

JOURNAL OF ENGINEERING RESEARCH & SCIENCES

JENRS



www.jenrs.com
ISSN: 2831-4085

Volume 4 Issue 1
January 2025

EDITORIAL BOARD

Editor-in-Chief

Prof. Paul Andrew
Universidade De São Paulo, Brazil

Editorial Board Members

Dr. Jianhang Shi

Department of Chemical and Biomolecular
Engineering, The Ohio State University, USA

Prof. Kamran Iqbal

Department of Systems Engineering, University
of Arkansas Little Rock, USA

Dr. Lixin Wang

Department of Computer Science, Columbus
State University, USA

Dr. Unnati Sunilkumar Shah

Department of Computer Science, Utica
University, USA

Dr. Qichun Zhang

Department of Computer Science, University of
Bradford, UK

Dr. Prabhash Dadhich

Biomedical Research, CellfBio, USA

Dr. Qiong Chen

Navigation College, Jimei University, China

Ms. Madhuri Inupakutika

Department of Biological Science, University of
North Texas, USA

Dr. Jianhui Li

Molecular Biophysics and Biochemistry,
Yale University, USA

Dr. Sonal Agrawal

Rush Alzheimer's Disease Center, Rush
University Medical Center, USA

Dr. Ramcharan Singh Angom

Biochemistry and Molecular Biology,
Mayo Clinic, USA

Dr. Anna Formica

National Research Council, Istituto di
Analisi dei Sistemi ed Informatica, Italy

Prof. Anle Mu

School of Mechanical and Precision
Instrument Engineering, Xi'an University
of Technology, China

Dr. Mingsen Pan

University of Texas at Arlington, USA

Dr. Żywiłek Justyna

Faculty of Management, Czestochowa
University of Technology, Poland

Dr. Diego Cristallini

Department of Signal Processing &
Imaging Radar, Fraunhofer FHR, Germany

Dr. Haiping Xu

Computer and Information Science Department, University of Massachusetts Dartmouth, USA

Editorial

The pursuit of technological advancement continues to reshape the boundaries of modern research, with recent developments illuminating new paths across software engineering, environmental sustainability, and maritime innovation. The selected research papers in this issue provide thought-provoking insights into how emerging technologies are being harnessed to solve complex challenges in their respective domains. From enhancing automated code comprehension to paving the way for a carbon-neutral economy and transforming maritime navigation through autonomous systems, these studies reflect the interdisciplinary nature and global relevance of contemporary scientific inquiry.

Recent efforts in automated code comprehension have leveraged advances in neural networks to bridge the gap between syntactic code structure and its semantic interpretation. This study introduces an enhanced hybrid method that combines code2vec embeddings with representations derived from large language models (LLMs), integrating both the structural and contextual understanding of Python code. Multiple fusion strategies are examined—such as concatenation, weighted sums, and attention mechanisms—to synthesize the strengths of each representation. The work emphasizes the advantages of this integrated approach for improving tasks like code search, documentation generation, and intelligent software development tools, ultimately pointing toward more robust and efficient development workflows [1].

In the context of global carbon neutrality goals, the exploration of recycled carbon fuels such as e-methane has emerged as a viable path forward. The study evaluates the feasibility of establishing international value chains for e-methane, focusing on optimal production site placement, resource accessibility, and the development of harmonized greenhouse gas (GHG) accounting rules. A notable contribution is the proposed carbon footprint (CFP) calculation method, which has been adopted in the ISO 6338-1:2024 standard. Additionally, the work highlights Japan's "Clean Gas Certificate" initiative as a case study, illustrating practical steps toward responsible certification and commercialization. These efforts address the crucial need for environmental accountability and standardization in carbon fuel markets during the 2030–2050 transition period [2].

Technological innovations in the maritime sector are being revolutionized by the emergence of Maritime Autonomous Surface Ships (MASS). Utilizing the Technological Innovation Systems (TIS) framework, this research dissects the complex ecosystem around MASS, examining market barriers and proposing strategies for successful deployment. By synthesizing insights from scholarly literature and expert interviews, the study identifies economic and systemic challenges, notably the high cost-benefit threshold, and responds with strategic recommendations such as niche market cultivation, engagement of early adopters, and hybridization models. These insights are instrumental in charting a roadmap for the mainstream adoption of autonomous maritime technologies in global shipping industries [3].

Collectively, these research contributions underline the vital interplay between technological innovation, environmental stewardship, and systemic transformation. They serve not only to advance their individual fields but also to inspire collaborative approaches to solving globally significant problems. The forward-looking perspectives offered here reaffirm the critical role of interdisciplinary research in driving sustainable and intelligent progress.

References:

- [1] L. H. Ngo, J. Rivalan, "Enhancing Python Code Embeddings: Fusion of Code2vec with Large Language Models," *Journal of Engineering Research and Sciences*, vol. 04, no. 01, pp. 1–7, 2025, doi:10.55708/js0401001.

- [2] R. Kuzuki, M. Kohara, N. Kizuki, S. Yoshida, Y. Nezasa, Y. Tsuji, "Developing E-Methane Value Chain and Proper Greenhouse Gas (GHG) Accounting Rules Incentivizing Recycled Carbon Fuels," *Journal of Engineering Research and Sciences*, vol. 04, no. 01, pp. 8–15, 2025, doi:10.55708/js0401002.
- [3] M. Kurtinaitis Joukes, R. Ortt, M. de Bruijne, L. M. Kamp, "How to analyze the introduction strategies for radically new technological innovations? The case of Autonomous Shipping," *Journal of Engineering Research and Sciences*, vol. 04, no. 01, pp. 16–30, 2025, doi:10.55708/js0401003.

Editor-in-chief

Prof. Paul Andrew

CONTENTS

<i>Enhancing Python Code Embeddings: Fusion of Code2vec with Large Language Models</i>	01
Long H. Ngo and Jonathan Rivalan	
<i>Developing E-Methane Value Chain and Proper Greenhouse Gas (GHG) Accounting Rules Incentivizing Recycled Carbon Fuels</i>	08
Ryota Kuzuki, Mitsuhiro Kohara, Noboru Kizuki, Satoshi Yoshida, Yuta Nezasa and Yuki Tsuji	
<i>How to Analyze the Introduction Strategies for Radically New Technological Innovations? The Case of Autonomous Shipping</i>	16
Anantharaj Thalaimalai Vanaraj and Reshi Razdan	

Enhancing Python Code Embeddings: Fusion of Code2vec with Large Language Models

Long H. Ngo^{*} , Jonathan Rivalan 

Smile France, Asnières-sur-Seine, 92600, France

^{*}Corresponding author: Long H. Ngo, Paris, France, Email: long.ngo@smile.fr

ABSTRACT: Automated code comprehension has recently become integral to software development. Neural networks, widely employed in natural language processing tasks, can capture the semantic meanings of language by representing it in vector form. Although programming code differs from natural language, we hypothesize that neural models can learn both the syntactic and semantic attributes inherent in code. This study presents an innovative approach to improve code representation and understanding for Python, building upon a previous work (code2vec extended with ASTminer). The novel method integrates embeddings from Large Language Models (LLMs) with code2vec vectors, aiming to align both semantic and syntactic information in code representations. We explore various fusion techniques, including simple concatenation, weighted sum, or attention-based mechanism, to combine LLM embeddings with code2vec vectors. We explore how semantic information from LLMs complements the structural information from code2vec, and discuss the potential impact of this synergy on software development practices. These findings open new directions for more accurate and adaptable code understanding models, with implications for improving documentation, code search, and overall software development efficiency.

KEYWORDS: Machine learning, Neural network, Large Language Model, Distributed representations, Code search

1. Introduction

In recent years, the field of automated code understanding has become increasingly crucial in software development. Certain problems in software development, such as assigning meaningful method names, highlight the need for concise semantic representations of code snippets. The core challenge lies in encoding code snippets in a way that captures semantically relevant information, transferable across multiple programs, and enables the prediction of properties like code labeling. This involves two key components: first, representing the code snippet in a manner that supports learning across different programs, and second, determining which parts of the representation are critical for property prediction and how to prioritize them.

This paper builds upon a previous study in [1] at the International Conference on Software Engineering and Artificial Intelligence (SEAI), by addressing its limitations and introducing novel methodology to improve code representation and search capabilities. Code representation presents challenges due to the need to capture both semantic and syntactic information in a format conducive to machine learning. The previous study, in [1], used a path-based representation technique to address two key tasks in software development: assigning semantic labels to code snippets and performing code searches. The path-based method represents a snippet by extracting paths from its abstract syntax tree (AST), capturing both the structure and semantics of the code, as demonstrated in previous research [2]. The first task focused on automating the prediction of a semantic

label for a given code snippet. This is a challenging task because it requires learning complex relationships between the content of a method and a semantic label. Specifically, it involves condensing numerous expressions and statements within the method into a single descriptive label, which demands sophisticated techniques for code representation and classification [3]. The second task aimed to develop an efficient and effective search mechanism for locating code snippets based on query requirements. This is a critical need for developers, who often need to find and reuse existing code. Successful code search must be able to match a query against a large codebase and retrieve relevant code snippets. Addressing both of these tasks effectively would significantly enhance software development productivity and efficiency.

However, the work in [1] also revealed limitations, particularly in aligning the docstrings with extracted paths from the input code snippets. It might miss out on high-level semantic relationships. This challenge highlighted the need for more sophisticated approaches to code representation.

In this extended version, we present several significant advancements:

- Integration of large language models (LLMs) with code2vec to create more comprehensive code embeddings.
- Development of fusion techniques to combine semantic embeddings from LLMs with syntactic vectors from code2vec.

- Extensive experimental settings with various fusion strategies, including concatenation, weighted sum, and attention-based mechanisms.

By addressing these areas, we aim to push the boundaries of code representation, enabling more accurate and efficient automated understanding of Python code. This work not only builds upon the previous findings but also opens new directions for research in the intersection of natural language processing and programming language analysis.

The remainder of this paper is structured as follows: Section 2 reviews related work in code representation and analysis. Section 3 describes the work where code2vec was extended with ASTminer. Section 4 describes the present work, including the adaptation of code2vec for Python and the proposed fusion techniques. Section 5 discusses the implications of our findings. Section 6 concludes the paper with a summary of contributions and directions for future work.

2. Related works

The field of code representation and analysis has seen significant advancements in recent years, driven by the application of deep learning techniques traditionally used in natural language processing (NLP). While the conventional approach in NLP involves treating text as a linear sequence of tokens processed through neural networks [4, 5], this method has also been widely adopted for source code representation [6, 7, 8, 9, 10]. However, recent research has highlighted the potential benefits of leveraging the inherent structure of programming languages. In [11, 12, 13], the authors have demonstrated that structured representations can significantly enhance performance in various code analysis tasks.

The ability to predict program properties through learning from large-scale codebases has been a focus of numerous studies [6, 8, 9, 12, 14]. This capability opens up a wide range of applications in software engineering, including predicting names for program entities [11, 13, 15], code completion [16, 17], code summarization [6], code generation [18, 19, 20]. These applications, among others [21, 22], showcase the potential impact of advanced code representation techniques on various aspects of software development and maintenance, with minimal or no semantic analysis.

A significant breakthrough in this field came with the introduction of code2vec [3]. This innovative approach uses neural networks to learn distributed representations of code, addressing one of the most challenging problems in software engineering. Unlike traditional methods that rely primarily on static code features, code2vec uses a large corpus of code snippets to learn representations that capture the temporal dynamics of code changes. This results in more accurate and robust code representations.

The code2vec model consists of two main components: 1) An input encoder that converts code snippets into token sequences. These sequences are then used to train a neural network to predict the next tokens in the sequence, which is similar to a large language model. 2) A code vectorizer that generates continuous-valued vector representations from the output of the encoder.

These components work together to capture both syntactic and semantic information from the code, utilizing attention mechanisms to prioritize the most relevant aspects of each snippet. The resulting vector representations have proven effective in various software engineering tasks, including code completion, bug detection, and program synthesis. The architecture of the code2vec network is illustrated in Figure 1.

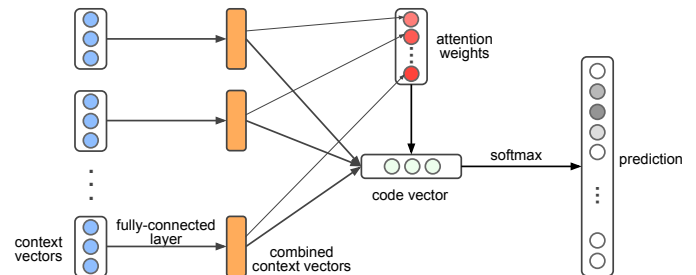


Figure 1: Original code2vec architecture. A fully-connected layer combines embeddings of each path-context. Attention weights are learned and used to compute a code vector, which predicts the label [3].

In [1], the authors adopted the code2vec neural network architecture [3], which was designed to learn low-dimensional vector representations, or embeddings, of the source code. In these vectors, the "meaning" of an entity is distributed across multiple components, allowing semantically similar entities to be mapped to nearby vectors. These embeddings enable efficient modeling of the relationship between code snippets and their semantic labels in a natural and efficient manner. Using the inherent structure of the source code, code2vec aggregated multiple syntactic paths into a single vector, leading to a more comprehensive and accurate representation of each code snippet.

The success of code2vec marks a significant departure from conventional code representation methods and has paved the way for further research in this area. Our work builds upon these foundations, extending the application of code2vec to Python and exploring novel ways to enhance its performance through integration with large language models.

3. Extended Code2vec with ASTminer for Python Code Embeddings

High-level view. In programming, a code snippet plays a pivotal role as it allows developers to write concise, reusable pieces of code applicable in different contexts. To better understand the construction of code snippets, it is essential to explore the concept of a bag of contexts. Each context within a code snippet is represented as a vector, referred to as the context vector. This vector is generated through a learning process that captures two important features of the context: (i) its semantic meaning and (ii) the attention it should receive. By aggregating these context vectors, a code vector is created, which can then be used for various downstream tasks.

3.1. Semantic labeling of code snippets

To extend code2vec for Python programming language, the ASTminer extractor [2] was employed instead of the JavaExtractor in the original code2vec [3]. ASTminer is an open source library that enables the extraction of path-based representations of codes, as shown in Figure 2. ASTminer parses the syntax tree of a code snippet to extract paths that summarize the structural and semantic essence of the code. Hence, it offers a reusable toolkit designed to simplify the task of modeling source code for machine learning algorithms, reducing the associated complexity and effort.

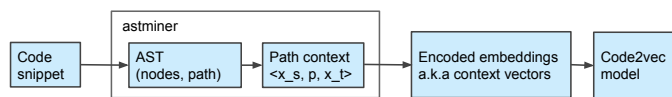


Figure 2: code2vec architecture for Python programming language.

The AST is essential in representing the syntactic structure of a program, selectively ignoring details such as punctuation, formatting or specific syntactic forms, while maintaining each node as a distinct syntactic unit. In the AST, each node represents a specific syntactic construct, such as variables, operations, or logical operators. The associated child nodes correspond to the subordinate elements related to the parent node. [2]

The ASTminer extractor uses this path-based representation to capture coding styles and structures, thereby encapsulating the logic of the code. The conversion of code into embeddings occurs in two steps: first, transforming the code into a vector, followed by combining these vectors with corresponding attention vectors. This prepares the model for further training. The architecture, depicted in Figure 3, mirrors the original code2vec embedding model.

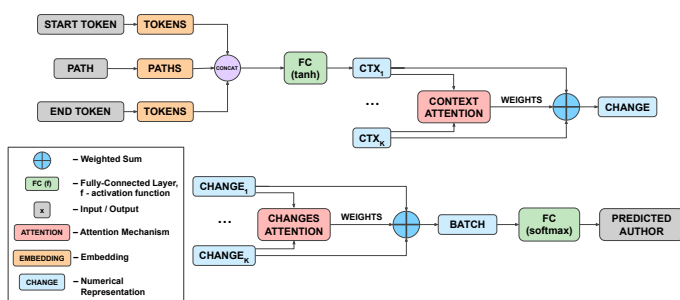


Figure 3: Overview of authorship attribution pipeline, which generates authorship-based embeddings of method changes. Furthermore, it highlights the significance of individual method changes in the authorship attribution. The method nodes and the corresponding attention weights are subsequently utilized to create author representation [23].

Following the approach in [23], the embeddings are constructed using the code and path tokens, then concatenated, and converted to a numerical representation using a fully connected (FC) activation function, as shown in Figure 3. These vectors are merged into batches where the weights of individual paths are combined, revealing the importance of each path. This pipeline is employed for predicting code titles, following the same principles as authorship attribution.

Analogous to [3], cross-entropy loss, which calculates the difference of the predicted distribution q and the “true”

distribution p , is used during training phase. p assigns a value of 1 to the actual label in the training sample and 0 otherwise. Hence, the cross-entropy loss for a single sample is equivalent to the negative log-likelihood of the true label, which can be expressed as follows:

$$\mathcal{H}p||q = -\sum_{y \in Y} p_y \cdot \log q_y = -\log q_{y_{true}} \quad (1)$$

where y_{true} is the actual label of the sample. The loss is the negative logarithm of $q_{y_{true}}$, the probability that the model assigns to y_{true} . As $q_{y_{true}}$ tends to 1, the loss approaches zero. The further $q_{y_{true}}$ goes below 1, the greater the loss becomes. Thus, minimizing this loss is equivalent to maximizing the log-likelihood that the model assigns to the true labels y_{true} . During training, gradient descent is utilized to iteratively update the parameters by minimizing the loss function. The learned parameters are refined using backpropagation. In other words, the parameters are iteratively adjusted by taking small steps in the direction that reduces the loss, thereby optimizing the model.

Based on the study conducted in [3], the network model can be applied effectively to predict categories for unseen code. This method involves generating a code vector by leveraging the weights and parameters learned during the training phase. The prediction is then made by determining the closest target label. This predictive model offers significant potential for improving the accuracy and efficiency of categorizing software code snippets, supporting tasks such as software maintenance, bug detection, and code optimization.

3.2. Code search

This study also builds upon previous works [24, 25, 26] to design a neural search system that uses joint embeddings of code snippets and queries. Each type of input (either code or natural language) is processed using separate encoders, trained to map inputs into a shared vector space. The system is trained to ensure that related code and queries are embedded closely in this space, enabling efficient search and retrieval of relevant code snippets. Although there are more sophisticated models that account for multiple interactions between queries and code, this architecture utilizes a single vector per query/snippet, simplifying indexing and search tasks [26]. Figure 4 provides an overview of the model architecture used for the Python language in the code search task.

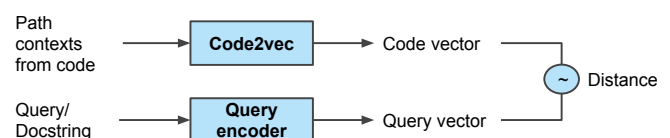


Figure 4: Text to code architecture based on code2vec model.

For this task, the extended code2vec model is used to encode Python code snippets. Queries in the form of docstrings are tokenized and passed through a query encoder, such as the Neural Bag of Words (NBoW) [27], Bidirectional RNN model [28], 1D Convolutional Neural Network [29], or Self-Attention [30]. Subsequently, the resulting token

embeddings are combined into a sequence embedding using a pooling function, either mean or max pooling, and an attention-based weighted sum mechanism. The training process used a collection of N code and docstring pairs, denoted by c_i, d_i , alongside the instantiation of a code encoder E_c and a query encoder E_q . The objective is to minimize the following loss function:

$$-\frac{1}{N} \sum_i \log \frac{\exp(E_c c_i^T E_q d_i)}{\sum_j \exp(E_c c_j^T E_q d_i)} \quad (2)$$

This loss encourages the inner product between matching code and docstring encodings pairs to be maximized, while minimizing the inner product between the query d_i and irrelevant code snippets c_j . During model evaluation, Mean Reciprocal Rank (MRR) is used to assess performance on the validation set. During the testing stage, the Annoy library¹, a high-speed, approximate nearest-neighbor indexing and search technique offered by Spotify, is used to index the entire CodeSearchNet Corpus. The index encompasses all functions in the corpus, including those without accompanying documentation. Annoy helps to achieve efficient and accurate retrieval [26].

4. Extended work: Code2vec and Embeddings Fusion

In this study, we propose a novel approach for enhancing the performance of code representation models by leveraging large language models (LLMs), such as Code Llama [31], Qwen2.5-coder [32], and code2vec for Python language. The integration of LLMs with traditional code representation techniques presents a promising way to enhance code understanding. Our approach combines the semantic richness of LLM embeddings with the structural insights of syntactic code2vec vectors. This fusion aims to create a more comprehensive and nuanced representation of code snippets. Figure 5 depicts the architecture of the embeddings fusion approach adapted with the code2vec model.

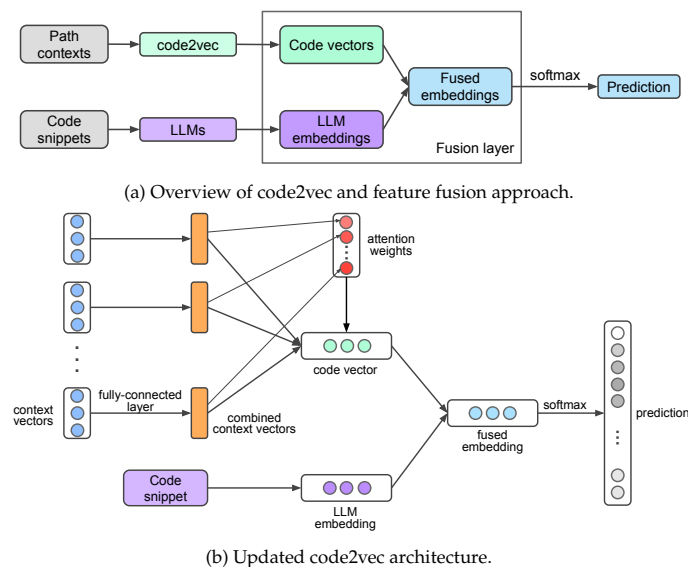


Figure 5: Proposed architecture of code2vec with fused embeddings using LLMs.

¹<https://github.com/spotify/annoy>

We utilize Code Llama - 7B [31], a variant of the Llama model fine-tuned on code, to generate semantic embeddings. Code Llama represents the cutting edge of publicly available large language models (LLMs) for code-related tasks. It offers the potential to enhance developer workflows by improving efficiency and speed, while also reducing the barriers for individuals new to programming. Additionally, Code Llama could serve as a valuable tool for productivity and education, assisting programmers in producing more reliable and well-documented software.

We used the CodeSearchNet corpus for training and evaluation. This dataset contains about 1.1 million Python functions, spanning diverse domains and programming paradigms. Each code snippet is processed through Code Llama to produce a high-dimensional vector capturing contextual and semantic information. Concurrently, we process the same code snippets through our adapted code2vec model for Python, which generates vectors representing the structural and syntactic features of the code.

To achieve this fusion, we explore various techniques to integrate LLM embeddings and code2vec code vectors into a single, unified representation. These techniques include:

1. Simple concatenation: A straightforward approach where Code Llama embeddings and code2vec vectors are simply combined end-to-end. This method preserves all information but increases dimensionality.

$$\text{fused_vector} = \text{llm_embedding} \parallel \text{code2vec_vector} \quad (3)$$

2. Weighted sum: We apply learnable weights to each embedding type before summation, allowing the model to adjust the importance of semantic versus syntactic information.

$$\text{fused_vector} = w_1 * \text{llm_embedding} + w_2 * \text{code2vec_vector} \quad (4)$$

3. Attention-based mechanisms: Inspired by transformer architectures, we implement a multi-head attention mechanism. This allows dynamic focus on different aspects of each embedding based on the specific code context. On the other hand, it allows the model to dynamically weight semantic and syntactic information based on task relevance, resolving the issue of fixed representation alignment by enabling the model to prioritize critical information.

$$\text{attention_output} = \text{MultiHeadAttention}(\text{llm_embedding}, \text{code2vec_vector})$$

$$\text{fused_vector} = \text{FeedForward}(\text{attention_output}) \quad (5)$$

Each method offers different advantages in terms of balancing the contribution of each type of embedding to the final representation. For instance, concatenation allows for a straightforward integration of both embeddings, while weighted sum can modulate the influence of each embedding type, and attention-based fusion offers the potential to dynamically adjust the focus on relevant features during training and inference. The fused representation is then employed in both the training and testing phases of code2vec, with the goal of improving performance on tasks such as code labeling and code search.

Figure 6 illustrates examples of two fusion techniques: concatenation and attention-based fusion. These visualizations provide insight into how each method integrates the embeddings and the potential impact on the resultant code representation.

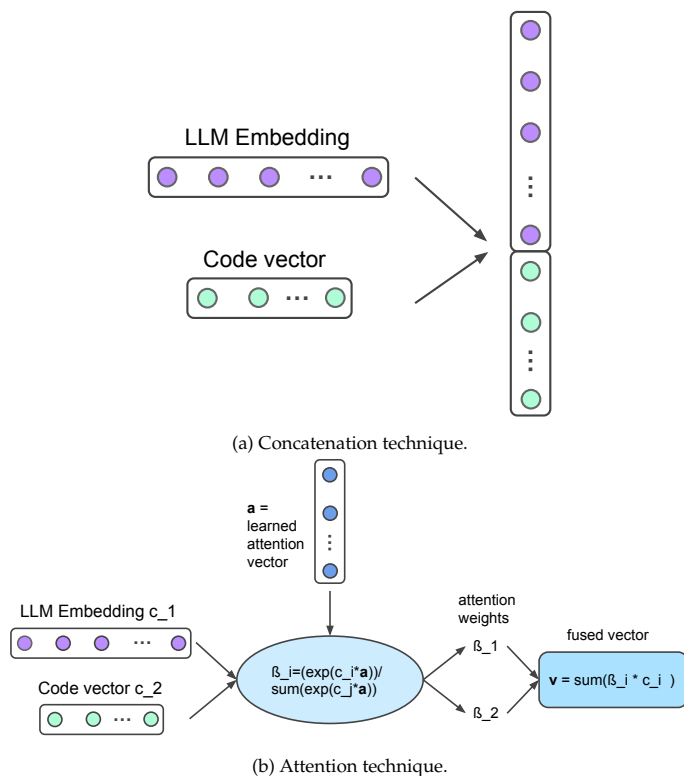


Figure 6: Examples of 2 feature fusion techniques used in our feature fusion approach.

5. Discussion

The integration of LLM embeddings with code2vec vectors shows a significant theoretical advancement in code representation techniques, which offers promising directions for combining semantic and syntactic information in code understanding tasks. Here, we discuss the implications of these findings and their potential impact on the field of automated code understanding.

Our findings highlight the complementary nature of LLM embeddings and code2vec vectors, demonstrating how these two approaches can be combined to enhance code representation. LLM embeddings capture rich semantic information and contextual nuances of code, while code2vec vectors excel at representing syntactic structure. The synergy between these two types of information allows our model to develop a more comprehensive understanding of code snippets. The relevance of semantic versus syntactic information may vary depending on the specific code snippet or task at hand. The multi-head attention mechanism in our fusion approach allows the model to dynamically focus on different aspects of the code representation. This adaptability is particularly valuable for handling various programming paradigms, coding styles, or specific tasks.

The integration of LLM embeddings, which are pre-trained on vast amounts of code data, with task-specific code2vec vectors opens up new possibilities for transfer learning in code analysis. This approach allows models

to benefit from the broad knowledge captured by LLMs while still maintaining the ability to fine-tune on specific tasks or codebases. This transfer learning capability could be particularly beneficial for organizations with limited labeled data or for tackling niche programming languages or domain-specific coding patterns. It suggests a path towards more generalizable code understanding models that can quickly adapt to new contexts.

The fusion of LLM embeddings and code2vec vectors also increases the computational complexity of the model. The attention-based fusion, in particular, adds a non-trivial amount of computation to the process. As we scale this approach to larger codebases or real-time applications, careful consideration must be given to balancing performance gains with computational efficiency. Future work could explore techniques for model compression or distillation to make this approach more feasible for resource-constrained environments or large-scale deployments.

The enhanced capabilities in code labeling and search have the potential to significantly impact software development practices. More accurate code labeling could lead to improved auto-documentation tools, helping maintain up-to-date and accurate code documentation. Enhanced code search capabilities could boost developer productivity by making it easier to find and reuse existing code snippets, potentially reducing duplication and improving code quality. Furthermore, these advancements could contribute to the development of more sophisticated code recommendation systems, assisting developers in writing more efficient, readable, and maintainable code.

6. Conclusion

This study has presented a significant extension to the previous work on code representation using code2vec and ASTminer for Python. By integrating large language model (LLM) embeddings with code2vec vectors, we have demonstrated a novel approach to capturing both semantic and syntactic information in code representations. While the current work focuses on architectural design and theoretical analysis, it lays important groundwork for future research in automated code understanding.

Our key theoretical contributions include: the development of fusion techniques, particularly an attention-based mechanism, to combine LLM embeddings with code2vec vectors; insights into the synergistic relationship between semantic information from LLMs and structural information from code2vec, and how this synergy can be leveraged in a novel architecture for better code understanding; theoretical foundation for improved code understanding models in future empirical studies.

These advancements address some of the limitations identified in the previous code2vec work, particularly in aligning semantic information (such as docstrings) with structural code representations. The fusion approach we developed offers a more robust and flexible way to represent code, adapting to the specific needs of different tasks and code structures.

Despite the promising aspects, several limitations and areas for further research remain:

- Language specificity: Our current work focuses on

Python. Further research is needed to assess the generalizability of this approach to other programming languages.

- Lack of empirical validation: The current study focuses on theoretical foundations and architectural design, without experimental results to validate performance claims. Comparative analysis with existing code representation techniques remains to be conducted.
- Model interpretability: While the attention mechanism provides some insight into the model's decision-making process, further work is needed to improve the interpretability of the fused representations.
- Fine-grained code understanding: Future research could explore how this approach performs on more fine-grained tasks, such as variable naming or type inference.
- Temporal aspects of code: Investigating how to incorporate information about code evolution and version history into our fused representations could provide additional valuable insights.

In conclusion, our work on fusing LLM embeddings with code2vec vectors represents a significant step forward in the field of automated code understanding. By leveraging both semantic and syntactic information, this approach opens up new possibilities for more accurate and versatile code understanding models, with potential far-reaching implications for software development practices and tools. As we continue to refine these techniques and explore their applications, we potentially anticipate a profound impact on the landscape of software development. Through rigorous experimental validation and continued refinement, these theoretical contributions can evolve into practical tools for improving code comprehension and development efficiency.

Conflict of Interest The authors declare no conflict of interest.

References

- [1] L. H. Ngo, V. Sekar, E. Leclercq, J. Rivalan, "Exploring code2vec and astminer for python code embeddings", "2023 IEEE 3rd International Conference on Software Engineering and Artificial Intelligence (SEAI)", pp. 53–57, IEEE, 2023, doi:[10.1109/SEAI59139.2023.10217505](https://doi.org/10.1109/SEAI59139.2023.10217505).
- [2] V. Kovalenko, E. Bogomolov, T. Bryksin, A. Bacchelli, "Pathminer: a library for mining of path-based representations of code", "Proceedings of the 16th International Conference on Mining Software Repositories", pp. 13–17, IEEE Press, 2019, doi:[10.1109/MSR.2019.00015](https://doi.org/10.1109/MSR.2019.00015).
- [3] U. Alon, M. Zilberstein, O. Levy, E. Yahav, "code2vec: Learning distributed representations of code", *Proceedings of the ACM on Programming Languages*, vol. 3, no. POPL, pp. 1–29, 2019, doi:[10.1145/3290353](https://doi.org/10.1145/3290353).
- [4] S. Hu, Y. Zuo, L. Wang, P. Liu, "A review about building hidden layer methods of deep learning", *Journal of Advances in Information Technology*, vol. 7, no. 1, 2016, doi:[10.12720/jait.7.1.58-63](https://doi.org/10.12720/jait.7.1.58-63).
- [5] Y. Sakai, Y. Eto, Y. Teranishi, "Structured pruning for deep neural networks with adaptive pruning rate derivation based on connection sensitivity and loss function", *Journal of Advances in Information Technology*, 2022, doi:[10.12720/jait.13.1.1-7](https://doi.org/10.12720/jait.13.1.1-7).
- [6] M. Allamanis, H. Peng, C. Sutton, "A convolutional attention network for extreme summarization of source code", "International conference on machine learning", pp. 2091–2100, PMLR, 2016, doi:[10.48550/arXiv.1602.03001](https://doi.org/10.48550/arXiv.1602.03001).
- [7] M. White, C. Vendome, M. Linares-Vásquez, D. Poshyanyk, "Toward deep learning software repositories", "2015 IEEE/ACM 12th Working Conference on Mining Software Repositories", pp. 334–345, IEEE, 2015, doi:[10.1109/MSR.2015.33](https://doi.org/10.1109/MSR.2015.33).
- [8] M. Allamanis, E. T. Barr, C. Bird, C. Sutton, "Learning natural coding conventions", "Proceedings of the 22nd ACM SIGSOFT International Symposium on Foundations of Software Engineering", pp. 281–293, 2014, doi:[10.1145/2635868.2635883](https://doi.org/10.1145/2635868.2635883).
- [9] M. Allamanis, C. Sutton, "Mining source code repositories at massive scale using language modeling", "2013 10th working conference on mining software repositories (MSR)", pp. 207–216, IEEE, 2013, doi:[10.1109/MSR.2013.6624029](https://doi.org/10.1109/MSR.2013.6624029).
- [10] D. Movshovitz-Attias, W. Cohen, "Natural language models for predicting programming comments", "Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)", pp. 35–40, 2013, doi:[10.18653/v1/P13-2007](https://doi.org/10.18653/v1/P13-2007).
- [11] U. Alon, M. Zilberstein, O. Levy, E. Yahav, "A general path-based representation for predicting program properties", *ACM SIGPLAN Notices*, vol. 53, no. 4, pp. 404–419, 2018, doi:[10.1145/3192366.3192412](https://doi.org/10.1145/3192366.3192412).
- [12] P. Bielik, V. Raychev, M. Vechev, "Phog: probabilistic model for code", "International conference on machine learning", pp. 2933–2942, PMLR, 2016, doi:[10.48550/arXiv.1602.05259](https://doi.org/10.48550/arXiv.1602.05259).
- [13] V. Raychev, M. Vechev, A. Krause, "Predicting program properties from 'big code'", *ACM SIGPLAN Notices*, vol. 50, no. 1, pp. 111–124, 2015, doi:[10.1145/2676726.2677009](https://doi.org/10.1145/2676726.2677009).
- [14] V. Raychev, P. Bielik, M. Vechev, "Probabilistic model for code with decision trees", *ACM SIGPLAN Notices*, vol. 51, no. 10, pp. 731–747, 2016, doi:[10.1145/2983990.2984041](https://doi.org/10.1145/2983990.2984041).
- [15] M. Allamanis, E. T. Barr, C. Bird, C. Sutton, "Suggesting accurate method and class names", "Proceedings of the 2015 10th joint meeting on foundations of software engineering", pp. 38–49, 2015, doi:[10.1145/2786805.2786849](https://doi.org/10.1145/2786805.2786849).
- [16] V. Raychev, M. Vechev, E. Yahav, "Code completion with statistical language models", "Proceedings of the 35th ACM SIGPLAN conference on programming language design and implementation", pp. 419–428, 2014, doi:[10.1145/2594291.2594321](https://doi.org/10.1145/2594291.2594321).
- [17] A. Mishne, S. Shoham, E. Yahav, "Typestate-based semantic code search over partial programs", "Proceedings of the ACM international conference on Object oriented programming systems languages and applications", pp. 997–1016, 2012, doi:[10.1145/2384616.2384698](https://doi.org/10.1145/2384616.2384698).
- [18] M. Amodio, S. Chaudhuri, T. W. Reps, "Neural attribute machines for program generation", *arXiv preprint arXiv:1705.09231*, 2017, doi:[10.48550/arXiv.1705.09231](https://doi.org/10.48550/arXiv.1705.09231).
- [19] Y. Lu, S. Chaudhuri, C. Jermaine, D. Melski, "Data-driven program completion", *arXiv preprint arXiv:1705.09042*, 2017, doi:[10.48550/arXiv.1705.09042](https://doi.org/10.48550/arXiv.1705.09042).
- [20] C. Maddison, D. Tarlow, "Structured generative models of natural source code", "International Conference on Machine Learning", pp. 649–657, PMLR, 2014, doi:[10.48550/arXiv.1401.0514](https://doi.org/10.48550/arXiv.1401.0514).
- [21] M. Allamanis, E. T. Barr, P. Devanbu, C. Sutton, "A survey of machine learning for big code and naturalness", *ACM Computing Surveys (CSUR)*, vol. 51, no. 4, pp. 1–37, 2018, doi:[10.1145/3212695](https://doi.org/10.1145/3212695).
- [22] M. Vechev, E. Yahav, et al., "Programming with 'big code'", *Foundations and Trends® in Programming Languages*, vol. 3, no. 4, pp. 231–284, 2016, doi:[10.1561/25000000028](https://doi.org/10.1561/25000000028).
- [23] V. Kovalenko, E. Bogomolov, T. Bryksin, A. Bacchelli, "Building implicit vector representations of individual coding style", "Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops", pp. 117–124, 2020, doi:[10.1145/3387940.3391461](https://doi.org/10.1145/3387940.3391461).
- [24] X. Gu, H. Zhang, S. Kim, "Deep code search", "Proceedings of the 40th International Conference on Software Engineering", pp. 933–944, 2018, doi:[10.1145/3180155.3180167](https://doi.org/10.1145/3180155.3180167).

- [25] B. Mitra, N. Craswell, *et al.*, "An introduction to neural information retrieval", *Foundations and Trends® in Information Retrieval*, vol. 13, no. 1, pp. 1–126, 2018, doi:[10.1561/15000000061](https://doi.org/10.1561/15000000061).
- [26] H. Husain, H.-H. Wu, T. Gazit, M. Allamanis, M. Brockschmidt, "Code-SearchNet challenge: Evaluating the state of semantic code search", *arXiv preprint arXiv:1909.09436*, 2019, doi:[10.48550/arXiv.1909.09436](https://doi.org/10.48550/arXiv.1909.09436).
- [27] I. Sheikh, I. Illina, D. Fohr, G. Linares, "Learning word importance with the neural bag-of-words model", "Proceedings of the 1st Workshop on Representation Learning for NLP", pp. 222–229, 2016, doi:[10.18653/v1/W16-1626](https://doi.org/10.18653/v1/W16-1626).
- [28] K. Cho, B. Van Merriënboer, D. Bahdanau, Y. Bengio, "On the properties of neural machine translation: Encoder-decoder approaches", "Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation", 2014, doi:[10.3115/v1/W14-4012](https://doi.org/10.3115/v1/W14-4012).
- [29] K. Yoon, "Convolutional neural networks for sentence classification [ol]", *arXiv Preprint*, 2014, doi:[10.48550/arXiv.1408.5882](https://doi.org/10.48550/arXiv.1408.5882).
- [30] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, "Attention is all you need", *Advances in neural information processing systems*, vol. 30, 2017, doi:[10.48550/arXiv.1706.03762](https://doi.org/10.48550/arXiv.1706.03762).
- [31] B. Roziere, J. Gehring, F. Gloeckle, S. Sootla, I. Gat, X. E. Tan, Y. Adi, J. Liu, R. Sauvestre, T. Remez, *et al.*, "Code llama: Open foundation models for code", *arXiv preprint arXiv:2308.12950*, 2023, doi:[10.48550/arXiv.2308.12950](https://doi.org/10.48550/arXiv.2308.12950).
- [32] B. Hui, J. Yang, Z. Cui, J. Yang, D. Liu, L. Zhang, T. Liu, J. Zhang, B. Yu, K. Dang, *et al.*, "Qwen2. 5-coder technical report", *arXiv preprint arXiv:2409.12186*, 2024, doi:[10.48550/arXiv.2409.12186](https://doi.org/10.48550/arXiv.2409.12186).

Copyright: This article is an open access article distributed under the terms and conditions of the Creative Commons At-

tribution (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).



LONG H. NGO has done his B.E. degree in electrical and electronics engineering from Ho Chi Minh City University of Technology in 2016. He has received his master's degree in imaging and networks from Université Sorbonne Paris Nord, France, in 2017. He has completed his PhD degree in machine learning and image processing from Université Sorbonne Paris Nord, France, in 2021.

His research interests include machine learning, image processing, computer vision, and NLP.



JONATHAN RIVALAN is R&D manager at Smile, French open source services integrator based in various locations in Europe. His main line of work targets workflow automation, digital infrastructures optimization and advanced HCI developments. He is currently participating and leading national and European R&D programs, which results are available as open source solutions, such as Rating Operator (metrics transformation towards KPIs) and Palindrome.js (multidimensional monitoring for distributed systems).

Developing E-Methane Value Chain and Proper Greenhouse Gas (GHG) Accounting Rules Incentivizing Recycled Carbon Fuels

Ryota Kuzuki*, Mitsuhiro Kohara, Noboru Kizuki, Satoshi Yoshida, Yuta Nezasa, Yuki Tsuji

Division of Int'l Certification & Standards Harmonization, Planning dept., The Japan Gas Association

*Corresponding author: Ryota Kuzuki, Division of Int'l Certification & Standards Harmonization, The Japan Gas Association

1-15-12, Toranomon, Minato-ku, Tokyo, 105-0001, Japan, Contact No: +81-3-35020116, Email: kuzuki.ryota@gas.or.jp

ABSTRACT: Many countries and industries in the world are aiming to transition to a carbon neutral society. One of the means to achieve this is the use of recycled carbon fuels (RCFs) such as e-methane. Feasibility of an international value chain is being studied in a variety of regions. In particular, with regard to e-methane, demonstration projects are progressing in several regions around the world as well as feasibility studies of cross-border value chain of e-methane. Ideally, e-methane production sites should be located to secure abundant and inexpensive hydrogen and CO₂ as feedstocks, besides close to natural gas transmission pipelines, liquefaction facilities, and easily accessible to LNG exporting terminals. In order to stimulate investment in commercialization, development of internationally applicable GHG accounting rules is need at first. It is essential that the CO₂ emitted in combustion is recognized under international accounting rules as having environmental value attribute that does not increase CO₂ in the atmosphere. As one of the examples, the Japan Gas Association and related organizations have developed a certification scheme called "Clean Gas Certificate" and has applied it to several domestic e-methane production projects in Japan. For the transition period (2030-2050), the use of recycled CO₂ captured from factories and thermal power plants emissions are expected to be major feedstock. We propose GHG calculation and accounting method in consideration of avoiding double counting of recycled CO₂ in the entire supply chain. A calculation formula of the carbon footprint (CFP) of e-methane was proposed to one of the ISO's working groups and it was successfully included in the ISO 6338-1:2024. Next, it is necessary to harmonize the proposed methods with internationally common GHG accounting rules.

KEYWORDS: e-methane, methanation, carbon capture utilization, GHG accounting rules, GHG intensity, recycled carbon fuels, environmental attribute certificates, Clean Gas Certificate

1. Introduction

Many countries and industries in the world are aiming to transition to a carbon neutral society. One of the means to achieve this is the use of recycled carbon fuels (RCFs) such as e-methane. In [1], our previous paper "Promoting Introduction of E-methane for Carbon Neutral City Gas Supply in Japan", focused on e-methane and overviewed about the gas industry's challenge of technological development and related initiatives such as net GHG emission calculation of e-methane. Following it, this paper describes the progress of the challenges of technological development, demonstrations, and feasibility studies for

commercialization. The core concept is CO₂ capture and utilization (CCU) which enables carbon reduction. As an example, by use of e-methane, Figure 1 shows an image of the principle of RCFs. By replacing fossil fuels with RCFs, CO₂ emissions can be reduced.

In order to stimulate investment in commercialization, development of internationally applicable GHG calculation and accounting rules is essential. The originality of this paper is that it makes a proposal based on the authors' practices, while taking into account existing international standards to address this issue. This paper also introduces a certification scheme which has

developed to enable transferring the environmental value attributes of RCFs such as e-methane.

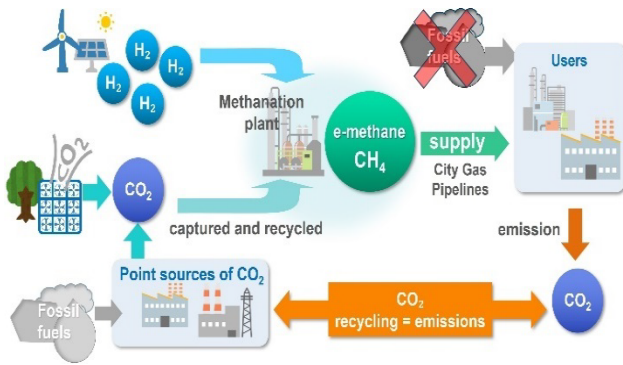


Figure 1: Image of methanation process for carbon neutral city gas supply

2. Expectations for e-methane and initiatives to date

In [2], the IPCC's 6th Assessment Report, released in April 2022, states "For greater compatibility with existing gas systems and appliances, the hydrogen can be combined with captured carbon dioxide to form methane and other synthetic fuels". In [3], a model analysis of IEA's "Net Zero by 2050" also estimates that 35% of global gas demand supplied by the gas grid in 2050 will be supplied by low-carbon gases, of which about 30% will be e-methane.

In September 2024, the IEA held a webinar focusing on e-methane for the first time, and in [4], Molnar presented survey looking over the global situation. Leading projects taken up by the IEA are shown in Figure 2. Among them, Tree Energy Solutions (TES) and Nordic Ren-Gas as well as the Japan Gas Association were invited as panelists and presented leading projects.

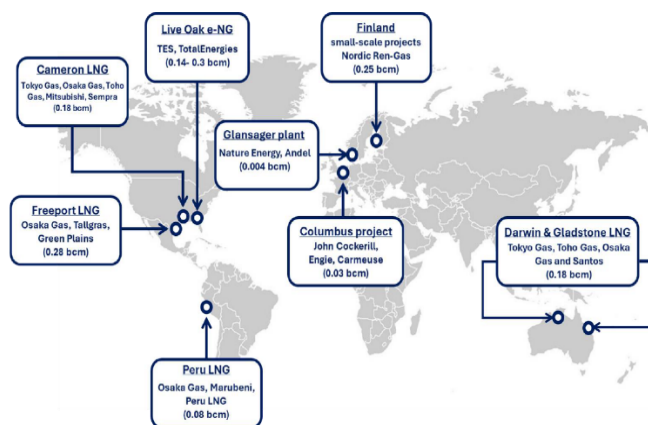


Figure 2: International partnerships drive e-methane projects (IEA's webinar on E-methane: a new gas for a net-zero future?)

In [5], Vercamen announced the formation of the e-NG Coalition, a framework for large companies to collaborate and advocate for their portfolio of e-methane projects in Japan, North America, Europe, and the Middle East.

In [6], Kujara, Nordic Ren Gas presented its plans of commercial production of e-methane at six sites in Finland to be operated from 2027, and announced that one of their

project was selected by the EU Hydrogen Auction bid as the most competitive cost of hydrogen.

In [7], the Japan Gas Association presented about its Action Plan in June 2021, which sets an interim goal for the entire city gas industry to establish a value chain including overseas imports of e-methane covering 1% in city gas supply (approximately 400 million m³) in 2030. And this includes a goal by 2050 of achieving carbon neutral city gas supply by e-methane as main feedstock (Figure 3).

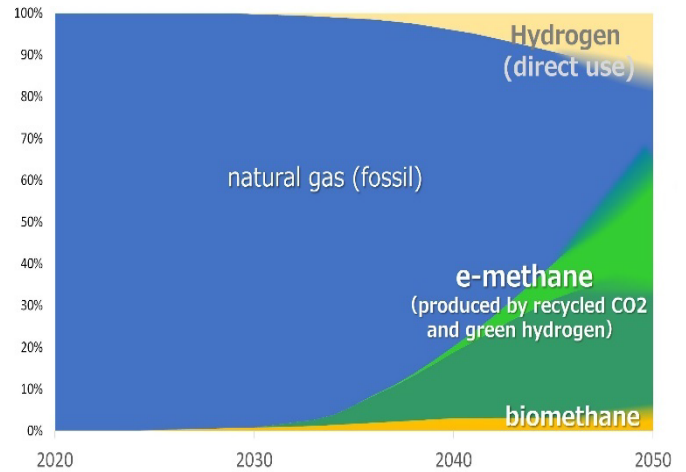


Figure 3: JGA's vision of transition of the share of feedstocks for Japan's city gas supply

In [8], European Biogas Association (EBA) published an Europe-wide study in September 2024. It reports a comprehensive survey on the e-methane projects in Europe. Figure 4 shows the increasing trend in number of e-methane production plants by country.

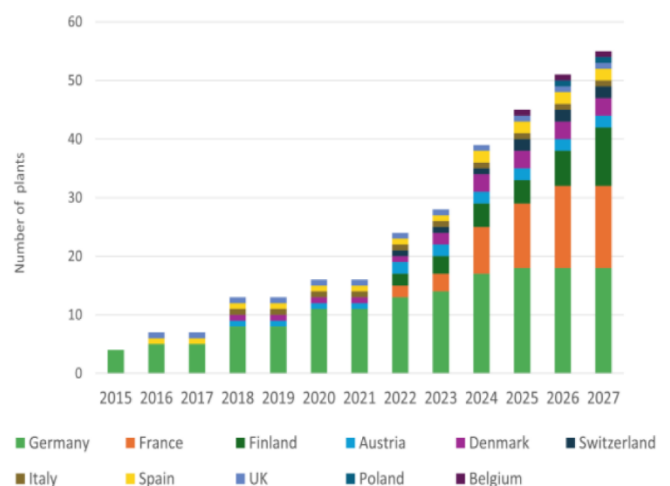


Figure 4: Growth in number of e-methane production plants in Europe (EBA; Mapping e-methane plants and technologies, 2024.9)

References above are appropriate to representing the latest situation. All of them suggest importance of proving the origin of the feedstock CO₂, but most of them are limited within biogenic CO₂, and there is few information of eligibility of fossil fuel-originated CO₂. Therefore, the attribution of emissions of fossil-originated CO₂ that is captured and used as feedstock are not discussed enough.

3. Harmonizing Internationally Applicable GHG Calculation and Accounting Rules for RCFs

3.1. Need for rules for cross-border transported RCFs

Ideally, the e-methane production site should have access to abundant and inexpensive hydrogen and CO₂ as feedstock and easy access to natural gas transmission pipelines, liquefaction, and LNG transportation. Several feasibility studies are currently underway in many countries and regions, including North America, Australia, Southeast Asia, Middle East, and South America. Figure 5 shows various feasibility studies (FSs) led by Japanese companies.

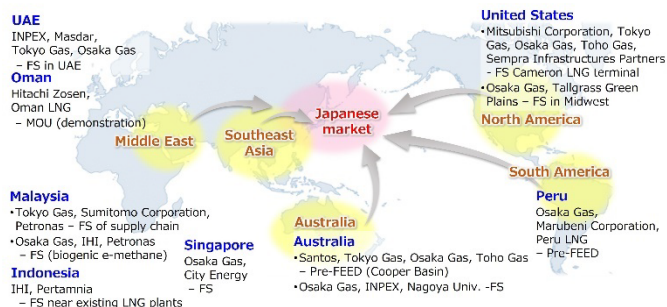


Figure 5: Various feasibility studies led by Japanese companies

As a result of site evaluations and feasibility studies conducted in various locations, they selected the area around the existing LNG export terminals as the most promising area because of their existing infrastructures for stable procurement of raw materials enabling early establishment of a value chain for e-methane. Some are expected for investment decisions. However, there is still no common international rule that can certify GHG calculation covering entire cross-border value chains and the transfer of environmental values of RCFs such as e-methane. Without them, there is concern that investment decisions would be delayed.

3.2. Existing international general rules for GHG calculation and accounting

During the transition period (from around 2030 to 2050), CO₂ captured from point sources is expected to become the major feedstock for producing e-methane. If the CO₂ emitted from point sources such as factories and power plants has already been accounted for and reported by the emitters, since the CO₂ would have been released into the atmosphere, when the e-methane is combusted, regardless of its origin (e.g. fossil or non-fossil), the emitted CO₂ should not be counted as additional emission in order to avoid double counting. This means that recycled carbon fuels (RCFs) such as e-methane have environmental value contributing CO₂ emission reduction, as stated in the 6th IPCC assessment report [2].

In order to promote social implementation of e-methane, common calculation method is needed to estimate the amount of CO₂ reductions by substituting

fossil fuels. Internationally accepted rules for calculating GHG emissions include the IPCC Inventory Guidelines at the national level, the GHG Protocol Standards and Guidelines and ISO 14064, ISO 14067 (carbon footprint) at the corporate and product level. Method to calculate environmental value of RCFs such as e-methane should be harmonized with these existing rules.

One of the representing analysis of GHG emission of RCFs covering entire supply chain is by S. Morimoto et. al.. In [9], Morimoto calculated lifecycle GHG emission of e-methane imported from abroad to be 0.95-1.65 kg-CO₂/kg-CH₄ (nearly equals to 17.0-29.6 g-CO₂/MJ) under a condition that hydrogen is produced by renewable power.

Another example is that specific default values are being set for over 100 fuel types, including RCFs, which are being discussed at the International Maritime Organization (IMO). In [10], when the CO₂ is derived from the point sources such as flue gas after combustion of fossil fuels, calculation formulas are provided to subtract the amount of CO₂ emitted during combustion from the amount already reported by the original emitter. The calculation method accumulates the CO₂ emitted throughout the supply chain.

3.3. Development of calculation formula of GHG intensity in ISO

Based on the standards and guidelines mentioned above, we proposed a calculation formula to one of the ISO's Working Group. As shown in Figure 6, referring ISO 14067, the proposed formula is to reflect in carbon footprint of e-methane within the entire boundary of well to consumption (WtC) in consideration of the environmental value that it does not increase CO₂ in the atmosphere if the CO₂ feedstock meets certain criteria.

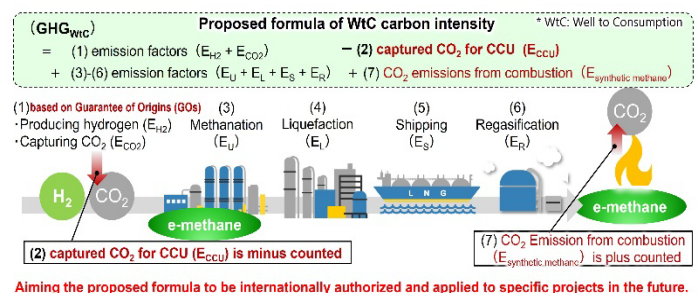


Figure 6: Proposed calculation formula of GHG intensity of e-methane

If the CO₂ feedstock is obtained by direct air capture (DAC) or biogenic origin, or if the CO₂ is already reported as GHG emissions before the e-methane production process, the captured CO₂ can be minus counted. This can offset the CO₂ emitted from combustion of the e-methane. In [11], this formula has been included in the ISO 6338-1:2024 (Calculations of greenhouse gas (GHG) emissions throughout the liquefied natural gas (LNG) chain Part 1: General) in January 2024.

This method was referenced when determining the upper limit of GHG emissions intensity of 49.3 g-CO₂/MJ, which is recognized as “low-carbon hydrogen, etc.” as stipulated in [12] “Basic Act on a Hydrogen Society”, it has been enacted by the Japanese government in May 2024.

3.4. Need for common rules to clarify the attribution of environmental value of RCFs

The next issue to be resolved is to develop common rule regarding who own the environmental value of e-methane, which does not increase atmospheric CO₂ even when combusted. In the case of biofuels, it can be commonly accepted that those who paid cost for biofuel own the environmental value. Recently, there is also a system in which the ownership of the environmental value is transferred by issuing certificates that separate the environmental value from the biofuels.

In the case of RCFs, unlike biofuels, the process of procuring CO₂ feedstocks and artificially synthesizing them is included in the value chain from production to consumption. The difference from biofuels and lack of common rules for RCFs causes uncertainty of the reward to the investment to developing a value chain for RCFs such as e-methane.

In Japan, discussions on how to handle CCU including e-methane are progressing. One of the most progressed discussions held in Japan was the Public-Private Council for Promoting Methanation led by the Ministry of Economy, Trade and Industry in 2021, and the Task Force for CO₂ Counting Rule under its auspices, have been held to discuss on this issue intensively.

In [13], the interim report issued by Ministry of Economy, Trade and Industry in 2022, as shown in Table 1, four cases, in which merits and demerits are included for the original emitters (captures’ side) or the final users, or both, depending on account for emissions or not.

Table 1: Examples of CO₂ accounting among stakeholders by cases [13]

		Original emitter (capturers’ side)	Users’-side
CO ₂ behavior related to e-methane production and utilization			
		Factories, Utilities	E-methane users
CO ₂ accounting for domestic production and consumption	Case 1	Account for emissions	No Emission
	Case 2	No emission	Account for emissions
	Case 3	Allocating emissions	Allocating emissions
	Case 4	No emission	No emissions

This report concluded that unless the environmental value is attributed to the users’ side, e-methane will not be selected. Therefore, it was proposed that the users’ side should account for as no emission. On the other hand, it suggested that complementary mechanism was necessary because there was a risk that production and supply of e-methane would not take place unless sufficient incentives

were provided to those who capture CO₂ rather than emitting it to the atmosphere.

There is a progress to create a system in which the original emitters (capturers’ side) can also claim environmental value for themselves, depending on the final use, by requiring proof of the captured CO₂ for CCU and proof of the final use.

Issues relating internationally applicable accounting rules has been raised in the intergovernmental frameworks. One of the examples is the Group of Seven (G7) Energy and Environment Ministerial Meeting held in Sapporo in April 2023. In [14], the communique stated that member countries would work together to resolve the issue (GHG accounting for RCFs).

As of now, the IPCC Inventory Guidelines nor the GHG Protocol, both are the most representative international rules, have not include any provisions regarding the treatment of CCU. As a result, even if RCFs is procured and used, users would not be rewarded since the CO₂ emitted during combustion still needs to be accounted for. As announced in [15] and [16], it is expected that discussions on CCU-related issues including RCFs will start from 2024 or 2025 and will be completed in a few years.

Meanwhile, as announced in [17], the EU’s Renewable Energy Directive (RED) Delegated Act, which is a common rule for all member countries, came into effect in June 2023. It stipulates the calculation method for the environmental value of RCFs and RFNBOs (liquid or gaseous renewable fuels of non-biological origin). It would be possible that this will not only have an impact within EU, but also spread to the rest of the worlds.

4. Development of Environmental Attribute Certificates for e-methane

From around 2023, discussions have emerged on internationally applicable counting rules that clarify the attribution of environmental value for RCFs including e-methane and the direction is to contribute to motivate developing value chain for RCFs including e-methane.

In certain regions, certification schemes for RCFs are almost being finalized.

4.1. Voluntary Scheme recognized by European Commission

In [18], the European Commission has announced that several existing international voluntary schemes have completed the positive technical assessment (September 2024) to be recognized as institutions responsible for third-party review in line with the Delegated Act [17]. For example, in [19], ISCC (International Sustainable Carbon Certificate) describes requirements for fossil-derived CO₂ to be eligible as feedstock for RCFs and RFNBOs as follows; “Post-industrial (fossil) CO₂ captured from

industrial processes, which use fossil sources to deliberately produce electricity, heat, or materials (e.g., cement, iron and steel, petrochemical industry) must be verified that the CO₂ was not deliberately produced for use in the above-mentioned production processes."

4.2. Clean Gas Certificate (Japan)

(1) Overview of Clean Gas Certificate

We developed and launched "Clean Gas Certificate" scheme with gas-related organizations in Japan that certifies the attribution of the environmental value of e-methane and biogas, which is considered as not increasing CO₂ in the atmosphere even when combusted. As shown in Figure 7, this is environmental attribute certificate (EAC) making it possible to transfer the environmental value separated from e-methane or biogas and to bundle it with fossil natural gas.

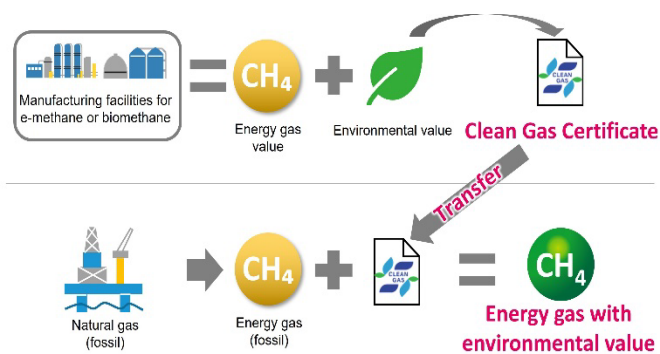


Figure 7: Image of function of "Clean Gas Certificate"

After a trial period in 2023, guidelines were developed, and the system was moved to actual operation as a private certification system from April 2024. In [20], details are posted on the website. With this scheme, e-methane and methane component in biogas are certified based on criteria. An sample image of the certificate to be issued is shown in Figure 8.

The attribute information included in the certificate are listed below. They have been decided based on surveys of some references of foreign certifying and registering systems for renewable fuels such as biogas.

- Serial number
- Issue date
- Production process
- Production period
- Name of the registered production facility
- Hydrogen suppliers' code
- CO₂ feedstock suppliers' code
- Name of the certificate issuer
- Type of clean gas (e-methane / biogas / mixed gas)



Figure 8: Sample image of "Clean Gas Certificate"

- Equivalent amount of clean gas (units: m³ and MJ)
- Use of produced gas (injection into city gas pipelines / private consumption)

(2) Implementation of Clean Gas Certificate

In Japan, various efforts are underway to commercialize e-methane, including technology development, demonstration on both the supply and demand sides, with support from the Japanese government. Figure 9 shows demonstration projects of e-methane production by Japan's major city gas companies.

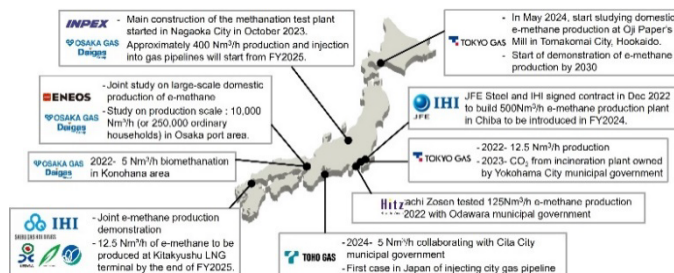


Figure 9: Demonstration projects of e-methane production in Japan

Some of the plants have already certified for Clean gas Certificates as below. Other projects planned in Japan are also expected to apply for the certification.

As announced in [21], in March 2022, Tokyo Gas Co., Ltd. began a methanation demonstration experiment at its research facility in Yokohama. Production capacity is 12.5Nm³/h. It procures electricity from renewable power sources and produce green hydrogen through water electrolysis. Having gained knowledge about the capabilities and issues of water electrolysis and the Sabatier methanation device, it has been working to improve the efficiency of the overall process. And in July 2023, it announced a new initiative to produce e-methane from CO₂ captured from the neighboring waste incineration plant owned by the Yokohama City municipal government, combined with the hydrogen obtained from its on-site water electrolysis device.

As announced in [22], in July 2022, Toho Gas has announced that it would work with the municipal government of Chita City in Aichi Prefecture to conduct a demonstration of methanation using CO₂ captured from biogas generated by sewage sludge treatment of the Chita City Southern Sewage Treatment Center and hydrogen produced by electricity generated from LNG cryogenic power generation. Production capacity is 5Nm³/h. The e-methane has been injected to the city gas supply network, which is for the first time in Japan.

As announced in [23], in December 2023, Saibu Gas announced that it started a demonstration project for methanation in Kitakyushu City in Fukuoka Prefecture using local raw materials in collaboration with nine other entities. Production is planned in FY2025 and with the capacity of 12.5Nm³/h. City gas customers, academia are

working together. Other city gas companies such as Hokkaido Gas, Hiroshima Gas, and Nihon Gas as well as the JGA also joined in this project aiming to transplant this model to other city gas supply areas.

As announced in [24], INPEX and Osaka Gas announced the start of a joint technology development project for the practical application of a methanation system. E-methane will be produced from FY2025 using CO₂ extracted from INPEX's gas field in Nagaoka City in Niigata Prefecture and will be injected into the company's existing gas pipeline. The e-methane production capacity is planned to be about 400 Nm³/h.

As one of the representative examples of its users, as announced in [25], Osaka Gas has been assigned to supply Clean Gas Certificate bundled city gas to the facilities located in the EXPO 2025 Osaka, Kansai during the period of the event.

4.3. Aiming for broader internationally applicable schemes

Broader than above cases, it is necessary to operate the scheme based on internationally applicable CCU-related accounting rules to be developed in the near future. It is necessary to build in on the existing LNG supply chain that include tracking information on the guarantee of origins (GOs) of the feedstocks, CO₂ and H₂, and the GHG intensity of e-methane shown in Figure 6. Figure 10 shows an image of e-methane certificates in circulation.

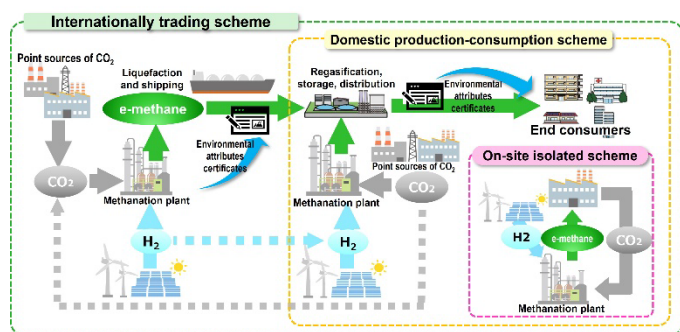


Figure 10: Image of environmental value transfer schemes in the existing LNG value chain including environmental attributes certificates

There are some notable developments regarding the IPCC Inventory Guidelines and GHG Protocol, which are the most commonly referenced GHG calculation and accounting rules in the world.

The IPCC has announced that it will hold expert meetings focusing on CO₂ removal and CCUS in the IPCC's Inventory Task Force through 2027, and will revise the inventory guidelines. As announced in [26], At the expert meeting in July 2024, "Address potential double-counting from the use of synthetic fuels (e.g. efuels) double counting" was raised as one of the issues to be considered.

The GHG Protocol is due for a major revision by the second half of 2026. In order to propose the addition of new rules on CCU to the standards and guidance, in [27], we submitted a proposal (open letter) to the secretariat in

January 2024, signed by 18 voluntary organizations. As published in [28], we also appealed for the need for rules related to CCU at a session attended by the representative of the GHG Protocol Secretariat at the annual general meeting hosted by a European organization that promotes CCU, and he responded positively to our appeal.

Since diverse stakeholders intervene, a platform is needed to manage the history of GHG emissions tracking, verification, and transfer of the value of reduction credits at stages along the value chain. For example, as announced in [29], Osaka Gas Co., Ltd., Mitsubishi Heavy Industries, Ltd. and IBM Japan, Ltd. are jointly conducting a proof-of-concept (PoC) of a system to visualize CO₂ emissions throughout the supply chain and the environmental value of e-methane (Figure 11). It is expected that such fundamental technology will be developed at an early stage to serve as a platform for tracking, verifying, and managing the transfer of environmental values.

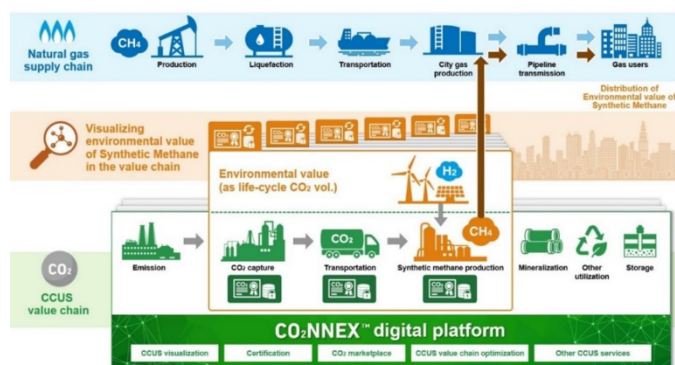


Figure 11: Example of digital platform tracking GHGs in e-methane supply chain

We will continue to monitor future discussions and work to harmonize our schemes with their revisions.

5. Conclusions

This paper focuses on e-methane, which is a leading RCFs produced from hydrogen derived from renewable energy and recycled CO₂, and which is attracting particular attention in Japan, Europe and North America because it can make the most of the existing LNG value chain. By leveraging the advantages of drop-in fuels, it is possible to reduce social costs and contribute to a smooth transition to a carbon-neutral society.

In this paper, we strongly advocate the need for international common rules for GHG calculation and accounting for RCFs through cross-border supply chains, which is a common issue for the social implementation of RCFs. In the early stages, the value chain of RCFs, including e-methane, is expected to grow within regional economic blocs such as Japan, EU and APEC, and in specific industries such as international shipping. As a result of certification of GHG intensity and in the carbon pricing system, market prices will be formed for RCFs. Based on the above circumstance, as a first step, we have

standardized the GHG intensity calculation formula for e-methane at the ISO and have also started to operate the Clean Gas Certificate scheme, which certifies the environmental value of e-methane produced at demonstration plants in Japan. Through these initiatives, we believe that the knowledge of trade-offs between GHG intensity and cost will contribute to form global RCFs market.

We especially expect academic world to scientifically clarify trade-off relationships that are more versatile. By sharing this knowledge, industrial sector is expected to work towards the formation of a global RCFs value chain. Everyone should collaborate to establish appropriate rules for a smooth transition to a carbon neutral society.

Conflict of Interest

The authors declare no conflict of interest.

References

- [1] R. Kuzuki, et.al., "Promoting Introduction of E-methane for Carbon Neutral City Gas Supply in Japan", *11th International Conference on Power Science and Engineering (ICPSE)*, 2022, doi: 10.1109/ICPSE56329.2022.9935370
- [2] IPCC, "The Sixth Assessment Report of the IPCC, Working Group III Report, Climate Change 2022 Mitigation of Climate Change, pp.6-57, 2022 https://www.ipcc.ch/report/ar6/wg3/downloads/report/IPCC_AR6_WGIII_FullReport.pdf
- [3] IEA, "Net Zero by 2050 A Roadmap for the Global Energy Sector", 2021, https://iea.blob.core.windows.net/assets/deebef5d-0c34-4539-9d0c-10b13d840027/NetZeroBy2050_ARoadmapfortheGlobalEnergySector_CORR.pdf
- [4] G. Molnar, "E-methane: a new gas for a net-zero future?", *Webinar on IEA Low-emissions Gases Work Programme*, 2024, https://iea.blob.core.windows.net/assets/9b86ac2a-2055-4eac-9f93-6ab379554d6d/IEA_E-methaneanewgasforanet-zerofuture.pdf
- [5] Y. Vercamen, "eNG's role for net-zero future", *Webinar on IEA Low-emissions Gases Work Programme*, 2024, <https://iea.blob.core.windows.net/assets/9b86ac2a-2055-4eac-9f93-6ab379554d6d/TESIEAe-methanewebinar05Sep2024.pdf>
- [6] S. Kujara, "E-methane market prospects in Northern Europe", *Webinar on IEA Low-emissions Gases Work Programme*, 2024, https://iea.blob.core.windows.net/assets/9b86ac2a-2055-4eac-9f93-6ab379554d6d/Ren_Gas_E-methane_presentation_IEA_2024_0902.pdf
- [7] The Japan Gas Association, Carbon Neutral Challenge 2050 (Press Release), 2020, https://www.gas.or.jp/pdf/en/newsrelease/newsrelease_20211101.pdf
- [8] European Biogas Association, "Mapping e-methane plants and technologies – The role of e-methane in the total energy mix", 2024, <https://www.europeanbiogas.eu/wp-content/uploads/2024/09/Mapping-e-methane-plants-and-technologies.pdf>
- [9] S. Morimoto et.al., "Scenario assessment of implementing methanation considering economic feasibility and regional characteristics", 2022, *Journal of CO₂ Utilization*, vol.58, 101935, doi.org/10.1016/j.jcou.2022.101935
- [10] R. Miyake, "Overview of IMO's guideline of marine fuel life cycle GHG intensity (in Japanese)", pp.26-28, *2024 Class NK Technical Seminar*, 2024, https://www.classnk.or.jp/hp/pdf/research/seminar/2024/seminar_005.pdf
- [11] ISO, "ISO 6338-1:2024 Calculations of greenhouse gas (GHG) emissions throughout the liquefied natural gas (LNG) chain Part 1: General", 2024, <https://www.iso.org/standard/87114.html>
- [12] CCR (carbon capture and reuse) Study Group; "A Study on the Carbon Intensity of Synthetic Methane (e-methane)", *12th Public-Private Council for Promoting Methanation* (only Japanese), 2024, https://www.meti.go.jp/shingikai/energy_environment/methanation_suishin/pdf/012_05_01.pdf
- [13] Interim report on CO₂ emission accounting in combustion of synthetic methane, Task Force for CO₂ Counting Rules of the Public-Private Council for Promoting Methanation (only Japanese), 2022, https://www.meti.go.jp/shingikai/energy_environment/methanation_suishin/co2_tf/pdf/20220322_1.pdf
- [14] G7 (Group of Seven) Climate, Energy and Environment Ministers' Communiqué Sapporo, 2023, 27, <https://www.meti.go.jp/press/2023/04/20230417004/20230417004-1.pdf>
- [15] IPCC, Decisions adopted by the Panel, *60th Session of the IPCC*, 2024, 9, https://www.ipcc.ch/site/assets/uploads/2024/02/IPCC-60_decisions_adopted_by_the_Panel.pdf
- [16] GHG Protocol, GHG Protocol Corporate Suite of Standards and Guidance Update Process, <https://ghgprotocol.org/ghg-protocol-corporate-suite-standards-and-guidance-update-process>
- [17] European Union, ANNEX Methodology for determining greenhouse gas emissions savings from renewable liquid and gaseous transport fuels of non-biological origin and from recycled carbon fuels, Official Journal of the European Union, Volume 66, (June 20 2023) 23-33
- [18] European Commission Website, "Voluntary schemes set standards for the production of sustainable fuels and gases" https://energy.ec.europa.eu/topics/renewable_energy/bioenergy/voluntary-schemes_en
- [19] ISCC, ISCC PLUS Ver 3.4.2, 2024, 13-15 https://www.iscc-system.org/wp-content/uploads/2024/03/ISCC-PLUS_v3.4.2.pdf
- [20] Clean Gas Certificate Committee Website (only in Japanese) <https://www.clean-gas-certificate.com/>
- [21] Tokyo Gas Co., Ltd., Towards a carbon-neutral, decarbonized society in 2050 - methanation demonstration tests begin (Press Release (in Japanese)), 2022, <https://www.tokyo-gas.co.jp/news/press/20220323-02.html>
- [22] Toho Gas Co.,Ltd., Methanation demonstration test using CO₂ from Biogas in collaboration with the municipal government of Chita City (Press Release (in Japanese)), 2022, https://www.tohogas.co.jp/corporate-n/press/1229823_1342.html
- [23] Saibu Gas CO., Ltd., etc.; Start of a methanation demonstration project using local materials (Press Release (in Japanese)), 2023, https://hd.saibugas.co.jp/news_release/detail/2023/nr040.html
- [24] INPEX Co.,Ltd and Osaka Gas Co.,Ltd, INPEX, Osaka Gas to Commence Technical Development Business on CO₂ Emission Reduction and Practical Application of Effective CO₂ Use Through One of the World's Largest Methanation Operations (Joint Press Release), 2021, https://www.osakagas.co.jp/company/press/pr2021/1300478_46443.html

- [25] Osaka Gas Co., Ltd., Contribution to the carbon neutrality of the 2025 Japan International Exposition through the first Clean Gas Certificate (Press Release (in Japanese)), 2024, https://www.osakagas.co.jp/company/press/pr2024/_icsFiles/afeldfile/2024/10/31/241101_1.pdf
- [26] IPCC, "IPCC Expert Meeting Report on Carbon Dioxide Removal Technologies and Carbon Capture, Utilization and Storage", 2024, https://www.ipcc-nggip.iges.or.jp/public/mtdocs/pdfiles/2407_CDR_CCUS_Report.pdf
- [27] The Japan Gas Association and 17 organizations and companies, "Proposal for Inclusion of Market-Based Accounting in Scope 1 of the GHG Protocol Update" (open letter to the GHG Protocol Secretariat), 2024, https://www.gas.or.jp/pdf/en/newsrelease/newsrelease_20240131.pdf
- [28] GHG Protocol Website "Event Recording: Reporting Emissions in Carbon Capture and Utilization (CCU) Applications", 2024, <https://www.youtube.com/watch?v=usB47sH-320>
- [29] Osaka Gas Co., Ltd., Mitsubishi Heavy Industries, Ltd., IBM Japan, Ltd., Three Partners to Jointly Undertake Proof of Concept (PoC) Applying "CO2NNEX™", a Digital Platform for CO₂ Accounting across its Supply Chain for Synthetic Methanes (Joint Press Release), 2022, <https://www.mhi.com/news/221021.html>

Copyright: This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).



Ryota Kuzuki has done his master's degree of precision engineering from the University of Tokyo in 1989. He has received Ph.D. degree in engineering from graduate school of Keio University in 2011.

He joined Tokyo Gas Co., Ltd. in 1989 and worked mainly for gas supply safety system design, corporate planning, and commercial and industrial sector's gas marketing. He currently serves as Division Head, Int'l Certification & Standards Harmonization, JGA



Mitsuhiro Kohara has done his master's degree in mechanical engineering from Graduate School of Engineering, Osaka University in 2001.

He joined Osaka Gas Co Ltd, where he was mainly involved in the design and construction of liquefied natural gas receiving terminals and was engaged in technical development. He currently serves as Manager of the Int'l Certification & Standards Harmonization Group, Planning Dept, JGA



Noboru Kizuki has done his master's degree in machinery engineering from National University in 2000.

He joined Osaka Gas Co., Ltd. in 2000, as an engineer of measuring and control system design and operation for LNG plants (2000-2010), Strategic Planning section (2010-2021). He currently serves as a Manager of Gas

Production Group, Engineering Dept. and Int'l Certification & Standards Harmonization, Planning Dept, JGA



Satoshi Yoshida has done his bachelor's degree of architecture from Waseda University. He has done his master's degree of urban planning from University of Washington. He is working for Japan's city gas industry since 1982 and currently serve as Senior Adviser to JGA.

He served as Senior Council Member for "Green Purchase Network", Japan (2007-2012), Chair, Sustainable Committee, International Gas Union (2012-2015).

He is an Executive Committee Member at International Gas Union since 2019. He is also a Senior Certified Architect registered in Tokyo.



Yuta Nezasa has done his master's degree in civil engineering from Graduate School of Engineering, Univ. of Nagoya, 2014.

He joined Toho Gas Co., Ltd., from 2014, he worked mainly for gas supply pipeline construction, disaster prevention planning, and corporate planning. He currently serves as Chief of Int'l Certification & Standards Harmonization, Planning Dept., JGA.



Yuki Tsuji has received his diploma from the School of Law at Keio University in 2014.

He joined the Japan Gas Association from 2014, and mainly worked for general affairs and residential sector's gas marketing. He currently serves as Chief of Int'l Certification & Standards Harmonization, Planning Dept., JGA

How to Analyze the Introduction Strategies for Radically New Technological Innovations? The Case of Autonomous Shipping

Mariah Kurtinaitis Joukes*, Roland Ortt, Mark de Bruijne, Linda M. Kamp

Faculty of Technology, Policy & Management, TU Delft, Delft, 2628 BX, The Netherlands

E-mail: mariahkurtinaitis@gmail.com (M. K. Joukes), I.R.Ortt@tudelft.nl (R. Ortt), M.L.C.deBruijne@tudelft.nl (M. d. Bruijne), L.M.Kamp@tudelft.nl (L. M. Kamp)

* Corresponding author: Mariah Kurtinaitis Joukes, E-mail mariahkurtinaitis@gmail.com

ABSTRACT: This article explores Maritime Autonomous Surface Ships (MASS) through the Technological Innovation Systems (TIS) framework to identify strategies for creating a market for this emerging technology. It examines the current state of TIS components for MASS and the factors shaping their development. Through a literature review and expert interviews, the article highlights key barriers, particularly the cost-benefit ratio, and suggests niche strategies to address barriers. Proposed strategies include the top niche strategy, lead user strategy, and hybridization niche strategy, aimed at overcoming challenges and facilitating broader market adoption of MASS in the future.

KEYWORDS: Autonomous ships, MASS, Technological Innovation System, niche strategy

1. Introduction

Radically new technological innovations represent an advance in technology that is so significant that either price/performance ratios are altered dramatically, or entirely new kinds of applications are made possible [1], [2]. Historical examples of such innovations in the automotive industry are hydrogen cars compared to cars with internal combustion engines. The diffusion of radical technological innovations differs from incremental innovations [3]. Radical innovations seldom start diffusing smoothly in an existing market. After the first introduction, an erratic process of introducing and withdrawing specific versions of the innovation in niche markets can be witnessed. It is remarkable how long it takes before radically new technological innovations start diffusing on a large scale [4-6]. This erratic process is often caused by a market that is full of barriers to large-scale diffusion. Another difference between radical and incremental innovations is the change in the socio-technical system required for large-scale diffusion. Radically new technological innovations often entail new complementary products and services, new customer

groups and adapted institutional arrangements, for example. This means that a new market arrangement must emerge before large-scale diffusion of that radically new technological innovation is possible. The ultimate market arrangement is often unknown early on, because of uncertain progress in the development of technologies, complex and unpredictable patterns of competition between old and other new technologies. The socio-technical system changes around a radically new technological innovation can be explored using a Technological Innovation System framework [7,8].

This article focuses on Maritime Autonomous Surface Ships (MASS) in order to illustrate how to analyze the wider system around an upcoming market and derive introduction strategies into that market. It takes the work of two previous articles as a starting point. In [9], we concluded that the diffusion of MASS is hampered by several barriers and MASS is thus applied in small scale niche applications only, such as survey and ferry transport. So, MASS has not yet started to diffuse on a large scale [9]. In [10], we analyzed the wider sociotechnical system around the upcoming market for MASS. We used the

Technological Innovation System (TIS) framework that can be used by maritime managers and researchers. This framework combines technological, economic, institutional and social aspects that are important for large-scale diffusion and relevant for strategy formulation. It contains seven so-called TIS building blocks of the socio-technical system from which new markets emerge [11]. We concluded that some important building blocks are still fully or partially absent that are required for large-scale diffusion of MASS [10].

However, a knowledge gap remains: to date, the literature has not explored the strategies that can be employed to spur the diffusion of MASS. The current article will investigate that. It extends the previous work by showing how several underlying conditions called influencing factors influence the status of the TIS building blocks for MASS. Based on the insight into the status of the TIS building blocks and the influencing factors we derive strategies to spur the diffusion of MASS. We focus on the application of MASS in commercial shipping rather than pleasure craft and on the societal and technological aspects that enable MASS as a whole, without detailing the enabling technologies that compose the various types of autonomous vessels. This leads to two research questions:

RQ1: What is the status of the building blocks that together form a Technological Innovation System (TIS) around MASS and how is that status influenced by influencing factors?

RQ2: How can potential (niche introduction) strategies be derived from the analysis of the TIS around MASS?

The research questions are scientifically relevant because the TIS framework provides a framework that can be used to analyze the context at the time a new market is emerging or does not even exist yet. That means the framework is an addition to the extant market research literature. The framework represents a (socio-technical) system perspective, but instead of most of the innovation system literature, it is not used to formulate policies for governments but to formulate strategies for companies [11]. The research questions are managerially relevant because the TIS framework and the strategies derived thereof provide an analytical tool for a situation in which standard market analysis instruments fail, and standard strategies are not applicable. As a result, many companies are found to fail in the situation of new and emerging markets for radically new technological innovations.

To answer these questions, we used a combination of literature research and two rounds of interviews with MASS researchers and practitioners. The unstructured and semi-structured interviews gave a broad perspective of the

barriers and a detailed view of the barriers and their underlying influencing factors. To answer research question 2, we used the approach outlined in the work of [11].

Our article is organized as follows: the theoretical notions are described in section 2, the method in section 3 after which results are presented in section 4. Conclusions, discussion and future research avenues are in section 5.

2. Theoretical background

In this section, we present a general TIS framework following the work of [11]. The framework is used to assess potential barriers and their influencing factors during the process of development and diffusion of radically new technological innovations, such as MASS. In practice, barriers to large-scale diffusion lead to a pattern of development and diffusion that differs from the well-known diffusion pattern introduced by [12]. Instead of a smooth S-shaped diffusion curve, a more erratic pattern can be witnessed for diffusion processes of radically new technologies [13]. After that, we explain how the framework can be used to derive possible strategies during an upcoming and turbulent market around a radically new technological innovation. In general, the notion of a TIS framework has been criticized in several ways [14]. One of the limitations revolves around the lack of attention for company strategies. Innovation system frameworks traditionally focus on governments and are mostly meant to inform policy formation rather than strategy formation. We tackle this limitation by creating a TIS framework from a company perspective that is shaped to facilitate strategy formulation.

2.1. A general TIS framework to formulate strategies

The framework is built up in a few steps and is composed of seven building blocks and seven influencing factors [11]. Each of these building blocks is essential for the emergence of a complete market and for the large-scale diffusion of a technological innovation in that market. Conversely, in the case of absence, incompleteness or incompatibility of one or more of these building blocks, large-scale diffusion is seriously hampered. The building blocks are:

1. Product performance and quality;
2. Product price;
3. Production system;
4. Complementary products and services;
5. Network formation and coordination;
6. Customers;
7. Innovation-specific institutions.

The explanation of each of the building blocks is summarized in Appendix 1.

These building blocks are important to formulate strategies. If all building blocks are complete and compatible, then large-scale diffusion is possible. In that situation, companies may consider adopting large-scale production and marketing strategies to sell the innovation. In contrast, if none of the building blocks is in place, complete and compatible, then introduction strategies are not a realistic option. If some building blocks are incomplete or incompatible, then these building blocks represent barriers to large-scale diffusion. In such cases, companies can sometimes adopt specific niche strategies that circumvent the barriers. This means that analysis of the building blocks is important to decide about specific aspects of introduction strategies: the timing of introduction (early versus late) and the scale of introduction (niche versus large-scale introduction).

Timing and scale of introduction can be derived from the status of the building blocks and are important aspects of an introduction strategy, but do not suffice. Especially in the face of an emerging and incomplete market it is important to deliberately formulate a niche strategy that fits the situation. That means we need more information to formulate the type of niche introduction strategy. To do so it is important to explore the causes of the incomplete or incompatible building blocks. These causes, referred to as influencing factors, are:

1. Knowledge and awareness of technology;
2. Knowledge and awareness of application and market;
3. Natural, human and financial resources;
4. Competition;
5. Macro-economic and strategic aspects;
6. Socio-cultural aspects;
7. Accidents and events.

The explanation of each of the influencing factors is summarized in Appendix 2.

If combinations of the influencing factors and building blocks are systematically explored several niche introduction strategies come to the fore. All these niche introduction strategies can be found in practice, as can be illustrated by historical cases of technological innovations. Different types of such niche introduction strategies are outlined in Table 1.

Table 1: General niche introduction strategies [15,16]

Niche Strategies	Definition
Demo Experiment and	Demonstrate the product in a controlled environment to enhance

Niche Strategies	Definition
Develop Niche Strategy	public knowledge about it and use the experiments to further develop the product quality. Relevant when knowledge of the technology is lacking.
Top Niche Strategy	Offer a limited number of products to the specific top end of the market that is not price-oriented until production capacity is feasible. Appropriate when knowledge of the technology is lacking and it affects its price.
Subsidized Niche Strategy	Subsidize the product for a particular segment if its use is considered socially important, but it is too expensive for the target customer. Applicable when the product is too expensive for the market, or when resources are scarce.
Redesign Niche Strategy	Introduce a simpler version of the product that can be produced with limited resources and sold at a lower price to explore the market. Appropriate when the product is too expensive to the market, when resources are scarce, or to fit socio-cultural aspects of suppliers and customers.
Dedicated System Niche Strategy	Offer the option of using the high-tech product independently from complementary products. Suitable when lack of knowledge about the technology affects the availability of complementary services.
Hybridization or Adaptor Niche Strategy	Combine the new product with the old one, allowing the use of all complementary products and services, or make the new product compatible with the existing complementary services. Applicable when resources are lacking, affecting the availability of complementary products.
Educate Niche Strategy	Educate customers and suppliers about the high-tech product to increase knowledge and awareness of the product capabilities. Useful when knowledge of the technology is lacking, affecting availability of suppliers and customers.
Geographic Niche Strategy	Introduce the high-tech product where laws and rules are less strict, or where customers, suppliers and

Niche Strategies	Definition
	resources are already available. Suitable when socio-cultural or macro-economic aspects affect the customers, and when institutional aspects pose a barrier to technology adoption.
Lead User Niche Strategy	Partner with lead users and innovators to co-develop products that innovators are willing to experiment with. Applicable when knowledge of the application is lacking, affecting a clear view of customer's applications.
Explore Multiple Markets Niche Strategy	Introduce the high-tech product in different customer applications to stimulate the explorative use in new functions. Relevant when knowledge of the application is missing, reducing the visibility of product application, usage pattern and benefits.

3. Methods

Given the lack of specific literature about the innovation management aspect of autonomous shipping and the complexity of the maritime industry, we deemed exploratory research the most appropriate method for this article. So, we based our analysis mainly on qualitative data in the form of literature review and semi-structured interviews. However, to safeguard a broad empirical basis using these two qualitative approaches, we interviewed a wide variety of experts and practitioners, and we interviewed them in two rounds. In the first round of interviews, open questions were posed regarding relevant actors and factors and possible barriers to large-scale use of MASS. In the second round, after consolidating the list of actors and factors, we returned to the experts to have their verdict regarding the barriers among those actors and factors for the purpose of validation.

The present article was based on the same research method as [10]. Therefore, the method section of the present article is very similar to the method section in [10].

We started our research with a literature review. This was necessary to understand the context in which autonomous shipping is being developed, as well as the barriers to MASS technology development already identified in the literature. In this way, we came to understand the factors that prevent or encourage MASS large-scale adoption by the maritime industry. This literature review also allowed us to adapt the general list of barriers to the specific context of MASS. For this literature review we made use of textbooks, conference proceedings, white papers, and scientific articles available in the TU

Delft Library database and on the internet. We used the keywords "Maritime", "Ships", "Autonomous", "Innovation", "Diffusion" and other similar terms in our search.

Because of the exploratory nature of this research, the broad variety of MASS usage in the maritime industry, the novelty of MASS technology and the lack of data regarding the barriers to its adoption by the maritime industry, next to the literature review we also made use of semi-structured interviews. The goal of semi-structured interviews is to have a specific set of questions for all interviewees but still allow room for further questions within the topic. This allows the interviewer to adapt the question list to fit the expertise of the particular interviewee [17]. We created a list of interview questions aimed at clarifying aspects of the literature available and acquiring more in-depth data about MASS and the TIS Framework actors and factors for the case of MASS.

3.1. Selection of interviewees

We performed 14 semi-structured interviews with MASS researchers, sailors, employees of companies that develop MASS technology, a ship owner, a Navy System Integrator involved in MASS projects and a shipyard manager. We chose these professionals based on their experience in MASS or their understanding of the maritime industry characteristics as shown in their LinkedIn profile, scientific articles or white papers written about the research topic. We contacted the interviewees using social media, or via the website of the companies for which they work. We asked each interviewee to refer to possible other interviewees, creating a snowball effect. The reason for choosing the only one of our interviewees not connected to the merchant marine sector, the Naval System Integrator, was not only his broad MASS knowledge acquired in the MASS joint industry project but also because the navy is known as the testbed for maritime technology, acting as an innovator according to the diffusion scale defined by [18]. We made sure to have a broad variety of interviewees in terms of their background and expertise. However, all our interviewees were based in the Netherlands. That can be a shortcoming but the effect of this is probably limited because the maritime industry in the Netherlands is among the industrial clusters that is closely involved with MASS and all our interviewees had an international perspective.

3.2. Data Analysis

We recorded all interviews, except when the interviewee preferred not to be recorded. In this specific case, we used our interview notes as material for further analysis. We created an automated transcript of the recorded interviews by making use of the program Otter AI, a program that transforms voice into text. After that, we reviewed the transcripts and interview notes according

to the content and we assigned codes to different categories for further data analysis. A complete list of keywords and databases used in the literature review of this article, the list with all interview questions, an example summary of the interviews, and the table with codes and categories of the interview analysis are available upon request.

Both the data from the literature and data from the interviews are mainly qualitative and hence our analysis is also qualitative. We started with a generalizable TIS framework. We used a combination of qualitative data and well-informed opinions (also qualitative in nature) voiced by experts to adapt that framework to the specific maritime industrial context. After adapting the framework, we could qualitatively assess the main barriers. For some of the barriers, more detailed and quantitative data could have been gathered but that would not have resulted in a better comparison of the relative importance of these barriers. Neither would it form a better basis to select relevant niche introduction strategies.

4. Results

As presented in the Theory section, many breakthrough technologies require the presence of specific actors and factors, the so-called TIS building blocks. This section presents and explains the (f)actors affecting MASS specifically, based on our desk research and interviews.

4.1. RQ1: What is the status of the building blocks that together form a Technological Innovation System (TIS) around MASS and how is that status influenced by influencing factors?

4.1.1. Quality & Performance

Quality and performance are inherent to each product. For MASS, its quality is defined by its capacity to increase ships' performance in terms of reliability, efficiency, and predictability. These performance measures were mainly rated by our interviewees as very important. However, MASS has not yet proven its performance in all these aspects, rendering this building block partially present [10].

4.1.2. Cost-Benefit Ratio

The cost-benefit ratio for MASS is calculated in terms of return on investments (ROI), which interviewees believe can be achieved by crew reduction and/or increased space for cargo because of MASS implementation onboard. Crew reduction is not (yet) regulated, making the business case for MASS adoption unclear to ship owners. Technology providers, on the other hand, ensure there is a valid return on investment to MASS, regardless of crew reduction. All interviewees

considered cost-benefit ratio an extremely important building block. The Royal Dutch Navy interviewee is the only exception since the focus of the navy lies on reducing the number of crew members and less on costs. It is valid to note that navy vessels are not obligated to follow merchant marine regulations and therefore can determine their crew size based on their needs and not on legal requirements. Most interviewees agree that a beneficial cost-benefit ratio is still missing in the MASS TIS [10].

4.1.3. Production System

MASS is not expected to ever reach mass production, since ships are not built in the same numbers as cars. Our interviewees consider that a MASS production system is the capability to produce standardized items and have a set of hardware and software standards, regulated by accredited institutions. These standards do not exist yet, but the equipment is currently being built and sold. Therefore, this building block was considered present and moderately important [10], since the implementation of standards in the form of technology battles is common within technology development [19].

4.1.4. Network Formation & Coordination

When customers evaluate the value of a technology, they consider the perceived value of having a network that can or will increase the product's value in the future [20]. Our interviewees recognize there is little collaboration among companies offering MASS and the large ones create a closed loop of products, which shields smaller companies from a possible network. Knowledge is also not shared among companies offering MASS. Despite not being considered important, 'network formation and coordination' was recognized as a building block to MASS diffusion and deemed partially present [10].

4.1.5. Customers

The customers for MASS technology are ship owners willing to purchase and install the technology in their vessels. Without customers, there is no technology diffusion. Companies offering MASS claim to have sold and installed equipment; other companies mentioned potential customers and current clients interested in the technology. However, there is no widespread interest in the technology (yet). Hence, the building block is extremely important and only partially present [10].

4.1.6. Innovation Specific Institutions

MASS is not yet regulated by the International Maritime Organization (IMO), which means its conventions that regulate equipment and manning onboard vessels worldwide also do not address MASS.

Most interviewees see this lack of regulation as a barrier to MASS diffusion. Given the strong weight of the IMO recognition of the technology and its impact on influencing factors described further in this section, this building block was evaluated as extremely important.

4.1.7. Safety

One interviewee described safety as a preconditional item that can only be increased onboard. From his perspective, every change on a vessel must increase safety, otherwise, it is not acceptable. All other interviewees and many researchers agree on this. This was the reason we decided to adapt the general list of TIS building blocks (see section 2) by including 'safety' as a separate building block. In other industries 'safety' may be seen as a kind of cost-benefit factor but in the maritime industry safety is a separate factor. What the interviewees cannot agree on is whether MASS will increase safety onboard or not [21-24]. The idea that MASS has the potential to reduce accidents onboard is widespread, still; few see this potential currently realized. Therefore, the building block is extremely important, but only partially present [10].

Summarizing the above analysis, we show the status of the TIS building blocks for MASS in Table 2 below.

Table 2: MASS Building blocks

MASS building blocks		
Quality & Performance	Very Important	● Partially Present
Cost-Benefit Ratio	Extremely Important	● Absent
Production System	Moderately Important	● Present
Network formation & Coordination	Slightly Important	● Partially Present
Customers	Extremely Important	● Partially Present
Innovation Specific Institutions	Extremely Important	● Partially Present
Safety	Extremely Important	● Partially Present

After the analysis of the building blocks, we further investigated the factors that influence the presence of these building blocks aiming at understanding the causes of the factors that currently influence the adoption of MASS.

4.1.8. Knowledge and awareness of technology

As a technology that is not yet fully proven, knowledge about MASS technological principles is not yet

fully mastered. Some operational aspects of MASS, such as the application of Collision Regulation (COLREG) rules issued by the IMO, are not yet grasped by AI systems, given it was made to be used by sailors. This opens room for interpretation. Communication difficulty between vessels and harbors or between conventional vessels and autonomous vessels was also pointed out by many interviewees as a difficult operational aspect to be solved. For lower levels of autonomy, where the crew is still onboard or when an operator is in the control room these aspects do not pose a problem, as the operator can correct possible AI mistakes. But it is an important influencing factor for unmanned autonomy. The lack of deep understanding of MASS technological principles is also a barrier to the regulators, who do not know enough about the advantages of disadvantages of the technology to regulate its use. This influencing factor is, then, partially hampering MASS adoption in the current MASS TIS [10].

4.1.9. Knowledge and awareness of application and market

Our interviewees noted that knowledge about the MASS application is a critical aspect, highlighting that even in the literature there are still misunderstandings about the difference between an autonomous and unmanned vessel, for example. Because of this reoccurring misunderstanding, many seafarers are afraid of being replaced by a machine. Ship owners are also not aware of how MASS can contribute to their operations. Most of the news about autonomous vessels is about unmanned autonomous projects, reinforcing the view that autonomous vessels are unmanned. Technology developers and researchers also lack knowledge about the maritime industry, disregarding the peculiarities of the industry such as its conservative thinking and the seafarers' knowledge as technology operators. Despite an increase in knowledge about MASS application in recent years, this influencing factor is still considered a partial barrier [10].

4.1.10. Natural, human, and financial resources

From the perspective of the ship owners, visualizing a good cost-benefit ratio is still difficult without the guarantee of crew reduction, and without this clear financial benefit, most companies do not want to have MASS only to increase safety. In this context, the problem seems to be the allocation of resources, rather than its availability.

The education of seafarers was also mentioned as a necessary aspect of MASS operation. The maritime schools follow IMO guidelines regarding the curriculum and as IMO has not yet certified this technology, learning how to

work with MASS is not yet in the curriculum. This aspect does not have a large influence because professional courses can be offered to increase the knowledge of officers after their graduation, as is currently done with Dynamic Positioning Officers.

The shortage of qualified seafarers [25,26] to operate the vessels is an encouraging factor for the development and adoption of MASS. Commercial vessels can enhance their profit margin, and the navy can maintain the vessels' operations even when there is little public interest in working for the military.

None of the interviewees mentioned natural resources as an influencing factor. The shortage of human resources has a positive influence on the adoption of MASS. Both the availability of resources and the lack of qualified seafarers influence MASS adoption.

4.1.11. Competition

At the time of the data collection, not many companies we interviewed were working on MASS solutions and very few could offer a complete integration of maritime equipment and autonomy. This could be one of the reasons companies offering MASS solutions do not experience competition among themselves, but with the status quo, namely the conventional ships. MASS can make barge transport cheaper and more efficient, saving fuel costs and reducing waiting time, competing with truck transport. The current focus on CO₂ emissions challenges truck transportation and creates an ideal window to make this modal shift to barges. However, when trucks become electric, and later autonomous, barges will most likely have reduced commercial value.

Even though many international companies offer MASS solutions to the market, the competition is still too small to have a positive influence on the price and quality. Consequently, MASS is still not able to compete on equal terms with the status quo, either in the form of conventional vessels or in a modal shift, which deems the competition a partial hampering factor.

4.1.12. Macroeconomic and strategic aspects

The maritime industry was recently at the bottom of its ordering cycle, and it will go further up in the coming 10 years, so more ships are expected to be ordered and built in the near future. The decision of which technology to add to a newly built ship considers an operational asset life of 20 years, ideally with as few modifications as possible. The lack of regulations about MASS increases the financial risk of investment for ship owners, which is an important aspect of the low-profit-margin risk-averse

maritime industry. The Dutch Navy has more flexibility regarding national and international regulations and is not pressured by profit margin, decreasing the risk of the investment. This is similar for navies in other countries. The ordering cycle stage and the reduced naval risks encourage both the military and the commercial vessels to invest in new technology. The difference is the level of risk each niche of the maritime sector needs to accept to design a strategy that suits the current macroeconomic scenario. Therefore, macroeconomic and strategic aspects represent an encouraging factor, only depending on the risk inclination of the investor.

4.1.13. Socio-cultural aspects

The maritime industry is not known for its openness to change; it is a conservative industry with a very strict task division on board. Society at large also influences the adoption of autonomous technologies, and people want the technology to be safer than humans; a mistake from a machine that ends up in the loss of human life is unacceptable, despite its frequency. Besides the overall fear of AI, many seafarers worry about their job and their safety, a social aspect of MASS that is also related to the knowledge and application of the technology. While society at large plays a role, ultimately the crew members are paramount to MASS acceptance; the crew operates the system, therefore if the members do not operate it correctly, the technology will not deliver its expected benefits. Benefits not delivered affect performance, which can affect adoption and stop (or delay) diffusion. Therefore, the socio-cultural influencing aspect is judged as a partially hampering factor.

4.1.14. Accidents and Events

National and international disasters and events have historically changed maritime regulations, such as the current international regulations which were developed after tragic accidents like the sinking of the Titanic, and the oil spill caused by the Exxon Valdez [27]. Accidents, however, are not unusual at sea; 65.8% of the accidents in the maritime industry are caused by human errors and MASS could decrease these numbers [28,29]. Despite not having one accident to name, the large number of accidents involving ships can be characterized as an encouraging factor to the adoption of MASS, given one of the selling points of the technology is to reduce accidents.

4.1.15. Complementary Products & Services

The Ortt & Kamp framework [11] introduces complementary products and services as a TIS building block, which is a factor capable of influencing the adoption

of a breakthrough technology on its own. Complementary products & services for MASS were identified as reliable internet, a standard for data transfer and differentiated insurance policies for vessels equipped with MASS [10]. None of these, however, is seen by the interviewees as capable of affecting MASS adoption directly. Therefore, for MASS we consider them as influencing factors instead of a TIS building block. According to our interviewees, differentiated insurance policies, in the form of a complementary service, can influence the cost-benefit ratio, a building block that is paramount for the adoption and subsequent diffusion of MASS [10]. Data Transfer Standards and 100% internet coverage can also improve technical performance, influencing the quality & performance building block. Given its absence, Complementary Products and Services are a hampering influencing factor in the MASS TIS [10].

4.1.16. Operational Aspects

Operational aspects are not mentioned as influencing factors in Ortt & Kamp's framework [11] but were mentioned by many interviewees. These aspects refer to factors that influence the regular operation of autonomous vessels, such as maintenance in the propulsion engines, especially diesel engines, cargo value and inspection, communication with other vessels and port facilities and finally stormy seas. These aspects are related to the unmanned operation of the vessels because as long as the vessel is locally operated, there will be a crew on board to give maintenance to the engines, ensure cargo safety, communicate, and perform maneuvers to ensure the vessel's water tightness. Consequently, this factor does not have an immediate influence on MASS adoption, but it will have an effect later in the large-scale diffusion of MASS if/when vessels become operated by Remote Operation Centers (ROCs).

The status of the different influencing factors for MASS is shown in Table 3 below.

Table 3: MASS Building blocks

MASS Influencing Factors	
Encouraging Factor Partial Barrier Barrier	
Knowledge and awareness of technology	
Knowledge and awareness of application and market	
Natural, human, and financial resources	

Competition	
Macroeconomic and strategic aspects	
Socio-cultural aspects	
Accidents and Events	
Complementary Products & Services	
Operational Aspects	
MASS Building Blocks	
Present Partially present Absent	
Quality & Performance	
Cost-Benefit	
Production System	
Network formation & Coordination	
Customers	
Innovation Specific Institutions	
Safety	

4.2. RQ2: How can potential (niche introduction) strategies be derived from the analysis of the TIS around MASS?

Now that we have knowledge of the status of the different TIS building blocks for MASS and how the influencing factors influence that status, we can derive niche strategies for the introduction of MASS in the market. The logic here is that the niche strategies are specifically derived to address incomplete or partially complete TIS building blocks in combination with their influencing factors.

4.2.1. Quality & Performance

The main issue with MASS lies in the fact that the technology is not yet proven. Little is known about commercial tests as companies only share Public Relations articles when they have positive outcomes. Most interviewees indicated that they did not know much about how MASS works.

Demonstrating, however, is not impossible and it was suggested by crew members interviewed as a way to prove MASS can deliver what it promises. The Demo experiment and Develop Niche strategy, explained in the theory

section, suggests a similar approach. In this niche strategy, the technology developers demonstrate the product in a controlled environment to further develop it and learn with it at the same time it enhances the awareness of the public [15]

Given the nature of MASS, it needs a demonstration that can be followed at a distance, like the blog created by the Mayflower developers [30] where it was possible to follow the vessel online, demonstrating its capabilities on a daily basis. Such live demonstrations can enhance trust in the system's capabilities. As a by-product, such online activity might also improve awareness of the technology and its application. To ensure that the demonstration can only take place digitally, this niche strategy is named the Digital Demo experiment and Develop Niche Strategy.

4.2.2. Cost-Benefit Ratio

One point was very clear for most interviewees: If the MASS suppliers want to sell, they need to present a clear business case and therefore a positive cost-benefit ratio. From the perspective of the companies offering technology, regulations must recognize MASS and allow crew reduction. If crew reduction is approved, the decision to adopt MASS becomes clear-cut for ship owners. IMO is working on a MASS code, with the non-mandatory part expected to be ready in 2025 and the mandatory code to enter force in 2032. Two strategies that can be used to overcome the cost-benefit ratio barrier, both described in the theory section [15]. The first one is the Top Niche strategy, which aims at a market segment that is not constrained by the price of the technology. The Royal Dutch Navy is more interested in crew reduction than in cost-benefit ratio and it does not have the obligation to comply with international rules regarding crewing [10]. Therefore, it is a suitable option for the Top Niche strategy. According to our interviewees, another maritime market segment known for its lack of budget constraints is the leisure yacht market segment, which can be open to MASS adoption without cost-benefit concern.

The second one is the Hybridization Niche Strategy, which entails making the new product compatible with existing complementary services or combining the new product with the old version [15]. A similar strategy is being pursued by two companies interviewed. Both offer MASS to existing vessels, while one ensures it sells a product and not a project, offering installation in every ship. The other focuses on retrofitting or installing the equipment in existing and operating vessels. Currently, the second company offers operational visibility to ship owners, not autonomy. The equipment installed, however, will learn with the vessel and, in the future, the company intends to add autonomy to the already sold product. The approach used by the second company resembles a combination of the previously mentioned Hybridization

Niche Strategy with the Lead User Niche Strategy, which is explained as a partnership with users that are willing to co-develop and experiment with technology [15].

4.2.3. Network Formation & Coordination

According to our interviewees, there is little collaboration between MASS providers and the network of complementary services, which directly affects the knowledge of the technology and its application. Interestingly, our interviewees understand the network as the exchange of information and ideas between the industry and its players. In the MASS context, a university of technology studying MASS technology could fulfil the public innovator role in the system. Despite not having control over what the researchers develop in the universities, companies offering MASS solutions could profit from a close relationship with academic researchers and enhance network formation for MASS. In general, research institutes can play an important coordinating and even leading role in the initial stages of development and diffusion of a radically new technology (e.g., [31])

4.2.4. Customers

Without customers, there is no diffusion and to purchase, customers need to be aware of the technology and have the resources to purchase it [11]. We found that potential customers lack awareness of MASS competencies, which affects their interest in purchasing and installing the technology in their vessels. To increase customers' interest, companies offering MASS solutions need to educate their market.

The Educate Niche Strategy is recommended to ensure customers have a better understanding of the technology [15]. The authors suggest organizing conferences and closing partnerships with universities to enhance the knowledge and awareness of innovative technologies. Organizing conferences can be rather costly to companies struggling with financial resources, and this is the situation for most small companies offering MASS solutions.

Another option to increase technology awareness is to partner with universities in projects related to MASS or make active use of the already existing autonomous vessel programs funded by national governments and institutions. Participation in these projects can have a twofold contribution to MASS adoption. On the one hand, smaller companies can piggyback on the resources of national programs to enhance awareness about their product. On the other hand, the participation of informational intermediaries, the universities and the joint industry projects, facilitates the network formation for MASS. Therefore, adopting the Educate Niche Strategy can help the companies solve the Customers' barrier to MASS adoption and diffusion.

4.2.5. Innovation Specific Institutions

Institutions, more specifically regulations regarding the number of crew members onboard, pose a large barrier to MASS adoption. The uncertainty about crew reduction possibilities hampers the adoption of MASS, except for military applications. Therefore, changes in regulations can have a substantial influence on MASS diffusion.

The Geographic Niche strategy is proposed when breakthrough technologies are restricted by regulations, suggesting introducing the technology in areas with more flexible rules [15]. Considering the power of the flag state to determine the number of crew members in vessels flying their flag and the concept of territorial waters [32], companies offering MASS can request a Minimum Safe Manning reduction with the flag state, and within territorial waters. Despite not having control of the results of these crew reduction requests, the companies offering MASS solutions stand better chances of success focusing on national entities, a solution that was also proposed by Rødseth [23] than waiting for an IMO international recognition.

4.2.6. Safety

Safety is not a building block encompassed by the Ort & Kamp framework [11], although it is paramount to MASS adoption [10]. Therefore, there is no entry strategy presented in our theory section designed to ensure the safety of a product in its market introduction.

Wróbel, Montewka, & Kujala [25] highlight the complexity of proving safety. Safety can be proven in different ways, for example by meeting requirements, by testing designs or by measuring actual accidents after implementation. All these ways have their shortcomings and are incomplete in terms of ensuring safety. According to Ventikos, Chmurski, & Louzis [33] the autonomous shipping systems should mimic human logic and actions in their decisions. For example, when an autonomous ship gives way to another ship, the maneuver should be clear and visible, rather than only optimal. Although helpful, the authors' suggestion does not influence MASS adoption because the input of humans in the AI decision-making process is not visible to crew members and ship owners that will operate and buy the technology respectively. Therefore, this study cannot propose a useful strategy yet to address safety concerns. The table with the main strategies that we found for the main barriers is shown in Table 4 below.

Table 3: Strategies to Surmount MASS Barriers

Strategies to Surmount MASS Barriers		
● Present	● Partially Present	● Absent
Quality & Performance	●	Digital Demo experiment and Develop Niche Strategy.
Cost-Benefit Ratio	●	Top Niche Strategy.

		Hybridization Niche Strategy
		Lead User Niche Strategy
Production System	●	
Network formation & Coordination	●	Build relationships with institutes researching MASS
Customers	●	Educate Niche Strategy
Innovation Specific Institutions	●	Geographic Niche Strategy
Safety	●	... (No strategy defined yet)

5. Discussion

In this article, we focus on MASS and in order to illustrate how to analyze the wider system around a radically new technological innovation. This analysis allows us to explore an upcoming market and derive introduction strategies into that market.

5.1. Answers to the research questions

Our first research question is 'What is the status of the building blocks that together form a Technological Innovation System (TIS) around MASS and how is that status affected by influencing factors?'. Our analysis revealed that the main barrier for MASS is the lack of cost-benefits, or a lacking viable business model to apply MASS. Other building blocks are partly complete except for the production system for MASS technology, which was found to be ready and complete. So, only one building block of the TIS is considered complete and all the other ones are incomplete. That means that the TIS analysis reveals that MASS is not even close to the start of large-scale diffusion. Similarly, several influencing factors were found to have a negative effect, among which are operational aspects and a lack of complementary products and services. The status of the influencing factors confirms the idea that MASS will not start to diffuse on a large scale soon, as also shown in [34].

Our second research question is 'How can potential (niche introduction) strategies be derived from the analysis of the TIS around MASS?'. The strategies presented in this article show different ways to deal with each incomplete or absent building block in the MASS TIS. Depending on the influencing factor that is found to be the cause of the incomplete building block, different niche strategies can be adopted. The lack of quality and performance of the technology and the lack of knowledge of how to apply the technology in practice, for example, can both be improved with a Demo Experiment and Develop Niche strategy. In this niche strategy, the technology developers demonstrate the product in a controlled environment to further develop it and learn to use it and at the same time it enhances the awareness of the public [15]. The biggest barrier for MASS turned out to

be the lack of a good business model, referred to as the cost-benefit ratio barrier. Two strategies are suggested to overcome the cost-benefit ratio barrier. The first one is the Top Niche strategy, which aims at a market segment that is not constrained by the price of the technology. An example of a niche segment in which this strategy may work is the Royal Dutch Navy, which is more interested in crew reduction than in financial costs. Another segment in which this strategy may apply is the high-end leisure yacht market segment. The second strategy to overcome the cost-benefit barrier is the Hybridization Niche Strategy, which entails making the new product compatible with existing complementary services or combining the new product with the old version. In a similar way, several other specific niche introduction strategies are suggested to deal with incomplete building blocks, or barriers in the market.

5.2. Contributions

The article contributes to the scientific domain of TIS in several ways. Firstly, the article presents an adapted TIS framework, specifically reflecting the industrial context of international cargo shipping. This focuses attention on the notion that a general framework needs to be modified for the context in which it is used. We also found several such modifications necessary in studies adapting the same general TIS to other industrial contexts (e.g., [35-36]). One of the adaptations made to the TIS framework in this article is that in the MASS, safety is considered a building block instead of an influencing factor. We found it to be affected by the influencing factors accidents and events, knowledge and awareness of the technology and socio-cultural aspects. With regard to strategy for this building block, it is possible to infer that a strategy to overcome the safety barrier must include measures related to the influencing factors accidents and events, socio-cultural aspects (the norms and values of behaviors at sea) and awareness of the technology. However, this study has not found any suitable strategies yet to overcome this barrier.

Secondly, another way to reflect the intricacies of a particular context is to indicate how important each TIS building block's status is. In this article, this has been done for the first time in a TIS-analysis. Building blocks represent aspects that generally, across a wide range of industrial and geographic contexts, are important for large-scale diffusion of a technological innovation. However, depending on the geographic or industrial context, their importance may vary considerably. In the case of MASS we found that the only complete building block was considered just moderately important, for example. In contrast, the building block that was fully incomplete, however, was considered extremely important by interviewees. This extra analysis in more detail confirms the overall conclusion that MASS is far from large-scale diffusion.

The article also represents managerial and societal contributions. The framework is a kind of dashboard that, when tracked carefully over time, indicates when MASS may begin to diffuse on a large scale. That is an important managerial instrument because it allows companies to prepare for upcoming large-scale diffusion. Conversely, the framework can also prevent unnecessary preparation for introducing or adopting MASS when the conditions for large-scale diffusion are simply not met by far. Societally, the analysis using the TIS framework can help governments and companies to address barriers, remove them and hence speed up the development and start of large-scale diffusion.

5.3. Limitations

The first limitation is that the research is time-constrained and in view of the speed of development of improved technological systems, the results can be outdated within a few years. On the other hand, some radically new technologies seem to remain in a status that is considered to be "just before large-scale diffusion" for decades. Inevitably the analysis needs to be updated in the future to explore whether the status of building blocks and influencing factors have changed. As a result of such changes, new niche introduction strategies may need to be considered.

A second limitation is that interviewees came from one country, the Netherlands. The reason is that we had access to the experts, researchers and practitioners in this country. The effect of this limitation may be limited because all interviewees showed themselves to have an international perspective.

The TIS framework considers the building blocks as separate, independent aspects. That is a third limitation. Actors and factors in a TIS are interrelated, which means that adopting a strategy to deal with a barrier might create a new barrier and hence the strategy may not have the desired effect. An example can be taken from the cost-benefit ratio building block and the Top Niche Strategy. If a company offering MASS solutions focuses only on this strategy as the path to diffusion, it has little chance of succeeding, as both the navy and the leisure yachts market segments are only a small percentage of the maritime sector. The strategy would be minimizing the importance of the customers, safety and institutions, all deemed to be also extremely important in the MASS TIS. Since many factors are considered extremely important, a pathway addressing them simultaneously has a higher probability of success on the road to MASS diffusion.

5.4. Future research

In future research a wider set of interviewees, from across different countries, can be approached. This may reveal different conditions in different regions, and how

these conditions call for different niche introduction strategies.

We used qualitative data and analysis methods. In future work it may be useful to track the changes of the building blocks and influencing factors over time. That may call for a quantification of measuring these building blocks and influencing factors. Using the current qualitative approach, we could only explain why a building block changes from red (a complete barrier) to orange or green (no barrier). Such a crude and qualitative assessment may suffice to reflect a current condition and choose a niche introduction strategy but will not suffice to carefully track changes over time.

Then there is the issue of safety concerns. As this building block was newly added to the TIS framework in this study, there is no suitable niche strategy yet to address this. This is another interesting topic for future research.

The interrelation of actors and factors in a TIS calls for more advanced models taking all the building blocks and influencing factors into account. The current TIS analysis represents an overview that reflects conditions at one point in time and ignores possible interactions. This suffices to select possible niche introduction strategies that fit these conditions. Inevitably the analysis using the TIS framework needs to be repeated over time to reflect changes in the conditions. If interrelationships are taken into account, then more dynamic models could be developed that also capture the causes of changes over time. Studying the development and diffusion of radically new technological innovations, it becomes clear how fascinating the dynamics in these processes can be.

6. Conclusion

Large-scale application of Maritime Autonomous Surface Ships (MASS), being a radically new technological innovation, is still hampered by several barriers. In this paper we have applied the TIS framework to identify the main barriers and their causes and to formulate niche strategies. The main barriers include the cost-benefit ratio, safety issues and quality and performance. Niche strategies such as the Top Niche Strategy can be used to introduce MASS in niche markets first, for example the navy. This could be a first step towards the introduction of this technology on a large scale.

References

- [1] M. L. Tushman, P. Anderson, "Technological discontinuities and organizational environments," *Administrative science quarterly*, vol. 31, no.3, pp. 439-465, 1986.
- [2] R. Garcia, R. Calantone, "A critical look at technological innovation typology and innovativeness terminology: a literature review," *Journal of Product Innovation Management*, vol. 19, no. 2, pp. 110-132, 2002.
- [3] R. D. Dewar, J. E. Dutton, "The adoption of radical and incremental innovations: an empirical analysis," *Management Science*, vol. 32, no. 11, pp. 1422-1433, 1986.
- [4] S. P. Schnaars, *Megamistakes*. Collier Macmillan, 1989.
- [5] R. Agarwal, B. L. Bayus, "The market evolution and sales takeoff of product innovations," *Management Science*, vol. 48, no. 8, pp. 1024-1041, 2002.
- [6] J. R. Ortt, "Understanding the Pre-diffusion Phases," in *Gaining Momentum*, Imperial College Press, pp. 47-80, 2010.
- [7] B. Carlsson, R. Stankiewicz, "On the nature, function and composition of technological systems," *Journal of evolutionary economics*, vol. 1, pp. 93-118, 1991.
- [8] A. Bergek, "Technological innovation systems: a review of recent findings and suggestions for future research," *Handbook of sustainable innovation*, vol. 30, pp. 200-218, 2019.
- [9] M. Kurtinaitis Joukes, J. R. Ortt, M. de Bruijne, "Autonomous shipping: current status and main barriers to large-scale diffusion," (Unpublished, 2023).
- [10] M. Kurtinaitis Joukes, J. R. Ortt, M. de Bruijne, "A system's perspective analysis of barriers to MASS large-scale diffusion", *Proceedings of the 29th International Conference on Engineering, Technology, and Innovation: Shaping the Future, ICE 2023. IEEE*, 2023.
- [11] J. R. Ortt, L. M. Kamp, "A technological innovation system framework to formulate niche introduction strategies for companies prior to large-scale diffusion," *Technological Forecasting and Social Change*, vol. 180, no. 121671, pp. 1-17, 2022.
- [12] E. M. Rogers, *Diffusion of Innovations*, The Free Press, 2005.
- [13] J. R. Ortt, J. P. L. Schoormans, "The pattern of development and diffusion of breakthrough communication technologies," *European Journal of Innovation Management*, vol. 7, no. 4, pp. 292-302, 2004.
- [14] H. Markard, M. Hekkert, S. Jacobsson, "The technological innovation systems framework: Response to six criticisms," *Environmental innovation and societal transitions*, vol. 16, pp. 76-86, 2015.
- [15] J. R. Ortt, D. J. Langley, N. Pals, "Ten niche strategies to commercialize new high-tech products," *2013 International Conference on Engineering, Technology and Innovation (ICE) & IEEE International Technology Management Conference*, pp. 1-12, 2013.
- [16] M. Kurtinaitis Joukes, "The Diffusion of Maritime Autonomous Surface Ships," (TU Delft repository, 2012). <http://resolver.tudelft.nl/uuid:b0bfb05f-83d2-4a47-a28d-bdb067793a3d>, 2021.
- [17] M. N. K. Saunders, P. Lewis, A. Thornhill, *Research methods for business students*, Pearson Education, 2015.
- [18] E. M. Rogers, "The diffusion of home computers among households in silicon valley," *Marriage Fam. Rev.*, vol. 8, no. 1-2, pp. 89-101, 1985.
- [19] G. Papachristos, G. van de Kaa, "A system dynamics model of standards competition," *IEEE Transactions on Engineering Management*, vol. 68, no. 1, pp. 18-32, 2020.
- [20] M. A. Schilling, *Strategic Management of Technological Innovation*, Mc Graw Hill Education, 2017.
- [21] A. Komianos, "The autonomous shipping era. operational, regulatory, and quality challenges," *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, vol. 12, no. 2, 2018.
- [22] "Future proofed?," *Nautilusfederation.org*. [Online]. Available: https://www.nautilusfederation.org/globalassets/federation/resources/downloads/report_auto_8.2.18.compressed.pdf. [Accessed: 27-Feb- 2023].

- [23] Ø. J. Rødseth, H. Nordahl, "Definitions for autonomous merchant ships," (Unpublished, 2017).
- [24] "The future of Smart Autonomy is here," Wartsila.com. [Online]. Available: https://www.wartsila.com/docs/default-source/whitepapers/smart_autonomy_e-report_2021-final.pdf?sfvrsn=31024843_3. [Accessed: 27-Feb-2023].
- [25] K. Wróbel, J. Montewka, P. Kujala, "Towards the assessment of potential impact of unmanned vessels on maritime transportation safety," *Reliab. Eng. Syst. Saf.*, vol. 165, pp. 155–169, 2017.
- [26] H. Leggate, "The future shortage of seafarers: will it become a reality?," *Maritime Policy & Management*, vol. 31, no. 1, pp. 3–13, 2004.
- [27] Galić S, Lušić Z, Pušić D. "Seafarers market," *International Journal of New Trends in Arts, Sports & Science Education*, vol. 1, no. 3, 2012.
- [28] A. A. Hebbbar, "Risk communication and maritime safety legislation," (World Maritime University Dissertations, 2006), Retrieved from http://commons.wmu.se/all_dissertations/346?utm_source=commons.wmu.se%2Fall_dissertations%2F346&utm_medium=PDF&utm_campaign=PDFCoverPages
- [29] "The future of Smart Autonomy is here," Wartsila.com. [Online]. Available: https://www.wartsila.com/docs/default-source/whitepapers/smart_autonomy_e-report_2021-final.pdf?sfvrsn=31024843_3. [Accessed: 27-Feb-2023].
- [30] Mayflower 400. 2020. The Mayflower Story. Retrieved from 1620-2020 Mayflower 400: Available at: <https://www.mayflower400uk.org/education/the-mayflower-story/>
- [31] O. Dedehayir, S. J. Mäkinen, J. R. Ortt, "Roles during innovation ecosystem genesis: A literature review," *Technological Forecasting and Social Change*, vol. 136, pp. 18–29, 2018.
- [32] J. N. Mansell, Flag state responsibility: Historical development and contemporary issues, Springer Science & Business Media, 2009.
- [33] N. P. Ventikos, A. Chmurski, K. Louzis, "A systems-based application for autonomous vessels safety: Hazard identification as a function of increasing autonomy levels," *Saf. Sci.*, vol. 131, no. 104919, p. 104919, 2020.
- [34] C. Kooij, "Towards unmanned ships, A task analysis to identify economically viable low manned ship concepts," (Ph. D Thesis, Delft University of Technology, 2021).
- [35] C. Verdegaaal, "Developing Effective Strategies for the Deployment of Sustainable Aircraft Technologies by the Aviation Industry," (Thesis report MSc MOT, Delft University of Technology, 2023).
- [36] M. P. Dwisatyawati, "Analysis on Mass Adoption of Solar Electric Vehicles in Indonesia. A Technological Innovation System Study," (Thesis report MSc COSEM, Delft University of Technology, 2022).

Copyright: This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).

7. Appendices

7.1. Appendix 1: Descriptions of the TIS building blocks

Building Blocks	Description
Product Performance and Quality	The newly developed technological product has sufficiently good quality and performance or is

	expected to have it shortly when compared to competing products. If the new technology suffers from low quality and is unable to meet the customer's needs and requirements large-scale diffusion will be hampered.
Product Price	The price of a product involves not only its monetary costs but also non-financial costs such as time, effort to implement the new product, switching costs, and costs to find new suppliers. For large-scale diffusion, the price should be reasonable when compared to other competitive alternatives of the same technology.
Production System	A production system that can deliver large quantities of high-quality products is paramount to the large-scale diffusion of technological innovation. Not only creating a production system but fine-tuning it to profit from the learning effect costs time and money, which can delay the diffusion process.
Complementary Products and Services	Complementary products and services support production, distribution, adoption and finally the disposal, if necessary, of the innovative technological product. Together, the network of complementary products and services can induce other innovations and motivate companies to align their strategy, which will ensure large-scale diffusion. The lack of these products and services forms a barrier that blocks the diffusion.
Network Formation and Coordination	A supply chain network with suppliers of parts, distributors, complementary services, and other actors is important for the diffusion of an innovative product. The lack of alignment in this network can impede the large-scale diffusion of the technology.
Customers	The customers are extremely important in the TIS and the diffusion of the innovative technology, without consumers, there is no diffusion. Customers must be able to afford the product,

	understand enough the product's capabilities to have the correct understanding of its risks and benefits, as well as use it. The development of technological innovation without the perspective of the future customer often results in issues that hamper diffusion, such as lack of integration with the customer's routine of use, alignment to other technologies already in use. Some technologies have to be adjusted later to fit the customer's wishes.		actors. Lack of application knowledge can affect customers as well as suppliers. Suppliers without application knowledge are not able to target the correct customer segment, and customers without knowledge of the purpose, how to use or buy an innovative product will be a barrier to product diffusion. Market analysis, experimentation, and learning by doing are ways of developing this type of knowledge.
Innovation-Specific Institutions	To form a TIS network, innovation-specific institutions must be present. These institutions refer to regulations, laws, standards, and government policies, which can either block or encourage the formation of the TIS. Factors such as stability of the political and legal systems, quality norms, and property rights produce trust in the system, which in turn, increases investments and facilitates the TIS formation.	Natural, Human and Financial Resources	All three types of resources, natural, human, and financial resources are required for the production system of the innovation itself as well as the complementary products and services. The lack of raw material, people with the necessary competencies or capital to invest in technological innovation can hamper technology diffusion. The lack of financial capital specifically blocks the development of radically new products and their future market formation.
		Competition	The presence of competition influences market formation, especially when there are alternative competing technologies requiring different components, production systems and complementary products. Such a situation leads to uncertainty, which hampers market formation, an important building block for large-scale diffusion. Competition also influences the relative price and performance, shaping the investment in production and complementary products and finally influencing customer adoption.
		Macroeconomic and Strategic Aspects	The macro-economic situation can largely influence the conditions for the TIS formation. While economic growth facilitates the formation of TIS with funds for incubators, for example. An economic recession can hamper the large-scale diffusion of a breakthrough innovation given the reduction of these incentives to innovation. The strategies and country policies are often affected by macroeconomic conditions, such

7.2. Appendix 2: Descriptions of the influencing factors

Influencing Factors	Description		
Knowledge and Awareness of Technology	This aspect refers to both fundamental and applied knowledge. They comprise knowledge about the product itself, its production system, and complementary products, as well as the knowledge about its application, which is necessary to develop, produce, maintain, and use the product. Actors in the TIS need to be aware of both types of knowledge, which can be developed through research and practical experimentation. The process of learning about technological innovation, mainly through education, is vital to increase both types of knowledge.		
Knowledge and Awareness of Application and Market	The price of a product involves not only its m This influencing factor relates to the use and applications of innovative products, as well as the market structure and its relevant		

	as the market structure, making both macroeconomic and strategic aspects not only dependent on each other but also of large influence in the TIS building blocks.
Socio-Cultural Aspects	The socio-cultural aspects are the norms and values held by the potential customers and stakeholders of innovative technologies. These norms are not necessarily laws, policies, or regulations, but informal common norms that impact the behavior of TIS actors, such as safety concerns. These rules change with time, and they can influence technology adoption in both ways, encouraging or blocking the technology large-scale adoption.
Accidents and Events	Events such as wars, or natural disasters, together with national or international accidents generate a great impact on the formation of TIS-es. Some of these events stimulate the formation of TIS for radically new technologies because they generate new needs, which have to be fulfilled with innovative technologies.