



Universiteit
Leiden
The Netherlands

Discovering emergent connections in quantum physics research via dynamic word embeddings

Frohnert, F.; Gu, X.; Krenn, M.; Nieuwenburg, E.P.L. van

Citation

Frohnert, F., Gu, X., Krenn, M., & Nieuwenburg, E. P. L. van. (2025). Discovering emergent connections in quantum physics research via dynamic word embeddings. *Machine Learning: Science And Technology*, 6(1). doi:10.1088/2632-2153/adb00a

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/4303538>

Note: To cite this publication please use the final published version (if applicable).



PAPER • OPEN ACCESS

Discovering emergent connections in quantum physics research via dynamic word embeddings

To cite this article: Felix Frohnert *et al* 2025 *Mach. Learn.: Sci. Technol.* **6** 015029

View the [article online](#) for updates and enhancements.

You may also like

- [Forecasting high-impact research topics via machine learning on evolving knowledge graphs](#)
Xuemei Gu and Mario Krenn
- [Accelerating scientific discovery with generative knowledge extraction, graph-based representation, and multimodal intelligent graph reasoning](#)
Markus J Buehler
- [Learning Gaussian Hierarchy Embedding for Relation Prediction in Knowledge Graph](#)
Ruobai Huxue, Jie Huang, Jiantao Wang et al.



PAPER

OPEN ACCESS

RECEIVED

11 November 2024

REVISED

26 January 2025

ACCEPTED FOR PUBLICATION

29 January 2025

PUBLISHED

7 February 2025

Original Content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Discovering emergent connections in quantum physics research via dynamic word embeddings

Felix Frohnert^{1,*} , Xuemei Gu² , Mario Krenn² and Evert van Nieuwenburg¹¹ $\langle a Q a^L \rangle$ Applied Quantum Algorithms, Universiteit Leiden, Leiden, The Netherlands² Max Planck Institute for the Science of Light, Erlangen, Germany

* Author to whom any correspondence should be addressed.

E-mail: f.frohnert@liacs.leidenuniv.nl**Keywords:** dynamic word embedding, quantum physics, forecasting research trajectoriesSupplementary material for this article is available [online](#)

Abstract

As the field of quantum physics evolves, researchers naturally form subgroups focusing on specialized problems. While this encourages in-depth exploration, it can limit the exchange of ideas across structurally similar problems in different subfields. To encourage cross-talk among these different specialized areas, data-driven approaches using machine learning have recently shown promise to uncover meaningful connections between research concepts, promoting cross-disciplinary innovation. Current state-of-the-art approaches represent concepts using knowledge graphs and frame the task as a link prediction problem, where connections between concepts are explicitly modeled. In this work, we introduce a novel approach based on dynamic word embeddings for concept combination prediction. Unlike knowledge graphs, our method captures implicit relationships between concepts, can be learned in a fully unsupervised manner, and encodes a broader spectrum of information. We demonstrate that this representation enables accurate predictions about the co-occurrence of concepts within research abstracts over time. To validate the effectiveness of our approach, we provide a comprehensive benchmark against existing methods and offer insights into the interpretability of these embeddings, particularly in the context of quantum physics research. Our findings suggest that this representation offers a more flexible and informative way of modeling conceptual relationships in scientific literature.

1. Introduction

The corpus of scientific literature is expanding at an ever-increasing rate. In particular, the field of quantum physics is witnessing a steady growth in published papers each year, growing at an increasing pace³. Thus, it is becoming increasingly challenging for individual researchers to gain a comprehensive understanding of the diverse domains within quantum physics. Naturally, this leads to the formation of various specialized subgroups within the community [1]. While specialization enables researchers to focus on specific areas, it also creates isolated knowledge silos where valuable insights may not be shared across different fields. This compartmentalization can result in parallel research trajectories, where solutions to structurally similar problems could remain confined within particular subfields. Overcoming this isolation can help researchers from leveraging advancements in adjacent fields that could accelerate progress in their own field [2–5].

Recent efforts to bridge the knowledge gaps between specialized subfields have led some researchers to leverage large language models (LLMs) trained on vast amounts of scientific literature. By feeding millions of papers into these models, the aim is to generate or rank novel research ideas directly [6–9]. While this approach holds significant promise, it poses substantial challenges in evaluation—it inherently requires human expertise to assess the quality and feasibility of the generated ideas.

³ Judged by the yearly number of *quant-ph* papers submitted to the arXiv, visualized in Appx.A1.

An alternative strategy involves forecasting the future trajectory of scientific research by predicting what scientists might work on next or identifying areas poised for high impact [10]. A common way to solve meta-scientific forecast tasks is to model the scientific progress as an evolving knowledge graph constructed from millions of publications, closely aligning with link prediction problems in network theory [11, 12]. Pioneered in the context of biochemistry, it has been demonstrated that knowledge graphs can be used to find more efficient research strategies to accelerate collective scientific discovery [13]. This idea has been expanded to the forecast of research directions in quantum physics [14] and artificial intelligence [15], the prediction of future impact of new scientific connections [16] and to the predictions of interesting research directions which human researchers might not discover [8]. However, representing scientific literature as a knowledge graph reduces the information down to binary connections between concepts, which may not fully capture the intricate dynamics of scientific advancement.

One way to address this issue is by using unsupervised word embeddings derived *directly* from scientific literature, analyzing how these embeddings evolve to predict future research trends and dynamics. This approach fundamentally differs from knowledge graphs by capturing information from scientific literature without human bias, representing words solely based on contextual similarities and mapping semantically related terms to closely related points in the embedding. Prominent examples involve the automated encoding of scientific literature in material science. Here, when the names and properties of materials are encoded, the word embedding can be used to find new materials with specific properties [17–19]. Another work demonstrated how automated embeddings can relate the surprise and impact of research ideas [20, 21].

In our work, we demonstrate how an unsupervised dynamic word embedding can efficiently predict future research directions in quantum physics⁴. Without relying on any hand-crafted features, a machine learning model that uses only the coordinates of the embedded latent space can predict the likelihood that previously unstudied research combinations will be explored in the future. Our method surpasses other approaches that do not rely on human-crafted features, showcasing a pathway for fully end-to-end prediction tasks within the science of science.

The rest of the paper is structured as follows: section 2 discusses the methodology, including the embedding generation process and the machine learning pipeline. Section 3 presents the benchmarking results of the trained ML models and the interpretation of the learned embedding. Finally, section 4 discusses the findings and provides an outlook for future research.

2. Methodology

2.1. Dataset

The text corpus \mathcal{D} consists of 66 839 abstracts from arXiv’s *quant-ph* category, covering the period ΔT from January 1994 to December 2023. The vocabulary of all words in the corpus is denoted as V . We define the set of quantum physics concepts C as multi-word tokens, each comprising one or more words $v \in V$ that represent distinct ideas or topics within the field. To identify these concepts, we process the corpus by matching words v to a predefined concept list as outlined in Appx.A. When a set of words matches a known quantum physics concept, it is replaced in the corpus with a single token, and the year of its appearance is recorded. This approach results in a final list of 10 235 distinct quantum physics concepts and a modified text corpus $\tilde{\mathcal{D}}$ with an updated vocabulary \tilde{V} containing the replaced tokens, where $C \subseteq \tilde{V}$. Figure 1(a) illustrates how an abstract may contain specific quantum physics concepts, such as *entanglement witnesses*, which are tracked as unique tokens.

2.2. Dynamic embedding model

The goal of the dynamic embedding model is to obtain a vector $w_{c,t} \in \mathbb{R}^n$ for each token $c \in C$ at each time step $t \in \Delta T$, encoding the semantic relationships between concepts in quantum physics literature and how these relationships change over time. To achieve this, we use the Word2Vec model with the Skip-gram approach [22–24]. Word2Vec is a neural network-based model that learns continuous vector representations of tokens (embeddings) from the text corpus $\tilde{\mathcal{D}}$. The Skip-gram model is designed to predict the context tokens surrounding a given target word within a window size k , learning embeddings that reflect semantic similarities based on context. Conceptually, Skip-gram is a two-layer neural network trained to maximize the conditional probability $P(w_{i+j,t}|w_{i,t})$ of observing context token $w_{i+j,t}$ given a target word $w_{i,t}$ [22]:

$$\mathcal{L}_{\text{skip}} = - \sum_{i \in \tilde{\mathcal{D}}} \sum_{\substack{-k \leq j \leq k \\ j \neq 0}} \log P(w_{i+j,t}|w_{i,t}). \quad (1)$$

⁴ The focus on quantum physics reflects our domain expertise, enabling rigorous interpretation of concepts and model results, but the methodology is broadly applicable.

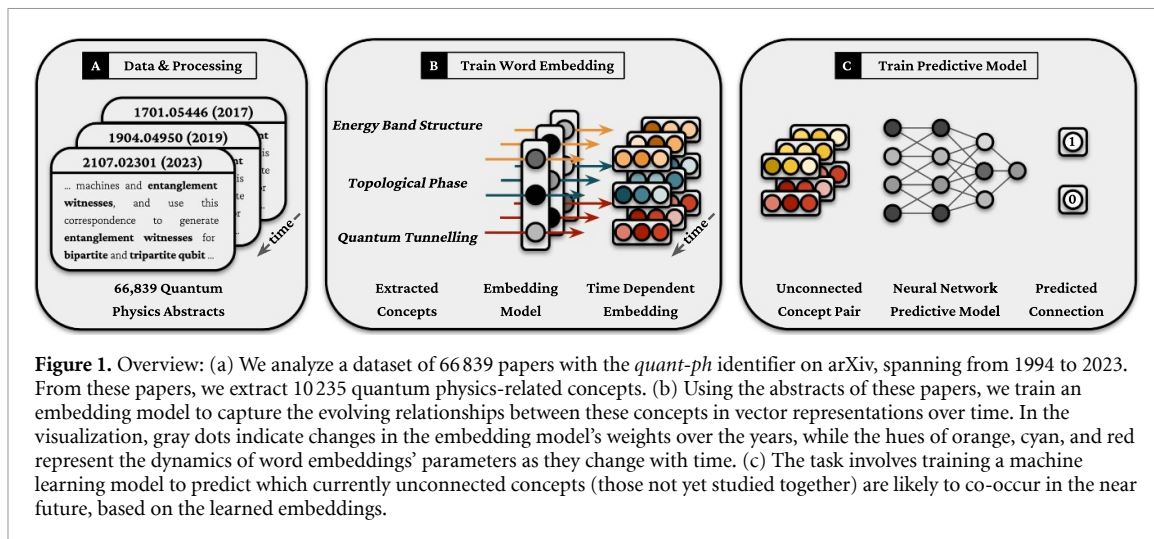


Figure 1. Overview: (a) We analyze a dataset of 66 839 papers with the *quant-ph* identifier on arXiv, spanning from 1994 to 2023. From these papers, we extract 10 235 quantum physics-related concepts. (b) Using the abstracts of these papers, we train an embedding model to capture the evolving relationships between these concepts in vector representations over time. In the visualization, gray dots indicate changes in the embedding model’s weights over the years, while the hues of orange, cyan, and red represent the dynamics of word embeddings’ parameters as they change with time. (c) The task involves training a machine learning model to predict which currently unconnected concepts (those not yet studied together) are likely to co-occur in the near future, based on the learned embeddings.

The architecture consists of an input layer, a projection (hidden) layer, and an output layer. To generate dynamically evolving embeddings, we train the model sequentially for each time interval in ΔT with step size δt , initializing each step with the trained model from the time interval, as illustrated in figure 1(b). This approach has proven to reveal semantic shifts over time [25]. This process generates embeddings for the set of concepts C from their first appearance in the text corpus up to the cutoff date. Since concepts emerge at different times in the corpus (see Appx.A1. inset), we use backtracking to maintain a consistent set of embeddings across all time intervals. For concepts that first appear after the corpus’s starting interval, we fill in the missing embeddings by using the embedding of their initial occurrence. This approach ensures that all concepts are trackable from the same initial time interval, providing a consistent basis for analysis over time.

2.3. Predictive model

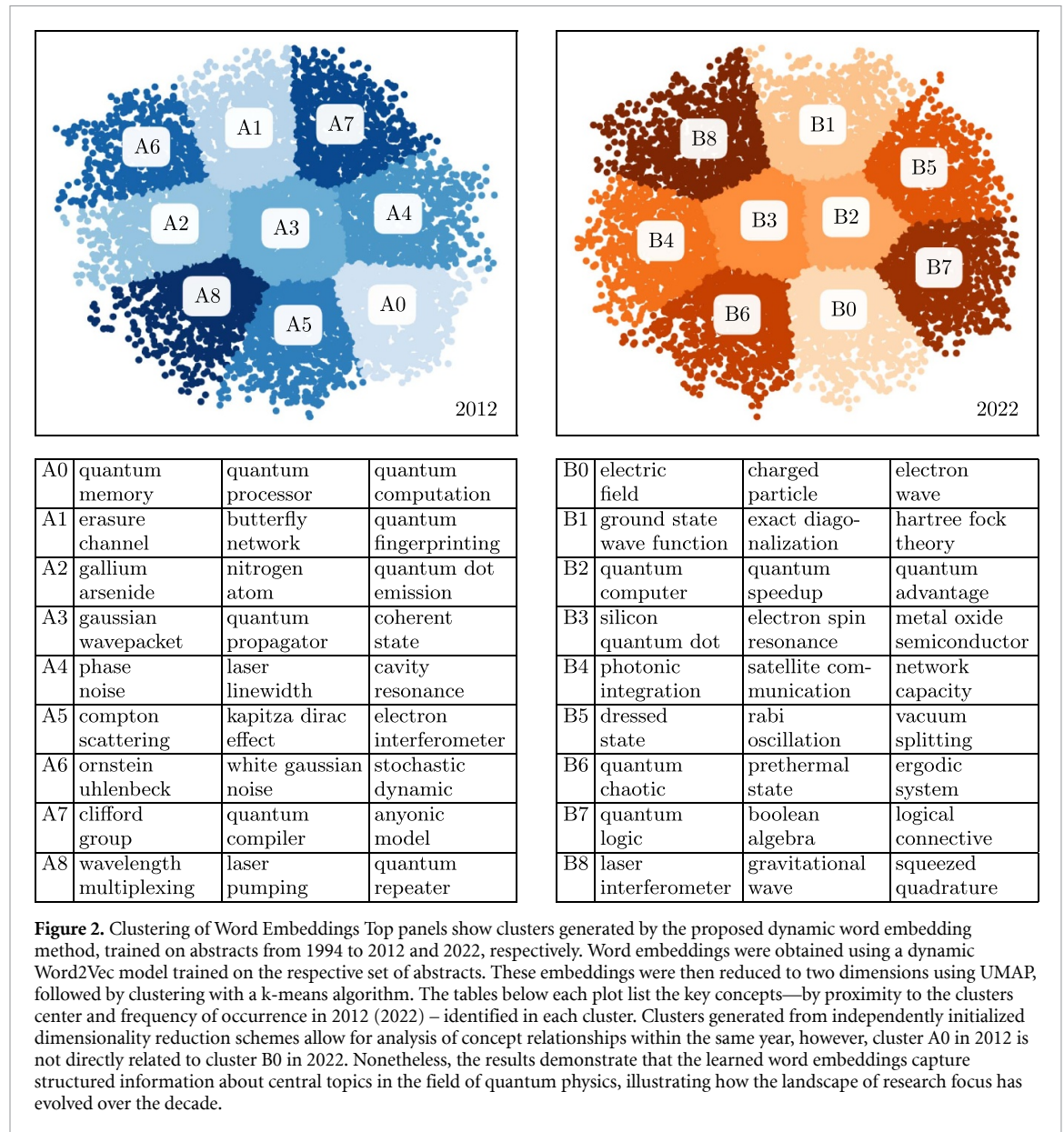
Given that the embedding features capture detailed information about relationships between various concepts and how these relationships evolve over time, we investigate training a machine learning model to predict whether concept pairs will eventually become connected based on their embedding data. Our approach aims to predict if pairs of concepts, which have not yet appeared together in the same abstract, will co-occur in future research. In our classification task, illustrated in figure 1(c), we categorize quantum physics concept pairs into two groups. For both groups, the concept pairs have not co-occurred in the same abstract during a training window Δt . We assign the label $y = 1$ if the concepts co-occur in the same abstract within a subsequent label window Λt , and $y = 0$ if they remain unconnected during this period. To perform this classification, we use a neural network-based classifier. The network is trained to differentiate between the two categories based on embeddings in the dataset \mathcal{X} , minimizing the binary cross-entropy loss [26] between true labels y and predicted labels \hat{y} :

$$\mathcal{L}_{\text{BCE}} = -\frac{1}{|\mathcal{X}|} \sum_{(x,y) \in \mathcal{X}} [y \log(\hat{y}) + (1-y) \log(1-\hat{y})]. \quad (2)$$

3. Results

3.1. Analyzing the dynamic embedding

We begin by analyzing the quality of the proposed dynamic embedding method to assess its effectiveness in capturing relationships among quantum physics concepts. To do this, we train the dynamic embedding model, as described in section 2.2, over two time frames: $\Delta t = [1994, 2012]$ and $\Delta t = [1994, 2022]$ with an embedding size of $n = 128$. To explore the distribution of the embeddings, we examine those from the final years, 2012 and 2022, each informed by data dating back to 1994. We apply Uniform Manifold Approximation and Projection (UMAP) [27] to reduce the 128-dimensional embeddings to two dimensions, allowing us to observe relationships among quantum physics concepts in this reduced space. In this, UMAP aims to preserve a notion of both local and global structures from the original feature space [27]. Additionally, we use k-means clustering [28] to identify nine distinct clusters within this reduced space, helping us analyze neighborhoods of related quantum physics concepts. Both UMAP and k-means hyperparameters were optimized empirically for the clearest visualization. The results of this clustering analysis are shown in figure 2, providing a visual representation of concept relationships and groupings based



on their embeddings. Below the figure, we include tables listing key concepts in each cluster, selected by proximity to the cluster center and their frequency in abstracts from 2012 (2022).

To evaluate embedding quality, we interpret the structural patterns within the clusters. Each cluster contains thematically related quantum physics concepts; for instance, cluster A0 in 2012 focuses on quantum computing, while A6 contains concepts about stochastic noise. Concept distances within clusters are also meaningful. For example, clusters A5 and A8, which are close, represent concepts related to experimental quantum physics, whereas A5 and A7, which are farther apart, include distinct topics. This clustering shows that the embeddings capture structured information about central quantum physics topics. To illustrate how research focus has shifted over time, we compare the clusters from 2012 to those from 2022. We observe emerging topics, such as gravitational waves and quantum advantage. While these clusters allow us to interpret the structure of the embeddings within a single year, the independent initialization of dimensionality reduction and evolving embeddings over time means that clusters from different years (e.g. A0 and B0) are not directly comparable.

The observed structure suggests that the embeddings encode meaningful information about relationships among quantum physics concepts and their evolution over time.

3.2. Comparing the performance of embedding methods

The primary objective of this study is to evaluate the effectiveness of the proposed dynamic embedding technique for classifying quantum physics concept pairs and to compare its performance against existing

Table 1. Comparing Performance of Embedding Techniques: Area under the curve (AUC) values for six embedding methods, comparing word embeddings and traditional knowledge graph approaches. Abbreviations are: the proposed dynamic word embeddings (Word-Dynamic), static word embeddings (Word-Static), static word embeddings with hand crafted features (Word-Hand), Knowledge Graph with hand-crafted features (Knowledge-Hand), Knowledge Graph with machine-learned (ProNE) features (Knowledge-Prone), and Knowledge Graph with machine-learned (Node2Vec) features (Knowledge-Node), Models are trained to predict concept co-occurrence in the period $\Delta t = [2018, 2020]$ based on embeddings from $\Delta t = [1994, 2017]$, the displayed performance is evaluated with a 3 year time shift.

Model	AUC	Reference	Knowledge graph	Dynamic embedding	Feature type
Word-Dynamic	0.88	Ours	No	Yes	Automatic
Word-Static	0.79	[23]	No	No	Automatic
Word-Hand	0.57	[17]	No	No	Handcrafted
Knowledge-Hand	0.82	[14]	Yes	No	Handcrafted
Knowledge-ProNE	0.79	[15]M7A	Yes	No	Automatic
Knowledge-Node	0.78	[15]M7B	Yes	No	Automatic

methods. All results are presented in table 1, where the performance of each model is evaluated using the area under the curve (AUC) metric [29]. The AUC measures the area under the receiver operating characteristic (ROC) curve, which plots the true positive rate (sensitivity) against the false positive rate across varying classification thresholds. Operationally, the AUC can be interpreted as the probability that, given a randomly selected positive and negative sample, the positive sample will be ranked higher than the negative one. This ranking is crucial for better classification performance, as it indicates that the model assigns higher confidence scores to true positive samples than to false positives. This property allows for more effective separation between classes across various threshold settings. AUC values closer to 1 signify strong discriminative ability, while a value of 0.5 represents random guessing. For the task at hand, a high AUC reflects the model's ability to distinguish between quantum physics concept pairs that are likely to co-occur and those that are not. The underlying ROC curves are presented in Appx.A2.

To ensure an equal comparison, all embedding models—whether utilizing our proposed embeddings or baseline methods—were trained and evaluated using the same predictive (neural network) architecture, as outlined in Appx.D. The only variable across experiments was the input representation, with the number of input neurons reflecting the dimensionality of the respective embeddings.

We trained and validated the models using embeddings generated from abstracts spanning the years $\Delta t = [1994, 2017]$, with label windows corresponding to the years $\Delta t = [2018, 2020]$. For testing, we utilized embeddings up to $\Delta t = [1994, 2020]$, validating predictions against the label window covering $\Delta t = [2021, 2023]$.

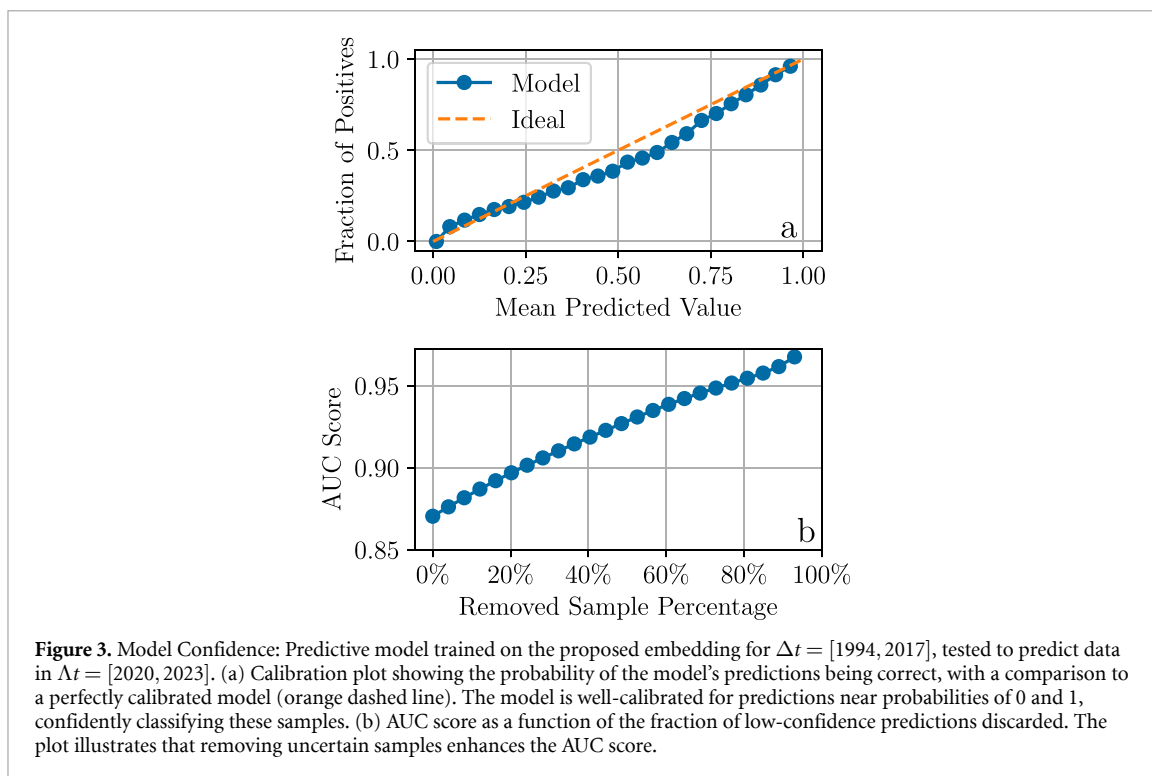
The results demonstrate the proposed dynamic embedding method consistently surpassing the baseline approaches. Since both features were derived from the same underlying dataset and the predictive models were trained with identical structures, the performance disparity can be attributed solely to the input embeddings.

We conclude that the model's capacity to utilize the temporal structure of these embeddings enriches the contextual information, leading to more accurate and robust predictions. In the following, we briefly outline the distinctions between each embedding technique, with a more detailed description provided in Appx.B.

(Ours) Dynamic word embedding (Word-Dynamic) – Our proposed dynamic embedding approach represents each concept as a vector $w_{c,t} \in \mathbb{R}^n$ for each token $c \in C$ at each time step d with $n = 128$. The time intervals in Δt are set to $\delta t = 1$ month. It captures the semantic relationships between concepts in quantum physics literature and how these relationships evolve over time. While this technique does not explicitly encode the connections between concepts, it provides the predictive model with comprehensive information about the structure within the embedding space. A study on the robustness of the embedding under varying hyperparameters is provided in Appx.E.

Static word embedding (Word-Static) – The static word embedding approach can be viewed as the standard Word2Vec method for generating embeddings $w_c \in \mathbb{R}^n$ with $n = 128$. Like the dynamic embedding, the training is performed in an unsupervised manner; however, it uses a single corpus of abstracts from a specified time frame, Δt . Thus, it does not incorporate the dynamic training introduced in section 2.2. The resulting structure encodes the relationships between concepts implicitly within an embedding space, but does not account for temporal changes in these relationships.

Static word embedding and handcrafted feature (Word-Hand) – This method utilizes the same static word embeddings but derives a handcrafted feature: the cosine similarity between word embedding vectors. In [17], this metric is used to predict potential undiscovered connections in material science, effectively



identifying novel relationships by assessing the proximity of concepts in the embedding space. While this method also implicitly encodes relationships between concepts, it simplifies the information into a single handcrafted feature.

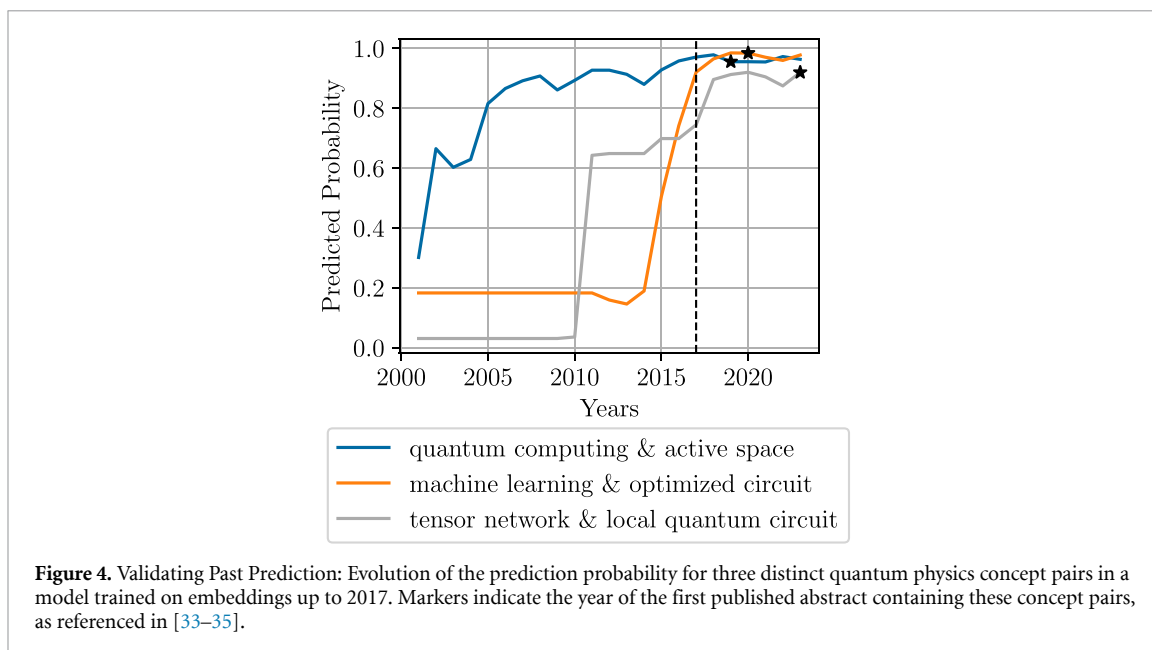
Knowledge graph and handcrafted feature (Knowledge-Hand) – This method does not use word embeddings but instead creates an evolving knowledge graph from the abstracts. In this graph, concepts are represented as nodes, and connections between concepts—i.e. if they co-occur in the same abstract—are modeled as edges. In [14], 15 handcrafted features are extracted from the knowledge graph to predict the connection between concepts. The approach formulates the task as a link prediction problem, training a machine learning model to classify whether two nodes, representing distinct concepts, will form a link within a specified timeframe. In general, the knowledge graph representation abstracts away all information aside from the connections, focusing explicitly on modeling the relationships between concepts.

Knowledge graph and machine learning feature (Knowledge-Node) – Instead of selecting handcrafted features from knowledge graphs for the classification task, machine learning-based features can also be used [15]. This model is based on Node2Vec embeddings [30], which eliminates the need for handcrafted features in link prediction within knowledge graphs.

Knowledge graph and machine learning feature (Knowledge-ProNE) – This model also uses machine-learned embeddings from the knowledge graph, in this case the ProNE method [31]. Thus, it also incorporates these automated embeddings into a neural network framework for link prediction between concepts, without the need for explicit feature engineering.

3.3. Analyzing the model calibration

In the previous section, we trained the predictive model with a logistic output by minimizing cross-entropy loss with binary labels, as introduced in equation (2), allowing the output to be interpreted as the probability that two quantum physics concepts will be linked in the future. This probability is generally well-calibrated but shows increased uncertainty around predicted probabilities near 0.5, as illustrated in figure 3(a). For example, among samples with an output value of 0.9, approximately 90% have a true label of 1. Predictions closer to 0 or 1 indicate higher confidence, while probabilities near 0.5 signal greater uncertainty and an increased likelihood of misclassification. Therefore, this probabilistic output thus enables selection of test samples, allowing uncertain samples to be filtered out. In figure 3(b), we see that removing low-confidence predictions, such as the 20% most uncertain pairs, increases the AUC score beyond 0.9 – a threshold indicating strong predictive performance [32]. For studying interesting concept combinations, discarding a substantial portion of samples is not inherently problematic. Focusing on a small subset of high-confidence



predictions makes investigating samples that warrant further analysis more feasible. This approach makes confidence-based selection an effective strategy for identifying this promising subset.

3.4. Analyzing the model predictions

An added benefit of dynamic word embeddings paired with a well-calibrated predictive model is their ability to capture and interpret the evolution of research concepts over time. Thus, as a final step, we evaluate whether our model—trained on quantum physics abstracts from prior years—could predict concept combinations that would later emerge in the literature.

In figure 4, we visualize the predicted probability—shown to be well-calibrated for confident cases in the previous section—of a model trained on dynamic embeddings from $\Delta t = [1994, 2017]$ and then evaluated on the embeddings for each year through 2023. We illustrate three pairs of concepts, showing how the model’s predictions evolved over time. The star marker in each line indicates the year when the concept pair first appeared together in a published abstract. The selected pairs highlight recent popular combinations in the field: (1) The predicted link between *quantum computing* and *active space* methods first appears in reference [33] given our text corpus and concept list. (2) The predicted connection of using *machine learning* for constructing *optimized circuits* for noise-resilience first appears in reference [34]. (3) The technique of using *tensor network* approaches in investigating *local quantum circuits* first appears in reference [35]. Importantly, the concept set C is highly granular and does not yet account for synonyms; we discuss the implications of this limitation in section 4. Nonetheless, this visualization shows that the model’s increasing probability predictions align well with the first recorded uses of these concept pairs in the corpus. While the pairs were chosen based on rising topics of interest, we note that high predicted probabilities also corresponded to these papers having (1) 88, (2) 118, and (3) 9 citations, respectively, as of this manuscript’s writing.

4. Discussion and outlook

In this work, we demonstrated the application of dynamic word embeddings for predicting research trends using neural networks, with a focus on quantum physics concepts. Our results show that features generated from an unsupervised embedding model—without the need for direct human input—show promising performance, leading to improved prediction accuracy compared to the baseline embeddings. Thus, the presented embedding method offers an interesting research direction for leveraging machine learning to investigate how experts across diverse fields approach or solve structurally similar problems.

Benchmark additional dynamic embedding approaches – For this initial study, we adopted the dynamic embedding method from reference [25]. A natural extension of this work would be to compare the effectiveness of various other dynamic embedding techniques for the specific prediction task at hand [36, 37].

Benchmark LLM embeddings – An interesting extension of our work, left for future studies, is the exploration of modern embedding models like LLaMA [38], BERT [39], or GPT [40]. These models generate

contextual embeddings that vary depending on the surrounding text, enhancing their representational power. This contextual representation has led to improvements in performance across various other natural language processing tasks. Importantly, these models are typically trained on large text corpora, presenting a key challenge: determining distinct pretraining cutoff dates in the past to effectively train and evaluate the predictive model, as discussed in section 2.3.

Reinforcement learning (RL) for discovering emergent connections – RL and causal modeling have shown promise in uncovering complex relationships by combining causal insights with adaptive learning [41]. Exploring the integration of RL into our framework could enable adaptive discovery of evolving conceptual relationships, offering a compelling direction for future work.

Extended model selection – A valuable future research direction involves comprehensive model selection and optimization, customizing the neural network architecture to suit the specific characteristics of the dynamic embedding features. This approach could further improve predictive performance and underscore the potential of dynamic embeddings. We focused on using the embedding from the final interval of the training window, as empirical investigations showed that incorporating multiple time slices did not significantly impact the final AUC score. A future study could explore whether a deep recurrent neural network architecture can leverage information from multiple time slices for enhanced predictive performance.

Application for quantum physics research – The ultimate goal of the proposed method is to forecast the future trajectory of scientific research by predicting potential directions that scientists might explore. Our method represents a step towards achieving high-quality predictions in this domain. In figure 4, we demonstrate that our method assigns high probabilities to fruitful concept combinations.

The predictions depend on the specific concepts extracted from each abstract, especially for very prominent concepts that have many synonyms (different phrases with the same semantic meaning). An interesting next step would be hierarchical grouping of related concepts (synonyms), meaning that although terms like *active space* and *quantum algorithm* may have co-occurred before, the model only flagged the exact combination of *quantum computing* and *active space* in the displayed year. This would reduce the granularity of the set of concepts but could enhance the model's ability to track when distinct concept areas begin to intersect.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://github.com/FelixFrohnertDB/axiv_nlp. The arXiv abstract corpus used in this study is publicly accessible through Kaggle.

Acknowledgment

The authors would like to thank Simon C. Marshall, Stefano Polla, and Patrick Emonts for useful discussions. This work was supported by the Dutch National Growth Fund (NGF), as part of the Quantum Delta NL program.

ORCID iDs

Felix Frohnert  <https://orcid.org/0000-0003-3717-6352>

Xuemei Gu  <https://orcid.org/0000-0002-0734-432X>

Mario Krenn  <https://orcid.org/0000-0003-1620-9207>

References

- [1] Jacobs J A 2014 *In Defense of Disciplines: Interdisciplinarity and Specialization in the Research University* (University of Chicago Press) (<https://doi.org/10.1086/682163>)
- [2] Evans J A and Foster J G 2011 Metaknowledge *Science* **331** 721–5
- [3] Fortunato S et al 2018 Science of science *Science* **359** 6379
- [4] Wang D and Barabási A-L 2021 *The Science of Science* (Cambridge University Press) (<https://doi.org/10.1017/9781108610834>)
- [5] Bonaventura M, Latora V, Nicosia V and Panzarasa P 2017 The advantages of interdisciplinarity in modern science (arXiv:1712.07910)
- [6] Wang Q, Downey D, Ji H and Hope T 2023 Scimon: scientific inspiration machines optimized for novelty (arXiv:2305.14259)
- [7] Baek J, Jauhar S K, Cucerzan S and Hwang S J 2024 Researchagent: iterative research idea generation over scientific literature with large language models (arXiv:2404.07738)
- [8] Gu X and Krenn M 2024 Generation and human-expert evaluation of interesting research ideas using knowledge graphs and large language models (arXiv:2405.17044)

- [9] Si C, Yang D and Hashimoto T 2024 Can llms generate novel research ideas? a large-scale human study with 100+ nlp researchers (arXiv:2409.04109)
- [10] Clauset A, Larremore D B and Sinatra R 2017 Data-driven predictions in the science of science *Science* **355** 477–80
- [11] Liben-Nowell D and Kleinberg J 2003 The link prediction problem for social networks *12th Int. Conf. on Information and Knowledge Management* (<https://doi.org/10.1145/956863.956972>)
- [12] Martínez V, Berzal F and Cubero J-C 2016 A survey of link prediction in complex networks *ACM Comput. Surv. (CSUR)* **49** 1–33
- [13] Rzhetsky A, Foster J G, Foster I T and Evans J A 2015 Choosing experiments to accelerate collective discovery *Proc. Natl Acad. Sci.* **112** 14569–74
- [14] Krenn M and Zeilinger A 2020 Predicting research trends with semantic and neural networks with an application in quantum physics *Proc. Natl Acad. Sci.* **117** 1910–6
- [15] Krenn M et al 2023 Forecasting the future of artificial intelligence with machine learning-based link prediction in an exponentially growing knowledge network *Nat. Mach. Intell.* **5** 1326–35
- [16] Gu X and Krenn M 2024 Forecasting high-impact research topics via machine learning on evolving knowledge graphs (arXiv:2402.08640)
- [17] Tshitoyan V, Dagdelen J, Weston L, Dunn A, Rong Z, Kononova O, Persson K A, Ceder G and Jain A 2019 Unsupervised word embeddings capture latent knowledge from materials science literature *Nature* **571** 95–98
- [18] Kim E et al 2020 Inorganic materials synthesis planning with literature-trained neural networks *J. Chem. Inf. Model.* **60** 1194
- [19] Shetty P and Ramprasad R 2021 Automated knowledge extraction from polymer literature using natural language processing *iScience* **24** 101922
- [20] Shi F and Evans J 2023 Surprising combinations of research contents and contexts are related to impact and emerge with scientific outsiders from distant disciplines *Nat. Commun.* **14** 1641
- [21] Sourati J and Evans J A 2023 Accelerating science with human-aware artificial intelligence *Nat. Human Behav.* **7** 1682–96
- [22] Mikolov T 2013 Efficient estimation of word representations in vector space (arXiv:1301.3781)
- [23] Mikolov T, Sutskever I, Chen K, Corrado G S and Dean J 2013 Distributed representations of words and phrases and their compositionality *Advances in Neural Information Processing Systems* vol 26
- [24] Sivakumar S, Videla L S, Kumar T R, Nagaraj J, Itnal S and Haritha D 2020 Review on word2vec word embedding neural net *IEEE Int. Conf. on Smart Electronics and Communication* (<https://doi.org/10.1109/ICOSEC49089.2020.9215319>)
- [25] Kim Y, Chiu Y I, Hanaki K, Hegde D and Petrov S 2014 Temporal analysis of language through neural language models (arXiv:1405.3515)
- [26] Ramos D, Franco-Pedroso J, Lozano-Diez A and Gonzalez-Rodriguez J 2018 Deconstructing cross-entropy for probabilistic binary classifiers *Entropy* **20** 208
- [27] McInnes L, Healy J and Melville J 2018 Umap: uniform manifold approximation and projection for dimension reduction (arXiv:1802.03426)
- [28] Lloyd S 1982 Least squares quantization in pcm *IEEE Trans. Inf. Theory* **28** 129–37
- [29] Fawcett T 2004 ROC graphs: notes and practical considerations for researchers *Machine Learning* vol 31 pp 1–38 (available at: https://www.researchgate.net/publication/284043217_ROC_Graphs_Notes_and_Practical_Considerations_for_Researchers)
- [30] Grover A and Leskovec J 2016 node2vec: scalable feature learning for networks *Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining* (<https://doi.org/10.48550/arXiv.1607.00653>)
- [31] Zhang J, Dong Y, Wang Y, Tang J and Ding M 2019 Prone: fast and scalable network representation learning *IJCAI* **19** 4278–84
- [32] Hosmer Jr D W, Lemeshow S and Sturdivant R X 2013 *Applied Logistic Regression* (Wiley) (<https://doi.org/10.1002/9781118548387>)
- [33] Bauman N P, Bylaska E J, Krishnamoorthy S, Low G H, Wiebe N, Granade C E, Roetteler M, Troyer M and Kowalski K 2019 Downfolding of many-body hamiltonians using active-space models: extension of the sub-system embedding sub-algebras approach to unitary coupled cluster formalisms *J. Chem. Phys.* **151** 014107
- [34] Cincio L, Rudinger K, Sarovar M and Coles P J 2021 Machine learning of noise-resilient quantum circuits *PRX Quantum* **2** 010324
- [35] Liu Y J, Shtengel K and Pollmann F 2024 Simulating 2D topological quantum phase transitions on a digital quantum computer (arXiv:2312.05079)
- [36] Bamler R and Mandt S 2017 Dynamic word embeddings *PMLR* (available at: <https://proceedings.mlr.press/v70/bamler17a.html>)
- [37] Horn F 2021 Exploring word usage change with continuously evolving embeddings *Proc. 59th Annual Meeting of the Association for Computational Linguistics and the 11th Int. Joint Conf. on Natural Language Processing: System Demonstrations* (<https://doi.org/10.18653/v1/2021.acl-demo.35>)
- [38] Touvron H et al 2023 Llama: open and efficient foundation language models (arXiv:2302.13971)
- [39] Devlin J, Chang M W, Lee K and Toutanova K 2019 Bert: pre-training of deep bidirectional transformers for language understanding (arXiv:1810.04805)
- [40] Brown T B et al 2020 Language models are few-shot learners (arXiv:2005.14165)
- [41] Méndez-Molina A, Morales E F and Sucar L E 2022 Causal discovery and reinforcement learning: a synergistic integration *PMLR Int. Conf. on Probabilistic Graphical Models* (available at: <https://proceedings.mlr.press/v186/mendez-molina22a.html>)