, technology advocacy groups' And how have they developed in the last years?

Part IV: Infrastructure
What are the main infrastructures in the field? Prompts to give: 'these could include the physical resources, technical artefacts and requirements for the technology to work' And how have they developed in the last years? If there are multiple options named for the same technology what would you prefer and why?

Closing & validation

- Okay that was the official part of the interview
- What you s AI d here and there, am I correct to read it like …. and incorporate it as such in my research?
- Do you have any questions for me?
- You will get a copy of the end-result in requested
- Ag AI n thank you for your time and cooperation

Semi-structured interviews #2

1. **Pre-interview**
   Send informed consent form

2. **Interview Questions Guide**

Introduction

Academic A/Governance G /Market M

- The letter indicate the focus on specific interviews during an interview.
- When answers are directed towards AI in general or LLM worldwide, follow questions are provided.

Preliminary scoping questions for the functions

F1. Entrepreneurial activities (M, A)

'How would you rate the amount of entrepreneurial activity that takes place for LLM use in healthcare in the Netherlands (weak, intermediate, strong)?'

- Are there sufficient and suitable types of actors contributing to and up-scaling?
- How much technological up-scaling takes place?
- And in AI in general? And in LLM internationally

F2. Knowledge development & Knowledge exchange (A, M)

'How has the amount of knowledge on LLM use in healthcare in the Netherlands changed in the last 5 years and is it (weak, intermediate, strong)?'

- Have you been aware of a lot of publishing in this field?
- Are there enough actors involved in knowledge development and are they competent?
- Is the knowledge sufficiently developed and aligned with needs of actors in the innovation system?
- Are there sufficient network connections between actors through which knowledge is exchanged

F3. Guidance of the search (A, G)

"Do actors and institutions provide a sufficiently direction for the future development of the technology? (strong, intermediate, weak)?"

- Who provide the strongest directionality
- And in AI in general? And in LLM internationally

F4. Market formation (M, G)

'Is there currently a market for these technologies? Are there incentives for (actors) to move in?'

- Is the size of the market sufficient to sust AI n innovation and entrepreneurial experimentation?
- And in AI in general? And in LLM internationally

F5. Resource mobilization (A,M,G)

'Are their huma/financial/data resources assigned to this technology? (Weak intermediate strong)

- Is there a plan in place of formulated to provide infrastructure for these technologies?
- Is there enough funding and from which source?
- And in AI in general? And in LLM internationally

F6. Creation of legitimacy (M, G)

'How is the process going of legitimacy of LLMs in healthcare?

- Do actors, formal and informal institutions sufficiently contribute to legitimacy?
- How much resistance is present towards the technology, project set up or permit procedure?
- And in AI in general? And in LLM internationally

F7. Synergies & positive externalities (A, M)
'*Are there standards in place or made for these technologies to comply with?*'

- Are there sufficient complementary goods av AI lable to facilitate the technology?
- Are new industry standards sufficient to facilitate the technology?
- And in AI in general? And in LLM internationally

Closing & validation

- Okay that was the official part of the interview
- What you s AI d ... and ...., am I correct to read it like .... and incorporate it as such in my research?
- Do you have any questions for me?
- You will get a copy of the end-result in requested
- Ag AI n thank you for your time and cooperation

APPENDIX B: SEARCH QUERIES

F1 Entrepreneurial activity

Techleap.nl

Netherlands: All locations

health platform: Sub-industry

medical devices: Sub-industry

biotechnology: Sub-industry

pharmaceutical: Sub-industry

artificial intelligence: Technologies

machine learning: Technologies

deep learning: Technologies

natural language processing: Technologies

Startups.eithealth.eu

HQ location: Netherlands

artificial intelligence: Technologies

machine learning: Technologies

deep learning: Technologies

natural language processing: Technologies

Crunchbase.com

AI

HQ location: "Netherlands" "US" "USA" "United States" "UK" "United Kingdom" "Germany" "France" "Switzerland" "Sweden" "Denmark" " Belgium"

Industry: Healthcare

Keywords: "Artificial Intelligence" "AI" "Natural Language Processing" "Machine Learning" "NLP" "Large Language Model" "LLM" "Genai"

HQ location: "Netherlands" "US" "USA" "United States" "UK" "United Kingdom" "Germany" "France" "Switzerland" "Sweden" "Denmark" " Belgium"

Industry: Artificial Intelligence

Keywords: "Medical" "Doctor" "Patient" "clinic" "clinical" "hospital" "health "healthcare" "care"

LLM, NLP

HQ location: "Netherlands" "US" "USA" "United States" "UK" "United Kingdom" "Germany" "France" "Switzerland" "Sweden" "Denmark" " Belgium"

Industry: Healthcare

Keywords: "Natural Language Processing" "NLP" "Large Language Model" "LLM" "Genai"

HQ location: "Netherlands" "US" "USA" "United States" "UK" "United Kingdom" "Germany" "France" "Switzerland" "Sweden" "Denmark" " Belgium"

Industry: "Natural Language Processing" "Generative AI"

Keywords: "Medical" "Doctor" "Patient" "clinic" "clinical" "hospital" "health "healthcare" "care"

F2 Knowledge development & Difussion

Scopus

TITLE-ABS-KEY-AUTH (("large language model" OR "generative AI" OR "generative artificial intelligence" OR "chatgpt" OR "ChatGPT" OR "GENAI" OR "natural language processing") AND ("patient" OR "Health" OR "Healthcare" OR "medicine" OR "medical" OR "clinical" OR "hospital")) AND PUBYEAR > 2016 AND PUBYEAR < 2025

Country / Territory: Netherlands

Country / Territory: Listed

TITLE-ABS-KEY-AUTH (("large language model" OR "generative AI" OR "generative artificial intelligence" OR "GENAI" OR "natural language processing" AND NOT "chatgpt" OR "ChatGPT") AND ("patient" OR "Health" OR "Healthcare" OR "medicine" OR "medical" OR "clinical" OR "hospital")) AND PUBYEAR > 2016 AND PUBYEAR < 2025

Country / Territory: Netherlands

Web of Science

TS=(("large language model" OR "generative AI" OR "generative artificial intelligence" OR "chatgpt" OR "ChatGPT" OR "GENAI" OR "natural language processing") AND ("patient" OR "Health" OR "Healthcare" OR "medicine" OR "medical" OR "clinical" OR "hospital")) PUBYEAR > 2016 AND PUBYEAR < 2025

Countries / Regions: Netherlands

Countries / Regions: Listed

TS=(("large language model" OR "generative AI" OR "generative artificial intelligence" OR "GENAI" OR "natural language processing" AND NOT "chatgpt" OR "ChatGPT") AND ("patient" OR "Health" OR "Healthcare" OR "medicine" OR "medical" OR "clinical" OR "hospital")) PUBYEAR > 2016 AND PUBYEAR < 2025

Countries / Regions: Netherlands

PubMed

("large language model"[Title/Abstract] OR "generative AI"[Title/Abstract] OR "generative artificial intelligence"[Title/Abstract] OR "GENAI"[Title/Abstract] OR "natural language processing"[Title/Abstract] OR "Chatgpt"[Title/Abstract] OR "ChatGPT"[Title/Abstract])

"Netherlands" "(possible top 20 countries)" [AD]

("large language model"[Title/Abstract] OR "generative AI"[Title/Abstract] OR "generative artificial intelligence"[Title/Abstract] OR "GENAI"[Title/Abstract] OR "natural language processing"[Title/Abstract] AND NOT "Chatgpt"[Title/Abstract] OR "ChatGPT"[Title/Abstract])

"Netherlands" [AD]

F5 Resource mobilization

Cordis.europe.eu

Collection: Results in brief

Collection: Project deliverables

Collection: Project publications

Collection: Report summaries

Collection: Projects

Domain of application: Health

Framework Programme: Horizon Europe

Framework Programme: Horizon 2020

Organisation Country: Netherlands

"Artificial Intelligence"

F6 Legitimacy

Advance-Lexis-com.proxy.library.uu.nl

("zorg" OR "medische" OR "arts" OR "medisch" OR "patient" OR "klinisch" OR "ziekenhuis" OR "gezondheid")

AND ("Artificiële intelligentie" OR "AI" OR "CHATGPT" OR "chat GPT" OR "generatieve artificiele intelligentie" OR "artificial intelligence" OR "Taalmodel" OR "Language model" OR "LLM" OR "Large Language Model" OR "kunstmatige intelligentie")

Language: Dutch

Date is after: 30 November 2022

Sources are: WebNews - Dutch, Trouw.nl, Artsenkrant (Dutch), NRC.nl, NRC, Trends Magazine, Noordhollands Dagblad, Trouw, AD/Algemeen Dagblad.nl, HP/De Tijd, De Gelderlander, De Twentsche Courant Tubantia, De Gooi- en Eemlander, De Gelderlander.nl, Provinciale Zeeuwse Courant, Brabants Dagblad, De Groene Amsterdammer, BN/DeStem, AD/Algemeen Dagblad, Reformatorisch Dagblad , De Stentor, De Volkskrant.nl, Brabants Dagblad.nl, Tubantia.nl, PZC.nl, De Stentor.nl Noordhollands Dagblad.nl, NU.nl, Het Parool, Dagblad van het Noorden, de Volkskrant, FD.nl, Het Parool.nl, BN De Stem.nl, Metronieuws.nl, Friesch Dagblad, Dagblad van het Noorden.nl

AND ("Artificiële intelligentie" OR "generatieve artificiele intelligentie" OR "artificial intelligence" OR "Taalmodel" OR "Language model" OR "LLM" OR "Large Language Model" OR "kunstmatige intelligentie" AND NOT (OR "AI" OR "CHATGPT" OR "chat GPT"))

APPENDIX C: CODING SCHEMES

Structural analysis

Exemplary quote	Structure (closed category)	Theme	Aggrated observational category
"Yes, the [UMC] is certainly very progressive in AI and the implementation of AI, as well as in the evaluation of different AI systems, to compare them, like, okay, there are now five different hand bone scans that you can do, which one is the best, you know, so you want to compare them with each other. So they are very big in that. Yes, I am not really, when you talk about LLMs, I don't really know of many other people who are also involved in this, except that it's starting to play a role for everyone, but as you said, it's nobody's core business yet, but I indeed think that from the innovation department, I would talk to the people from the AI team from the hospital	**Actors**	**UMCs have recently developed departments**	**UMCs as actors**
" but look, most of the innovation at the university medical centers is of course just the research that they regularly do, but also within large UMCU (University Medical Center Utrecht), you also have such an innovation lab, and there these kinds of things happen, then you can just try them out."	**Actors**	**UMCs are a source of knowledge**	
"and then you actually enter a whole trajectory of Medical Devices and how that's terribly mismanaged. And that makes it very difficult. So, if we just look at a timeline. You have a product already finished, so you've really done your research and everything. Well, then you want to get into a hospital. Well, then you first need to get your certificate that you are a Medical device and if you do it right and that you are compliant with all regulation. Well, then you first need to go through a notified body, and a notified body is designated by the EU in a country, and it's a private company to do the first part of the work for the certificate. Well, the EU does not designate enough companies, and there's too much demand to get your certificate within those private companies. So, the waiting time to get the first step of your certificate is a year in the Netherlands, so you apply. You wait a year to get to the notified body.	**Institutions**	**MDR system**	**MDR as institutional factor**
"It depends on the purpose of the AI. So, if it's part of the Medical Device Regulation, you have to look at whether the AI is co-responsible for the medical decision being made. And then, what class does it belong to? Is it a 1, 2a, or 2b? You have to comply with that. There are standards you could meet."	**Institutions**	**MDR categories and their implications**	

Functional analysis

Exemplary quote	Function (closed category)	Theme	Aggrated observational category
"For example, [EHR] has set up 1100 people to implement ChatGPT into their system. On a server in Europe, you've tackled one thing, with data that is [EHR] and not from others, and everything under one name, so it is traceable. That can also be significantly traced, so that's interesting. And they have trained it themselves to say, but we think it doesn't make mistakes or something similar. And you can always check it yourself. Well, yes. So they have introduced a few of those elements. It looks really good and it's also a lot of fun. Those are the first examples of professional use of LLM in health and care in the Netherlands. Yes."	**F1 Entrepreneurial activity**	**EHR providers are active in AI/LLMs**	**Incumbent activity**
'No, because we are one of the strategic partners of []; they provide a lot of technology and software. We deliver a very specialized system, but we are also in discussions with [], as well as with [], actually with all the major technology players, to see how we can connect or integrate the AI technology they provide with our EPD (Electronic Patient Record). The beauty of this is that we don't really have to make a choice; we simply say, we let healthcare itself determine which suppliers they want to select. So that could be [], [], it doesn't really matter to the hospitals, and we then ensure that the integration on the back end is already organized.'	**F1 Entrepreneurial activity**	**EHR providers are active in AI/LLMs**	
"I mean, a department can just have a good researcher who knows a lot about it, even more than we do, maybe someone who is specialized in a particular model, but then they often don't know, we are kind of a generalist in that sense, ... And that broad knowledge, that is what most departments lack."	**F2 Knowledge development & Diffusion**	**Knowledge in hospital varies**	**Hospital Knowledge development**
"Yes, but you notice that some small hospitals, and it depends very much on the region and the funding, but they are sometimes further ahead than university medical centers or TKZ hospitals. There are really a few among them. Yes, the mindset of the Board of Directors and the knowledge of the people in-house, and you notice that some hospitals have just really attracted very good people who really know what they're doing, and that pays off."	**F2 Knowledge development & Diffusion**	**Knowledge development differences between UMCs and TKZs**	

Exemplary quote	Function (closed category)	Theme	Aggrated observational category
"Europe is not entirely leading the way in this regard, in terms of guidances or guidelines in the medical device regulation. You do see, for example, the FDA in America, which is actively involved in this. And a few years ago, they introduced a framework for how to deal with these types of issues. And can't you move to a method where you agree on a bandwidth with the regulator beforehand? If my AI performs within that bandwidth, then it's just okay."	**F3 Guidance of the search**	**MDR regulation is rigid in Euirope and the Netherlands**	**MDR regulations guidance**
"Now we are fully engaged in a trajectory and also working with the [] to develop policy for AI tools. I think it then becomes less scary for people or less of a barrier for people to go through. I think initially it was just unclear what exactly you had to do for the MDR, when exactly you complied with such things."	**F3 Guidance of the search**	**MDR mitigation by data-teams**	
"LLMs are piloting in multiple hospitals. In the coming weeks, we will also start piloting at the []. So it can just be used with patients during actual consultations. Doctors can test it to see what they think of it.	**F4 Market formation**	**LLM are in pilote phase**	**Current status of market formation**
"Looking at AI models, also AI models in healthcare and those applications in healthcare, you would of course expect that a number of years later, the number of applications that actually enter the market, are used and are scaled up, that it will follow a similar trend, so that it also increases significantly. But I think that's still lagging behind, so for me at the moment there are still many great ideas being tried out, but actually implementing them in healthcare, well, that's still not very common."	**F4 Market formation**	**Low amount of commercially implemented AI applications**	