Noémi Lemonnier

Ethical Issues using Artificial Intelligence in Healthcare

Master's thesis of mathematical information technology

April 29, 2021

University of Jyväskylä

Department of Mathematical Information Technology

Author: Noémi Lemonnier

Contact information: noemilemonnier@gmail.com

Supervisors: Ville Vakkuri and Vagan Terziyan

Title: Ethical Issues using Artificial Intelligence in Healthcare

Project: Master's thesis

Study line: Cognitive Computing and Collective Intelligence

Page count: 79+10 = 89 pages

Abstract: This thesis presents the results of the Systematic Mapping Study of Artificial Intelligence (AI) Ethics in healthcare. AI ethics thrives to reduce ethical issues to create moral, fair, and safe AI applications. This thesis aims to provide a more precise view of AI ethics in healthcare. In healthcare, the four main ethical issues mentioned throughout various published research are transparency, justice and fairness, accountability and responsibility, and privacy and security. As the AI industry is constantly expanding, AI ethics in healthcare will become a growing concern for society. Additionally, the AI field can lack information, clarity, and structure. Thus, identifying the origin of ethical issues and providing solutions for them will be relevant for the day-to-day and academic spheres. A clear proposition to lessen these ethical issues in the research domain has yet to be mentioned as most research focus on highlighting issues without providing concrete solutions. This thesis contributes to the research field by analyzing the relationships between the different stakeholders involved and their respective ethical issues. A total of 56 papers were analyzed and the results were 15 empirical conclusions that highlighted the current literature and its gaps, for example, which stakeholder is mentioned less and research more to limit further moral issues.

Keywords: Artificial Intelligence, Healthcare, AI Ethics, Medicine, Systematic Mapping Study

Glossary

AI	Artificial Intelligence
SMS	Systematic Mapping Study
SLR	Systematic Literature Review
ML	Machine Learning
NLP	Natural Language Processing
DL	Deep Learning
CV	Computer Vision
CNN	Convolutional Neural Networks
ANN	Artificial Neural Networks
RNN	Recurrent Neural Networks
HER	Electronic Health Records
BCI	Brain-Computer Interfaces
RPA	Robot Process Automation
XAI	Explainable Artificial Intelligence
EC	Empirical Conclusion

List of Figures

List of Tables

Contents

1 Introduction

In the last decade, our society has evolved drastically especially regarding technologies. One, in particular, that has grabbed the public and experts' attention by its potential, is Artificial Intelligence (AI). Despite having various meanings, the one selected for AI regarding this thesis is "a discipline that combines computer science, engineering and related fields to build machines capable of mimicking human cognitive processes". (Murphy, et al., 2021) AI holds numerous promises to help and improve society's daily life. For example, it is actively used by social media, healthcare, and other domains to predict customer behaviors or analyze images. Despite having numerous benefits, AI also bears disadvantages such as ethical issues. Society is concerned about unemployment, privacy and surveillance, bias, and discrimination. (Pazzanese, 2020) Currently, health systems across nations are going through high demand, pressure, and stress as the coronavirus pandemic is spreading. AI has been used in healthcare to help medical experts during this difficult time. Therefore, the research on AI ethics in healthcare is a currently relevant topic.

Although the competition to provide the best AI solutions is constantly growing, many possible moral questions are not considered. Hence, society is slowly expressing a need for ethical legislation of AI. One that has emerged during the 20th century is AI ethics and expresses moral concerns, principles, and values related to the use of intelligent machines. AI ethics will be defined later on. However, it is the ethics of technology that regroups some of these principles. Despite AI ethics questioning how AI systems are designed, made, used, and treated (Jobin, Ienca & Vayena, 2019), the use of AI in healthcare has created various ethical issues such as accountability, privacy, and transparency issues. (Davenport & Kalakota, 2019) In 2017, the first FDA DL application was approved for healthcare (Kaul et al., 2020) and the European Parliament established the Civil Law Rules on Robotics: European Parliament resolution of 16 February 2017 which included guidance to AI in healthcare. (Gerke et al., 2020) Thus, this thesis aims to provide a more precise view of certain ethical impacts that AI has on healthcare.

1.1 Research Questions

To achieve the objective of this thesis, a research method is necessary. Thus, the chosen research method is Systematic Mapping Study (SMS) which will help to map the current academic literature of AI ethics in healthcare as notions related to AI keep changing rapidly. This methodology choice is explained in the following section. As for the main research question, *what is the current state of ethical issues created by using AI in healthcare,* it is split into the following:

> *[R1] What is the current state of stakeholders involved in using AI in healthcare in the research field?*
>
> *[R2] How are the ethical issues using AI in healthcare mitigated in the research field?*
>
> *[R3] What are the current gaps in the research field?*

These sub-questions will be explored in more detail in the SMS chapter.

1.2 Research Method

For this thesis, the research method needed to provide a broad view of the current state of academic literature. As Petersen et al. (2008) explained, the Systematic Mapping Study (SMS) results present the quantity and type of the relevant literature reviewed as well as the current gaps in the academic literature. Therefore, an SMS was selected over a literature review as the topic is still emerging and changes rapidly. SMS is a methodology that provides an overview of the type of reports and results published by categorizing them. (Kitchenham et al., 2012) As shown in Figure 1, it consists of many steps and each has different outcomes.

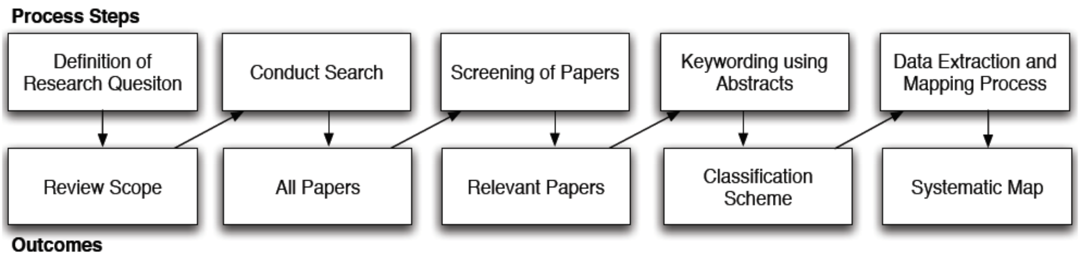

Figure 1. SMS Steps and Outcomes from Petersen et al. (2008)

The first step is to define the research questions that will reduce the quantity and types of research and results within that field; this will set the study scope. Then, the conducted search will define the search string across different databases. This search string comes from the keywords of the research questions. Next, the screening of papers defines which articles are to include and exclude from the research. Keywording allows one to reduce the time required for creating the classification scheme and ensures that the scheme takes the existing research into account. Finally, after the classification scheme is completed and the relevant sources are sorted, the data extraction starts. It will highlight the areas that will need further research by analyzing the frequency of types of existing publications. (Petersen et al., 2008) These steps are more closely analyzed in Chapter 3.

Also, SMS is a type of Systematic Literature Review (SLR). Both, SMS and SLR, are secondary studies. It means that they are data gathered from previously conducted research, compared to primary studies that are self-conducted data. SLR tends to collect, select, and analyze primary studies to answer a specific research question in contrast to SMS that provides a broader view by categorizing the current literature. SLR tends to be based on empiric data, while SMS is based on constructive data. (Kitchenham et al., 2011)

For this reason, SMS methodology is a better option to understand the current ethical issues involving AI in healthcare. The main benefit of SMS is that the research content can be reused for an SLR which would be less time-consuming. Although, that means the SMS methodology was done correctly and information is up to date. (Kitchenham et al., 2011) It also is a very time-consuming study which means that not everyone can provide good results with SMS methodology. Therefore, the inclusion and exclusion criteria are important to limit the amount of literature to ensure the quality of the research. (Petersen et al., 2008)

1.3 Thesis Outline

The second chapter of the thesis presents the theory needed for the SMS methodology. It briefly goes over the evolution of AI, the type of intelligent machines used in healthcare, and ethics. The history provides an overview of how fast AI has been growing and its potential. Also, it describes the type of AI currently used and in which domain. Finally, the ethics section provides some of the current issues faced by developers, medical experts, patients, AI autonomous machines, and governments and institutions. In summary, this background gives the reader an understanding of the current state of the field of AI ethics in healthcare.

The third chapter covers the literature search process. This continues the SMS methodology description by going over the research questions and process, the primary search, and the inclusion and exclusion criteria. The number of articles from 2017-2021 was 428 after the literature search.

The fourth chapter is about the classification after the inclusion and exclusion criteria and the process which made the selected papers went through three processes of screening. Once the screening finished, the final sample was of 56 papers.

The fifth section goes over the selected studies (n=56) by displaying the classification schema and results. It presents analyzed and structured results including the bubble plot visualization as well as empirical conclusions.

The sixth chapter presents the discussion of the findings. Finally, the last chapter presents the conclusion. The research questions are compared to the results, limitations of the research as well as possibilities for future research are mentioned.

2 Background

This chapter intends to provide an overview of the current primary themes related to this topic and the present state of AI ethics in healthcare in the academic literature. Furthermore, this section aims to provide the necessary knowledge to understand this thesis in more detail. First, it explores the history of AI in medicine followed by the types of AI used in healthcare, including their benefits and disadvantages. Next, AI Ethics is explained as well as the ongoing ethical issues related to healthcare. Finally, this section ends with a summary of the presented theory.

2.1 History of AI in Medicine

One of the first to mention the concept of AI was Alan Turing in 1950. He defined it as the ability to stimulate critical thinking in computers to achieve cognitive tasks. (Amisha et al., 2019) Although AI has various interpretations, Guan (2019) defined it as "giving human intelligence to a physical or virtual machine". Despite starting with simple "if, else" statements, it evolved to include complex algorithms to imitate the human brain and can take different forms in technologies such as Machine Learning (ML), Natural Language Processing (NLP), Computer Vision (CV), and Deep Learning (DL). (Kaul, Enslin, & Gross, 2020) These are explained in detail later in this section.

In the last decades, AI has been more present and accepted in medicine due to the progress of DL and ML. (Kaul et al., 2020) As modern medicine is confronted with collecting, analyzing, and applying an enormous amount of data, (Ramesh et al., 2004) predictive tools can be used by clinicians for diagnosis and prediction of therapeutic response, and preventive medicine. (Le Berre et al., 2020) These tools supported by AI provide more accuracy, efficiency in the workflow and clinical operations, and facilitating patients' monitoring and outcomes. (Kaul et al., 2020)

Additionally, the progress of AI in healthcare has not been linear. According to Kaul et al. (2020), it is can be categorized by periods: 1950s-1970s, 1970s-2000s, and 2000s-today. The first category contains the beginning of AI. During that time, developers were only interested in developing machines able to display critical thinking. (Kaul et al., 2020) In terms of technologies, that period contains multiple innovations. The first industrial robot arm was created to help the assembly line at General Motors (Moran, 2007). Also, a new chatbot called Eliza was introduced and it used NLP to imitate an online human discussion (Weizenbaum, 1966) as well as, Shakey, the first mobile robot, was presented and it was able to comprehend instructions. (Kuipers et al., 2017)

The following period, the 1970s-2000s, is known as “AI Winter” as interest and funding greatly reduced during these years. (Kaul et al., 2020) Despite this, pioneers in the field still collaborated and created different AI tools. For example, a consultation program for glaucoma using the CASNET model was created to provide advice on patient management given a specific disease based on its database. (Weiss et al., 1978) It also became possible to use computer analysis in diagnosing strong abdominal pain. (Ramesh et al., 2004) MYCIN was appraised for being able to provide a list of potential pathogens and the correct antibiotics according to the patient’s case. (Kulikowski, 2019) Finally, DXplain was created to provide possible diagnostics based on given symptoms (Amisha et al., 2019) In healthcare, different AI tools started to emerge in clinical settings such as fuzzy expert systems, Bayesian networks, artificial neural networks, and hybrid intelligent systems. (Amisha et al., 2019)

Finally, the last period contains the most impressive advancements for AI as NLP and DL evolved. Some technologies included during that time are IBM Watson, virtual assistances, Pharmabot, Mandy, and Convolutional Neural Network (CNN). IBM Watson is a supercomputer that uses DeepQA, a mix of NLP and search algorithms, to provide answers to any question, (Ferrucci et al., 2013) Different virtual assistances that use NLP were introduced to society such as Siri from Apple and Alexa from Amazon. Pharmabot was a chatbot used to help children and parents with medication (Comendador et al., 2015) and Mandy was a chatbot used to discuss with a patient to assess their needs and forward them to medical experts. (Ni et al., 2017) Finally, CNN was developed to be used in image processing classification. (Hoogenboom et al., 2020) Additionally, in 2016, healthcare applications had received the most funds compared to other sectors. (Amisha et al., 2019) To resume, Figure 2 provides an overview of the development and use of AI in medicine.

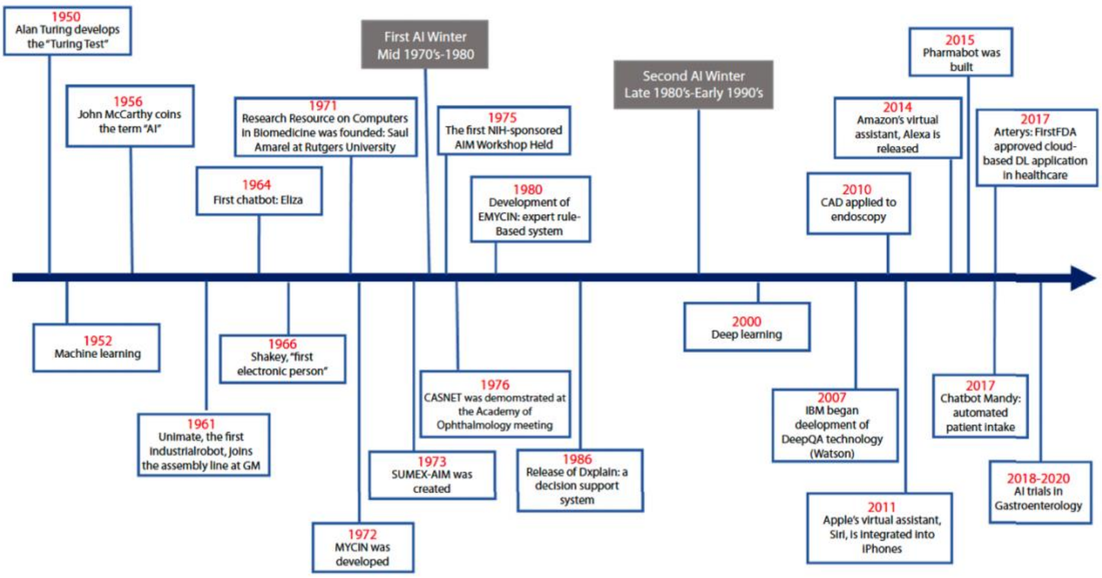

Figure 2. Timeline of the development and use of AI in medicine from Kaul et al. (2020)

2.2 Types of AI in Healthcare

AI can take several forms in healthcare. It can be virtual, physical, or a mix of both. Virtual AI includes ML, NLP, rule-based expert systems, and robot process automation. Physical AI includes physical robots and brain-computer interfaces (BCIs). (Guan, 2019) Additionally, most AI technologies mentioned in the following chapter use DL algorithms.

2.2.1 Machine Learning, Artificial Neutral Network & Deep Learning

First, ML can be defined as an AI field where a computer program can learn from training models with data to perform tasks without receiving explicit instructions. (Dalal, 2020) It is the most common approach of AI and has different levels of complexity. (Davenport & Kalakota, 2019) In healthcare, ML has many applications and uses. Usually, precision medicine utilizes traditional ML, and disciplines like radiology, oncology (Davenport & Kalakota, 2019), genetics, and molecular medicine require more complex forms of ML. (Guan, 2019) Traditional ML differs from DL as it regroups different methods such as regression, trees, cluster, and classification. Traditional ML is based on a strict set of rules to provide results while DL uses neural networks. (Paterakis et al, 2017)

While ML regroups diverse approaches and techniques, Artificial Neural Networks (ANN) is its most popular one in medicine. (Ramesh et al., 2004) Naraei et al. (2016) describe ANN as a tool used for data classification. It is composed of interconnected computer processors able to perform parallel computations for data processing and knowledge representation. (Ramesh et al., 2004) ANN is very versatile and can conform to any given data. (Naraei et al., 2016) It learns from historical examples and its own experience. Then, it proceeds to analyze unrelated information, handle unclear knowledge, store the general outcome model and apply it to another set of data. (Ramesh et al., 2004) It helps the computer program to learn.

Various domains in healthcare use ANN such as in the clinical diagnosis, image analysis in radiology and histopathology, data interpretation in an intensive care setting, waveform analysis, diagnosing cytological and histological specimens, analyze cancer data (Ramesh et al., 2004), for smart health records, and crowdsourcing data. (Dalal, 2020) For many researchers and medical experts, ANN helps to find and identify intricate relationships between variables in a complex setting that they could not have found without ANN. Its main issue is to use pre-existing information that can potentially contain any human bias. (Ramesh et al., 2004) Similarly, DL is another technique of ML. (Davenport & Kalakota, 2019) It can be explained as:

> *"A form of representation learning—in which a machine is fed with raw data and develops its own representations needed for pattern recognition—that is composed of multiple layers of representations. These layers are typically arranged sequentially and composed of a large number of primitive, nonlinear operations, such that the representation of one layer (beginning with the raw data input) is fed into the next layer and transformed into a more abstract representation. As data flows through the layers of the system, the input space becomes iteratively warped until data points become distinguishable. In this manner, highly complex functions can be learned." (Esteva et al., 2019)*

A great benefit of DL is to be able to multitask. It can run on large datasets while continuously improving the data gathered. Also, it can take different types of data as input; thus, DL outperforms many ML technologies. (Esteva et al., 2019) Figure 3 provides a visualization of how DL transforms different sources of information into results.

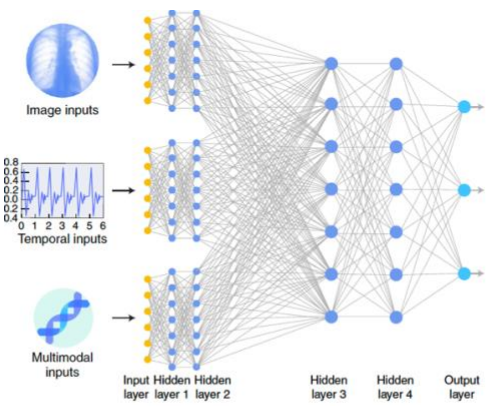

Figure 3. Example of Deep learning from Esteva et al., (2019)

In healthcare, DL is used in various specialties such as radiology in pattern imaging analysis, speech recognition, diagnosis (Davenport & Kalakota, 2019), the discovery of drugs and manufacturing, personalized medicine, and many others. (Dalal, 2020) Unfortunately, one of its issues is explaining its reasoning behind the obtained result because it is almost impossible for developers and medical experts to do so. (Davenport & Kalakota, 2019)

2.2.2 Natural Language Processing

NLP is defined as a way for computers to comprehend human language. Also, it is utilized in different fields such as speech recognition, text analysis, and many more. (Davenport & Kalakota, 2019) NLP proceeds by transforming writings into machine-readable structured data. It can do so by using ML methods and its algorithms. (Jiang et al., 2017) For example, Recurrent Neural Networks (RNN) is a type of DL algorithm effective at processing sequential inputs like language, speech, and time-series data. (Sutskever et al., 2014)

In recent years, many successes were attributed to NLP such as machine translation, text generation, and image captioning. (Esteva et al., 2019) In healthcare, it is commonly used for the creation, understanding, and classification of clinical documentation and published research. Also, it can analyze unstructured clinical notes on patients, prepare reports, transcribe patient interactions and conduct conversational AI. (Daven-port & Kalakota, 2019) The combination of using DL and language technologies allows the creation and sustainability of domain applications such as Electronic Health Records (EHR). (Esteva et al., 2019) EHR is gaining popularity and is becoming omnipresent. It is evaluated that within a decade, the EHR of a large medical organization can comprehend up to 10 million patients' medical transactions that each produces a maximum of 150,000 bits of data. (Shickel et al., 2017) It is significant progress for medical experts as it represents 200,000 years of doctor knowledge and 100 million years of patients' information. (Rajkomar et al., 2018)

There are also chatbots used for patient interaction, mental health and wellness, and telehealth. (Davenport & Kalakota, 2019) Soon, many believe that automatic speech recognition (Shickel et al., 2017) and information extraction technologies combined will create reliable clinical voice assistants that will be able to take notes of patients' visits. This improvement would allow doctors to reduce time on documentation and increase time spent with patients. (Esteva et al., 2019) Patients have expressed one issue regarding the probability of chatbots revealing confidential information, complex health conditions, and poor usability. (Davenport & Kalakota, 2019)

2.2.3 Others

In this section, several AI technologies are briefly presented including rule-based expert systems, physical robots and BCI, and Computer Vision (CV).

Rule-based expert systems are established on a collection of "if-then" rules that require experts' knowledge in a particular field to set those boundaries. (Davenport & Kalakota, 2019) They are mostly used concerning clinical decision support for heart failure diagnosis and treatment plans. One main problem with those computerized systems is the lack of guidelines to provide automated decision support and alerts. (Seto et al., 2012)

Robot Process Automation (RPA) is an inexpensive program that helps to automate digital assignments easily. (Davenport & Kalakota, 2019) In healthcare, RPA holds many benefits such as increasing efficiency, providing support to the front desk, improving data privacy, and being cost-effective. (Ratia et al., 2018) In the same domain, it is used generally for administrative and repetitive tasks like prior authorization, updating patient records, and billing. (Davenport & Kalakota, 2019)

Physical robots are set to perform specific tasks and to be care or surgical robots. (Davenport & Kalakota, 2019) They are usually used for elderly care and in medical procedures. AI Robots have different problems like the technology not being advanced enough to achieve their goals, its robustness, and several legal and ethical issues. (Turja et al., 2017)

BCI is a system that receives, decodes, and interprets brain signals to a given output such as a device or feedback to the user. Their primary function in healthcare is to improve patient's lives suffering from neurological disorders. (Guan, 2019) It is essential to understand that both end-users and BCI form a team. As the user generates data, the BCI can start decoding once the training dataset is completed. It is usually used in healthcare to improve a disabled person's day-to-day life. Its main problems are related to the privacy and confidentiality of patients. (Shih et al., 2012)

Finally, CV is a tool that can analyze images and video by using classification, detection, and segmentation. It is mainly used for medical imaging for diagnosis in dermatology, radiology, ophthalmology, and pathology. (Russakovsky et al., 2015) CV can achieve this by applying Convolutional Neural Networks (CNN) which is a form of DL that evaluates the information that expresses natural spatial invariance. (Esteva et al., 2019) In simple words, the CNN process can be divided into two steps when breaking down a picture: first, it learns the natural statistics (lines, curves, colors, etc.) in the image by allowing its algorithm to process large quantities of information; second, its algorithm analyzes higher-level layers to find similarities between learned diagnostics. (Choi et al., 2017) CV also faces different issues such as lack of clinical context and difficulties to obtain large labeled datasets. (Ronneberger et al., 2015)

To summarize the various technologies and approaches in AI, Figure 4 presents the different relationships between them. DL and ANN are part of ML that is a subset of AI. NLP, voice recognition, CV, robotics & motions are different AI technologies that also employ ML with their respective algorithms. (Merkell, 2020)

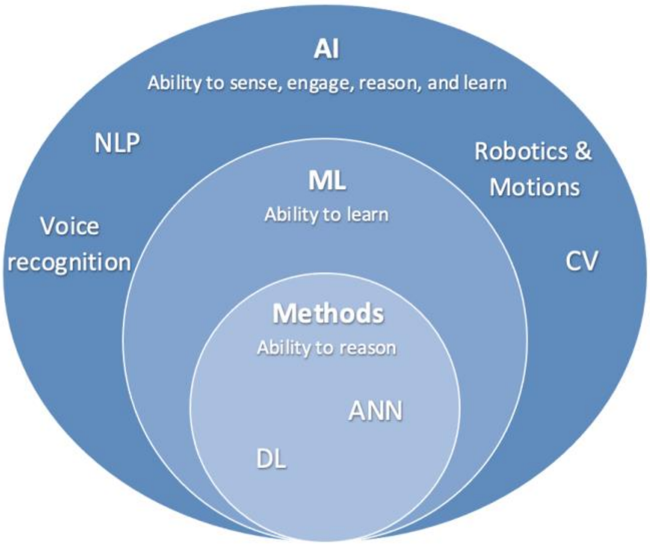

Figure 4. Overview of AI technologies' structure based on Merkell (2020)

2.3 Ethics

As new AI advancements previously mentioned emerged, the public and policymakers' interest grew stronger regarding them. (Jameel et al., 2020) AI tools have clear benefits as well as disadvantages. For example, they create new ethical issues and challenge current norms such as transparency issues, but they also can assist society in simple or more complex tasks. (Müller, 2020) Despite the AI's potential to solve complex problems, few publications discussed ethical issues involved using AI. Some papers mentioned and proposed machine ethics as a solution, but many criticized it. (Vakkuri & Abrahamsson, 2018) As society wants to exploit AI to its fullest and improve daily life, AI will require to follow certain essential ethics as human individuals to limit most possible accidents. It is noteworthy to mention that AI has the ability to make decisions that can have an ethical impact. (Jameel et al., 2020) Thus, AI ethics has gained momentum recently to provide possible guidelines to these new issues.

As it involves different domains such as computer science and philosophy, the academic discussion about AI ethics has been diverse. Therefore, establishing a unified definition for it is quite challenging. (Vakkuri & Abrahamsson, 2018) In the author's opinion, AI technologies must follow some guidelines in terms of ethics, as too many ambiguous areas can lead to burdensome issues. The European Parliament also supports such statement as it has published the Civil Law Rules on Robotics: European Parliament resolution of 16 February 2017 with guidance to AI in healthcare. (Gerke et al., 2020)

Furthermore, ethics is defined as a study in philosophy that attempts to classify things into the notions of right and wrong and to help one develop its morality. (Velasquez et al., 2010) Ethics is divided into the following:

1. Normative ethics, which includes theories about what we should do and why.
2. Metaethics, which is more focused on ethics theories themselves.
3. Applied ethics, which includes how to use normative theories to given issues. (McCartney, 2015)

The last category is usually work- or organization-related. Thus, one of its sub-fields is computer ethics, which itself contains AI ethics. (McCartney, 2015) In the last two decades, information and computer ethics have merged. (Floridi, 2009) Computer ethics is defined as theories evaluating the nature and social impact of computers and reasoning ethical policies behind their uses. These theories are part of a complex and dynamic field because computers evolve each year with newer technologies. (Moor, 1985) While information ethics studies ethical issues behind the validity, availability, and accuracy of online information. (Floridi, 2009)

As mentioned previously, in this thesis, AI ethics is seen as a sub-study of computer and information ethics. (Moor, 2006) It contains different theories such as machine ethics which has been researched quite a lot such as Anderson & Anderson's study (2007). Compared to computer and information ethics that focuses on how individuals use a machine, machine ethics studies machines' behavior towards human and machine users. (Anderson & Anderson, 2007)

The essential point of machine ethics is that machines are implicit and explicit ethical agents. This is useful to understand the relevancy of AI machines as stakeholders in this thesis. First, implicit because they have software inside them to avoid potential unethical behaviors. Second, explicit since they can make the best choice in case of a moral issue. (Moor, 2006) Hence, to perform at their best ability and following ethical guidelines, AI machines will need to have moral guidelines, and this is where AI ethics becomes critical. In this thesis, machine ethics is a sub-part of AI ethics. Below, Figure 5 summarizes how all the mentioned ethics are related to each other.

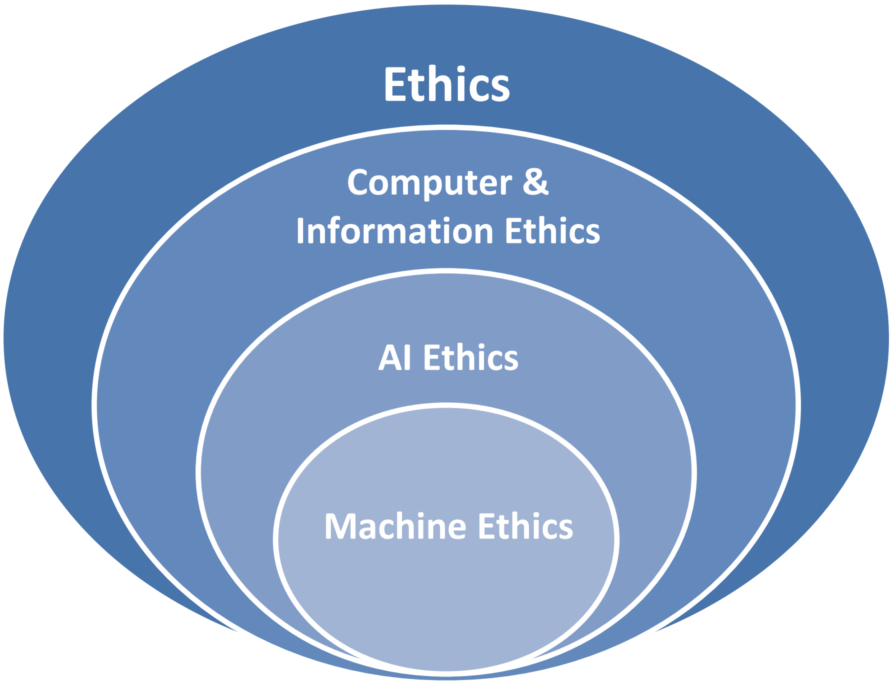

Figure 5. Summary of Classification of Mentioned Ethics

The field of AI is very active, in constant evolution, and involves various domains such as computer science, mathematics, information science, and others. (Russell, & Norvig, 2015) Therefore, defining AI ethics can be quite complex and limiting, but as AI tools are gaining more importance within society, a need for ethical guidelines emerges. for this thesis, AI ethics will be described as a term, used in AI sub-sectors, as a response to ethical problems in terms of causes, consequences, and possible solutions. AI ethics has many research fields such as explainable AI (XAI), responsible AI, and machine ethics.

Thus, developers should aim to limit ethical issues to create moral, fair, and safe AI applications by considering AI ethics guidelines. (Leslie, 2019) Moreover, AI ethics seems to have many points of view, but the five recurring central themes throughout different studies including the research of Jobin et al. (2019), Reddy et al. (2019), and Davenport & Kalakota (2019), with recurrent similar definitions:

1. Transparency (which involves XAI)
2. Justice & Fairness
3. Security
4. Accountability & Responsibility (which involves responsible AI)
5. Privacy

Transparency is described as explainability, understandability, and interpretability. Justice and fairness are defined as consistency, inclusion, equality, equity, non-bias, non-discrimination, diversity, plurality, accessibility, reversibility, remedy, redress, challenge, access, and distribution. Security is defined as non-maleficence, safety, harm, protection, precaution, prevention, integrity (bodily or mental), non-subversion. (Jobin et al., 2019) Responsibility is explained as accountability, (Davenport & Kalakota, 2019) liability, and acting with integrity. (Jobin et al., 2019) Privacy is characterized by personal or private data. (Reddy et al., 2019)

For example, some will use a guideline called FAST that regroups fairness, accountability, sustainability, and transparency. Developers usually use it in AI projects to maintain the ethical aspect. Figure 6 provides a brief understanding of each point.

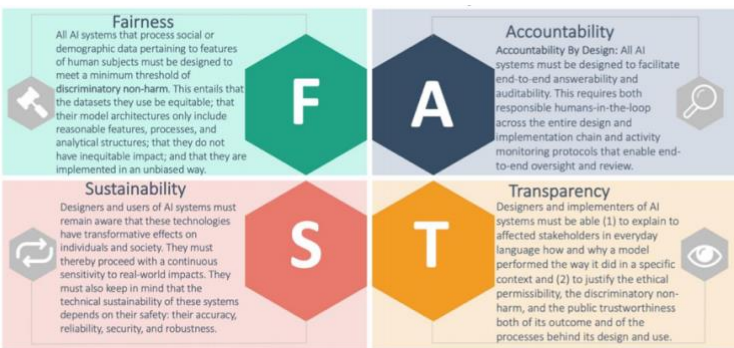

Figure 6. FAST Theorem from Leslie (2019)

2.4 AI Ethics in Healthcare

As the AI field keeps evolving, AI tools, such as decision-support systems, are slowly replacing and amplifying human cognitive activities in diverse fields. With this, growing concerns are emerging on how to ensure these systems can act within a certain set of values that are aligned with its users, developers, and society. (Dignum, 2020) As Bartoletti (2019) mentioned, healthcare is considered one of the most attractive and promising fields for AI technologies. For example, medical experts are now using AI imaging to detect cancer faster and earlier than before. However, since healthcare takes care of people's health, any technology of this domain must comply with laws, regulations, and privacy rules. (Bartoletti, 2019) In general, AI technologies can generate various ethical issues in healthcare, such as AI bias, privacy issues, patient-clinician trust issues (Reddy et al., 2019), transparency, accountability, and permission problems. (Davenport & Kalakota, 2019) Despite them, AI technologies have the potential to democratize expertise, globalize healthcare, and make healthcare available in remote areas. (Gerke et al., 2020) AI ethics aims to highlight some of these problems for medical experts, developers, and entities to find possible solutions. This thesis focuses on the four issues mentioned in the previous chapter: transparency, justice and fairness, accountability and responsibility, privacy and security (Jobin et al., 2019), and their possible solutions.

2.4.1 Transparency

Transparency is defined as an information exchange between a receiver and an object, where the object is in charge of giving the results of an operation to the receiver. (Woudstra, 2020) This concept is essential as it allows to create and maintain trust amongst stakeholders. In healthcare, the trust between medical experts and patients is crucial to ensure a successful implementation of AI. (Gerke et al., 2020) Unfortunately, many AI algorithms, especially ML and DL, are near impossible to interpret or explain by developers and medical experts. (Whittaker et al., 2018) This problem, also known as the black-box issue, appears when AI leads to opaque decision-making processes. Then, patients' trust can decrease and lead to transparency issues. Equally, overreliance on this technology can reduce the discussion and contact between clinicians and patients and create transparency concerns. (Reddy et al., 2019) In AI ethics, the transparency aspect creates concern with putting into place and maintaining a framework for defining various types of transparency and for the audition of algorithms. (Weller, 2017) Additionally, XAI is also part of the transparency aspect in AI ethics and can be explained as an AI tool able to provide a report regarding the algorithm responsibility between stakeholders. (Gunning & Aha, 2019) It consists of two main aspects: to be able to provide human-readable explanations on its intent, reasoning, and decision-making process and to be able to pinpoint whose responsibility it is in case of a bias for example. (Miller, 2019)

For example, Corti is an AI software that uses ML to assist emergency dispatchers in making decisions during a cardiac arrest. Its algorithms are considered "black box" as even its inventor cannot explain or deduce how the software has reached its conclusion. (Gerke et al., 2020) As well as reducing trust, AI models can impair the recommendations given by the technology and the identification of any biases. (Reddy et al., 2019) Therefore, transparency and fairness go hand in hand. As the AI machine learns from a data set, it takes it as the truth and cannot detect biases. The quality of the given dataset is crucial because the program will reproduce that flaw. (Gebru et al., 2020) Transparency and accountability problems also go hand in hand as understanding AI technology's thought process is near impossible. For example, who will be to blame if an incident happens: the technology, the developer, or the user. (Davenport & Kalakota, 2019)

To summarize, the lack of explainability of results provided by AI will decrease the trust of patients towards medical experts as well as the trust of medical experts in AI. (Reddy et al., 2019) Moreover, it will impair the detection of any bias. (Gerke et al., 2020) A solution currently used to decrease these transparency issues is XAI. (Reddy et al., 2019) Although, the technology is not expected to provide the detailed reasoning behind its decision. (van Lent et al., 2004) XAI aims to increase transparency, to be able to trace the given result, and to improve the AI model. (Dave et al., 2020)

2.4.2 Justice & Fairness

Justice and fairness are defined as consistency, inclusion, equality, equity, non-bias, non-discrimination, diversity, plurality, accessibility, reversibility, remedy, redress, challenge, access, and distribution. (Jobin et al., 2019) In this context, justice and fairness are covered by exploring the data consistency and inclusion, and the potential biases. In healthcare, AI data analysis is used to make predictions. There are three types of concerns regarding AI: the given data had was unfair, human cognitive bias such as intuitive judgment, and statistical bias such as when data exhibits a systematic error. Usually, these biases happen when unfair conclusions are made by the influence of irrelevant aspects to the matter. (Gerke et al., 2020) In AI ethics, the fairness aspect is a concern as the algorithm needs to be equally efficient for all involved users without introducing in the future possible discrimination, especially regarding decision-support systems. (Mehrabi et al., 2019)

Data biases are divided into three categories: behavioral bias, which is about content sharing and news spreading; population bias, which is about different gender, ethnicity, age, etc.; and linking bias, which is the different influences on the study during data collection. (Jameel et al., 2020) AI biases arise when the data used in training AI models, is not representative of the target population, is inadequate, or incomplete. It can lead to an over or under-estimation of risks as overestimating risks of criminal recidivism for a racial group. (Reddy et al., 2019)

If an ML tool would be given a biased database, it would fail to recognize the issue and would codify and automate it. (Gebru et al., 2020) For example, a decision support system aims to assist medical experts in identifying the best treatment for patients with skin cancer, but it has only trained with a database based on Caucasian patients. It can lead to the tool having issues suggesting recommendations because of the labeled data being under-inclusive. Another example, IBM Watson for Oncology runs AI algorithms to evaluate data from existing medical records and provide medical experts with possible treatment recommendations for the given patient. Recently, it was accused of providing inaccurate cancer treatments during test cases as its labeled dataset was composed of a few created cancer cases. Therefore, given datasets must be reliable and accurate because the better will the labeled data be, the better the AI technology will behave. (Gerke et al., 2020) Despite the difficulty of finding a big labeled dataset, developers need to be aware of such bias when they attempt to minimize them at all stages of product development. They should consider the risks of biases when selecting the ML technologies and the dataset's quality and diversity. (Gerke et al., 2020)

2.4.3 Accountability & Responsibility

Accountability and responsibility are important issues in AI ethics, especially in healthcare. There is a difference between the two terms. According to Dignum (2020), responsibility refers to developers' duty to develop an accurate and ethical AI technology, to educate on how to use it correctly, medical experts' usage of the tool, and the AI machine's capabilities of providing answers and errors. Accountability refers to one responding for their action and is related to liability. For example, who would be accountable if a self-driven car hits a pedestrian? (Dignum, 2020) Thus, it is difficult to establish accountability for AI systems. (Davenport & Kalakota, 2019) Responsibility is associated with autonomy and personhood. In AI ethics, some systems have a certain level of technical autonomy without questioning responsibility. (Alexander & Ripstein, 2001)

Regarding AI ethics, the responsibility aspect is focused on algorithmic accountability. Wieringa (2020) defines it as an accountability relationship in which one individual provides a statement for the algorithm that they may or may not have created. That one individual can be anyone involved in the making and deployment of the algorithm. (Wieringa, 2020) Currently, there is a global approval that accountability, liability, and the rule of law are basic requirements that new technologies should take into account. In the case of robots, it has not yet been agreed on how responsibility and accountability should be applied. (Coeckelbergh, 2010) In Europe, the European Parliament published the Civil Law Rules on Robotics: European Parliament resolution of 16 February 2017 with guidance to AI in healthcare. This resolution challenges the legitimacy of present liability rules and maps the accountability of emerging digital technologies such as AI. (Gerke et al., 2020) In the United States, if a medical expert would use AI technology and an incident would happen with a patient, they would be held accountable. It is considered that the clinician should only use the AI tool as a recommendation, and they are the ones making the final decision. (Gerke et al., 2020) Therefore, to avoid this problem, physicians should adopt it as a confirmatory tool instead of simply following the recommendation. Also, some suggested product liability against the developers in case of misdiagnosis. It would require stricter accountability of the manufacturer for defects. (Gerke et al., 2020)

Some solutions proposed through different research were to identify the appropriate stages (approval, introduction, and deployment) for which monitoring and evaluating are critical to ensure the safety and quality of AI-enabled services, (Reddy et al., 2019) and to keep the current medical malpractice regulation that aims to meet deterrence and compensation of the victims. For example, vaccine manufacturers place money in a fund and the system automatically pays those harmed by the vaccines. AI manufacturers could follow a similar procedure to compensate patients. (Gerke et al., 2020)

2.4.4 Privacy & Security

Privacy and security are crucial notions in AI, especially for patients. Privacy is defined as "the right to be let alone", information privacy, privacy as an aspect of personhood, control over information about oneself, and the right to secrecy. As the digital world is now omnipresent, all data collection and storage are also digital that can later become an issue. (Müller, 2020) The usage of AI health apps and chatbots increases; one can now use a wearable device to collect data from steps to heartbeat measures. (Gerke et al., 2020) While AI increases smart data collection and analysis, the value of medical information reaches up to billions of dollars. (Gerke et al., 2020) Hence, the public has become wary of data collection, unethical use of data, and transparency issues (Bartoletti, 2019), and documentation indicates that society is troubled by companies or governments selling individual data for revenue. (Gerke et al., 2020) Unfortunately, it is complicated to control who is collecting information in the digital sphere. (Whittaker et al. 2018) For example, the Royal Free NHS Foundation Trust was accused of a privacy breach because participants were not properly informed during a clinical safety testing that their data was shared with Google DeepMind. It was an exchange between the two companies, so one obtained real labeled data, the other used DeepMind for free for five years. (Gerke et al., 2020)

Privacy is crucial for patients as it is bound to their autonomy, personal identity, and well-being. Patients are concerned that even anonymized data could be reidentified with few data points. (Reddy et al., 2019) Sometimes, patients' data is collected without their awareness of its final purpose. Explicit consent from the patients is essential. (Gerke et al., 2020) It will be essential for stakeholders to understand the difference between personal data and sensitive information. (Bartoletti, 2019) As well, genetic privacy puts at risk not only one person but anyone related to that individual. (Gerke et al., 2020) Privacy breaches can happen at any moment if the system has not proper security, hence why security and privacy are strongly related. (Reddy et al., 2019)

Thus, all artificial intelligent systems should be equipped against privacy breaches to avoid any psychological and reputational harm to patients. (Reddy et al., 2019) Also, stakeholders should review when informed consent is required in healthcare. (Gerke et al., 2020) Bartoletti (2019) suggest that developers follow a clear set of steps for the deployment of algorithms:

- Data Privacy Impact Assessments to verify the possibilities of privacy issues.
- Algorithmic Impact Assessments to protect labeled datasets from bias.
- Maintain audit trails to trace who is doing what, which data is used, and what changes are made.
- Procurement law in healthcare to certify that bought AI systems follow strict procedures such as how the dataset was trained and if they have been analyzed and assigned a trust mark.

2.5 Conclusion

To summarize, the theory explained previously concerned the history of AI in medicine, the types of AI used in healthcare, and AI ethics. They aimed to provide the reader a better picture of what is currently happening in this field.

The first section covered the evolution of AI in terms of technologies in medicine. The notion of AI was first mentioned by Alan Turing in 1950 (Ramesh et al., 2004) and it began with a simple “if, else” rules and then, evolved into complex algorithms able to mimic human reasoning. In healthcare, AI has many uses such as diagnosis, prediction of therapeutic response, image processing, and preventive medicine (Le Berre et al., 2020) as well as many benefits, for example, providing more accuracy, efficiency in the workflow and for clinical operations, and facilitating the patients monitoring and outcomes. (Kaul et al., 2020) It highlighted AI's non-linear growth rate and the rise of interest in AI in the last decade.

The second part presented the different forms of AI used in healthcare. ML, DL, ANN, NLP, decision support systems, RPA, physical robots, BCI, and CV were explored briefly. It looked into how they worked and their benefits and disadvantages. ML is the most common approach to AI (Davenport & Kalakota, 2019) and contains ANN and DL, which are algorithms used to help the computer program learn. (Ramesh et al., 2004) Its crucial issue is to be based on historical information that can have any human bias. (Ramesh et al., 2004) Then, NLP uses ML in combination with other algorithms (Merkell, 2020) to help with documentation. (Daven-port & Kalakota, 2019) It allows medical experts to reduce time on administration and increase time spent with patients. (Esteva et al., 2019)

Finally, the last section covered AI Ethics. Generally, AI technologies generate many ethical issues in healthcare like transparency, justice and fairness, accountability and responsibility, and privacy and security issues. (Jobin et al., 2019) For transparency, the lack of explainability of results provided decreases patients' trust towards medical experts and of medical experts in AI (Reddy et al., 2019). It will impair the detection of any bias. (Gerke et al., 2020) For justice and fairness, the trained dataset should not contain any bias. They will need to follow strict regulations to ensure their validity and accuracy. (Jameel et al., 2020) For accountability and responsibility, it is about who will be held liable in case of an incident. For privacy and security, it is about a privacy breach, data collection, and consent from patients. Currently, there are entities such as the Institute of Electrical and Electronics Engineers (IEEE) and the British Standards Institution (BSI) that have established standards, especially on technical issues like data security and transparency. (Müller, 2020)

3 Literature Search for Primary Studies

In recent years, society has begun to question extensively moral problems regarding the use of AI. Therefore, many have begun more intensive research regarding them. As it is a relatively new field, it can lack information, clarity, and structure. A clear proposition to reduce these ethical issues in the research domain has yet to be found as most research focus on highlighting issues without providing concrete solutions.

As mentioned in the "Introduction" chapter, the thesis is conducted using SMS. It provides a broad view of the current state of the academic literature. Moreover, Figure 7 presents SMS steps as well as the outcomes adapted from Petersen et al. (2008). Since this thesis explores a new perspective of AI ethics in healthcare, SMS results observed evaluate the different stakeholders and how current ethical issues in healthcare are managed. This chapter covers the SMS methodology applied for this topic and the research process.

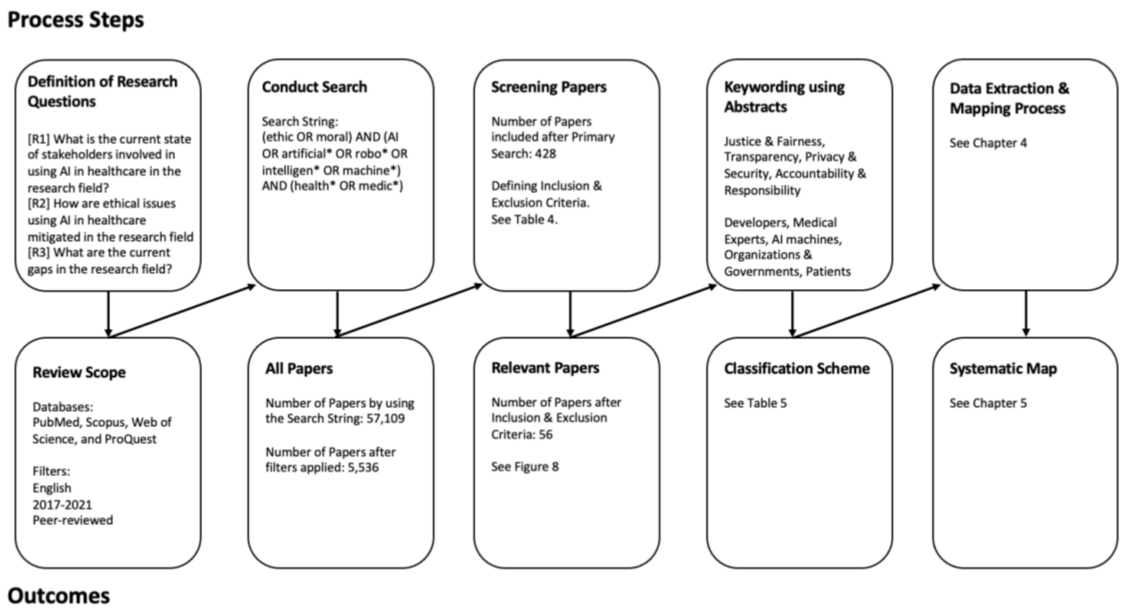

Figure 7. SMS Process based on Petersen et al. (2008)

3.1 Research Questions & Research Process

As this industry is constantly expanding, AI's ethical impact in healthcare is becoming a growing concern for society. Already, some distrust medical experts due to the lack of transparency as explaining AI machines' results is nearly impossible. Patient and clinician trust is crucial because the public needs to trust healthcare experts and machines. (Reddy et al., 2019) Thus, identifying the source of these ethical issues and their solutions is relevant for the day-to-day and academic spheres.

A study is necessary to have a comprehensive view of AI ethics in healthcare. The SMS results present the quantity and type of the relevant literature reviewed and the current gaps in the academic literature. (Petersen et al., 2008) Figure 7 highlights the different steps of the research method and they are cumulative. Therefore, they must be done in the correct order and explained once completed.

The main research question of this thesis is: *What is the current state of ethical issues by using AI in healthcare?* is divided into:

> *[R1] What is the current state of stakeholders involved in using AI in healthcare in the research field?*
>
> *[R2] How are ethical issues using AI in healthcare mitigated in the research field*
>
> *[R3] What are the current gaps in the research field?*

The objective of these research questions is to understand the present state of AI ethics in healthcare and the current gaps in the literature. Therefore, the literature must be related to AI ethics and healthcare, and the following steps mentioned in Figure 7 are applied to the relevant papers. Petersen et al. (2008) mentioned that SMS does not value the quality of the articles. Thus, the number of papers can be quite large at first glance. (Petersen et al., 2008) In this thesis, peer-reviewed papers will ensure the quality of the literature. After the process was completed once, there was a total of 428 from the four databases. Once these papers were analyzed and passed through the inclusion and exclusion criteria, 56 remained. The following sections cover the SMS process in more detail.

3.2 Primary Search

This section presents the literature search that includes the search strings process. When looking at Figure 7, the following step after establishing the research questions is to conduct the research and form search strings. Since a global overview is needed, the primary inclusion and search strings cannot be too narrow, and the search must also include different databases. (Petersen et al., 2008) A manual screening must be done because the search strings' results across the various databases still contain irrelevant literature.

The formulation of search strings is crucial as it will define the result of the primary search. It should be done in a way to maximize the number of papers. One methodology mentioned in the study of Kitchenham et al. (2011) is PICO. According to James et al. (2016), it is defined as:

- Population, which refers to the subject of the research;
- Intervention, which refers to what is impacting the subject;
- Comparison, which refers to a similar subject; and
- Outcomes, which refer to the search results related to the subject.

For this thesis, the population is defined as all papers related to AI and healthcare. The intervention is AI ethics which was the focus; no comparison was used. The outcome is to view the current state of academic literature; hence, only peer-reviewed articles were included. If PICO is applied to the central question of this thesis, "*What is the current state of ethical issues by using AI in healthcare?",* the keywords here are ethics, AI, and healthcare. Following are synonyms for each to increase the search:

- Ethics: moral, ethic
- AI: AI, artificial, robotic intelligent, machine
- Healthcare: health, healthcare, medicine, medicare

Therefore, the final string is:

- (ethic OR moral) AND (AI OR artificial* OR robo* OR intelligen* OR machine*) AND (health* OR medic*)

Additionally, the search strings were limited to the document title and abstract. Thus, the number of irrelevant results decreased. For this thesis, four databases were selected: PubMed, Scopus, Web of Science, and ProQuest. PubMed was selected as it is one of the largest medicine-related databases. Scopus, Web of Science, and ProQuest were chosen as they are large multidisciplinary centers of literature.

At first glance, there was a total of 57,109 papers across four centers of information. Then, the inclusion criteria were applied to reduce the amount of literature. They were the publication date (2017-2021), language (English), and type of publication (peer-reviews). Since the field of AI in healthcare has been growing rapidly in the last decade, the search focuses on papers published after 2017. As mentioned in the "History of AI in Medicine" section, it is in 2017 that the first FDA DL application was approved for healthcare (Kaul et al., 2020) and that the European Parliament established the Civil Law Rules on Robotics: European Parliament resolution of 16 February 2017 which included guidance to AI in healthcare. (Gerke et al., 2020) After the three filters were applied to each database, the number of papers was narrowed to 5,536. Table 1 presents the number of results retrieved, followed by the three filters to each database.

Date	Database	Search String	Before Filters	Language	Document Type	Year
05.04.2021	Scopus	TITLE-ABS-KEY(AI OR artificial* OR robo* OR intelligen* OR machine*) AND TITLE-ABS-KEY(health* OR medic*) AND TITLE-ABS-KEY(ethic* OR moral*)	8556	7603	1239	399
05.04.2021	ProQuest	noft((AI OR artificial* OR auto* OR intelligen* OR machine* OR robo*)) AND noft((ethic* OR moral*)) AND noft((health* OR medic*))	21588	20926	10093	3995
05.04.2021	PubMed	((ethic*[Title/Abstract] OR moral[Title/Abstract]) AND (AI[Title/Abstract] OR artificial*[Title/Abstract] OR robo*[Title/Abstract] OR intelligen*[Title/Abstract] OR	1724	1543	353	194
05.04.2021	Web Of Science	(TS=((AI OR artificial* OR auto* OR intelligen* OR machine* OR robo*) AND (ethic* OR moral*)AND (health* OR medic*)))	25241	21720	3960	948
Total			57109	51792	15645	5536

Table 1. Results after Filters

Table 2 presents the results of the primary search that includes the search strings, the databases, and the number of papers from 2017-2021. It also presents the filtered results, followed by the related papers that met the inclusion criteria and the last column, the number of papers without duplicates. The manual screening removed papers that only included keywords in the abstract, that were not available in full text online, and that were not related to the field of AI ethics and medical ethics.

Database	Search String	Total Papers	After filters	After Inclusion Criteria	Duplicate Removal
Scopus	TITLE-ABS-KEY(AI OR artificial* OR robo* OR intelligen* OR machine*) AND TITLE-ABS-KEY(health* OR medic*) AND TITLE-ABS-KEY(ethic* OR moral*)	8556	399	175	175
ProQuest	noft((AI OR artificial* OR auto* OR intelligen* OR machine* OR robo*)) AND noft((ethic* OR moral*)) AND noft((health* OR medic*))	21588	3995	232	230
PubMed	((ethic*[Title/Abstract] OR moral[Title/Abstract]) AND (AI[Title/Abstract] OR artificial*[Title/Abstract] OR robo*[Title/Abstract] OR intelligen*[Title/Abstract] OR machine*[Title/Abstract]) AND (health*[Title/Abstract] OR medic*[Title/Abstract]))	1724	194	105	105
Web Of Scie	(TS=((AI OR artificial* OR auto* OR intelligen* OR machine* OR robo*) AND (ethic* OR moral*)AND (health* OR medic*)))	25241	948	243	243

Table 2. Primary Search Results

The final number of papers is in Table 3 that displays the number of papers left after each step of the process. After three filters were applied (language, document type, 2017-2021), 5,536 papers remained. Then, the manual screening was done according to inclusion criteria (n=755). Finally, the duplicates in each database were removed (n=753) and the removal of duplicates across all databases left only 428 papers. It resumes the first round of the screening process.

Search Process Step	Number of Papers
Results with the search string	57109
Filtered papers	5536
Manually included papers	755
After deletion of duplicates in separate datasets	753
After deletion of duplicates cross datasets	428

Table 3. Number of Papers per Search Process Step

3.3 Inclusion & Exclusion Criteria

The SMS can be used in various ways with different quality results; therefore, it is crucial to establish inclusion and exclusion criteria. After the primary search, the sample was narrowed to 428. The next step of the SMS process, as seen in Figure 7, is the screening process with the inclusion and exclusion criteria presented in Table 4. It ensures that the sample is analyzed so only relevant papers to the research questions, are kept. (Petersen et al., 2008) Also, a single reviewer processed these criteria.

Inclusion	Exclusion
[I1] Paper focused on AI Ethics	[E1] AI Ethics, healthcare mentioned only in the introduction or/and abstract
[I2] Published between 2017-2021	[E2] Papers not related to healthcare
[I3] In English	[E3] Papers with empirical data
[I4] Peer-reviewed articles	
[I5] Available in Full Access	
[I6] White literature	

Table 4. Inclusion & Exclusion Criteria

A paper must fulfill all the inclusion criteria and none of the exclusion criteria to be kept in the sample. As the main research question of this thesis is "*What are the current ethical issues by the use of AI in healthcare",* it was necessary that all papers must focus on AI Ethics (I1). Hence, all papers that were not related to healthcare are excluded (E2). The selected articles must have white literature to maintain a good quality level of sources (I6). White literature is explained as articles published by high control and credible entities, thus, any papers such as blogs, websites, high school papers, and others with black or grey literature are excluded. (Bellefontaine & Lee, 2013)

The rest criteria of inclusion stated that the paper must have been published from 2017 to 2021 (I2) to have the most recent and relevant findings, be written in English (I3), it must be a peer-reviewed article (I3), and to be available in its integrity (I5). These criteria were checked within the database search parameters. Then, for the exclusion criteria, if the paper only mentioned one of the keywords in the search string in its abstract but was not relevant to the rest of the research (E1), it was excluded. Finally, as this thesis focuses on qualitative information and is looking for non-empirical papers for possibly new theories, empirical papers were not relevant (E3). Therefore, those papers were excluded.

3.3.1 Additional Rounds of Screening

To summarize the first round of screening, the sample of papers went through I1, I2, I3 and was narrowed down to 5,536. Then, a manual screening with E1 and E2 was conducted. (n=755) Duplicates from each database were removed (n=753) and removal of duplicates across databases was done. (n=428) The second screening reviewed all the collected papers to see if they are all focused on AI Ethics (I1) and to exclude them if they are not related to healthcare (E2), hence the sample was reduced to 108. Finally, the third and last screening processed the remaining papers by verifying if they had white literature. It would ensure the quality of the article and the SMS results. Figure 7 summarizes the three screening processes and the number of articles left or/and removed each time.

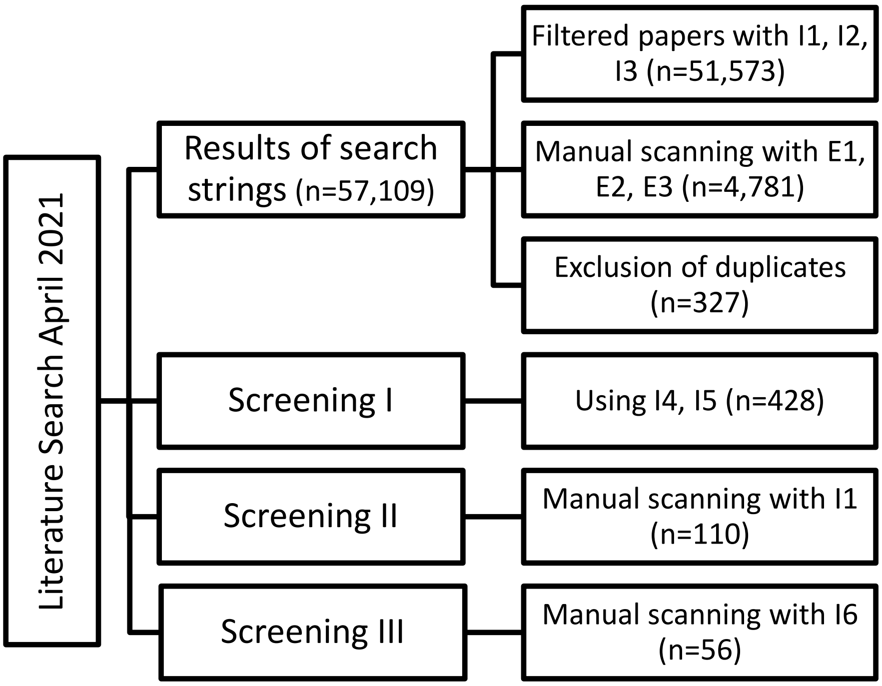

Figure 8. Literature Search April 2021

The final sample for the SMS included 56 papers. The following chapter will cover the SMS classification.

4 Classification

The classification aims to provide continuous and evolving schema throughout the research. As seen in Figure 8, it contains multiple steps to follow. Keywording reduces the time needed to build the classification scheme. It also ensures that the scheme considers the current literature into account. Firstly, one must look at the abstract and identify the concepts, keywords, and context of the paper. Then, after all, papers are reviewed and have keywords attached to them, one can build a set of keywords to create categories. Once the final set of categories is chosen, then the map can be done. (Petersen et al., 2008)

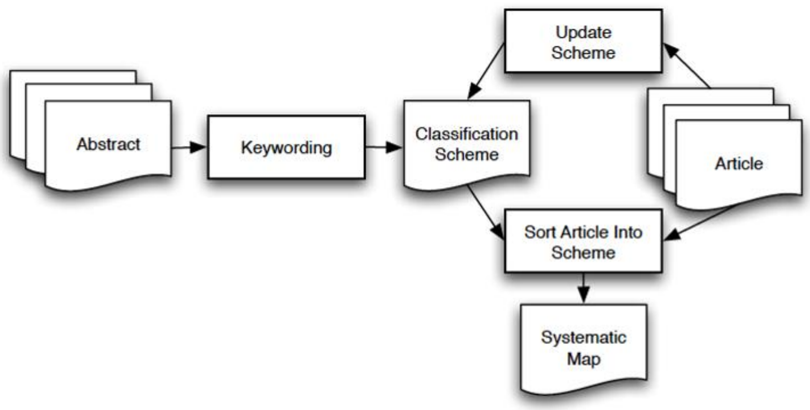

Figure 9. Classification Scheme from Petersen et al. (2008)

For this thesis, the last screening (n=56) was based on literature quality, as well as the focus of each paper. The focus was to look at the abstract, the title, and the keywords used. It was the start of keywording; thus, it helps to place papers in different categories. The following section discusses the process of building the classification scheme and the results.

4.1 Classification Schema

The classification scheme is formed by four facets: research type facet, contribution facet, study facet, and stakeholder facet. The selection of the research type and contribution facets was based on Peterson et al. (2008). The other two were adapted to this thesis to answer the research questions. The summary of this classification can be found in Table 5 based on Mehta et al. (2019). The facets are explained below.

1. Research facet

   It represents the type of research done. (Petersen et al., 2008) In Table 5, this section contains the different research approach and their descriptions based on the work done by Wieringa et al (2005). The research facet is composed of:

   - Solution Proposal, which refers to the proposition of a new solution that the author defends by proving its relevance.
   - Validation Research, which refers to investigating by following a methodology their or someone's paper.
   - Philosophical Paper, which refers to a new conceptual framework as it presents a new view on a problem.
   - Discussion Paper, which refers to resuming the current state of the subject to start or continue the discussion about it.
   - Opinion Paper, which refers to a paper expressing the author's opinion without necessarily following a methodology.
   - Experience Paper, which refers to the implementation of technology in practice. (Mehta et al., 2019)

2. Contribution facet

   It tries to identify the concrete contribution of the paper. In Table 5, this section contains the different contributing factors and their descriptions based on the work of Shaw (2003). It is composed of:

   - Procedure, which refers to producing a new and more efficient way to accomplish something, for example, a framework or an architecture.

- Model, which refers to using a mathematical or conceptual model to solve a problem.
- Tool, which refers to developing an algorithm, a program to solve an issue.
- Specific Solution, which refers to using a specific solution for a specific dilemma.
- Report, which refers to providing some advice, a recommendation with less mathematics or factual proofs.

3. Study focus facet

   It outlines the four ethical issues that are described below. It will allow visualizing which moral issue is less mentioned through the sample. This facet is composed of:

   - Transparency, which refers to the lack of understandability of AI results from developers and medical experts.
   - Justice & Fairness, which refers the data consistency and inclusion, and the potential biases that can affect all stakeholders.
   - Accountability & Responsibility, which refers to the liability issue between developers, medical experts, and AI technologies.
   - Privacy & Security, which refers to the control over the information of oneself and how protected is the data.

4. Stakeholder facet

   It represents all stakeholders involved in AI in healthcare and is composed of:

   - AI Machines, which refer to all technologies able to make decisions by themselves.
   - Developers, which refer to individuals creating or improving an AI tool.
   - Medical experts, which refer to individuals within the healthcare system that use AI tools with their patients.
   - Patients, which refer to individuals experiencing AI tools in healthcare.

Facet	Category	Description
Research facet (Wieringa et al., 2005)	Solution Proposal	Novel relevant solution
	Validation Research	Investigation and evaluation of concepts
	Philosophical Paper	New Conceptual Framework
	Discussion Paper	Information to initiate discussion
	Opinion Paper	Expressing One's Opinion
	Experience Paper	Description of Implementation in Practice
Contribution Facet (Shaw, 2003)	Procedure	Propose a new efficient way
	Model	Use of mathematic or conceptual model
	Tool	Use of a computational tool
	Specific Solution	Specific solution for a specific issue
	Report	Generic advice, recommendation
Focus Facet	Transparency	Lack of understandability
	Justice & Fairness	Data consistency and inclusion & biases
	Accountability & Responsibility	Liability for developers, machines, experts
	Privacy & Security	Anonymization and protection of data
Stakeholders Facet	AI Autonomous Machines	Technologies able to make decision
	Developers	People creating AI technologies
	Medical experts	Clinicians interacting with AI technologies
	Patients	Individual interacting with AI in healthcare
	Governments & Organizations	Any international, national entity

Table 5. Classification Scheme based on Mehta et al. (2019)

The classification process was completed by the author alone, thus its results can vary from one individual to another and is highly opinion-based. Despite it, each paper was categorized in the best fitting facet.

4.2 Results

Once the classification scheme is completed, the data extraction and mapping process is seen in Figure 7 begins. During that step, each paper is classified and throughout that process, the classification might evolve to better fit the obtained results. In this case, it did as no philosophical paper was found; therefore, it was removed. Also, multiple answers could be found for the focus and stakeholders facets as the papers would go over several ethical issues regarding various roles. Hence, all were recorded and compiled accordingly.

Figure 9 regroups all the results in the different facets as mentioned, the focus and stakeholders facet has a total of over 56 as all issues and roles covered were counted to have a better global view of which stakeholder and ethical issue were the less explained. Most papers were discussion papers (36%) as they wanted to bring awareness to current moral issues such as justice & fairness (36%) and privacy & security (27%). Over half of the papers provided recommendations (55%) instead of concrete solutions such as tools (4%) or a conceptual model to follow (13%). Authors considered more carefully medical experts (32%) and little mentioned AI tools (10%).

Facet	Category	Total	Percentage
Research facet	Solution Proposal	13	23%
	Validation Research	6	11%
	Discussion Paper	20	36%
	Opinion Paper	4	7%
	Experience Paper	13	23%
Contribution Facet	Procedure	11	20%
	Model	7	13%
	Tool	2	4%
	Specific Solution	5	9%
	Report	31	55%
Focus Facet	Transparency	29	22%
	Justice & Fairness	48	36%
	Accountability & Responsibility	19	14%
	Privacy & Security	36	27%
Stakeholders Facet	AI Autonomous Machines	14	10%
	Developers	24	17%
	Medical experts	46	32%
	Patients	31	22%
	Governments & Organizations	28	20%

Table 6. Results of Classification Scheme

For the research facet, it classifies the kind of article done by the authors. Figure 10 provides the summary of that facet. Most of the papers were discussion papers (36%) as the authors wanted to bring awareness and start the discussion on particular subjects. Then, solution proposal and experience papers (23%) were both equally covered as one was trying to provide a new solution, and the other analyzes real-world practices. Finally, opinion (7%) and validation research (11%) papers were less popular within the subject.

EC1: Opinion paper is the least popular type of research for AI ethics in healthcare.

EC2: Discussion paper is the most popular type of research for AI ethics in healthcare with representing 36% of the final sample.

Figure 10. Research Facet Results

For the contribution facet, it was consisting of evaluating what did the paper offered to the research community. Figure 10 provides the summary of that facet. Over half of the sample were contributing by providing a report. (55%) They gave recommendations or advice to the different stakeholders mentioned throughout their article. The following popular one was to contribute by giving a procedure (20%) which is to provide a new efficient way to solve the mentioned issue. Then, some were providing a conceptual model

(15%) as a solution and a small number of papers provided a specific solution to a problem (9%) or a tool (4%) such as software.

EC3: Over half of the final sample contributed by giving a report which included recommendations and advice on a given problem.

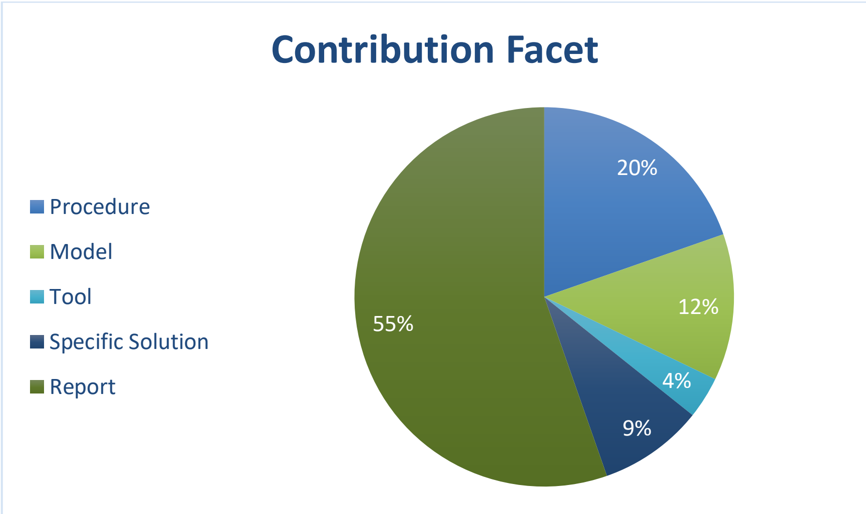

Figure 11. Contribution Facet Results

For the focus facet, its objective was to provide quantitative data on which moral issue in AI ethics is covered the most and the least. Figure 11 provides the summary of that facet. The most stated ethical dilemma was justice & fairness (37%) that regrouped any bias such as data bias, AI bias, human cognitive bias, and legal aspect. Then, privacy & security (28%) was also generously covered regarding patient data protection, data anonymization, and more. Followed closely by transparency (23%), it touched anything related to black-box issues, interpretability, and understandability of algorithms and results of AI tools. Finally, accountability & responsibility (12%) was less discussed amongst papers.

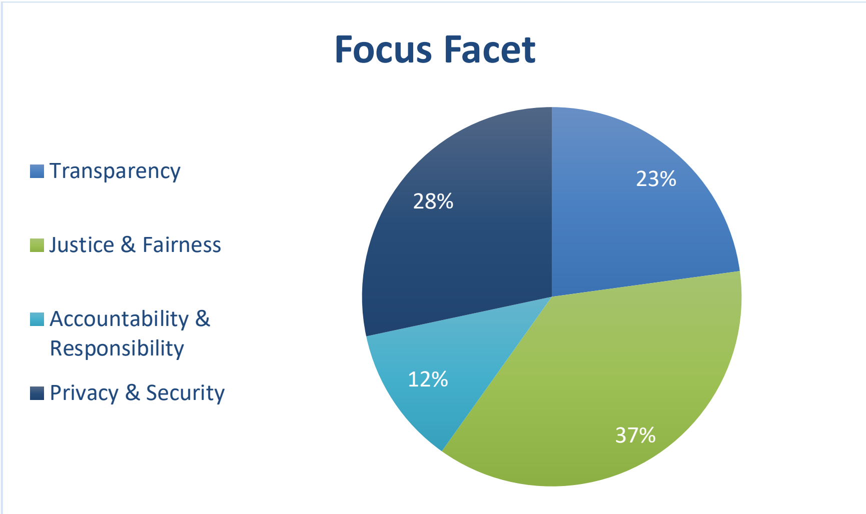

Figure 12. Focus Facet Results

EC4: In the focus facet, accountability & responsibility was the least discussed category through papers.

For the stakeholders facet, it needed to give an overview of the involved stakeholders in using AI in healthcare. Figure 12 provides that global summary. Generally, papers did explore ethical issues regarding medical experts (31%), patients (22%), and organizations (20%). More rarely, they mentioned developers (17%) and AI machines (10%).

EC5: AI tools were explored by 10% of the final sample regarding their morality in ethical issues in healthcare.

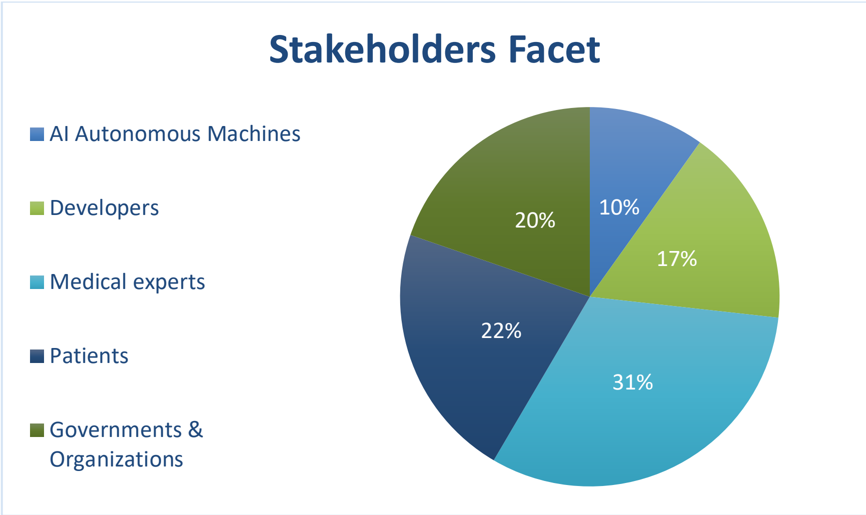

Figure 13. Stakeholders Facet Results

Additionally, each paper was analyzed to evaluate if it provides a complete solution to all stakeholders' ethical issues mentioned, a partial solution, or none. Table 7 provides a general overview of the categorization, followed by Figure 13 that provides the percentage from the final sample (n=56). Only 27% of the papers provided a complete solution for each stakeholder and moral problem mentioned. The rest either provided some solution or none at all.

Category	Description	Total
None	No solution proposed for ethical issues	21
Partially	Partial solution proposed for ethical issues	20
Fully	Complete solution proposed for ethical issues	15

Table 7. Type of Solutions Provided

EC6: 37.5% of the papers did not provide any solution to ethical issues related to AI in healthcare.

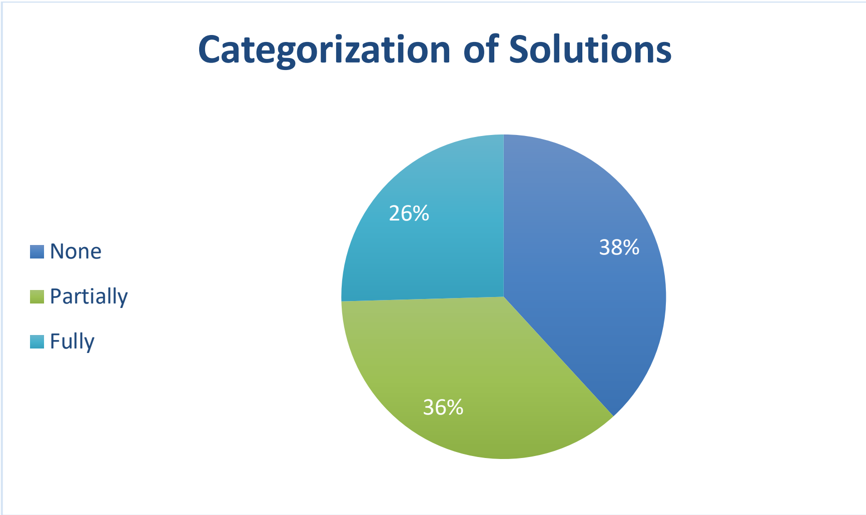

Figure 14. Categorization of Solutions

Finally, Table 8 presents the number of papers published per year starting 2017 to 2021. It highlights a significant increase in the academic literature starting in 2019.

EC7: An increase of published articles regarding AI ethics in healthcare is seen in the academic literature regarding AI ethics in healthcare starting 2019.

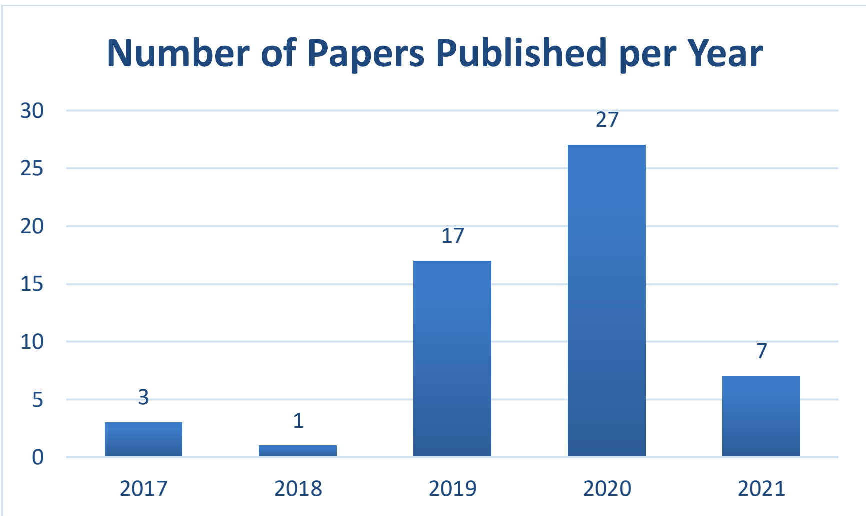

Table 8. Number of Papers Published per Year

4.3 Overview of Final Sample

This section provides all information related to the final sample (n=56) of papers seen in Table 8. Only the first author's last name is mentioned to keep some space for the focus and stakeholders facets. The classification includes the four facets, and the solution column is also present.

Paper	Classification				Solution
1st Author	Research	Contribution	Focus	Stakeholders	Solution
Abramoff et al. (2020)	Solution Proposal	Model	Responsibility; Justice & fairness; Transparency; Privacy & Security	Patient, Medical Experts, Machines, Developers	Fully
Ahmed et al. (2020)	Solution Proposal	Procedure	Security & Privacy; Justice & Fairness	Machines	Fully
Alami et al. (2020)	Validation Research	Report	Responsibility; Privacy & Security; Transparency	Organizations, Developers, Medical Experts, Patients	Partially
Amann et al. (2020)	Solution Proposal	Specific Solution	Transparency; Justice & Fairness	Developers, Medical Experts; Patients; Organizations	Fully
Asan et al. (2020)	Discussion Paper	Model	Transparency; Justice & Fairness	Medical Experts, Machines	Partially
Bærøe et al. (2020)	Experience Paper	Model	Justice & Fairness	Organizations	None
Beil et al. (2019)	Validation Research	Specific Solution	Transparency; Justice & Fairness	Medical Experts, Patients, Organizations	Fully
Bezemer et al. (2019)	Solution Proposal	Report	Transparency; Justice & Fairness; Responsibility	Medical Experts	Partially
Bonderman et al. (2017)	Experience Paper	Report	Transparency	Developers	None
Burwell et al. (2017)	Discussion Paper	Report	Responsiblity, Privacy & Security	Patients, Machines	Partially
Car et al. (2019)	Experience Paper	Report	Privacy & Security	Patients	None

Coin et al. (2020)	Discussion Paper	Procedure	Responsibility, Privacy & Security, Justice&Fairness	Patients, Medical experts, Organizations	Fully
Cui et al. (2021)	Experience Paper	Report	Privacy & Security	Patients, Medical experts	None
Cuzzolin et al. (2020)	Solution Proposal	Procedure	Transparency, Privacy & Security, Justice&Fairness	Patients, Machines	Partially
Du-Harpur et al. (2020)	Discussion Paper	Report	Justice & Fairness; Transparency	Machines	None
Esmaeilzadeh (2020)	Solution Proposal	Model	Justice & Fairness; Transparency; Security & Privacy; Responsibility	Developers, Medical experts; Patients; Organizations	Fully
Favaretto et al. (2020)	Solution Proposal	Procedure	Justice & Fairness; Security & Privacy	Medical experts; Patients; Organizations	Fully
Galbusera et al. (2019)	Experience Paper	Report	Justice & Fairness; Security & Privacy	Medical experts; Patients; Organizations	None
Garcelon et al. (2020)	Validation Research	Report	Justice & Fairness; Security & Privacy	Organizations; Medical experts	Partially
Geis et al. (2019)	Opinion Paper	Procedure	Accountability; Justice&fairness; Transparency; Security&Privacy	Developers, Medical experts; Patients; Organizations	
Graham et al. (2019)	Discussion Paper	Report	Justice & Fairness; Transparency	Organizations; Medical experts; Patients	None
Grote et al. (2019)	Discussion Paper	Report	Transparency; Responsibility	Medical experts; Patients; Organizations	None
He et al. (2020)	Experience Paper	Report	Justice & Fairness; Security & Privacy; Transparency	Medical experts; Patients; Organizations; Developers	None
Jiang et al. (2021)	Solution Proposal	Procedure	Justice & Fairness; Security & Privacy; Responsibility	Medical experts; Patients; Organizations; Developers	Fully
Kellmeyer (2019)	Discussion Paper	Report	Justice & Fairness; Security & Privacy; Responsibility	Medical experts; Patients	None

Kögel et al. (2019)	Experience Paper	Report	Responsibility, Privacy & Security, Justice&Fairness; Transparency	Medical experts; Patients; Organizations; Developers	None
Lawrie et al. (2019)	Experience Paper	Report	Justice & Fairness; Security & Privacy	Medical experts; Patients; Organizations	Fully
Lebcir et al. (2021)	Discussion Paper	Report	Justice & Fairness	Medical experts	None
Lepri et al. (2021)	Solution Proposal	Tool	Justice & Fairness; Transparency; Security & Privacy; Responsibility	Developers	Fully
Loftus et al. (2020)	Solution Proposal	Model	Justice & Fairness; Responsibility	Medical experts	Partially
Mackey et al. (2020)	Experience Paper	Procedure	Justice & Fairness; Transparency	Medical experts; Organizations	Partially
Mooney et al. (2017)	Solution Proposal	Report	Justice & Fairness; Transparency; Security & Privacy	Medical experts; Developers; Patients	Partially
Nagaraj et al. (2020)	Solution Proposal	Model	Justice & Fairness; Security & Privacy	Patients	Fully
Park et al. (2019)	Discussion Paper	Specific Solution	Transparency, Justice & Fairness	Patients, Medical experts	Partially
Patel et al. (2020)	Experience Paper	Report	Justice & Fairness	Medical experts, Organizations	Partially
Pedersen et al. (2020)	Experience Paper	Report	Justice & Fairness; Transparency; Security & Privacy	Medical experts; Patients; Organizations	None
Pépin et al. (2019)	Experience Paper	Specific Solution	Security & Privacy	Medical experts; Patients; Organizations	Partially
Pot et al. (2021)	Discussion Paper	Report	Justice & Fairness	Medical experts; Patients; Organizations, Developers	None
Price et al. (2019)	Discussion Paper	Report	Justice & Fairness; Security & Privacy	Medical experts; Patients; Organizations	None
Rowe et al. (2019)	Validation Research	Specific Solution	Justice & Fairness; Security & Privacy	Developers	Partially

Saba et al. (2019)	Discussion Paper	Report	Security & Privacy; Responsibility	Medical experts; Patients; Machines	None
Schwendicke et al. (2020)	Opinion Paper	Report	Justice & Fairness; Security & Privacy; Transparency	Developers; Medical experts	None
Shuaib et al. (2020)	Opinion Paper	Report	Justice & Fairness; Security & Privacy; Responsibility	Medical experts	Partially
Six Dijkstra et al. (2020)	Discussion Paper	Report	Justice & Fairness; Security & Privacy; Transparency	Developers; Medical experts; Patients	None
Smith et al. (2020)	Opinion Paper	Report	Justice & Fairness	Developers; Medical experts; Machines	None
Stanfill et al. (2019)	Discussion Paper	Report	Justice & Fairness; Security & Privacy; Responsibility	Developers; Medical experts; Machines; Organizations	None
Starke et al. (2020)	Discussion Paper	Procedure	Justice & Fairness; Security & Privacy; Responsibility; Transparency	Developers; Medical experts; Machines; Organizations	Partially
Sumiyama et al. (2021)	Experience Paper	Tool	Justice & Fairness; Transparency	Medical experts	None
Thapa et al. (2021)	Validation Research	Model	Justice & Fairness; Security & Privacy; Responsibility; Transparency	Developers; Medical experts; Machines; Organizations	Fully
Thieme et al. (2020)	Discussion Paper	Report	Justice & Fairness; Security & Privacy	Patients; Medical experts	Partially
Ursin et al. (2021)	Discussion Paper	Procedure	Justice & Fairness; Security & Privacy; Responsibility; Transparency	Developers; Medical experts; Machines; Organizations	Partially
Vaisman et al. (2020)	Discussion Paper	Report	Justice & Fairness; Security & Privacy	Medical experts; Developers; Patients	Fully
Vellido (2019)	Discussion Paper	Report	Justice & Fairness; Security & Privacy; Transparency	Developers; Medical experts; Machines; Organizations	Partially
Vollmer et al. (2020)	Solution Proposal	Procedure	Justice & Fairness; Responsibility; Transparency	Developers; Medical experts;	Fully

				Machines; Organizations	
Wahl et al. (2018)	Discussion Paper	Report	Justice & Fairness	Medical experts; Patients	Partially
Zhang et al. (2020)	Validation Research	Procedure	Justice & Fairness; Responsibility; Transparency	Medical experts; Developers	Partially

Table 9. Classification of the Final Sample

5 Results of Systematic Mapping Study

The utility of having different facets in SMS is to highlight the frequency of each category within the subject. By doing so, it emphasized the current state of the academic field in AI ethics in healthcare. Usually, it is recommended to display the results in a bar graph or a bubble graph because they allow identifying the possible gaps and future research. (Petersen et al., 2008) This chapter will go over the bubble charts created to represent the current state of the academic literature and an additional table.

5.1 Bubble Plot Visualization

To build this bubble plot, it was required to identify the number of papers per category. In this case, two bubble plots were made to analyze the most covered ethical dilemma and its research and contribution facet and the most mentioned stakeholder and its research and contribution facet. Table 9 and Table 10 presents the distribution of the papers. One author may have mentioned at least one ethical dilemma and stakeholders; therefore, the final amount is above 56. As recommended by Petersen et al., a bubble plot was completed for each table respectively to provide a clearer view of what has been done in the academic literature.

	Transparency	Justice & Fairness	Accountability & Responsibility	Privacy & Security
Solution Proposal	3	13	7	9
Validation Research	1	5	3	4
Discussion Paper	3	17	6	12
Opinion Paper	0	4	2	3
Experience Paper	1	9	1	8
Procedure	1	11	6	8
Model	1	7	4	4
Tool	0	2	1	1
Specific Solution	3	4	0	2
Report	3	24	8	21

Table 10. Ethical Issues and Research & Contribution Facets

	AI Autonomous Machines	Developers	Medical experts	Patients
Solution Proposal	4	7	9	8
Validation Research	1	4	5	2
Discussion Paper	8	7	18	13
Opinion Paper	1	3	4	1
Experience Paper	0	3	10	8
Procedure	5	6	9	5
Model	3	3	5	3
Tool	0	1	1	0
Specific Solution	0	2	4	4
Report	6	12	27	20

Table 11. Stakeholders and Research & Contribution Facets

Figure 14 presents the current state of literature for ethical issues (privacy & security, accountability & responsibility, justice & fairness, and transparency). It is related to Table 9. Inside each bubble is the number of papers. For example, 21 papers explored privacy & security and provided a report as a contribution. Only one experience paper dealt with transparency.

EC8: 76% of discussion papers covered privacy & security and justice & fairness.

EC9: 80% of papers covering privacy & security and justice & fairness contributed by providing a report.

EC10: Justice & Fairness is the most explored category of ethical issues.

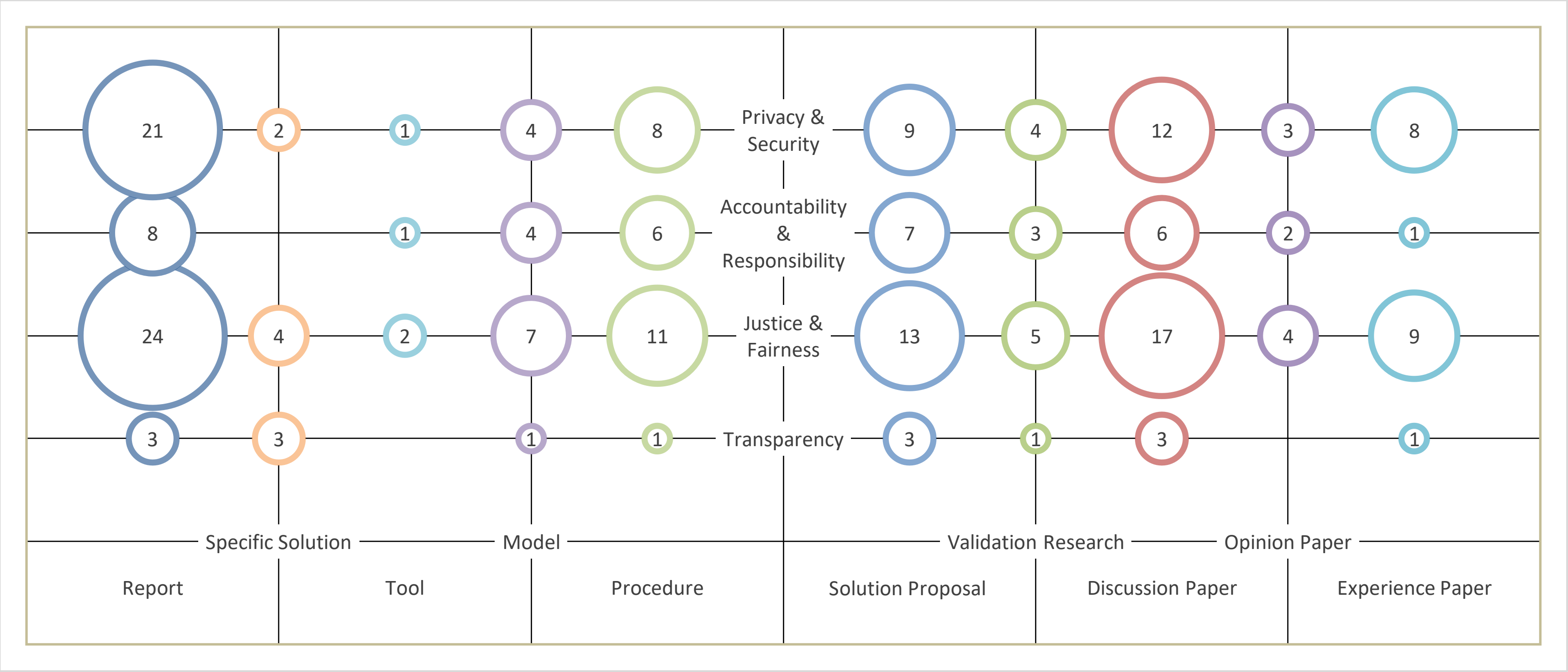

Figure 15. Bubble Plot with Ethical Issues

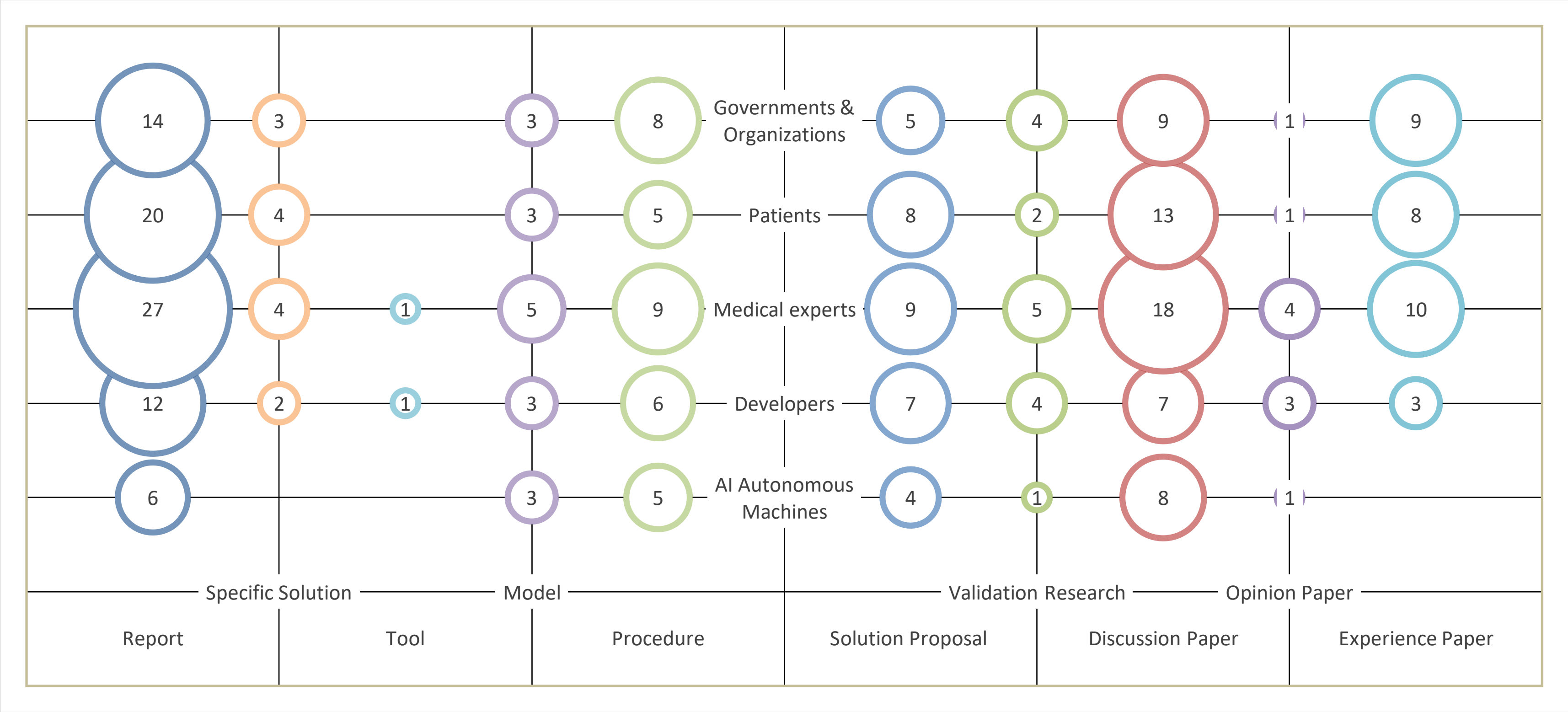

Figure 16. Bubble Plot with Stakeholders

Figure 15 provides an overview of the academic literature for stakeholders involved in AI in healthcare (AI autonomous machines, developers, medical experts, patients, and organizations). It is related to Table 10. Inside each bubble is the number of papers. For example, 27 papers explored explore the role of medical experts and provided a report for the contribution facet.

EC11: Over 50% of the contribution facet is categorized as a report.

EC12: Medical experts are the most considered stakeholders regarding ethical issues.

Additionally, all authors providing a complete solution to all ethical problems mentioned were analyzed. Figure 16 displays the number of papers per contribution facet category. It provided information on the current solutions proposed in the academic literature and which are most popular.

EC13: The most popular type of contribution for papers that provided a complete solution is a procedure.

EC14: Only 26% of the final sample provided a complete solution to ethical issues mentioned by the authors.

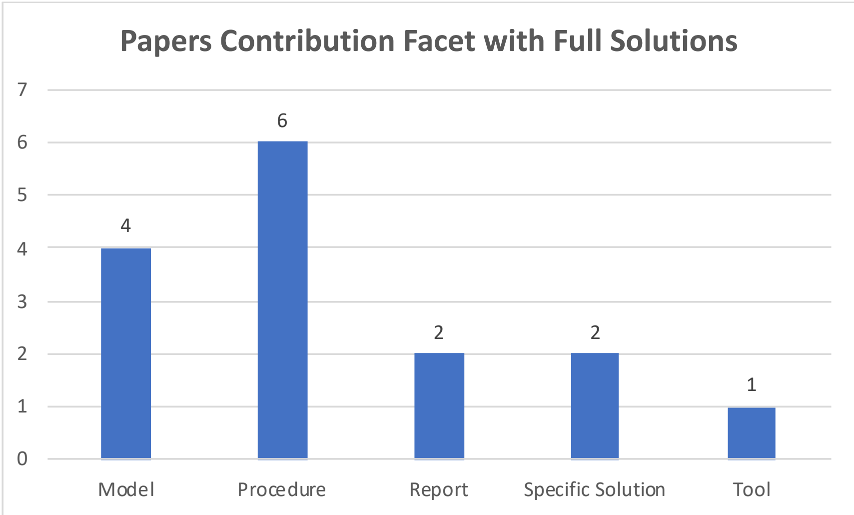

Figure 17. Papers' Contribution Facet that provided Full Solution

The evolution of AI ethics in healthcare can be evaluated by comparing the publication dates of the papers with their contribution and research facets. Figure 17 highlights which area has been gaining more attention in recent years as well as which type was the most or the least done and the type of solutions proposed.

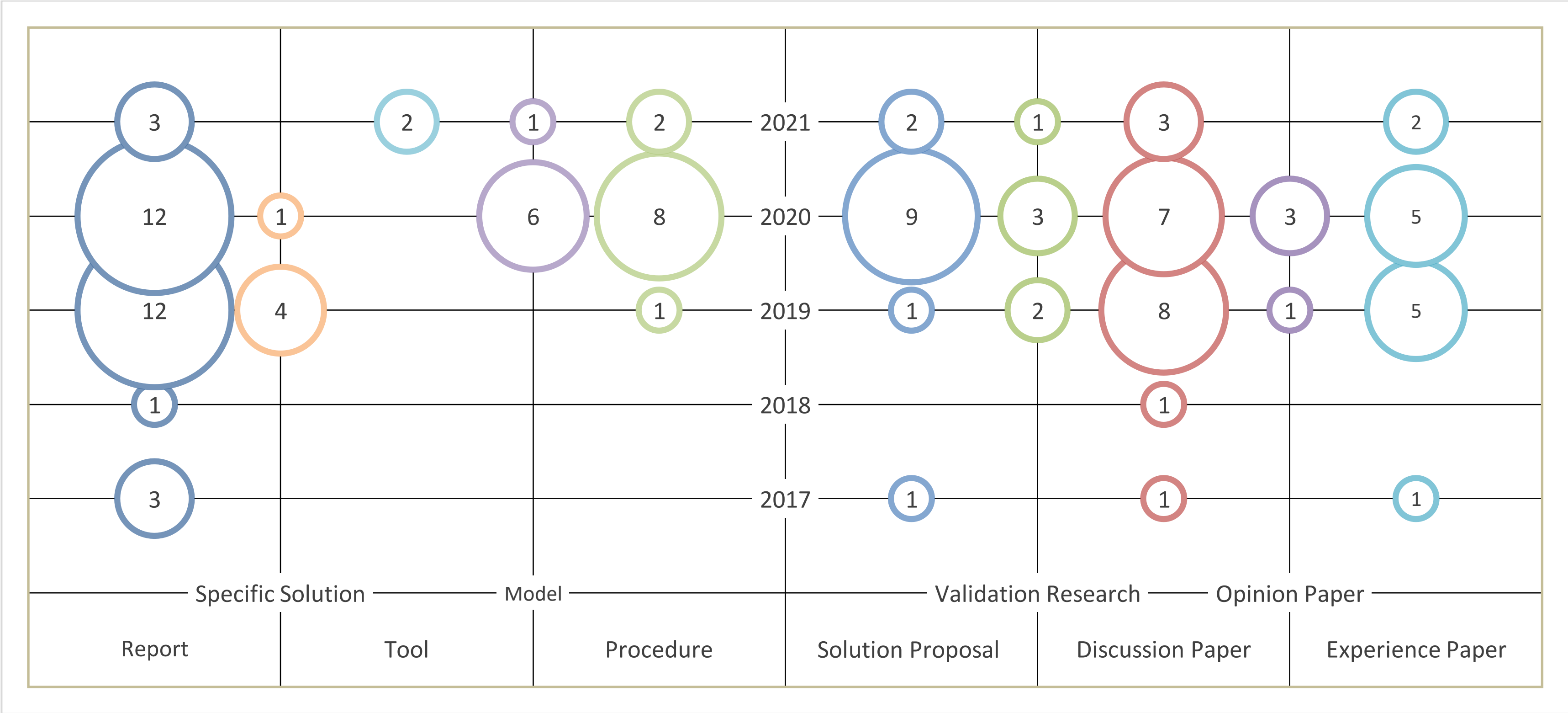

Figure 18. Research and Contribution Facets According to Year

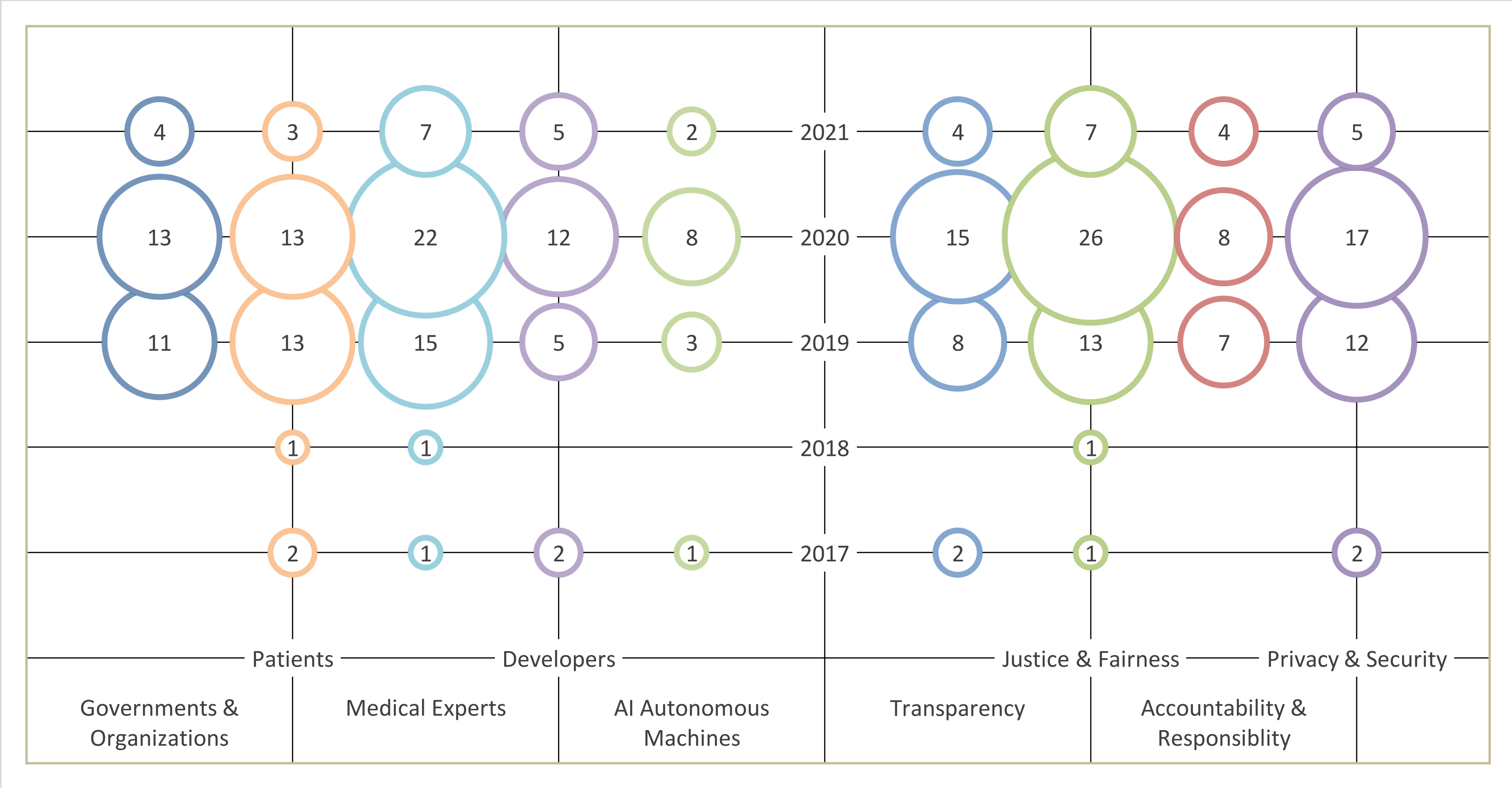

Figure 19. Study Focus and Stakeholders Facet According to Year

Regarding the stakeholders and ethical issues, it is also interesting to see their evolution in terms of attention from academic literature. Figure 18 provides a bubble plot that shows a significant increase starting in 2019.

EC15: The number of articles exploring justice & fairness issues doubles between 2019 and 2020.

5.2 Summary of Empirical Conclusions

This section provides a summary in Table 12 of all empirical conclusions (EC) mentioned previously. These will provide additional help to future research that has a similar subject.

Number	**Description**
EC1	Opinion paper is the least popular type of research for AI ethics in healthcare.
EC2	The discussion paper is the most popular type of research for AI ethics in healthcare with representing 36% of the final sample
EC3	Over half of the final sample contributed by giving a report which included recommendations and advice on a given problem
EC4	In the focus facet, accountability & responsibility was the least discussed category through papers.
EC5	AI tools were explored by 10% of the final sample regarding their morality in ethical issues in healthcare.
EC6	37.5% of the papers did not provide any solution to ethical issues related to AI in healthcare.
EC7	An increase of published articles regarding AI ethics in healthcare is seen in the academic literature regarding AI ethics in healthcare starting 2019.
EC8	76% of discussion papers covered privacy & security and justice & fairness.
EC9	80% of papers that covered privacy & security and justice & fairness contributed by providing a report.
EC10	Justice & Fairness is the most explored category of ethical issues
EC11	Over 50% of the contribution facet is categorized as a report.
EC12	Medical experts are the most considered stakeholders regarding ethical issues.

EC13	The most popular type of contribution for papers providing a complete solution is a procedure.
EC14	Only 26% of the final sample provided a complete solution to ethical issues mentioned by the authors.
EC15	The number of articles exploring justice & fairness issues has doubled between 2019 and 2020.

Table 12. Summary of Empirical Conclusions

6 Discussion

This chapter presents a discussion of the findings by comparing the results and the current literature. The summary of all ECs can be found in Table 12.

6.1 Current State of Stakeholders involved in using AI in Healthcare

The stakeholders involved in using AI in healthcare in the research field have been identified as medical experts, patients, organizations & governments, developers, and AI autonomous machines. Figure 12 is a pie chart that represents the percentage of papers that have explored a particular stakeholder regarding moral issues. Two EC mentioned previously highlighted that (EC12) medical experts were the most popular stakeholders considered and that (EC5) AI tools were the least considered. Abramoff et al. (2020) and Gecke et al. (2020) also agree that fairness and transparency are crucial for a successful clinician-patient relationship.

AI autonomous machines are rarely mentioned as they are not considered by authors as a concrete stakeholder. Hence, the developers and medical experts have higher importance. Authors seemed to prioritize medical experts and patients as (EC8) 76% of discussion papers covered privacy & security and justice & fairness. Both are directly involved with these stakeholders as displayed in Figure 15. (EC4) Developers and AI autonomous tools would require more research as they are highly involved in the focus facet, accountability & responsibility, and (EC4) it was the least discussed category through papers. Reddy et al. (2019) state that AI machines should not deceive humans by mimicking voice or human appearance, this was not found in the results as few researchers covered that stakeholder. Also, in Figure 18, the number of articles for AI autonomous machines and developers has doubled from 2019 to 2020.

6.2 Solutions for AI Ethical Issues in Healthcare

The main ethical issues selected for this thesis were: justice & fairness, privacy & security, accountability & responsibility, and transparency which is also found in the studies of Ahmed et al. (2020) and Alami et al. (2020). Each article within the final sample could go over multiple moral issues, hence, the larger numbers in those tables and charts. A contribution facet was chosen to identify proposed solutions for them. It was divided into the following: procedure, model, tool, specific solution, and report. A paper could only have one type of solution for all ethical problems stated. From the results, (EC3, EC11) 55% of the final sample of articles was categorized as a report. It meant that most of them only provided recommendations or advice to follow. Additionally, not all papers provided a solution, in Table 7 and Figure 13, (EC6) 37.5% of them did not give any solution, and (EC14) only 26% of them provided a complete solution to ethical issues related to AI in healthcare. Analyzing authors that provided a complete solution in Figure 16, (EC13) the most popular type of contribution for these papers is a procedure. Also, (EC9) 80% of papers covering privacy & security and justice & fairness contributed by providing a report as solutions. Abramoff et al. (2020) have proposed a solution to evaluate AI autonomous machines for developers. Similarly, Asan et al. (2020) have proposed a validation and verification plan to ensure the reliability of AI.

6.3 Current gaps in the Academic Literature

Throughout chapters 5 and 6, some gaps are highlighted in the different facets of the classification scheme. In Table 8, (EC7) the number of papers can be seen increasing starting 2019, and (EC15) 2020 contains almost half of the final sample. Additionally, (EC16) the number of articles exploring justice & fairness has doubled between 2019 and 2020. Despite these encouraging statistics, it has not been the case for all categories. For example, developers & AI autonomous machines were the least stated. They would require more research as they are highly involved in accountability & responsibility, and (EC4) it was the least discussed category through papers. Therefore, regarding the stakeholders facet, developers and AI tools need more attention.

Concerning the research facet, (EC1) opinion papers and validation research are the least done for AI ethics in healthcare as seen in Figure 9. Meanwhile, (EC2) discussion papers are the most popular and represent 36% of the final sample. For the contribution facet, only a handful of authors have provided a complete and concrete solution. Most of them have proposed recommendations. More research should focus on providing tools, models, and procedures to solve these ethical issues. Finally, the study focus has highlighted in Figure 11 that (EC10) justice & fairness is the most explored and that accountability & responsibility is the least analyzed. Many authors such as Garcelon et al. (2020) believe that large sets of data are difficult to find for research and it would require data fragmentation, data quality, and data privacy to ensure that quality.

6.4 Overall Results

For the practical implementation, almost all authors have mentioned that collaboration between medical experts and developers is necessary to provide a secure, explainable, and fair AI tool. For theoretical implications, authors must consider developers and AI machines more regarding ethical issues. Medical experts and patients are crucial but would not use AI tools without developers and machines.

As mentioned previously, AI ethics in healthcare has been gaining more attention within the last three years as society is concerned for privacy & safety and justice & fairness the most. From medical experts and developers, accountability & responsibility is important. At the moment, only medical experts are liable if the machine misdiagnosis a patient because it should be used as a decision support tool. More research would be needed to validate the current use of AI in healthcare from an ethical point of views such as experience and validation papers.

Authors also need to provide more concrete solutions to help developers, medical experts, and organizations & governments to improve the administration behind AI machines. Implementation of AI tools in healthcare will require the cooperation of medical experts, developers, organizations, and patients to successfully use them. Otherwise, they will bring more ethical issues than solutions.

7 Conclusions

Finally, to understand the background of the subject a quick summary of AI history in medicine, the types of AI used in healthcare, and AI ethics in healthcare was done. This thesis was concluded using the SMS method to visualize the current state of academic literature regarding AI ethics in healthcare. It provides an overview of the type of reports and results published by categorizing them. (Kitchenham et al., 2012) The classification was done according to four facets: contribution, research, study focus, and stakeholders. A total of 56 papers were classified and analyzed. It created various bubble plots that displayed different relationships and findings. Additionally, the SMS results highlighted different gaps in the academic literature. The aim purpose of this thesis was to provide a clearer view of the ethical impacts that AI has on healthcare. Each article was analyzed regarding solutions (if any) proposed. The summary of all ECs can be found in Table 12. The most interesting findings were that justice & fairness is the most explored category of ethical issues and that medical experts are the most considered stakeholders regarding ethical issues.

7.1 Limitations

This thesis has certain limitations that need to be mentioned. The research method, SMS, has biases on its own such as lack of quality in the results, lack of quality in the articles selected, and it does not show that evidence is missing or insufficient. (Kitchenham et al., 2012) By limiting the ethical issues to justice & fairness, responsibility & accountability, security & privacy, and transparency, some papers have been excluded and might have provided data regarding other important aspects. As the primary studies search was done by inclusion and exclusion criteria, this could have affected the accuracy of the obtained final sample of papers. To keep the sample size to a minimum, at the last stage of selection, articles were analyzed one by one to ensure their validity regarding the thesis. This has been time-consuming and required attentive work. It was based on criteria but also the author's judgment. The classification scheme was also done accordingly to the author's judgment based on each category's definition. Throughout the SMS process, human bias could have happened.

7.2 Future Research

Finally, future research can be done using this thesis. It has the potential to gather a deeper understanding of the different stakeholders and ethical issues in healthcare. The thesis has followed clear and explicit guidelines using SMS methodology thus, it can be reused. As its findings have highlighted the different gaps in academic literature, a study could potentially research more into them. One future research would be to update the literature by the end of 2021. Future research could look deeper into a complete solution for the different ethical issues according to specific stakeholders that were less explored in this thesis.

Bibliography

Alexander, L., & Ripstein, A. (2001). Equality, Responsibility, and the Law. *Law and Philosophy*, *20*(6), 617. https://doi.org/10.2307/3505159

Anderson, M. & Anderson, S. L. (2007). Machine Ethics: Creating an Ethical Intelligent Agent. *AI Magazine*, 28, 15-26.

Amisha, F., Malik, P., Pathania, M., & Rathaur, V. (2019). Overview of artificial intelligence in medicine. *Journal of Family Medicine and Primary Care*, 8(7), 2328. https://doi.org/10.4103/jfmpc.jfmpc_440_19

Bartoletti, I. (2019). AI in Healthcare: Ethical and Privacy Challenges. *Artificial Intelligence in Medicine,* 7–10. https://doi.org/10.1007/978-3-030-21642-9_2

Bellefontaine, S. P., & Lee, C. M. (2013). Between Black and White: Examining Grey Literature in Meta-analyses of Psychological Research. *Journal of Child and Family Studies*, *23*(8), 1378–1388. https://doi.org/10.1007/s10826-013-9795-1

Choi, K., Fazekas, G., Sandler, M., & Cho, K. (2017). Convolutional recurrent neural networks for music classification. *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* 2392–2396. https://doi.org/10.1109/icassp.2017.7952585

Coeckelbergh, M. (2010). Robot rights? Towards a social-relational justification of moral consideration. *Ethics and Information Technology*, *12*(3), 209–221. https://doi.org/10.1007/s10676-010-9235-5

Comendador, B. E. V., Francisco, B. M. B., Medenilla, J. S., Nacion, S. M. T., & Serac, T. B. E. (2015). Pharmabot: A Pediatric Generic Medicine Consultant Chatbot. *Journal of Automation and Control Engineering*, 3(2), 137–140. https://doi.org/10.12720/joace.3.2.137-140

Dalal, K. R. (2020). Analysing the Implementation of Machine Learning in Healthcare. *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 133–137. https://doi.org/10.1109/icesc48915.2020.9156061

Dave, D., Naik, H., Singhal, S., & Patel, P. (2020, November 6). *Explainable AI meets Healthcare: A Study on Heart Disease Dataset*. ArXiv.Org. https://arxiv.org/abs/2011.03195

Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthcare Journa*l,6(2), 94-98. doi:10.7861/futurehosp.6-2-94

Dignum, V. (2020). Responsibility and Artificial Intelligence. *The Oxford Handbook of Ethics of AI*, 213–231. https://doi.org/10.1093/oxfordhb/9780190067397.013.12

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. *Nature Medicine*, *25*(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z

Ferrucci, D., Levas, A., Bagchi, S., Gondek, D., & Mueller, E. T. (2013). Watson: Beyond Jeopardy! Artificial Intelligence, 199–200, 93–105. https://doi.org/10.1016/j.artint.2012.06.009

Floridi, L. (2009). Foundations of Information Ethics. *The Handbook of Information and Computer Ethics*, 1–23. https://doi.org/10.1002/9780470281819.ch1

Gebru, T., Morgenstern, J., Vecchione, B., Wortman Vaughan, J., Wallach, H., Daumeé III, H., and Crawford, K. (2018) *Datasheets for Datasets*, arxiv:1803.09010

Gerke, S., Minssen, T., & Cohen, G. (2020). Ethical and legal challenges of artificial intelligence-driven healthcare. *Artificial Intelligence in Healthcare*, 295–336. https://doi.org/10.1016/b978-0-12-818438-7.00012-5

Guan, J. (2019). Artificial Intelligence in Healthcare and Medicine: Promises, Ethical Challenges, and Governance. *Chinese Medical Sciences Journal*, 34(2), 76-83. doi:10.24920/003611

Gunning, D., & Aha, D. (2019). DARPA's Explainable Artificial Intelligence (XAI) Program. *AI Magazine*, *40*(2), 44–58. https://doi.org/10.1609/aimag.v40i2.2850

Hoogenboom, S. A., Bagci, U., & Wallace, M. B. (2020). Artificial intelligence in gastroenterology. The current state of play and the potential. How will it affect our practice and when? *Techniques and Innovations in Gastrointestinal Endoscopy*, 22(2), 42–47. https://doi.org/10.1016/j.tgie.2019.150634

Jameel, T., Ali, R., & Toheed, I. (2020). Ethics of Artificial Intelligence: Research Challenges and Potential Solutions. *2020 3rd International Conference on Computing, Mathematics and Engineering Technologies (ICoMET)*, 1–6. https://doi.org/10.1109/icomet48670.2020.9073911

James, K. L., Randall, N. P., & Haddaway, N. R. (2016). A methodology for systematic mapping in environmental sciences. *Environmental Evidence*, *5:7*(1), 1–13. https://doi.org/10.1186/s13750-016-0059-6

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. *Stroke and Vascular Neurology*, 2(4), 230–243. https://doi.org/10.1136/svn-2017-000101

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Johnson, A. T. (2020). Ethics in the Era of Artificial Intelligence. *IEEE Pulse*, *11*(3), 44–47. https://doi.org/10.1109/mpuls.2020.2993667

Kaul, V., Enslin, S., & Gross, S. A. (2020). History of artificial intelligence in medicine. *Gastrointestinal Endoscopy*, 92(4), 807-812. doi:10.1016/j.gie.2020.06.040

Kitchenham, B. A., Budgen, D., & Pearl Brereton, "Mapping study completeness and reliability - a case study," 1*6th International Conference on Evaluation & Assessment in Software Engineering (EASE 2012), Ciudad Real*, 2012, pp. 126-135, doi: 10.1049/ic.2012.0016.

Kitchenham, B. A., Budgen, D., & Pearl Brereton, O. (2011). Using mapping studies as the basis for further research – A participant-observer case study. I*nformation and Software Technology*, 53(6), 638–651. https://doi.org/10.1016/j.infsof.2010.12.011

Kuipers, B., Feigenbaum, E. A., Hart, P. E., & Nilsson, N. J. (2017). Shakey: From Conception to History. *AI Magazine*, 38(1), 88–103. https://doi.org/10.1609/aimag.v38i1.2716

Kulikowski, C. A. (2019). Beginnings of Artificial Intelligence in Medicine (AIM): Computational Artifice Assisting Scientific Inquiry and Clinical Art – with Reflections on Present AIM Challenges. *Yearbook of Medical Informatics*, 28(01), 249–256. https://doi.org/10.1055/s-0039-1677895

Le Berre, C., Sandborn, W. J., Aridhi, S., Devignes, M. D., Fournier, L., Smaïl-Tabbone, M., Danese, S., & Peyrin-Biroulet, L. (2020). Application of Artificial Intelligence to Gastroenterology and Hepatology. *Gastroenterology*, 158(1), 76–94. https://doi.org/10.1053/j.gastro.2019.08.058

Leslie, D. (2019). Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector. *The Alan Turing Institute*. https://doi.org/10.5281/zenodo.3240529

McCartney, S. (2015, April 17). *2.1 Major Ethical Systems – Ethics in Law Enforcement*. Pressbooks. https://opentextbc.ca/ethicsinlawenforcement/chapter/2-1-major-ethical-systems/

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2019, August 23). *A Survey on Bias and Fairness in Machine Learning*. ArXiv.Org. https://arxiv.org/abs/1908.09635

Mehta, N., Pandit, A., & Shukla, S. (2019). Transforming healthcare with big data analytics and artificial intelligence: A systematic mapping study. *Journal of Biomedical Informatics*, *100*(1033113), 1–14. https://doi.org/10.1016/j.jbi.2019.103311

Merkell, S. (2020, March 17). Part 1: Artificial Intelligence Defined. *Deloitte Sweden*. https://www2.deloitte.com/se/sv/pages/technology/articles/part1-artificial-intelligence-defined.html

Miller, T. (2019). "But why?" Understanding explainable artificial intelligence. *XRDS: Crossroads, The ACM Magazine for Students*, *25*(3), 20–25. https://doi.org/10.1145/3313107

Moor, J. (1985). WHAT IS COMPUTER ETHICS? Metaphilosophy, 16(4), 266-275. Retrieved April 1, 2021, from http://www.jstor.org/stable/24436819

Moor, J. (2006). The Nature, Importance, and Difficulty of Machine Ethics. *IEEE Intelligent Systems*, *21*(4), 18–21. https://doi.org/10.1109/mis.2006.80

Moran, M. E. (2007). Evolution of robotic arms. *Journal of Robotic Surgery*, 1(2), 103–111. https://doi.org/10.1007/s11701-006-0002-x

Müller, V. C. (2020, April 30). *Ethics of Artificial Intelligence and Robotics (Stanford Encyclopedia of Philosophy)*. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/ethics-ai/#BackFiel

Murphy, K., Di Ruggiero, E., Upshur, R., Willison, D. J., Malhotra, N., Cai, J. C., Malhotra, N., Lui, V., & Gibson, J. (2021). Artificial intelligence for good health: a scoping review of the ethics literature. *BMC Medical Ethics*, 22(1). Retrieved from Scopus. https://doi.org/10.1186/s12910-021-00577-8

Naraei, P., Abhari, A., & Sadeghian, A. (2016). Application of multilayer perceptron neural networks and support vector machines in classification of healthcare data. *2016 Future Technologies Conference (FTC)*, 848–852. https://doi.org/10.1109/ftc.2016.7821702

Ni, L., Lu, C., Liu, N., & Liu, J. (2017). MANDY: Towards a Smart Primary Care Chatbot Application. *Communications in Computer and Information Science*, 780, 38–52. https://doi.org/10.1007/978-981-10-6989-5_4N. G. Paterakis, E. Mocanu, M. Gibescu, B. Stappers, and W. van Alst, "Deep learning versus traditional machine learning methods for aggregated energy demand prediction," *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, 2017, pp. 1-6, doi: 10.1109/ISGTEurope.2017.8260289.

Paterakis, N. G., Mocanu, E., Gibescu, M., Stappers, B., and Van Alst, W. "Deep learning versus traditional machine learning methods for aggregated energy demand prediction," 2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2017, pp. 1-6, doi: 10.1109/ISGTEurope.2017.8260289.

Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008). Systematic mapping studies in software engineering. *EASE'08: Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering*, 68–77. https://doi.org/10.14236/ewic/EASE2008.

Rajasekar, S., Philominathan, P., & Chinnathambi, V. (2013, October). Research Methodology. https://arxiv.org/pdf/physics/0601009.pdf

Rajkomar, A., Oren, E., Chen, K., Dai, A. M., Hajaj, N., Hardt, M., Liu, P. J., Liu, X., Marcus, J., Sun, M., Sundberg, P., Yee, H., Zhang, K., Zhang, Y., Flores, G., Duggan, G. E., Irvine, J., Le, Q., Litsch, K., . . . Dean, J. (2018). Scalable and accurate deep learning with electronic health records. *Npj Digital Medicine*, 1(18), 1–10. https://doi.org/10.1038/s41746-018-0029-1

Ramesh, A. N., Kambhampati, C., Monson, J. R. T., & Drew, P. J. (2004). Artificial intelli-gence in medicine. *Annals of The Royal College of Surgeons of England*, 86(5), 334–338. https://doi.org/10.1308/147870804290

Ratia, M., Myllärniemi, J., & Helander, N. (2018). Robotic Process Automation - Creating Value by Digitalizing Work in the Private Healthcare? *Proceedings of the 22nd International Academic Mindtrek Conference*, 222–227. https://doi.org/10.1145/3275116.3275129

Reddy, S., Allan, S., Coghlan, S., & Cooper, P. (2019). A governance model for the applica-tion of AI in health care. *Journal of the American Medical Informatics Associa-tion*, 27(3), 491-497. doi:10.1093/jamia/ocz192

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Lecture Notes in Computer Science*, *9351*, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, *115*(3), 211–252. https://doi.org/10.1007/s11263-015-0816-y

Russell, S., & Norvig, P. (2015). *Artificial Intelligence: A Modern Approach* (3rd ed.). Pearson.

Seto, E., Leonard, K. J., Cafazzo, J. A., Barnsley, J., Masino, C., & Ross, H. J. (2012). Developing healthcare rule-based expert systems: Case study of a heart failure telemonitoring system. *International Journal of Medical Informatics*, *81*(8), 556–565. https://doi.org/10.1016/j.ijmedinf.2012.03.001

Shailaja, K., Seetharamulu, B., & Jabbar, M. A. (2018). Machine Learning in Healthcare: A Review. *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 910–914. https://doi.org/10.1109/iceca.2018.8474918

Shaw, M. (2003). Writing good software engineering research papers. *25th International Conference on Software Engineering, 2003. Proceedings.*, 1–11. https://doi.org/10.1109/icse.2003.1201262

Shickel, B., Tighe, P. J., Bihorac, A., & Rashidi, P. (2018). Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis. *IEEE Journal of Biomedical and Health Informatics*, *22*(5), 1589–1604. https://doi.org/10.1109/jbhi.2017.2767063

Shih, J. J., Krusienski, D. J., & Wolpaw, J. R. (2012). Brain-Computer Interfaces in Medicine. *Mayo Clinic Proceedings*, *87*(3), 268–279. https://doi.org/10.1016/j.mayocp.2011.12.008

Sutskever, I., Vinyals, O. & Le, Q. V. Sequence to sequence learning with neural networks. *Advances in Neural Information Processing Systems*, 3104–3112 (2014).

Turja, T., Rantanen, T., & Oksanen, A. (2017). Robot use self-efficacy in healthcare work (RUSH): development and validation of a new measure. *AI & SOCIETY*, *34*(1), 137–143. https://doi.org/10.1007/s00146-017-0751-2

Vakkuri, V., & Abrahamsson, P. (2018). The Key Concepts of Ethics of Artificial Intelligence. *2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, 1–6. https://doi.org/10.1109/ice.2018.8436265

van Lent, M., Fisher, W., & Mancuso, M. (2004, July). An Explainable Artificial Intelligence System for Small-unit Tactical Behavior. *In Proceedings of the 16th Conference on Innovative Applications of Artificial Intelligence,* 900–907. https://doi.org/10.5555/1597321.1597342

Velasquez, M., Andre, C., Shanks, T., & Meyer, M. J. (2010, January 1). *What is Ethics?* Markkula Center for Applied Ethics. https://www.scu.edu/ethics/ethics-resources/ethical-decision-making/what-is-ethics/

Weiss, S., Kulikowski, C. A., & Safir, A. (1978). Glaucoma consultation by computer. *Computers in Biology and Medicine,* 8(1), 25–40. https://doi.org/10.1016/0010-4825(78)90011-2

Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45. https://doi.org/10.1145/365153.365168

Weller, A. (2017, July 29). *Transparency: Motivations and Challenges*. ArXiv.Org. https://arxiv.org/abs/1708.01870

Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., Myers West, S., Richardson, R., Schultz, J., & Schwartz, O. (2018, December). *AI Now Report 2018*. AI Now Institute. https://ainowinstitute.org/AI_Now_2018_Report.pdf

Wieringa, R., Maiden, N., Mead, N., & Rolland, C. (2005). Requirements engineering paper classification and evaluation criteria: a proposal and a discussion. *Requirements Engineering*, *11*(1), 102–107. https://doi.org/10.1007/s00766-005-0021-6

Wieringa, M. (2020). What to account for when accounting for algorithms. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 1–18. https://doi.org/10.1145/3351095.3372833

Woudstra, F. (2020, April 9). *Ethical Guidelines for Transparent Development and Implementation of AI - an Overview*. Filosofie in Actie. https://www.filosofieinactie.nl/blog/2020/4/9/ethical-guidelines-for-transparent-development-and-implementation-of-ai

Appendices

A Final Sample of the SMS Process N=56

Publication Year	Author	Title	DOI
2021	Pot, Mirjam; Kieusseyan, Nathalie; Prainsack, Barbara	Not all biases are bad: equitable and inequitable biases in machine learning and radiology	10.1186/s13244-020-00955-7
2020	Coin, Allen; Mulder, Megan; Dubljevic, Veljko	Ethical Aspects of BCI Technology: What Is the State of the Art?	10.3390/philosophies5040031
2020	Thieme, Anja; Belgrave, Danielle; Doherty, Gavin	Machine Learning in Mental Health: A Systematic Review of the HCI Literature to Support the Development of Effective and Implementable ML Systems	10.1145/3398069
2019	Beil, Michael; Proft, Ingo; van Heerden, Daniel; Sviri, Sigal; van Heerden, Peter Vernon	Ethical considerations about artificial intelligence for prognostication in intensive care	10.1186/s40635-019-0286-6
2019	Park, Seong Ho; Kim, Young-Hak; Lee, Jun Young; Yoo, Soyoung; Kim, Chong Jai	Ethical challenges regarding artificial intelligence in medicine from the perspective of scientific editing and peer review	10.6087/kcse.164

2019	Koegel, Johannes; Schmid, Jennifer R.; Jox, Ralf J.; Friedrich, Orsolya	Using brain-computer interfaces: a scoping review of studies employing social research methods	10.1186/s12910-019-0354-1
2019	Galbusera, Fabio; Casaroli, Gloria; Bassani, Tito	Artificial intelligence and machine learning in spine research	10.1002/jsp2.1044
2021	Thapa, Chandra; Camtepe, Seyit	Precision health data: Requirements, challenges and existing techniques for data security and privacy	10.1016/j.compbiomed.2020.104130
2021	Ursin, Frank; Timmermann, Cristian; Steger, Florian	Ethical Implications of Alzheimer's Disease Prediction in Asymptomatic Individuals Through Artificial Intelligence	10.3390/diagnostics11030440
2020	Abràmoff, Michael D; Tobey, Danny; Char, Danton S	Lessons Learned About Autonomous AI: Finding a Safe, Efficacious, and Ethical Path Through the Development Process	10.1016/j.ajo.2020.02.022
2020	Smith, Maxwell J; Axler, Renata; Bean, Sally; Rudzicz, Frank; Shaw, James	Four equity considerations for the use of artificial intelligence in public health	10.2471/BLT.19.237503
2020	Bærøe, Kristine; Miyata-Sturm, Ainar; Henden, Edmund	How to achieve trustworthy artificial intelligence for health	10.2471/BLT.19.237289
2020	Vaisman, Alon; Linder, Nina; Lundin, Johan; Orchanian-Cheff, Ani; Coulibaly, Jean T; Ephraim, Richard K D; Bogoch, Isaac I	Artificial intelligence, diagnostic imaging and neglected tropical diseases: ethical implications	10.2471/BLF.19.237560

2020	Vollmer, Sebastian; Mateen, Bilal A; Bohner, Gergo; Király, Franz J; Ghani, Rayid; Jonsson, Pall; Cumbers, Sarah; Jonas, Adrian; McAllister, Katherine S L; Myles, Puja; Granger, David; Birse, Mark; Branson, Richard; Moons, Karel G M; Collins, Gary S; Ioannidis, John P A; Holmes, Chris; Hemingway, Harry	Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness	10.1136/bmj.l6927
2020	Amann, Julia; Blasimme, Alessandro; Vayena, Effy; Frey, Dietmar; Madai, Vince I; the Precise4Q consortium	Explainability for artificial intelligence in healthcare: a multidisciplinary perspective	10.1186/s12911-020-01332-6
2020	Shuaib, Abdullah; Husain Arian; Shuaib, Ali	The Increasing Role of Artificial Intelligence in Health Care: Will Robots Replace Doctors in the Future?	10.2147/IJGM.S268093
2020	Esmaeilzadeh, Pouyan	Use of AI-based tools for healthcare purposes: a survey study from consumers' perspectives	10.1186/s12911-020-01191-1
2020	Favaretto, Maddalena; Shaw, David; De Clercq, Eva; Joda, Tim; Elger, Bernice Simone	Big Data and Digitalization in Dentistry: A Systematic Review of the Ethical Issues	10.3390/ijerph17072495
2019	Geis, J Raymond; Brady, Adrian; Wu, Carol C; Spencer, Jack; Ranschaert, Erik; Jaremko, Jacob L; Langer, Steve G; Kitts, Andrea Borondy; Birch, Judy; Shields, William F; Robert van den Hoven van Genderen; Kotter, Elmar; Gichoya, Judy Wawira; Cook, Tessa S;	Ethics of artificial intelligence in radiology: summary of the joint European and North American multisociety statement	10.1186/s13244-019-0785-8

	Morgan, Matthew B; Tang, An; Safdar, Nabile M; Kohli, Marc		
2019	Lawrie, Stephen M; Fletcher-Watson, Sue; Whalley, Heather C; McIntosh, Andrew M	Predicting major mental illness: ethical and practical considerations	10.1192/bjo.2019.11
2019	Graham, Sarah; Depp, Colin; Lee, Ellen E.; Nebeker, Camille; Tu, Xin; Kim, Ho-Cheol; Jeste, Dilip V.	Artificial Intelligence for Mental Health and Mental Illnesses: an Overview.	10.1007/s11920-019-1094-0
2019	Price, W. Nicholson 2nd; Cohen, I. Glenn	Privacy in the age of medical big data.	10.1038/s41591-018-0272-7
2018	Mooney, Stephen J.; Pejaver, Vikas	Big Data in Public Health: Terminology, Machine Learning, and Privacy.	10.1146/annurev-publhealth-040617-014208
2020	Loftus, Tyler J.; Tighe, Patrick J.; Filiberto, Amanda C.; Efron, Philip A.; Brakenridge, Scott C.; Mohr, Alicia M.; Rashidi, Parisa; Upchurch, Gilbert R. Jr; Bihorac, Azra	Artificial Intelligence and Surgical Decision-making.	10.1001/jamasurg.2019.4917
2020	Schwendicke, F.; Samek, W.; Krois, J.	Artificial Intelligence in Dentistry: Chances and Challenges.	10.1177/0022034520915714
2020	Du-Harpur, X.; Watt, F. M.; Luscombe, N. M.; Lynch, M. D.	What is AI? Applications of artificial intelligence to dermatology.	10.1111/bjd.18880
2017	Bonderman, Diana	Artificial intelligence in cardiology.	10.1007/s00508-017-1275-y

2019	Vellido, Alfredo	Societal Issues Concerning the Application of Artificial Intelligence in Medicine.	10.1159/000492428
2018	Wahl, Brian; Cossy-Gantner, Aline; Germann, Stefan; Schwalbe, Nina R.	Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings?	10.1136/bmjgh-2018-000798
2020	Pépin, Jean-Louis; Bailly, Sébastien; Tamisier, Renaud	Big Data in sleep apnoea: Opportunities and challenges.	10.1111/resp.13669
2020	Pedersen, Mangor; Verspoor, Karin; Jenkinson, Mark; Law, Meng; Abbott, David F.; Jackson, Graeme D.	Artificial intelligence for clinical decision support in neurology.	10.1093/braincomms/fcaa096
2017	Burwell, Sasha; Sample, Matthew; Racine, Eric	Ethical aspects of brain computer interfaces: a scoping review.	10.1186/s12910-017-0220-y
2020	Six Dijkstra, Marianne W. M. C.; Siebrand, Egbert; Dorrestijn, Steven; Salomons, Etto L.; Reneman, Michiel F.; Oosterveld, Frits G. J.; Soer, Remko; Gross, Douglas P.; Bieleman, Hendrik J.	Ethical Considerations of Using Machine Learning for Decision Support in Occupational Health: An Example Involving Periodic Workers' Health Assessments.	10.1007/s10926-020-09895-x
2021	Lepri, Bruno; Oliver, Nuria; Pentland, Alex	Ethical machines: The human-centric use of artificial intelligence.	10.1016/j.isci.2021.102249
2021	Cui, M.; Zhang, D.Y.	Artificial intelligence and computational pathology	10.1038/s41374-020-00514-0
2021	Lebcir, R.; Hill, T.; Atun, R.; Cubric, M.	Stakeholders' views on the organisational factors affecting application of artificial	10.1136/bmjopen-2020-044074

		intelligence in healthcare: A scoping review protocol	
2021	Jiang, L.; Wu, Z.; Xu, X.; Zhan, Y.; Jin, X.; Wang, L.; Qiu, Y.	Opportunities and challenges of artificial intelligence in the medical field: current application, emerging problems, and problem-solving strategies	10.1177/03000605211000157
2021	Sumiyama, K.; Futakuchi, T.; Kamba, S.; Matsui, H.; Tamai, N.	Artificial intelligence in endoscopy: Present and future perspectives	10.1111/den.13837
2021	Kellmeyer, P.	Artificial intelligence in basic and clinical neuroscience: Opportunities and ethical challenges	10.1515/nf-2019-0018
2020	Nagaraj, S.; Harish, V.; McCoy, L.G.; Morgado, F.; Stedman, I.; Lu, S.; Drysdale, E.; Brudno, M.; Singh, D.	From Clinic to Computer and Back Again: Practical Considerations When Designing and Implementing Machine Learning Solutions for Pediatrics	10.1007/s40746-020-00205-4
2020	Zhang, J.; Oh, Y.J.; Lange, P.; Yu, Z.; Fukuoka, Y.	Artificial intelligence chatbot behavior change model for designing artificial intelligence chatbots to promote physical activity and a healthy diet: Viewpoint	10.2196/22845

2020	Rowe, J.P.; Lester, J.C.	Artificial Intelligence for Personalized Preventive Adolescent Healthcare	10.1016/j.jadohealth.2020.02.021
2020	He, M.; Li, Z.; Liu, C.; Shi, D.; Tan, Z.	Deployment of Artificial Intelligence in Real-World Practice: Opportunity and Challenge	10.1097/APO.0000000000000301
2020	Alami, H.; Lehoux, P.; Auclair, Y.; de Guise, M.; Gagnon, M.-P.; Shaw, J.; Roy, D.; Fleet, R.; Ahmed, M.A.A.; Fortin, J.-P.	Artificial intelligence and health technology assessment: Anticipating a new level of complexity	10.2196/17707
2020	Patel, R.; Ashcroft, J.; Darzi, A.; Singh, H.; Leff, D.R.	Neuroenhancement in surgeons: benefits, risks and ethical issues	10.1002/bjs.11601
2020	Asan, O.; Bayrak, A.E.; Choudhury, A.	Artificial Intelligence and Human Trust in Healthcare: Focus on Clinicians	10.2196/15154
2020	Cuzzolin, F.; Morelli, A.; Cîrstea, B.; Sahakian, B.J.	Knowing me, knowing you: Theory of mind in AI	10.1017/S0033291720000835
2020	Garcelon, N.; Burgun, A.; Salomon, R.; Neuraz, A.	Electronic health records for the diagnosis of rare diseases	10.1016/j.kint.2019.11.037
2020	Grote, T.; Berens, P.	On the ethics of algorithmic decision-making in healthcare	10.1136/medethics-2019-105586
2020	Mackey, T.K.; Cuomo, R.E.	An interdisciplinary review of digital technologies to facilitate anti-corruption, transparency and accountability in medicines procurement	10.1080/16549716.2019.1695241

2020	Starke, G.; De Clercq, E.; Borgwardt, S.; Elger, B.S.	Computing schizophrenia: Ethical challenges for machine learning in psychiatry	10.1017/S0033291720001683
2020	Ahmed, Z.; Mohamed, K.; Zeeshan, S.; Dong, X.	Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine	10.1093/database/baaa010
2019	Stanfill, M.H.; Marc, D.T.	Health Information Management: Implications of Artificial Intelligence on Healthcare Data and Information Management	10.1055/s-0039-1677913
2019	Car, J.; Sheikh, A.; Wicks, P.; Williams, M.S.	Beyond the hype of big data and artificial intelligence: Building foundations for knowledge and wisdom	10.1186/s12916-019-1382-x
2019	Saba, L.; Biswas, M.; Kuppili, V.; Cuadrado Godia, E.; Suri, H.S.; Edla, D.R.; Omerzu, T.; Laird, J.R.; Khanna, N.N.; Mavrogeni, S.; Protogerou, A.; Sfikakis, P.P.; Viswanathan, V.; Kitas, G.D.; Nicolaides, A.; Gupta, A.; Suri, J.S.	The present and future of deep learning in radiology	10.1016/j.ejrad.2019.02.038
2019	Bezemer, T.; De Groot, M.C.H.; Blasse, E.; Ten Berg, M.J.; Kappen, T.H.; Bredenoord, A.L.; Van Solinge, W.W.; Hoefer, I.E.; Haitjema, S.	A human(E) factor in clinical decision support systems	10.2196/11732