Samu Kumpulainen

Artificial General Intelligence - a systematic mapping study

Master's Thesis in Information Technology

April 20, 2021

University of Jyväskylä

Faculty of Information Technology

Author: Samu Kumpulainen

Contact information: `samu.p.kumpulainen@student.jyu.fi`

Supervisors: Vagan Terziyan, and Kari Kärkkäinen

Title: Artificial General Intelligence - a systematic mapping study

Työn nimi: Yleinen tekoäly - systemaattinen kirjallisuuskartoitus

Project: Master's Thesis

Study line: Mathematical Information Technology

Page count: 47+9

Abstract: The research field of artificial general intelligence is growing more popular in recent years, but it is complex and fragmented, and therefore difficult to enter for the new researchers. In this thesis, a systematic mapping study was conducted on the field of artificial general intelligence. The goal of the study was to gain insight about the recent developments in the study field, and achieve an overview of the area. In the study there were 92 accepted articles from years 2015-2019 from five different publication forums. Key findings show small but steady amount of researched being published yearly, with focus on novel solution proposals and philosophical papers. Most popular research topics are cognitive architectures, theory of universal AI, AI safety and ethics, and different types of learning methods. Most of the AGI research is conducted in European countries but USA is, by considerable margin, the most active country in the research.

Keywords: AGI, AI, artificial intelligence, artificial general intelligence, systematic literature mapping, mapping study

Suomenkielinen tiivistelmä: Yleisen tekoälyn tutkimuskenttä on viime vuosina kasvattanut kiinnostustaan, mutta aihealueen monimutkaisuuden ja jakautuneisuuden vuoksi uusien tutkijoiden voi olla hankala päästä siihen sisään. Tässä tutkielmassa suoritettiin systemaattinen kirjallisuuskartoitus yleisen tekoälyn tutkimuksesta. Tavoitteena oli saada selville tutkimuskentän viimeaikaiset ilmiöt, sekä luoda yleiskuva nykyisestä tutkimuksesta. Tutkielmassa käytiin

läpi 92 artikkelia viidestä eri julkaisukanavasta vuosilta 2015-2019. Tuloksista nähdään, että tutkimuksen määrä on pieni, mutta uusia tutkimuksia julkaistaan tasaiseen tahtiin vuosittain. Tutkimus koostuu pääosin uusista ratkaisuehdotuksista, sekä filosofisista artikkeleista. Suosituimpia tutkimusaiheita ovat kognitiiviset arkkitehtuurit, tekoälyn laskettavuuden teoria, tekoälyn turvallisuutta ja etiikkaa koskevat kysymykset, sekä erilaiset oppimismenetelmät. Suurin osa yleisen tekoälyn tutkimuksesta tulee Euroopan maista, mutta julkaisumäärältään suurin yksittäinen maa on Yhdysvallat.

Avainsanat: tekoäly, yleinen tekoäly, kirjallisuuskartoitus

List of Figures

List of Tables

Contents

1 Introduction

This thesis is a systematic research mapping on the field of Artificial General Intelligence (AGI). The goal of the thesis is to identify the themes and topics researched in the AGI field in recent years and discover the types of research gaps that exist in the field. While developing a system that displays general, human-like intelligence was the original goal of artificial intelligence research, it has not been a very popular approach to research AI in mainstream research segment since the 1980s. Instead, the more contextually targeted intelligent solutions, known as 'narrow AI', have grown in popularity. However, recently the wider and more general approach to artificial intelligence has been regaining interest. This kind of systematic mapping study is needed as the research field is complex and there is no clear existing presentation of the current trends and focal points. Creating this kind of overview is a valuable asset for future research, as it enables focusing the research on less ventured areas. It can also be useful in introducing the study field to new researchers.

This thesis is structured as follows: chapter 2 introduces the Artificial General Intelligence, focusing on the history of AI and its definition. Chapter 3 describes the research method of this thesis, systematic literature mapping. In chapter 4, the conducted mapping process is reported, with the results being presented in chapter 5. Chapter 6 summarizes and concludes the thesis.

2 Artificial General Intelligence

In this chapter, the history of artificial intelligence (AI) is shortly described as an introduction to the subject. Then, the definitions of intelligence, and building on that, Artificial General Intelligence, are introduced.

2.1 History of Artificial Intelligence

Even though the idea of autonomous machinery has been around since ancient Greece (Bramer 2009), AI originated in the 1940s. At the time, American science fiction author Isaac Asimov wrote numerous novels and short stories about conscious robots and technology's relation to humankind. His work has inspired countless people in the fields of AI and computer science (Haenlein and Kaplan July 2019). Also in the 1940s, mathematician Alan Turing's work on Britain's code breaking efforts lead to the creation of first electromechanical computer, The Bombe (Haenlein and Kaplan July 2019). Turing later gave lectures and wrote an article titled *"Computing Machinery and Intelligence"* (1950), in which he presented several ideas later prevalent in the AI field, including the "imitation game", a test to measure the intelligence of a machine (Russell and Norvig 2009). This later became known as the Turing test.

The term Artificial Intelligence was coined in 1956 during a two-month workshop *Darthmouth Summer Research Project on Artificial Intelligence*, organized by John McCarthy and Marvin Minsky (Haenlein and Kaplan July 2019). The participants of the workshop would later become the most prominent figures of AI research. During DSRPAI two researchers, Allen Newell and Herbert Simon, presented Logic Theorist, their existing reasoning program, capable of proving multiple mathematical theorems (Russell and Norvig 2009). Based on this work the two later created General Problem Solver, GPS, which could solve simple puzzles like Towers of Hanoi using human like recursive approaches (Newell, Shaw, and Simon 1959). The early days of AI research produced many similar results in different areas. IBM's Arthur Samuel created AI programs that learned to play checkers at a strong amateur level (Russell and Norvig 2009). John McCarthy's 1958 paper titled "Programs with

common sense", describes Advice Taker, a complete but hypothetical AI system with general knowledge about the world and deductive processes to manipulate it. The paper is still thought to be relevant today. McCarthy's system was able to acquire new skills in previously unknown areas without being reprogrammed.

During these years also work on the neural networks started to gain interest. The initial work of McCulloch and Pitts (1943), later demonstrated by Hebb (Russell and Norvig 2009), showed that a neural network is capable of learning. In the 1960s Rosenblatt's work on perceptrons and Widrow and Hoff's LMS algorithm were some of the biggest advances in the area (Widrow and Lehr 1995). The next great discovery that would propel the neural networks into the focal point of AI research happened in the mid-1980s when the backpropagation algorithm originally presented by Bryson and Ho in 1969 was rediscovered by multiple independent groups (Russell and Norvig 2009). Backpropagation is one of the most widely used algorithms for training neural networks these days for its relative power and simplicity (Rumelhart et al. 1995).

History of artificial intelligence contains occasional periods of reduced interest and funding. These so called "AI winters" are a result of high expectations collapsing under criticism. The first period that can be considered an AI winter started in the 1970s, and Russell and Norvig (2009) present the following possible reasons for it: Firstly, the early programs knew nothing about their context, and solved the problems via syntactic manipulations. This was especially apparent on machine translation projects. As a language cannot be fully understood without knowing the full context of the sentences and other nuances of the language, accurate translation proved to be a difficult task. Failed translation efforts lead to funding cuts in the US. The second difficulty pointed out by (Russell and Norvig 2009) was the sheer complexity of the target problems. As the early AI programs were focused on simple tasks, finding a solution by trial and error was possible in practice. As the problems became more complex, "combinatorial explosion" issue became more apparent. The issue was also discussed in British scientist James Lighthill's report on the state of the AI (1973). The report is considered to be one of the main reasons why the British government decided to cut AI funding in all but two universities. Lastly, the limitations of the data structures used in AI fields, such as perceptrons, restricted the capabilities of the solutions. According to Russell

and Norvig (2009) this led to funding cuts also in the neural network research.

During and after the first AI winter, there was a considerable amount of research relating to expert systems (Russell and Norvig 2009). These systems perform their tasks in a way similar to human experts in the specific, narrow domain, relying on a knowledge encoded into a set of rules (Myers 1986). This style of AI research was inspired by the success of DENDRAL (Buchanan, Feigenbaum, and Lederberg 1968), a system developed at Stanford by Ed Feigenbaum, Bruce Buchanan and Joshua Lederberg. DENDRAL's purpose was to use data from mass spectrometer to infer the structure of a given molecule. MYCIN, developed in the 1970s (Shortliffe et al. 1975), incorporated domain knowledge acquired through expert interviews, with the uncertainty of medical evaluation taken into account via certainty factors (Russell and Norvig 2009).

Expert systems gained commercial interest, leading to increased research and adoption in the industry. Government investments in Japan led to increased funding in United States and Britain, leading to an AI boom in the 1980s (Russell and Norvig 2009). After the boom, at the end of the 1980s, the second AI winter arrived. Participation in AI conferences dropped, several of the new AI companies met their end, as did the AI research divisions in larger hardware and software companies (Nilsson 2009). The imminent burst of the bubble was foreseen by several leading researchers, but their warnings did not have considerable effect (Nilsson 2009).

According to Russell and Norvig (2009), around this time the AI field started to adopt the scientific method. This means the earlier ways of proposing completely new theories based on vague evidence or oversimplified examples had been replaced by basis on existing theories, repeatable experiments, and real-world examples. This newly discovered open-mindedness then led to a completely new ways of looking at the AI research. AI solutions based on existing theories, such as speech recognition based on hidden Markov models, enables the researchers to build on the rigorous mathematical theory behind it (Russell and Norvig 2009). Work of Judea Pearl (1988) and Peter Cheeceman (1985) on the probabilistic reasoning led to it being accepted back into the AI field. Later Pearl's Bayesian networks have been used to handle uncertainty in AI problems. They are graphical models that join probabilistic information and dependencies to events, enabling inference using probabilistic methods (Goertzel

and Pennachin 2007).

In the 21st century, artificial intelligence research has been steadily growing. According to (Liu et al. 2018), not only has the amount of publications in the field been increasing, but also the collaboration between researchers. The study also deduces that that AI has become more open-minded and popular, as the rate of self-references is reducing. One reason for the rising popularity on the field is the success that narrow AI solutions have presented in multitude of problems. For example, in the classical game of Go, a program called AlphaGo developed by Google-owned DeepMind, defeated the world champion Lee Sedol in 2015 (Silver et al. 2016). Due to Go's computationally complex nature this was a impressive feat which was previously thought to be impossible. Later DeepMind developed even more advanced versions of AlphaGo, called AlphaGo Zero, and generalized AlphaZero, which could even play Shogi and Chess on superhuman level (Silver et al. 2018).

In recent years, a majority of the field has been focusing on the narrow AI approaches (Goertzel and Pennachin 2007). However, the interest in the classical, strong AI has also been increasing. This can be seen in the publications from many influential AI researchers. Authors like John McCarthy (2007), Nils Nilsson (2005) and Marvin Minsky (2007) have voiced their opinions to apply a higher level of effort to pursue a more general AI solution. There are several terms used regarding these efforts. **Human-level Artificial Intelligence** (HLAI) aims to reach "human-level intelligence" and common sense, a goal that according to Marvin Minsky (2004) can be reached by not using any single method, but a combination of different resources and methods. Similar term is **Artificial General Intelligence** (AGI) presented by Ben Goertzel and Casio Pennachin (2007). The goal of AGI is similar to HLAI, to create an AI system that can express general intelligence instead of being locked into a single domain. On the next chapter this general approach is presented in more detail, as it is the focus of this thesis.

2.2 Definition

In order to define AGI, or artificial intelligence in general, one must first consider the definitions of intelligence in general. There exists several different definitions, in many different

branches of science. Legg and Hutter (2007) list over 60 definitions collected from various academic sources. These include, for example, *"the general mental ability involved in calculating, reasoning, perceiving relationships and analogies, learning quickly, storing and retrieving information, using language fluently, classifying, generalizing, and adjusting to new situations."* (Columbia Encyclopedia, sixth edition, 2006), *"that facet of mind underlying our capacity to think, to solve novel problems, to reason and to have knowledge of the world"* (Anderson 2006), and *"Intelligence is the ability for an information processing system to adapt to its environment with insufficient knowledge and resources."* (Wang 1995). Based on the aforementioned collection of definitions, Legg and Hutter (2007) have formed the following definition: *"Intelligence measures an agent's ability to achieve goals in a wide range of environments"*. This gives us a single definition which encompasses the common traits in intelligence definitions.

Artificial General Intelligence, sometimes referred as "strong AI", according to Goertzel and Pennachin (2007) means *"AI systems that possess a reasonable degree of self-understanding and autonomous self-control, and have the ability to solve a variety of complex problems in a variety of contexts, and to learn to solve new problems that they didn't know about at the time of their creation."*. It can be seen that an agent fulfilling this definition also possesses the intelligence defined in the previous chapter. In this thesis terms artificial general intelligence and human-level artificial intelligence are treated as synonyms, as they pursue more or less the same goal of general intelligence. Goertzel and Pennachin (2007) suggest that the term AGI is more fitting to the area than HLAI as human-like approaches are not necessarily used.

The reason general intelligence is specified instead of plain intelligence is that there is a need to differentiate it from the domain specific artificial intelligence, also known as "narrow AI" or "weak AI", that has become prevalent in AI research in the recent past. Terms weak AI and strong AI were coined by John Searle in 1980 (Searle et al. 1980). Searle makes a difference with AI being a powerful tool, and being an actually intelligent system, respectively. Narrow AI means smart solutions that may learn and improve their performance through training, but they are only focused on specific types of problems in specific contexts. Examples of such AI§ include chess engines, autonomous vehicles, and natural language processing. These solutions may outperform human capabilities, but only in their limited tasks. When presented

a problem outside their domain, they usually perform poorly. The strong AI instead is able to perform well in every type of scenario, displaying intelligence and understanding. However, this does not necessarily mean that the AI would possess what humans would consider as consciousness (Searle et al. 1980). As the above definition by Goertzel and Pennachin describes, AGI is able to function on different context and tasks without separate human intervention and reconfiguration.

As the AGI community is diverse and there are a multitude of opinions on the best approaches and the goals that should be pursued in the research, several possible roadmaps have been presented in an attempt to create a common basis for the discussion and research of human-level artificial general intelligence. In the paper *Mapping the Landscape of Human-Level Artificial General Intelligence* (Adams et al. March 2012) a high level roadmap with AGI's initial required capabilities and scenario-based milestones is suggested, building on previous work and workshops organized in 2008 and 2009. Presented scenarios can be used to measure the progress and capabilities of AGI restricting the progress of different approaches to a single test situation (Adams et al. March 2012). A more concrete example is provided by Ben Goertzel and Gino Yu, who outline creation of an AGI-oriented cognitive architecture based on existing CogPrime architecture (Goertzel and Yu 2014). A simultaneous development of multiple AGI-style applications is suggested to maintain the generality of intelligence. Cog-Prime is implemented with the OpenCog framework, developed by OpenCog Foundation and AI researcher Ben Goertzel. OpenCog is an attempt to create an open source framework for artificial general intelligence (*The Open Cognition Project;* B. Goertzel 2012).

One motivation behind this thesis is to find out if these roadmaps and the approaches presented in them have actually had any effect in the direction that AGI research has gone, or have they been ineffective in their attempt to organize the complex path to AGI.

3   Systematic literature mapping process

Systematic literature mapping is a secondary study method that helps to identify the focal points and research gaps in the subject area, providing an overview of previous research (Petersen et al. 2008). It is also known as a systematic mapping study or scoping study (Kitchenham and Charters 2007). This chapter introduces the mapping method, describing each phase of the mapping process. The key differences with a more common study method, systematic literature review (SLR) are also presented, as well as the reasoning behind this method choice.

3.1   Research method description

Systematic mapping is a common research method used in fields such as evidence-based medicine, but have until recently been rare in software engineering (Petersen et al. 2008). Kitchenham, Dybå, and Jorgensen (2004) suggested that adopting an evidence-based approach to software engineering research might benefit the industry by ensuring approaches used are backed by evidence. Aggregating evidence is done by systematic literature reviews and similar secondary studies, such as mapping studies (Kitchenham et al. 2010). Although the study field of artificial intelligence is not directly part of software engineering research, this evidence-based approach can also be utilized in AI research. Some example articles that employ mapping studies in AI are: the study on possibilities of AI and big data analytics in healthcare by Mehta, Pandit and Shukla (2019), which even uses Petersen's aforementioned guidelines, and the study by Wiafe et al. (2020), which studies AI approaches in cybersecurity.

The systematic literature mapping in this thesis is following the method guidelines presented by Petersen et al (2008), later updated by Petersen, Vakkalanka and Kuzniarz (2015). The mapping process overview can be seen in Figure 1. It consists of five separate phases: **definition of research questions, conducting a search, screening of papers, keywording, and data extraction and mapping.** Each phase produces a subresult to be used in the next one. This process results in a systematic map of the area. This can and should be further visualized

using, for example, bubble graphs as it is a powerful way to achieve a quick overview of the field. (Petersen et al. 2008). This also enables easier recognition of research gaps and focus points in the target area. As researcher bias is one of the weak points of secondary studies, adhering to strict protocol and guidelines is required to minimize it (Brereton et al. 2007).

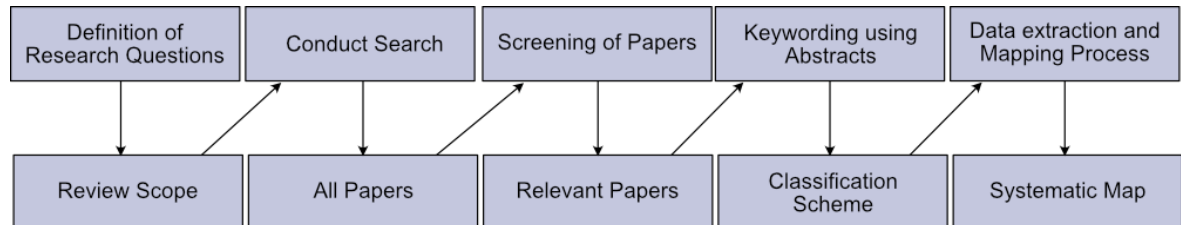| Definition of Research Questions	Conduct Search	Screening of Papers	Keywording using Abstracts	Data extraction and Mapping Process
Review Scope	All Papers	Relevant Papers	Classification Scheme	Systematic Map

Figure 1. Process model (Petersen et al. 2008)

The process begins by defining focused research questions that are aligned with the goal of the study. The goal of the study is often to create a general overview of the research area, and to identify the type and quantity of research (Petersen et al. 2008). Unlike in more focused systematic literature reviews, the research questions of mapping studies are less focused and cover a broader scope (Kitchenham et al. 2010). For example, possible research questions on studies could be: *"Which are the most investigated quality aspects of the software requirements specifications techniques?"* (Condori-Fernandez et al. 2009) or *"What efforts have been reported in literature for Software Engineering curriculum design, implementation, assessment and management?"* (Qadir and Usman 2011). These topic-oriented questions are often combined with research questions regarding the meta-level information, such as publication year, venue, and research methods (Petersen, Vakkalanka, and Kuzniarz 2015).

The next phase of the mapping is the initial material search, which can be conducted in multiple ways. Search strings can be formulated from the research questions, and used on academic databases and search engines (Petersen et al. 2008). For example, databases such as IEEE Explore and ACM, as well as aggregators like Google Scholar can be utilized. As the search phrases should be research question driven, following a criteria such as PICO might prove to be helpful, as suggested by Kitchenham and Charters (2007). PICO (Population, Intervention, Comparison, Outcomes) provides a frame to consider research questions'

elements and identify keywords. As the goal is to achieve a broad overview of the research, study outcomes are not taken into account, as this could result in biased results (Petersen et al. 2008). The search can also be conducted manually on specific journals and conference proceedings that cover the target area (Petersen et al. 2008). This approach is used in this thesis, as it enables targeting specific reputable and well-known publication venues.

After the initial material has been gathered, it is further refined by excluding papers not relevant to answering the research questions (Petersen et al. 2008). Separate criteria for both inclusion and exclusion is used to find the papers fit for further analysis. According to Petersen, Vakkalanka and Kuzniarz (2015), the criteria may refer to relevance to the topic, publication venue, time period, language restrictions, and evaluation requirements. Considering evaluation requirements should be avoided with systematic maps as it might limit recent trends. Once the criteria is decided, it is applied to the titles and abstracts of the articles (Petersen, Vakkalanka, and Kuzniarz 2015). In case of unclear or poor quality abstracts, the introduction and conclusion sections of the article may be studied (Petersen et al. 2008). As no full-text reading is required, screening of papers can be performed rapidly.

Once the final set of papers is narrowed down and determined, keywording is performed. As described by Peterson (2008), The keywording process starts first by analyzing the research papers' abstracts by searching possible frequent keywords and prevalent concepts from them. The keywords of each paper are then combined together to achieve a more general set of concepts. In some cases having a more detailed inspection of the article might be required (Petersen et al. 2008; Petersen, Vakkalanka, and Kuzniarz 2015). After the final set of keywords is chosen, they are clustered into categories representing the article population (Petersen et al. 2008). This emergent classification schema can then be used in the data extraction phase.

Different research facets can be used in the classification. For example in addition to the topical scheme emerging from the keywords, a topic-independent facet reflecting the research approach can be used. One example of the latter is the classification of research approaches by Wieringa et. al (2006). Wieringa's classification categorizes scientific papers into six categories such as validation research, solution proposals, and opinion papers. Using existing topic-independent categorization also enables the comparison of different research fields

(Petersen, Vakkalanka, and Kuzniarz 2015).

In the final phase of the mapping process, the data extraction is performed by sorting the papers into the classification schemes present (Petersen et al. 2008). The emergent schema may evolve during the data extraction process, changing the categories to match the article population more accurately (Petersen, Vakkalanka, and Kuzniarz 2015). The categorization based on the chosen facets results in a frequency table, i.e. the mapping, which can then be presented with via visualization and summary statistics. Visualization using for example bubble plots is preferred, as it is a powerful way to represent the information and map of the field (Petersen et al. 2008).

3.2 Differences with systematic literature review

Because the systematic literature mapping as a study methods is less known in the field of software engineering, the typical differences between it and more common secondary study method, systematic literature review, is presented here. There is many common factors in both of the methods, and as Kitchenham et al. (2010) states, *"the distinction between mapping studies and conventional SLRs can be somewhat fuzzy"*. They also present the view that mapping studies are just a different type of systematic literature review.

Similar to other secondary study methods, mapping study aims to summarize and present the research performed in the past. Whereas systematic review focuses on very narrow research questions, mapping study usually has multiple broader questions(Kitchenham et al. 2010). Difference in the scope breadth can also be seen in the search strings, as the initial material search for mapping is likely to return a large number of studies (Kitchenham and Charters 2007; Petersen et al. 2008). Mappings are usually conducted to achieve an overview of the research area, and therefore the depth of the studies is not as great as in the SLRs. The mapping focuses more on the thematic analysis of the articles instead of an in-depth analysis of their results or gathering empirical evidence based on their results, which is often the goal of a traditional SLR (Petersen et al. 2008). Therefore, the quality of the objects of study is not relevant, unless the quality itself is the aspect to be investigated.

Since both methods have their strengths and weaknesses, using them complementarily can be

an effective combination. As suggested by Petersen (Petersen et al. 2008) and Kitchenham (Kitchenham et al. 2010), good approach is to first get an overview of the research area with systematic map, and then applying conventional literature review to a specific focus area. Results of the mapping can provide information on the quantity of available evidence, which can help target the follow-up SLRs with more precision.

3.3 Method choice

As a thesis topic, Artificial General Intelligence (AGI) is a challenging and broad area. By focusing on the AGI research articles as an object of study, useful research can still be performed without requiring too much prior knowledge and expertise on the topic from the author. The reason behind the choice of systematic literature mapping is that it fits specifically well in creating an overview of the study area. As the research questions of this thesis are broad, mapping study is a more suitable approach than a conventional SLR.

Due to the complexity of the topic, it can be difficult to enter as a newcomer. Kitchenham et al. (2010) suggest that a systematic mapping of the field can be useful to researchers new to the area. It can also be useful in introducing it to people unfamiliar with the field, both in academia and on the industry's side. As can be seen from the history of AI research in section 2.1, AGI is also a topic that is known for it's fluctuating interest and popularity, therefore seeing the temporal trends and state of the current research can be of interest. Mapping study is a good way to achieve that.

The mapping process guidelines by Petersen et al. (2008; 2015) were chosen to be followed as they were the most used ones in software engineering community (Petersen, Vakkalanka, and Kuzniarz 2015). Also the most popular guideline for systematic literature review by Kitchenham et al. (2007) can be utilized to some degree because of the similarity of the methods. These guidelines provide clear methodologies for the study. In addition, the process can further utilize the scientific paper classification scheme presented by Wieringa et al. (Wieringa et al. 2006), as it provides a useful and straightforward research facet to the categorization phase of the process.

To minimize the possible validity threats in this thesis and to try to prove its rigor, the re-

search method and it's execution is described in detail in chapters 3 and 4. However, as the study was conducted by one Master's student with limited time, the complete negation of researcher bias especially in the article exclusion and data extraction phases is not possible. As pointed out by Wohlin et al. (2013), the reliability of secondary studies such as this mapping should not be taken for granted.

4 Conducting the literature mapping

In this chapter the systematic literature mapping is conducted following the process described in section 3.1. To ensure that the study is conducted with rigor, documenting it in detail is essential. This chapter only describes the process, the actual results of the mapping are presented in chapter 5. Following the guidelines mentioned in section 3.1, the following sections focus on each of the phases of the study.

A search for similar literature studies was performed in University of Jyväskylä's thesis archive JYX, Google Scholar, and Scopus, and no results were found. Studies that were found were mostly conventional SLRs with much more narrow scope. It was concluded that there exists no similar mapping study on the topic.

4.1 Research questions

The following research questions aim to cover the goal of the study, to create an overview of the scientific research performed on the field of AGI.

RQ1: How much, and what kind of research is done in the field of AGI?

RQ2: Where and when were the studies published?

RQ3: What are the major research topics in the field and have they changed over time?

The first question focuses on the type and volume of the published AGI related research. The second question focuses on the publication venue and time of the articles, showing temporal trends and popular forums for the topic. Research question three aims to find out the most popular subtopics and their change in the study's time frame. The answers to these questions will be presented in chapter 5.

4.2 Material search

This thesis utilized both manual and broad automatic search in the material gathering phase. The targeted publication venues are either very topic specific, or high-ranked in the artificial

intelligence community. The following venues are used in this thesis:

- **Journal of Artificial general intelligence** (journal)
- **Artificial Intelligence** (journal)
- **Journal of Artificial Intelligence Research** (journal)
- **International Conference on Artificial General Intelligence** (conference)
- **International Joint Conference on Artificial Intelligence** (conference)

The inspected publications were selected based on their relatedness to AGI, as well as their perceived popularity and quality. One basis for the selection of publications was the Finnish Publication Forum (*Publication Forum*) ranking. In this thesis Publication Forum is referred to as JUFO. Publication Forum ranks peer-reviewed scientific publication venues on grades 1-3, with 1 being the standard level, and 3 being the highest.

The reasoning behind venue choices were as follows: Journal of Artificial general intelligence (JAGI) and International Conference on Artificial General Intelligence (ICAGI) are very topic specific, and therefore should provide the most relevant information within the field. These main forums are likely to contain most of the research articles relevant for this thesis.

Artificial Intelligence journal (AIJ) and Journal of Artificial Intelligence Research (JAIR) were chosen because they are considered the leading publications of the AI field by Finnish experts, having the highest JUFO ranking. Both also have high citescore (7.7 and 6.4 respectively) considering they do not focus on any specific subfield of AI. International Joint Conference on Artificial Intelligence (IJCAI) was included because it is a popular conference that produces a large amount of articles, and it is also ranked high by JUFO (rank 2). These more general sources may not contain as many papers as the topic-specific publications, but can show the relative popularity of the field within the AI research.

Out of these, the Journal of Artificial General Intelligence (JAGI) was searched manually due to a small number of articles, but otherwise automatic search was performed using search engines Scopus and Springer Link.

The search phrase for the automatic search was derived from the research questions, follow-

ing the aim of this study. Included were the close concepts similar to AGI, such as HLAI and superintelligence. The logical search phrase used was " *"artificial general intelligence" OR agi OR "human-level ai" OR hlai OR superintelligence* ". The phrase is very broad and non-restricting, but this helps the system yield results even in more general venues.

The publication interval of the papers was limited to years **2015-2019** to reduce the material to a reasonable quantity. This choice prevents seeing temporal trends that might not show in such a short interval, but a longer time frame was not viable considering this is only a master's thesis. Consequently this also keeps the focus on the most recent developments in the field.

4.3   Choosing papers for inclusion

The initial material search was performed during October 2020. The goal of the search process was to narrow down the papers using the following criteria. All literature accepted for this thesis needed to be:

- presenting and/or relating to Artificial General Intelligence, Human-level AI, or similar concepts of general intelligence,
- written in English,
- published between 2015-2019,
- a peer-reviewed journal or conference paper, and
- available digitally free or as a member of University of Jyväskylä.

Piece of literature would be excluded if it was:

- an editorial,
- not wholly available digitally or
- a duplicate of an already included paper.

For a paper to be found relevant to the topic, it needed to clearly reference the concept of (artificial) general intelligence in its title or abstract, with intention of relating the article to it. If the concept is only mentioned as a sidenote or as an example without further relevance, it should be excluded. If a clear decision couldn't be made from these parts, the article could

be further examined. In case of duplicate articles or articles concerning the same study, only the latest one is included in the study.

The initial search yielded total of 187 papers. Almost all of the initial results fulfilled the meta-level criteria as the used search engines (Scopus and Springer Link) offered extensive filtering capabilities. The preselection of venues ensured the articles were peer-reviewed and accessible, but for greater clarity these criteria were also listed above. These results were then inspected in two separate phases. In the first phase, the titles and abstracts of the articles were examined, and then the article was marked as either excluded or potential in a spreadsheet. After this initial inspection, 122 papers were considered as a potential research subject, and would be subject of further inspection.

In the second phase, the potential papers were examined again, with the research questions and inclusion criteria considered more thoroughly, to further exclude papers that were not relevant to the topic. During this phase, a small example set of excluded and included papers was sent to the thesis supervisor to confirm their categorization. During this phase it became apparent that without stricter inclusion criteria the amount of accepted papers would be excessively high. It was decided that to be included, the article would have to explicitly relate itself to general intelligence. In the end, 92 papers were included and would be used as the mapping sample. Table 1 shows the number of accepted papers and their publication venues. In the "Total articles" column, the total number of articles published during the observed period 2015-2019 is shown. "Total results" means the results from the automated search, before going through the papers manually.

Publication forum	Total articles	Total results	Potential	Accepted	Accepted (%)
JAIR	312	5	3	2	40%
IJCAI	3923	12	8	4	33.33%
AIJ	382	3	3	1	33.33%
JAGI	15	15	9	6	40%
ICAGI	152	152	99	79	51.97%
All	**4784**	**187**	**122**	**92**	**49.20%**

Table 1: Publication forums and their results

4.4 Keywording

Once the articles were narrowed down to the stated relevant topic criteria, they were inspected more thoroughly and the emergent concepts or keywords were extracted. These keywords were then further combined to topic categories until no further merging could be done without making them too general and losing too much information in the process. The resulting categorical scheme of 22 distinct topic categories is illustrated in Table 3 in the results chapter.

4.5 Data extraction and mapping

In the actual data extraction phase, the included research articles were sorted into two different categorization schemes. The first one was the topic categorization that emerged from the papers themselves. The second one was the topic-independent categorization presented by Wieringa et al. (2006). The latter enables us to make comparisons with AGI and other fields of research that aren't necessarily related. It also provides information on the kind of research done, answering RQ1. Wieringa's scheme is presented in Table 2.

Class	Class description
Evaluation research	Investigates existing problem or technique in practice with sound research methods.
Validation research	Investigates a solution proposal not yet implemented in practice using sound research methods. Prototypes, simulations or mathematical proof are among possible methods.
Solution proposal	Paper proposes a new solution to a problem, with arguments for its relevance. Solutions must be novel or significantly improve existing ones.
Philosophical papers	Conceptual frameworks and new viewpoints on the topic.
Opinion papers	Papers presenting author's personal opinion on the topic.
Personal experience papers	Present author's experience on something that has been done in practice. Doesn't necessarily provide evidence.

Table 2: Article classification by Wieringa et al. (2006)

The papers could belong to multiple categories on both schemes. This enables inspection of connections between the topics that are often met together. The results of this mapping

process are presented in the following chapter.

4.6 Material control

The initial inspection of articles was performed online. Once the paper exclusion phase was complete, the final set of articles was downloaded locally for further local inspection, and the data and metadata collected from the articles was saved in a spreadsheet document. After inputting the article data to the spreadsheet, it could easily be read by python scripts that would then either create CSV files to be used by LaTeX, or plot a graph that would be added to the thesis as is. Graphs in this thesis were plotted with either LaTeX package *Pgfplots* or with Python scripts, using graphic libraries such as *Matplotlib* and *Plotly*. Python was also used to automate parts of the keywording process, and to quickly inspect topics, classes, and other article metadata when needed. All the text and source files, as well as the data files of this thesis were stored in a git repository in Github. The repository can be found in https://github.com/Ozame/aigrad.

5 Results and analysis

In this chapter the results of the literature mapping are presented. Answers to research questions are discussed, and observations made during the study are shown. As suggested by Petersen et al. (2008), the results are visualized with graphs for easier interpretation.

5.1 Publication years and venues

As the field of Artificial General Intelligence is relatively small when compared to other branches of AI research (see Table 1), it is very focused on specific publication forums. This can be seen evidently from the sample papers' publication venues. Figure 2 shows that almost 86% of the 92 articles were published as the proceedings of the International Conference on Artificial General Intelligence. The second largest concentration of articles was on the Journal of Artificial General Intelligence, as was expected it being the only non-conference publication on the topic. It must be also noted that even though Artificial Intelligence Journal and Journal of Artificial Intelligence Research are esteemed journals that contribute dozens of AI articles every year, these articles only constituted to less than four percent of this study's papers. International Joint Conference on Artificial Intelligence, that published impressive 3923 papers as proceedings during the inspected period, only produced 4 accepted papers about the topic. It is clear that the publication of AGI research on the most popular forums is very marginal.
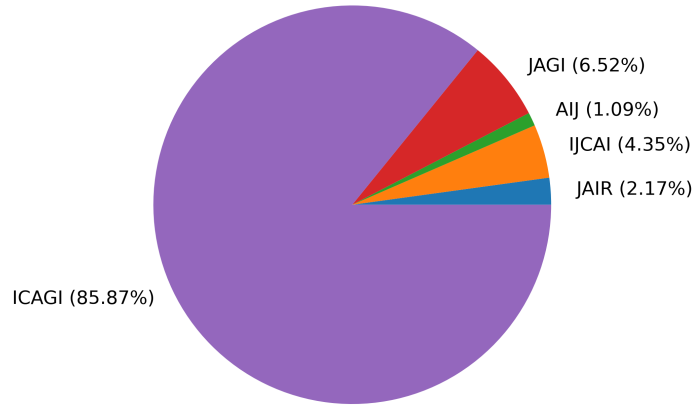
Figure 2. Article distribution between the publication forums

In Figure 3 the amount of yearly published articles and their venues are visualized in a bar plot. As the inspected time period on this thesis is short, only five years, predicting temporal trends accurately is not possible. However, it can be seen that every year there is a steady publication pace of at least 13 articles, with the average being 18.4 a year. In year 2018 25 different articles were published, which was the highest number during the inspection period.

ICAGI is the *"only major conference series devoted wholly and specifically to the creation of AI systems possessing general intelligence at the human level and ultimately beyond"* (*AGI Conference*). It has been organized yearly since 2008.

Therefore to answer research question RQ2, it can be said that while there are AGI studies being published in different publication forums, a vast majority of the research is released as yearly conference proceedings of International Conference on Artificial General Intelligence. Research community is active but small, as only average of 18,4 articles fitting this scope are published yearly.
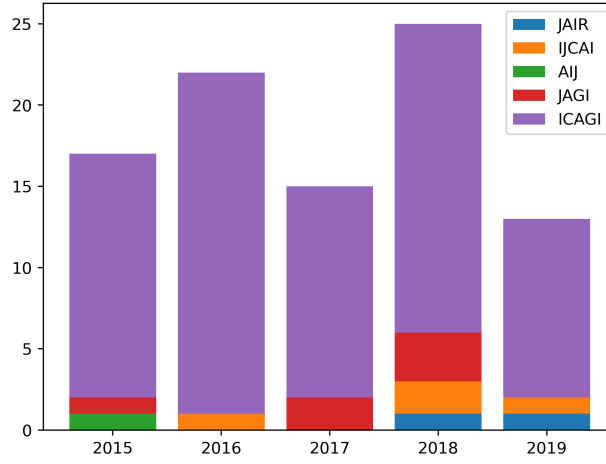
Figure 3. Yearly articles by publication forum

5.2   Common research topics

The main goal of this thesis was to achieve an overview of the AGI field. Through the mapping process, 22 different topic categories were found, and they are presented in Table 3. The order of topics is based on the themes, with similar topics close to each other. A single paper can relate to multiple topics. These topics show us how the AGI research is focused on the top level. From the amount of different topics it can be seen that the research area is broad and is not focused on only a few different approaches.

From the studied articles, four different themes could be identified. These themes are just high-level observations by the author based on the found topics, and therefore a single topic does not necessarily directly belong to only one theme. Their purpose is merely to further abstract the research field. **Building AGI systems** is of course naturally a very prominent theme. These papers were more focused on the possible theory and implementations that could be used in building of an AGI system: how it reasons and plans, how capabilities like perception could be implemented, and how to handle issues like combinatorial explosion. One other reoccurring theme was **Learning**. Different approaches to more common machine learning techniques like reinforcement learning and algorithms related to it like Q-learning and Deep Q-Networks are demonstrated. In addition to this, research on cumulative learning,

lifelong learning, and episodic memory is very prominent. The third theme was **Agent in-teraction**, with topics focusing on how the intelligent agents interact with their environment, and what kind of environments could be used to train and evaluate them. Interactions with other agents in multi-agent systems, and with humans through actions and emotions, were also researched. Lastly there are **Non-technical topics**, such as philosophical questions, AI ethics, emotion in AI and meta-level research. These kind of papers often discussed new theoretical frameworks to be used in AGI research and implementations.

#	Topic	Topic description
1	Cognitive architectures	Cognitive architectures and their descriptions.
2	AGI Design	General ideas on how AGI or its components could be designed and implemented.
3	Reasoning and Inference	Approaches on temporal and causal reasoning and inference techniques.
4	Planning and decision making	Utilizing existing knowledge in planning and making decisions.
5	Probabilistic approaches	Probabilistic approaches e.g. Bayesian techniques and uncertainty handling.
6	Category theory	Approaches relating to category theory.
7	Universal AI	Concepts relating to Universal AI (Hutter 2004): universal induction, AIXI, compression.
8	Physical robots	Physical robots and interaction with physical environment.
9	Computer vision and perception	Topics concerning vision and perception systems of an agent.
10	Nature-inspired approaches	Artificial animals, homeostatic agents and other nature-inspired ideas.
11	Reinforcement learning	Topics directly relating to reinforcement learning, e.g. Q-learning, rewarding techniques.
12	Recursive self-Improvement	Relating to fast self-improvement of an agent and intelligence explosion.
13	Experiential learning	Cumulative learning, artificial pedagogy and other topics related to how agent builds on existing knowledge.
14	Agent environment	Descriptions of environments and how agents interact within them.
15	Multi-agent systems	Topics relating to agent-to-agent interaction and cooperation.
16	Human-computer interaction	How human and agent interact and communicate and their relation to each other.
17	AI safety	Approaches on how to safely create and interact with AGI, and what safety issues arise alongside general intelligence.
18	Philosophical aspects	Philosophical questions relating to artificial intelligence, e.g. AI ethics and morality.
19	Human-like qualities	Approaches with basis on human qualities like emotion and empathy.
20	AGI research	Secondary studies about AGI research.
21	AI evaluation	How to evaluate and measure AI intelligence and performance.
22	Game playing	Game playing as a tool in development and evaluation of general agents.

Table 3: Emergent topic categories

In Figure 4, the frequencies of research topics are presented over the years. The most researched topic is **Cognitive architectures**. Cognitive architecture means an abstract model of cognition as well as its software implementation, which aims to be a system showing intelligent behaviour through artificial intelligence (Lieto et al. 2018). In AGI research there are few cognitive architectures that are standing out in the field. Goertzel's OpenCog framework (See section 2.2) was seen in 8 different papers, e.g Goertzel's idea of bridging the gap between theory and practice in AGI design (Ben Goertzel 2017) and Potanov's attempt to create semantic vision system by combining OpenCog with YOLOv2 object detection system (Potapov et al. 2018). In addition to OpenCog, Non-Axiomatic Reasoning System (NARS), developed by Pei Wang, was part of many articles. For example, besides the introduction of its implementation (Hammer, Lofthouse, and Wang 2016), there were papers describing its approach to emotion (Wang, Talanov, and Hammer 2016) and inferential learning (Wang and Li 2016). While many of these cognitive architecture articles were mainly focused on presenting authors' system's implementation, some were offering ideas that could be used in any other approaches to AGI. These articles, among other papers focusing on not-implementation-specific ideas, were also categorized in to a **AGI design** category which included 6 papers.

The second largest topic category was **Universal AI**. The theory of universal AI was created by Marcus Hutter (2004) and it describes a complete mathematical model for general artificial intelligence, named AIXI (Hutter 2012). Although incomputable, this theory and the topics related to it are targets of vigorous research. In this category there is 14 articles concerning AIXI, Solomonoff's universal induction, functional programming and compression. For example in 2019 paper by Franz, Gogulya and Löffler (2019), a monolithic inductive approach to AGI was presented, taking advantage of AIXI and incremental compression techniques. In (Martin, Everitt, and Hutter 2016), "death" was formally defined to generally intelligent agents like AIXI and discoveries were made regarding agents' behaviour in such situations.

Reinforcement learning (RL) is the third largest category with 11 articles relating directly to it. As can be seen from the topic heatmap in figure 5, it is technique associated with wide range of other topics. In 2016 paper Susumu Katayama presents a new RL algorithm

idea with similarity to AIXI (Katayama 2016). It involves usage of MAGICHASKELLER, research group's inductive programming system. RL is one of the main paradigmas of machine learning, and widely used in narrow AI approaches, but it is also a very important part of many general approaches to AI.

Experiential learning means that an agent utilizes its previous experiences through its actions, incrementally increasing its knowledge (Thórisson et al. 2019). This type of learning enables agent to generalize its abilities continuosly, making it one of the necessary requirements of AGI. 11 articles with this topic were found, with some defining new concepts such as cumulative learning (Thórisson et al. 2019), some researching novel techniques like imitation learning (Katz et al. 2016), and some focusing on how to teach cumulatively learning agents systematically via artificial pedagogy (Bieger, Thórisson, and Steunebrink 2017).

Many papers with less technical topics such as **AI safety** and **Philosophical aspects** could be found in the study, both prevalent in 11 papers. An especially targeted topic was AI safety, which concerns problems like how can we make sure AGI has the same desirable ethical values as its creators, how can a superintelligent agent be contained safely, and how we can be sure that the AGI accomplishes its goals without the possibility of using harmful unintended shortcuts. These questions are generally referred respectively as *the alignment problem, the containment problem, and the problem of perverse instantiation*. In (Babcock, Kramár, and Yampolskiy 2016), necessary requirements are identified for AGI containers to solve the containment problem. In 2019 paper, Aliman and Kester present a novel ethical framework called *Augmented Utilitarianism* to alleviate the problem of perverse instantiation (Aliman and Kester 2019). As can be seen from the quantity of papers relating to AI safety, it is clear that this is one of the main issues that needs to be solved in the creation of AGI.

The philosophical papers that were not directly concerned with AI safety were either discussing the ethics of AI or concepts like *understanding* and *intelligence* itself. For example, in (Thórisson and Kremelberg 2017), the relationship between understanding and common sense is discussed, and in (Weinbaum and Veitas 2016), a new way of looking at the intelligence, titled *Open-Ended Intelligence*, is introduced as a novel approach to AGI.
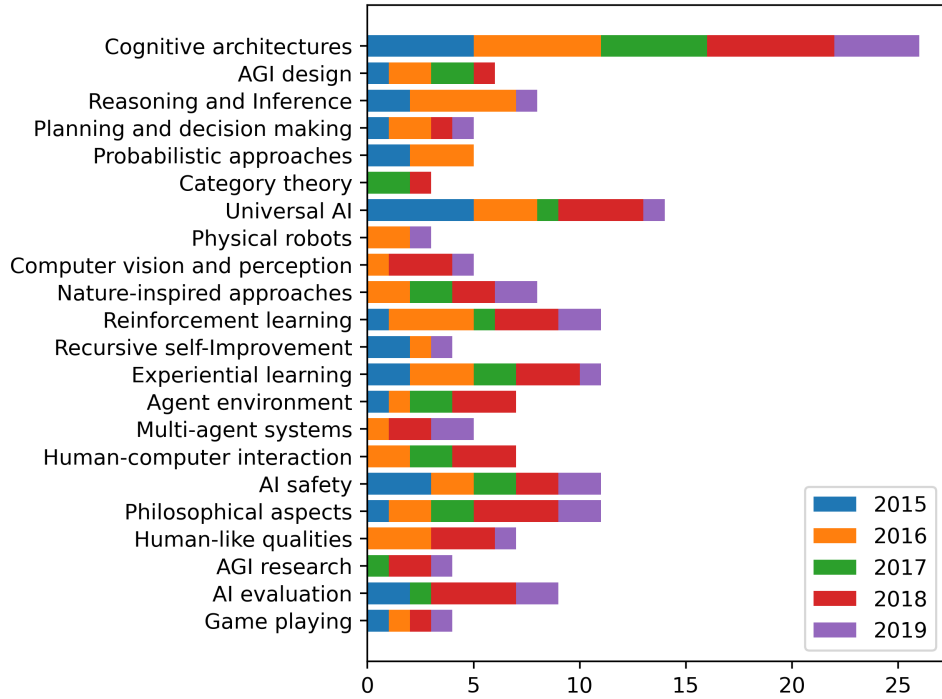
Figure 4. Topic category frequencies

5.3 Temporal trends

Due to the limited time span, it is difficult to observe long term trends in AGI research. However, some short term observations can be made from the yearly publications presented in Figure 4.

The last year included in the study, 2019, had a major drop in published articles in comparison with the previous year, from 25 to 13. In 2019, many topics that were prominent on the previous years, e.g. universal AI, agent environment, and human-computer interaction, dropped to only one or zero articles published. Surprisingly there are some topics that don't have any relating papers published in recent years, but each have one on 2019. This is the case of recursive self-improvement, physical robots, and reasoning and inference, each area has only had one paper published since 2016.

Few of the most common topics like AI safety, philosophical aspects, and cognitive architectures have a regular publication pace, with approximately the same number of articles

published every year, even in the slower years of 2017 and 2019. Other popular topics such as experiential learning and universal AI show much more fluctuation in the yearly number of articles published.

Interestingly articles about probabilistic approaches have not been published since 2016, which suggests that the interest towards that particular approach is decreasing.

5.4   Connections between topics

Figure 5 shows the relations between the topic categories. As could be expected, cognitive architectures can be associated with as many as 15 other categories. As the aim of cognitive architecture is often to create a versatile general agent, it is reflected on the way the research is done. Interestingly nature-inspired approaches is often associated with agent environment and reinforcement learning topics. One explaining factor is the subject of artificial animals, also known as animats, that were discussed in three articles. Animats are homeostatic reinforcement learning agents that interact with their environment.

The relation between AI evaluation and universal AI can also be observed, with three related articles. As universal AI deals with computability and similar subjects, their mathematical evaluation might be more viable than other approaches. Two of the three articles were authored by Ondřej Vadinský (2018a, 2018b), and focused on the Algorithmic Intelligence Quotient test, designed for intelligent agent evaluation.

A clearly visible focus area on the topic relations is the area with topics 15-20. The heatmap shows close connections between less technical topics such as human-computer interaction, AI safety, philosophical aspects, and human-like qualities, and also multi-agent systems and AGI research. This shows how discussing AI safety also requires discussing how humans and computers interact, and how abstract and difficult concepts like ethics, values and emotions can be represented and conveyed to the machine. Two out of three of the found secondary research articles were targeting this focus area, so there is an undeniable interest in these topics.
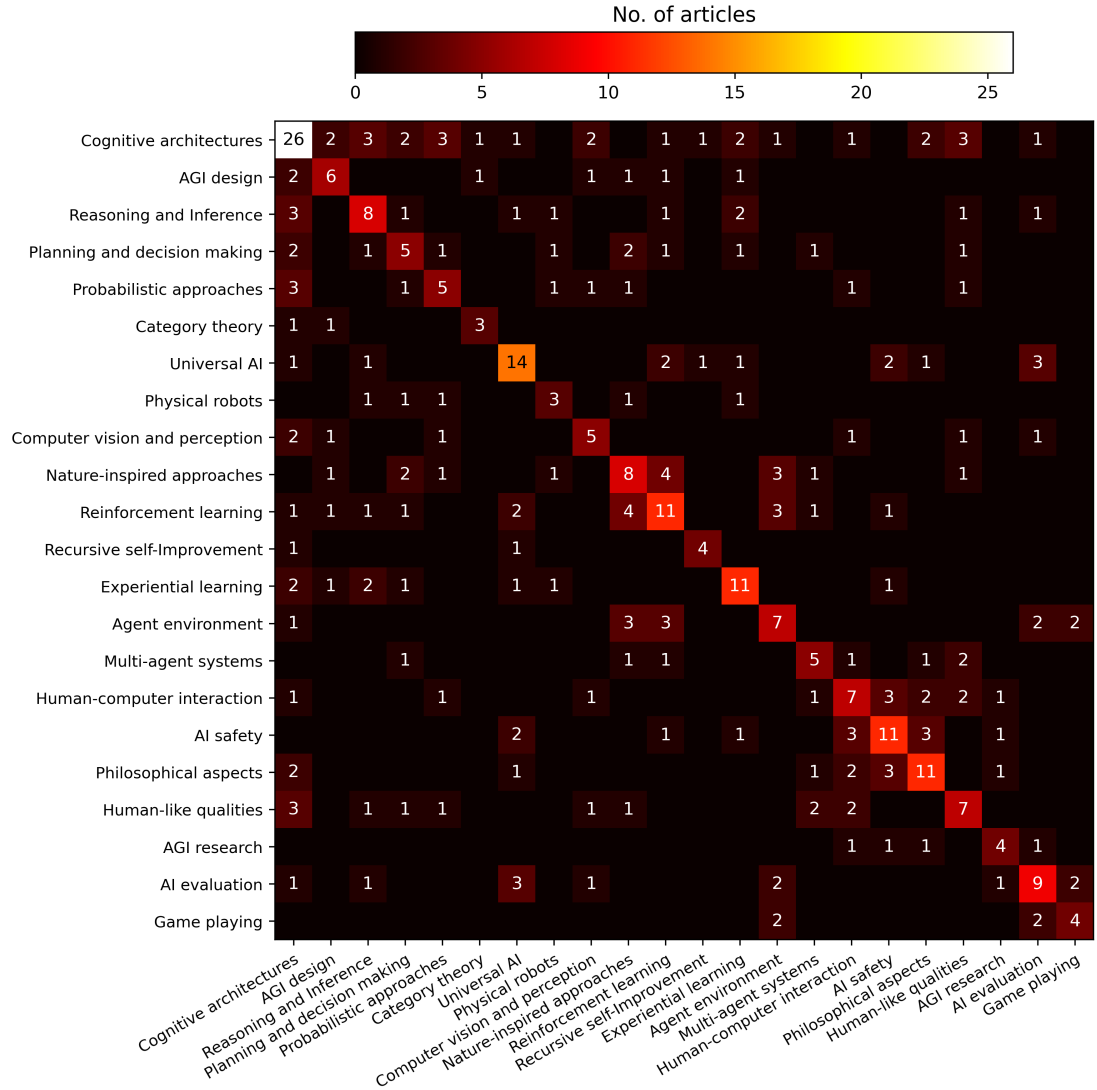
Figure 5. Topic's relations to each other

5.5   Types of AGI research

The bubble graph in Figure 6 shows us the relation of the articles' topics and their Wieringa classification. Here we can observe the specific foci of the field in two different facets.

It can be clearly seen that most of the research in the field is solution proposals. This means that the research consists predominantly of new approaches to different problems. This focus on the new ideas combined with the almost complete lack of evaluation research shows that the field is still very young, as there is not much practical applications to investigate. It is also

possible that often this kind of evaluation research could be very valuable and therefore kept private and unpublished, but as the sample articles are mostly from academia, that should not be the case here.

There is some validation research, which means investigation of not-yet-implemented solution proposals. Often solution proposal articles provided some proof in form of methodological analysis, prototype or experiments, which makes them also validation research papers.

Also common were philosophical papers, meaning papers that sketch a new way of looking at the subject, or that present a conceptual framework to be used in future research. Especially on the topics of philosophical aspects and AI safety this was a dominating research type, with multiple ethical frameworks and safety guidelines presented in the articles. Especially on these topics the lack of practical applications makes evaluation and validation research difficult, as there is currently no competent AGI.

To answer the second part of the research question RQ1, it can be said that the research in the AGI field is mostly proposals of new solutions and approaches, and philosophical papers discussing new conceptual frameworks and views. The absence of evaluation research shows that the research of existing practical solutions is nominal, probably due to the fact that the field is challenging and still focused more on the theoretical side. The lack of opinion papers also indicates that the performed research is objective.

The research gap in evaluation research could be target for future research. Finding examples of AGI solutions used in practice, and investigating their effectiveness against traditional approaches or more narrow AI solutions would be an interesting way to survey the state of the field in more detail. Especially the usage of popular cognitive architectures in real-world situations could be a good subject for a more focused systematic literature review.

Topic-wise, category theory and also physical robots are the least researched. This is interesting especially when considering the wide usage of robots in manufacturing and other industries. There are also recent suggestions that AGI will never be realized as it cannot experience the world as humans can, attaining tacit knowledge (Fjelland 2020). Considering this, it would make sense to invest in a future research that would aim to enable AIs to experience the physical world through robotics. This was also suggested by David Kremel-

berg in one of the studied articles (Kremelberg 2019), where he argues that embodiment is a necessity for general intelligence.
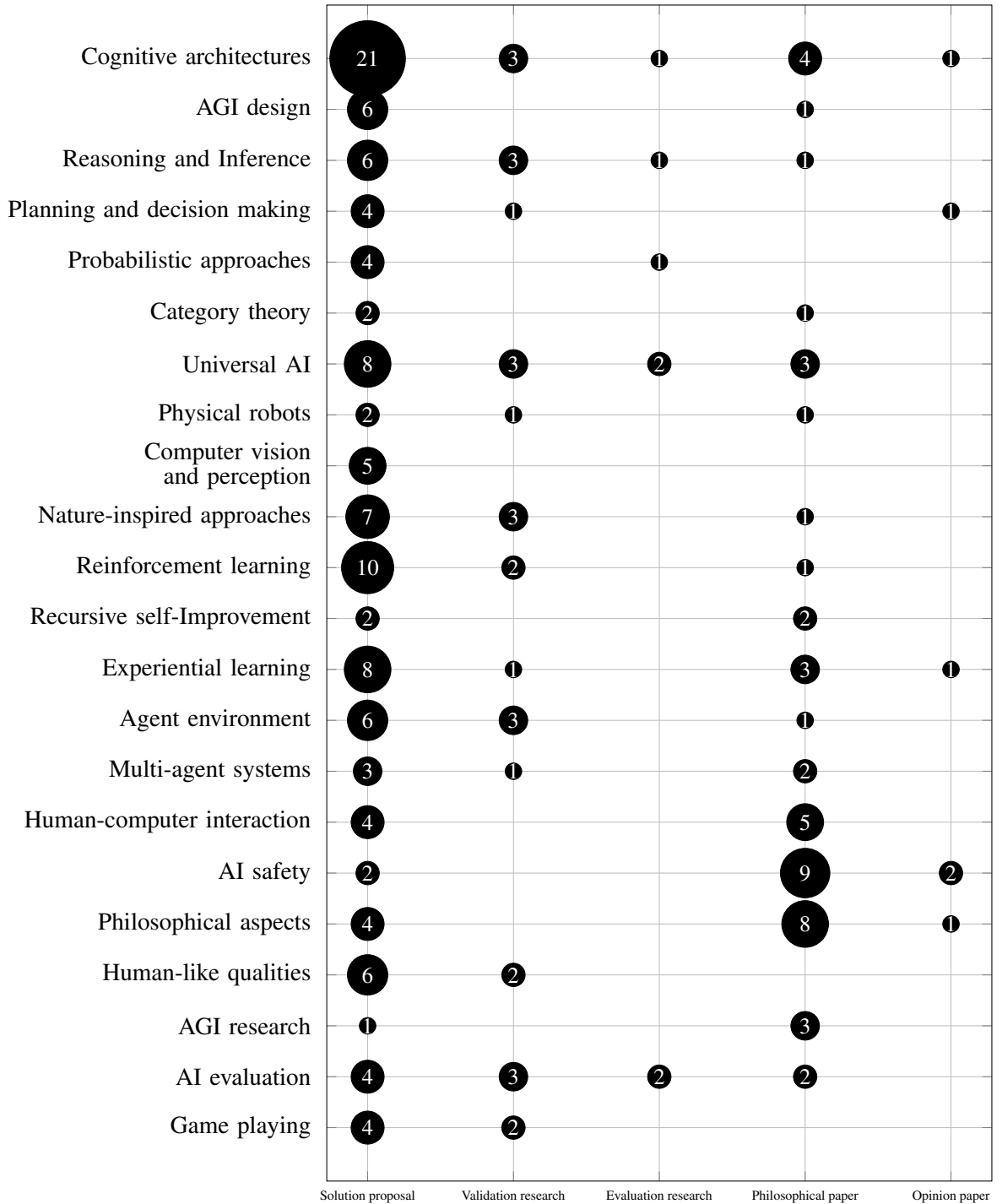

Figure 6. Article distribution between topics and article types

5.6   Research locations

AI is an issue that has been discussed in mainstream media a lot in the recent years. Even leaders of many countries have voiced their opinions about the future prospects of AI research and utilizing AI in society. Because of this, although not the main focus of this thesis, the affiliations of studied articles were also mapped geographically, and presented in Figure 7. The figure shows how the research of artificial general intelligence is focused around the world in different countries.

With 59 papers published, most of the articles are affiliated with researchers in European countries, although the largest single country in AGI research is the USA, with 36 published articles. Surprisingly, Iceland and Netherlands are the runner-ups with 10 articles each. It can be seen that economic powers like China, Russia and Japan are still in the 10 largest countries in the field, with 7, 6, and 5 papers respectively. However, their amount of published articles when compared to that of the USA is relatively low. Here it is important to notice that these numbers do not necessarily reflect the total amount of AI research done, as AGI was the focus of the mapping process. Some countries collaborate more than others, especially Iceland, Switzerland, Netherlands and the USA which often have articles involving other nationalities.

It can be observed that there are some authors whose contribution to AGI research is quite noticeable, based on the number of papers published. In Iceland, there's Kristinn R. Thórisson, in China there is Ben Goertzel, in the USA there's Pei Wang and Patrick Hammer, and also many others. Naturally their work is often relating to same subjects, in this case, cumulative learning, OpenCog, and NARS, respectively.

The complete list of countries can be found in appendix B. There was also a paper with no affiliated country, released by Microsoft's research group.

Figure 7. Article affiliation map
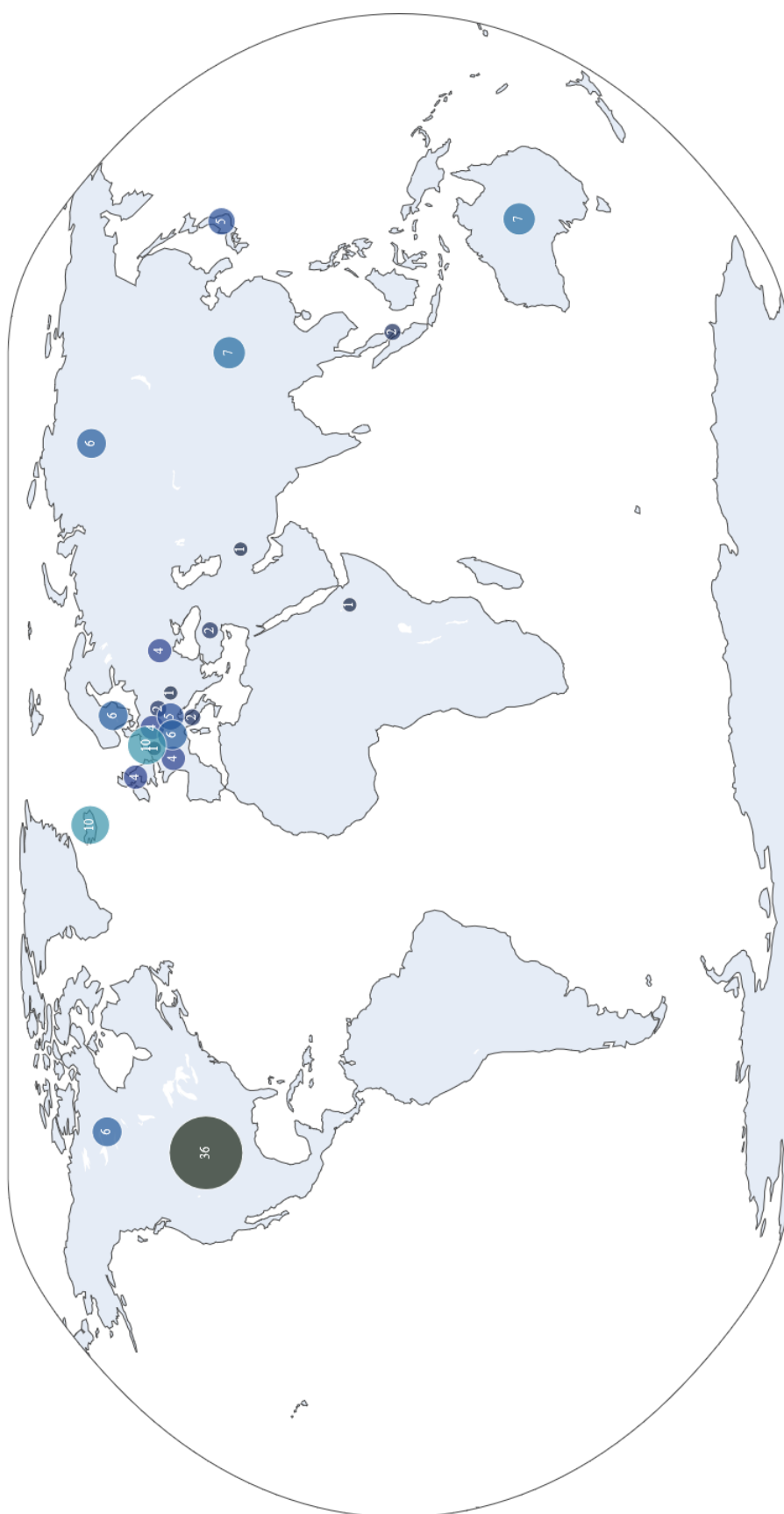
6 Conclusion

In this thesis, a systematic literature mapping study was conducted on the field of artificial general intelligence. The goal was to create a general overview of the complex AGI research field and to uncover its current state.

92 peer-reviewed articles from scientific journals and conference proceedings were inspected. With three journals and two conferences examined, it was discovered that a majority of AGI research is published as the proceedings of the International Conference on Artificial General Intelligence and in the Journal of Artificial General Intelligence, with shares of 85.87% and 6.52%, respectively. The AGI research is focused on these two venues, as the three more general forums constitute only 7.61% of the publications.

During the inspected years 2015-2019, an average of 18,4 articles were published yearly, with some fluctuation in particular years. While popular topics remain relatively well represented each year, there are topics like probabilistic approaches, that haven't been seen in the articles since 2016.

Through the mapping process, 22 distinct topical categories were found. Major themes in the research were development of AGI systems, different types of learning, interaction of agents, and philosophical questions about AI. Topics that stood out the most were cognitive architectures, universal AI, reinforcement learning, experiental learning, and AI safety and ethics. Cognitive architecture frameworks and implementations like OpenCog and NARS are heavily researched, with 26 articles directly relating to them. Universal AI, which comprises subjects like universal induction and AIXI, is the second most researched topic with 14 relating papers. It is also interesting to see that as the dangers of AI and "intelligence explosion" are subjects often discussed in the media, AI safety is also one of the most researched topics in the field of AGI.

When viewed through scientific paper classification by Wieringa et al. (Wieringa et al. 2006), the current AGI research is mostly solution proposals, and presenting new ideas and approaches to problems. The lack of evaluation research shows that there is not yet much practical applications to evaluate. The ultimate goal of AGI is not yet realized and the complete

absence of articles presenting industry applications is a concerning sign of slow progress.

Philosophical papers presenting conceptual frameworks and new views are also common in the field. AI safety, meaning how to control and contain superintelligent agent, and ethical questions about values and the relationship between the man and the machine are clearly areas that require further discussion.

When placing the AGI research on the geographic map, it is apparent that most of the explorations in the field is performed by researchers in Europe and the United States of America. In Europe, nations standing out are Iceland and Netherlands, both publishing more articles on the subject than Russia, China and Japan. However, this may not reflect the amount of other AI research besides AGI.

When concerning future AGI research, the gaps observed through the mapping would suggest that there is a need for more research on practical applications of AGI, if there are any. This would truly show the current state of progress, and could help the growth of interest in the area. It is also seen that there are only few studies combining robotics with AGI, and as there are suggestions that having a physical body is required for human-like intelligence, it would make sense to further investigate this subject as well.

In this mapping study, it was observed that while AGI research is definitely not the most popular subfield of AI at the moment, there is steady amount of articles being published on the topics regularly in its main publication forums, with wide variety of different issues. It is obvious that even though there have been major breakthroughs in AI in recent years, the ultimate goal of general intelligence is not yet close to realization.

Due to the fact that this was a master's thesis with limited time resources, there are things that could be done to improve the rigour of this study. For example, having a second researcher take part during the exclusion and data extraction phases would ensure no papers are wrongly excluded, and prevent possible misclassifications of articles. If this study were to be continued in the future, it would also be very interesting to inspect a longer timespan to catch long term trends not visible in the inspected five year period.

Bibliography

Adams, Sam, Itmar Arel, Joscha Bach, Robert Coop, Rod Furlan, Ben Goertzel, J. Storrs Hall, et al. March 2012. "Mapping the Landscape of Human-Level Artificial General Intelligence". *AI Magazine* 33, number 1 (): 25–42. https://doi.org/10.1609/aimag.v33i1.2322. https://www.aaai.org/ojs/index.php/aimagazine/article/view/2322.

AGI Conference. Accessed: 2021-01-14. https://www.agi-conference.org/.

Aliman, Nadisha-Marie, and Leon Kester. 2019. "Augmented Utilitarianism for AGI Safety". In *International Conference on Artificial General Intelligence,* 11–21. Springer.

Anderson, M. 2006. "Intelligence". *MS Encarta online encyclopedia.*

Babcock, James, János Kramár, and Roman Yampolskiy. 2016. "The AGI containment problem". In *International Conference on Artificial General Intelligence,* 53–63. Springer.

Bieger, Jordi, Kristinn R Thórisson, and Bas R Steunebrink. 2017. "The pedagogical pentagon: A conceptual framework for artificial pedagogy". In *International Conference on Artificial General Intelligence,* 212–222. Springer.

Bramer, Max. 2009. *Artificial Intelligence. An International Perspective: An International Perspective.* Volume 5640. Springer.

Brereton, Pearl, Barbara A Kitchenham, David Budgen, Mark Turner, and Mohamed Khalil. 2007. "Lessons from applying the systematic literature review process within the software engineering domain". *Journal of systems and software* 80 (4): 571–583.

Buchanan, B, E Feigenbaum, and J Lederberg. 1968. "Heuristic DENDRAL-A program for generating explanatory hypotheses in organic chemistry."

Cheeseman, Peter C. 1985. "In Defense of Probability." In *IJCAI,* 2:1002–1009.

Condori-Fernandez, N., M. Daneva, K. Sikkel, R. Wieringa, O. Dieste, and O. Pastor. 2009. "A systematic mapping study on empirical evaluation of software requirements specifications techniques". In *2009 3rd International Symposium on Empirical Software Engineering and Measurement,* 502–505.

Fjelland, Ragnar. 2020. "Why general artificial intelligence will not be realized". *Humanities and Social Sciences Communications* 7 (1): 1–9.

Franz, Arthur, Victoria Gogulya, and Michael Löffler. 2019. "WILLIAM: A monolithic approach to AGI". In *International Conference on Artificial General Intelligence,* 44–58. Springer.

Goertzel, B. 2012. "CogPrime : An Integrative Architecture for Embodied Artificial General Intelligence".

Goertzel, Ben. 2017. "From abstract agents models to real-world AGI architectures: bridging the gap". In *International Conference on Artificial General Intelligence,* 3–12. Springer.

Goertzel, Ben, and Cassio Pennachin. 2007. *Artificial general intelligence.* Volume 2. Springer.

Goertzel, Ben, and Gino Yu. 2014. "From here to AGI: A roadmap to the realization of human-level artificial general intelligence". *2014 International Joint Conference on Neural Networks (IJCNN),* 1525–1533.

Haenlein, Michael, and Andreas Kaplan. July 2019. "A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence:" *California Management Review* (). https://doi.org/10.1177/0008125619864925. https://journals.sagepub.com/doi/10.1177/0008125619864925.

Hammer, Patrick, Tony Lofthouse, and Pei Wang. 2016. "The OpenNARS implementation of the non-axiomatic reasoning system". In *International conference on artificial general intelligence,* 160–170. Springer.

Hutter, Marcus. 2004. *Universal artificial intelligence: Sequential decisions based on algorithmic probability.* Springer Science & Business Media.

————. 2012. "One decade of universal artificial intelligence". In *Theoretical foundations of artificial general intelligence,* 67–88. Springer.

Katayama, Susumu. 2016. "Ideas for a reinforcement learning algorithm that learns programs". In *International Conference on Artificial General Intelligence,* 354–362. Springer.

Katz, Garrett, Di-Wei Huang, Rodolphe Gentili, and James Reggia. 2016. "Imitation learning as cause-effect reasoning". In *International Conference on Artificial General Intelligence,* 64–73. Springer.

Kitchenham, Barbara, and Stuart Charters. 2007. *Guidelines for performing systematic literature reviews in software engineering.* Technical report.

Kitchenham, Barbara, Rialette Pretorius, David Budgen, O Pearl Brereton, Mark Turner, Mahmood Niazi, and Stephen Linkman. 2010. "Systematic literature reviews in software engineering–a tertiary study". *Information and software technology* 52 (8): 792–805.

Kitchenham, Barbara A, Tore Dyba, and Magne Jorgensen. 2004. "Evidence-based software engineering". In *Proceedings. 26th International Conference on Software Engineering,* 273–281. IEEE.

Kremelberg, David. 2019. "Embodiment as a necessary a priori of general intelligence". In *International Conference on Artificial General Intelligence,* 132–136. Springer.

Legg, Shane, Marcus Hutter, et al. 2007. "A collection of definitions of intelligence". *Frontiers in Artificial Intelligence and applications* 157:17.

Lieto, Antonio, Mehul Bhatt, Alessandro Oltramari, and David Vernon. 2018. *The role of cognitive architectures in general artificial intelligence.*

Lighthill, James. 1973. "Artificial intelligence: A general survey". In *Artificial Intelligence: a paper symposium,* 1–21. Science Research Council London.

Liu, Jiaying, Xiangjie Kong, Feng Xia, Xiaomei Bai, Lei Wang, Qing Qing, and Ivan Lee. 2018. "Artificial intelligence in the 21st century". *IEEE Access* 6:34403–34421.

Martin, Jarryd, Tom Everitt, and Marcus Hutter. 2016. "Death and suicide in universal artificial intelligence". In *International Conference on Artificial General Intelligence,* 23–32. Springer.

McCarthy, John. 2007. "From here to human-level AI". *Artificial Intelligence* 171 (18): 1174–1182.

McCulloch, Warren S, and Walter Pitts. 1943. "A logical calculus of the ideas immanent in nervous activity". *The bulletin of mathematical biophysics* 5 (4): 115–133.

Mehta, Nishita, Anil Pandit, and Sharvari Shukla. 2019. "Transforming healthcare with big data analytics and artificial intelligence: A systematic mapping study". *Journal of biomedical informatics* 100:103311.

Minsky, Marvin. 2007. *The emotion machine: Commonsense thinking, artificial intelligence, and the future of the human mind.* Simon / Schuster.

Minsky, Marvin L, Push Singh, and Aaron Sloman. 2004. "The St. Thomas common sense symposium: designing architectures for human-level intelligence". *Ai Magazine* 25 (2): 113–113.

Myers, Ware. 1986. "Introduction to expert systems". *IEEE Computer Architecture Letters* 1 (01): 100–109.

Newell, Allen, John C Shaw, and Herbert A Simon. 1959. "Report on a general problem solving program". In *IFIP congress,* 256:64. Pittsburgh, PA.

Nilsson, Nils J. 2005. "Human-level artificial intelligence? Be serious!" *AI magazine* 26 (4): 68–68.

———. 2009. *The quest for artificial intelligence.* Cambridge University Press.

Pearl, Judea. 1988. "Probabilistic reasoning in intelligent systems: Networks of plausible inference". *Morgan Kauffmann Publishers Inc.*

Petersen, Kai, Robert Feldt, Shahid Mujtaba, and Michael Mattsson. 2008. "Systematic mapping studies in software engineering". In *12th International Conference on Evaluation and Assessment in Software Engineering (EASE) 12,* 1–10.

Petersen, Kai, Sairam Vakkalanka, and Ludwik Kuzniarz. 2015. "Guidelines for conducting systematic mapping studies in software engineering: An update". *Information and Software Technology* 64:1–18.

Potapov, Alexey, Innokentii Zhdanov, Oleg Scherbakov, Nikolai Skorobogatko, Hugo Latapie, and Enzo Fenoglio. 2018. "Semantic image retrieval by uniting deep neural networks and cognitive architectures". In *International Conference on Artificial General Intelligence,* 196–206. Springer.

Publication Forum. Accessed: 2020-10-20. Federation of Finnish Learned Societies. https://www.julkaisufoorumi.fi/en.

Qadir, M. M., and M. Usman. 2011. "Software Engineering Curriculum: A systematic mapping study". In *2011 Malaysian Conference in Software Engineering,* 269–274.

Rumelhart, David E, Richard Durbin, Richard Golden, and Yves Chauvin. 1995. "Backpropagation: The basic theory". *Backpropagation: Theory, architectures and applications,* 1–34.

Russell, Stuart, and Peter Norvig. 2009. *Artificial Intelligence: A Modern Approach.* 3rd. Prentice Hall Press. ISBN: 0136042597.

Searle, John, et al. 1980. "Minds, Brains and Programs". *Behavioral and Brain Science* 3:417–424.

Shortliffe, Edward H, Randall Davis, Stanton G Axline, Bruce G Buchanan, C Cordell Green, and Stanley N Cohen. 1975. "Computer-based consultations in clinical therapeutics: explanation and rule acquisition capabilities of the MYCIN system". *Computers and biomedical research* 8 (4): 303–320.

Silver, David, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. 2016. "Mastering the game of Go with deep neural networks and tree search". *nature* 529 (7587): 484–489.

Silver, David, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, et al. 2018. "A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play". *Science* 362 (6419): 1140–1144.

The Open Cognition Project. Accessed: 2020-09-24. https://wiki.opencog.org/w/The_Open_Cognition_Project.

Thórisson, Kristinn R, Jordi Bieger, Xiang Li, and Pei Wang. 2019. "Cumulative learning". In *International Conference on Artificial General Intelligence,* 198–208. Springer.

Thórisson, Kristinn R, and David Kremelberg. 2017. "Understanding and Common Sense: Two Sides of the Same Coin?" In *International Conference on Artificial General Intelligence,* 201–211. Springer.

TURING, A.M. 1950. "Computing machinery and intelligence". *Mind* 59 (236): 433–460.

Vadinský, Ondřej. 2018a. "Towards general evaluation of intelligent systems: lessons learned from reproducing AIQ test results". *Journal of Artificial General Intelligence* 9 (1): 1–54.

———. 2018b. "Towards General Evaluation of Intelligent Systems: Using Semantic Analysis to Improve Environments in the AIQ Test". In *International Conference on Artificial General Intelligence,* 248–258. Springer.

Wang, Pei. 1995. "On the working definition of intelligence". *Center for Research on Concepts and Cognition CRCC, Indiana University.*

Wang, Pei, and Xiang Li. 2016. "Different conceptions of learning: function approximation vs. self-organization". In *International Conference on Artificial General Intelligence,* 140–149. Springer.

Wang, Pei, Max Talanov, and Patrick Hammer. 2016. "The emotional mechanisms in NARS". In *International Conference on Artificial General Intelligence,* 150–159. Springer.

Weinbaum, David, and Viktoras Veitas. 2016. "Open-Ended Intelligence - On the Role of Individuation in AGI". In *AGI.*

Wiafe, Isaac, Felix Nti Koranteng, Emmanuel Nyarko Obeng, Nana Assyne, Abigail Wiafe, and Stephen R Gulliver. 2020. "Artificial intelligence for cybersecurity: a systematic mapping of literature". *IEEE Access* 8:146598–146612.

Widrow, Bernard, and Michael A Lehr. 1995. "Perceptrons, Adalines, and backpropagation". *Arbib* 4:719–724.

Wieringa, Roel, Neil Maiden, Nancy Mead, and Colette Rolland. 2006. "Requirements engineering paper classification and evaluation criteria: a proposal and a discussion". *Requirements engineering* 11 (1): 102–107.

Wohlin, Claes, Per Runeson, Paulo Anselmo da Mota Silveira Neto, Emelie Engström, Ivan do Carmo Machado, and Eduardo Santana de Almeida. 2013. "On the reliability of mapping studies in software engineering". *Journal of Systems and Software* 86 (10): 2594–2610. ISSN: 0164-1212. https://doi.org/https://doi.org/10.1016/j.jss.2013.04.076. http://www.sciencedirect.com/science/article/pii/S0164121213001234.

Appendices

A Accepted papers

Year	Authors	Title	Class	Categories
2015	Davis E.	Ethical guidelines for a superintelligence	Opinion paper	AI safety, Philosophical aspects
2015	Thomas Manzini, Simon Ellis, James Hendler	A Play on Words: Using Cognitive Computing as a Basis for AI Solvers in Word Puzzles	Solution proposal	Game playing
2015	Vitaly Khudobakhshov, Andrey Pitko, Denis Zotov	Programming Languages and Artificial General Intelligence	Solution proposal	Cognitive architectures, Probabilistic approaches
2015	Zoltán Tősér, András Lőrincz	The Cyber-Physical System Approach Towards Artificial General Intelligence: The Problem of Verification	Solution proposal	AGI design, Reinforcement learning
2015	Jordi Bieger, Kristinn R. Thórisson, Pei Wang	Safe Baby AGI	Opinion paper	AI safety, Experiential learning
2015	Roman V. Yampolskiy	On the Limits of Recursively Self-Improving AGI	Philosophical paper	Recursive self-Improvement
2015	Roman V. Yampolskiy	Analysis of Types of Self-Improving Software	Philosophical paper	Recursive self-Improvement
2015	Pei Wang, Patrick Hammer	Assumptions of Decision-Making Models in AGI	Opinion paper	Cognitive architectures, Planning and decision making
2015	Arthur Franz	Toward Tractable Universal Induction Through Recursive Program Learning	Solution proposal	Universal AI
2015	Can Eren Sezener	Inferring Human Values for Safe AGI Design	Philosophical paper	AI safety, Universal AI
2015	Pei Wang, Patrick Hammer	Issues in Temporal and Causal Inference	Philosophical paper, Solution proposal	Cognitive architectures, Reasoning and Inference
2015	Ute Schmid, Marco Ragni	Comparing Computer Models Solving Number Series Problems	Evaluation research	Universal AI, Reasoning and Inference, AI evaluation
2015	Kristinn R. Thórisson, Jordi Bieger, Stephan Schiffel, Deon Garrett	Towards Flexible Task Environments for Comprehensive Evaluation of Artificial Intelligent Systems and Automatic Learners	Philosophical paper	Agent environment, AI evaluation

2015	Ben Goertzel, Nil Geisweiller, Eddie Monroe, Mike Duncan, Selamawit Yilma, Meseret Dastaw, Misgana Bayetta, Amen Belayneh, Matthew Ikle, Gino Yu	Speculative Scientific Inference via Synergetic Combination of Probabilistic Logic and Evolutionary Pattern Recognition	Evaluation research	Cognitive architectures, Probabilistic approaches
2015	Susumu Katayama	Towards Human-Level Inductive Functional Programming	Solution proposal	Universal AI
2015	Eric Nivel, Kristinn R. Thórisson, Bas Steunebrink, Jürgen Schmidhuber	Anytime Bounded Rationality	Solution proposal	Cognitive architectures, Experiential learning
2015	Benja Fallenstein, Nate Soares, Jessica Taylor	Reflective Variants of Solomonoff Induction and AIXI	Solution proposal	Universal AI
2016	Johnson M., Hofmann K., Hutton T., Bignell D.	The malmo platform for artificial intelligence experimentation	Solution proposal	Agent environment, Game playing
2016	John Licato, Maxwell Fowler	Embracing Inference as Action: A Step Towards Human-Level Reasoning	Solution proposal	Reasoning and Inference
2016	Ben Goertzel	Probabilistic Growth and Mining of Combinations: A Unifying Meta-Algorithm for Practical General Intelligence	Solution proposal	Cognitive architectures, Probabilistic approaches
2016	Sean Markan	System Induction Games and Cognitive Modeling as an AGI Methodology	Solution proposal	Experiential learning, AGI design
2016	James Babcock, János Kramár, Roman Yampolskiy	The AGI Containment Problem	Philosophical paper	AI safety
2016	Jarryd Martin, Tom Everitt, Marcus Hutter	Death and Suicide in Universal Artificial Intelligence	Philosophical paper	AI safety, Universal AI, Reinforcement learning
2016	Ben Goertzel, Misgana Bayetta Belachew, Matthew Ikle, Gino Yu	Controlling Combinatorial Explosion in Inference via Synergy with Nonlinear-Dynamical Attention Allocation	Validation research	Cognitive architectures, Human-like qualities, Reasoning and Inference
2016	Andrew Macfie	Analysis of Algorithms and Partial Algorithms	Solution proposal	Recursive self-Improvement

2016	Nutchanon Yongsatian-chot, Stacy Marsella	Integrating Model-Based Prediction and Facial Expressions in the Perception of Emotion	Solution proposal	Human-computer interaction, Human-like qualities, Computer vision and perception, Probabilistic approaches
2016	David Weaver Weinbaum, Viktoras Veitas	Open-Ended Intelligence	Philosophical paper	Philosophical aspects, Multi-agent systems
2016	Kristinn R. Thórisson, David Kremelberg, Bas R. Steunebrink, Eric Nivel	About Understanding	Philosophical paper, Solution proposal	Cognitive architectures, Human-computer interaction, Philosophical aspects
2016	Garrett Katz, Di-Wei Huang, Rodolphe Gentili, James Reggia	Imitation Learning as Cause-Effect Reasoning	Solution proposal, Validation research	Experiential learning, Reasoning and Inference, Physical robots
2016	Claes Strannegård, Abdul Rahim Nizamani	Integrating Symbolic and Sub-symbolic Reasoning	Solution proposal	Reasoning and Inference, Reinforcement learning
2016	Frédéric Alexandre, Maxime Carrere	Modeling Neuromodulation as a Framework to Integrate Uncertainty in General Cognitive Architectures	Solution proposal	AGI design, Nature-inspired approaches
2016	Craig Sherstan, Adam White, Marlos C. Machado, Patrick M. Pilarski	Introspective Agents: Confidence Measures for General Value Functions	Solution proposal	Experiential learning, Planning and decision making, Reasoning and Inference
2016	Susumu Katayama	Ideas for a Reinforcement Learning Algorithm that Learns Programs	Solution proposal	Universal AI, Reinforcement learning
2016	Pei Wang, Max Talanov, Patrick Hammer	The Emotional Mechanisms in NARS	Solution proposal	Cognitive architectures, Human-like qualities
2016	Pei Wang, Xiang Li	Different Conceptions of Learning: Function Approximation vs. Self-Organization	Solution proposal	Cognitive architectures

2016	Alexey Potapov, Sergey Rodionov, Vita Potapova	Real-Time GA-Based Probabilistic Programming in Application to Robot Control	Solution proposal	Planning and decision making, Probabilistic approaches, Nature-inspired approaches, Physical robots
2016	Patrick Hammer, Tony Lofthouse, Pei Wang	The OpenNARS Implementation of the Non-Axiomatic Reasoning System	Solution proposal	Cognitive architectures
2016	Arthur Franz	Some Theorems on Incremental Compression	Solution proposal, Validation research	Universal AI
2016	Tom Everitt, Marcus Hutter	Avoiding Wireheading with Value Reinforcement Learning	Solution proposal	Reinforcement learning
2017	Claes Strannegård, Abdul Rahim Nizamani, Jonas Juel, Ulf Persson	Learning and Reasoning in Unknown Domains	Solution proposal	Cognitive architectures, Agent environment
2017	Paul S. Rosenbloom, Abram Demski, Volkan Ustun	The Sigma Cognitive Architecture and System: Towards Functionally Elegant Grand Unification	Solution proposal	Cognitive architectures
2017	Marcus Abundis	A Priori Modeling of Information and Intelligence	Philosophical paper	Nature-inspired approaches
2017	Christian Rodriguez, Giselle Marston, William Goolkasian, Ashley Rosenberg, Andrew Nuxoll	The MaRz Algorithm: Towards an Artificial General Episodic Learner	Solution proposal	Experiential learning
2017	Steven Phillips	A General (Category Theory) Principle for General Intelligence: Duality (Adjointness)	Solution proposal	Category theory
2017	Nadisha-Marie Aliman	Malevolent Cyborgization	Philosophical paper	AI safety, Human-computer interaction
2017	Ben Goertzel	From Abstract Agents Models to Real-World AGI Architectures: Bridging the Gap	Solution proposal, Philosophical paper	Cognitive architectures, AGI design
2017	Timothy Van Gelder, Richard De Rozario	Pursuing Fundamental Advances in Human Reasoning	Philosophical paper	AGI research, AI evaluation
2017	Kristinn R. Thórisson, David Kremelberg	Understanding and Common Sense: Two Sides of the Same Coin?	Philosophical paper	Philosophical aspects

2017	Ben Goertzel	A Formal Model of Cognitive Synergy	Solution proposal	Cognitive architectures, AGI design, Category theory
2017	Jordi Bieger, Kristinn R. Thórisson, Bas R. Steunebrink	The Pedagogical Pentagon: A Conceptual Framework for Artificial Pedagogy	Philosophical paper	Experiential learning
2017	Tobias Wängberg, Mikael Böörs, Elliot Catt, Tom Everitt, Marcus Hutter	A Game-Theoretic Analysis of the Off-Switch Game	Philosophical paper	AI safety, Human-computer interaction
2017	Claes Strannegård, Nils Svangård, Joscha Bach, Bas Steunebrink	Generic Animats	Solution proposal	Agent environment, Nature-inspired approaches, Reinforcement learning
2017	Pei Wang, Xiang Li, Patrick Hammer	Self-awareness and Self-control in NARS	Solution proposal	Cognitive architectures, Philosophical aspects
2017	Arthur Franz	On Hierarchical Compression and Power Laws in Nature	Solution proposal, Validation research	Universal AI
2018	Machado M.C., Bellemare M.G., Talvitie E., Veness J., Hausknecht M., Bowling M.	Revisiting the arcade learning environment: Evaluation protocols and open problems for general agents	Solution proposal, Validation research	Agent environment, AI evaluation, Game playing
2018	Everitt T., Lea G., Hutter M.	AGI safety literature review	Philosophical paper	AI safety, AGI research
2018	Yu H., Shen Z., Miao C., Leung C., Lesser V.R., Yang Q.	Building ethics into artificial intelligence	Philosophical paper	Human-computer interaction, AGI research, Philosophical aspects
2018	Ondřej Vadinský	Towards General Evaluation of Intelligent Systems: Lessons Learned from Reproducing AIQ Test Results	Validation research	Universal AI, AI evaluation
2018	Claes Strannegård, Nils Svangård, David Lindström, Joscha Bach, Bas Steunebrink	Learning and decision-making in artificial animals	Solution proposal, Validation research	Nature-inspired approaches, Reinforcement learning, Agent environment

2018	Naoto Yoshida	Homeostatic Agent for General Environment	Solution proposal, Validation research	Nature-inspired approaches, Reinforcement learning, Agent environment
2018	Andreas M. Hein, Hélène Condat	Can Machines Design? An Artificial General Intelligence Approach	Solution proposal	Human-like qualities
2018	Eray Özkural	The Foundations of Deep Learning with a Path Towards General Intelligence	Philosophical paper	Universal AI, Philosophical aspects
2018	Nadisha-Marie Aliman, Leon Kester	Hybrid Strategies Towards Safe "Self-Aware" Superintelligent Systems	Solution proposal, Philosophical paper	AI safety, Human-computer interaction
2018	Jeremy O. Turner, Steve Dipaola	Transforming Kantian Aesthetic Principles into Qualitative Hermeneutics for Contemplative AGI Agents	Philosophical paper	Philosophical aspects
2018	Anton Kolonin	Resource-Constrained Social Evidence Based Cognitive Model for Empathy-Driven Artificial Intelligence	Solution proposal	Human-like qualities, Multi-agent systems, Human-computer interaction
2018	Alexey Potapov, Sergey Rodionov, Maxim Peterson, Oleg Scherbakov, Innokentii Zhdanov, Nikolai Skorobogatko	Vision System for AGI: Problems and Directions	Solution proposal	AGI design, Computer vision and perception
2018	Pei Wang, Patrick Hammer	Perception from an AGI Perspective	Solution proposal	Cognitive architectures, Computer vision and perception
2018	Giacomo Spigler	The Temporal Singularity: Time-Accelerated Simulated Civilizations and Their Implications	Philosophical paper	Multi-agent systems
2018	Kristinn R. Thórisson, Arthur Talbot	Cumulative Learning with Causal-Relational Models	Solution proposal	Experiential learning
2018	Andrej Dameski	A Comprehensive Ethical Framework for AI Entities: Foundations	Solution proposal	Philosophical aspects
2018	Ondřej Vadinský	Towards General Evaluation of Intelligent Systems: Using Semantic Analysis to Improve Environments in the AIQ Test	Evaluation research	Universal AI, AI evaluation
2018	Xiang Li, Patrick Hammer, Pei Wang, Hongling Xie	Functionalist Emotion Model in NARS	Solution proposal	Cognitive architectures, Human-like qualities

2018	Jordi E. Bieger, Kristinn R. Thórisson	Task Analysis for Teaching Cumulative Learners	Solution proposal, Philosophical paper	Experiential learning
2018	Rafik Hadfi	Solving Tree Problems with Category Theory	Philosophical paper	Category theory
2018	Alexey Potapov, Inno-kentii Zhdanov, Oleg Scherbakov, Nikolai Skorobogatko, Hugo Latapie, Enzo Fenoglio	Semantic Image Retrieval by Uniting Deep Neu-ral Networks and Cognitive Architectures	Solution proposal	Cognitive archi-tectures, Computer vision and perception
2018	Patrick Hammer, Tony Lofthouse	Goal-Directed Procedure Learning	Solution proposal	Cognitive architec-tures, Planning and decision making, Reinforcement learning
2018	Alex Glushchenko, Andres Suarez, Anton Kolonin, Ben Goertzel, Claudia Castillo, Man Hin Leung, Oleg Baskov	Unsupervised Language Learning in OpenCog	Solution proposal	Cognitive ar-chitectures, AI evaluation
2018	Khin Hua Ng, Zhiyuan Du, Gee Wah Ng	DSO Cognitive Architecture: Implementation and Validation of the Global Workspace En-hancement	Validation research	Cognitive architec-tures
2018	Robert Wünsche	Adaptive Compressed Search	Solution proposal	Universal AI, Experi-ential learning
2019	Schofield M., Thielscher M.	General game playing with imperfect informa-tion	Solution proposal, Validation research	AI evaluation, Game playing
2019	Sampat S.	Technical, hard and explainable question answer-ing (the-QA)	Solution proposal	Computer vision and perception, AI evalu-ation
2019	Mohammadreza Alidoust	AGI Brain: A Learning and Decision Mak-ing Framework for Artificial General Intelligence Systems Based on Modern Control Theory	Solution proposal, Validation research	Planning and decision making, Multi-agent systems, Human-like quali-ties, Nature-inspired approaches
2019	David Kremelberg	Embodiment as a Necessary a Priori of General Intelligence	Philosophical paper	Physical robots
2019	Nadisha-Marie Aliman, Leon Kester	Augmented Utilitarianism for AGI Safety	Philosophical paper	AI safety, Philosoph-ical aspects

2019	Ross Gruetzemacher, David Paradice	Toward Mapping the Paths to AGI	Solution proposal	AGI research
2019	Kristinn R. Thórisson, Jordi Bieger, Xiang Li, Pei Wang	Cumulative Learning	Philosophical paper	Cognitive architectures, Experiential learning
2019	Arthur Franz, Victoria Gogulya, Michael Löffler	WILLIAM: A Monolithic Approach to AGI	Solution proposal	Cognitive architectures, Universal AI, Recursive self-Improvement
2019	Patrick Andersson, Anton Strandman, Claes Strannegård	Exploration Strategies for Homeostatic Agents	Solution proposal	Nature-inspired approaches, Reinforcement learning
2019	Qiong Huang, Doya Kenji	An Experimental Study of Emergence of Communication of Reinforcement Learning Agents	Solution proposal	Multi-agent systems, Reinforcement learning
2019	Nil Geisweiller, Ben Goertzel	An Inferential Approach to Mining Surprising Patterns in Hypergraphs	Solution proposal, Validation research	Cognitive architectures, Reasoning and Inference
2019	Bill Power, Xiang Li, Pei Wang	Generalized Diagnostics with the Non-Axiomatic Reasoning System (NARS)	Solution proposal	Cognitive architectures
2019	Nadisha-Marie Aliman, Leon Kester, Peter Werkhoven, Roman Yampolskiy	Orthogonality-Based Disentanglement of Responsibilities for Ethical Intelligent Systems	Solution proposal, Philosophical paper	AI safety, Philosophical aspects

B Article count per country

Country	Article count
USA	36
Iceland	10
Netherlands	10
Australia	7
China	7
Canada	6
Sweden	6
Switzerland	6
Russia	6
Japan	5
Austria	5
UK	4
France	4
Germany	4
Ukraine	4
Singapore	2
Czech Republic	2
Turkey	2
Italy	2
Iran	1
Hungary	1
Ethiopia	1
Belgium	1
None (Microsoft)	1