University of Zurich[UZH]

20. October 2020

The Third Wave AI

A Framework to Classify Artificial Intelligence

Master's Thesis

William Martini
Dürnten, Switzerland

13-765-334
william.martini@uzh.ch

Advisor: Dario Stähelin

Prof. Dr. Gerhard Schwabe
Department of Informatics
University of Zurich
https://www.ifi.uzh.ch/imrg

Abstract

The Master's thesis takes a closer look at the term *Third Wave AI*. The Defense Advanced Research Projects Agency (DARPA) introduced a framework, where the evolution of artificial intelligence (AI) is categorised with waves. According to DARPA, the next wave will be the third one, which is based on contextual adaptation. Based on the initial literature review, the issue of the legitimacy of DARPA's framework was raised. To address this topic, different types of analyses were performed. The public discourse analysis was used to identify central entities and dominating topics. They showed a lack of common definitions and intermixing of *Third Wave AI* related terms. Since the results revealed the need for adjustments, a new framework was suggested, challenged and revised in a sensemaking workshop. For further improvements, interviews with industry and AI experts were conducted. They validated the identified problems and many suggestions for enhancements were collected. With these insights, the framework's final version was created. The main component consists of a technical and functional layer, which are each segregated in dimensions representing the requirements. On the functional layer, the dimensions human-centered, explainable and contextual adaption were set as prerequisites. The technical layer consists of DARPA's requirements, which are perceiving, learning, abstracting and reasoning. Furthermore, the main component was expanded with a zoom-out and zoom-in. For the framework's application, a spider diagram is presented. By introducing a new approach with the corresponding definitions, the identified needs for action are remedied, and the legitimacy of *Third Wave AI* will be improved.

Zusammenfassung

Die Masterarbeit beschäftigt sich mit dem Begriff *Third Wave AI*. Die Defense Advanced Research Projects Agency (DARPA) führte ein Framework ein, in dem die Entwicklung der künstlichen Intelligenz (KI) mit Wellen beschrieben wird. Laut DARPA wird die nächste Welle die Dritte sein, welche durch die kontextuelle Anpassung begründet wird. Auf der Grundlage einer ersten Literaturrecherche wurde die Frage nach der Legitimität des DARPA-Frameworks aufgeworfen. Um dieser nachzugehen, wurden verschiedene Arten von Analysen durchgeführt. Die Analyse des öffentlichen Diskurses diente dazu, zentrale Entitäten und dominierende Themen zu identifizieren. Es offenbarte sich ein Mangel an allgemeingültigen Definitionen und eine Vermischung von *Third Wave AI* Begriffen. Da die Ergebnisse einen Anpassungsbedarf aufzeigten, wurde ein neues Framework vorgeschlagen. Dieses wurde in einem Sensemaking Workshop infrage gestellt und überarbeitet. Zur weiteren Verbesserung wurden Interviews mit Experten aus der Industrie und dem Feld der KI durchgeführt. Diese validierten die identifizierten Probleme und es wurden viele Verbesserungsvorschläge gesammelt. Mit den zahlreichen Erkenntnissen wurde die endgültige Version des Frameworks erstellt. Das Hauptframework besteht aus einer technischen und einer funktionellen Ebene. Diese sind in Dimensionen unterteilt, welche die Anforderungen an die jeweilige Ebene darstellen. Auf der funktionalen Ebene wurden die Dimensionen *human-centered*, *explainable* und *contextual adaption* als Anforderungen festgelegt. Die technische Schicht besteht aus den DARPA-Anforderungen Wahrnehmen, Lernen, Abstrahieren und Begründen. Darüber hinaus wurde das Hauptframework um einen Zoom-out und Zoom-in erweitert. Für die Anwendung des Frameworks wurde ein Netzdiagramm ausgearbeitet. Durch die Einführung eines neuen Frameworks mit den entsprechenden Definitionen wird der identifizierte Handlungsbedarf adressiert und die Legitimität von *Third Wave AI* verbessert.

Acknowledgements

First of all, my thanks go to Dario Stähelin for his supervision, guidance and commitment during this project. Additionally, I would like to thank Dr. Mateusz Dolota for his valuable inputs and support. Thanks also go to Prof. Dr. Gerhard Schwabe for his support, inputs and for allowing me to work on this topic.

At this point, I would also like to express my special thanks to Gabriele Schwarz, Sandra Völler, Bharath Ramesh, Gniewosz Leliwa, Markus Christen, Pascal Hagedorn, Peter Voss, Gniewosz Leliwa, Steve Furber, Thomas Neuhaus and Umberto Michelucci for their time and willingness to answer my interview questions.

Special thanks also go to Juan Martini and Vanessa Rüegsegger for proofreading and the support over the past few months.

Last but not least, I want to thanks Florian Ruosch and Patrick Düggelin for being valuable discussion partners and contributing their expertise.

Contents

List of Tables

List of Figures

List of Abbreviations

AGI	artificial general intelligence
AI	artificial intelligence
ASI	artificial super intelligence
CSV	comma-separated values
DARPA	Defense Advanced Research Projects Agency
HAI	human-centered AI
HITL	human-in-the-loop
IoT	Internet of Things
PoC	proof of concept
QDA	qualitative data analysis
XAI	explainable AI

1

Introduction

Whenever somebody speaks of Third Wave AI, they refer mainly to the video called *A DARPA Perspective on Artificial Intelligence* by Launchbury (2017). It is the only source, where Defense Advanced Research Projects Agency (DARPA) explains their framework. Therefore, all used information in this master thesis about the framework is exclusively from Launchbury (2017). In the video, the artificial intelligence (AI) evolution is categorised in three waves. An overview of the framework is illustrated in Table 2.1. The first and second wave describe the historical evolution of AI. In contrast, DARPA predicts with their Third Wave AI how AI will evolve in the next years. Since it is only a presumption, the term should be used with caution and should be critically reviewed.

However, when the term Third Wave AI is applied, reference is always made to the video of DARPA. For example, this is the case with Beinart (2019), Daws (2018), Perez (2017) and Scott (2018). They use the DARPA's framework but without questioning it nor comparing it to other frameworks. For this reason, the question arises how legitimate this framework is. This fundamental question is the subject of this master thesis. To answer it, the following research questions will be addressed.

(i) Which frameworks exist to classify AI?

(ii) What public discourse exists about Third Wave AI?

(iii) How can the Third Wave framework be defined?

(iv) How can the framework be applied, for instance, to support collaboration between humans and AI?

Since a lack of a historical categorisation of AI exists, there is a vast potential for a generally usable framework, which can be applied to categories AI. A framework for AI like Industry 4.0 is missing in the information systems research. To answer the research questions, there should be either the DARPA's adjusted framework or a completely new one, which should serve the purpose of providing a terminological foundation. This foundation would help the academia and industry, making AI more approachable since the historical categorisation is easy to understand and intuitive.

The structure of the thesis is based on the order of the research questions. (i) is addressed in *Related Work* (Chapter 2). Additionally, the relevant literature about the term Third Wave AI is reviewed, related theories are shown, and the definitions used in this thesis are introduced. The chapter *Methodology* (Chapter 3) considers the chosen approach and how the data collection and analysis has been performed especially for the analysis of the public discourse. The findings are disclosed in *Results* (Chapter 4). This includes all collected outcomes on the quantitative and qualitative analysis of the public discourse, the sensemaking workshop and the interviews. With these results, the research questions (ii) and (iii) are discussed, and a comprehensive framework is introduced. Furthermore, a concrete application example is presented. The application is used to answer (iv). In *Discussion* (Chapter 5), the added value of this Master's thesis for the academia and the industry is demonstrated. Moreover, the framework and the legitimacy of the term Third Wave AI are discussed in the context of all collected results and related literature. Finally, in the last chapter *Limitations and Future Work* (Chapter 6), the thesis' limitations are reviewed, and how follow-up projects can build upon this thesis.

2

Related Work

As controversial as the definition of AI are the frameworks. Therefore an overview of the related literature is given. This chapter is split into four sections. First, DARPA's framework will be reviewed. To be able to compare DARPA's framework to others, an analysis of different frameworks will be performed. Afterwards, multiple definitions will be introduced to build a foundation for further analysis and discussions. Finally, a closer look will be taken at other theories, which are not related to AI. The theories will allow an approach as open-minded as possible.

2.1 Third Wave AI

DARPA's waves can be compared to the Industry 4.0 framework. Each level involves a fundamental change in the industry. The waves presented in Table 2.1 are reasoned by fundamental changes in AI. The individual characteristics are rated on a scale from 0 to 100. 0 means the wave does not contain the characteristic and 100 it masters it on a human-like level. The first wave contains rule-based systems, where the developer of the AI says exactly how the AI should react in certain situations. Since the developer predetermines everything, the decisions of the AI are comprehensible. However, these AIs only consider environmental factors, which it is instructed to and can neither learn new behaviour nor can abstract information. The second wave is based on statistical learning. The statistical approach allows the Second Wave AI to react better on new inputs and to perform increasingly accurate predictions with more data. However, this is at the expense of comprehensibility, since the learned behaviour is hardly traceable. As in the first wave, the second wave can only abstract poorly and is always programmed for one purpose only. Additionally, there are several challenges, which are faced by the second wave and can not be solved with statistical learning. For example, one of the biggest challenges is the data needed for training. The training data has to be of high quality and in huge quantities. DARPA characterises the third wave by contextual adaptation.

	First Wave Handcrafted Knowledge	**Second Wave** Statistical Learning	**Third Wave** Contextual Adaptation
Perceiving (0-100)	25	75	75
Learning (0-100)	0	75	75
Abstracting (0-100)	0	25	50
Reasoning (0-100)	75	25	75
Example	TurboTax	Tay (Twitter Bot)	-

Table 2.1: The three waves of AI (Launchbury, 2017)

This kind of AI should perform better in the areas of abstracting and reasoning than the second wave. Such characteristics should be achieved primarily by enabling AI to explain its decisions. These AIs should also be able to abstract better and thus require less data for training.

To achieve contextual adaptation, DARPA attempts to use contextual explanatory models. Third Wave AI systems should be able to build models based on real-world phenomena. The example provided by DARPA is a system for identifying cats. It does not recognise the cat by probability as it would be the case with statistical AIs. Moreover, the system identifies the cat based on its characteristics like its nose, its ears, its tail and its fur. These characteristics are part of the contextual explanatory model, which lead to the identification of the cat as a cat. The big benefit of these systems is the use of these explainable models. Since the explanation does already exist, it can be provided to the user. Such an explanation would help the user to understand why the system thinks the presented animal is a cat. Unfortunately, DARPA does neither present an implemented solution nor deliver a concept on how to create such contextual explanatory models.

A framework, which expands DARPA's framework with an additional wave, is presented by Scott (2018). The first three waves remain identical with the exception of Scott (2018) adding more detailed descriptions of each wave. He defines contextual adaptation as one of the key characteristics but puts it equally to explainability. The mix-up of contextual adaptation and explainability frequently occurs in the literature. In consequence of the intermixing, it was decided not to include the framework in the analysis. It would be impossible to build upon the framework if there is a need for such fundamental changes. Despite contextual adaptation, he adds communication in natural language as a characteristic of Third Wave AI. Scott (2018) predicts the fourth wave will be motivated by artificial general intelligence (AGI), which could lead to an artificial super intelligence (ASI). Additionally, it is the only framework which states a concrete schedule. He estimates Third Wave AI between 2020 and 2030.

2.2 AI Frameworks

There are countless heterogeneous frameworks with different approaches and focuses. To gain an overview for this Master's thesis, the most relevant one has been collected and summarized in a factsheet. The factsheet can be found in the Appendix A. In total, the factsheet contains ten distinct frameworks. For every framework, four out of the Five Ws are listed. *Why* was left out because to answer this question, a very detailed analysis of each framework would have been necessary, and the added value would have been marginal. The other Ws help to compare the frameworks based on their context. For this reason, the document type of each framework is listed. The overview consists of four frameworks from books, two from papers, one from an article, two from websites and one from Medium. The main focus was set on academic frameworks, but to achieve the broadest possible overview, frameworks from non-academic sources like Medium were added. Additionally, for every framework, the approach has been identified to reflect the focus of each framework. Interestingly, the approaches are very different. Some of the approaches are located on a very technical or philosophical layer. The frameworks from non-academic source are distinguished as more application-oriented, like Zilis and Cham (2016).

A framework which shows parallels to DARPA's framework is the one of Doloc (2019). It categorizes AI based on the AI winters. They call it first, second and third AI and the two last phases are initialized by passing an AI winters. However, all other frameworks collected in the factsheet differ from DARPA's approach. Frameworks as Dodhiawala et al. (1989) or Yampolskiy (2013) mainly focus on providing requirements for their categories. Whereas DARPA focuses on one concrete technology, the contextual explanatory models. Additional frameworks, which use DARPA's waves and have been collected during the public discourse, are listed in Appendix C. They will be discussed in subsection 4.1.2.

2.3 Definitions of Relevant Terms

To answer the research question, it is necessary to not only look at the definition of Third Wave AI. The related terms' definitions also have a great impact on how Third Wave AI is discussed. Therefore, the following definitions are introduced to provide a foundation for the analysis. The most used terms in the context of Third Wave AI are contextual adaptation and explainable AI (XAI). DARPA's uses contextual adaptation as the descriptive term for Third Wave AI without providing a definition. Therefore, it was required to introduce a definition for this Master's thesis. The used definition is provided by Popovic (2001) and is not related to AI. He defines contextual adaptation "to react to changes in the environment and adapt the application behaviour accordingly" (p. 2). Contextual adaptation consists of two steps. First, the system has to perceive its context, situation and environment. Especially, it should recognise changes around itself. Second, the AI should adapt to these changes and continuously repeat these two steps.

There are many definitions of XAI available, but for this thesis, two were applied. For Xu (2019) does the "explainable AI (XAI) enables users to understand the algorithm and parameters used" (p. 44). Additionally, Barredo Arrieta et al. (2020) describe XAI as "one that produces details or reasons to make its functioning clear or easy to understand" (p. 85). The two definitions were used because they complement each other. Xu (2019) places the user in focus since the goal of XAI should be to provide an explanation to a user, which he would understand. Furthermore, an important point mentioned in the definition is the relevance of disclosure. The used parameters play an enormous role in the outcome. Barredo Arrieta et al. (2020) add with their definition the focus on how the provided explanation should be. They specify that the explanation should be easy to understand. This requirement is necessary to build an XAI, which is actually used. Providing an explanation to the user is not enough, since it remains crucial that the explanation can be quickly and easily understood.

Besides contextual adaptation and XAI, other terms are also relevant for the analysis such as human-centered AI (HAI) and human-in-the-loop (HITL). The definition of HAI used for this project was introduced by Xu (2019). His framework contains three elements: *ethically aligned design*, *technology enhancement* and *human factors design*. All these three components lead to HAI. *Ethically aligned design* includes ethical topics like fairness and justice. The premise of this component is to ensure that AI does not replace humans. *Technology enhancement* aims to achieve human-like intelligence. *Human factors design* includes requirements as explainability, comprehensibility or usability.

HITL is a term, which appears quite often in the related literature. Nevertheless, very different definitions are used. In the first part of this thesis, the following definition was used:

"HITL describes the process when the machine or computer system is unable to solve a problem, needs human intervention like involving in both the training and testing stages of building an algorithm, for creating a continuous feedback loop allowing the algorithm to give every time better results." (Bisen, 2020)

This definition is a very narrow one and focuses only on how to solve edge cases with a human. In the initial phase of the thesis, this framework fulfilled its purpose. Due to the results shown in Section 4.2, it was necessary to change the definition. Therefore, another definition was used for the second part. The new definition is based on interactive machine learning. (Holzinger, 2016) describes interactive machine learning as "algorithms that can interact with agents and can optimize their learning behaviour through these interactions, where the agents can also be human" (p. 1). The key point of this definition is the optimisation through interaction. In this thesis, HITL is taken even one step further. The system should recognise the needs of an individual human and optimise itself based on this information.

2.4 Related Theories

In order to broaden the discussion, it is advisable to look beyond the AI literature. Therefore, two theories were added to the analysis to put Third Wave AI in a more global context. The first theory is the sociotechnical systems of (Eric Trist et al., 1993). Their theory is about social and technical systems. Since the theory was originated from English coal mines, the social system consisted of the workers and the technical system of the production process. In general, the social system focus on the attributes of humans like skills and the technical system focus on processes and tasks (Militello et al., 2013). The goal is to achieve joint optimisation by optimising the interface between social and technical system since both parts need positive outcomes from an adjustment (Appelbaum, 1997). An example of such a joint optimisation is the user interface of computer information systems (Maguire, 2014). Since the interaction between humans and AI plays an important role, there should be many publications about this topic. However, this is not the case. This lack could be reasoned by the research's current focus on only technical aspects of AI. The contextual explanatory models of DARPA's framework are an improvement of the technical system. XAI has the potential to be a joint optimisation by enhancing both social and technical system. For the social system, the explanation helps the user to understand and to use the AI. However, the consideration of the social system alone is not carried out. For this reason, DARPA considered the user or human only to a negligible extent and therefore, ignored the question of which user needs should be satisfied with the solution.

The second theory is the Ladder of Causation of Pearl and Mackenzie (2018). The theory contains a model with three levels, which build upon themselves. The lowest level is called *association* and includes the activities seeing and observing. Therefore, systems or organisms on this level are able to recognise patterns. On the next level *intervention*, the activities evolve from seeing to doing and intervening. *Intervention* leads to the ability to become aware of the consequences of an action. The final level is *counterfatuals* with the activities being imagining, retrospection and understanding. Pearl and Mackenzie (2018) describe this level with the question: "What if I had done ..." (p. 32). Such organisms can envision alternative outcomes of events that took place and thus can predict the outcome of future events.

Pearl and Mackenzie (2018) set robots on the lowest step of the ladder. If AI would be assimilated to the robots or viewed as a robot's component, this would imply AI interacting on the same level. Looking at Table 2.1, the high value for perceiving supports the classification. However, the second wave is also rated with 75, despite the perceiving being limited to configured parameters. As a result, the AI recognises only patterns based on the predefined information. Additionally, current AI systems observe exclusively during the training of the model. Therefore, the step *association* is only taken halfway. Despite that, the main focus of the Second Wave AI is doing, which is an activity of the second level *intervention*. Thereby the reaching of the *association* step is skipped. This situation could indicate one reason why AI's benefits are limited, and they are only able to solve simple tasks. How Third Wave AI could improve this issue is discussed in Chapter 5.

3

Methodology

The Master's thesis consisted of two main phases. The goal of the first main phase was to answer (i), (ii) and (iii) based mainly on an analysis of the public discourse. By analysing the public discourse, it was possible to identify how the term Third Wave AI is used and if there is a need for action, for example, to make the term more applicable. The second main phase, interviews were conducted to answer the research question (iv). With the interviews, an attempt was made to perform a real-world comparison. These interviews helped to analyse whether the framework is well-known among the experts or which adjustments are necessary for use in business practice. Based on the results from the public discourse analysis, an improved framework was suggested. The first version of the improvements was challenged with a sensemaking workshop. The revised framework was then used in the interviews. The interviews' goal consisted of getting insights into the practice, validating and developing the proposed framework further. The used methodologies in each phase for the data collection and analysis is described in the following sections.

3.1 Data Collection

This Master's thesis uses multiple techniques for data collection. To analyse the public discourse, a quantitative and qualitative analysis were performed. For the quantitative analysis, Sopus was used. Scopus' function *analyse search results* and *export* were utilised to identify and export the relevant data as comma-separated values (CSV). The exact query of each analysis will be exhibited in the corresponding part of subsection 4.1.1. Scopus proved to be the best choice because it provides easy to use functions and the data, which is needed for the analysis. Other tools like Publish and Perish were tried, but they often included a huge amount of irrelevant or wrong results. Therefore, it was not possible to obtain any insights as to the evolution of Third Wave AI. The CSVs were exported on the 30th of June 2020.

Third	wave	AI
Four	revolution	Artificial intelligence
Contextual Adaptation	stage	machine learning
Dritte	Welle	DARPA
Vier	Revolution	John Launchbury
Kontextuelle Adaption	Phase	Six Kin Development
		KI
		Künstliche Intelligenz
		Maschinelles Lernen

Table 3.1: First draft of the keywords

For the qualitative analysis, an analogical methodology was performed like Zavolokina et al. (2016). The reason for this decision is the similar structure of both projects, and the same type of results should be achieved. For the data collection, the relevant keywords were identified based on the collected insights out of the qualitative analysis and the initial literature review. The first draft shown in Table 3.1 was mainly based on DARPA's video and needed some adjustments which are shown in Table 3.2.

third*	wave	ai
3rd	waves	artificial intelligence
contextual adaption		a.i.
		a. i.
		DARPA

Table 3.2: Final keywords used for the data collection

One of the biggest adjustments consisted of excluding German keywords. Initial tests revealed a lack of German literature about Third Wave AI. Therefore, the decision was made to exclude all other languages besides English from the analysis. Furthermore, the keywords *Four* and *Six Kin Development* were identified as not beneficial for the analysis. They were included in the first draft because of Scott (2018), which introduced a more comprehensive framework with an additional wave. The first three waves are based on DARPA. Once again, the initial tests revealed with these keywords, no additional relevant documents were added in the results. *John Launchbury* was also identified as not valuable because the use during the tests did not change the results. The terms *revolution*, *stage* and *machine learning* were removed because they added to much noise to the results. Nevertheless, the only terms which were added are *3rd*, *a.i.* and *a. i.* to take into account the different ways of spelling.

Scientific Literature	**News Articles**	**Grey Literature**	**Blogs**
Scopus	AI News	Google	Medium
ACM Digital Library	Fortune	DARPA	Twitter
AIS Electronic Library	Forbes	company websites	GitHub
IEEE Xplore Digital Library	Analytics Insight		
Microsoft Academic	SiliconANGLE		
Google Scholar			

Table 3.3: Databases classified by categories

The keywords from Table 3.2 were used to build queries to search in the databases shown in Table 3.3. For example, for Scopus, the following query was used:

TITLE-ABS-KEY((third OR "3rd" OR three OR "contextual adaption") AND (wave OR waves) AND (ai OR "artificial intelligence" OR a.i. OR "A. I." OR "DARPA")) AND NOT (harmonic OR surface OR wavelet OR acoustic OR frequency OR 5g OR "brain wave" OR ultrason* OR waveform OR "pulse wave" OR "wave filter" OR electromagnetic) AND (LIMIT-TO(LANGUAGE, "English"))*

The search was performed on titles, abstracts and keywords. To decrease the high number of irrelevant documents, several keywords were added to the exclusion criteria. The keywords consist mainly of terms related to wave. By including the term wave to the query, a significant number of documents not related to Third Wave AI was added. There are many types of waves, like water or acoustic waves. In natural sciences, AI is used for optimisation or other use cases. Some examples are presented in subsection 4.1.1. The publications about such topics are not relevant to this analysis. Therefore, the keywords shown in the query were used to reduce the number of these kinds of publications. Not all irrelevant publications could be eliminated because the keywords had not to be too general. Otherwise, relevant documents would have been excluded. Due to the low number of results, this would have had a more negative impact than including irrelevant ones.

As analogous as possible queries were used for each database if feasible. Especially the non-scientific databases like GitHub did not have such search options, and the search was manually performed as close as possible. The search in the scientific literature databases was performed on titles, abstracts and keywords. For the other types of databases, the provided features like in GitHub searching over Readmes or topics were used.

The databases illustrated in Table 3.3 were used for the data collection to analyse the public discourse. To achieve broad coverage, publications out of scientific literature, news articles, grey literature and blogs were collected. In Table 3.3, the used databases are classified by these categories. The databases were selected based on experience, initial tests and examples from literature like Zavolokina et al. (2016). For the collection of news articles, many other platforms were analysed, but only the ones are listed where relevant articles were found. The company websites are summarised in the table because otherwise, the table would be too large. Many company websites were checked based on their relevance in AI or the industry like Accenture, KPMG, Gartner, IBM and Intel. For the collection of tweets, different tools were tested like COSMOS and ATLAS.ti. All the reviewed tools did not provide the data needed for the analysis. Therefore, a manual collection based on hashtag was performed with TweetDeck.

The sensemaking workshop was structured based on the defined goals. The goals consisted of the collection of feedback on the acquired results, the suggested framework's validation, the identification of need for action and the creation of a foundation for this thesis. Two members of the Information Management Research Group and Prof. Dr. Schwabe participated. The workshop lasted two hours and was divided into two parts. First, the chosen definitions and results of the qualitative analysis were discussed to validate if they are legitimate and comprehensible. In the second part, the first version of the new framework was presented. A use case was chosen to provide a concrete example to discuss the framework. The article of Preetipadma (2020) was employed for the use case. It describes an application where neuromorphic computing is used to detect substances in the air. In advance of the workshop, the article was sent to the participants. They were asked to read it as preparation. The start question raised was whether the participants think that the use case is a Third Wave AI. One hour was reserved just for the discussion about the use case. The discussion was recorded and analysed afterwards.

For the last data collection, interviews were conducted. The goal was to conduct ten interviews. Therefore, twenty potential interview partners were inquired. The final ten interview partners are listed in Appendix B, and they are categorised into four groups. The first group *Third Wave AI expert* consists of experts who are part of the public discourse about Third Wave AI. They either published about Third Wave AI or claim to provide Third Wave AI application. The second group is called *industry experts*. They do not have a relation to term third wave AI but can provide information about the requirements AI have to meet in the future. Therefore, these industry experts are working in an innovative environment and having many years of experience. Based on the results of this group, it will be possible to determine if the framework represents the needs of the industry or if there are adjustments necessary. The third group consists of *AI experts* who use AI in their everyday life. With their experience, the framework can be challenged, and insights can be collected about the practical application of AI. Finally, the last group are the *neuromorphic experts*. Neuromorphic computing is a term which frequently appeared in connection with Third Wave AI during the analysis of the public discourse. Therefore, these experts were interviewed to reveal what role neuromorphic computing will play in the future of AI. The reason for these four groups was to collect a wide range of input and to reduce the risk of only analysing third wave in its bubble. Four industry experts, two Third Wave AI experts, two AI experts and two neuromorphic experts were interviewed.

The potential interview partners were identified based on their publications and their experience. To ensure a wide range of insights, not only Swiss interview partners were inquired but also international ones. Therefore, based on the interview partner's preference, the interviews were conducted in German or in English. Four interviews were held in English and six in German. Additionally, six of the questioned interview partners are originated from Switzerland, one from Poland, one from England, one from America and one from Australia.

The interviews were set up semi-structured, and a template was designed with questions as guidance. For every interview, the template was calibrated to the interview partner and the new insights. An example of a used interview guide can be found in Appendix H. For the questions, multiple interview techniques were used such as the Five Whys based on Serrat (2017), what if-questions and critical incidents based on Flanagan (1954). All interviews were recorded for later analysis and to allow to entirely focus on the interviews without being forced to take notes in parallel. For each interview, a duration of one hour was scheduled. Per interview, only one interview partner was questioned with one exception where two interview partners were questioned at the same time. For that interview, the duration was increased to one and a half hour which allowed that each interview partner was given enough time to answer the questions.

3.2 Data Analysis

Multiple tools were used for the data analysis due to the different approaches and types of data. The CSV-files from Scopus were imported in Excel and used to build multiple diagrams to get a better understanding of the context of Third Wave AI. Similar diagrams were built like Scopus provides with its *analyse search results* function, but for example, multiple diagrams of different searches were merged in one to allow a better comparison.

Not only for the data collection, the same approach as Zavolokina et al. (2016) was used but also for the qualitative analysis of the public discourse. They used tools of the content analysis like identification of central entities and dominating topics. These tools were also implemented in this Master's thesis. A comprehensive content analysis is performed to get insights into the ongoing public discourse (Cukier et al., 2004). To perform the qualitative analysis, the qualitative data analysis (QDA) software ATLAS.ti was used to encode all the collected documents. In total, 32 documents were encoded with 57 codes. The codes were chosen based on relevant terms or evaluative codes like criticism or challenge. Based on the related work, the relevant terms were identified like neuromorphic computing or HITL. A total of over 700 citations were encoded. In Figure 3.1 an example of a encoded paragraph is exhibited. These citations were then used to build networks with ATLAS.ti. The networks provide insights about documents, which are related and can be enriched with semantics. The resulting networks are discussed in subsection 4.1.1.

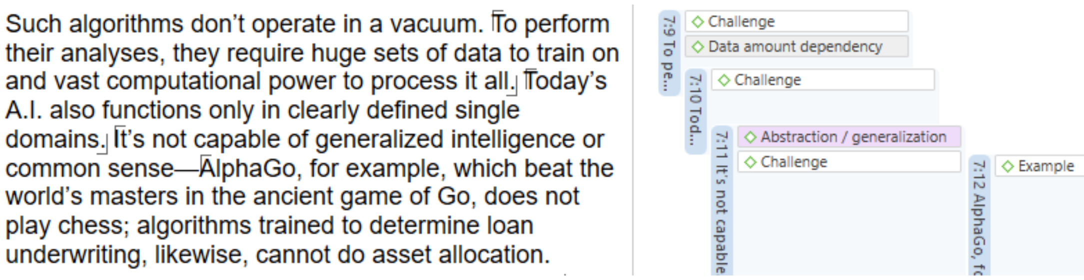


Figure 3.1: Example of how the literature was encoded in ATLAS.ti on a paragraph of Lee (2018)

For the analysis of the interviews, they were automatically transcribed by a service provider. Using an automated transcribe service was necessary to reduce the time effort. The interviews in German were transcripted with f4x and the ones in English with

Happy Scribe. To reduce the costs, the interviews were edited, and only the questions and answers were transcribed. Nevertheless, over ten hours in total were automatically transcribed. The transcripts were then reviewed, and the relevant parts were corrected. All the relevant parts were collected, and the final version of the framework was built. The results are presented in Section 4.2.

4

Results

As already mentioned in Chapter 3, there are many diverse results due to the contrasting methodologies. This chapter is structured accordingly to the subphases and is therefore chronological. It starts with Section 4.1 about the analysis of the public discourse. The section is split into two subsections *Quantitative Analysis* and *Qualitative Analysis*.

4.1 Public Discourse

4.1.1 Quantitative Analysis

Figure 4.1 displays the number of scientific publications per year about Third Wave AI. The used query is the same, which was already shown in Section 3.1. Therefore, the documents shown in the chart are the same as the one analysed in the qualitative analysis. The start point of Figure 4.1 is set to 1990 because the publications before 1990 are of limited relevance. 43 documents are found by the query before 1990, and the oldest one is from 1914. Documents as Curtis (1914) do not refer to AI but were included because the American Institute of Electrical Engineers published it. They are shortened to A. I. E. E. and are included in the abstract.

In total, 246 documents were published, which fulfil the query. Figure 4.1 illustrates a substantial increase around 2017, when DARPA published its framework. Interestingly, the highest peak was neither during 2017 nor one year later. 2019 the most documents were published, with a total of 30. The data was collected on the 30th of June 2020. Therefore the number of publications in 2020 has not reached its final quantity. However, the number increased from 7 to 19 on the 11th of October 2020. It did not reach to the publication number in 2019, but this may change by the end of the year. The chart illustrates, DARPA's framework shaped the term Third Wave AI. There is a lack of evidence-based on the quantitative analysis if DARPA is the first one using the

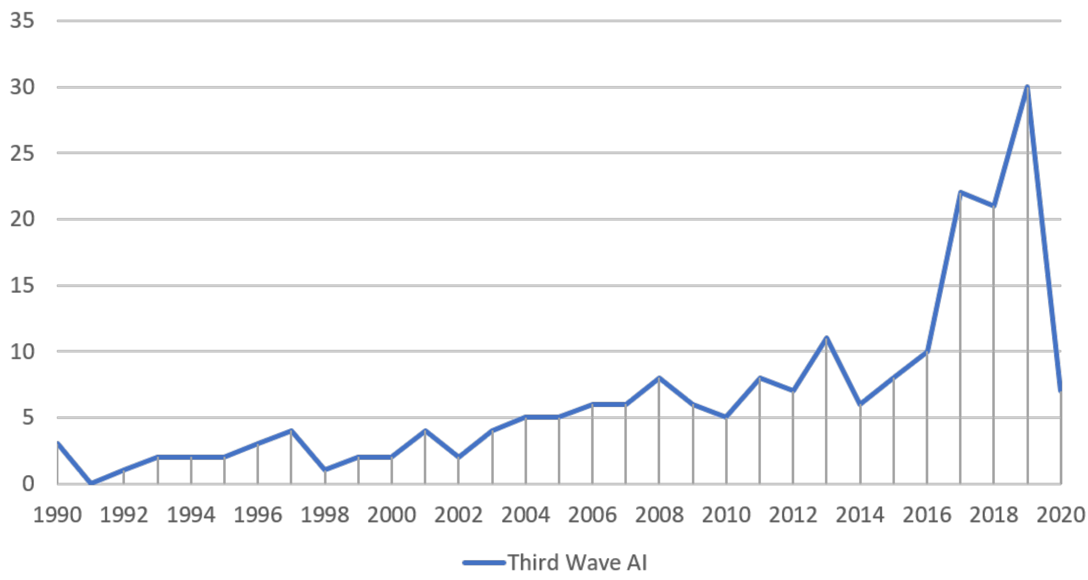


Figure 4.1: Publications per year about Third Wave AI based on the data of "Scopus" (2020)

third wave in the context of AI and linking it to contextual adaptation. The continuous increase before 2017 can be explained by the growing utilisation of AI in research fields involving waves. For example, Cecotti and Graser (2011), Diskin et al. (2002) and Ambjorn (2008) are all publications about the application of AI in such research fields. All these publications belonged to the query's result, which could not be excluded with keywords.

To compare the discussed results, the chart shown in Figure 4.2 was built. It displays the publications per year and keywords. Again the chart only shows results after 1990. The following query was used:

TITLE-ABS-KEY((X) AND (ai OR "artificial intelligence" OR a.i. OR "A. I."))
AND (LIMIT-TO(LANGUAGE,"English"))

The *(X)* was replaced for each line by the corresponding keywords shown in the legend. All keywords are related to Third Wave AI and were chosen based on the related literature. Four categories were analysed with six different queries. The first two belong to the term contextual adaptation and related terms, which DARPA uses for describing Third Wave AI. Explainability and trust are also terms used by DARPA and enjoys great popularity in research. Therefore, they are the perfect candidates to compare the

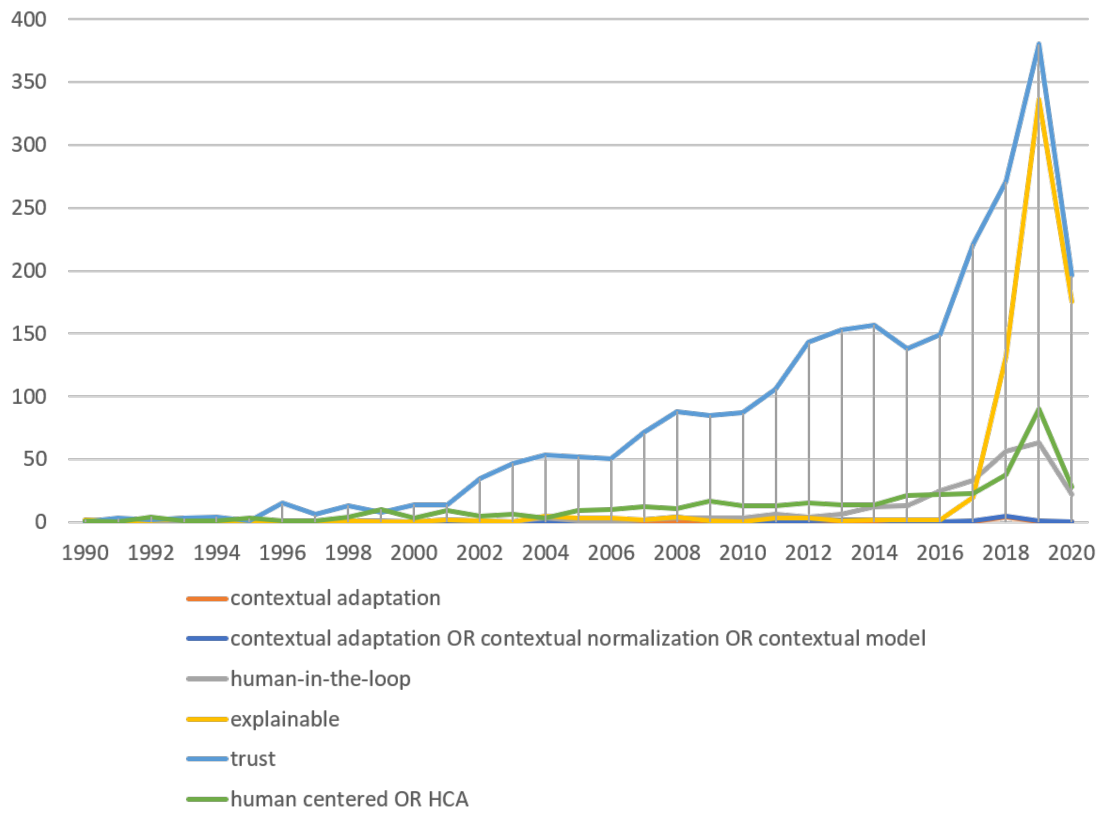


Figure 4.2: Publications per year and keywords based on the data of "Scopus" (2020)

amounts of publications and to be able to assess the relevance of the other terms. HITL and HAI were added because in the initial literature review, they turned out to be terms with high relevance in the context of how AI should develop.

As expected, explainable and trust are the ones with the highest number of publications. HITL and HAI are relatively close to each other but have a significantly lower quantity of documents than explainable. Contextual adaptation has by far the lowest numbers. All terms have their spike in 2019 besides contextual adaptation, and it is likely that in 2020 the numbers will reach the same or even higher quantities. Contextual adaptation reaches its peak one year before the others. Overall, the academic focus is clearly on XAI and partially on human-related AI topics. DARPA's terms and Third Wave AI in general, got low numbers from which can be assumed, something is holding back people contributing to these topics.

In Figure 4.3, the results of contextual adaptation is broken down into groups based on the subject of the publications. The most dominant part is computer science, followed by mathematics and engineering. To conclude, the focus remains mainly on technical aspects of contextual adaptation. The other subjects are represented each by only one publication. Interestingly, there are no publications found about social sciences, and many other categories are missing like medicine. Therefore, a clear need for non-technical discussions can be identified.

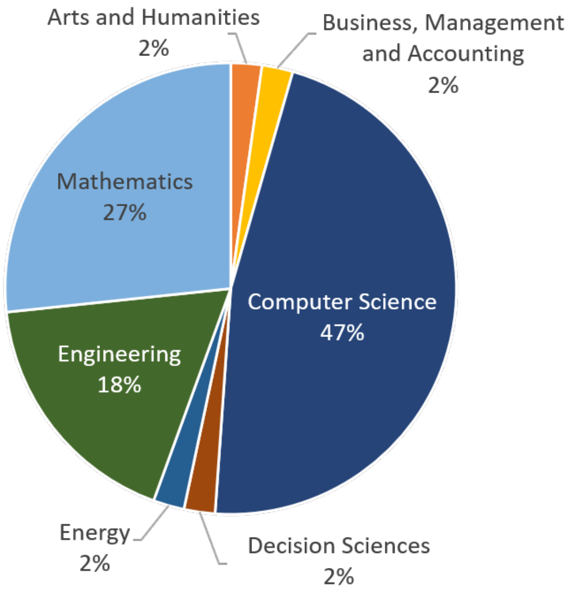


Figure 4.3: Documents about contextual adaptation by subject based on the data of "Scopus" (2020)

4.1.2 Qualitative Analysis

In contrast to the quantitative analysis, Third Wave AI is a term which becomes more and more relevant. Thus, there is a need for academic discourse. The call for papers of Buxmann et al. (2019) addresses this deficiency and many other topics like human-machine interaction. It represents the need for scientific publications in the non-technical field. Buxmann et al. (2019) use DARPA's framework to give an idea of how the future of AI could look. With DARPA's framework, they make contextual adaption and explainability a subject of discussion. However, the term contextual adaptation was never used in the publication since they call it reasoning capabilities. They also set the main focus on explainability because they state that the Third Wave AI with human-like communication and contextual adaptation is not achieved. When comparing the use of the term Third Wave AI in this publication and the definition of DARPA, they are divergent. The reason could be the lack of other publications than DARPA's video or the comprehensibility of DARPA's definition. The call for papers is for the first issue in 2021, and the issue will be very suitable for further analysis.

DARPA's framework is not the only one, which utilizes waves to describe the evolution of AI. As shown in Table C.1, there are multiple frameworks which use very similar approaches and even three waves. Szu et al. (2019) and Xu (2019) define almost identical first and second wave like DARPA. However, the third wave differs in each framework. Szu et al. (2019) defines the third wave by the achieving of human fuzzy linguistic thinking and Xu (2019) by breakthroughs in existing technologies and the inclusion of a human-centered approach. In contrast to DARPA, both frameworks do not predict a revolution on the technical layer. The question arises what kind of revolution to initialise Third Wave AI is needed to reach the same level as the first and second wave. The analysed frameworks reflect very contrasting opinions. Bai et al. (2019) define three stages by the AI's capabilities. Their three stages are weak, strong and super AI. They claim we are currently in the first stage of AI. All the authors introduce their own framework without building upon other frameworks or definitions. However, it is obvious that especially the definitions of the first and second wave are too close to another. It can be presumed that the frameworks used similar foundations. However, it is impossible to identify neither the frameworks shown in Table C.1 have been inspired by DARPA nor DARPA has used the same unknown source for inspiration.

There are frameworks, which also work with waves, but in a different way than DARPA. These frameworks are listed in Table C.3. They either use an additional wave or completely different approaches. The most relevant framework with four waves is Scott (2018), which was introduced and discussed in Section 2.1. Lee (2018) uses also four waves to describe the evolution of AI. It is the only framework which defines the third wave similar to DARPA, but all other waves differ significantly. *Perception A.I.* is the term Lee (2018) uses and describes an AI which has advanced sensing and uses this data for creating new applications. In essence, it is the same as contextual adaptation. The first is called *Internet A.I.* and the second *Business A.I.*. The final form based on Lee (2018) is *Autonomous A.I.* which combines all waves by mastering the perception, adaptation and many more characteristics.

Table C.2 also contains frameworks, which use waves but reflect on a more bird's eye view. Carroll (2020) and Paul R. Daugherty and Wilson (2018) define AI as the third wave. Toffler (1980) goes even a step further and defines the whole information age as the third wave. However, all the presented frameworks have in commune identifying AI as a revolution and agreeing on the relevance of AI. The wave representation of the AI evolution has proved to be suitable based on many different frameworks using it. Most frameworks agree on the first and second wave as DARPA describes them. However, there is no perfect framework, and most of them are only guesses on how the future of AI could look. There is still a need for an overall valid framework which does not determine technology and instead focus on the requirements of the Third Wave AI.

Another evidence for the relevance of the term Third Wave AI was found while analysing Twitter. On 23th of June 2020, Adler (2020) announced Toyota AI Ventures' investment in a company called Third Wave. The company works on what they call shared autonomy ("Third Wave", n.d.). They work on a new generation of forklifts with shared autonomy which is "a breakthrough in machine learning, computer vision and robotic material handling" (Voss, 2017). Also based on their website, the forklifts can request help which is a HITL scenario. With the information available, it is not possible to determine if Third Wave's solution is a Third Wave AI when using DARPA's framework. They do not use the same terms than DARPA like contextual adaptation and make no association with the framework. The question arises, why they have chosen the name Third Wave. Nevertheless, using Third Wave as a name for an AI startup increases the opacity of the term Third Wave AI. This problem will increase when the company Third Wave grows and gains in importance.

Besides the tweets about the company Third Wave, the remaining ones consist mainly of comments and reactions to DARPA's framework like the responses to DARPA (2019). In general, the tweets about Third Wave AI are quite neutral, and most ones do not judge over DARPA's view, like the example displayed in Figure 4.4. The analysis of GitHub did not show any relevant results. Therefore, it can be assumed, the term Third Wave AI is not used in the developer community. Additionally, there is likely a project, which has a third wave nature and could be labelled as such based on DARPA's framework. Since the term Third Wave AI is not established, nobody uses the term as a tag in their project. The reasons why it is not established could not be identified only with the analysis of the project on GitHub. However, it can be assumed that the reasons are the same as already shown. The lack of definition and application make DARPA's framework unusable for software engineers.

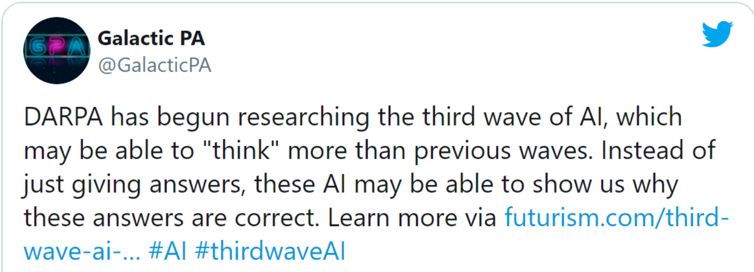


Figure 4.4: Tweet about Third Wave AI from Galactic PA (2018)

To gain more structured insights, especially in the academic field, literature was encoded. With the produced data, different networks were built to visualise the insights. For the identification of central entities and dominating topics as Zavolokina et al. (2016), the literature was encoded with the code challenge to identify the different types of challenges regarding AI. The resulting network is shown in Appendix D. All citations encoded with *challenge* are displayed. The citations, which belong to the same topic, were combined to a new code. These codes are displayed in grey and include topics like privacy, trust, emotion or handling uncertainty. Additionally, the relations to the code second wave are displayed. With these relations, it is possible to identify which are general challenges of AI and which are second wave challenges based on the literature. The challenges presented in DARPA's presentation are overlapping with the ones mentioned in other publication. There are only three additional challenges added by other publications. To identify if the relevant keywords are part of the challenges, they were added to the network and are coloured.

With the network, the dominating challenges are shown. Only four dominating challenges out of ten are addressed with DARPA's framework. The three challenges are data amount dependency, data quality dependency, individually unreliable and handling of uncertainty. DARPA's Third Wave AI will solve some of the dominating challenges indirectly, like the black-box challenge. With the explanation produced by the contextual exploratory models, the AI can be made more transparent, which reduce the problem of AIs being a black-box. Other challenges are not affected at all, as emotions or privacy. An ideal framework would address more of the dominating challenges because it would lead to an increase in the relevance of the framework. Therefore, there is a need for adjustments based on the challenges.

Besides the challenges, the documents were encoded if they contain criticism or suggestions for improvements. The generated network in Figure 4.5 shows only two documents which fulfil these codes. Perez (2017) is the only publication which criticises DARPA's framework. He states that the framework is too elementary because the second wave represents too many different approaches. By comparing DARPA's framework to his own, he demonstrates the simplification by mapping Third Wave AI to only one of his five steps. These steps are part of his roadmap on how deep learning will evolve. His framework is summarised in Appendix A. Additionally to Perez (2017), Scott (2018) is shown in Figure 4.5 as the only publication which builds upon DARPA's framework. The resulting framework was already introduced in Section 2.1. The network displays one of the biggest problems of DARPA's framework, the lack of discourse. Many publications use the framework without questioning it, which appears as very problematic.

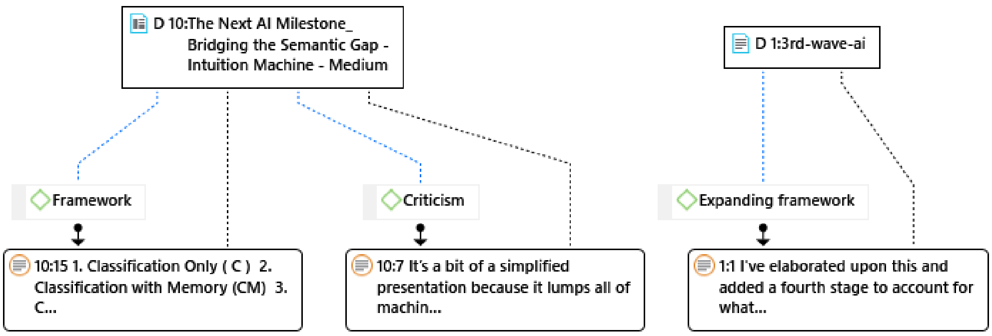


Figure 4.5: Criticism and expanded frameworks based on the encoding

The Figure 4.6 illustrates another problem of the framework. For this network, the documents were encoded based on if they do not cite DARPA's definition or provide any definition at all. All the documents at least address the Third Wave AI or use it. The documents, which are encoded with *no citation DARPA*, use a framework or a definition, which is very similar to DARPA's but without citing DARPA or another source. Based on the network, eight documents fulfil these criteria and were published after DARPA. Some publications build upon DARPA's framework without properly declaring it. The reason for this behaviour could result from the missing publication of DARPA besides the video and the presentation slides. Additionally, three documents use Third Wave AI without any definition. There are multiple reasons for using Third Wave AI as a term without introducing a definition. However, the most likely reason is that in the author's view, the term is well established, and therefore for him, no definition is needed. This discrepancy indicates the Third Wave AI being on the way of getting established. Nevertheless, additional academic discourse is needed to achieve the acknowledgement of Third Wave AI as a scientific term.

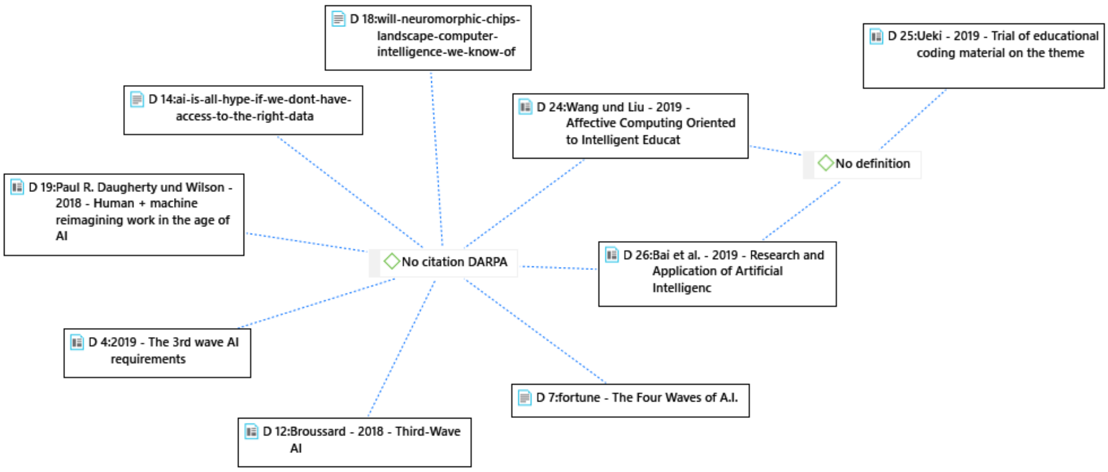


Figure 4.6: Problematic use of the term Third Wave AI based on the encoding

Another already mentioned problem is the mix-up of contextual adaptation and XAI. To analyse this problem, the network in Appendix E was built. It displays the affected terms and the documents which contain citations related to the terms. Besides contextual adaptation and AI, the related terms contextual reasoning and contextual explanatory models were added. The citations were not only encoded when it mentioned the corresponding term but as well if similar terms were used or the term was described and fitted the established definitions. The citations are arranged as clusters to help to visualise the different categories of relations. The biggest cluster is the publications which only address contextual adaptation. Based on DARPA, contextual adaptation is achieved with using contextual explanatory models which also enables XAI. That is why DARPA uses the term contextual adaptation as the descriptive term for Third Wave AI. Therefore, when a publication is talking about XAI in the context of Third Wave AI, contextual adaptation and contextual explanatory models should be addressed too. However, there is only one document besides DARPA's video, which addresses all three terms. There are just six documents, which cover contextual explanatory models at all. Three documents mention contextual adaptation and XAI without contextual explanatory models.

Looking more closely at the citations in the documents, the problem becomes even more apparent. The following example is one which was identified with the encoding.

"Third Wave AI systems will feature dramatic improvements, most notably in their ability for contextual adaptation. They will understand context and meaning, and be able to adapt accordingly. Third Wave AI will not only recognize the cat, but will be able to explain why it's a cat and how it arrived at that conclusion – a giant leap from today's "black box" systems." (Scott, 2018)

Scott (2018) describes in the first part contextual adaptation similar to the used definition in this thesis. In the second part, he focuses directly on explainability without introducing contextual explanatory models. Skipping this clarification step leads to the impression that XAI is a result of contextual adaptation which is not correct. With contextual explanatory models, contextual adaptation and XAI can be achieved, but none of them is a result of the other. The omission of contextual explanatory models is therefore problematic and occurs quite often in the literature as it is shown in Appendix E.

In summary, the qualitative analysis shows that DARPA's framework is used divergent, and there are many varied interpretations. This circumstance is likely due to the lack of definitions and additional explanations provided by DARPA. The used terms are not very precise, especially contextual adaptation. This imprecision increases the risk of misinterpretations. Additionally, the missing sources make the framework even more incomprehensible. The intermixing is intensified by this incomprehensibility and other very similar frameworks.

As already shown with the challenges, DARPA's framework addresses only a small number of challenges. Xu (2019) is convinced that currently, the biggest issue with AI research is the focus only on technical aspects. Therefore, he introduces the definition of HAI, which was presented in Section 2.3. DARPA's framework focuses only on the technical aspects of AI, which is also the reason why only a small amount of the identified challenges are addressed. For example, privacy is a mainly non-technical challenge which needs to be solved with transparency and privacy-friendly engineering. The analysis highlights the needs for frameworks with more than just a technical classification.

In general, the analysed publications express a need to shift the focus from only technically looking on AI. Like the following example, many publications suggest areas which should be considered furthermore.

"We also know that a significant challenge for future AI is contextual adaptation, i.e., systems that incrementally help to construct explanatory models for solving real-world problems. Here it would be beneficial not to exclude human expertise, but to augment human intelligence with artificial intelligence." (Goebel et al., 2018, p. 3)

Goebel et al. (2018) suggest adding human expertise to Third Wave AI. Human expertise can be a part of HITL or HAI. The suggestion of Goebel et al. (2018) is based on the basic assumption that human intelligence will not be replaced by AI, but rather enhanced. In the analysed documents are many similar suggestions on how Third Wave AI should look. To get an overview of all suggestion, the network shown in Appendix F was elaborated. It displays the suggested characteristics for Third Wave AI. To identify the central entities, citations which suggest very similar characteristics were grouped under the corresponding term. These dominating terms are shown in different colours. Additionally, terms, which come from the same areas are associated with each other by having an identical colour. The other citations include topics like usable AI or human-like thinking. Some of these topics could also be added to the dominating topics, but are not assignable. For example, usable AI can be a a requirement of HITL or HAI or a result of contextual adaptation or XAI. Therefore, it was decided to leave the citation as an entity.

The network shows clearly how narrow DARPA's framework is. It only addresses six out of thirteen characteristics. The six characteristics are XAI, contextual explanatory models, contextual reasoning, contextual adaptation, complex adaptive systems and contextual models. If the topics with identical colours are counted as one, DARPA only deals with three out of nine groups. Therefore, an adjustment of DARPA's framework is needed to address at least more or ideally all of the dominating topics. By targeting the six technical characteristics, DARPA provides a perfect foundation for such adjustments.

4.1.3 Sensemaking Workshop

To validate the qualitative analysis' results, a sensemaking workshop was carried out. In order to provide a basis for discussion, a proposal for modifications was developed. As a starting point, the network with the characteristics of Appendix F was used. In the first step, it was reduced to the dominating topics, which is shown in Figure 4.7.

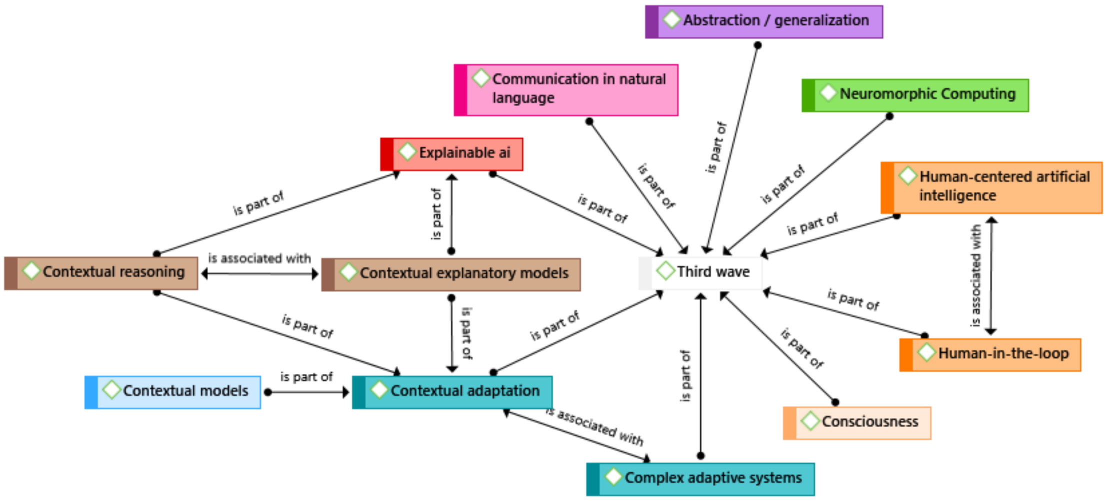


Figure 4.7: Reduced characteristics of Third Wave AI

Also, links were added to the network to visualise the relations between the terms. With the links and the colours, it was possible to reduce the network even further. The resulting terms were contextual adaptation, XAI, HITL, communication in natural language, abstraction/generalisation, consciousness and neuromorphic computing. It became clear that the suggested adjustment would not only be an extension of DARPA's framework. An additional framework which builds upon DARPA's and fits the purpose better. However, to achieve a framework, it was necessary to reduce the terms even more and reduce it to its essence. Based on the identified needs in the qualitative analysis, it was decided to focus on functional aspects of Third Wave AI. As a result, the terms neuromorphic computing and contextual explanatory models were removed. HAI was chosen over HITL since HAI is the broader term and includes HITL based on the used definitions. Communication in natural language is a requirement to master HAI and HITL since it constitutes the foundation for a flawless interaction between human and AI. The reason for removing abstraction and generalisation was they are more requirements on a technical layer than a functional one. The aspects which make them functional requirements, for example, the ability to transfer learning, are addressed with contextual adaptation. Consciousness was also removed because the term is on another level than remaining ones. Achieving consciousness in AI could be a result of mastering contextual adaptation and will be a breakthrough on its own. How contextual adaptation could lead to consciousness is discussed in Chapter 5.

The remaining terms are contextual adaptation, XAI and HITL. The framework should be as simplistic as possible yet comprehensive. Contextual adaptation and XAI are already part of DARPA's framework but not as functional requirements. Instead, they are more results of the contextual explanatory models. To shift the focus even more, HITL was added to complete the functional requirements. However, each requirement on its own will not lead to a Third Wave AI. It will be the combination of all three, which improves AI significantly and potentially leads to revolutionise AI. The resulting first version of the framework is shown in Figure 4.8. The arrows have been added to visualise the synergy.

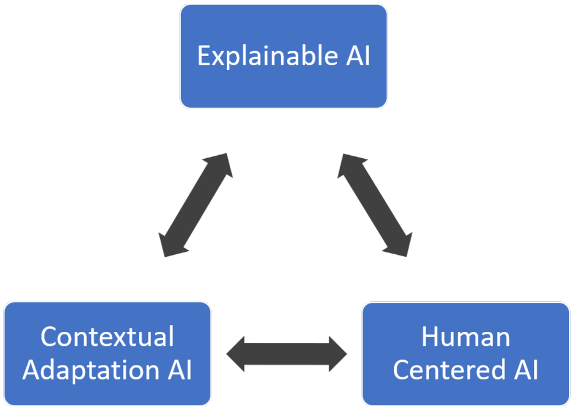


Figure 4.8: First version of the framework

The participants validated the used definitions and agreed on DARPA's first and second wave. They also confirmed the problematic use of contextual adaptation and XAI and confirmed the necessity for a new framework. All agreed on the need to shift focus from the only technical perspective. All participants declared the use case as not fulfilling Third Wave AI. However, the suggested framework was strongly criticised, and many inputs were given. One participant questioned the framework's primary purpose. It was unclear for him how the framework could be applied. Another identified problem was that the definitions were still lacking. For the participants, it was not clear what the terms represented. This ambiguity led to discussions on terminology, which were not leading to a meaningful outcome. The used terms were generally questioned but on different levels. However, all participants agreed that terms should be used without AI at the end.

For all workshop participants, the paradigm shift was undefined, which is needed to establish a new wave. The reason for that could be the too simplistic approach, and therefore the lack of comprehensiveness. Defining Third Wave AI only from a functional point of view, is too narrow. The first two waves are based on technical paradigm shifts. Therefore, it is not expedient to remove the technical core form the framework. It was suggested to work with different layers to provide more comprehensiveness without cluttering the framework. For example, a participant suggested a social layer. The visual representation was also criticised, and suggestions were provided on how to improve it. Especially, to visualise the different layers, multiple suggestions were made, like using a similar approach as the St. Galler Management-Modell of Rüegg-Stürm (2003). Furthermore, literature was suggested to compare the framework with other approaches, which are not AI or computer science approaches. One of these suggestions was Pearl and Mackenzie (2018) and their Ladder of Causation, which was explained in Section 2.4.

All suggestions were collected and based on the gained insights, a new version of the framework was built up. The version is illustrated in Figure 4.9. Not only the framework's visualisation was changed, but also the terminology was revised. The most significant change provides the introduction of layers. A functional and technical layer was visualised with different colours. In the centre is the technical layer with DARPA's contextual explanatory models shaped like a circle. The functional layer surrounds the technical layer. Different shapes were tried, some drafts are displayed in Appendix G. However, the cube turned out to be the best option. It consists of three dimensions, which are HITL, XAI and contextual adaptation. Based on the results of the interview, it was decided to replace HAI with HITL. In this stage, the definition of Xu (2019) was used for HAI. Hence, the term did not fit into the framework, and HITL was the best alternative. Using the cube helps to understand that all dimensions are needed to shape the object, which belongs to Third Wave AI. With this version, it was possible to achieve comprehensiveness without reducing the comprehensibility.

4.2 Interviews

Each interview was very divergent, and therefore the results are meaningfully heterogeneous. Besides challenging the suggested framework, the goal was also the validation of the so far collected results. For this purpose, DARPA's framework was presented with the Table 2.1, and the interview partners were asked to describe their first impression. Most of the interviewee agreed with DARPA's first and second wave. INT10 stated the

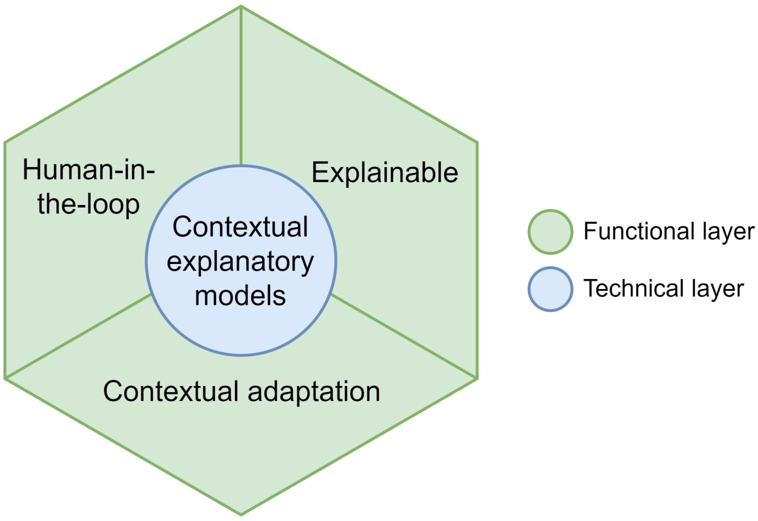


Figure 4.9: Interview version of the framework

second wave being not as advanced as it is illustrated in Table 2.1. For him, statistical algorithms can only be used for simple tasks at the moment, and they are not able to abstract and generalise. Another criticism of DARPA's framework was the following one:

"One thing I would say as a critique of this table is, it seems to think the second wave has good learning capabilities. I would say 75 is overrating that. I would say learning is at 25 in the second wave. Because, as you say, it depends critically on big data. And furthermore, it's pretty much incapable of continuing to learn." (INT7)

Compared to the criticism about the first and second wave, Third Wave AI was challenged significant more often. INT1 suspected a marketing function of the term. Other interviewees made similar assumptions. The lack of a concrete definition of Third Wave AI could justify such guesses. Many interview partners had problems understanding DARPA's definition, especially the industry experts. "This is the problem of definition, that the third wave is not so well-defined that it immediately pops into the head of our interlocutor" (INT5). The example shows why the lack of definition leads to the framework's non-utilisation in the industry. For most interviewees, Third Wave AI was not tangible, and therefore, the willingness to use the term is negligible. also explained why, in his opinion, Third Wave AI is not extensively discussed in the academic world. He stated that the framework has to get much deeper to make it suitable for researchers and academia.

Despite the criticisms about DARPA's framework, the feedback on the suggested framework was overall positive. INT3 appreciated the comprehensibility and the simplicity of the framework. The majority of interviewees shared this opinion. INT8 challenged the framework with the following use case:

"What's about Alexa? It adapts to you. It knows how you speak. It wants you to speak for five minutes and tries to recognize you. And it gives you some explanations about why she thinks that's a stupid question or why she doesn't know it. It definitely has human-in-the-loop." (INT8)

Interestingly, the explanation of why Alexa is not a Third Wave AI was as provided in another interview: "The input is simply categorized. And, you know, there's no disambiguation, there's no reasoning. It's a purely statistical thing" (INT4). The statement remains particularly accurate for Third Wave AI when looking at the Table 2.1 where reasoning is rated with 75. Alexa is hardly contextual adaptive based on the definition used in this thesis. It minimally perceives its environment and context. For example, when someone asks Alexa two questions, its answers will not be related to each other, even if it is a follow-up question. One option to make Alexa more contextual adaptive is using the first question to answer the second.

The functional layer was well received, and INT8 recognised even the cube. INT3 appreciates the focus on the functional layer because according to her experience, the technical layer was rarely the biggest problem, but the functional layer was. One of the examples INT3 mentioned was data protection issues which are addressed with HITL in the framework. To achieve HAI and furthermore HITL data protection is essential. Since a human will not work and trust a system which does not protect his data.

Besides the appreciation, the interview partners validated the individual dimensions of the functional layer as the right ones. INT7 confirmed the importance of HAI because if the user's only task consists of keeping an eye on the AI, he will rapidly trust the system, and therefore lose attention. The majority approved contextual adaptation's importance for Third Wave AI. Though many different terms were used to describe contextual adaptation. "I think what you call contextual adaptation is probably the same as what I am calling online learning. So, the ability to absorb new information and adjust performance in the light of the current environment rather than historical training" (INT7). Besides online learning, INT9 and INT10 mentioned the term transfer learning in relation to contextual adaptation. Transfer learning and generalization will be an important enabler for contextual adaptation. The ability to use accumulated knowledge to solve a new or different problem will constitute a huge improvement on AI. INT10 challenged the necessity of XAI as a dimension. He argued that XAI is only necessary in cases where AI fails, and an explanation is needed for the traceability. Opinions are divided on this issue. Since XAI can also be used to establish trust or as a method of persuasion. Nevertheless, Third Wave AI will not reach such a level of completeness. The term HITL remains the most controversial one. Most interviewees initially had problems understanding the term. "I was wondering what "human-in-the-loop" is doing there, but when you explained it to me, I think we see it the same way, as a knowledge provider" (INT5). In addition to the ambiguity of the term, INT4 brought up the apprehension that the HITL's application will slow down AI systems. The feedback suggested reconsidering the use of HITL. Responses like the following example confirmed the suspicion: "And possibly human-centric AI captures that better than human-in-the-loop" (INT7).

Not only the individual dimensions of the functional layer were appreciated, but also the combinations met with approval. However, multiple concerns were mentioned by the interviewees. INT1 and INT4 draw attention to the area of tension between contextual adaptation and XAI. Both stated that increasing contextual adaptation would lead to lower explainability. To achieve better contextual adaption, on the one hand, far more input would be needed, and on the other hand, the system should also learn from its output. Either way, they both significantly increase the system's complexity and therefore make them hard to explain. Nevertheless, INT1 describes the collaboration of all three dimensions perfectly. The more contextual adaptive a system is, the higher is the difficulty of making it explainable and the more important the HITL gets (INT1).

The opinions appeared to be most controversial on the technical level. Especially, each Third Wave AI expert had a specific vision on how the technical layer looks. INT4 uses cognitive architecture to achieve Third Wave AI. He describes the cognitive architecture as follows: "The focus is on adaptive system systems that can basically learn immediately, that can reason, and that can adapt to circumstances in the real world. And I identify that as a cognitive architecture" (INT4). Based on Table 2.1 and the description above, cognitive architecture fulfils the requirements perceiving, learning, abstracting, and reasoning introduced by DARPA. However, for INT5, another approach will lead to Third Wave AI system. "The third wave means we should have good learning capabilities and good reasoning capabilities. This is exactly what we do in Samurai. We combine the first wave and the second wave to maximize those capabilities" (INT5). The combination of the first and second wave is a widespread approach. INT1 mentioned it as a guess about what Third Wave AI could be about. The argument for this approach is using the capabilities of each wave to accomplish a huge improvement. The Table 2.1 allows demonstrating the increase in values by the combination (INT5). When the highest value of each wave and category is picked, the results match up with DARPA's values for Third Wave AI. The only exception is abstracting, wherewith this method only 25 is reached, but 50 is requested. However, both technologies have the potential to provide Third Wave AI on the technical layer, but stay far away from being a Third Wave AI on the functional layer. Therefore, there is no technology, which turned out to be a clear favourite and could be added to the framework. INT9 suggested avoiding to commit to a specific technology. In one of the last interview, the technical layer was challenged by a completely different approach. "I would expect to focus more on parameters such as learning, abstracting and reasoning, because I think those are what define the waves" (INT5).

For HITL, two additional subcategories were suggested. The following statement raised the first subcategory: "This is how we understand the third wave; it is not only learning from labelled data but also from expertise, for example, from cyberbullying experts or linguists" (INT5). HITL uses a human to solve edge cases, but it is undefined if the human in the loop is the user or a completely independent person. If the user is asked to help to solve the edge case, likely, the user does not have the required knowledge either. There is a risk that the user consciously or unconsciously solves the edge case incorrectly. Hence, it would be better to use experts for solving edge cases. Experts-in-the-loop would drastically improve the quality of an AI. INT2 explained a similar approach where new questions to a voice assistant are passed on to an employee who enters the missing knowledge. To achieve expert-in-the-loop, systems must be built accordingly. "And for me, the most important task of the third wave is to invite those experts to the process of development, but not merely as advisors in the process of labelling data, but to really provide knowledge to this system" (INT5). To include experts as a part of the system, corresponding tools need to be provided. Especial when keeping in mind that the majority of experts are not computer scientists, and therefore need a way to enter and maintain the knowledge with basic informatics knowledge. The second subcategory is the user-in-the-loop suggested by INT9. In combination with expert-in-the-loop, it would help to better differentiate between the various applications of HITL.

The two terms determine who is tried to be integrated into the AI system.

Furthermore, an additional layer was suggested multiple times. INT8 addressed the need for a hardware layer:

> *"So, another aspect I would add is cloud computing and edge processing. So, I think these are valuable because if you want humans-in-the-loop, you can't run everything. You can't run the algorithm on, say, the mobile phone, you just connect to the internet, and some server runs the model for you."* (INT8)

An important point remains to not only focus on technical and functional aspects, but also on hardware, which is essential for the other layers. Advancements on the hardware layer will positively affect all layers and dimensions. For example, accelerated hardware improves HITL since, among others, it reduces loading times, which increases user satisfaction, and therefore raises the willingness to use the system. Accordingly, it is important to specify the requirements for the hardware layer. Especially, keeping in mind that AI may remain limited. "It will always have to make decisions with limited information and limited resources" (INT4). Therefore the requirements for hardware are important. INT8 highlighted the parallels to biology: "Biology does it better. It does it at low power. It does not cost that much time, the low latency aspect of biology" (INT8).

A buzzword used by one interview partner was Internet of Things (IoT). "Now different answers are connected. Different systems are connected, so maybe models can interact with models. It's not just human-in-the-loop that teaches. It's maybe some other model that performs better, which can teach this model" (INT8). IoT, as explained in the statement, has a big potential. A term, which describes the approach better is Internet of AI. By connecting AIs to each other, a whole new level of AI ecosystem could be achieved. Another buzzword is neuromorphic computing. Both neuromorphic experts agreed on neuromorphic computing as not representing a technical dimension because it does not belong to the technological approaches (INT7; INT8). Moreover, it is a hardware approach, which will allow new technological approaches. "The more we learn how to raise the level of abstraction in building neural systems, the more we may be able to map those problems onto more conventional computing platforms as well" (INT7).

Multiple dimension were suggested during the interviews. INT1 proposed a social, legal and ethical layer. On the opposite, for INT2, the social or ethical dimension are included in the HITL. Because if the human is as an integral component of AI, these dimensions are already taken into account since they are results of humans. If the inclusion in HITL is possible depends on the used definition. However, INT1 and INT7 agreed that these kinds of dimensions are not located in the functional layer, but more on a layer above. INT9 prioritised ethics as a needed dimension since ethical changes play an extremely prominent role, such as gender-sensitive language. Another dimension which was suggested by INT3 is society-in-the-loop. Since HITL and HAI focus more on the individual human should society-in-the-loop take into account different types of societies. Societies significantly influence our decisions and actions. Therefore it is essential to integrate societies into AI.

A completely neglected dimension is security. It was brought up only by INT6 despite its importance. Security is the only dimension, which affects all layers and should be an integral part of the framework. Interestingly, when looking again at the qualitative analysis of public discourse, security was not addressed in the reviewed literature. Even reviewing the identified challenges of AI in Appendix D, the only challenge which relates to security is trust. Security forms an indispensable requirement for trust. If an AI system contains a security vulnerability and for example, loses user data, would have a dramatic impact on the confidence in AI. Therefore, security is a topic, which is neglected by the literature and should be addressed with the framework.

4.2.1 Final Framework

With all collected insights, a final version of the framework was built, which is illustrated in Figure 4.10. Multiple drafts were tried out and adapted to the feedback obtained from various interested parties. Some of these drafts are displayed in Appendix G. However, the framework's essence has been maintained. It still consists of a functional and a technical layer. The functional layer kept the subdivision into three dimensions and XAI as well as contextual adaptation are preserved. Based on the interviewees' input, HITL was exchanged with HAI. The advantage of using HAI consists of the term being more general than HITL. However, the exchange needed to clarify distinct definitions for both terms. Therefore, the definitions presented in Section 2.3 were introduced. HAI's definition was made more universally valid by removing *ethically aligned design* and *technology enhancement* from the HAI framework of Xu (2019). Therefore, HAI only focuses on *human factors design* and how one or many humans can become an integral part of HAI. Thus, the term is on the same level than XAI and contextual adaptation. The brought up term expert-in-the-loop was considered as an option for the functional layer. The goal is to integrate not only the user in the system but also the expert as a knowledge provider. Despite its importance, it was decided to integrate the expert-in-the-loop into HAI.

Furthermore, the visualisation was improved by increasing the overall quality and adding shadows. During the development, a special focus was always placed on visualisation. The goal consisted of not only building a usable framework, but also an appealing one. Since only one interview partner straightaway recognised the cube, it has been considered to enhance the 3D effect. Shadows were the right tool to improve spatial representation. Now the functional layer surrounds the technical one by placing the planes of the cube behind. The technical layer was transformed from a circle into a sphere with the shadows. As a result, the technical layer forms the cube's core.

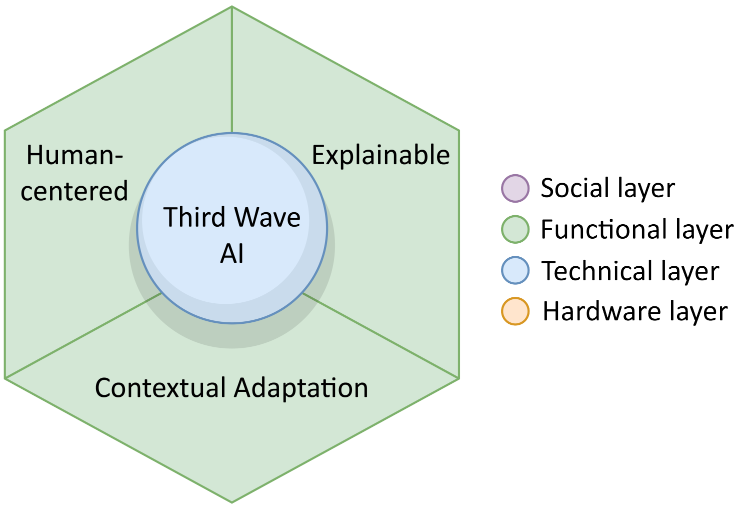


Figure 4.10: Final version of the framework

The biggest change can be suspected when looking at the technical layer. As proposed by INT9, contextual explanatory models were replaced by Third Wave AI. The technical core is undetermined because based on the results, there is no technology, which will doubtless become the Third Wave AI's initiator. The spatial representation helps to explain the framework. By the spatial extension, the interplay of the components can be visualised. The functional layer frames the technical core, which represents the requirements for achieving Third Wave AI. Not only the technology will lead to Third Wave AI. Instead, it must comply with the boundary condition of the functional dimensions.

Otherwise, the technology's benefit will be minimal, and will most likely reveal the same challenges and shortcomings as the previous waves. Besides the functional layer being essential for technical one, vice versa, there is also a dependency. The technical core is quintessential for the technical layer. As mentioned multiple times, there is no technology, which currently achieved Third Wave AI. However, an update of technology is required to accomplish all functional dimensions.

On the functional layer, the improved illustration of the cube gives a better idea of the dimension's interplay. The cube needs each dimension for stability, and therefore, the significant benefit of Third Wave AI will only be achieved when all dimensions are involved and united. For example, DARPA always uses the already mentioned in Section 2.1 process of identifying a cat to explain their contextual explanatory models. These systems will identify a cat by its nose, its ears, its tail and its fur. When someone takes the system to maybe Indonesia, where some cats have a stumpy tail. An example of a cat with a stumpy tail is displayed in Figure 4.11. A statistical model would fail without adjusting the model. However, the combination of the three functional dimensions of Third Wave AI can overcome this challenge. The systems would falsely identify the first cat as another animal. With XAI, the user can figure out based on the explanation, what has gone wrong and why the system identified the cat not as a cat. Then he can give a response and tell the system that the cats in Indonesia can have stumpy tails. The system can then learn based on this response and adapt to the change of environment. Same with XAI in combination with the other dimension. The explanation will only be as good as the user who understands it. Therefore, the explanation needs to be adaptive to the user, to the situation, to the context and the environment. Also, to make a real HAI system, it needs to be explainable and adaptive to the user. To achieve a perfect integration of the user, he needs to be able to interact with the system naturally. Overall, every dimension dependents on the others to achieve completeness.

Despite the many suggestions, it was decided not to add a layer to the main framework. Otherwise, the framework's simplicity would be endangered. Nevertheless, in order to achieve the requested comprehensiveness, additional views of the framework turned out to be the best option. For visualising the synergy between functional and social requirements, a zoom-out was built, which is presented in Figure 4.12. The added social layer consists of three dimensions. Again a three-dimensional cube was chosen where the main framework is placed. As the functional cube, the social cube should provide the boundary conditions for the functional and technical layer. The social challenges have an impact

Figure 4.11: Sample of a cat with stumpy tail in Indonesia

on each layer and specify how they should or should not be shaped. Ethical design and society-in-the-loop were chosen based on the interviewee's responses show in Section 4.2. Another argument for adding ethical design was the removal of ethically aligned design from HAI's definition of Xu (2019). Additionally, HITL was annexed to the social layer. HITL forms the counterpart of society-in-the-loop. A system, which only focuses on society could cause the neglect of individual persons. The disregarding would result in the system no longer being used. Since the user is not interested in utilising a system that does not take his interests into account. Therefore, it is essential to include HITL as a representation of individual needs. User-in-the-loop was another discussed option instead of HITL. The problem with user-in-the-loop is again the ambiguous utilization of the term. Evers et al. (2014) use the term in the context of human-computer interaction and define it as a user participation's enabler for self-adaptive applications. Based on the definition used in this thesis, user-in-the-loop would be a part of HAI. Otherwise, Schoenen and Yanikomeroglu (2014) use the term user-in-the-loop totally contrasted. They name their approach of demand shaping accordingly. However, none of the definitions includes the mentioned aspects, and hence, HITL remained the best option to go.

Besides the zoom-out, a zoom-in shown in Figure 4.13 was added. The zoom-in should allow a closer look at the technical layer and the newly introduced hardware layer. At first, it was planed to visualize the different technologies. However, it turned out not being the right approach. Since the other layers consist of requirements, the techno-

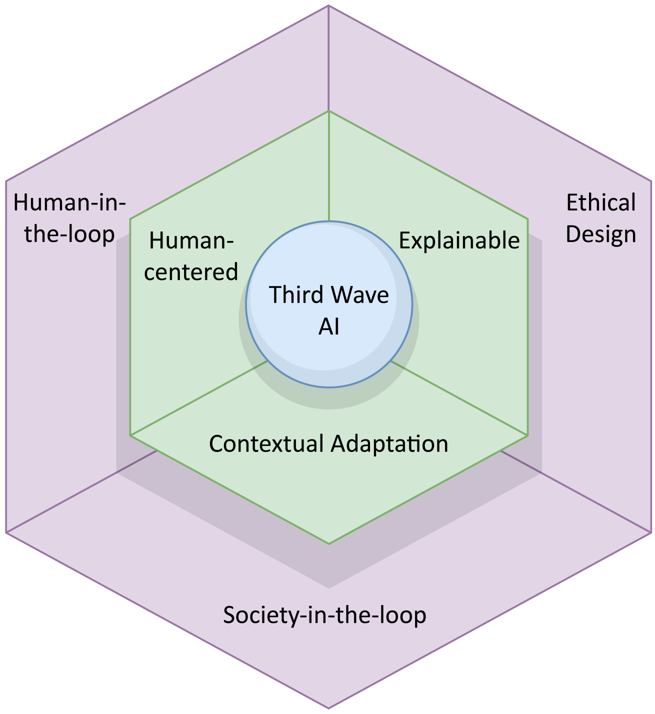


Figure 4.12: Zoom-out of the framework

logical and hardware layer would consist of examples. Additionally, since there were so many different technologies and hardware, which could enable Third Wave AI, it was not possible to limit the dimensions in a meaningful way. Choosing only the most important dimensions would have led to the framework becoming subjective and incomplete, as there is no sufficiently dominant technology. Furthermore, the feedback from INT5 questioned the technical layer by bringing up the requirements of DARPA. These requirements fit into the framework much better. Therefore, the four requirements perceiving, learning, abstracting and reasoning were added as technical dimensions. With these dimensions, the framework can build up on DARPA's without being committed to their contextual explanatory models. In addition to the technical layer, the hardware layer was introduced as a feedback's result from the interviews. For the hardware layer, the requirements stated by INT8 were added. Low power and low latency were brought up as requirements for neuromorphic computing, which were also mentioned by other interviewees. Additionally, low latency will be a requirement for achieving the real-time AI of Dodhiawala et al. (1989). The real-time AI is one of the AI framework, which was analysed in Appendix A. Positioning the hardware layer as an oval into the technical layer indicates that the hardware belongs to the core of the technology. It enables different technologies to perform at best. Therefore, the zoom-in represents the synergy between hardware and technology. The hardware dimensions are the most generic dimensions of the frameworks. Therefore it is likely that these dimensions need some adjustments. Security can be imagined in the framework as a plane cutting through all levels.

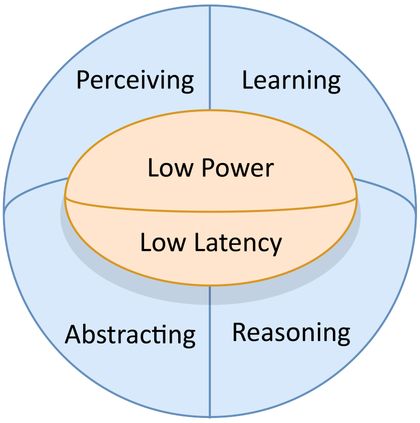


Figure 4.13: Zoom-in of the framework

As already listed, INT6 mentioned the absence of security in the framework. This absence is still the case since it would be a dimension, which has to be a part of each layer. Since security must be taken into account on the hardware, technical, functional and social layer. Each layer has its weaknesses, which need to be secured. Defining security as a layer would also not be sufficient since security does not influence the other layers. There is not a security approach, which solves social and technical challenges. Therefore, it is impossible to visualise security in this framework. It does not mean that security will not be important for Third Wave AI. As already mentioned in Section 4.2, security is indispensable for AI and will become more and more important as AI increasingly gains relevance in our daily lives.

4.2.2 Application

The framework can be used for many applications. One is to categorise technologies, approaches, trends or even theories. In Figure 4.14, the technologies mentioned in the interviews were incorporated. On the technical layer, DARPA's contextual explanatory models, INT4's cognitive architecture and INT5's combination of the first and second wave were added. This structure of the technical layer would have been embedded in the framework if INT5 had not brought up the requirements. Looking at the hardware oval, it is split into three different hardware approaches. They were added because of INT8's response. Cloud, edge and neuromorphic computing are all terms, which will or already have impacted AI. Furthermore, all parts of the technical and hardware layer have the potential to shape the Third Wave AI fundamentally. Although this will be the case for some, not all will be involved in the wave. Maybe none of them will be a component of Third Wave AI.

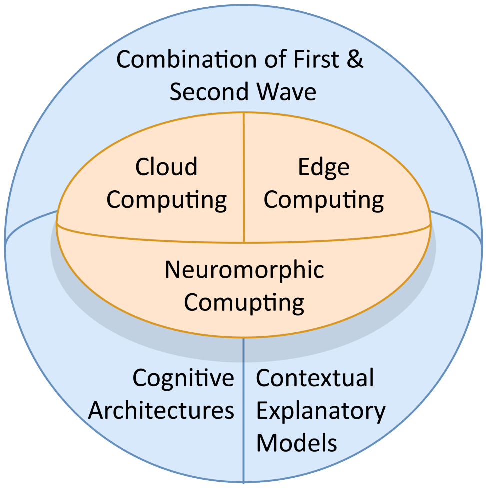


Figure 4.14: Examples which could fulfil the technical and hardware layer

The interviewees suggested heterogeneous users for the framework. INT8 recommended it to PhD candidates or graduates. For INT1, the managers belong to the potential user group. Mainly users were suggested which have or need a broad view of AI. This pattern was also apparent in the proposed applications. INT2 and INT6 could image using it for conceptual design of applications. INT9 submitted the use of the framework as a tool for the design thinking process. Most suggestions included using the framework for creating new ideas. However, INT6 saw potential in using it for the evaluation of services. Additionally, INT3 could envisage using it for explaining the layer's synergies and requirements besides the technical ones. The framework could also be an important reminder of the key aspects to consider when implementing an AI system in a company (INT1). Not only the industry was suggested as a field of application. INT6 brought up that the framework could be used for the political discourse. Additionally, INT6 suggested building sustainable solutions with the framework's use. INT10 proposed an alternative application for team leaders. The framework could be used as guidance for putting together a team. Using the dimensions as requirements for the team composition instead of the system. Therefore, the framework would be used to identify which skills are needed and then to look for potential employees with the corresponding skillset.

INT7 brought up a totally different application: "It's almost drawn as something that could be used as the basis of a spider diagram" (INT7). This statement turned out to be an excellent idea. The framework is very suitable to build a spider diagram. Each dimension forms a node, which can be rated from zero to ten. A potential Third Wave AI can be analysed with the spider diagram, and it can be identified if dimensions are neglected. For the technical requirements, the values of DARPA can be used as a baseline. Converting the DARPA's values to the spider diagram results in a required score for perceiving of 7.5, learning of 7.5, abstracting of 5 and reasoning of 7.5. For many of the suggested applications, the spider diagram could be used and would provide additional value. For example, when designing an application, the framework guides the design process. Finally, the application is evaluated with the spider diagram to show if requirements have been met or if there is a need for action. Another option of how the spider diagram can be used is to compare different solutions. In Figure 4.15 an example of an evaluation is demonstrated. Besides the sample, the DARPA's baseline is displayed.

For the sample analysis, the company Samurai Labs was used. They build a AI, which tries to make the internet a safer place. It identifies potential dangers like cyberbullying, sexual harassment or indications of suicide attempts (INT5). All the information used for the evaluation were collected during the interview INT5. Therefore, the analysis is subjective and based on minimal information. During the interview, a rudimentary demo was presented, but it was not possible to try out the application. Looking at the requirements of the main framework in Figure 4.15, Samurai Labs does not perform particularly well. The highest value besides human-centered is achieved in reasoning with six and explainable with five. The reason why explainable is ranked lower than reasoning is the way how explanations are provided. All other requirements are lower than five. Overall, it can be summarized to become a Third Wave AI every dimension needs to be improved. Especially when comparing the AI to the baseline, DARPA's values are much higher. On the hardware layer, low power and low latency were set to zero since no information about hardware was mentioned. The highest numbers are achieved on the social layer. The only exception was ethical design since it was barely

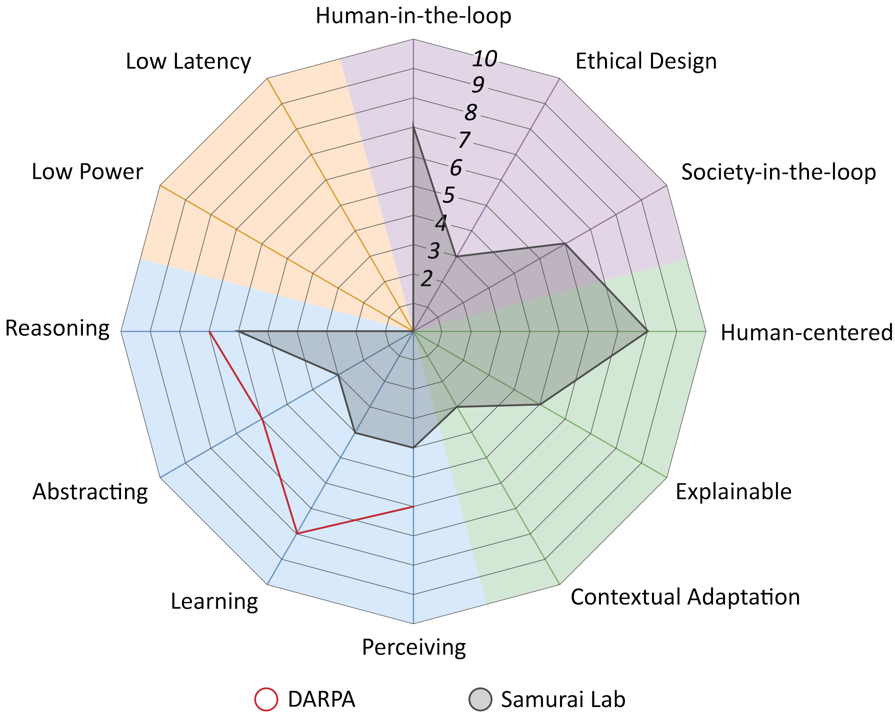


Figure 4.15: Sample analysis of Samurai Labs

touched in the presentation. The solution has many positive aspects and a noble goal. However, the outputs are very sensitive and could be abused. For example, insurance companies could use mental information to make the insurance's pricing dependent on it. HAI and HITL were ranked as the best performing capabilities. The reason for the high values is the solution's focus on people. On the one hand, the victims, which are affected by the danger and on the other hand, the authorities and experts, which try to prevent the dangers or need to intervene in time. Therefore, Samurai Labs need to master these dimensions and is on the right track.

However, only the demo and presented information were used for this evaluation. It can be safely assumed that their advanced solutions achieve much higher values. To put the analysis into a more general context, another AI needed to be analysed. However, this is impossible, as there is no application, where sufficient information is available. In Appendix I is provided a template, which can be used to evaluate a Third Wave AI system.

5

Discussion

Comparing the framework to sociotechnical systems, a reasonable degree of similarity can be identified. The technical layer can be assimilated to the technical system. Additionally, the social layer has components very similar characteristics to the social system. Especially, HITL intersects with the social system by focusing on the user's needs. To identify the individual needs, Maslow's hierarchy of need could be used. However, for the integration of Maslow's hierarchy of need, further work is necessary. To achieve the goal of improving both systems, joint optimisation is required. Looking at Figure 4.12 can be recognised that besides both systems, the joint optimisation is also included in the framework. The functional layer forms the transition from the technical to the functional system and vice versa.

As already mentioned in Section 2.4, XAI is such a joint optimiser by connecting and providing improvements for both layers. Equal HAI and contextual adaptation also cover the role of a joint optimiser. HAI with the focus on user integration is the perfect example of a joint optimisation. For the social system, the benefits consist of the AI taking the user into account and providing the desired information. For example, this can be achieved by offering a suitable and easy to work with user interface like Maguire (2014) suggested for computer information systems in general. The benefit for the technical system with a well-built user interface is not only the user being more willing to use the system, which constantly improves the system. Moreover, the AI gets all information it needs in an optimal form. Contextual adaptation also affects technical and social system by improving both. If a system adapts to the user, the accuracy of the results will improve, which increases the user's satisfaction and make him more willing to accept a decision. For the technical system, an improvement in accuracy results in being more reliable and therefore, an increase in performance.

Additionally, the combination of all layers has a significant impact on the social and technical system. When reviewing the example of the cats with stumpy tails, introduced in subsection 4.2.1, the enhancement through the functional layer makes the user more satisfied since he can identify what went wrong (XAI), provide feedback over the user interface (HAI) and only has to do this once since the system adapts (contextual adaptation). On the technical layer, the AI learns new knowledge, which it can use for a similar situation. It benefits from the user's experience without having to acquire the knowledge itself. Therefore the system perceived new information, learnt from the feedback and can utilise the knowledge for reasoning and abstracting.

Overall the functional layer can be considered as one large joint optimiser, or each dimension is a joint optimiser on its own. Either way, with the comparison to the sociotechnical systems, it can be demonstrated that the framework meets the goal of shifting the focus from only technical to a more comprehensive one. Additionally, the chosen layers and representation is supported by the sociotechnical systems theory. The functional layer is the transitional layer, which is implemented in the framework with the functional cube wrapping the technical core and being enclosed by the social cube.

When a system masters contextual adaptation, it will be aware of all changes in the environment and context. This awareness leads to the ability that the system recognises the changes caused by its decision. Pearl and Mackenzie (2018) describe an example where a robot turns on the vacuum cleaner while the user is asleep. The user reacts by telling the robot that it should not do that. Pearl and Mackenzie (2018) argue that the robot needs to understand cause-and-effect relation. Otherwise, because of the user's response, it would never vacuum again. It can only tell that the action appeared as unpleasant and not why it was unpleasant. Therefore, the robot needs to master contextual adaptation. It must be aware of its context and environment and which effects its actions cause. In the described example, the robot would have recognised that its action waked up somebody and based on the person's feedback, it identifies it as an unpleasant consequence. When it knows that waking up a human is unwise and vacuuming leads to waking up a human, it can adapt to this knowledge. As a result, next time, it will not vacuum when a human is sleeping in the same room.

As discussed in Section 2.4 the Ladder of Causation by Pearl and Mackenzie (2018) plays an important role for AI. Achieving contextual adaptation can resolve the lack of fulfilling the first step *association*. Especially continuous contextual adaptation will improve climbing the Ladder of Causation. Since the first step of contextual adaptation is observing the context and environment, it needs to fulfil the *association* step. The second step is consists of adapting to the collected information, which represents the step *intervention*. When the system understands the effects of its action, even the third step *counterfactuals* is involved.

Hintze (2016) calls this level of AI type III. He describes these systems under the application of the theory of mind. Based on the theory of mind, the system should be capable of understanding that other entities' behaviour is affected by emotions and thoughts (Hintze, 2016). As a result, the system will be aware of its impact, which will, for example, improve the decision-making process. INT4 indicates that the AI is at this point able to answer the following questions "What do they expect to hear? What do they already know? What are they confused about?". However, the system can then use a wrong advice or information to cause a user deciding for the right option. At this point, the system embodies the capability to manipulate humans or other systems. Supposed an AI achieves such a level of intelligence, the question of Pearl and Mackenzie (2018) arises about the human's capability to build a system that can distinguish good from bad. In that case, it will be essential to answering this question with yes. Otherwise, these systems will be unpredictable. When using the framework to build such advanced AIs, the dimension should be fulfilled, which requires a human-centered approach and also mastering the social layer. Therefore, the system has to be able to distinguish good from bad, at least from a human's point of view. Especially the society-in-the-loop will play an important role. If a system only takes the user into account for its decision, the consequences of the decision could be bad for another human or even the society. For that reason, it is indispensable to keep all dimensions in mind when building such new systems. The framework can help to identify a lack of focus on dimensions or layers by providing a thought-provoking guide. With the spider diagram, solutions can be evaluated to make sure the focus is not only on technical and hardware requirements but also on functional and social layer.

A system of such a level of intelligence will not only be aware of the effects it causes, but as a result, it will also be aware of its capabilities. The system accomplishes self-awareness and can then proactively work together with other AIs. It can decide if it needs help for a certain task and which system or human will provide it. "Having the ability to know what you don't know is also an important of necessity" (INT4). Hintze (2016) categorises the state of self-awareness as type IV and final form of AI. The self-awareness is also essential for the last level based on the Ladder of Causation. The final level is *counterfactuals* and consist of the activities imagining, retrospection and understanding (Pearl & Mackenzie, 2018). As already mentioned, understanding will be achieved when the system understands the effects of its behaviour. For retrospection, self-awareness is necessary and provides the foundation for imagination. A system needs to be aware of itself and in which position or situation is to imaging itself in another position or situation. For example, it needs to be aware of other systems to imaging itself as another system. If an AI is capable of doing these activities, an ASI is engineered with human-like intelligence or even outperform humans. Nevertheless, mastering contextual adaptation on the described level could be the fourth or even the 50th waves. The only fact which can be said with certainty is the research remains quite far away from achieving an ASI (Lee, 2018; Scott, 2018).

Not only contextual adaptation has its challenges, but also explainability. The biggest criticism to similar approaches is often mentioned when discussing XAI is for making a system explainable, it needs to be of lower complexity (Bloomberg, 2018). Systems with lower complexity will be not as mighty, and therefore the added value will be accordingly reduced (Bloomberg, 2018). The functional layer will always compete with the technical layer. All the layers have issues facing each other. For example, when a developer builds an application, he wants to provide as many features as possible. Accordingly, he focuses on the functional layer to provide the best contextual adaptation. To achieve that goal, it would require to use all available data from the user and to collect every information while the user interacts with the application. All the data is used for the adaptation to the user and to improve the AI, but will lead to privacy and fairness issues. The juxtaposition of the social and functional levels is not a problem of AI but of software development in general, which is a similar discussion as explainability. To achieve the biggest value, all layers have to be taken into account.

Kaski et al. (2019) pose the question if the focus in Europe on ethical issues slows down the progress of AI. Based on their conducted interviews, they conclude a balance between ethical issues and the benefits of AI needs to be established. The suggested framework has the benefit that it does not focus only on the social layer but also on the functional, technical and hardware layer. By using the framework, a balance can be accomplished, and an evaluation can be performed based on the application. Not every system will need to fulfil every dimension perfectly. The example used by Kaski et al. (2019) is 4G usage. Users do not need an explanation to trust and to use the technology. As with 4G, there will be AI systems for certain tasks which may not have to fulfil the requirements of Third Wave AI. As already mentioned, INT10 had put forward the hypothesis that explanation is only needed if the system appears as imperfect. In his opinion, the demand for explanation will only arise in the event of AI's misconduct, and at that point, the systems will not fail, and the demand will disappear. The requirements of the waves after the third one have to be defined when Third Wave AI is achieved. It is possible that dimensions will disappear or exchanged with new ones.

The minimum requirement for XAI is to provide an explanation to the user, which he understands. However, to master XAI, the AI needs to be able to fully adapt to users. For achieving such a perfect adaptation to the user, the system is required to understand the user. For INT4, "AI needs at least to be able to explain it in the way that people can understand. And that also requires quite high-level cognitive capability. It requires a theory of mind". As already mentioned, when AI achieves the theory of mind, it is capable of recognising other entities' emotions and thoughts, which affect their behaviour (Hintze, 2016). The theory of mind would immensely improve XAI. At that point, an AI could adapt its explanation to the user's state of mind. Furthermore, it could identify which information is needed to make a decision comprehensible based on the user's education. As explained for contextual adaptation, mastering such skills will lead to the point where the systems are able to manipulate the user.

The framework provides a foundation for further discussions about the future of AI. With the presented results, this thesis shows that the term Third Wave AI needs clarification, especially when looking at the interviews, the framework of DARPA is not sufficient. The ambiguity of the used terms like contextual adaptation or contextual exploratory models is addressed by introducing and adjusting existing definitions. Also, the terms were put into context and positioned with the framework. This overview emphasises a bird's eye view over all layers and dimensions. Furthermore, DARPA's only technical view was extended by a social, functional and hardware view. Overall, the framework defines precise requirements, which are needed to achieve a new wave of AI. In combination with the spider diagram, it provides a guideline and offers the possibility to rate existing solutions. With the spider diagram, it is possible to identify an eventual need for action or to uncover the full potential of a solution.

To validate the improvement of the introduced framework, the identified challenges by the public discourse analysis illustrated in Appendix D and discussed in subsection 4.1.2

were reviewed. The resulted network is presented in Appendix J. Each challenge was categorised with colours based on the layer, which addresses it. 27 challenges were identified in the analysis, and the new framework effects 19 of them. Compared to DARPA with nine addressed challenges, the framework doubles the quantity. Therefore, the framework provides a significant improvement versus to DARPA's initial framework. Looking at the challenges, each layer is present, which proofs that the choice of the levels is justified. There are still challenges left uninvolved. Some of them like consciousness will as already discussed, be part of following waves.

The fundamental question of this thesis is how legitimate DARPA's framework is. By answering the four research questions, the outcome can be summarised as follows: DARPA's framework is a great starting point. However, the legitimacy of the term is limited by the lack of definitions and the applicability. The analysis of the public discourse and the interviews confirmed these problem areas. With the suggested framework in subsection 4.2.1 and the established definitions in Section 2.3, these issues are addressed, and a solution is proposed. The provided definitions and the resulting clarification help to reduce the intermixing of terms. The framework will help to establish Third Wave AI not only as a marketing term. It makes the Third Wave AI applicable not only for the industry but also for academia. This academic applicability is essential to provide a foundation for discussions. The discourse about the future of AI is becoming more and more critical, and this framework helps to take bird's eye view for the industry, but also for academia. Hence, it supports the process of understanding the dimensions' synergies, which is necessary to identify the role of AI, and for building long-term sustainable solutions.

6

Limitations and Future Work

Despite the potential added value by the suggested framework, additional validation is required to investigate if the final version of the framework fulfils the needs. As pointed out in Section 4.2, the improved version was already shown during the last interviews, but often there was not enough time to evaluate different versions of the framework. To keep the interviews comparable, the interview version was discussed first according to the guide. In case time was left at the end, the newest version was presented and challenged. Another interview round with the same interview partner would resolve this described issue. It would be a great way to validate if the final version handles the suggestions collected in the first round. Furthermore, an increased number of interview partners for the second round would help test the framework more extensively. An additional option would be to conduct a study in the form of a survey to achieve an even broader validation.

As already mentioned, security is a strongly neglected topic. Therefore a necessity for future work exists about security and AI. Even in the suggested framework, security is not present neither as layer nor as dimension. There could be a way to integrate security in the framework but would only solve the problem to a limited extent. It is recommended to address the topic of security in AI in the next project. The result of such a project could be a security framework. Therefore, the suggested framework could help to build a security-related framework since it guides trough addressing the different layers. There exists a high risk in such a project to focus only on the security of the technical layer without the suggested framework. However, security in the other layers become more and more crucial. For example, on the social layer, there is a danger of AI being harmful to societies. An AI could only consider one society and therefore would turn against others. Such a situation must be prevented with the appropriate security measures.

Another limitation of the suggested framework is based on the collected results of this Master's thesis, no suggestion can be given on how much each domain must be fulfilled to achieve Third Wave AI. Especially when using the spider, the needed score is not identified. Therefore, the spider diagram appears to be the most useful for comparisons of different approaches or AIs. The results show that all dimensions have to be targeted to achieve Third Wave AI. However, when the highest score ten is achieved on every dimension, it is likely not a Third Wave AI anymore. Instead, it could be a Fourth or Fifth Wave AI. To determine the matching score for each wave, additional research must be carried out. Future work could test the spider diagram with as many use cases as possible. Many of the solutions and projects mentioned in this Master's thesis would be well suited for such analysis like Aigo or SingularityNET. By comparing all the solutions to each other, the leaders could be identified in each dimension. Based on these leaders, it would be possible to determine the score needed in each dimension.

A proof of concept (PoC) is as well necessary to achieve improvements to the framework. For example, carrying out a PoC with an innovation consulting company would help to assess the acceptance and probability of success. Especially for the spider diagram, such a test would support to refine it and for instance display requirements for the implementation of the spider diagram as a web application. Realising the spider diagram as a web application would strongly improve the applicability of the framework, particularly for the industry. The reason for the improvement of usability with a web application consists of creating a spider diagram, either Excel or a drawing tool needs to be used. The drawing tool like Adobe Illustrator, which was used for Figure 4.15 has the benefit of providing the tool to create a visually appealing spider diagram but at the expense of efficiency. When using Excel, the efficiency is much higher by using the provided spider diagram feature. However, the feature does not provide enough option to generate a diagram appealing enough for a professional presentation. Additionally, Excel is not the best option for a workshop or a brainstorming session. Therefore, a web application with collaboration or brainstorming functionalities would massively improve the applicability. Furthermore, by conducting a PoC, new applications besides the spider diagram could also be identified. Another benefit of a PoC would imply a best practice, which could be elaborated. The best practice would improve the usability of the framework in the industry.

Besides the lack of validation, another limitation is the public discourse, which was only analysed in English. In future work, it would be interesting to expand the analysis to among languages. With such an analysis, differences in focus per languages could be identified, and conclusions could be drawn about regional differences. The framework could be confronted with such differences, and based on the results, new dimensions could be implemented, or additional applications could be identified.

The framework provides a foundation for the academic discourse, but for future work, it will be important to not only argue with the framework but also to discuss the framework. This critical analysis is essential to not encounter the same problems as DARPA's framework. With the academic discourse, additional applications can be identified. These applications are not limited to AI or computer science. The framework can be used for interdisciplinary discussion, which will be essential because AI will influence every field. Especially when working on HITL and trying to achieve experts-in-the-loop, the interdisciplinary exchange will be crucial.

Based on the analysis of this Master's thesis, there is an absence of Third Wave AI application on the market at the moment. Therefore, the framework relies on a prediction of how the future of AI could and should evolve. To keep the framework up-to-date, it is necessary to review the framework continually and to adapt it to discoveries and breakthroughs. The framework can be used to analyse these updates, but if they lead to the demand for changing the defined requirements, then it will be crucial to fulfil these demands. Otherwise, the framework would quickly become obsolete and would no longer be suitable for further analysis.

References

Adler, J. (2020, June 23). *Transforming forklifts with cloud robotics: Our investment in third wave* [Medium] [Library Catalog: medium.com]. Retrieved June 23, 2020, from https://medium.com/toyota-ai-ventures/transforming-forklifts-with-cloud-robotics-our-investment-in-third-wave-512a2fe3230

Ambjorn, C. (2008, May). Seatrack web forecasts and backtracking of oil spills - an efficient tool to find illegal spills using AIS [ISSN: 2150-6035], In *2008 IEEE/OES US/EU-baltic international symposium*. 2008 IEEE/OES US/EU-Baltic International Symposium. ISSN: 2150-6035. https://doi.org/10.1109/BALTIC.2008.4625512

Appelbaum, S. H. (1997). Socio-technical systems theory: An intervention strategy for organizational development [Publisher: MCB UP Ltd]. *Management Decision*, *35*(6), 452–463. https://doi.org/10.1108/00251749710173823

Bai, S., Fang, D., & Zhang, Q. (2019). Research and application of artificial intelligence technology in the field of risk perception [Publisher: IOP Publishing]. *Journal of Physics: Conference Series*, *1302*, 022002. https://doi.org/10.1088/1742-6596/1302/2/022002

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Beinart, M. (2019, September 23). *DARPA releasing new project opportunities for 'third wave' AI research*. Retrieved March 26, 2020, from https://www.defensedaily.com/darpa-releasing-new-project-opportunities-third-wave-ai-research/cyber/

Bisen, V. S. (2020, May 20). *What is human in the loop machine learning: Why & how used in AI?* [Medium] [Library Catalog: medium.com]. Retrieved July 21, 2020, from https://medium.com/vsinghbisen/what-is-human-in-the-loop-machine-learning-why-how-used-in-ai-60c7b44eb2c0

Bloomberg, J. (2018, September 16). *Don't trust artificial intelligence? time to open the AI 'black box'* [Forbes] [Library Catalog: www.forbes.com Section: Innovation]. Retrieved June 10, 2020, from https://www.forbes.com/sites/jasonbloomberg/2018/09/16/dont-trust-artificial-intelligence-time-to-open-the-ai-black-box/

Buxmann, P., Hess, T., & Thatcher, J. (2019). Call for papers, issue 1/2021: AI-based information systems. *Business & Information Systems Engineering*, *61*(4), 545–547. https://doi.org/10.1007/s12599-019-00606-2

Carroll, W. M. (2020, February 1). *Emerging technologies for nurses: Implications for practice* [Google-Books-ID: UJawDwAAQBAJ]. Springer Publishing Company.

Cecotti, H., & Graser, A. (2011). Convolutional neural networks for p300 detection with application to brain-computer interfaces [Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *33*(3), 433–445. https://doi.org/10.1109/TPAMI.2010.125

Cukier, W., Bauer, R., & Middleton, C. (2004). Applying habermas' validity claims as a standard for critical discourse analysis (B. Kaplan, D. P. Truex, D. Wastell, A. T. Wood-Harper, & J. I. DeGross, Eds.). In B. Kaplan, D. P. Truex, D. Wastell, A. T. Wood-Harper, & J. I. DeGross (Eds.), *Information systems research: Relevant theory and informed practice*. Boston, MA, Springer US. https://doi.org/10.1007/1-4020-8095-6_14

Curtis, L. F. (1914). The effect of delta and star connections upon transformer wave forms [Conference Name: Transactions of the American Institute of Electrical Engineers]. *Transactions of the American Institute of Electrical Engineers*, *XXXIII*(2), 1273–1282. https://doi.org/10.1109/T-AIEE.1914.4765182

DARPA. (2019, August 12). *DARPA on Twitter* [Twitter] [@DARPA]. Retrieved October 12, 2020, from https://twitter.com/DARPA/status/1160950659866345474

Daws, R. (2018, September 28). *DARPA introduces 'third wave' of artificial intelligence* [AI news]. Retrieved March 26, 2020, from https://artificialintelligence-news.com/2018/09/28/darpa-third-wave-artificial-intelligence/

Diskin, M. G., Austin, E. J., & Roche, J. F. (2002). Exogenous hormonal manipulation of ovarian activity in cattle. *Domestic Animal Endocrinology*, *23*(1), 211–228. https://doi.org/10.1016/S0739-7240(02)00158-3

Dodhiawala, R., Sridharan, N. S., Raulefs, P., & Pickering, C. (1989). Real-time AI systems: A definition and an architecture, In *In proceedings of the international joint conference on artificial intelligence.*

Doloc, C. (2019, October 29). *Applications of computational intelligence in data-driven trading* [Google-Books-ID: wQazDwAAQBAJ]. John Wiley & Sons.

Eric Trist, Beulah Trist, & Hugh Murray. (1993). *The social engagement of social science, volume 2: A tavistock anthology–the socio-technical perspective.* University of Pennsylvania Press. Retrieved October 17, 2020, from https://sfx.ethz.ch/sfx_lib4ri?url_ver=Z39.88-2004&ctx_ver=Z39.88-2004&ctx_enc=info:ofi/enc:UTF-8&rfr_id=info:sid/sfxit.com:opac_856&url_ctx_fmt=info:ofi/fmt:kev:mtx:ctx&sfx.ignore_date_threshold=1&rft.object_id=3710000000631476&svc_val_fmt=info:ofi/fmt:kev:mtx:sch_svc&

Evers, C., Kniewel, R., Geihs, K., & Schmidt, L. (2014). The user in the loop: Enabling user participation for self-adaptive applications. *Future Generation Computer Systems*, *34*, 110–123. https://doi.org/10.1016/j.future.2013.12.010

Flanagan, J. C. (1954). The critical incident technique [Place: US Publisher: American Psychological Association]. *Psychological Bulletin*, *51*(4), 327–358. https://doi.org/10.1037/h0061470

Galactic PA. (2018, August 3). *Galactic PA on Twitter* [Twitter] [@GalacticPA]. Retrieved October 12, 2020, from https://twitter.com/GalacticPA/status/1025403530096640000

Goebel, R., Chander, A., Holzinger, K., Lecue, F., Akata, Z., Stumpf, S., Kieseberg, P., & Holzinger, A. (2018). Explainable AI: The new 42? (A. Holzinger, P. Kieseberg, A. M. Tjoa, & E. Weippl, Eds.). In A. Holzinger, P. Kieseberg, A. M. Tjoa, & E. Weippl (Eds.), *Machine learning and knowledge extraction*, Cham, Springer International Publishing. https://doi.org/10.1007/978-3-319-99740-7_21

Hintze, A. (2016, November 14). *Understanding the four types of AI, from reactive robots to self-aware beings* [The conversation] [Library Catalog: theconversation.com]. Retrieved May 6, 2020, from http://theconversation.com/understanding-the-four-types-of-ai-from-reactive-robots-to-self-aware-beings-67616

Holzinger, A. (2016). Interactive machine learning for health informatics: When do we need the human-in-the-loop? *Brain Informatics*, *3*(2), 119–131. https://doi.org/10.1007/s40708-016-0042-6

Kaski, S., Ailisto, H., & Suominen, A. (2019, June 19). International AI experts: Towards the third wave of artificial intelligence [Publisher: Ministry of Economic Affairs and Employment], In *Leading the way into the age of artificial intelligence : Final* *report of finland's artificial intelligence programme 2019*. Publisher: Ministry of Economic Affairs and Employment. Retrieved June 19, 2020, from https://cris.vtt.fi/en/publications/international-ai-experts-towards-the-third-wave-of-artificial-int

Launchbury, J. (2017, February 15). *A DARPA perspective on artificial intelligence*. Retrieved February 26, 2020, from https://www.youtube.com/watch?v=-O01G3tSYpU#action=share

Lee, K.-F. (2018, October 22). *The four waves of a.i.* [Fortune] [Library Catalog: fortune.com]. Retrieved March 10, 2020, from https://fortune.com/2018/10/22/artificial-intelligence-ai-deep-learning-kai-fu-lee/

Maguire, M. (2014). Socio-technical systems and interaction design – 21st century relevance. *Applied Ergonomics*, *45*(2), 162–170. https://doi.org/10.1016/j.apergo.2013.05.011

Militello, L., Arbuckle, N., Saleem, J., Patterson, E., Flanagan, M., Haggstrom, D., & Doebbeling, B. (2013). Sources of variation in primary care clinical workflow: Implications for the design of cognitive support. *Health informatics journal*. https://doi.org/10.1177/1460458213476968

Paul R. Daugherty, & Wilson, H. J. (2018). *Human + machine: Reimagining work in the age of AI*. Boston Massachusetts, Harvard Business Review Press.

Pearl, J., & Mackenzie, D. (2018). *The book of why: The new science of cause and effect* (First edition). New York, Basic Books.

Perez, C. E. (2017, October 13). *The next AI milestone: Bridging the semantic gap* [Medium]. Retrieved March 26, 2020, from https://medium.com/intuitionmachine/the-first-rule-of-agi-is-bc8725d21530

Popovic, A. (2001). *Contextual adaptation*. ETH Zürich. Retrieved July 22, 2020, from https://www.vs.inf.ethz.ch/edu/WS0102/UI/slides/ui_09context.pdf

Preetipadma. (2020, April 10). *Neuromorphic chips: The third wave of artificial intelligence* [Analytics insight] [Library Catalog: www.analyticsinsight.net Section: Cloud Computing]. Retrieved June 10, 2020, from https://www.analyticsinsight.net/will-neuromorphic-chips-landscape-computer-intelligence-we-know-of/

Rüegg-Stürm, J. (2003, February 1). *Das neue st. galler management-modell. grundkategorien einer integrierten managementlehre. der HSG-ansatz.* (2., durchges. A). Bern, Haupt Verlag.

Schoenen, R., & Yanikomeroglu, H. (2014). User-in-the-loop: Spatial and temporal demand shaping for sustainable wireless networks [Conference Name: IEEE Communications Magazine]. *IEEE Communications Magazine*, *52*(2), 196–203. https://doi.org/10.1109/MCOM.2014.6736762

Scopus [Scopus]. (2020). Retrieved June 30, 2020, from https://www.scopus.com/search/form.uri?display=basic

Scott, J. (2018, August 27). *Third wave AI: The coming revolution in artificial intelligence*. Retrieved February 26, 2020, from https://www.sixkin.com/posts/3rd-wave-ai/

Serrat, O. (2017). The five whys technique (O. Serrat, Ed.). In O. Serrat (Ed.), *Knowledge solutions: Tools, methods, and approaches to drive organizational performance*. Singapore, Springer. https://doi.org/10.1007/978-981-10-0983-9_32

Szu, H. H., Chang, L. C., Chu, H., & Kolluru, R. (2019). The 3rd wave AI requirements [Publisher: MedCrave Publishing]. *MOJ Applied Bionics and Biomechanics*, *Volume 3*. https://doi.org/10.15406/mojabb.2019.03.00094

Third wave. (n.d.). Retrieved October 12, 2020, from https://www.thirdwave.ai/

Toffler, A. (1980). *The third wave* [Google-Books-ID: ViRmAAAAIAAJ]. Morrow.

Voss, P. (2017, September 27). *The third wave of AI — LinkedIn*. Retrieved August 3, 2020, from https://www.linkedin.com/pulse/third-wave-ai-peter-voss/

Xu, W. (2019). Toward human-centered AI: A perspective from human-computer interaction. *Interactions*, *26*(4), 42–46. https://doi.org/10.1145/3328485

Yampolskiy, R. V. (2013). Turing test as a defining feature of AI-completeness (X.-S. Yang, Ed.). In X.-S. Yang (Ed.), *Artificial intelligence, evolutionary computing and metaheuristics: In the footsteps of alan turing*. Berlin, Heidelberg, Springer. https://doi.org/10.1007/978-3-642-29694-9_1

Zavolokina, L., Dolata, M., & Schwabe, G. (2016). FinTech – what's in a name? *ICIS 2016 Proceedings*. https://aisel.aisnet.org/icis2016/DigitalInnovation/Presentations/12

Zilis, S., & Cham, J. (2016, November 7). *The current state of machine intelligence 3.0* [O'reilly media] [Library Catalog: www.oreilly.com]. Retrieved May 10, 2020, from https://www.oreilly.com/content/the-current-state-of-machine-intelligence-3-0/

A

Overview of AI Frameworks

AI Frameworks

Implementation approach

Who:	Stuart Russel & Peter Norvig
Where:	Artificial Intelligence: A Modern Approach (Book)
When:	(2016)
What:	

Thinking Humanly	**Thinking Rationally**
"The exciting new effort to make computers think . . . *machines with minds*, in the full and literal sense." (Haugeland, 1985) "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman, 1978)	"The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Acting Humanly	**Acting Rationally**
"The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990) "The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	"Computational Intelligence is the study of the design of intelligent agents." (Poole *et al.*, 1998) "AI . . . is concerned with intelligent behavior in artifacts." (Nilsson, 1998)

Problem approach

Who:	Roman V. Yampolskiy
Where:	Turing Test as a Defining Feature of AI-Completeness (book part)
When:	(2013)
What:	

Definition 1: A problem C is **AI-Complete** if it has two properties:

1. It is in the set of AI problems (Human Oracle solvable).
2. Any AI problem can be converted into C by some polynomial time algorithm.

Definition 2: AI-Hard: A problem H is AI-Hard if and only if there is an AI-Complete problem C that is polynomial time Turing-reducible to H.

Definition 3: AI-Easy: The complexity class AI-easy is the set of problems that are solvable in polynomial time by a deterministic Turing machine with an oracle for some AI problem. In other words, a problem X is AI-easy if and only if there exists some AI problem Y such that X is polynomial-time Turing reducible to Y. This means that given an oracle for Y, there exists an algorithm that solves X in polynomial time.

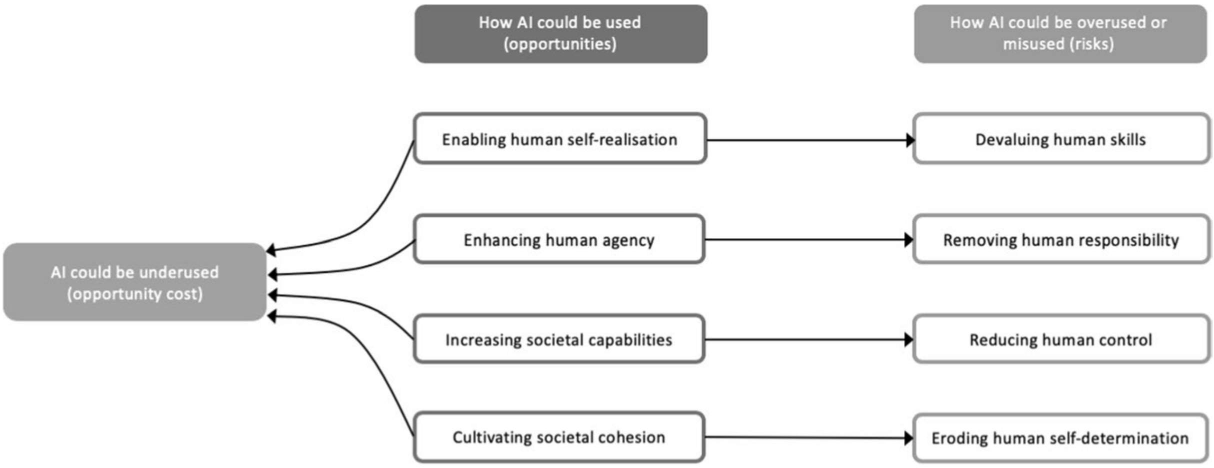
Ethical / opportunity approach

Who: Luciano Floridi, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, Francesca Rossi, Burkhard Schafer, Peggy Valcke & Effy Vayena

Where: AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations (paper)

When: (2018)

What:



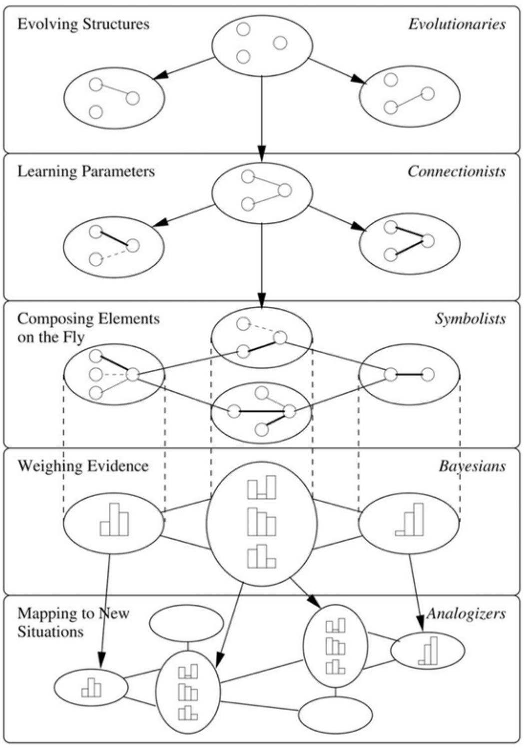
Machine learning approach

Who: Pedro Domingos

Where: The master algorithm: how the quest for the ultimate learning machine will remake our world (book)

When: (2015)

What:



(Trazzi, 2020)

Development approach

Who: Arend Hintze

Where: Understanding the four types of AI, from reactive robots to self-aware beings (article)

When: (2016)

What:

- Type I AI: Reactive machines
  The most basic types of AI systems are purely reactive, and have the ability neither to form memories nor to use past experiences to inform current decisions. Deep Blue, IBM's chess-playing supercomputer, which beat international grandmaster Garry Kasparov in the late 1990s, is the perfect example of this type of machine.
- Type II AI: Limited memory
  This Type II class contains machines can look into the past. Self-driving cars do some of this already. For example, they observe other cars' speed and direction. That can't be done in a just one moment, but rather requires identifying specific objects and monitoring them over time.
- Type III AI: Theory of mind
  Machines in the next, more advanced, class not only form representations about the world, but also about other agents or entities in the world. In psychology, this is called "theory of mind" – the understanding that people, creatures and objects in the world can have thoughts and emotions that affect their own behavior.
- Type IV AI: Self-awareness
  The final step of AI development is to build systems that can form representations about themselves. Ultimately, we AI researchers will have to not only understand consciousness, but build machines that have it.

Real-time AI system

Who: Rajendra Dodhiawala, N. S. Sridharan, Peter Raulefs & Cynthia Pickering

Where: Real-time AI systems: A definition and an architecture (Paper)

When: (1989)

What:

Four aspects of real-time performance to determine if an AI system is real-time or not.
The four aspects of real-time performance are:

- speed
- responsiveness
- timeliness, and
- graceful adaptation.

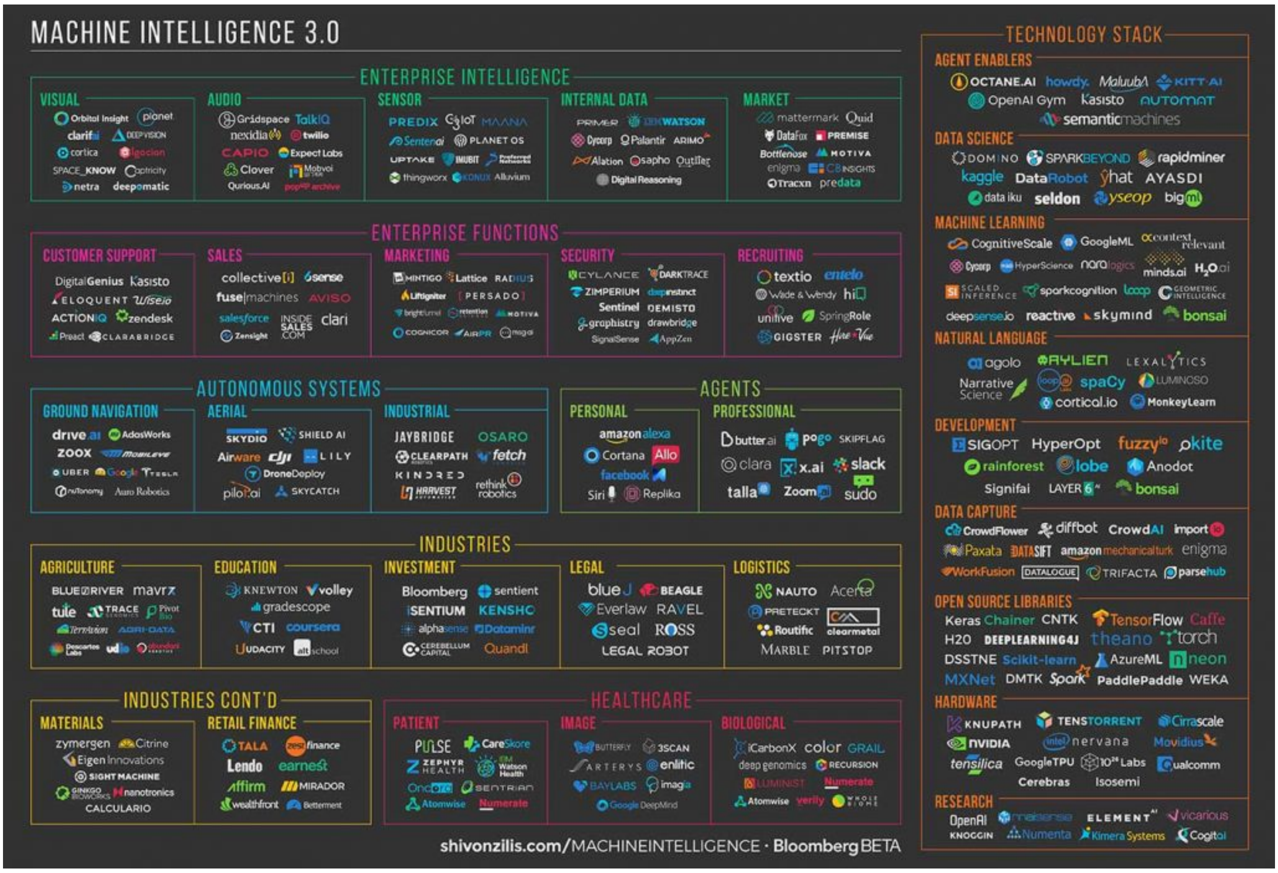
Application area approach

Who: Shivon Zilis & James Cham

Where: The current state of machine intelligence 3.0 (website)

When: (2016)

What:



Deep learning / capability approach

Who: Carlos E. Perez

Where: The Five Capability Levels of Deep Learning Intelligence (medium)

When: (2018)

What:

- Classification Only
  This level includes the fully connected neural network (FCN) and the convolution network (CNN) and various combinations of them. These system take a high dimensional vector as input and arrive at a single result, typically a classification of the input vector. You can consider these systems as being stateless functions, meaning that their behavior is only a function of the current input. Generative models are one of those hotly researched areas and these also belong to this category. In short, these systems are quite capable by themselves.
- Classification with Memory
  This level includes memory elements incorporated with the C level networks. LSTMs are example of these with the memory units are embedded inside the LSTM node. Other variants of these are the Neural Turing Machine (NMT) and the Differentiable Neural Computer (DNC) from DeepMind. These systems maintain state as they compute their behavior.

- Classification with Knowledge (DARPA's third wave)
  This level is somewhat similar to the CM level, however rather than raw memory, the information that the C level network is able to access is a symbolic knowledge base. There are actually three kinds of symbolic integration that I have found, a transfer learning approach, a top-down approach, a bottom up approach. The first approach uses a symbolic system that acts as a regularizer. The second approach has the symbolic elements at the top of the hierarchy that are composed at the bottom by neural representations. The last approach has it reversed, where a C level network is actually attached to a symbolic knowledge base.
- Classification with Imperfect Knowledge
  At this level, we have a system that is built on top of CK, however is able to reason with imperfect information. An example of this kind of system would be AlphaGo and Poker playing systems. AlphaGo however does not employ CK but rather CM level capability. Like AlphaGo, these kind of systems can train itself by running simulation of it against itself.
- Collaborative Classification with Imperfect Knowledge
  This level is very similar to the "theory of mind" where we actually have multiple agent neural networks combining to solve problems. Theses systems are designed to solve multiple objectives. We actually do se primitive versions of this in adversarial networks, that learn to perform generalization with competing discriminator and generative networks Expand that concept further into game-theoretic driven networks that are able to perform strategically and tactically solving multiple objectives and you have the making of these kind of extremely adaptive systems. We aren't at this level yet and there's still plenty of research to be done in the previous levels.

Application approach

Who:	Kai-Fu Lee
Where:	Venture capitalist and author Kai-Fu Lee on why artificial intelligence is moving from sci-fi to the mainstream (website)
When:	(2018)
What:	

The first stage is "Internet A.I." Powered by the huge amount of data flowing through the web, Internet A.I. leverages the fact that users automatically label data as we browse: buying vs. not buying, clicking vs. not clicking. These cascades of labeled data build a detailed profile of our personalities, habits, demands, and desires: the perfect recipe for more tailored content to keep us on a given platform, or to maximize revenue or profit.

The second wave is "business A.I." Here, algorithms can be trained on proprietary data sets ranging from customer purchases to machine maintenance records to complex business processes—and ultimately lead managers to improved decision-making. An algorithm, for example, might study many thousands of bank loans and repayment rates, and learn if one type of borrower is a hidden risk for default or, alternatively, a surprisingly good, but overlooked, lending prospect. Medical researchers, similarly, can use deep-learning algorithms to digest enormous quantities of data on patient diagnoses, genomic profiles, resultant therapies, and subsequent health outcomes and perhaps discover a worthy personalized treatment protocol that would have otherwise been missed. By scouting out hidden correlations that escape our linear cause-and-effect logic, business A.I. can outperform even the most veteran of experts.

The third wave of artificial intelligence—call it "perception A.I."— gets an upgrade with eyes, ears, and myriad other senses, collecting new data that was never before captured, and using it to create new applications. As sensors and smart devices proliferate through our homes and cities, we are on the verge of entering a trillion-sensor economy. This includes speech interfaces (from Alexa and Siri to future super¬smart assistants that remember everything for you) as well as computer-vision applications—from face recognition to manufacturing quality inspection.

The fourth wave is the most monumental but also the most difficult: “autonomous A.I.” Integrating all previous waves, autonomous A.I. gives machines the ability to sense and respond to the world around them, to move intuitively, and to manipulate objects as easily as a human can. Included in this wave are autonomous vehicles that can “see” the environment around them: recognizing patterns in the camera’s pixels (red octagons, for instance); figuring out what they correlate to (stop signs); and then using that information to make decisions (applying pressure to the brake in order to slowly stop the vehicle). In the area of robotics, such advanced A.I. algorithms will be applied to industrial applications (automated assembly lines and warehouses), commercial tasks (dishwashing and fruit-harvesting robots), and eventually consumer ones too.

Another third wave framework

Who:	Cris Doloc
Where:	Applications of Computational Intelligence in Data-Driven Trading (book)
When:	(2019)
What:	

The second AI wave started in the 1980s and it was reignited by two sources: an expansion of the algorithmic tool kit, and a substantial increase in funding. Alexey Ivakhnenko and David Rumelhart popularized Deep Learning techniques which allowed computers to learn using experience.

The third wave of AI (see Figure 3.2) was marked by events like the one when IBM’s *Deep Blue* computer system defeated the reigning world chess champion and

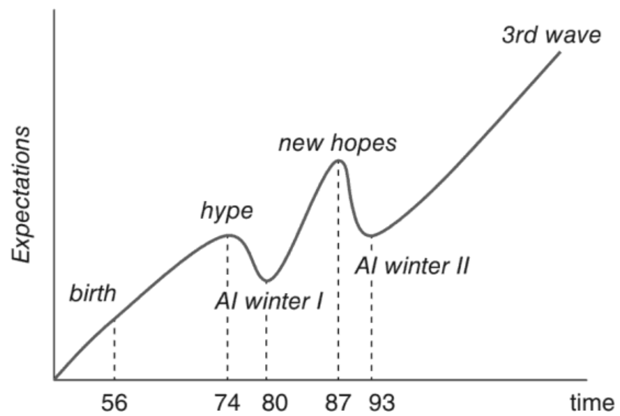


FIGURE 3.1 The AI cycles.

References

Dodhiawala, R., Sridharan, N. S., Raulefs, P., & Pickering, C. (1989). Real-time AI systems: A definition and an architecture. *In Proceedings of the International Joint Conference on Artificial Intelligence*, 256–261.

Doloc, C. (2019). *Applications of Computational Intelligence in Data-Driven Trading*. John Wiley & Sons.

Domingos, P. (2015). The master algorithm: How the quest for the ultimate learning machine will remake our world. *Choice Reviews Online*, *53*(07), 53-3100-53–3100. https://doi.org/10.5860/CHOICE.194685

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, *28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Hintze, A. (2016, November 14). *Understanding the four types of AI, from reactive robots to self-aware beings*. The Conversation. http://theconversation.com/understanding-the-four-types-of-ai-from-reactive-robots-to-self-aware-beings-67616

Lee, K.-F. (2018, October 22). *The Four Waves of A.I.* Fortune. https://fortune.com/2018/10/22/artificial-intelligence-ai-deep-learning-kai-fu-lee/

Perez, C. E. (2018, August 1). *The Five Capability Levels of Deep Learning Intelligence*. Medium. https://medium.com/intuitionmachine/five-levels-of-capability-of-deep-learning-ai-4ac1d4a9f2be

Russell, S., & Norvig, P. (2016). *Artificial Intelligence: A Modern Approach, Global Edition*. Pearson Education Limited. http://ebookcentral.proquest.com/lib/uzh/detail.action?docID=5186203

Trazzi, M. (2020, April 6). *The 5 Tribes of the ML world*. Medium. https://medium.com/42ai/the-5-tribes-of-the-ml-world-670ebce96b4c

Yampolskiy, R. V. (2013). Turing Test as a Defining Feature of AI-Completeness. In X.-S. Yang (Ed.), *Artificial Intelligence, Evolutionary Computing and Metaheuristics: In the Footsteps of Alan Turing* (pp. 3–17). Springer. https://doi.org/10.1007/978-3-642-29694-9_1

Zilis, S., & Cham, J. (2016, November 7). *The current state of machine intelligence 3.0*. O'Reilly Media. https://www.oreilly.com/content/the-current-state-of-machine-intelligence-3-0/

B

Interview Partners

Code	Name	Position	Company	Category	Language
INT1	Markus Christen	Managing Director	Digital Society Initiative	Industry expert	German
INT2	Pascal Hagedorn	Co-Innovation Architect	SAP	Industry expert	German
INT3	Gabriele Schwarz	Innovation Mentor	Innosuisse	Industry expert	German
INT4	Peter Voss	Founder/ CEO/ Chief Scientist	AGI Innovations & Aigo.ai	Third Wave AI expert	English
INT5	Gniewosz Leliwa	Director of AI Research & Co-founder	Samurai Labs	Third Wave AI expert	English
INT6	Sandra Völler & Thomas Neuhaus	CEO & Head of Digital Consulting	AGILITA AG	Industry expert	German
INT7	Steve Furber	ICL Professor of Computer Engineering	University of Manchester	Neuromorphic expert	English
INT8	Bharath Ramesh	Lecturer in Neuromorphic Engineering	Western Sydney University	Neuromorphic expert	English
INT9	-	Software Engineer	AI Startup	AI expert	German
INT10	Umberto Michelucci	Head of AI Center of Excellence	Helsana	AI expert	German

Table B.1: Interview partners and corresponding codes used in the master's thesis

C

Overview of AI Frameworks with Waves

	First Wave	**Second Wave**	**Third Wave**
(Szu et al., 2019)	Rule based systems	”learn-able Rule Based system” with supervised learning with labeled data ”from A to B”	Human Fuzzy Linguistic Thinking using Unsupervised Learning of Homo sapiens at a constant temperature T0
(Xu, 2019)	Early symbolism and connectionism school, production systems, knowledge inference, preliminary expert systems	Statistical model in speech recognition and machine translation, artificial neural network in pattern recognition, expert systems	Breakthroughs in applications of deep learning in speech recognition, pattern recognition, big data, high-performance computers: Technological enhancement and application + a human-centered approach
(Doloc, 2019)	Birth of AI	Extension of the algorithmic toolkit and substantial increase in funding (i.a. deep learning)	IBM’s Deep Blue
(Bai et al., 2019)	weak artificial intelligence	strong artificial intelligence	super artificial intelligence

Table C.1: Overview of different AI frameworks with three waves

	First Wave	**Second Wave**	**Third Wave**
(Toffler, 1980)	the agricultural revolution	the industrial revolution	the information age
(Carroll, 2020)	Machine automation	Combinatorial innovation and cross-over	AI, Robotics and IoT
(Paul R. Daugherty & Wilson, 2018)	ushered in by Henry Ford, involved standardized processes	automated processes	collaboration of human and machine (AI)

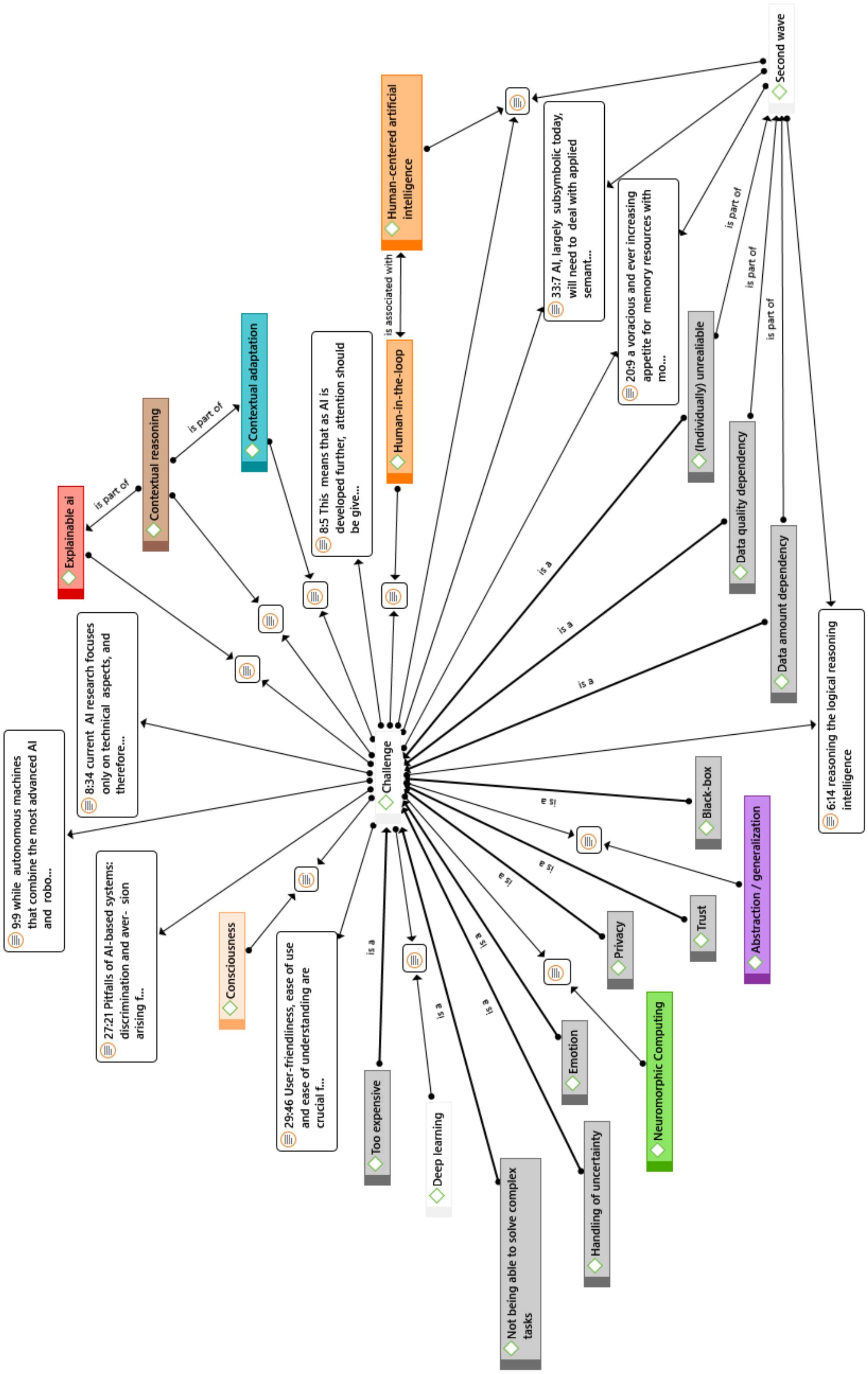
Table C.2: Overview of different frameworks with three waves

	First Wave	Second Wave	Third Wave	Fourth Wave
(Lee, 2018)	Internet A.I.	Business A.I.	Perception A.I.	Autonomous A.I.
(Scott, 2018)	• OGFAI - "good old fashioned ai" • Symbolic, heuristic, rule based • Handcrafted knowledge, "expert systems"	• Statistical learning, "deep" neural nets, CNNs, RNNs • Advanced text, speech, language and vision processing	• Contextual adaptation, able to explain decisions • Can converse in natural language • Requires far fewer data samples for training • Able to learn and function with minimal supervision	• AGI (artificial general intelligence), possibly leading to ASI (artificial superintelligence) and the "Technological Singularity."

Table C.3: Overview of different frameworks with four waves

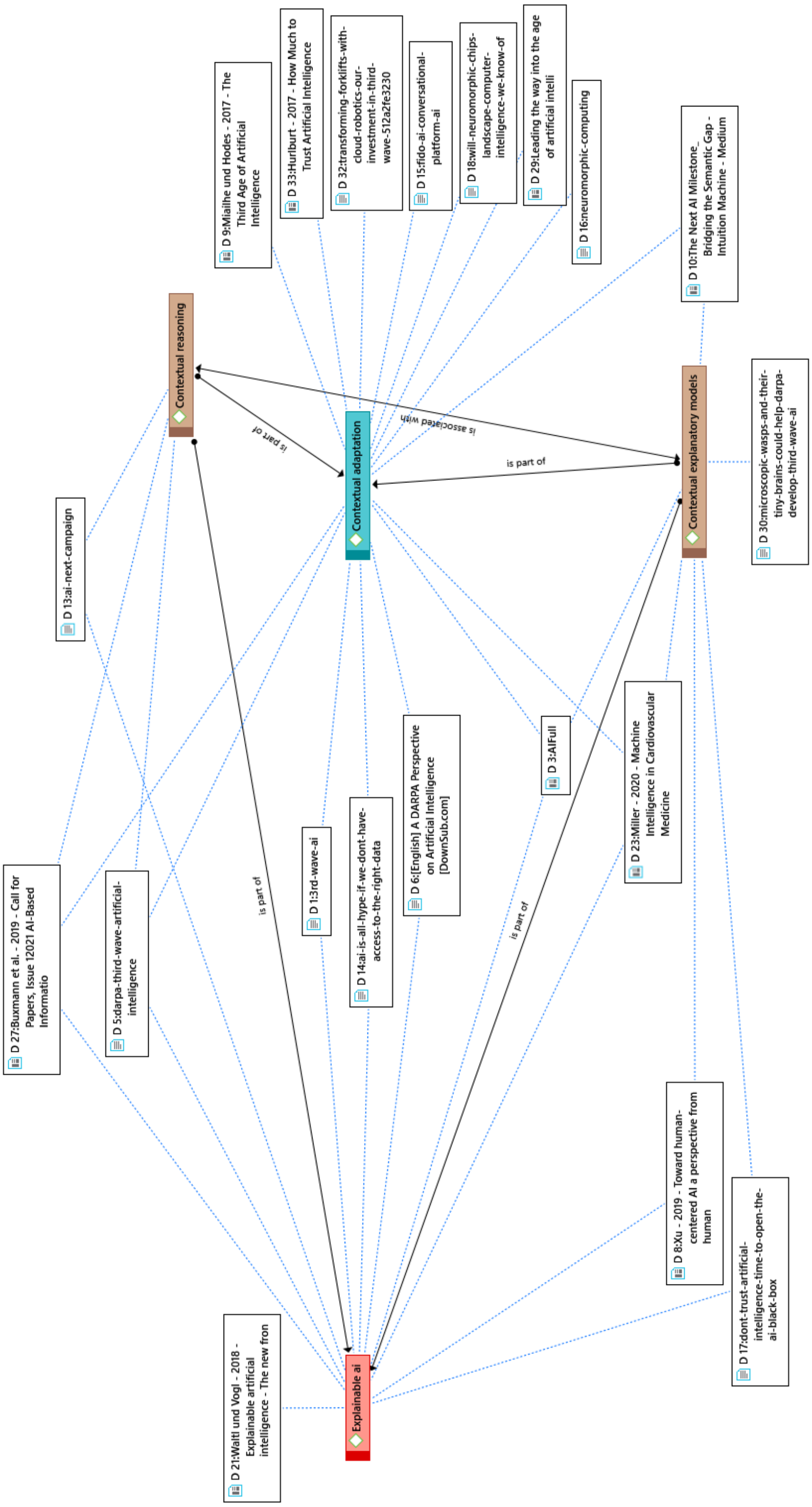
D

Challenges of AI


Second wave
Human-centered artificial intelligence
33:7 AI, largely subsymbolic today, will need to deal with applied semant...
20:9 a voracious and ever increasing appetite for memory resources with mo...
(Individually) unrealiable
Data quality dependency
Data amount dependency
6:14 reasoning the logical reasoning intelligence
Black-box
Abstraction / generalization
Trust
Privacy
Neuromorphic Computing
Emotion
Handling of uncertainty
Not being able to solve complex tasks
Deep learning
Too expensive
29:46 User-friendliness, ease of use and ease of understanding are crucial f...
Consciousness
27:21 Pitfalls of AI-based systems: discrimination and aver- sion arising f...
9:9 while autonomous machines that combine the most advanced AI and robo...
8:34 current AI research focuses only on technical aspects, and therefore...
Explainable ai
Contextual reasoning
Contextual adaptation
8:5 This means that as AI is developed further, attention should be give...
Human-in-the-loop
Challenge
is part of
is associated with
is a

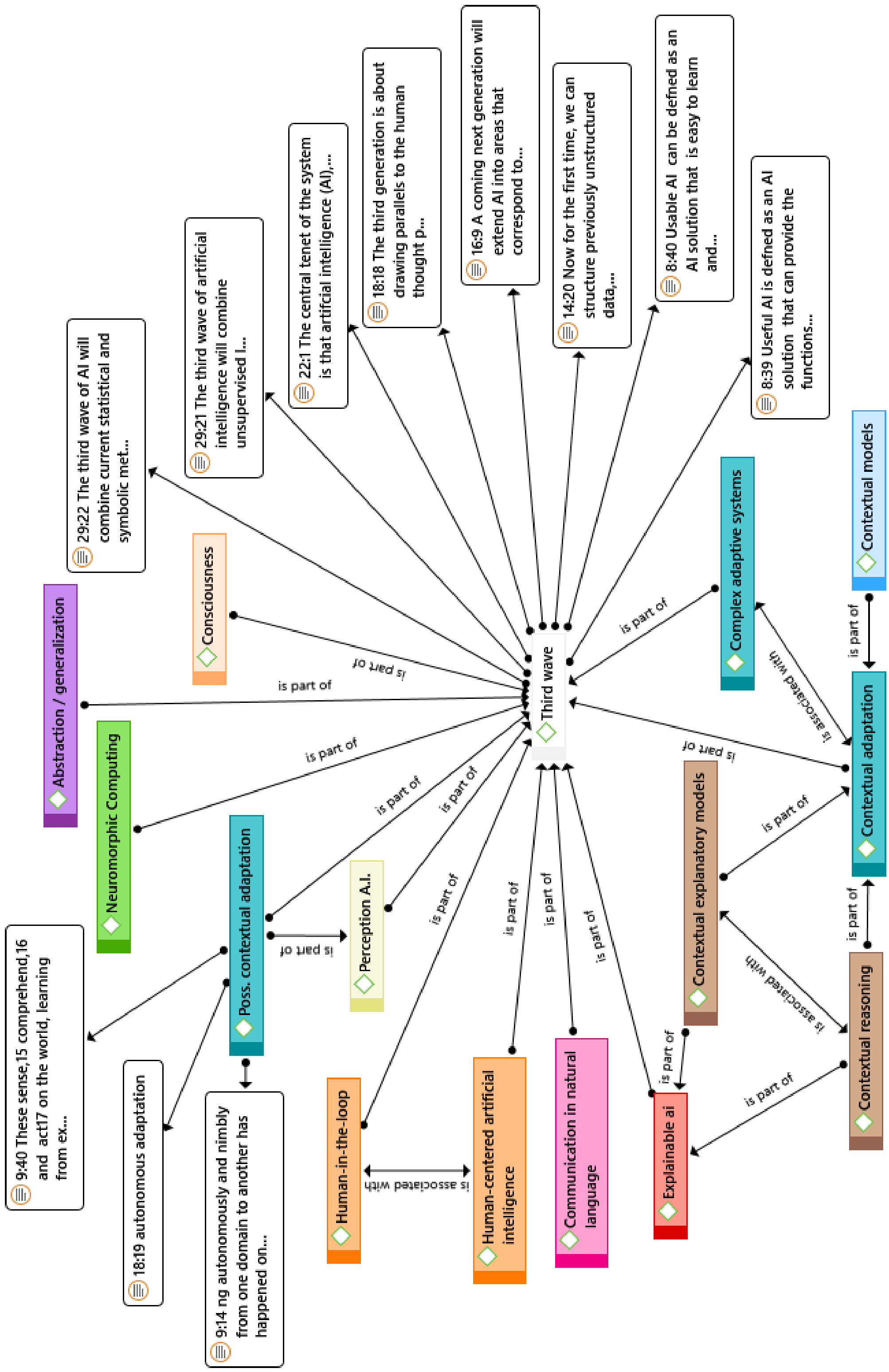
E

Intermixing Explainable AI and Contextual Adaptation


Contextual reasoning
Contextual adaptation
Contextual explanatory models
Explainable ai
is part of
is associated with
is part of
is part of
is part of
D 9:Mialhe und Hodes - 2017 - The Third Age of Artificial Intelligence
D 33:Hurlburt - 2017 - How Much to Trust Artificial Intelligence
D 32:transforming-forklifts-with-cloud-robotics-our-investment-in-third-wave-512a2fe3230
D 15:fido-ai-conversational-platform-ai
D 18:will-neuromorphic-chips-landscape-computer-intelligence-we-know-of
D 29:Leading the way into the age of artificial intelli
D 16:neuromorphic-computing
D 10:The Next AI Milestone_ Bridging the Semantic Gap - Intuition Machine - Medium
D 30:microscopic-wasps-and-their-tiny-brains-could-help-darpa-develop-third-wave-ai
D 13:ai-next-campaign
D 27:Buxmann et al. - 2019 - Call for Papers, Issue 12021 AI-Based Informatio
D 5:darpa-third-wave-artificial-intelligence
D 1:3rd-wave-ai
D 14:ai-is-all-hype-if-we-dont-have-access-to-the-right-data
D 6:[English] A DARPA Perspective on Artificial Intelligence [DownSub.com]
D 3:AIFull
D 23:Miller - 2020 - Machine Intelligence in Cardiovascular Medicine
D 8:Xu - 2019 - Toward human-centered AI a perspective from human
D 17:dont-trust-artificial-intelligence-time-to-open-the-ai-black-box
D 21:Waltl und Vogl - 2018 - Explainable artificial intelligence - The new fron

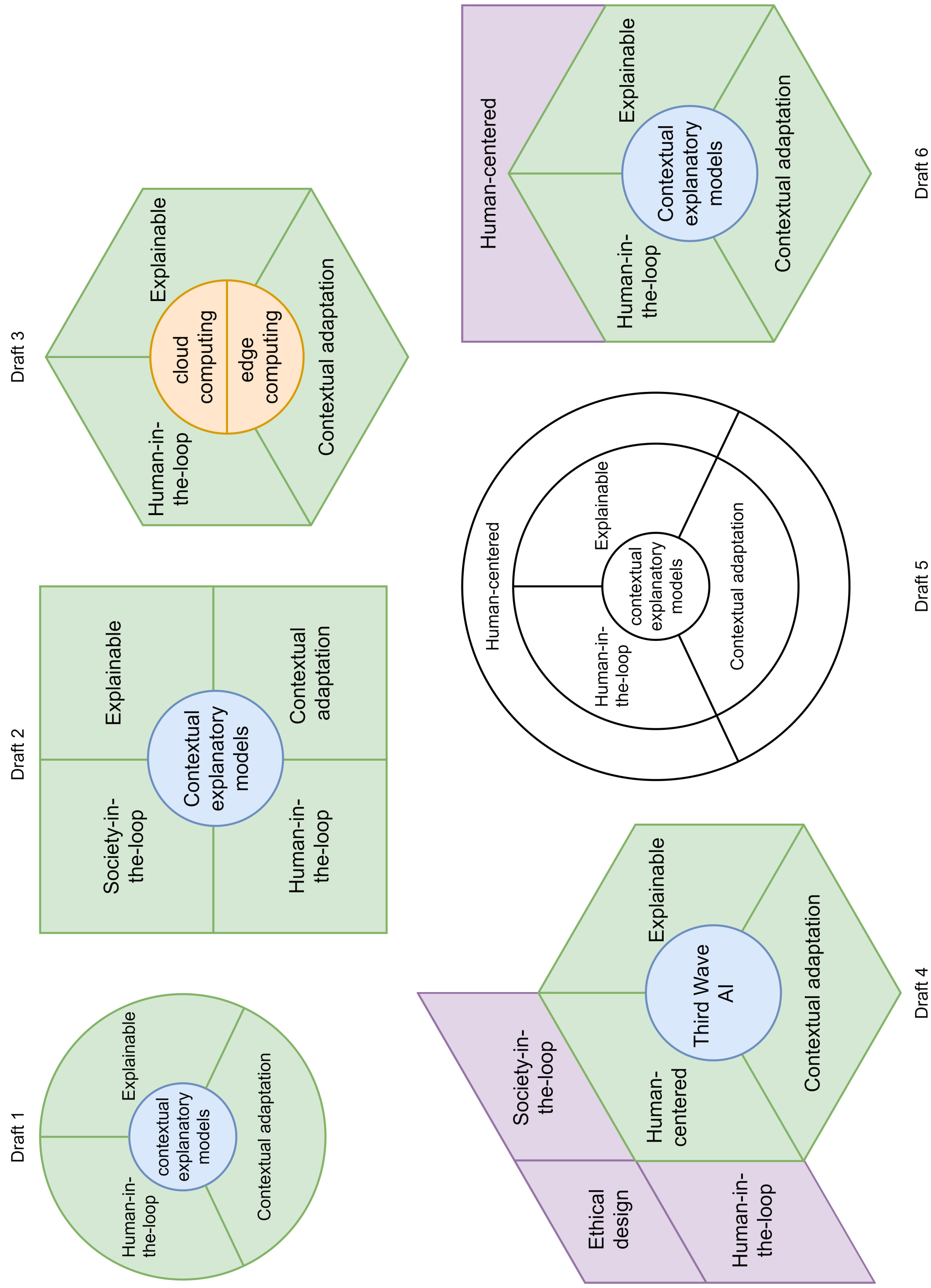
F

Characteristics of Third Wave


Third wave
29:22 The third wave of AI will combine current statistical and symbolic met...
29:21 The third wave of artificial intelligence will combine unsupervised l...
22:1 The central tenet of the system is that artificial intelligence (AI),...
18:18 The third generation is about drawing parallels to the human thought p...
16:9 A coming next generation will extend AI into areas that correspond to...
14:20 Now for the first time, we can structure previously unstructured data,...
8:40 Usable AI can be defned as an AI solution that is easy to learn and...
8:39 Useful AI is defned as an AI solution that can provide the functions...
Consciousness
Abstraction / generalization
Neuromorphic Computing
Poss. contextual adaptation
Perception A.I.
9:40 These sense,15 comprehend,16 and act17 on the world, learning from ex...
18:19 autonomous adaptation
9:14 ng autonomously and nimbly from one domain to another has happened on...
Human-in-the-loop
Human-centered artificial intelligence
Communication in natural language
Explainable ai
Contextual explanatory models
Contextual reasoning
Contextual adaptation
Complex adaptive systems
Contextual models
is part of
is associated with

G

Different Drafts of the Framework


Draft 1
contextual explanatory models
Explainable
Human-in-the-loop
Contextual adaptation
Draft 2
Contextual explanatory models
Explainable
Society-in-the-loop
Contextual adaptation
Human-in-the-loop
Draft 3
cloud computing
edge computing
Explainable
Human-in-the-loop
Contextual adaptation
Draft 4
Third Wave AI
Explainable
Human-centered
Contextual adaptation
Society-in-the-loop
Ethical design
Human-in-the-loop
Draft 5
contextual explanatory models
Explainable
Human-in-the-loop
Contextual adaptation
Human-centered
Draft 6
Contextual explanatory models
Explainable
Human-in-the-loop
Contextual adaptation
Human-centered

H

Example of an Interview Guide

Interview Guide

1. Administrative matters (3 min)

1.1. Thank you for your time

1.2. Remind that the session is recorded

1.3. Explaining how the content of the interview is used

1.4. Check if the interview can be used anonymous or named

1.5. Giving a short overview of the master thesis

2. Introduction (7 min)

2.1. What is your worst experience with AI?

2.1.1. How could this experience have been prevented?

2.1.2. Can this experience be used to identify the need for action in today's systems in general?

2.1.3. Which changes are necessary to use AI sustainably and successfully?

2.2. The reason why I want to do an interview with you, is that in an article about the third wave of AI they talked about neuromorphic computing and intel's new chips. They used project SpiNNaker as an example for the neuromorphic approach. And the term neuromorphic computing came up quite often during my research. So, I thought that it would be interesting to talk to experts in this research field. Also, they claim in the article that neuromorphic computing will be the next generation of AI.

2.3. Can you briefly explain to me SpiNNaker?

2.3.1. Do you think that SpiNNaker or neuromorphic computing in general will be the next generation of AI and why?

2.3.2. What are the challenges at the moment with neuromorphic computing?

2.4. How do you categorize AI?

2.4.1. e.g. machine learning, deep learning, rule-based system, statistical models, by algorithm, by application, hard or soft AI, supervised and unsupervised

2.4.2. What does the next generation of AI look like? Does one exist at all?

2.4.3. What paradigm shift is needed to revolutionize AI?

3. Third Wave (10 min)

3.1. What do you know about Third Wave AI?

3.2. Explanation Third Wave (slide)

3.3. What is your first impression of Third Wave?

3.3.1. Do you think that neuromorphic computing will lead to these contextual exploratory models?

3.3.2. Will neuromorphic computing be the next wave and does neuromorphic computing fulfill DARPA's requirements?

3.3.3. Do you agree with the definition of the first two waves?

3.3.4. What do you understand under contextual adaptation?

3.3.5. Where do you see the problems with this framework?

3.3.6. What should be changed?

3.4. Do you know a Third Wave application?

3.4.1. What makes this application a Third Wave application?

4. New framework (35 min)

4.1. Explanation framework

4.2. What is your first impression of this framework?

4.2.1. What is your opinion about the new layer? Are they comprehensible? Are the adjustments necessary?

4.2.2. Do you see the focus in the functional or technological layer and why?

4.2.3. Are there other layers or dimensions that could be essential for the next wave? E.g. social, societal, hardware, ethical

4.2.4. Would neuromorphic computing or SpiNNaker fulfil this framework?

4.2.5. Where do you see the problems with this framework?

4.2.6. What should be changed? From your point of view is something missing or is something should be removed?

4.2.7. Where do you see the potential of this framework?

4.3. How would you use the framework in concrete examples?

4.3.1. e.g. explaining AI, for the evaluation of start-ups, designing new AI solutions, brainstorming

4.3.2. If not applicable How does the framework have to be adapted in order to be able to use it?

4.3.3. If applicable: Why this example?

4.3.4. if applicable: Why would using the framework in this example provide more value?

4.3.5. Would you use the framework only in theory or also in practice? Why?

4.3.6. Which adjustments are necessary for a (better) usage?

4.3.7. To whom would you recommend this framework and for what propose?

4.4. Do you know a software, solution or product that can be classified as Third Wave according to this framework?

  4.4.1. If not: Why do you think there are no applications that meet the criteria?

  4.4.2. If no: Do you know an application that will revolutionize AI for other reasons? And what are the reasons?

  4.4.3. If no: Which applications come close to the category Third Wave?

  4.4.4. If yes: Why did you choose this application?

  4.4.5. If yes: How does this application differ from other AI applications?

  4.4.6. Which framework did you use for classification and why?

  4.4.7. How will applications that meet the framework be a revolution?

  4.4.8. Do you have an idea for an application that meet the framework?

5. Conclusion (5 min)

5.1. Will you use the term Third Wave AI in the future and if so, based on which framework?

5.2. Could you define your Third Wave in two / three sentences?

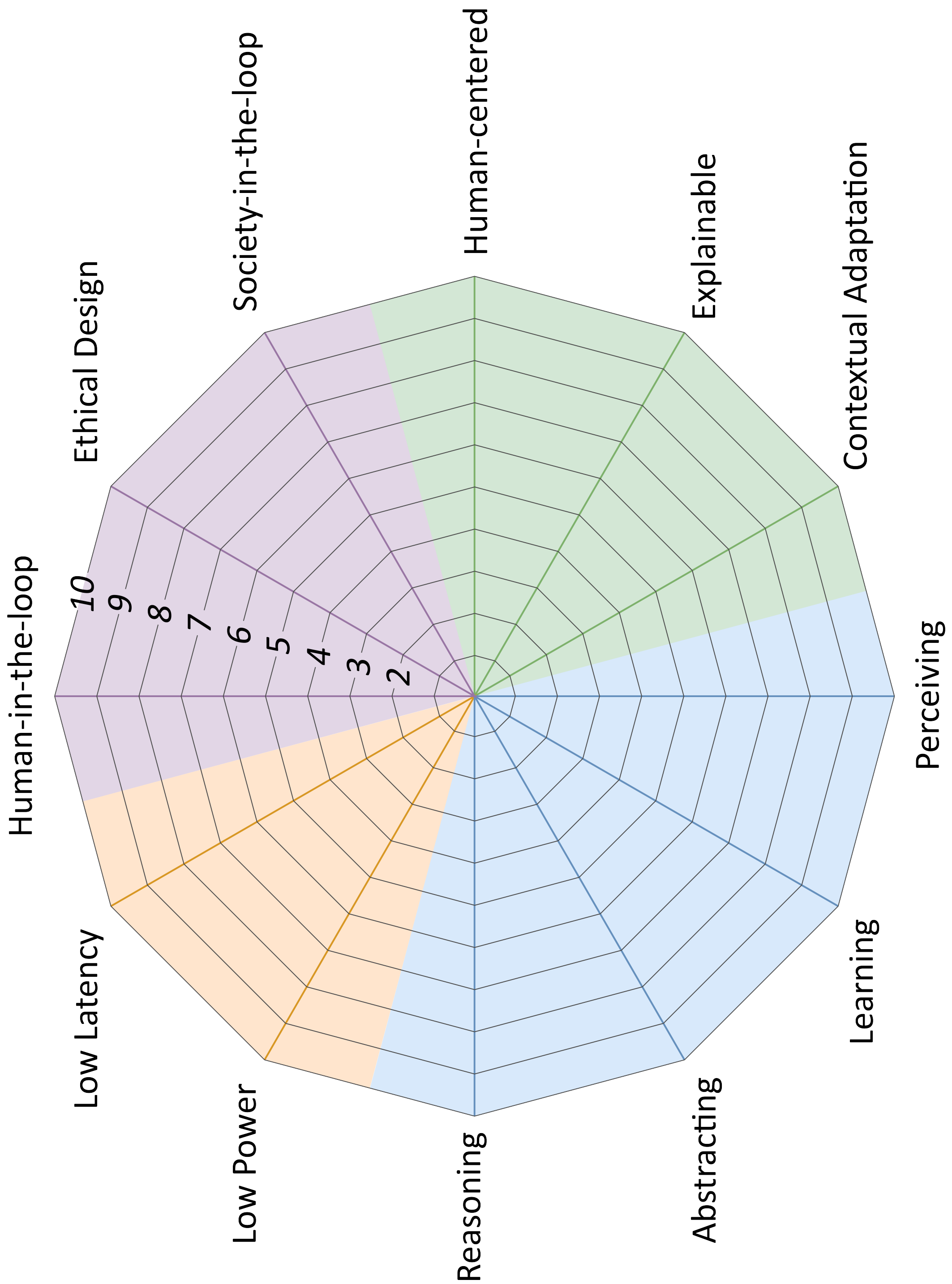
5.3. Would your bad experience from the beginning have been prevented by using the framework?

5.4. Was there something you expected but I didn't ask?

5.5. Expressing thankfulness

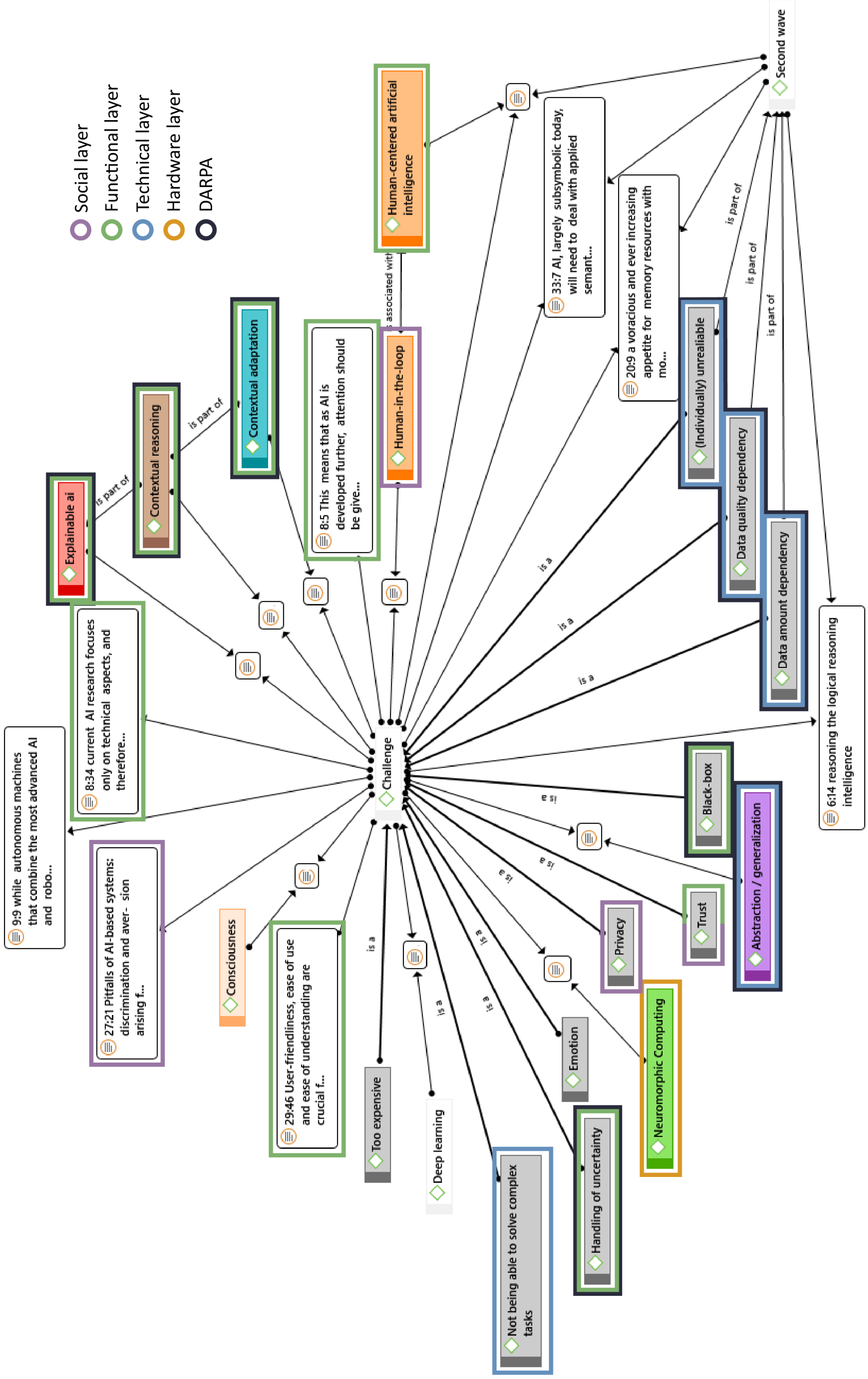
I

Spider Diagram Template


Human-centered
Explainable
Contextual Adaptation
Perceiving
Learning
Abstracting
Reasoning
Low Power
Low Latency
Human-in-the-loop
Ethical Design
Society-in-the-loop
10
9
8
7
6
5
4
3
2

J

Addressed Challenges of AI



Social layer
Functional layer
Technical layer
Hardware layer
DARPA
Second wave
33:7 AI, largely subsymbolic today, will need to deal with applied semant...
20:9 a voracious and ever increasing appetite for memory resources with mo...
Human-centered artificial intelligence
is associated with
Human-in-the-loop
8:5 This means that as AI is developed further, attention should be give...
Contextual adaptation
Contextual reasoning
Explainable ai
is part of
is part of
is part of
is part of
is part of
(Individually) unrealiable
Data quality dependency
Data amount dependency
is a
is a
is a
8:34 current AI research focuses only on technical aspects, and therefore...
9:9 while autonomous machines that combine the most advanced AI and robo...
Challenge
6:14 reasoning the logical reasoning intelligence
Black-box
Abstraction / generalization
Trust
Privacy
27:21 Pitfalls of AI-based systems: discrimination and aver- sion arising f...
Consciousness
29:46 User-friendliness, ease of use and ease of understanding are crucial f...
Too expensive
Deep learning
Emotion
Neuromorphic Computing
Handling of uncertainty
Not being able to solve complex tasks
is a