Received: 2 August 2023 | Revised: 25 September 2023 | Accepted: 12 October 2023 | Published online: 20 October 2023

RESEARCH ARTICLE

Archives of Advanced Engineering Science
2024, Vol. 2(2) 79–92
DOI: 10.47852/bonviewAAES32021471

BON VIEW PUBLISHING

Machine Learning Insights into Hypersonics Research Evolution: A 21st Century Perspective

Ashkan Ebadi[1,*], Alain Auger[2] and Yvan Gauthier[1]

[1]*Digital Technologies, National Research Council Canada, Canada*

[2]*Science and Technology Foresight and Risk Assessment Unit, Defence Research and Development Canada, Canada*

Abstract: In recent years, the field of hypersonics has witnessed substantial growth in research and development activities, driven by its diverse range of applications spanning both military and commercial sectors. Governments and private companies in several countries have made substantial investments in hypersonic technologies to gain a competitive edge, secure or enhance strategic capabilities, and bolster deterrence measures. In this rapidly evolving landscape, the ability to swiftly and accurately identify emerging technologies becomes paramount. Leveraging the advancements in information technology and computer science, which enable the analysis of vast datasets and the extraction of concealed trends and patterns, this study aims to provide valuable insights to decision-makers in the hypersonics domain. Our focus is on scientific publications related to hypersonics, encompassing the years 2000–2020. We employ state-of-the-art natural language processing and machine learning techniques to comprehensively characterize the research landscape. The urgency of this endeavor lies in the necessity for organizations to remain at the forefront of hypersonic research. By algorithmically identifying and tracking 12 key latent research themes and examining their temporal evolution, we offer a structured and objective analysis of the field. Our methodology eliminates subjectivity from the assessment, facilitating consistent comparisons both across topics and across different time intervals. In addition, through our extensive publication similarity analysis, we uncover nuanced patterns that shed light on the cyclical nature of research trends over the two decades under investigation. This comprehensive examination of the hypersonics research landscape not only underscores its critical significance but also provides a robust foundation for informed decision-making. As such, our study serves as a valuable resource for stakeholders seeking to navigate the dynamics of the rapidly advancing field of hypersonics effectively.

Keywords: hypersonics, research evolution, temporal change, natural language processing, machine learning, structural topic modeling

1. Introduction

The term “hypersonic” is used to characterize objects that move faster than 5 times the speed of sound (Van Wie, 2021). Hypersonic technology is not a new research topic and has been investigated for more than six decades (Blankson & Pyle, 1993; Czysz & Vandenkerckhove, 2001), although international interest varied over the years (Van Wie, 2021). With applications in several domains, including military (Malinowski, 2020) and commercial (Ingenito et al., 2011), hypersonic technology is argued to be crucial for both national defence and space exploration purposes (Gu & Olivier, 2020). Given the competition between states, hypersonics is now of increased importance (Van Wie, 2021), and it has become a hot research topic in the scientific community as well as in international relations.

Over 60 years of hypersonics technology, investigation has resulted in both experimental and operational systems (Van Wie, 2021). High-speed vehicles have attracted a growing interest in the aviation field, supporting manned and unmanned operations and explorations in low earth orbit, calling for high flexibility, affordability, and high safety standards (Viviani & Pezzella, 2019). In addition to space exploration missions, interest has been recently growing in hypersonics travels for commercial and civilian applications, with many start-ups involved in the development of hypersonic aircrafts (Viviani & Pezzella, 2019). Therefore, many researchers in basic and applied technologies have focused on hypersonic research and development (R&D) due to the potential advantages of an operational hypersonic vehicle (HV) (Ding et al., 2022; Viviani & Pezzella, 2019).

There are several examples of activities/initiatives that have contributed to engineering advancement, for example: (1) the X-15 research aircraft (1959–1968), (2) the Apollo reentry capsule (1966–1975), (3) the Space Shuttle program (1981–2011), and (4) the SpaceX Falcon 9 launch system (2010–present) (Van Wie, 2021). From a military perspective, aviation and space have been always critical for nations from a national defence perspective (Richman et al., 2019). As a result of this global attention, new categories of hypersonic capabilities have emerged or are being explored. This includes, but is not limited to, hypersonic boost-glide systems, interceptor missiles, reusable aircrafts, hypersonic cruise missiles, and gun-launched projectiles (Van Wie, 2021). Technology emergence is particularly noticeable in the area of

***Corresponding author:** Ashkan Ebadi, Digital Technologies, National Research Council Canada, Canada. Email: ashkan.ebadi@nrc-cnrc.gc.ca

offensive hypersonic strike systems, followed by the emergence of hypersonic defence capabilities to counter these offensive systems (Van Wie, 2021). To date, the United States, China, and Russia have the most advanced hypersonic military capabilities (McFarland, 2023; Tiron, 2021).

With efforts in hypersonics research accelerating, applications of the technology in different markets are becoming operational (Deloitte, 2020). Beyond military applications, hypersonic technology has potential for commercial applications, including reusable hypersonic systems and aircrafts and fully reusable space-access vehicles (Van Wie, 2021). For instance, reusable hypersonic aircrafts are expected to be developed for routine commercial flights (Carioscia et al., 2019; Viola et al., 2023). Such vehicles will require additional technological advancement to address issues inherent to interacting components such as propulsion, structures, and guidance and control systems (Van Wie, 2021).

Governments are investing in hypersonics technology as part of their national defence strategies. In the United States only, more than \$2.6 billion is annually invested, with a 26% compound annual growth rate (CAGR) since 2014 (Deloitte, 2020). The international hypersonic market is predicted to develop at a CAGR of 9.5% during the period of 2022–2028, with the global market value exceeding \$12.5 billion by 2028 (Facts & Factors, 2022). In addition, venture capital investment in hypersonic technologies is expected to grow in the coming years in response to market demand (Deloitte, 2020). The research activities have also been fueled by significant investments. For instance, the United States has made a more than 7-fold funding increase in hypersonic research between 2015 and 2020 (McDonald, 2021). The reusability and speed of HVs have been key driving factors of these investments (Piplica, 2021), especially on the commercial side, and hypersonic air travel is now expected to become operational within the decade (McDonald, 2021).

For the scientific community, peer-reviewed scientific publications represent the main channel to communicate results and achievements. Due to the availability of large-scale publication data, quantitative analysis of such data has been widely carried out to extract insights from the corpus of publications and the relationships between authors and articles (Ebadi et al., 2020). A few studies employed traditional methods to examine the research landscape in the fields of hypersonics and its related technologies. For example, Kostoff et al. (1999) applied bibliometrics and database tomography to extract technical insights from a database of papers on hypersonic/supersonic flow. In another study, Anderson (1984) conducted a survey study examining contemporary developments in hypersonics aerodynamics. While the paper offers insights into potential future advancements within the field, it is characterized by subjectivity and relies on the author's personal experience and viewpoint. Bakos (2008) reported on the hypersonic research state in the United States, focusing on engine designs. Previous studies, either outdated or employing conventional approaches, have left a gap in providing a current and comprehensive understanding of the subject.

Recent advancements in the field of computer science provide a unique opportunity to go beyond traditional methods, explore massive high-dimensional datasets, and obtain new insights. For example, topic modeling (TM) is an unsupervised machine learning technique that can automatically organize a collection of documents, in our case scientific publications, into a set of high-level themes, so-called topics (Papadimitriou et al., 2000), leveraging context clues to infer hidden themes in the unstructured text data. In contrast to conventional approaches that demand substantial manual labor (Gatti et al., 2015), TM can automatically distill and reveal hidden semantic themes from massive text datasets (Blei et al., 2003). For example, content analysis (Krippendorff, 2018), as an instance of conventional approaches, involves manually coding and categorizing text based on predefined criteria. While it can provide rich insights, it is labor-intensive, time-consuming, and subject to human bias and error. It may also struggle to handle large volumes of text data efficiently. However, TM is automated and can handle large text datasets without the need for extensive manual coding. It is less prone to bias and can uncover latent themes that may not be apparent through manual coding alone. TM has been widely used for extracting latent topics from the collection of documents in various domains, e.g., psychology (Bittermann & Fischer, 2018), education and information technologies (Ozyurt & Ayaz, 2022), big data (Mohammadi & Karami, 2022), transportation research (Das et al., 2016), and the recent COVID-19 pandemic (Ebadi et al., 2021), to name a few.

Considering the economic and national defence implications of hypersonic technologies as well as their fast-evolving nature, it is crucial for policy analysts and decision-makers to have a comprehensive understanding of the related R&D landscape. However, this often translates into a manual investigation of a huge set of datasets/sources, which, apart from being time-consuming and not scalable, may also suffer from a lack of comprehensiveness, biases, and subjectivity. The high availability of digitalized research data, on the one hand, and recent advances in analytics and machine learning research, on the other hand, present new opportunities for large-scale advanced data analytics. These advanced approaches and tools can assist human analysts by revealing hidden patterns and new insights from massive datasets (Ebadi et al., 2022; Keller & von der Gracht, 2014).

The motivation for this proposed work stems from the burgeoning significance of hypersonic technology in both military and commercial domains. With increasing investments worldwide, there is a pressing need for a systematic and data-driven approach to track and understand the evolving landscape of hypersonics research. The main objective of this study is to obtain a better understanding of the hypersonics research landscape and its evolution over time. To that end, we leverage natural language processing (NLP) and machine learning to analyze quantitatively the scientific publications about hypersonics within the period from 2000 to 2020. The contribution of this work lies in its pioneering effort to address a critical knowledge gap in the domain of hypersonics research. To the best of our knowledge, no prior study has undertaken such a comprehensive and data-driven analysis of hypersonics publications spanning two decades. Our work offers several substantial contributions: (1) Characterization of the research landscape: This study provides an in-depth understanding of the hypersonics research landscape, offering policy-makers and strategic planners a holistic view of the field's evolution. By identifying and categorizing research themes, we create a high-level representation that serves as a foundational resource for decision-makers. (2) Objective insights: By automating the extraction of key research themes and their temporal evolution, our approach eliminates subjectivity from the analysis, ensuring that the insights are based on data-driven observations resulting from a repeatable process. This objectivity enhances the reliability and relevance of the findings. (3) Identification of research trends: The study uncovers temporal patterns and trends within the hypersonics research community, shedding light on the emergence, consolidation, and potential future directions of key technologies. This information is invaluable for those tasked with making informed decisions about research priorities and investments. And, (4) strategic decision support: The insights generated by our methodology empower policy-makers and strategic planners to make well-informed decisions regarding resource allocation, technology development, and international collaborations. It enables them to stay ahead of the curve in the rapidly evolving field of hypersonics. By harnessing the power of NLP and machine learning, this study aims to provide decision-makers with timely and objective insights into emerging technologies, enabling them to stay competitive and strategically relevant in this rapidly advancing field.

The remainder of this paper is as follows. Section 2 describes data and techniques in detail. Section 3 presents the findings of the research. Findings are then discussed and the conclusions are presented in Section 4. Finally, some limitations of the research and future directions are presented in Section 5.

2. Data and Methodology

2.1. Data

The scope of this research covers all research publications about hypersonics that are accessible through Elsevier's Scopus, i.e., a comprehensive abstract and citation database of peer-reviewed scientific journals, books, and conference proceedings. Initially, the bibliographic data of hypersonics-related publications, published within the period of 2000–2020, were retrieved from Elsevier's Scopus by running "hypersonic*" as the search query (data were collected in August 2021). The collected bibliographic data included meta-data about each of the extracted papers, such as the title, abstract, date of publication, list of authors, and their affiliations. We filtered out publications for which neither title nor abstract was available. This initially resulted in a total of 17,075 publications on which we applied several preprocessing steps that will be explained in detail in Section 2.2.1.

2.2. Methodology

The methodology has five main components: (1) text preprocessing, (2) correspondence analysis (CA), (3) lexical complexity analysis, (4) publications similarity analysis, and (5) structural topic modeling (STM). These components are introduced and discussed in detail below.

2.2.1. Text preprocessing

The abstract of a publication provides a condensed representation of its content. Titles may also provide some complementary informative keywords and/or keyphrases that may not be necessarily present in the abstract section. Since we did not have access to the entire texts of the publications, we decided to integrate for each paper both the title and abstract (and not just use one of them) to obtain a better representation of the publications' content. As the first preprocessing step, we created a new column in the dataset by combining the title and abstract of each publication and removing duplicated records. Next, we applied several preprocessing steps to prepare the data for analysis. These steps involved converting text to lowercase, removing stop words based on a customized English stop words list, correcting special characters, and removing punctuations. We tokenized the processed data and created a document term-frequency matrix in which each row represents one publication, columns represent the tokenized terms, and each cell value is the number of appearances of a given term in a given publication.

2.2.2. Correspondence analysis

Correspondence factor analysis (Greenacre & Blasius, 2006) was first applied to the preprocessed data (the document term-frequency matrix) to verify the existence of a temporal trend in publications. CA provides a graphical representation of cross-tabulations by mapping data on visually understandable dimensions in which noises are filtered out, making it easier to visualize general patterns in the data (Blasius & Greenacre, 1998; Doré & Ojasoo, 2001). Using the generated term-frequency matrix (as explained in Section 2.2.1), we created a term-frequency matrix for each year of the examined period by filtering on the year of publication. Next, terms with a frequency of less than 50 in each year were filtered out in order to reduce/remove noise. Highly frequent terms with the appearance in more than 60% of the publications in each year were also removed as they were common terms and not specific enough to provide insight into the general temporal pattern. The filtered term-frequency matrix was then mapped to a 2-dimension space by extracting the first two principal components. Results are discussed in Section 3.2.

2.2.3. Lexical complexity analysis

As a preliminary step in understanding the evolution of the scientific terminology in the examined case technology, we investigated the linguistic complexity of the publications. Specifically, using abstracts of publications in the dataset, we analyzed their textual complexity from two aspects, i.e., textual readability and richness.

To assess textual readability, we calculated the following five measures per paper and averaged over each year of the examined time interval: (1) number of words per paper, (2) the number of sentences, (3) sentence length, defined as the number of words divided by the number of sentences, (4) ratio of difficult words, defined as words that are not common and have at least 2 syllables, and (5) the Flesch–Kincaid score (Kincaid et al., 1975) that indicates how difficult a text is to understand using sentences, words, and syllables as the core components. Lower Flesch–Kincaid scores demonstrate texts that are easier to read while higher numbers indicate more difficult texts.

The following three measures were calculated to assess lexical richness: (1) type-token ratio (TTR), (2) measure of textual lexical diversity (MTLD), and (3) hypergeometric distribution diversity (HD-D). As the most well-known measure of lexical diversity (Lissón & Ballier, 2018) and as a measure of vocabulary variation, TTR is defined as the ratio of the total number of unique words, i.e., types, divided by the total number of words, i.e., tokens, in a given text. The MTLD measure, developed by McCarthy (2005), divides the text into segments that are of variable lengths since fragmentation is done based on the TTR values of the segments, until it reaches the default TTR size value (0.72). The HD-D measure (McCarthy & Jarvis, 2007) is based on the hypergeometric distribution representing the probability of finding a certain number of words from a random sample of a certain size. Results are discussed in Section 3.3.

2.2.4. Publications similarity over time

After verifying the existence of a temporal trend in the corpus (cf. Section 2.2.2), we investigated research similarities. This preliminary analysis would shed light on the evolution rate of the research terminology in the examined case technology, i.e., hypersonics. Assessing document similarity is a challenging NLP task. When two documents share similar semantic contexts, they can be considered similar, and manually discerning this similarity among a vast number of documents can be quite very difficult, if not impossible. To simplify this process, we employ a Doc2Vec model (Le & Mikolov, 2014), which is an unsupervised machine learning technique used to transform a document into a vector representation, trained on the corpus to learn publication-level embeddings. We set the parameters of the Doc2Vec model as follows: vector_size = 30, max_epochs = 200, alpha = 0.025, and min_alpha = 0.00025. The vector_size refers to the dimensionality of the embedding vectors, alpha is the initial learning rate, and the learning gradually decreases to reach the minimum value defined by min_alpha as the training process advances. Finally, cosine similarity was applied to the publications' embedding vectors, and results were aggregated for each year of the examined period to evaluate publications' similarity over time. Results are discussed in Section 3.4.

2.2.5. Structural topic modeling

To complement the previous steps, TM was applied to extract latent research themes from the corpus. Unlike conventional methods (e.g., content analysis (Krippendorff, 2018)) that require considerable manual effort (Gatti et al., 2015), as an unsupervised machine learning technique, TM can summarize huge text data collections and extract latent semantic themes automatically (Blei et al., 2003). In other words, TM is an approach to clustering few words across the corpus into topics (Bhat et al., 2020). TM has been widely used in the literature to analyze the research landscape of various domains, including but not limited to transportation (Sun & Yin, 2017), cancer (Mosallaie et al., 2021; Stout et al., 2018), manufacturing (Yoon et al., 2019), and even the recent COVID-19 pandemic (Ebadi et al., 2021).

In this work, we used STM (Roberts et al., 2019) since it has two main properties that were critical for our research objectives: (1) it allows incorporating publication-level covariates of interest for which we considered publication date as a covariate to analyze the temporal evolution of the domain and (2) it allows topics to be correlated, which helped us to better understand the dynamics of the hypersonics research as well as the structure of the topics at the corpus level. The STM model was built on the entire dataset with an annual granularity to capture hidden temporal patterns necessary to draw the landscape of hypersonics research and analyze its evolution. Similar to other TM approaches, STM requires no data labeling and the topics emerge automatically in an unsupervised setting (Ebadi et al., 2020).

STM requires the number of topics to be set as a fixed parameter. There is no universal quantitative approach for finding the optimal number of topics in topic models (Lucas et al., 2015). Since a fully automated approach to finding the optimal number of topics may result in inaccurate findings (Maskeri et al., 2008), we used a multi-layer semi-automated approach to determine the number of topics. For this purpose, several baseline latent Dirichlet allocation models (Blei et al., 2003) were first built by varying the number of topics in the [3, 20] range, and the coherence of the generated topics was calculated. We used the C_v score (Röder et al., 2015), which is an intrinsic evaluation metric, for topic coherence calculation. The C_v scores helped us to narrow down the optimal range for the number of topics further. It was found that the optimal number of topics should be in the range of [10, 15]. Next, we manually checked the top keywords and keyphrases assigned to each topic in the generated models and analyzed topic-word distributions. Based on this manual topic verification step, it was concluded that the optimal number of topics for the examined corpus is 12, and hence, the STM model was built with 12 topics. Since STM does not assign a representative label to the extracted topics automatically and in order to ensure the quality of the model, the extracted topics were finally reviewed by senior scientists with expertise in hypersonics research from Defence Research and Development Canada. They assigned a meaningful and concrete label to the topics. Figure 1 shows the conceptual flow of the study.

Figure 1
The conceptual flow of the study

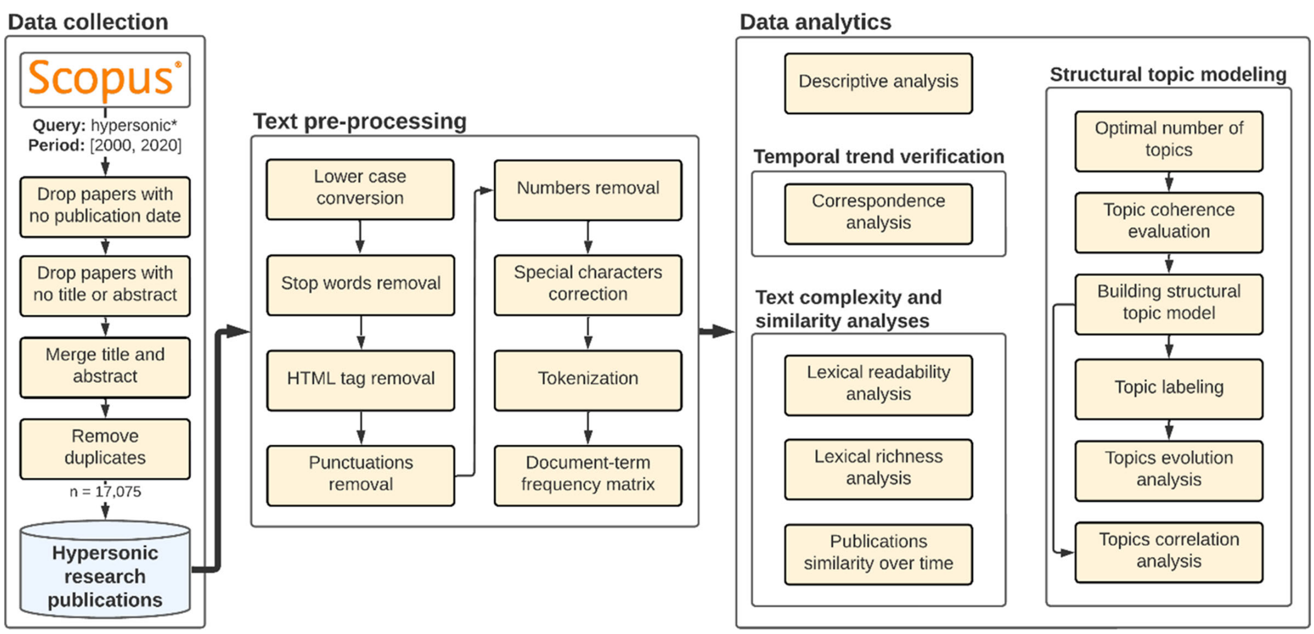
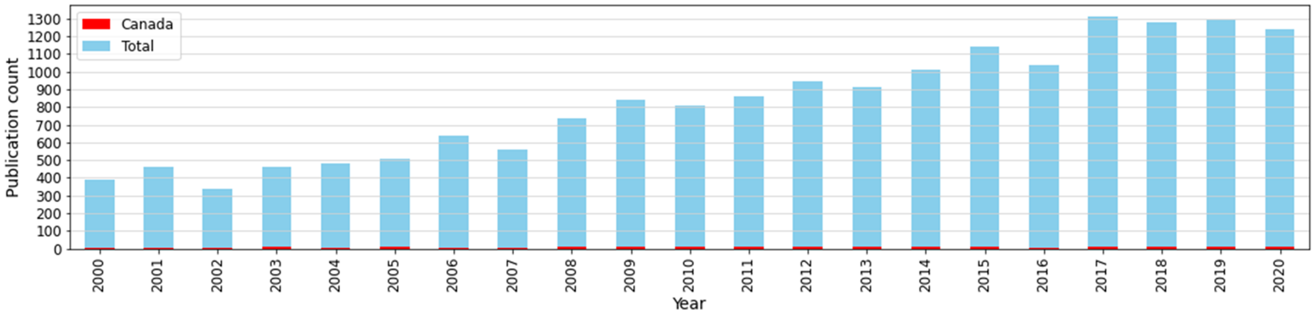


Figure 2
The trend of hypersonics publications between 2000 and 2020



3. Results

3.1. Number of publications trend

Figure 2 shows the hypersonics publications distribution over the examined period. We marked an article as a Canadian publication if at least one of the authors had a Canadian affiliation. Canadian publications are color-coded in red in Figure 2 and blue bars represent the total number of publications. The number of publications follows an increasing trend, and almost tripled between the beginning and the end of the period. Canadian publications maintained a constant output since 2008, except for a drop in 2016. The overall publication rate plateaued between 2017 and 2020. The observed growth in the number of publications is in line with several studies (Bornmann & Mutz, 2015; Ebadi et al., 2020; Zeng et al., 2019) that reported such an increasing trend in multiple domains. One may note that the growth could be also partially due to higher coverage of publications in Scopus during the final years of the period under review.

3.2. Temporal trend verification

Before analyzing the temporal evolution of the hypersonics research themes, the existence of the temporal trend was verified by performing a CA on the frequent terms extracted from publications in each year of the examined time interval (as explained in Section 2.2.2). In Figure 3, years of publications and their representative frequent terms are depicted with orange and blue points, respectively. As seen in Figure 3, the research terminology has evolved over time, forming a relatively U-shaped curve. Based on the distance of the publication date variables, i.e., the orange points in the figure, from the axes and the origin, it is clear that the frequent terms have evolved over time, which confirms the existence of a temporal trend in the examined corpus.

3.3. Lexical complexity analysis

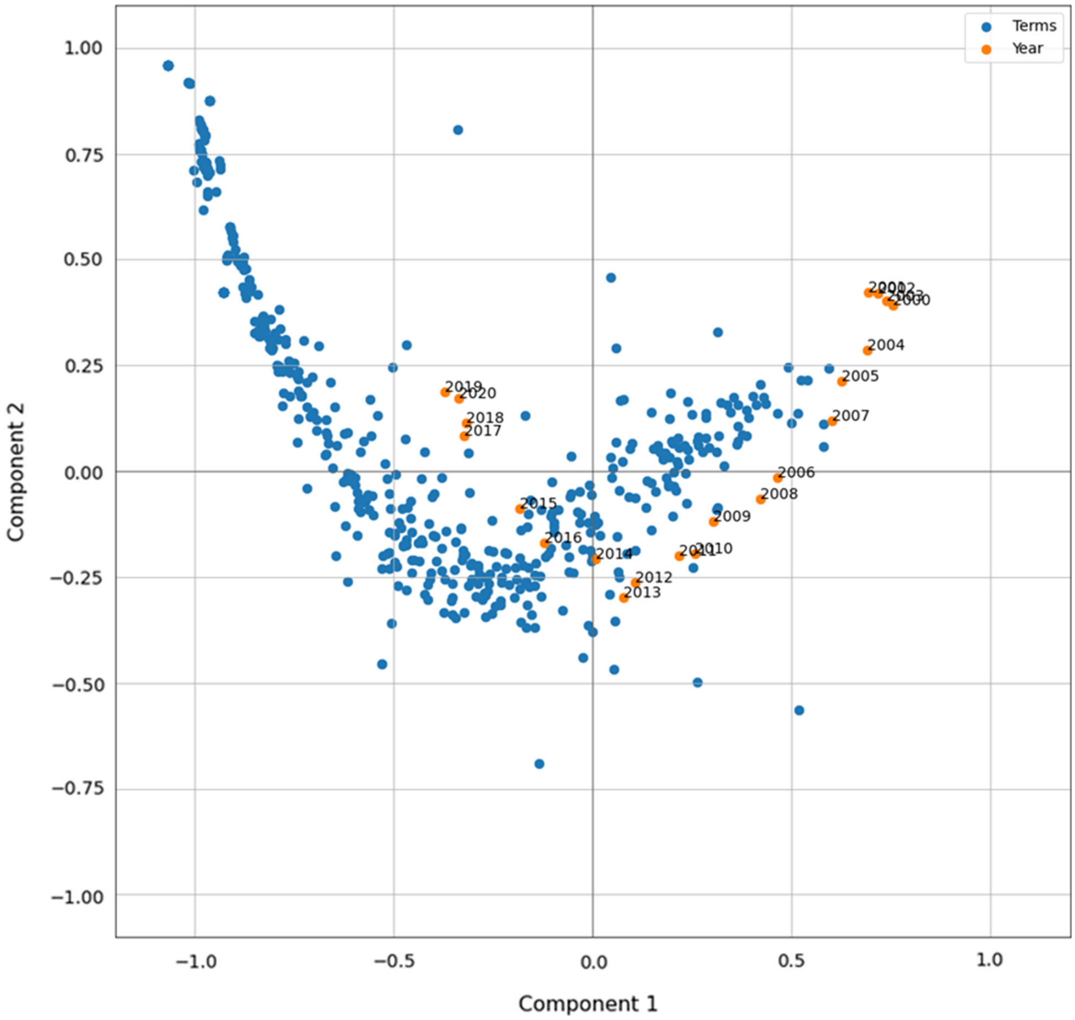
Figure 4 shows the results for the readability measures. From Figure 4(a), it is observed that the average number of words in the abstract section of the publication has followed an increasing trend after 2010, after some fluctuations at the beginning of the examined time interval, and reaching its peak in 2020. The number of sentences (Figure 4(b)) has increased sharply after 2013, while experiencing a smoother increasing trend beforehand. However, as observed in Figure 4(c), the average length of the sentences has decreased. From Figure 4(a), (b), and (c), it can be said that researchers tend to use more sentences of shorter lengths recently that contain relatively more common (easy) words. Figure 4(d) shows the trend of the ratio of the difficult words used in the publications' abstracts. The difficult word ratio has also a decreasing trend with a steeper slope after 2013. The last sub-figure, i.e., Figure 4(e), depicts the Flesch–Kincaid score. Intuitively, this score reflects the number of years of education that is required to understand the text. Figure 4(d) and (e) shows an increase in lexical readability over time, indicating that researchers have gradually used fewer difficult words in the abstract section to (maybe) engage a broader audience.

Figure 5 shows measures of lexical richness. As seen in Figure 5(a), after a relatively constant trend in the beginning years, the average TTR has followed a decreasing trend, indicating that the vocabulary in hypersonics research publications has become less varied, and more uniform. One problem reported for TTR is that it does not generalize well for long texts (Johansson, 2008); however, this is not the case here as we performed the analysis only on the abstracts of publications. Figure 5(b) shows the average HD-D score. The intuition behind the HD-D score is that if a sample contains many forms of a specific word, drawing a sample containing at least one form of that word would be very likely (Fergadiotis et al., 2013). The trend of the MTLD is depicted in Figure 5(c). HD-D and MTLD results also confirm that the diversity of the research vocabulary has decreased over time. This could be an indicator that the field and its terminology are stabilizing, however; further investigation is required to confirm this observation. HD-D value increases with text length while MTLD value decreases with text length (Treffers-Daller et al., 2018). The fact that both mentioned measures indicate a decreasing trend for lexical richness also verifies that text length is not an influencing factor in our performed analysis.

3.4. Publications similarity over time

Figure 6(a) shows the between-years similarity of hypersonics publications. Papers published in 2020 have on average the highest similarity with the publications in the previous years. For publications in 2020, the highest similarity is observed with publications in 2000, 2002, and 2004. The observation that papers published in 2020 exhibit the highest similarity with publications from the early 2000s may be attributed to several factors. Firstly, it could indicate that the foundational concepts and knowledge developed in the early 2000s have remained relevant and formed the basis for subsequent research in hypersonics. This suggests a

Figure 3
Correspondence analysis on the frequent terms that appeared in the hypersonics publications from 2000 to 2020



strong continuity in the field's core ideas over two decades. In addition, it may partially indicate modification of earlier results and/or reinterpretation of theories and conceptual frameworks. Conversely, the notably low similarity of publications from 2001 to other years may suggest a distinct research direction or unique scientific ideas pursued during that year. This could be due to groundbreaking discoveries or novel methodologies that set 2001 apart from the surrounding years. It may be worthwhile to delve deeper into the specific topics and research areas covered in 2001 to understand why they diverged from the broader trends in hypersonics research during that period. Figure 6(b) depicts research similarities in consecutive years. As observed, the similarity score is relatively high, fluctuating in the range of 0.60 and 0.65. In addition, the similarity between hypersonics publications has followed an increasing trend. These observations may partially shed light on the evolution of hypersonics research and its cycles. Also, the increasing trend in similarity between hypersonics publications across the examined years could reflect a growing consensus and convergence of research themes in the field. This might indicate that over time, researchers have collectively focused on key topics, leading to greater alignment in the content and approaches used in their publications. While the findings may initially seem counter-intuitive, they offer valuable insights into the evolution and dynamics of hypersonics research. The high similarity with early 2000s publications underscores the field's enduring foundations, while the distinctiveness of 2001 suggests a year of unique contributions. The increasing similarity trend highlights the evolving landscape of hypersonics research over the analyzed two-decade period. We further investigate the evolution of the examined field in the next section.

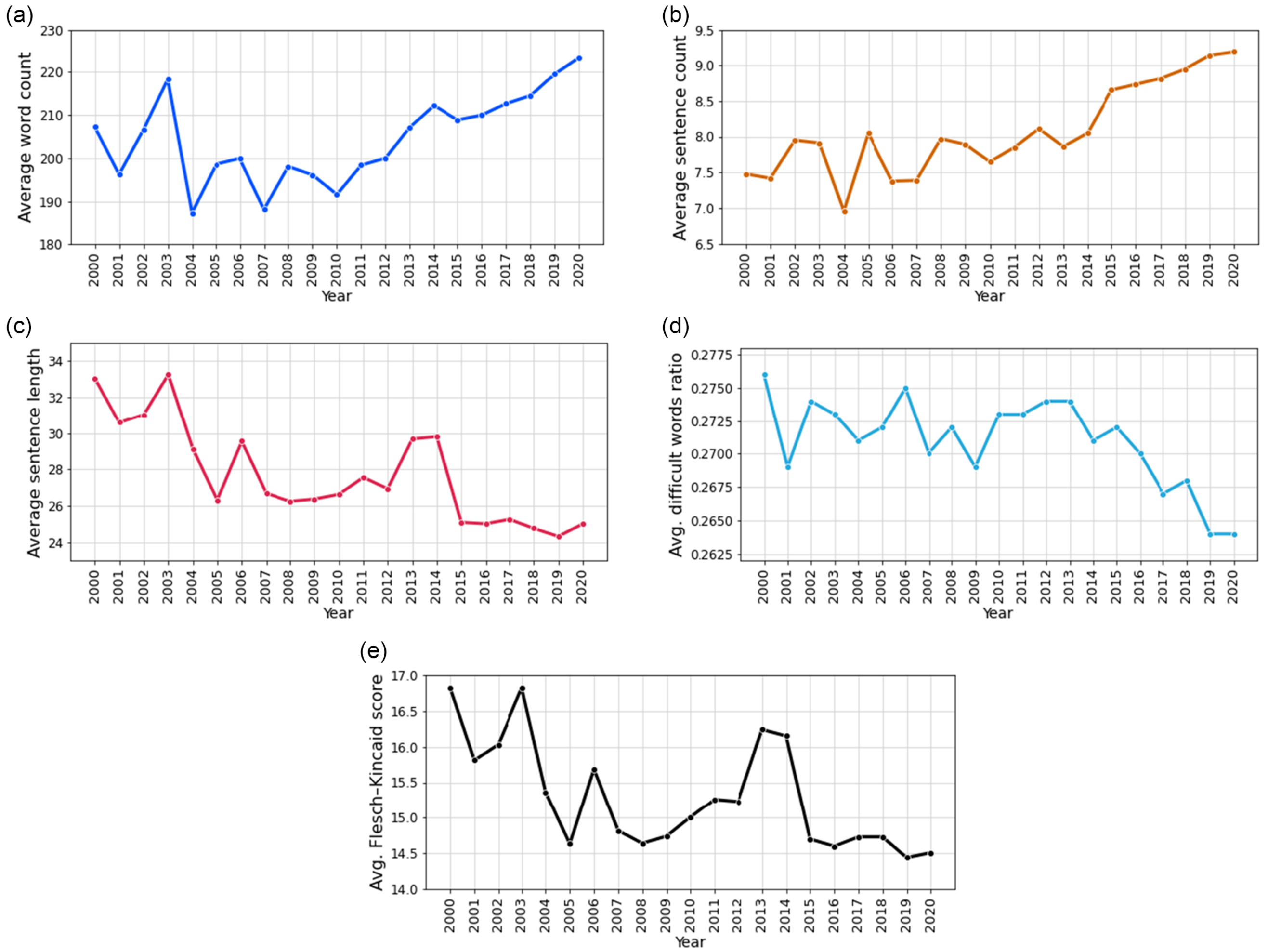
3.5. Hypersonics research themes and their temporal evolution

In this section, we first extract the main research themes in the hypersonics field and analyze their correlation and quality. Then, we analyze the temporal evolution of the research themes in more detail.

3.5.1. Hypersonics research themes

As explained in the Section 2.2.5, the STM technique (Roberts et al., 2019) was used to extract 12 main research themes in the field of hypersonics, using the year of publication as the covariate. The respective keyword sets of the extracted research themes were carefully reviewed and verified by senior scientists from Defence Research and Development Canada with domain expertise in hypersonics, and a representative label was assigned to each of

Figure 4
Lexical readability (a) average word count, (b) average sentence count, (c) average sentence length, (d) average ratio of difficult words, and (e) average Flesch–Kincaid score



the extracted themes as follows, terms in parentheses represent the respective high-level category:

- **Topic 1**: Cooling materials (*aerothermodynamics*)
- **Topic 2**: Fuel flow/propulsion systems (*aerothermodynamics*)
- **Topic 3**: Flight systems design (*aerothermodynamics*)
- **Topic 4**: Physico-chemical modeling (*aerothermodynamics*)
- **Topic 5**: Hypersonic velocity (*materials and structures*)
- **Topic 6**: Shock wave/boundary layer interaction-SWBLI (*materials and structures*)
- **Topic 7**: Flight vehicles tracking control (*guidance, navigation and control*)
- **Topic 8**: Wing panel modeling (*modeling, simulation and analysis*)
- **Topic 9**: Wind tunnel testing (*modeling, simulation and analysis*)
- **Topic 10**: Hypersonic materials (*materials and structures*)
- **Topic 11**: Hypersonic inlets (*vehicles, propulsion and fuels*)
- **Topic 12**: Guidance against hypersonic targets (*guidance, navigation and control*)

As seen in the above list, 4 research themes belong to aerothermodynamics, 3 belong to materials and structures, 2 to guidance, navigation and control, and modeling, simulation, and analysis, and 1 to the vehicles, propulsion, and fuels categories, respectively. In extracting topics, we followed the approach proposed by Bischof and Airoldi (2012) to improve the quality of the keywords, by extracting keywords that were both frequent and exclusive. Figure 7 shows the extracted topics' quality assessed through a combination of semantic coherence of the keywords of a given topic and exclusivity of the keywords to the topic. The size of the markers in the figure represents the proportion of the publications covered by the topic. Semantic coherence is calculated based on the number of times the keywords of a topic co-occur in publications. By this definition, a topic with a higher semantic coherence score could be considered as being more interpretable to humans. On the other hand, a topic is exclusive if keywords with a high co-occurrence likelihood conditioned on the topic have a low likelihood conditioned on other topics (Kuhn, 2018). As seen in the figure, the extracted topics have high semantic coherence and exclusivity, and most of them are located in the upper right region of the plot. Among the extracted topics, Topic 1, i.e., cooling materials, and Topic 4, i.e., physico-chemical modeling, have the highest exclusivity and coherence, respectively.

It should be noted that these extracted research themes only represent the main research areas of interest in the field of

Figure 5
Lexical richness (a) average type-token ratio (TTR), (b) average hypergeometric distribution diversity (HD-D) score, and (c) average measure of textual lexical diversity (MTLD)

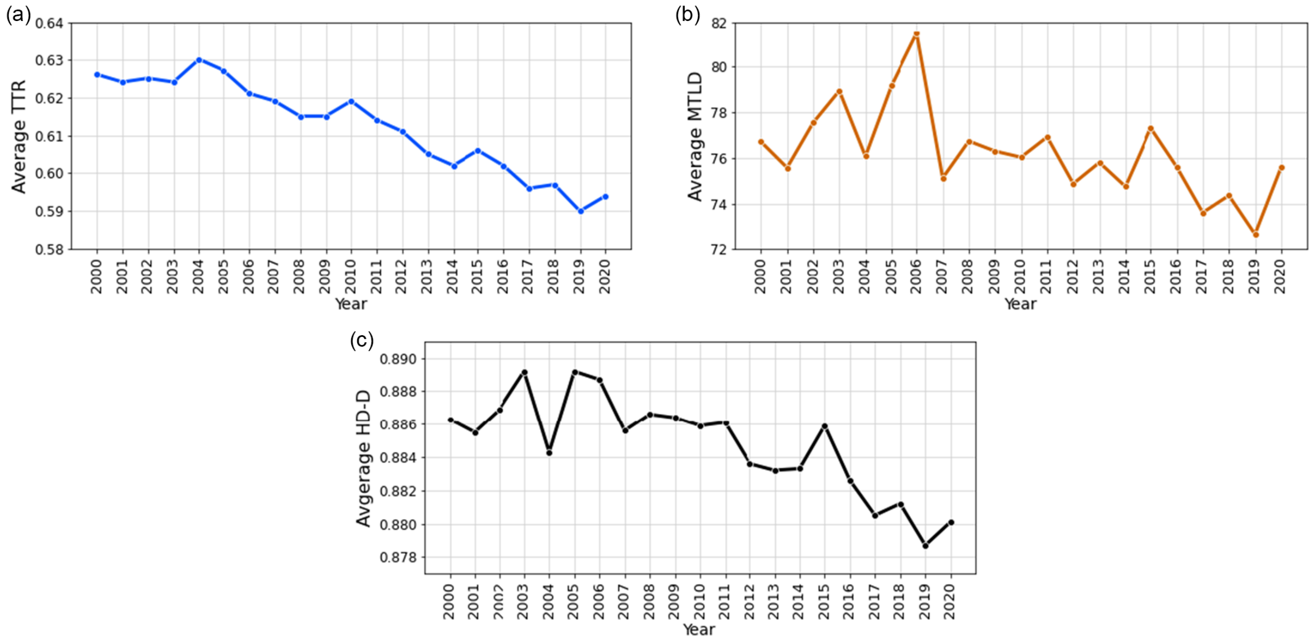
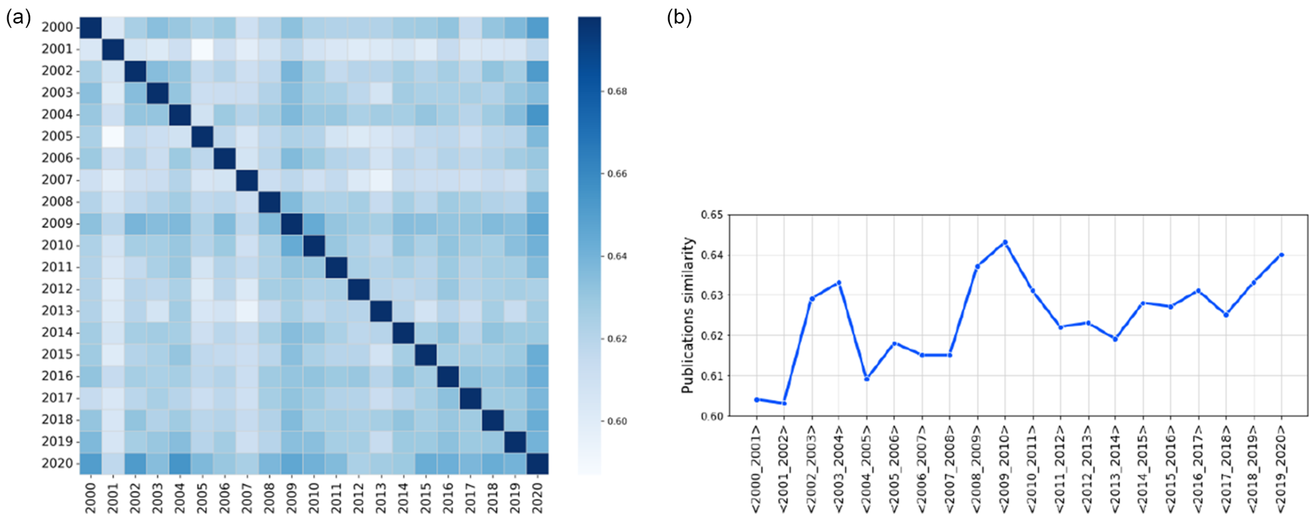


Figure 6
Publications similarity (a) between years research similarity heatmap and (b) consecutive years similarity



hypersonics at an abstract level, based on the collected scientific publications, and by no means do they capture all the details in the examined research area. STM allows topics to be correlated. To analyze interrelationships between the extracted topics and indicate topics that are likely to co-occur within the same publication, topic correlation analysis was performed. As seen in Figure 8(a), most of the extracted topics either do not correlate or are negatively correlated meaning that it is very unlikely for them to occur in the same paper. Figure 8(b) shows the topics correlation graph in which two nodes, i.e., topics, are connected if they are positively correlated. The thickness of an edge linking two nodes represents the correlation coefficient. As observed, the correlation coefficient is not high. The four topic pairs that are positively correlated in descending order are: <Topic 12, Topic 7>, <Topic 5, Topic 10>, <Topic 1, Topic 10>, and <Topic 6, Topic 9>. One may note that these correlations might have been affected by journal coverage and indexing over time. However, for instance, a positive correlation between Topic 12, i.e., guidance against hypersonic targets, and Topic 7, i.e., tracking control of hypersonic flight vehicles (HFVs) is reflected in the literature as well in, e.g., guidance and control design of interceptors (Song & Zhang, 2015), guidance algorithms for hypersonic reentry vehicles (Wu et al., 2023; Zang et al., 2019), and guidance and control for hypersonic missiles (Liu et al., 2017),

Figure 7
Quality of the extracted research themes based on semantic coherence and exclusivity of their representative keywords

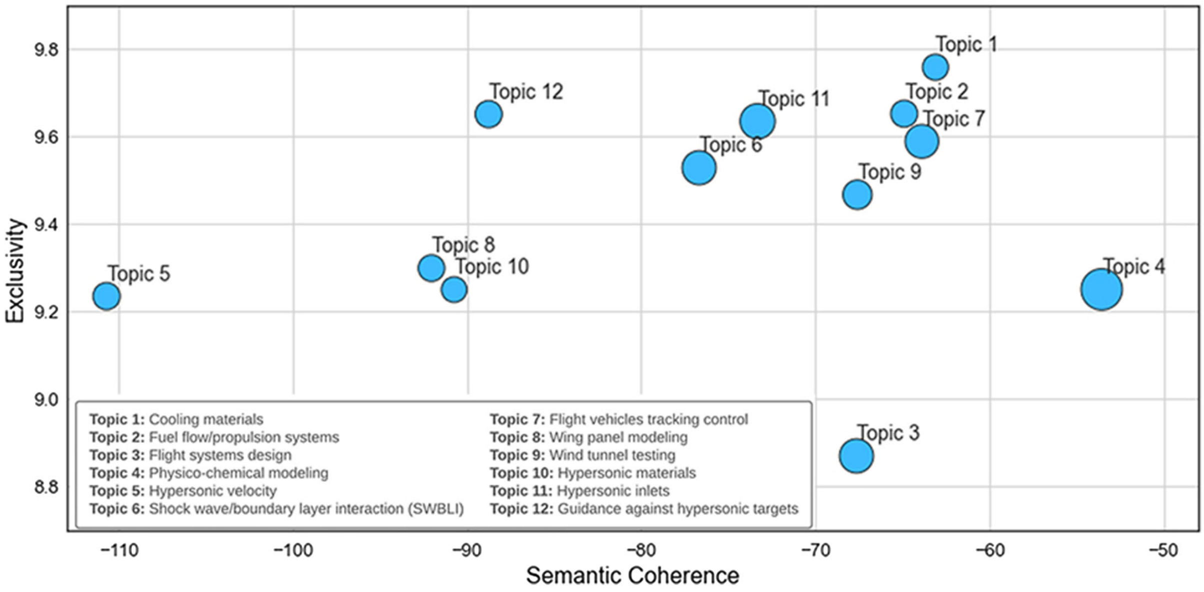
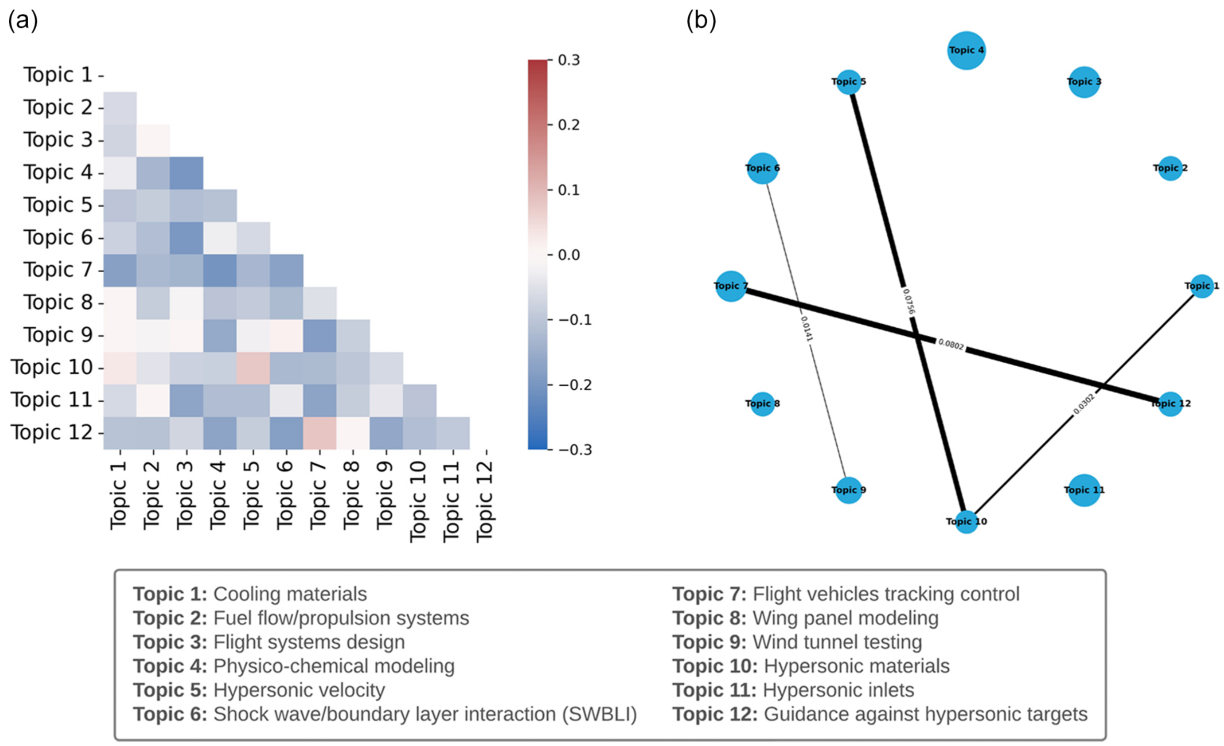


Figure 8
Topics correlations (a) correlation heatmap and (b) positive correlations graph



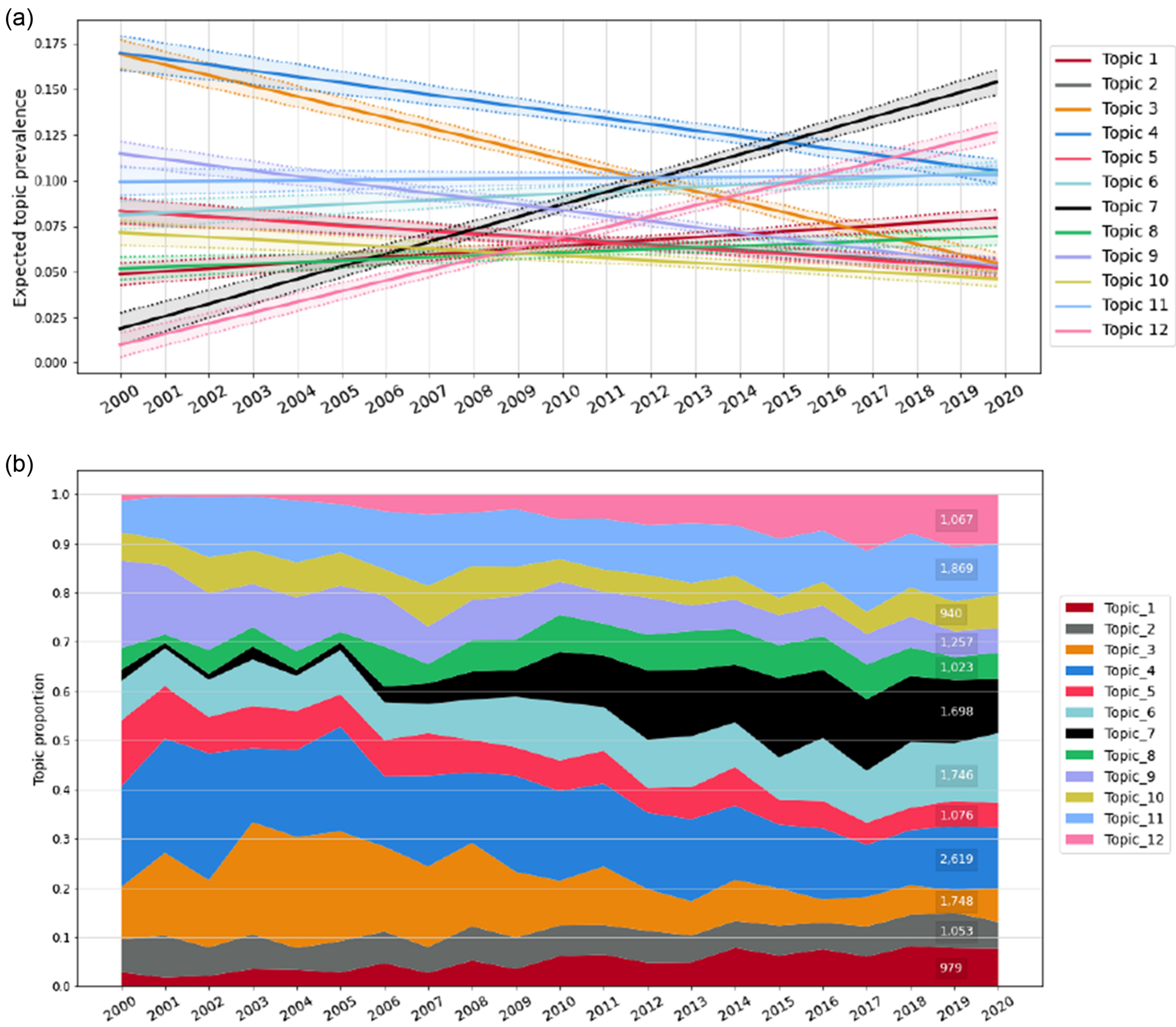
to name a few. The uncertainties in HFVs dynamics due to their specific characteristics such as fast time-varying flight environment and strong dynamics coupling (Li et al., 2020), as well as a non-linearities enhancement as a result of the integration of the propulsion system and the body, have made HFVs' controller design a complicated task (Bu et al., 2015) that requires expertise in several sub-fields. The topic correlation analysis indicates the existence of more distinct research areas, despite the observed diversity and overlaps between a few topics.

3.5.2. Evolution of the research themes

After building the STM model and extracting the main research topics in order to analyze the evolution of topics over time, the proportion of each publication was regressed on the date of publication. More specifically, the conditional expectation of topic prevalence given the characteristics of the publications and their date of publication was estimated. Figure 9(a) shows the results. Shaded areas in the figure represent the 95% confidence interval. As seen in the figure, 6 topics, i.e., “Topic 1," “Topic 6," “Topic 7," “Topic 8," “Topic 11," and “Topic 12," followed an increasing trend over time and the other 6 topics' prevalence decreased. Among topics with an increasing trend, “Topic 7," i.e., flight vehicles tracking control, and “Topic 12," i.e., guidance against hypersonic targets, had the steepest increase. On the other hand, “Topic 3," i.e., flight systems design, and “Topic 4," i.e., physico-chemical modeling, had the sharpest decrease among topics with a decreasing trend. The other observation is that at the beginning of the period, researchers' focus was mostly on “Topic 3" and “Topic 4," however, the attention was shifted more to “Topic 7" and “Topic 12" in the final period. In addition, from the figure, it

Figure 9

Temporal evolution of hypersonics research topics (a) topic prevalence in publications from 2000 to 2020. The shaded areas between the dotted lines indicate the 95% confidence interval and (b) dominant topic distribution across publications over time. The numbers on the curves in the figure represent the total number of publications dominated by the respective topic



Topic 1: Cooling materials
Topic 2: Fuel flow/propulsion systems
Topic 3: Flight systems design
Topic 4: Physico-chemical modeling
Topic 5: Hypersonic velocity
Topic 6: Shock wave/boundary layer interaction (SWBLI)
Topic 7: Flight vehicles tracking control
Topic 8: Wing panel modeling
Topic 9: Wind tunnel testing
Topic 10: Hypersonic materials
Topic 11: Hypersonic inlets
Topic 12: Guidance against hypersonic targets

is seen that researchers' focus on "Topic 11," i.e., hypersonic inlets, was almost constant over time. "Topic 7," "Topic 12," and "Topic 4" are the top-3 most prevalent topics in 2020.

Figure 9(b) shows the results for the distribution of dominant topics across publications over time. In STM, each publication can cover more than one topic as topics are assigned to publications with a probability. To calculate the distribution of dominant topics across publications, the publication-topic probability matrix was considered. For each publication, only the topic with the highest probability was assigned. As seen in the figure, the results are in line with our previous observations in Figure 9(a). "Topic 3" and "Topic 4" were the most dominating topics in the beginning years. Although "Topic 4" proportion has decreased over time, it still remains one of the most dominating topics in the final period in terms of publication covered, placing it among the top-3 dominating topics in 2020 next to "Topic 6," i.e., SWBLI, and "Topic 7." Among these topics, "Topic 7" was the one with the lowest proportion in the beginning years. Similarly, the attention to "Topic 12" has increased drastically over time.

4. Discussion and Conclusion

The field of hypersonics is wide, and complex, and has been witnessing tremendous developments in recent years. With defence and commercial applications, it is said that the technology has the potential to transform the industry (Deloitte, 2020). In this study, we focused on hypersonic-related scientific publications within the period from 2000 to 2020 and employed machine learning and NLP techniques to characterize the research landscape and its evolution over time. We measured publication similarity over time and identified some patterns that are indicative of cycles during two decades of research. We also used structural topic modeling to identify 12 key topics in hypersonics research. The identified topics offer comprehensive and logical coverage of the research field and are relatively similar to the topics being used to structure review papers (Sziroczak & Smith, 2016) or scientometric studies (Senay, 2017) around hypersonics. The fact that topics are algorithmically extracted removes subjectivity from the exercise and enables consistent comparisons between topics and between time intervals.

The number of publications has been growing in almost all scientific disciplines. It was observed that the field of hypersonics has been no exception with publications increasing over time. However, our findings suggest that the focus on research topics has been shifting gradually, due to the dynamic nature of the science and in order to reflect changing needs. Based on the results, although the focus was mainly on the design of hypersonic flight systems and physico-chemical modeling in hypersonic flow simulation, in the beginning, non-linear adaptive tracking control of HFVs and robust guidance against hypersonic targets attracted more attention in the final year. Moreover, research on tracking control of HFVs has experienced the sharpest increase over the examined period, illustrating one of the challenges of hypersonic flight. The importance of optimal performance for hypersonic flight control is also reflected in the literature. For instance, a line of research is investigating a shift from basic control performance including, e.g., stability, and robustness, to the design of fuzzy and reinforcement learning-based optimal tracking controllers for HFVs (Bu & Qi, 2020; Hu et al., 2022).

Aerothermodynamics, propulsion, and structures are identified as the three major challenge areas associated with HVs (Sziroczak & Smith, 2016). These research areas are reflected well in our extracted topics. In addition to these lines of research, our findings also demonstrate an emphasis on modeling and simulation research projects. Of note, recent developments in computer science, e.g., deep learning techniques, have opened up new directions for HVs' trajectory simulation and optimization (Shi & Wang, 2020; Shi & Wang, 2021; Wang et al., 2022). As another example, such powerful machine learning techniques have been also used for hypersonic flight control recently (Xu et al., 2014), a category that is represented well in our extracted topics. In fact, flight control systems have been influenced by artificial neural networks and rapidly evolved during the last two decades (Emami et al., 2022).

The fact that most steps of the workflow can be automated is significant. It means that insights can be generated quickly, even for complex research fields involving thousands of papers every year, to assist analysts and decision-makers in better understanding the research dynamics and help with R&D strategies. This is of particular importance for R&D that requires fast progress, e.g., COVID-19 research (Ebadi et al., 2021), or disruptive technology development potentially affecting strategic stability (Sechser et al., 2019), economic development (Rifkin, 2011), and national security. With defence and commercial applications, hypersonics research will continue to bring innovation to aeronautic and space industries (Deloitte, 2020), but other examples of such disruptive technologies abound, e.g., deep learning (LeCun et al., 2015), quantum computing (Gyongyosi & Imre, 2019), mRNA vaccines (Sahin et al., 2014). The nations and organizations able to best understand and monitor the research landscape around them will have a competitive advantage.

5. Limitations and Future Work

We used scientific publications in the period from 2000 to 2020 to characterize the landscape of hypersonics research. Other data sources, e.g., patents, and time intervals could be used to perform complementary investigations. It should be noted that our findings may only reflect the focus of researchers in the field of hypersonics at a very high level. Future works may consider other levels of abstraction using our proposed methodology. Temporal fusion and/or division of research themes can be also investigated in future research. In addition, this work was based on the analysis of uni-grams (tokens) and could benefit from an *n*-gram approach in the future. Finally, we performed the analysis using titles and abstracts of publications as we did not have access to the full text of the collected publications. Future work may consider the full body of publications.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data are retrievable by running the search query mentioned in the manuscript on the data source, i.e., Scopus. In addition, the raw data can be made available upon request. Please contact the corresponding author.

Code Availability Statement

Unfortunately, we are not able to share the code due to the intellectual property involved. However, interested readers can contact the corresponding author if they have any questions regarding the code.

Authors' Contributions

Conceptualization: AE. Data collection: AE. Data curation: AE. Methodology and design of experiments: AE. Experimentation and data analysis: AE. Interpretation of results: AA, YG, and AE. Wrote the manuscript: AE, AA, and YG.

References

Anderson, Jr., J. (1984). A survey of modern research in hypersonic aerodynamics. In *17th Fluid Dynamics, Plasma Dynamics, and Lasers Conference*, *1578*.

Bakos, R. (2008). Current hypersonic research in the USA. *Advances on Propulsion Technology for High-Speed Aircraft*, *10*, 1–26.

Bhat, M. R., Kundroo, M. A., Tarray, T. A., & Agarwal, B. (2020). Deep LDA: A new way to topic model. *Journal of Information and Optimization Sciences*, *41*(3), 823–834. https://doi.org/10.1080/02522667.2019.1616911

Bischof, J., & Airoldi, E. M. (2012). Summarizing topical content with word frequency and exclusivity. In *Proceedings of the 29th International Conference on Machine Learning*, 201–208.

Bittermann, A., & Fischer, A. (2018). How to identify hot topics in psychology using topic modeling. *Zeitschrift für Psychologie*, *226*(1).

Blankson, I., & Pyle, J. (1993). NASA's hypersonic flight research program. In *31st Aerospace Sciences Meeting*, 308.

Blasius, J., & Greenacre, M. (1998). *Visualization of categorical data*. USA: Academic Press.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*(4), 993–1022. https://doi.org/10.1016/b978-0-12-411519-4.00006-9

Bornmann, L., & Mutz, R. (2015). Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. *Journal of the Association for Information Science and Technology*, *66*(11), 2215–2222. https://doi.org/10.1002/asi.23329

Bu, X., & Qi, Q. (2020). Fuzzy optimal tracking control of hypersonic flight vehicles via single-network adaptive critic design. *IEEE Transactions on Fuzzy Systems*, *30*(1), 270–278. https://doi.org/10.1109/TFUZZ.2020.3036706

Bu, X., Wu, X., Zhu, F., Huang, J., Ma, Z., & Zhang, R. (2015). Novel prescribed performance neural control of a flexible air-breathing hypersonic vehicle with unknown initial errors. *ISA Transactions*, *59*, 149–159. https://doi.org/10.1016/j.isatra.2015.09.007

Carioscia, S. A., Locke, J. W., Boyd, L. D., Lewis, M. J., HalionSun, R. P., & Smith, H. (2019). *Commercial development of civilian supersonic aircraft*. USA: Institute for Defense Analyses.

Czysz, P., & Vandenkerckhove, J. (2001). Transatmospheric launcher sizing. In S. N. B. Murthy & E. T. Curran (Eds.), *Scramjet propulsion* (pp. 979–1103). American Institute of Aeronautics and Astronautics.

Das, S., Sun, X., & Dutta, A. (2016). Text mining and topic modeling of compendiums of papers from transportation research board annual meetings. *Transportation Research Record*, *2552*(1), 48–56. https://doi.org/10.3141/2552-07

Deloitte. (2020). *The rise of hypersonics*. Retrieved from: https://www2.deloitte.com/us/en/pages/energy-and-resources/articles/rise-of-hypersonics.html

Ding, Y., Yue, X., Chen, G., & Si, J. (2022). Review of control and guidance technology on hypersonic vehicle. *Chinese Journal of Aeronautics*, *35*(7), 1–18. https://doi.org/10.1016/j.cja.2021.10.037

Doré, J. C., & Ojasoo, T. (2001). How to analyze publication time trends by correspondence factor analysis: Analysis of publications by 48 countries in 18 disciplines over 12 years. *Journal of the American Society for Information Science and Technology*, *52*(9), 763–769. https://doi.org/10.1002/asi.1130

Ebadi, A., Auger, A., & Gauthier, Y. (2022). Detecting emerging technologies and their evolution using deep learning and weak signal analysis. *Journal of Informetrics*, *16*(4). https://doi.org/10.1016/j.joi.2022.101344

Ebadi, A., Tremblay, S., Goutte, C., & Schiffauerova, A. (2020). Application of machine learning techniques to assess the trends and alignment of the funded research output. *Journal of Informetrics*, *14*(2). https://doi.org/10.1016/j.joi.2020.101018

Ebadi, A., Xi, P., Tremblay, S., Spencer, B., Pall, R., & Wong, A. (2021). Understanding the temporal evolution of COVID-19 research through machine learning and natural language processing. *Scientometrics*, *126*(1), 725–739. https://doi.org/10.1007/s11192-020-03744-7

Emami, S. A., Castaldi, P., & Banazadeh, A. (2022). Neural network-based flight control systems: Present and future. *Annual Reviews in Control*, *53*, 97–137. https://doi.org/10.1016/j.arcontrol.2022.04.006

Facts & Factors. (2022). *Hypersonic technology market size, share global analysis report, 2022-2028*. Retrieved from: https://www.fnfresearch.com/hypersonic-technology-market.

Fergadiotis, G., Wright, H. H., & West, T. M. (2013). Measuring lexical diversity in narrative discourse of people with aphasia. *American Journal of Speech-Language Pathology*, *22*(2), 397–408. https://doi.org/10.1044/1058-0360(2013/12-0083).

Gatti, C. J., Brooks, J. D., & Nurre, S. G. (2015). A historical analysis of the field of OR/MS using topic models. *arXiv Preprint:1510.05154*.

Greenacre, M., & Blasius, J. (2006). *Multiple correspondence analysis and related methods*. USA: CRC Press.

Gu, S., & Olivier, H. (2020). Capabilities and limitations of existing hypersonic facilities. *Progress in Aerospace Sciences*, *113*. https://doi.org/10.1016/j.paerosci.2020.100607

Gyongyosi, L., & Imre, S. (2019). A survey on quantum computing technology. *Computer Science Review*, *31*, 51–71. https://doi.org/10.1016/j.cosrev.2018.11.002

Hu, X., Dong, K., Yang, T., & Xiao, B. (2022). Reinforcement learning based optimal tracking control for hypersonic flight vehicle: A model free approach. In *IEEE 20th International Conference on Industrial Informatics*, 711–717.

Ingenito, A., Gulli, S., Bruno, C., Colemann, G., Chudoba, B., & Czysz, P. A. (2011). Sizing of a fully integrated hypersonic commercial airliner. *Journal of Aircraft*, *48*(6), 2161–2164. https://doi.org/10.2514/1.C000205

Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: A developmental perspective. *Working papers/Lund University, Department of Linguistics and Phonetics*, *53*, 61–79.

Keller, J., & von der Gracht, H. A. (2014). The influence of information and communication technology (ICT) on future foresight processes—Results from a Delphi survey.

Technological Forecasting and Social Change, *85*, 81–92. https://doi.org/10.1016/j.techfore.2013.07.010

Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (automated readability index, fog count and Flesch reading ease formula) for navy enlisted personnel*. USA: Institute for Simulation and Training.

Kostoff, R. N., Eberhart, H. J., & Toothman, D. R. (1999). Hypersonic and supersonic flow roadmaps using bibliometrics and database tomography. *Journal of the American Society for Information Science*, *50*(5), 427–447. https://doi.org/10.1002/(SICI)1097-4571(1999)50:5<427::AID-ASI5>3.0.CO;2-P.

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. USA: Sage publications.

Kuhn, K. D. (2018). Using structural topic modeling to identify latent topics and trends in aviation incident reports. *Transportation Research Part C: Emerging Technologies*, *87*, 105–122. https://doi.org/10.1016/j.trc.2017.12.018

Le, Q., & Mikolov, T. (2014). Distributed representations of sentences and documents. In *International Conference on Machine Learning*, 1188–1196.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Li, Z., Zhou, W., & Liu, H. (2020). Robust controller design of non-minimum phase hypersonic aircrafts model based on quantitative feedback theory. *The Journal of the Astronautical Sciences*, *67*(1), 137–163. https://doi.org/10.1007/s40295-019-00187-y.

Lissón, P., & Ballier, N. (2018). Investigating lexical progression through lexical diversity metrics in a Corpus of French L3. Discours. *Revue de linguistique, psycholinguistique et informatique. A Journal of Linguistics, Psycholinguistics and Computational Linguistics*, *23*.

Liu, X., Huang, W., & Du, L. (2017). An integrated guidance and control approach in three-dimensional space for hypersonic missile constrained by impact angles. *ISA Transactions*, *66*, 164–175. https://doi.org/10.1016/j.isatra.2016.10.008

Lucas, C., Nielsen, R. A., Roberts, M. E., Stewart, B. M., Storer, A., & Tingley, D. (2015). Computer-assisted text analysis for comparative politics. *Political Analysis*, *23*(2), 254–277. https://doi.org/10.1093/pan/mpu019

Malinowski, P. (2020). Hypersonic weapon as a new challenge for the anti-aircraft defense command and control system. *Safety & Defense*, *6*(2), 89–99. https://doi.org/10.37105/sd.87

Maskeri, G., Sarkar, S., & Heafield, K. (2008). Mining business topics in source code using latent Dirichlet allocation. In *Proceedings of the 1st India Software Engineering Conference*, 113–120.

McCarthy, P. M. (2005). *An assessment of the range and usefulness of lexical diversity measures and the potential of the measure of textual, lexical diversity (MTLD)*. Doctoral Dissertation, The University of Memphis.

McCarthy, P. M., & Jarvis, S. (2007). VOCD: A theoretical and empirical evaluation. *Language Testing*, *24*(4), 459–488. https://doi.org/10.1177/0265532207080767

McDonald, S. (2021). *Defense spending act makes hypersonic a top priority, calling for billions in investment*. Retrieved from: https://www.cnbc.com/2021/12/29/defense-act-makes-hypersonic-top-priority-calls-for-investing-billions.html

McFarland, J. (2023). The development of hypersonic weapons in the US, China and Russia: An incipient arms race. *The RUSI Journal, 168*(1–2), 10–18.

Mohammadi, E., & Karami, A. (2022). Exploring research trends in big data across disciplines: A text mining analysis. *Journal of Information Science*, *48*(1), 44–56. https://doi.org/10.1177/0165551520932855

Mosallaie, S., Rad, M., Schiffauerova, A., & Ebadi, A. (2021). Discovering the evolution of artificial intelligence in cancer research using dynamic topic modeling. *COLLNET Journal of Scientometrics and Information Management*, *15*(2), 225–240. https://doi.org/10.1080/09737766.2021.1958659

Ozyurt, O., & Ayaz, A. (2022). Twenty-five years of education and information technologies: Insights from a topic modeling based bibliometric analysis. *Education and Information Technologies*, *27*(8), 11025–11054. https://doi.org/10.1007/s10639-022-11071-y.

Papadimitriou, C. H., Raghavan, P., Tamaki, H., & Vempala, S. (2000). Latent semantic indexing: A probabilistic analysis. *Journal of Computer and System Sciences*, *61*(2), 217–235. https://doi.org/10.1006/jcss.2000.1711

Piplica, A. J. (2021). *Why invest in hypersonic technology?* Retrieved from: https://www.defenseone.com/ideas/2021/03/why-invest-hypersonic-technology/172361/

Richman, M., Kenyon, J., & Sega, R. (2019). High speed and hypersonic science and technology. In *41st AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, 4099*.

Rifkin, J. (2011). *The third industrial revolution: How lateral power is transforming energy, the economy, and the world*. UK: Palgrave MacMillan.

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software*, *91*, 1–40. https://doi.org/10.18637/jss.v091.i02

Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, 399–408. https://doi.org/10.1145/2684822.2685324

Sahin, U., Karikó, K., & Türeci, Ö. (2014). mRNA-based therapeutics—Developing a new class of drugs. *Nature Reviews Drug Discovery*, *13*(10), 759–780. https://doi.org/10.1038/nrd4278

Sechser, T. S., Narang, N., & Talmadge, C. (2019). Emerging technologies and strategic stability in peacetime, crisis, and war. *Journal of Strategic Studies*, *42*(6), 727–735. https://doi.org/10.1080/01402390.2019.1626725

Senay, M. (2017). *Hypersonic/hypervelocity vehicles and weapons scientometric study*. Canada: National Research Council Canada.

Shi, Y., & Wang, Z. (2020). A deep learning-based approach to real-time trajectory optimization for hypersonic vehicles. In *AIAA Scitech 2020 Forum*.

Shi, Y., & Wang, Z. (2021). Onboard generation of optimal trajectories for hypersonic vehicles using deep learning. *Journal of Spacecraft and Rockets*, *58*(2), 400–414. https://doi.org/10.2514/1.A34670

Song, H., & Zhang, T. (2015). Fast robust integrated guidance and control design of interceptors. *IEEE Transactions on Control Systems Technology*, *24*(1), 349–356. https://doi.org/10.1109/TCST.2015.2431641

Stout, N. L., Alfano, C. M., Belter, C. W., Nitkin, R., Cernich, A., Lohmann Siegel, K., & Chan, L. (2018). A bibliometric analysis of the landscape of cancer rehabilitation research (1992–2016). *Journal of the National Cancer Institute*, *110*(8), 815–824. https://doi.org/10.1093/jnci/djy108

Sun, L., & Yin, Y. (2017). Discovering themes and trends in transportation research using topic modeling. *Transportation*

Research Part C: Emerging Technologies, *77*, 49–66. https://doi.org/10.1016/j.trc.2017.01.013
Sziroczak, D., & Smith, H. (2016). A review of design issues specific to hypersonic flight vehicles. *Progress in Aerospace Sciences*, *84*, 1–28. https://doi.org/10.1016/j.paerosci.2016.04.001
Tiron, T. (2021). *Hypersonic weapons: Who has them and why it matters*. Retrieved from: https://www.bloomberg.com/news/articles/2021-10-27/hypersonic-weapons-who-has-them-and-why-it-matters-quicktake#xj4y7vzkg
Treffers-Daller, J., Parslow, P., & Williams, S. (2018). Back to basics: How measures of lexical diversity can help discriminate between CEFR levels. *Applied Linguistics*, *39*(3), 302–327.
Van Wie, D. (2021). Hypersonics: Past, present, and potential future. *Johns Hopkins APL Technical Digest*, *35*(4), 335–341.
Viola, N., Fusaro, R., Ferretto, D., & Vercella, V. (2023). Research, development and production costs prediction parametric model for future civil hypersonic aircraft. *Acta Astronautica*, *204*, 58–72. https://doi.org/10.1016/j.actaastro.2022.12.036
Viviani, A., & Pezzella, G. (2019). Introductory chapter: Hypersonic vehicles-past, present, and future insights. In G. Pezzella & A. Viviani (Eds.), *Hypersonic vehicles-past, present and future developments*. IntechOpen.
Wang, J., Wu, Y., Liu, M., Yang, M., & Liang, H. (2022). A real-time trajectory optimization method for hypersonic vehicles based on a deep neural network. *Aerospace*, *9*(4), 188. https://doi.org/10.3390/aerospace9040188
Wu, T., Wang, H., Liu, Y., Li, T., & Yu, Y. (2023). Learning-based interfered fluid avoidance guidance for hypersonic reentry vehicles with multiple constraints. *ISA Transactions*, *139*, 291–307. https://doi.org/10.1016/j.isatra.2023.04.004
Xu, B., Pan, Y., Wang, D., & Sun, F. (2014). Discrete-time hypersonic flight control based on extreme learning machine. *Neurocomputing*, *128*, 232–241. https://doi.org/10.1016/j.neucom.2013.02.049
Yoon, Y. S., Lee, J., & Park, K. (2019). Extracting promising topics on smart manufacturing based on latent Dirichlet allocation (LDA). In *International Conference on Information and Communication Technology Convergence*, 1237–1242.
Zang, L., Lin, D., Chen, S., Wang, H., & Ji, Y. (2019). An on-line guidance algorithm for high L/D hypersonic reentry vehicles. *Aerospace Science and Technology*, *89*, 150–162. https://doi.org/10.1016/j.ast.2019.03.052
Zeng, A., Shen, Z., Zhou, J., Fan, Y., Di, Z., Wang, Y., . . . , & Havlin, S. (2019). Increasing trend of scientists to switch between topics. *Nature Communications*, *10*(1), 1–11.

How to Cite: Ebadi, A., Auger, A., & Gauthier, Y. (2024). Machine Learning Insights into Hypersonics Research Evolution: A 21st Century Perspective. *Archives of Advanced Engineering Science*, *2*(2), 79–92. https://doi.org/10.47852/bonviewAAES32021471