

**How discursive ideas could shape the future of artificial intelligence
in international relations: An English School analysis**



BA in International Studies – Spring 2024

Elias Eland Kristensen – 73723
Supervisor: Michelle Pace

Number of characters: 80.958
Harvard citation system

If we don't change direction soon, we'll end up where we're going – Irwin Corey.

Abstract

What are the discursive ideas around emerging artificial intelligence in international relations, and what could these mean for the future? Through a combination of quantitative topic modeling and manual qualitative discourse analysis, the project analyzes 55 articles from the epistemically influential *Foreign Affairs Magazine*. The project identifies five distinct discursive ideas around emerging AI and IR. By applying the English School theory, the project assesses the normativity of the discursive ideas, which, following the project's constructivist philosophy of science, could shape the future of international society. The project argues that most of the discursive ideas point towards a bifurcation or dissolution of the current nascent global governance efforts around AI. Furthermore, the project finds that the ideals of order and zero-sum power dominate the ideals around global justice.

Abbreviations

- 4IR: Fourth industrial revolution
- AI: Artificial intelligence
- AIT: Artificial intelligence technology
- CS: Computer science
- DARPA: Defense Advanced Research Projects Agency
- ES: English School
- GPAI: Global Partnership on Artificial Intelligence
- IO: International organizations
- IR: International relations
- LDA: Latent Dirichlet Allocation

Contents

Introduction	4
<i>Research puzzle and question</i>	4
<i>Research strategy</i>	5
Identifying and defining the project's key concepts	6
<i>Artificial intelligence</i>	6
<i>Constructivism and discursive ideas</i>	7
Literature review	8
<i>Artificial intelligence in international relations</i>	8
<i>Discourse, international relations, and artificial intelligence</i>	10
Methodology – Mixed method discourse analysis	11
<i>Using Foreign Affairs articles as empirical sources</i>	12
<i>Data collection</i>	12
<i>Textual Corpus</i>	13
<i>Topic model analysis – Latent Dirichlet Allocation</i>	14
<i>Qualitative discourse analysis</i>	15
Theory – the English School	16
<i>International society as a dynamic entity</i>	17
<i>Variations in international society – between pluralist order and solidarist justice</i>	17
<i>Application of the English School concepts</i>	18
The current structure of global AI governance	19
Analysis	20
<i>Identifying and interpreting the discursive ideas</i>	20
<i>Summary of the analytical findings</i>	28
Analytical discussion – The future of international society?	29
Conclusion	31
Bibliography	31
List of appendices	35

Introduction

Within recent years, the world has experienced great advancements within artificial intelligence (AI), impacting people's lives and the world as we know it (Kissinger, Schmidt & Huttenlocher 2022). One area highly affected by emerging AI is international relations (IR). A 2023 report from the geopolitical thinktank the RAND Corporation, stated: "*As AI continues to advance geopolitics may never be the same*" (Pavel et al. 2023:2). Advancements in cyberwarfare capabilities, the emergence of autonomous weapons systems and deepfake enabled misinformation have the potential to upend global power dynamics (Cummings et al. 2018). Meanwhile, advancements in AI and use-case artificial intelligence technologies (AIT) are projected to contribute \$25.6 trillion to the global economy, thereby supercharging economic competition (Lazard Geopolitical Advisory 2023:4). With these possible AI related changes within IR as the motivation, the following section lays out the project's research puzzle and question.

Research puzzle and question

With AI poised to change key aspects of contemporary IR, multiple groups of political actors like politicians, experts, and the mass media, have presented a plethora of ideas around how best to approach the coming wave of AI (Suleyman & Bhaskar 2023; Burton 2023). Therefore, the project seeks to explore the emerging intersection of IR, AI, and discourse. Asking, what characterizes the discourse around AI and IR, and what could these mean for the possible futures of IR (Ndzendze & Marwala 2023)? In adopting this analytical focus, the project moves beyond trying to analyze the more materialistic and tangible aspects of AI. In doing so, the project heeds the call from constructivist-oriented IR scholars, to explore the equally important and underdeveloped discursive dimensions of AI in IR (Burton 2023; Gu 2023).

To highlight why discursive ideas matter when talking about AI and IR, the project adopts a thick constructivist understanding around the role of discourses in society. Therefore, the project moves beyond understanding discursive ideas as a mirror to reality and instead acknowledges their potential to constitute reality (Bareis & Katzenbach 2022). By acknowledging this potential power of discursive ideas, the current ideas around AI in IR becomes a key subject of study, as these hold the power to shape the future of IR. This understanding of ideational power is summarized within the context of AI by Henry Kissinger et al. who in their book *The Age of AI* (2022)

write: “*While the advancement of AI may be inevitable its ultimate destination is not*” (Kissinger, Schmidt & Huttenlocher 2022:15).

Against this backdrop, and to analyze and interpret the current discursive ideas around AI and IR, the project analyzes articles on this exact subject from the influential magazine *Foreign Affairs*. Furthermore, the project uses the English School (ES) theory, as it provides a framework for assessing what the discursive ideas could mean for the future of international society. Summarizing the introductory considerations, the project poses the following research question:

- What characterizes the current discursive ideas around AI in IR, and what could these mean for states and the international society’s future approach to AI?

Research strategy

The research question encompasses two overarching and interconnected dimensions. The first dimension is focused on analyzing the characteristics and substance of the different discursive ideas. The second dimension builds on this and analyses the ideas’ possible implications for state behavior and the structure of international society. As the second dimension is dependent on the first, the first dimension should be understood as the primary part of the research question, followed by a supplementary, but still highly relevant, part.

Furthermore, as the research question encompasses two analytical dimensions, the project also contains two parallel research objectives. These objectives help qualify the later choices made in the project, clarifying what dimension of the research question they primarily address. The first research objective is descriptive, as it covers the part of the research question focused on characterization. This objective will be addressed, among other things, through the project’s methodological design that seeks to quantitatively identify discursive commonalities across the *Foreign Affairs* articles.

Following the descriptive objective, the second research objective is more interpretive in nature. This research objective echoes the project’s constructivist philosophy of science, as discursive ideas are understood as having a huge potential for societal change. To examine this, the project draws upon the ES, as it provides a theoretical framework, for interpreting how the different ideas may influence international society.

Identifying and defining the project's key concepts

In answering the research question above, the project briefly provides a clarification of its central concepts. This is necessary, as the research question's concepts are widely contested within the academic literature (Bostrom 2014; Schmidt 2008). Furthermore, clarifying the project's understanding of these concepts, helps to specify the project's scope and direction.

Artificial intelligence

Establishing a definition of AI, relevant to all aspects of AI research is an insurmountable task, therefore selecting a definition often steers the research in a certain direction (Krafft et al. 2019). With recent advancements in AI, Alan Turing's traditional benchmark test for determining machine intelligence has been beaten, with some AIs succeeding in imitating humans (Bostrom 2014). This has accelerated the computer science (CS) debate, around what constitutes an intelligent system and thereby how best to define AI (Wang 2019). However, as the exact philosophical understanding of computational/artificial intelligence isn't crucial to the project's analysis, further engagement with this debate is avoided, by adopting Pei Wang's common definition as a baseline: *"Intelligence is the capacity of an information-processing system to adapt to its environment while operating with insufficient knowledge and resources"* (Ibid:17).

The decision to delimit the project from the technical CS literature places the project within what the EU's Joint Research Centre calls the transversal taxonomy of AI (Samoili et al. 2020). This establishes an understanding of AI that, while grounded in the larger CS debate, is more suitable for social science research. This is because the primary focus of the transversal taxonomy isn't the actual AIs, but their derived technologies (referred to as AITs) and their possible impact on society, which is more in line with the projects research objectives (Samoili et al. 2020; Crawford 2021).

Furthermore, by delimiting the project from the more technical literature, it avoids engaging with the futurist debate around artificial general intelligence (AGI), argued as the likely far-future development of AI (Bostrom 2014). This delimits the project temporally and ensures a focus on the near-term future, which fits better with the general IR literature, which is rarely focused on the far future (Bilgin 2016).

Constructivism and discursive ideas

It is important to note that the understanding of AI, as outlined above, primarily serves as a practical necessity to narrow the project's scope. Therefore, while using the definition of AI above, the project acknowledges that AI as a term is fluid and highly contested within the political discourse (Samoili et al. 2020). This observation underlines the project's thick constructivist philosophy of science. Understanding IR through a constructivist philosophy of science, results in ideas becoming a highly relevant subject of analysis, as they are highlighted as a driving force behind changes in international politics (Ndzendze & Marwala 2023).

The primary reason for adopting the thick version of constructivism goes back to the project's underlying understanding of power. Thick constructivism is well suited for analyzing normative clashes between competing discursive ideas, in the case of this project around how to best approach AI in IR (Kratochvíl & Tulmets 2017). Thus, power, in the context of this project, is when discursive ideas become embedded, thereby becoming constitutive of political outcomes (Schmidt 2008). Lastly, adopting a thick constructivist philosophy of science harmonizes with the project's later use of the ES, which, as will later be shown, also highlights the importance of ideas when theorizing about the future of IR. Furthermore, the analytical emphasis on discourse, places the project within the European tradition of IR constructivism, where linguistic discourse is the primary constructivist factor (Ndzendze & Marwala 2023).

While discursive ideas are highlighted as important in constructivist IR, it is seldom clarified what to look for when analyzing them (Ibid.). Therefore, an early operationalized framework for what constitutes a discursive idea is necessary.

The project understands discursive ideas as being articulated through four components. The first two components are the setting and characters: Who are articulated as the main characters (for example states, individuals, or IO) and what characterizes the setting these exists within? Third, is the political narrative, which links the characters and setting together, thereby defining the nature of their internal relationships. Lastly, an idea is predicated on a set of underlying normative ideals, which are often conveyed through policy proposals or characterizations of best practices (Schmidt 2008; Guenduez & Mettler 2022). As normativity is difficult to analyze on its own, the project uses the theory of the ES as a means of analysis.

Literature review

The following sections contains the project's literature review. To better understand the nuances within the literature, the review first presents the overarching IR approaches to analyzing AI. Subsequently, the review will zoom in on the literary juncture of discursive ideas, AI, and IR, to identify common approaches, literary commonalities and to illuminate possible epistemic gaps in the literature.

Artificial intelligence in international relations

With the recent developments in AI, highlighted in the project's introduction, renewed interest has propelled AI/AITs into the center of tech-oriented IR. This renewed interest is illustrated in Henry Kissinger's seminal article *How the Enlightenment Ends* in *The Atlantic* (2018). In it, Kissinger argues that human society is unprepared for the emergence of AI, and he further argues for giving AI priority within the IR research (Kissinger 2018). Kissinger's insights were substantiated by a Chatham House report from the same year, systematically assessing key pillars of IR, namely warfare, diplomacy, security, and the global economy, concluding that all of these will be upended by AI (Cummings et al. 2018). Some IR scholars take it a step further, and argue that the current emergence of general-purpose technologies, like AI, quantum computing, big data, and the internet of things merit the analogy to a fourth industrial revolution (4IR) (Schwab 2016; Philbeck & Davis 2018). And, as with previous industrial revolutions, the 4IR has the potential to redistribute and upend global power dynamics (Pavel et al. 2023).

Within the IR literature, the studies of emerging AI are highly influenced by the field's established theories (Ndzendze & Marwala 2023). Thereby, to get an overview of the literature inspired by the possibility for change outlined above, one needs to understand how realism, liberalism and the critical theories approach the topic.

Central to the realist literature on AI is the concept of power. The understanding of power within realism diverges from this project's constructivist understanding. Where this project highlights the ideational dimensions of power, realism argues that only material capabilities translate to power. (Mearsheimer 2021). The emerging potential for military AI power is explored through the joint realist concepts of AI balance of power and AI supremacy/sovereignty (Ndzendze & Marwala 2023). As AI development accelerates, derived military AITs emerge, which have the potential to change warfare and further upend military power (Horowitz 2019). Therefore, realist

scholars have begun to view AI innovation as an important strategic capability, with states trying to achieve either AI supremacy or AI sovereignty (depending on great power status) (Regan & Davidovic 2023; Pavel et al. 2023). Realist scholars have also highlighted that the logic AI supremacy and power is visible in the early history of AI development, with institutions like the US's Defense Advanced Research Projects Agency (DARPA) playing a key role in advancing foundational technologies like machine learning and image recognition (Crawford 2021; Fouse, Cross & Lapin 2020).

Central to the liberalist literature on AI in IR is the Janus-faced nature of enlightenment modernity, which AI is seen as being a part of. As John Ikenberry writes: “(...) *modernity is a continuous process, global in scope, which generates a stream of opportunities and dangers. In this modernizing world, the human condition can be improved through cooperation in the pursuit of enlightened interests*” (Ikenberry 2020:33). With this understanding of AI, the key challenges identified by liberalist IR scholars are not about state power, but instead how to maximize the upsides and mitigate the downsides around emerging AI through international governance. Thereby, expanding beyond state-centric analysis and encompassing international organizations (IO) as analytical subjects (Tallberg et al. 2023). Following the possible negative consequences of emerging AI, another key challenge identified by liberalist scholars is AI's influence on democracy. Some scholars argue that if not governed correctly, AI may end up undermining democracy and strengthening authoritarianism through misinformation and public manipulation (Roberts et al. 2021). Thereby, challenging the Kantian underpinning for the liberal peace thesis, that posits that democracies don't engage in war with other democracies (Ikenberry 2020).

Analyzing AI as a threat to global stability and justice is also at the nexus of the critical IR theories. Central within contemporary critical IR literature is Kate Crawford's magnum opus *Atlas of AI* (2021). In her book, Crawford identifies several problematic areas surrounding the current AI development and implementation. These areas range from unsustainable resource extraction to problems surrounding the current use of AITs (Crawford 2021). As later research has also shown, states and companies in the global North extract non-consensual training data, rely on cheap labor for micro-tasks, and extract rare earths minerals from countries in the global South (Mohamed, Png & Isaac 2020). While the critical IR literature is primarily focused on the global in-

justices outlined above, it is worth noting that these issues not only exist on the state level but have also been identified across both gender and race (D’Ignazio & Klein 2020; Benjamin 2019).

The review of the traditional IR theories’ approach to AI shows that these schools still hold analytical and theoretical value. To acknowledge this and to be able to engage with the mainstream IR literature outlined above, the project employs the ES theory. As will be shown later, the ES is able to engage with both the realist state-centrism and the liberalist focus on IOs and global governance, while still emphasizing the importance of discursive ideas. Thereby staying true to the project’s thick constructivist philosophy of science.

Discourse, international relations, and artificial intelligence

The second part of the literature review is centered around articles sharing the research topic of this project, namely the intersection of AI/AITs, IR, and discourse (Appendix 2). To identify epistemic gaps and assess the general approaches and subject matter, the project sorted the identified articles based upon their analytical subject. As shown in Table 1, four overarching discursive analytical subjects were identified within the literature. While their subject matters differ, most articles share a similar constructivist philosophy of science. This is captured by Ali Guenduez and Tobias Mettler who argue that: “(...) *how AI is framed and communicated in policies is key to any subsequent implementation*” (Guenduez & Mettler 2022).

Table 1: Variations in the analytical subject of articles analyzing AI, IR, and discourse

X = Explicit focus

(X) = Secondary/implicit focus

	<i>Governmental discourse</i>	<i>AI developers and companies</i>	<i>Mass media and public discourse</i>	<i>Epistemic communities</i>
<i>Analytical subject</i>	70% (10%)	20% (20%)	30% (20%)	10% (20%)

Source: Table 2 in appendix 2

With AI emerging as an important technology, an increasing number of countries have begun to develop AI strategies (Roberts et al. 2023). This has been accompanied by an increase in academic articles analyzing commonalities and differences within the different states’ strategies (what Table 1 calls governmental discourse) (Gu 2023; Guenduez & Mettler 2022; Radu 2021).

One article finds that the Nordic states’ AI strategies deemphasize strategic competition and instead focus on values like trust, transparency, and openness (Robinson 2022). This ideal is also identified among other European states by Jascha Bareis and Christian Katzenbach, who identify

a clear rift between the liberal and humanist oriented legislation of France and Germany versus the realist-oriented legislation being implemented by the US and China (Bareis & Katzenbach 2022). This rift is also identified by Ali Guenduez and Tobias Mettler who, in their large comparative study of 35 national AI strategies, identify the different roles government can take. Great powers like China and the US, but also smaller states like the UAE, Israel, or Taiwan, discursively present themselves as leaders or enablers when it comes to AI, where many European states, present a more regulatory role (Guenduez & Mettler 2022).

While focusing on governmental discourses, often through the analysis of national AI strategies, some of the article's hint at the underlying importance of epistemic communities in shaping both policy outcomes and discursive ideas more broadly (Gu 2023; Roberts et al. 2023). Meanwhile, as shown by Table 1, analyzing the discourse of the epistemic communities, consisting of AI experts and scholars, is a somewhat overlooked subject, when compared to the discourses of the mass media or AI developers (Ding & Kong 2019). The only author the project was able to identify who explicitly analyzed the epistemic communities' discursive ideas was Hendrick Schopmans (2022). While this project delimited itself from the more technical literature by staying within the transversal taxonomy, Schopman's article does the opposite. He identifies the different discursive within the CS epistemic community, which makes it a highly technically oriented paper (Schopmans 2022). Thereby, by focusing more on the possible societal impact of the foreign policy epistemic community's discursive ideas, the project should contribute with new insights to the literature on AI, IR, and discourse.

Methodology – Mixed method discourse analysis

In analyzing what characterizes the different discursive ideas around AI within IR, the project applies a sequential mixed method approach (Creswell 2003). The project combines quantitative textual analysis through Latent Dirichlet Allocation (LDA) topic modelling with a manual qualitative discourse analysis (Jacobs & Tshötschel 2019). However, before presenting the methodological considerations, the section below briefly explores the benefits and possible pitfalls around using *Foreign Affairs* Magazine articles as empirical sources.

Using Foreign Affairs articles as empirical sources

The decision to analyze *Foreign Affairs* articles draws on insights gained through the project's literature review. As shown, multiple authors have already explored the discursive ideas of AI in IR through governmental legislation and mass media, leaving the ideas of epistemic communities as a somewhat neglected subject. However, while being an underexplored area academically, the epistemic communities' importance is clear, as some contemporary AI policies are influenced by various epistemic communities (Gu 2023).

One of the key challenges within the first research objective is to identify a collection of texts that can illuminate the discursive ideas around AI in IR. *Foreign Affairs* articles are one of the few large-scale collections of texts in English that explicitly tackle both IR and AI, while stemming from an important epistemic community. Since its inception in 1922, *Foreign Affairs* has been one of the premier publications within global politics. However, it is important to note the magazine's Western bias historically, as this influences the project's later conclusion (Foreign Affairs 2024). Despite its bias, *Foreign Affairs* is a central forum for discussing emerging trends in IR, making it a great source for analyzing how influential IR thinkers ideate emerging AI. In summary, using *Foreign Affairs* articles as empirical sources contributes to the identified research gap around the discursive ideas of the epistemic communities operating within the transversal taxonomy of AI.

Data collection

The following section builds upon the use of *Foreign Affairs* articles as empirical source and explores the considerations and selection criteria of the individual articles comprising the project's empirical data. As *Foreign Affairs* as a whole is about different facets of IR, the key determinant for text selection was a focus on the topic of AI. To identify such articles, the project used *Foreign Affairs*' own keyword search function. By selecting 'artificial intelligence' as a keyword, the website provides the reader with all articles marked by their contributing authors as being explicitly on AI, ideally sorting out articles that only passingly mention AI (Ibid.). Thereby ensuring that the selected texts contain discursive ideas around IR and AI.

Due to the project's preliminary quantitative textual analysis all of the identified articles were selected as empirical sources. This should ensure that the project's data captures the highest degree of width in the discursive ideas and minimizes possible selection bias (Isoaho, Gritsenko &

Makela 2021). Furthermore, prioritizing width early on in the methodological process foreshadows a second quantitatively enabled systematic selection, which will be explored later in this section.

The broad preliminary selection yielded 59 separate texts, spanning from January 2015 to February 2024. However, while 59 texts were initially identified, an early presorting resulted in the elimination of four texts. These texts were omitted, as they weren't actual articles, but instead brief advertisements for *Foreign Affairs*' podcast, making their content irrelevant. After presorting, the project ended up with 55 separate articles consisting of 921.083 tokens.

Textual Corpus

The next methodological step, following the selection of articles and prior to the quantitative analysis, was the creation of a textual corpus. Creating a textual corpus is the process of turning the selected articles into data, which can function as an input for a computer-based analysis. The textual corpus was created following standard procedures, as laid out by Grimmer and Stewart (2013) and Benoit (2020).

Firstly, to conduct a computational analysis, the gathered texts must be compiled into a file-format compatible with your analytical program of choice. In the case of this project, an excel file compatible with R-Studio, the software for quantitative analysis used in this project.

Secondly, to improve the quality of the textual corpus, the texts are tokenized, meaning you remove all irrelevant stop words and punctuation, as well as lowercasing all words (Appendix 1:34-47). This is done to increase analytical accuracy. For example, lowercasing the words ensures that a capital word in the beginning of a sentence is calculated the same way as a later identical noncapital word.

The last step to improve the quality of the analysis is to calculate the collocations of words. Calculating collocations is a way to force the software to view pairs or groups of connected words, that frequently appear together, as being one word (Ibid:49-56). Calculating collocations is especially crucial in the case of this project, as both 'artificial intelligence' and 'United States' gets their meaning through being collocated. Therefore, if collocations weren't calculated in the creation of the textual corpus, 'artificial intelligence' and 'United States' would appear as separate, which could result in the US being overlooked as an important word-pair in the later analysis.

Overlooking US would be problematic as it is one of the key characters within the IR of AI (Burton 2023).

Topic model analysis – Latent Dirichlet Allocation

In its sequential methodological strategy, the project initially conducted a quantitative textual analysis based upon LDA topic modeling. The following section will first explore the project's considerations around doing LDA topic modeling, as well as explain what topic modelling is. Lastly, the section will touch upon some of the key practical choices made as part of the quantitative analysis.

The overarching reason for using topic modelling as a methodology to explore the discursive ideas around AI and IR is that it's well-suited to illuminate commonalities across large amounts of textual data (Isoaho, Gritsenko & Makela 2021). Topic modeling, specifically Latent Dirichlet Allocation (LDA) which is the algorithm the project uses, works by assigning a pre-determined number of topics to the model. Then the model calculates which topics the most frequently used words belong to, based upon a probabilistic process. In other words, the LDA model assumes that each text contains multiple topics and calculates what topic best fits the text based upon the actual words of the text. This 'best fit' is achieved by the model running multiple iterations and presenting the best (Blei 2012).

The reasons for using topic modelling as a methodology, beyond the fact that it is great at identifying latent themes, can be condensed into two primary considerations. Firstly, to ensure that the analysis captures the full spectrum of discursive ideas in *Foreign Affairs*, the textual corpus has to be quite large, whereby conducting a similar analysis manually is pragmatically infeasible. Secondly, addressing the first more descriptive research objective, a topic model analysis like LDA is a suitable method for defining and superficially describing the different discursive ideas, as it provides an introductory overview of the most frequently used words within each topic.

The key methodological choices associated with a topic model analysis are also twofold. The first choice is selecting the topic model algorithm to use. As shown above, this project uses LDA as its algorithm. The reason for choosing LDA is that it is the simplest topic modelling algorithm, and since the project doesn't analyze metadata or different types of textual data, the project's research design is compatible with LDA (Guenduez & Mettler 2022; Blei 2012). The second important choice is selecting the predetermined number of topics that the model searches for. To

determine this, the project draws on methodological insights from the literature review, examining other authors who have adopted a similar quantitative analytical approach on a similar subject matter. These authors respectively identify five, and six different topics through their analyses of both national and transnational AI legislation, as well as mass media discourse on AI (Gu 2023; Guenduez & Mettler 2022; Ding & Kong 2019). This hints at a correlation in the number of discursive ideas around AI and IR. Therefore, after trying four and six topics and getting worse results, the project decided on five topics as the input for the LDA-analysis (Appendix 1:65-69). The following table contains the output of the five-topic LDA-analysis, listing the 15 most frequent and relevant words for each of the topics.

Table 2: The 15 most frequent and relevant words within each identified topic

<i>Topic 1</i>	<i>Topic 2</i>	<i>Topic 3</i>	<i>Topic 4</i>	<i>Topic 5</i>
AI	China	Countries	AI	Military
Digital	United States	Many	Data	AI
Governments	Chinese	Technology	Can	U.S
Information	China's	Workers	Human	Nuclear
International	U.S	New	People	United States
New	Technology	Services	One	Intelligence
Public	Beijing	People	Even	Defense
Risks	Manufacturing	Jobs	World	Systems
Use	Government	Technological	Systems	New
Control	Innovation	Percent	Machines	War
Global	Technologies	Also	Machine Learning	Washington
Including	AI	World	Power	Used
Deepfakes	Industry	Economy	World	States
Disinformation	Chips	Labor	Make	Ukraine
Surveilnace	Advantage	Trade	Robots	Operations

It is important to note that a quantitative analysis can't meet the interpretive research objective of the project alone. Reiterating the project's sequential methodological design, topic modeling serves as a great foundation for further qualitative analysis (Blei 2012). The LDA-model is able to identify the key texts within each of the different topics, which can then be analyzed manually (Appendix 1:74-119). Therefore, besides providing an overview, the LDA analysis also provides a systematic way to select the most relevant text, which again goes back to why the broadest array of texts was selected from *Foreign Affairs* initially.

Qualitative discourse analysis

To determine which texts are most relevant for further analysis, the project identified the three texts from each topic with the highest correlation score (Ibid:74-119). The topic correlation score specifies what percentage of an article's words fit within a given topic. This selection process

ensures that the further analysis is conducted on the most important texts within each topic, thereby making the process of identifying and analyzing the discursive idea within each topic easier and more accurate (Isoaho, Gritsenko & Makela 2021).

Table 3: The most relevant texts within topic

<i>Topic</i>	<i>Title</i>	<i>Author(s)</i>	<i>Year</i>	<i>Topic correlation</i>
Topic 1	Defending the Year of Democracy	Duffy, Kat & Harbath, Katie	2024	71,5%
	The Premature Quest for International AI Cooperation	Schaake, Marietje	2023	68,7%
Topic 2	The Race to Regulate Artificial Intelligence	Bradford, Anu	2023	63,4%
	China’s Sputnik Moment?	Wang, Dan	2021	67,2%
	China’s Hidden Tech Revolution	Wang, Dan	2023	60,2%
Topic 3	America Can’t Win the Tech Race Alone	Thomas, Christopher & Kreps, Sarah	2024	58%
	The Global Economy’s Next Winners	Lund, Susan; Manyika, James & Spence, Micheal	2019	70,1%
	People Over Robots	Pritchett, Lant	2023	63,4%
Topic 4	The Coming AI Economic Revolution	Manyika, James & Spence, Micheal	2023	44,1%
	She, Robot	Greiner, Helen	2015	77,7%
	Ready for Robots?	Cukier, Kenneth	2019	71,7%
Topic 5	Tech World	Drum, Kevin	2018	61,7%
	How Ukraine Is Remaking War	Kahn, Lauren	2022	54,8%
	How to Compete in Cyberspace	Nakasone, Paul & Sulmeyer, Micheal	2020	53,7%
	A Force for the Future	Horowitz, Micheal; Kahn, Lauren & Samotin, Laura	2020	51,1%

By conducting a manual discourse analysis, the specific context of the identified words can be understood, providing a more nuanced and ideally complete picture of the different discursive ideas (Jacobs & Tshötschel 2019). To describe and interpret the discursive ideas, the project draws on the previously operationalized understanding of the four underlying components of a discursive idea, which serve as the key pillars of the project’s analysis. Namely, the characters, the setting, the narrative that links together the setting and characters, and the ideas underlying normativity. Lastly, to better understand the underlying normativity and the ideas’ possible implications for states and the international society, the analysis draws upon the ES.

Theory – the English School

The following section outlines the project’s use of the ES theory. As previously argued, the reason for using the ES, is that it is able to encompass both the realist state-centrism and the liberalist focus on global governance, while aligning with the project’s constructivist position. In using

the ES, the project addresses the second research objective and interpret the discursive ideas' possible implications.

International society as a dynamic entity

The ES emerged in the 1950s with notable authors like Hedley Bull and Martin Wight. These authors sought to break with, what they viewed as, the false dichotomy between realism and liberalism (Dunne 2021). Central to the ES understanding of IR is the concept of the ever-present international society (Buzan 2004). Although understood as being ever-present, the exact nature and structure of the international society are not constant (Wendt 1992). Where structural IR theorists argue that states as rational agents explain international corporation, the ES argues that the inherent nature of international society is dynamic, changing as different ideas emerge and disappear (Buzan 2004; Ikenberry 2020; Mearsheimer 2021). Returning to the project's research question, this assumption underlines the importance of understanding the ideas around AI in IR, as they may come to shape international society.

It is important to note that while authors like Bull and Wight analyze international society as encompassing all aspects of inter-state relations, recent scholars have narrowed the theory's analytical scope, by applying it to more narrow aspects of inter-state relations (Buzan 2004). This narrowed analytical scope helps clarify the project's use of the ES, as it primarily focusses on the global governance around AI rather than global governance in general.

Variations in international society – between pluralist order and solidarist justice

To analyze the variations in international society the ES presents a set of theoretical terms that conceptualize international society as a spectrum, with ideas as the driving force for developments in either direction. This spectrum exists between what the ES labels the pluralist and solidarist positions (Ibid.).

At the pluralist end of the spectrum lies what the ES calls the international system of states. The international system parallels the realist/neo-realist assumptions, as it is characterized by states pursuing their own interests and power being the main currency within a world defined by anarchy (Dunne 2021). Drawing on the literature review, concepts like AI supremacy and the AI balance of power are staples of an international system. While realists understand the balance of power and similar concepts as universal, the ES and constructivist scholars argue that these are only ideas that have embedded themselves within IR (Wendt 1992). Building on the realist as-

sumptions underscoring the pluralist position, is the normative ideal of order through sovereignty (Buzan 2004).

At the other end of the spectrum is the solidarist position, which argues for a very high degree of cooperation and joint engagement among states. Challenging the pluralist idea of an universally anarchic world, the solidarist position mirrors the liberalist theory, which posits that shared interest among states can overcome the anarchical state through the establishment of international institutions (Ibid.). These institutions ensure collective enforcement of international rules and norms, as well as the universal protection of human rights and global equality. By emphasizing the importance of global rights and equality the solidarist position's normative ideal of global justice trumps the pluralist ideal of order and sovereignty. The values of the solidarist position, is captured within the ES concept of world society (Dunne 2021).

Between the two positions of solidarism and pluralism, the ES presents the concept of international society, which serves as a middle ground where a lesser degree of international engagement is present to ensure some degree of global justice, while the states still retain a high degree of sovereignty (Buzan 2004).

Application of the English School concepts

The key argument for including the ES is that it provides the project with a spectrum of concepts through which to interpret the potential implications of the different identified discursive ideas. In doing so, the concepts of the international system, the international society, and a world society, each with their varying emphasis on the normative ideals of order and justice, serve as idealistic examples of how IR around AI could look. Through these, the project can analyze how the different discursive ideas around AI in IR in *Foreign Affairs* intersect with either of the ES positions. Analyzing not only how the different discursive ideas depict international society, but also, in line with the project's constructivist position, what direction these ideas could push international society in. While the ES provides a normative frame of reference for the analysis, combining it with a constructivist philosophy of science, helps articulate the social nature of political outcomes within IR (Reus-Smit 2002).

The current structure of global AI governance

Before proceeding with the analysis, and to better contextualize the possible impacts of the discursive ideas, the project briefly outlines the current state of AI governance. In doing so, the project establishes an analytical point of departure, from which possible changes can be understood and compared.

To account for the current state of global AI governance, the project draws on insights from the literature review, as well as secondary academic sources. While this section outlines the current structure of global governance, it remains predominantly agnostic regarding the actual contents of the different initiatives, as these would be secondary to the research objectives. Similarly, the structure of global engagement among companies is omitted, as the project's primary focus is on inter-state relations.

As the second part of the literature review shows, contemporary AI governance is largely characterized by the emergence of national AI strategies. Therefore, it's clear that states have begun to approach AI as a policy area to be regulated/governed. However, while many states have deployed their own national AI strategies, often with parallel ambitions around international engagement, the current structure of global AI governance is nascent and broadly fragmented (Butcher & Beridze 2019).

The current global governance around AI is predominantly polycentric, with two primary centers of governance emerging. The first center of power emerging is the UN, with Secretary General António Guterres highlighting AI/AITs as important areas for the UN's efforts. The UN's approach to AI primarily takes place within its well-established IOs and structures. As an example, the UN Interregional Crime and Justice Research Institute (UNICRI) opened a center for AI in Robotics in 2017. A similar effort, under the UN umbrella, was launched in 2017 when the International Telecommunications Union (ITU) hosted their first annual *AI for Good Global Summit*. This summit centered on how a multitude of global stakeholders can align emerging AITs with the UN's sustainable development goals. At the same time, the UN is tackling the military aspects of AI, with the UN Office for Disarmament Affairs, suggesting restrictions on the use of battlefield AITs (Ibid.)

Where the UN incorporates AI into its existing institutional governance structure, the second center for global AI governance has emerged more recently. A key new initiative within contemporary AI governance is the G7's 2017 Global Partnership on Artificial Intelligence (GPAI). Despite originating within the G7, the GPAI quickly expanded, as more states joined, and it is currently existing as part of the OECD secretariat. Almost simultaneously, in 2019, the OECD adopted its own set of parallel AI principles (the OECD principles on AI). Recently, both the OECD principles and the GPAI was endorsed by the G20, unifying China, Russia, and the US in a structure of joint commitments around the development of AI. While the OECD principles and the GPAI currently hold wide support among important state actors within IR, they don't hold any actual regulatory power, and must therefore rely on soft power and non-binding guidelines (Schmitt 2022).

Currently, the only binding non-national example of AI governance is within the EU. While not in effect yet, the EU's AI Act is highlighted as the most comprehensive regulatory proposals for AI (Larsen 2022). With the regulatory efforts of the UN and the OECD being non-mandatory, the EU's extensive regulatory regime seems to be an outlier when assessing the current structure of global governance. However, while being nascent and fragmented, recent expansions in the structure of global governance have demonstrated that the field is rapidly developing as states have shown a willingness to engage (Butcher & Beridze 2019).

Analysis

The following section will analyze the different discursive ideas around AI as well as interpret what they could mean for international society, the current structure of which was outlined above. Through the preliminary quantitative topic model analysis and the subsequent manual analysis five competing discursive ideas around AI in IR were identified, each of which (corresponding to the identified LDA topics) will be analyzed and discussed in detail.

Identifying and interpreting the discursive ideas

Discursive idea 1: Global governance and regulatory power

The characters of the first discursive idea can be divided into two primary groups. The first group, comprised of state governments, is identified through the quantitative analysis, with

“governments” being the third most frequent word (Table 2). Along with governments, the second group of characters, are the big trans-national technology companies and their user-platforms: “(...) such as Facebook, Google, Instagram, Telegram, TikTok, WhatsApp and YouTube” (Duffy & Harbath 2024:3). While the companies and their platforms are mentioned as important characters, they are also portrayed as enabling secondary hostile actors through their integration of AI (Bradford 2023:9).

This mirrors the idea’s setting, which is defined by rising availability of powerful generative AITs that could threaten democratic governments and elections. Thereby, the setting is characterized through emerging “digital” “risks” (Table 2). In parallel to these emerging risks, the setting highlights 2024 as a critical year for preserving/defending democracy globally, with several key democracies like the European Union, the United Kingdom, the US, and India holding elections (Duffy & Harbath 2024:8).

This latent tension between technology and democracy embedded in the setting, is what comprises the primary narrative of the discursive idea. Some of the frequently used words, are words associated with the negative dimensions of AI: “deepfakes”, “disinformation” and “surveillance” (Table 2). These exemplify the narrative that AITs, developed and distributed by global tech companies, are being utilized by hostile secondary actors to undermine democratic governments globally. Furthermore, the narrative moves beyond the threat to democracies and argues that emerging AITs could also threaten economic security, civil liberties, and human rights around the world (Schaake 2023).

Built on the narrative above, the discursive idea’s immediate normative imperative is one of regulation and monitoring when it comes to AI: “All have recognized the urgent need for government regulation that ensures AI applications operate within the confines of the law and that safeguards national security, human rights, and economic competition” (Ibid:2). Within the theory of the ES, the normative imperative of global regulation is traditionally associated with the concept of world society and its emphasis on global justice. However, the exact normative prescriptions of the first discursive idea diverge substantially from the ES notion of a world society, by deemphasizing justice as a normative ideal.

This deemphasis on global justice is articulated through what the idea calls: “The race to regulate Artificial intelligence” (Bradford 2023:1). Here the normative call for regulating AI, isn’t

grounded in the ideal of justice, but instead that regulation is yet another medium for achieving state power. Thereby, if the US wants to maintain the current order, it needs to establish a competing regulatory regime: *“As the EU and China lead the race to regulate AI, Washington needs to decide whether it wants to have a role in building the digital world of the future”* (Ibid:8). The quote above is accompanied by a clear characterization of global AI governance as a zero-sum game (Ibid:9-10). This widely contradicts the notion of a world society, where the US and China would pursue a unified regulatory framework aimed at strengthening justice globally.

Comparing the above to the current structure of global AI governance, the US and China’s joint engagement through the OECD, shouldn’t necessarily be viewed as an attempt to achieve sustainable or just AI development, but instead a push for regulatory power. This argument is further collaborated by the normativity of the discursive idea, which highlight the primacy of national regulation over what the UN and the OECD is currently pursuing: *“Political leaders should first hammer out the preconditions and content of those laws [national AI regulation] – and only then fit agencies to oversee regulation”* (Schaake 2023:3-4).

All in all, the characters, setting and narrative all point towards a strong normative imperative of AI regulation, which is traditionally associated with the development of a world society. However, when analyzing the idea’s actual normative suggestions, it is clear that justice is deemphasized, pointing more towards an international society, where the steering principle isn’t justice but order.

Discursive idea 2: Great power competition

When interpreting the most frequent words summarized in Table 2, the characters of the second discursive idea are clear, as six of the words are associated with either the US (*“United States”* and *“U.S.”*) or China (*“China”*, *“Chinese”*, *“China’s”*, and *“Beijing”*). Furthermore, the frequent use of a word like *“government”* indicates that the characters are not merely the US and China, but their respective governments. This observation is collaborated by the content of the three most relevant articles, which center around the two countries’ policies around AI and technology in general (Wang 2021; Wang 2023). While the articles mention other states and companies, the main agency is retained for the US and China: *“The United States and China are such dominant players in this field, too [the field of AI], that public- and private-sector technology actors else-*

where will need to pick a side (...)” (Thomas & Kreps 2024:4). Creating a binary ‘this or that’ effect, that goes against the unified ideal of a world society.

Following the above, states and companies beyond the US and China are portrayed as part of the discursive idea’s setting, rather than being meaningful characters. Besides being populated by malleable states, the setting is characterized by a zero-sum understanding of IR, presented through the logic, that US policy inaction will lead to Chinese dominance (Wang 2023:14). This echoes the realist concepts of anarchy and AI balance of power presented in the literature review.

Expanding on the competitive characters and setting, the discursive idea’s narrative is, that when it comes to AI and other emerging technologies (i.e. photovoltaics, 5G, drones, and semiconductor/chips) the US and China are locked into a self-perpetuating spiral of escalating competition. With China being characterized as an “*manufacturing superpower*” which threatens the US’ current technological supremacy (Ibid:8). Meanwhile, China’s emergence is also characterized as a response to the US’s anti-China economic policies: “*(...) the sweeping nature of the Trump administration’s sanctions did not suggest a careful selection process. Rather, they gave the impression that the United States would punish any Chinese company that achieved success*” (Wang 2021:10). Highlighting the aforementioned spiral of mutually created competition, which is central to the idea’s narrative.

With the narrative of self-perpetuation competition, the normative suggestions aren’t focused on lowering tensions, but instead for the US to accelerate its drive for strategic advantage. Both by increasing domestic capabilities when it comes to AI and technology (Wang 2023:4). But also, by engaging with other states and companies to create a technological ecosystem surpassing and even dominating the one China is characterized as creating (Thomas & Kreps 2024:8).

Interpreting the normative suggestions above through the concepts of the ES, the discursive idea highlights sovereignty and power as the clear objectives to be achieved by the US. However, interpreting the above as a true pluralist international system would be faulty, as the second discursive idea argues that the US needs engagement with its allies to maintain its’ lead over China (Ibid:5). Therefore, you could understand the normative ideal as a clearly subdivided international society, with one American and one Chinese led ecosystem around AI and technology. Such an approach would devastate the OECD’s current attempts at joint engagement, which have succeeded in uniting both China and the US under, the “OECD principles on AI”. Thereby, trying to

prevent a bifurcation structure of competing international societies. Furthermore, bifurcation could mean that the ideals of justice, somewhat present within the concept of international society, are reserved for actors within a given ecosystem. Potentially creating global inequalities in civil and human rights in the face of emerging AI.

Discursive idea 3: Ensuring justice in a changing global economy

Where the previous discursive ideas dealt with AI in a more general and societal sense, the third discursive idea focusses primarily on the economic aspects of emerging AI. This is illustrated in the frequent use of words like “workers”, “jobs”, “economy”, “labor”, and “trade” (Table 2). Besides highlighting an economic context, the frequent use of words like “workers” and “jobs” shows that one group of characters are the global working populous, a group the previous ideas have widely omitted. While the global populous exist as a clear character, the main agents are still arguably states and their governments, as these are portrayed as the key enactors of economic policies (Pritchett 2023).

The setting is one characterized by economic turmoil due to two underlying factors. Firstly, the global economy is argued to be experiencing a series of destabilizing economic conditions, primarily due to a deceleration in productivity growth. Secondly, further contributing to the economic turmoil, are unfavorable demographic conditions among large parts of the industrialized world. The discursive idea posits that these demographic conditions may result in that: “(...) by the 2040s, there will be millions too few native-born people in developed countries available to perform all the essential nonroutine, manual tasks in the economy” (Ibid:6). All in all, the setting portrays an economically bleak reality for the states and the global population.

The discursive idea’s narrative primarily focusses on what may happen when AI or AITs emerge within the economic reality that characterizes the setting. The discursive idea’s narrative posits that the policy currently accompanying emerging AI is characterized by protectionism and de-risking: “The era of building global supply chains entirely on the basis of efficiency and comparative advantage has clearly come to a close” (Manyika & Spence 2023:4). Where the previous discursive ideas posed their normative positions as a logical extension of their respective narratives, the third discursive idea takes the opposite approach in that its normativity opposes the laid-out narrative.

The third idea's normative assertion is that the current narrative of the world turning inward through rising protectionism and onshoring shouldn't be allowed to continue, as this would be an economic blunder: *"They are shutting themselves off from the outside world at the very moment when they should be welcoming it"* (Lund, Manyika & Spence 2019:2). Implicit in this quote is the assumption that to maximize the economic upsides of emerging AI globally, the industrialized countries can't turn their backs on globalization. However, the third discursive idea also acknowledges that this can't be done exclusively through the free markets' forces: *"Despite its enormous promise, AI is unlikely to trigger an economy-wide jump in productivity, or to support sustainable and inclusive growth, if its use is left to market forces"* (Manyika & Spence 2023:12-13).

Therefore, building on the normativity above, the third discursive idea posits a strengthening of the current global economic governance, especially when it comes to emerging AI (Ibid:15). This could be interpreted as a call to expand the structure of global AI governance already taking shape within the OECD and through the GPAI. One could argue that the normative position of the third discursive idea points toward a broadly solidarist world society, when it comes to the narrower economic governance of AI and AITs. This argument is collaborated by the idea's normative emphasis on global justice, which, to be exact, is limited to economic justice within this discursive idea. The idea posits that increased economic AI governance, doesn't merely serve as a vessel for reversing the current negative economic trends, but also as a tool to ensure increased economic justice globally, as it is argued that the two are interconnected: *"The extent to which AI can be developed and used in an equitable way worldwide will determine the magnitude of its effect on the global economy"*, (Ibid:16).

Discursive idea 4: Developing safe and ethical AI

Before analyzing the key texts within the fourth discursive idea, it is worth noting that the idea is somewhat of an outlier compared to the others. As shown in Table 3, the texts comprising the other identified discursive ideas are fairly new, with all but one being published after 2020. However, the texts with the highest correlation score within the fourth discursive idea were all published before 2019, which, when considering the blistering speed of recent advancements in AI, is practically ancient. This could suggest that some of the arguments within the discursive idea may have become somewhat obsolete in the face of recent technological advancements.

The key characters within the discursive idea are actors closely associated with the technical development of AI, for example “*technological entrepreneurs*” or “*venture capitalists*” (Rose 2015:5). This primary focus on characters engaged in technological development is collaborated by the general overview provided in Table 2. Here some of the most frequently used words are associated with the technical aspects of AI, like “*data*”, “*systems*”, “*machines*”, “*machine learning*”, and “*robots*”. However, as the fourth discursive idea has a more futurist focus than the previous ideas, AITs (especially robotics) and AI in general are portrayed as having a large degree of agency in their own right, challenging the state-centrism of traditional IR theories. This non-human agency is highlighted by quoting figures like Elon Musk, who describes emerging AI as: “*Summoning the demon*” (Cukier 2019:4).

The setting of the fourth discourse idea mirrors the scholars highlighted in the literature review who argue that due to advancements in technological capabilities, society is currently experiencing a new industrial revolution: “*Right now, the world is at the dawn of a second Industrial Revolution, this time a digital revolution*” (Drum 2018:3). Therefore, the setting is permeated by a huge potential for change in the near future.

The narrative of the discursive idea takes this potential for change and transforms it into an almost deterministic reality: “*You can't stop technology; the world's going to continue to move forward*” (Rose 2015:7). However, going back to Musk's quote outlining the emerging agency of AI, if technological progress results in the release of technological demons, the narrative of technological determinism is far from a positive one. This narrative of technology as an emerging threat is further captured in the rubric: “*Welcoming our new robot overlords*” (Drum 2018:5). Thus, the narrative portrays the idea's characters as being locked into a deterministic trajectory, which ultimately may end up posing an existential threat to humanity.

Whereas the previous ideas' normativity responded to threats from other human actors, for example competing states or hostile, the fourth idea instead responds to the non-human technological threat outlined above. As a result, the idea goes further on global governance and the establishment of a world society than the previous ideas. One such suggestion is a mandatory global licensing system for AI developers, which aim to ensure ethical and safe AI. Such a system would move far beyond the current state of non-mandatory AI governance within the OECD and UN. However, as is argued by the discursive idea, such a system only works if states are willing

to pool a huge part of their sovereignty (Cukier 2019:6). Thereby, following the world society's normative position of sacrificing order for justice.

Discursive idea 5: Global instability and emerging threats

Similar to previous ideas, the primary characters of the fifth discursive idea are states, and their governments. This is illustrated through the frequent appearance of words like: “*United States*”, “*Washington*” and “*Ukraine*” (Table 2). Furthermore, when qualitatively analyzing the texts of the discursive idea, the frequent use of the word *Ukraine* happens in connection with Russia's invasion, indicating that Russia should also be understood as a noteworthy character (Kahn 2022).

The setting is portrayed as one where the US holds global military and technological supremacy: “*The United States is the world's preeminent technological powerhouse*” (Horowitz, Kahn & Samotin 2022:2). While the setting emphasizes US hegemony, it also highlights AI's potential to upend global power dynamics: “*Every day more actors execute more sophisticated attacks against more civilian and military targets*” (Nakasone & Sulmeyer 2020:2). Taken together, the setting is characterized as being in flux, and while the US may hold a historic advantage due to their technological supremacy, its hegemonic status is in no way future-proof.

This dynamic is further emphasized through the discursive idea's narrative. Firstly, the narrative centers around AI/AITs in the context of “*war*”, “*defense*”, “*intelligence*”, and most importantly “*military*” (Table 2). Within this context, and as previously hinted at, the narrative is one of adversarial relations between the different characters. These adversarial relationships are described as occurring both in the material world and cyberspace, and are both, according to the setting, supercharged by emerging AI/AITs.

To illustrate AI proliferation and power in the material realm of warfare, the Russia-Ukraine war is invoked as an example: “*(...) the first conflict in which both sides are using artificial intelligence, particularly machine- and deep-learning algorithms*” (Kahn 2022:4). The idea posits that a key reason for the unexpected Ukrainian military success is their deployment of battlefield AIT, for example AI targeting algorithms and counterespionage software (Ibid.). This underscores the discursive idea's narrative around AI's potential to change power dynamics among adversarial characters.

A similar narrative is presented when it comes to cyberspace, where the US is being challenged by adversarial characters like China, Russia, Iran, and North Korea (Nakasone & Sulmeyer 2020:2). Furthermore, this challenge is enabled not only by large scale investments in AI R&D, but also its implementation into active operations: *“Over the last few years, China has invested roughly the same amount as the United States has in AI research and development, but it is more aggressively integrating the technology into its military strategy, planning and systems* (Horowitz, Kahn & Samotin 2022:3).

The discursive idea’s normative response to the narrative laid out above is broadly chaptered in the quote: *“If the United States fails to develop effective AI, it could find itself at the mercy of increasingly sophisticated adversarial actors”* (Horowitz, Kahn & Samotin 2022:6). Investing in AI, especially to develop militarily/defense-oriented AITs and implementing them at scale and speed, is widely presented as the normative imperative for the US.

Analyzed through the ES, the normative ideal moves beyond the ideal of order highlighted in the second discursive idea around great power competition, towards the ideal of an international system. The ideals of the international system can be seen in the discursive idea’s normative emphasis on the US maintaining its current advantages order and prioritizing sovereignty above all else. The emphasis on sovereignty is exemplified through the fear embedded in the narrative of possibly losing to an adversarial character like China. Also embedded in the discursive idea is the more realist assumption, that power function as the main currency in IR, and in a setting characterized by flux, due to rapid AI development and proliferation, power enabled by technology is more important than ever as it can: *“(…) expand the competitive space between the United States and its adversaries”* (Nakasone & Sulmeyer 2020:3).

Summary of the analytical findings

The following table summarizes the project’s analytical findings, highlighting the characteristics of each of the identified discursive ideas. This addresses the project’s first descriptive research objective, as it presents a complete overview of the different identified discursive ideas around AI in IR. Similarly to the analysis, the summary is based upon the underlying components to what constitutes a discursive idea. Additionally, the table serves as the basis for the project’s subsequent analytical discussion.

	<i>Characters</i>	<i>Setting</i>	<i>Narrative</i>	<i>Normativity</i>
1) Global governance and regulatory power	State governments, large tech-companies, and their platforms	Rising availability of generative AITs and 2024 as an important year for democracy	AITs as a threat to democracies, civil liberties, and human rights	International society: Regulation as an imperative but curtailing justice
2) Great power competition	The US and Chinese governments	Zero-sum competition	A self-perpetuating spiral of technological competition between the characters	International society: Separate Chinese and US ecosystems
3) Ensuring justice in a changing global economy	State governments and the global populous	Economic turmoil caused by low productivity and demographic challenges	Protectionism in the face of emerging technologies	World society: Deep economic integration, governance, and justice
4) Developing safe and ethical AI	Technological developers and AI as actor in its own right	A new industrial revolution	A negative technological determinism	World society: A global system of containment
5) Global instability and emerging threats	State governments	US technological hegemony upended by the emergence of AI/AITs	Adversarial relations between states supercharged by emerging AI	International system: US should maintain the current order through technological supremacy

Analytical discussion – The future of international society?

While Table 4 above addresses the first research objective by characterizing the discursive ideas' normative implications separately, the following section expands on this, by discussing their possible implications jointly, thereby, addressing the second research objective.

When comparing the discursive ideas, it is evident that the first, second and fifth ideas are broadly similar, as they all exist within the topic of national security. This is highlighted through the ideas' use of similar characters, emphasizing state governments, accompanied by narratives centered around increasing adversarial inter-state relations caused by emerging AI. Within the IR literature, national security concerns are often labeled as 'high politics', which is prioritized above all else by states when conducting foreign policy. Therefore, it could be argued that when the discursive ideas above present a narrative of national security threats, mainly from a US perspective, their normative suggestions could be more likely to dictate international society and state behavior, as they get perceived as 'high politics' (Youde 2016). Collectively, these ideas' normativity oppose the concept of a world society, which could potentially result in a fragmented international society, where competition and supremacy are the key logics defining emerging AI.

Following the argument above, if emerging AI becomes a policy area viewed by states as an area of security related ‘high politics’, where states, in line with the ES, prioritize order and sovereignty above global justice, it could hinder current efforts to address the negative implications of AI, illuminated through the critical IR literature. This trend could be especially impactful, as the current nascent efforts by the OECD and the UN to ensure sustainable AI are non-binding, meaning that there are no regulatory safeguards to prevent the discursive ideas around security and competition to define the future of state behavior. Ultimately, the combination of the discourse around AI as ‘high politics’ and the current shallow and non-binding global governance could result in a fragmented and nationally oriented approach to AI.

While the three previously mentioned discursive ideas emphasize the ‘high politics’ aspects of AI, idea number three, is more aligned with the ‘low politics’ of economic considerations due to its exclusive economic focus. Here, ‘low politics’ refers to IR policy disassociated from national security concerns. ‘Low politics’ issues are therefore often understood as being secondary to ‘high politics’ (Ibid.). Following this policy dichotomy between ‘low’ and ‘high’ politics, it is worth asking whether states can have their cake and eat it too, or do the discursive ideas around AI as a national security area overrule the normative calls for an economic world society? The narrative of the third discursive idea points to the latter being the case. As the analysis highlights, the third discursive idea’s narrative aligns with the normative position of the previously mentioned security-oriented ideas, emphasizing current rising protectionism and competition in the context of AI. Therefore, it could be argued that global security concerns are already overruling normative calls for a world society, and that this trend may continue if the normativity of the security-oriented ideas become even more dominant, due to their emphasis on ‘high politics’.

A similar thing could be expected to happen with the world society ideal of the fourth discursive idea. Unlike the previous economically oriented idea, the fourth discursive idea aligns more closely with the three security-oriented ideas, as it presents a narrative where poorly developed AI could be perceived as an existential threat to humanity. Thereby, being a highly securitized issue. However, unlike the previous ideas, the security threat doesn’t stem from other contemporary state actors, but from AI emerging as a threat through a longer time horizon. Some scholars have argued that this makes the threat less tangible and immediate, thereby making it harder for political actors to respond to it (Bostrom 2014). This could result in the discursive ideas around

the immediate security threats gaining preference, resulting in a short-term international society or even an international system, rather than the long-term world society being presented as the fourth discursive idea's ideal.

Conclusion

The project set out to examine what characterized the different discursive ideas around AI and IR and interpret what these could mean for states and the future of international society. Through its analysis, the project was able to identify five competing discursive ideas within the foreign policy epistemic community writing in *Foreign Affairs*, each with a distinct narrative and normative ideals around how to approach emerging AI. The discursive ideas range from viewing AI/AITs as enablers of state power within a military and security contexts, to advocating for increased global justice through the establishment of a world society, emphasizing global regulation and integration. However, as the analytical discussion showed, the discourses that highlight the immediate 'high politics' around national security and therefore advocate for an international system or society as their normative ideal, may dominate the ideas arguing for a world society. This could shatter the current nascent efforts of global AI governance, thereby furthering an international society where competition and relative gains around AI are the prevailing logics. As a final remark, it is important to note that the project's conclusions are limited in scope due to the predominantly Western context of the selected *Foreign Affairs* articles. Therefore, the identified discursive ideas and their possible implications are by no means a conclusive overview of all prevailing ideas around AI, as others may exist beyond *Foreign Affairs* and new ideas may emerge as AI technological development continues to accelerate.

Bibliography

- Bareis, J. & Katzenbach, C. (2022), Talking AI into Being: The Narratives and Imaginaries of National AI Strategies and Their Performative Politics, *Science, Technology, & Human Values*, 47(5):855-881.
- Benjamin, R. (2019), *Race After Technology: Abolitionist Tools for the New Jim Code*, John Wiley & Sons, New York.

- Benoit, K. (2020), Text as Data: An Overview, In Curini, L. and Franzese, R. (Eds.) *Handbook of Research Methods in Political Science and International Relations*, Sage Publications, Thousand Oaks, pp. 461–497.
- Bilgin, M. (2016), The state of future in international relations, *Futures: The journal of policy, planning and futures studies*, 82(2016):52–62.
- Blei, D. M. (2012). Probabilistic topic models, *Communications of the ACM*, 55(4):77–84.
- Bostrom, N. (2014), *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, Oxford.
- Burton, J. (2023), Algorithmic extremism? The securitization of artificial intelligence (AI) and its impact on radicalism, polarization and political violence, *Technology in Society*, 75(2023):1-10.
- Butcher, J. & Beridze, I. (2019), What is the State of Artificial Intelligence Governance Globally? *The RUSI journal*, 164(5–6):88–96.
- Buzan, B. (2004), *English school theory and the social structure of globalisation*, Cambridge University Press, Cambridge.
- Crawford, K. (2021), *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*, Yale University Press, Yale.
- Creswell, J. W. (2003), *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, Sage Publications, Thousand Oaks.
- Cummings, M. L., Roff, H. M., Cukier, K. Parakilas, J. & Bryce, H. (2018), *Artificial Intelligence and International Affairs: Disruption Anticipated*, Chatham House: The Royal Institute of International Affairs, London.
- D’Ignazio, C. & Klein, C. F. (2020), *Data feminism*, MIT Press, Cambridge.
- Ding, H. & Kong, Y. (2019), Constructing Artificial Intelligence in the U.S. and China: A Cross-Cultural, Corpus-Assisted Study, *China media research*, 15(1):93-106.
- Dunne, T. (2021), The English School, In Dunne, T., Kurki, M. & Smith, S. (Eds.) *International Relations Theories: Discipline and Diversity*, Oxford University Press, Oxford, pp. 108-128.
- Foreign Affairs (2024), *About Foreign Affairs*, <https://www.foreignaffairs.com/about-foreign-affairs> [04.04.2024].
- Fouse, S., Cross, S. & Lapin, Z. J. (2020), DARPA's Impact on Artificial Intelligence, *The AI Magazine*, 41(2):3-8.
- Grimmer, J. & Stewart, B. M. (2013), Text as data: The promise and pitfalls of automatic content analysis methods for political texts, *Political Analysis*, 21(3):267–297.
- Gu, F. (2023), Corpus-based Critical Discourse on AI Policy: A Comparison Between North America and Developing Countries in East Asia, *Asian Journal of Social Science Studies*, 8(3):14-25.

- Guenduez, A. A. & Mettler, T. (2022), Strategically constructed narratives on artificial intelligence: What stories are told in governmental artificial intelligence policies? *Government Information Quarterly*, 40(2023):1-13.
- Horowitz, M. C. (2019), When speed kills: Lethal autonomous weapon systems, deterrence and stability, *Journal of Strategic Studies*, 42(6):764-788.
- Ikenberry, J. G. (2020), *A World Safe for Democracy: Liberal Internationalism and the Crisis of Global Order*, Yale University Press, Yale.
- Isoaho, K., Gritsenko, D. & Makela, E. (2021), Topic Modeling and Text Analysis for Qualitative Policy Research, *Policy Studies Journal*, 49(1):300-324.
- Jacobs, T. & Tschötschel, R. (2019), Topic models meet discourse analysis: a quantitative tool for a qualitative approach, *International Journal of Social Research Methodology*, 22(5):469–485.
- Kissinger, H. (2018), How the Enlightenment Ends, *The Atlantic*, 16.05.2018.
- Kissinger, H., Schmidt, E. & Huttenlocher, D. (2022), *The Age of AI: And Our Human Future*, John Murray, London.
- Krafft, P. M., Huang, K., Young, M., Bugingo, G. & Katell, M. (2019), Defining AI in Policy versus Practice, *Ithaca: Cornell University Library*, 1-11.
- Kratochvíl P. & Tulmets, E. (2017), Constructivist approaches to the study of the European Neighbourhood Policy, In Shumacher, T., Marchetti, A. & Demmelhuber, T. (Eds.) *The Routledge Handbook on the European Neighbourhood Policy*, Routledge, London, pp. 70-80.
- Larsen, B. C. (2022), *The Geopolitics of AI and the Rise of Digital Sovereignty*, Brookings Institute, Washington D.C.
- Lazard Geopolitical Advisory (2023), *Research Brief: Geopolitics of Artificial Intelligence*, Lazard Geopolitical Advisory, New York.
- Mearsheimer, J. J. (2021), Structural Realism, In Dunne, T., Kurki, M. & Smith, S. (Eds.) *International Relations Theories: Discipline and Diversity*, Oxford University Press, Oxford, pp. 51-67.
- Mohamed, S., Png, M-T. & Isaac, W. (2020), Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence, *Philosophy & Technology*, 33(4):659-684.
- Ndzendze; B. & Marwala, T. (2023), *Artificial Intelligence and International Relations Theory*, Palgrave Macmillan, Singapore.
- Pavel, B., Ke, I., Spirtas, M., Ryseff, J., Sabbag, L., Smith, G., Scholl, K. & Lumpkin, D. (2023), *AI and Geopolitics: How might AI affect the rise and fall of nations?* RAND Corporation, Santa Monica.
- Philbeck, T. & Davis, N. (2018), The Fourth Industrial Revolution, *Journal of International Affairs*, 72(1):17-22.
- Radu, R. (2021), Steering the governance of artificial intelligence: national strategies in perspective, *Policy & society*, 40(2):178–193.

- Regan, M. & Davidovic, J. (2023), Just preparation for war and AI-enabled weapons, *Frontiers in big data*, 6(1020107):1-6.
- Reus-Smit, C. (2002), Imagining society: Constructivism and the English School, *British Journal of Politics and International Relations*, 4(3):487-509.
- Roberts, H., Cowles, J., Morley, J., Taddeo, M., Wang, V. & Floridi, L. (2021), The Chinese Approach to Artificial Intelligence: An Analysis of Policy, Ethics, and Regulation, *AI & Society* 36(1): 59–77.
- Roberts, H., Cowls, J. Hine, E., Morley, J., Wang, V., Taddeo, M. & Floridi, L. (2023), Governing artificial intelligence in China and the European Union: Comparing aims and promoting ethical outcomes, *The Information Society*, 39(2):79-97.
- Robinson, S. C. (2020), Trust, transparency, and openness: How inclusion of cultural values shapes Nordic national public policy strategies for artificial intelligence (AI), *Technology in Society*, 63(2020):1-15.
- Samoili, S., López Cobo, M., Gómez, E., De Prato, G., Martínez-Plumed, F., & Delipetrev, B. (2020), *AI Watch: Defining Artificial Intelligence: Towards an operational definition and taxonomy of artificial intelligence*, Publications Office of the European Union, Luxembourg.
- Schmidt, V. A. (2008), Discursive institutionalism: The explanatory power of ideas and discourse, *Annual review of political science*, 11(1):303–326.
- Schmitt, L. (2022), Mapping global AI governance: a nascent regime in a fragmented landscape, *AI and ethics*, 2(2):303–314.
- Schopmans, H. R. (2022), From Coded Bias to Existential Threat: Expert Frames and the Epistemic Politics of AI Governance, *AAAI/ACM Conference on AI, Ethics, and Society*, pp. 627-640.
- Schwab, K. (2016), *The Fourth Industrial Revolution*, Penguin, New York.
- Suleyman, M. & Bhaskar, M. (2023), *The Coming Wave: AI, Power and the 21st Century's Greatest Dilemma*, The Bodley Head, London.
- Tallberg, J., Erman, E., Furendal, M., Geith, J. Klamberg, M. & Lundgren, M. (2023), The Global Governance of Artificial Intelligence: Next Steps for Empirical and Normative Research, *International Studies Review*, 25(3):1-23.
- Wang, P. (2019), On Defining Artificial Intelligence, *Journal of Artificial General Intelligence*, 10(2):1-37.
- Wendt, A. (1992), Anarchy is what States Make of it: The Social Construction of Power Politics, *International Organization*, 26(2):391-425.
- Youde, J. (2016), High Politics, Low Politics, and Global Health, *Journal of global security studies*, 1(2):157–170.

List of appendices

- Appendix 1: Code repository.
- Appendix 2: Resources for the literature review.