Received: 15 July 2022 | Revised: 21 September 2022 | Accepted: 24 October 2022

DOI: 10.1002/inf2.12397

RESEARCH ARTICLE

WILEY

Capturing 2D van der Waals magnets with high probability for experimental demonstration from materials science literature

Haiyang Song[1] | **Yinghe Zhao**[1] | **Eleanor Turner**[2] | **Yu Wu**[1] | **Yuan Li**[1] | **Menghao Wu**[3] | **Guang Feng**[4] | **Huiqiao Li**[1] | **Tianyou Zhai**[1]

[1]State Key Laboratory of Materials Processing and Die & Mould Technology, School of Materials Science and Engineering, Huazhong University of Science and Technology, Wuhan, the People's Republic of China

[2]Department of Chemical & Biological Engineering, Monash University, Clayton Campus, Clayton, Victoria, Australia

[3]School of Physics, Huazhong University of Science and Technology, Wuhan, the People's Republic of China

[4]State Key Laboratory of Coal Combustion, School of Energy and Power Engineering, Huazhong University of Science and Technology, Wuhan, the People's Republic of China

Correspondence
Yinghe Zhao and Tianyou Zhai, State Key Laboratory of Materials Processing and Die & Mould Technology, School of Materials Science and Engineering, Huazhong University of Science and Technology, Wuhan, Hubei 430074, the Peoples' Republic of China.
Email: zhaoyh@hust.edu.cn and zhaity@hust.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Numbers: U21A2069, 22103026, 21825103; Fundamental Research Funds for the Central Universities, Grant/Award Number: 2020kfyXJJS050

Abstract

2D van der Waals (vdW) magnets have opened intriguing prospects for next-generation spintronic nanodevices. Machine learning techniques and density functional theory calculations enable the discovery of 2D vdW magnets to be accelerated; however, current computational frameworks based on these state-of-the-art approaches cannot offer probability analysis on whether a 2D vdW magnet can be experimentally demonstrated. Herein, a new framework can be established to overcome this challenge. Via the framework, 2D vdW magnets with high probability for experimental demonstration are captured from materials science literature. The key to the successful establishment is the introduction of the theory of mutual information. Historical validation of predictions substantiates the high reliability of the framework. For example, half of the 30 2D vdW magnets discovered in the literature published prior to 2017 have been experimentally demonstrated in the subsequent years. This framework has the potential to become a revolutionary force for progressing experimental discovery of 2D vdW magnets.

KEYWORDS
2D vdW magnets, mutual information, neural networks

1 | INTRODUCTION

2D van der Waals (vdW) magnets can manifest magnetic phenomena in the 2D limit, and they have garnered substantial research interest from the materials community in recent years.[1–15] 2D vdW magnets can not only serve as an ideal platform for probing the fundamental physics of magnetism, but they can also enable the integration of magnetic layers into nanodevices. This application has a revolutionary impact on information technologies.[1–15]

However, there are a few challenging issues on the road to the practical applications of 2D vdW magnets, such as poor ambient stability and low magnetic transition temperature.[16–20] New 2D vdW magnets may provide unprecedented opportunities for tackling these issues and bring fascinating physical phenomena, providing novel platforms for studying the intrinsic mechanisms of magnetism. Machine learning (ML) techniques and density functional theory (DFT) calculations have been widely utilized to accelerate the discovery of 2D vdW magnets.[21–25]

Current computational frameworks based on these state-of-the-art approaches are unable to provide probability analysis on whether a 2D vdW magnet can be demonstrated by experiments. The endeavor for preparation of new 2D vdW materials or experimental measurement of magnetism in 2D materials is often laborious, cumbersome, and time-consuming, which therefore, makes the discovery of 2D vdW magnets driven by traditional trial-and-error experiments extremely inefficient. Thus, a computational framework that can serve to focus experimental resources on high-priority experimental tasks is urgently required. By mining the knowledge underlying scientific literatures, experienced researchers could detect those materials with high probability for experimental demonstration; to acquire this skill, a beginner devotes many years of substantial effort to literature reading. In addition, with the rapid development of materials science, the amount of literature data related to materials science has far surpassed what the human brain can cope with, making it easy to miss a large number of potentially valuable functional materials.

In this work, by mimicking materials scientists' comprehension of intrinsic relations underlying literature data via mutual information (MI), we successfully establish a framework that can capture 2D vdW magnets with high probability for experimental demonstration from a vast body of materials science literature. Workflow of the framework of materials discovery driven by the MI-based knowledge mining is schematically presented in Figure 1, and it can be regarded as a well-designed four-step process: (i) literation collection, (ii) text processing, (iii) the training of an MI-based neural network model, and (iv) the MI-driven output of promising new functional materials. Historical validation of predictions was performed to assess the reliability of this framework. Remarkably, the framework can discover 2D vdW magnets from huge volumes of literature data several years before their experimental discovery, which validates its high reliability.

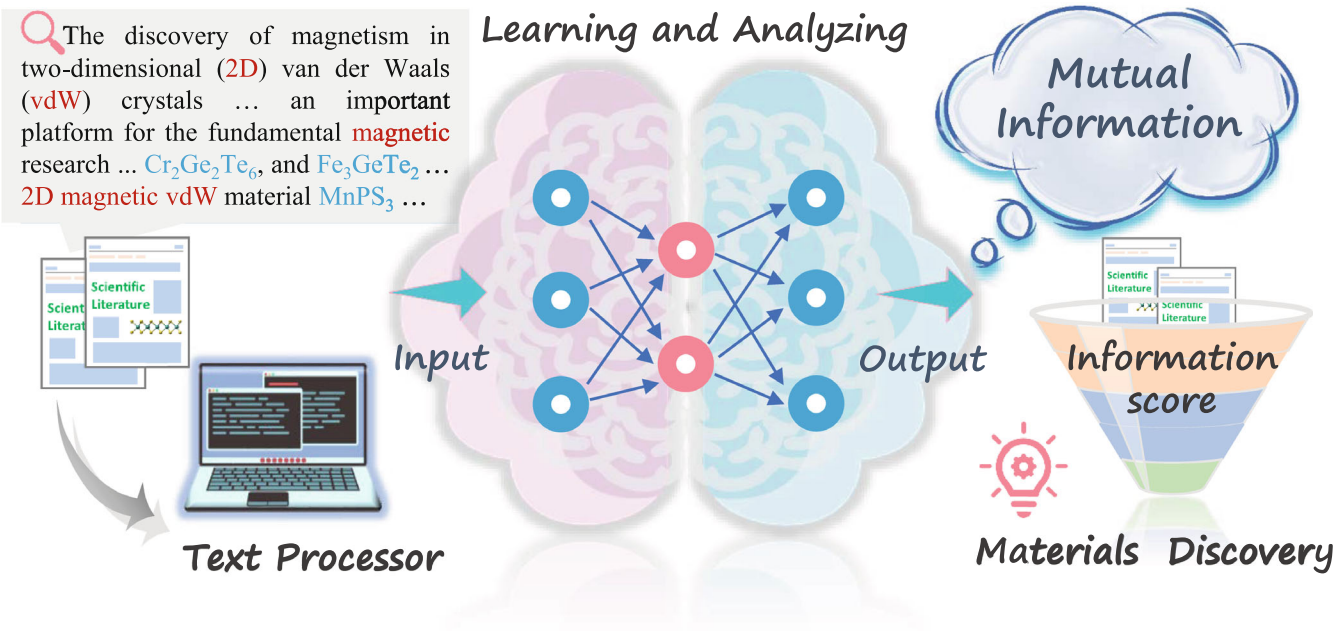

FIGURE 1 Workflow of the framework of MI-driven materials discovery. To begin, the titles and abstracts of millions of publications related to materials science are downloaded in bulk and analyzed (see Section 4.1), and the processed data is then trained using a three-layer neural network (see Section 4.2). Finally, information score (IS) is employed in the search for promising functional materials.

2 | RESULTS AND DISCUSSION

2.1 | Establishment of the framework of MI-driven materials discovery

The rapid development of materials science in the past decades produces vast amounts of relevant information. An important manifestation in this regard is an explosion of academic papers related to materials science. Unfortunately it is restricted by the storage capacity and computing power of human brain. Researchers have long been overwhelmed by the vast and complex information space of materials science. Besides, the latent materials-related information outside human cognition is particularly hard to mine since humans find it challenging to think outside the box. An intractable issue in materials science is how to mine potentially valuable information from vast relevant information. Computers are endowed with the storage capacity and computing power significantly greater than the human brain, but they lack the logical ability to comprehend the information relations underpinning materials science expertise like materials scientists can. Developing an approach capable of rationally characterizing the intrinsic information relations hidden in the vast and complex information space of materials science is key to addressing the above issue.

The theory of MI, which originated from Shannon's information theory,[26] was originally proposed for measuring the relation of signals in communication systems, and later widely used in characterizing the complex information relations in bioinformatics[27–29] and computer science.[30–32] The Supporting Information provides a brief introduction to the mathematical theory of MI and its relation-characterizing capacity. We found that the complex materials–property relations underlying massive literature data can be well quantified by IS, which is derived from the theory of MI. The IS-quantified materials–property relations can help us discover new functional materials.

According to Harris' distribution hypothesis,[33] the relationship between the words w and v can be reflected using another word c. Based on this hypothesis, the model takes the following form:

$$\mathcal{F}\left(\vec{w}, \vec{v}, \vec{c}\right) = \frac{p(c|w)}{p(c|v)}$$

where $\vec{w}, \vec{v}, \vec{c} \in \mathbb{R}^d$, and $\mathcal{F}$ depends on some, as of yet, unspecified parameters. Although the number of possibilities for $\mathcal{F}$ is large, a unique choice can be selected by enforcing a few desiderata, that is,

$$\vec{w}^T \vec{c} = \log \frac{p(w, c)}{p(w)p(c)} = I(w, c)$$

Detailed mathematical derivation is shown in the Supporting Information. The above model is the factorization of a pointwise mutual information (PMI) matrix.

However, the PMI matrix is so huge that we must reduce its dimension. We found that a neural network model called as SGNS (skip-gram with negative sampling)[34] can solve the problem and achieve our goal. Mathematically, the objective function of SGNS is optimized based on PMI (see the Supporting Information for detailed mathematical derivation). In other words, the SGNS-characterized information relation is equivalent to that characterized by PMI. Therefore, the PMI matrix can be solved via SGNS. Then, IS between the words w and v can be computed via the above model weights, and is mathematically expressed as:

$$IS(w, v) = \vec{w}^{\mathrm{T}} \cdot \vec{v}$$

$IS(w, v)$ is determined by the production between the weight vector $\vec{w}^{\mathrm{T}}$ and the vector $\vec{v}$ from the hidden layer. The specific calculation process of IS is schematically represented in Figure 2. All words in the processed materials science literature are stored into a vocabulary table and turned into a one-hot high-dimensional vector for training. The dimensionality of the vector is compressed after training the PMI-based neural network, and the IS between the two words is determined based on the training results. IS quantifies the relationship between two target words underlying massive amounts of literature data. The greater IS means the stronger relationship.

2.2 | MI-driven discovery of 2D vdW magnets from massive literature data

We first focused on exploring the reliability of the established framework in quantifying the materials–property relations, since the reliability is the most important precondition for achieving materials discovery. The reliability is well substantiated by the exploration of the relationships between the subfields of materials science. According to whether a chemical reaction is involved, the subfields of materials science can be divided into two categories: physics-related and chemistry-related subfields. We selected 10 typical subfields related to physics: "FET" (field-effect transistor), "ferroelectric", "magnetic", "photodetector",

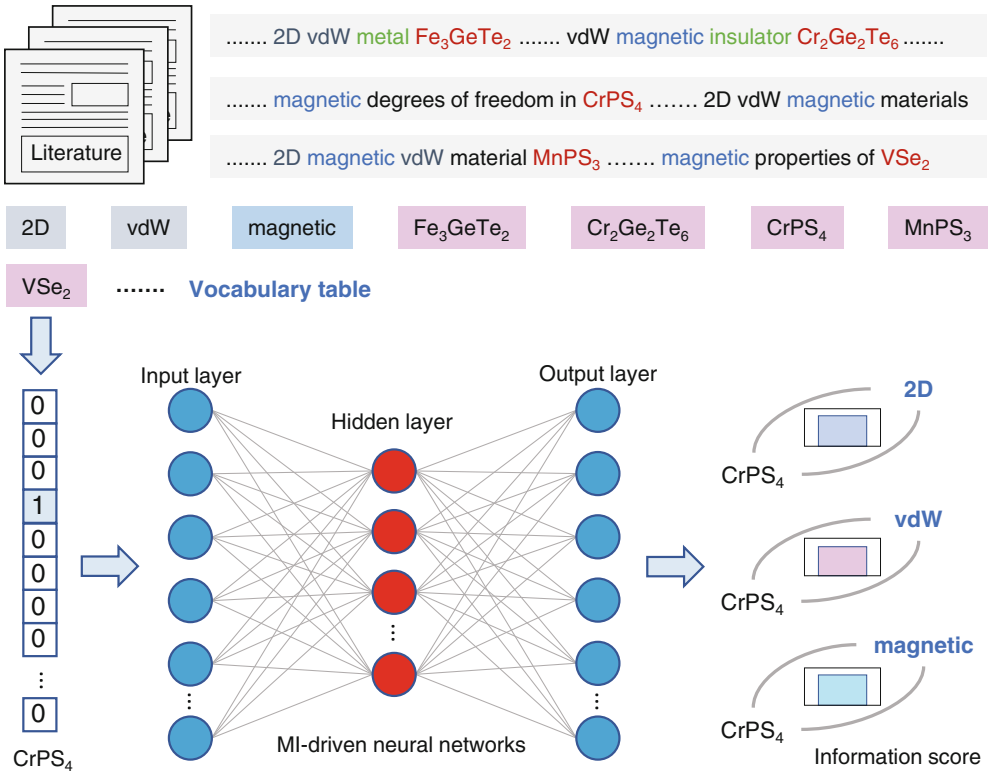

FIGURE 2 Schematic diagram for the calculation process of IS. The process comprises three main steps: (i) the production of the initial one-hot vector (for each word, such as $CrPS_4$, the one-hot vector is a vector with zeros everywhere except its index), (ii) training through the PMI-based neural network, and (iii) the calculation of IS based on the training outcomes.

"photodiode", "phototransistor", "photovoltaic", "piezoelectric", "solar cell", and "thermoelectric". In addition, 5 typical chemistry-related subfields were selected, namely, "fuel cell", "HER" (hydrogen evolution reaction), "LIB" (Li-ion battery), "OER" (oxygen evolution reaction), and "ORR" (oxygen reduction reaction). We chose 15 subfields as the research subjects in all. The maximal information coefficient (MIC)[35] reflects the relationships. More specifically, IS values between materials and the target subfield (or property) are calculated first, followed by MIC between two sets of IS values which is used to show the strength of the relationship between two target subfields. The greater MIC is, the stronger the relationship is. The computational details of MIC are presented in Section 4.2.1.

The MIC-reflected relationships are consistent with materials scientists' experience (Figure S1), thus validating the reliability in quantifying the materials–property relation. The consistency can be well illustrated by the following several examples: (i) there is a strong relationship between "photodetector" and "phototransistor", but they both exhibit weak relationships with "HER", "OER", "ORR", "LIB", and so on chemistry-related subfields; (ii) a strong relationship is presented between "OER" and "HER" or "ORR"; and (iii) "ferroelectric" has the strongest relationship with "piezoelectric" compared with the other subfields. It is worth noting that the MIC-reflected relationship is not obtained simply by examining the semantic similarity. For example, "fuel cell" should be more semantically similar to "solar cell" than the other subfields, but the MIC between "fuel cell" and "solar cell" is low. The semantic similarity between "solar cell" and "photovoltaic" is modest, yet the MIC between them is high. Similarly, while "fuel cell" exhibits a low degree of semantic similarity with "ORR", the MIC-based analysis reveals a relatively strong relationship between the two terms.

The number of 2D vdW magnets that have been experimentally documented is low (about 30; see Table S1 for a complete list). Accordingly, we computed IS between the group consisting of these 30 materials and the 15 subfields presented in Figure S1, 5 attributes related to structural features (i.e., "COF", "MOF", "perovskite", "polymer", and "vdW"), and 3 dimension-related attributes (i.e., "0D", "1D", and "2D"). Obviously, "2D", "vdW", and "magnetic" exhibit remarkably stronger relations with the group than the other attributes and subfields (Figure 3A). The results reveal that, just like experienced researchers, the framework of MI-driven materials discovery can recognize the significant commonality of these materials.

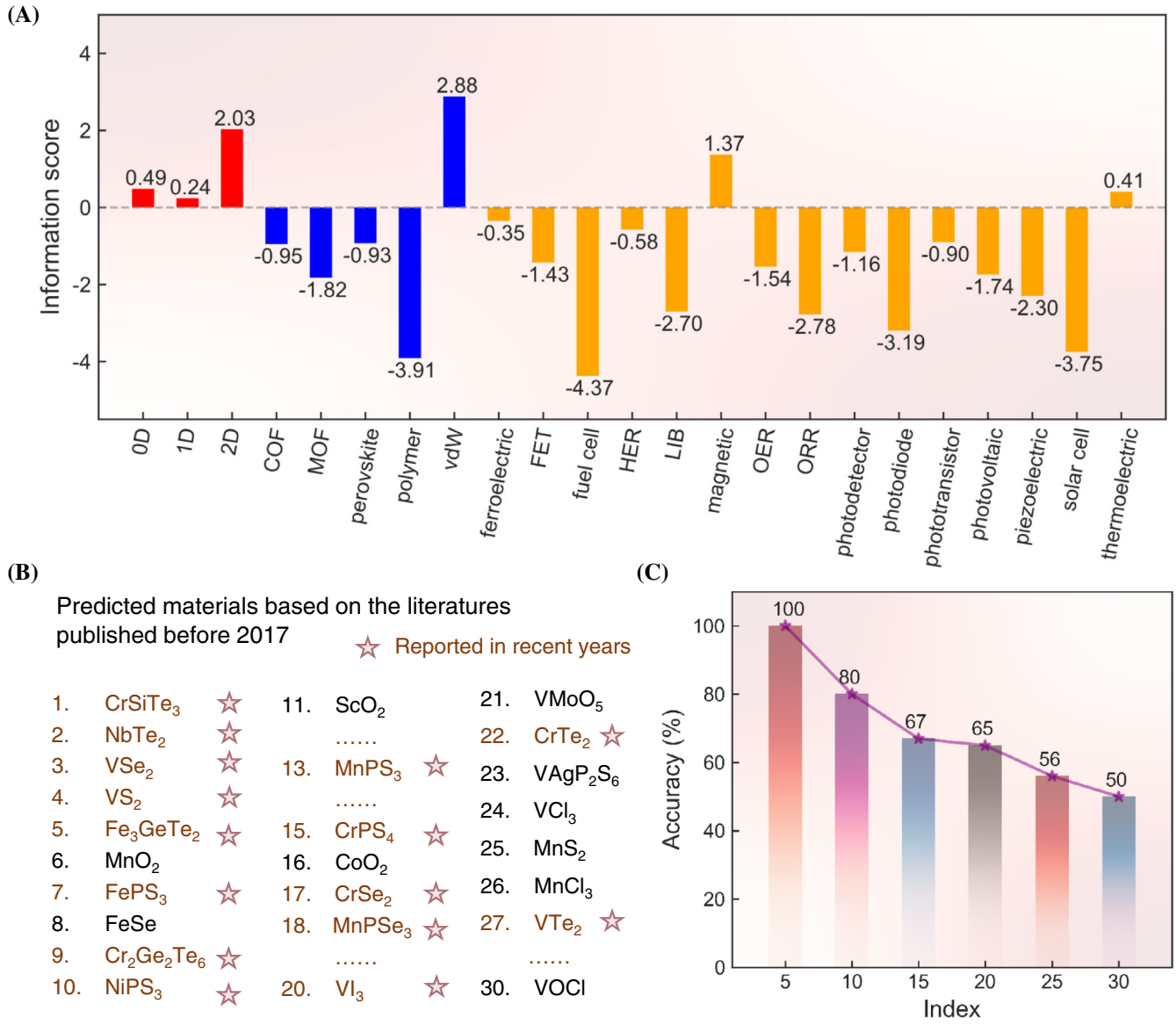

FIGURE 3 (A) IS between the group of materials (30 experimentally documented 2D vdW magnets) and 15 subfields of materials science, 5 structure-related attributes, as well as 3 dimension-related attributes. (B) Top 30 predictions for 2D vdW magnets obtained via the framework of MI-driven materials discovery from materials science literature published before 2017. (C) Percentages of 2D vdW magnets (they have been experimentally documented after 2017) in top 5, 10, 15, 20, 25, and 30 candidate materials presented in (B).

More importantly, with the help of the framework, new 2D vdW magnets with high probability for experimental demonstration can be captured from large volumes of literature data. Spontaneous magnetic ordering is common in bulk materials, but it was not thought to be assured in 2D materials. The experimental evidence of the magnetic ordering in 2D materials was originally reported in 2017,[36,37] answering a long-standing question concerning the presence of the magnetic ordering in the 2D limit. By computing IS with the multivariate feature {"2D", "vdW", "magnetic"}, 30 promising 2D vdW magnets were screened from vast amounts of literature published prior to 2017 (notably, before 2017, they had not been experimentally documented as 2D vdW magnets). After 2017, 80% of the top 10 candidate materials have been experimentally identified as 2D vdW magnets (Figure 3B,C). This demonstrates that the framework of MI-driven materials discovery can discover those 2D vdW magnets with high probability for experimental demonstration. Finally, 10 promising 2D vdW magnets for experimental demonstration were screened from the literature published prior to 2022 to provide valuable candidates for materials scientists (Figure 4).

The above results of historical validation of predictions for 2D vdW magnets have shown that candidate materials with high IS for 2D vdW magnets are more

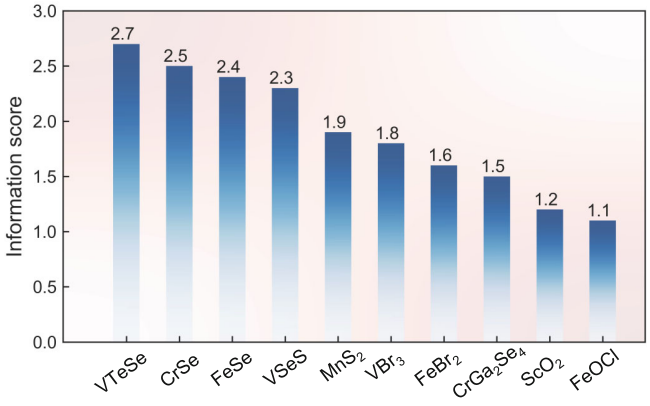

FIGURE 4 Top 10 candidate 2D vdW magnets with high probability for experimental demonstration, which were captured from materials science literature published before 2022 via the framework of MI-driven materials discovery.

likely to be prepared and reported by experimental scientists in future. For example, as shown in Figure 3C, 8 of the top 10 predictions have been experimentally validated in recent years. The resounding success implies that the top 10 predictions in 2022 (Figure 4) can also have high possibility for experimental demonstration in the future. Among the top 10 predictions, VSeTe, VSeS, and MnS_2 have been predicted by DFT calculations as 2D magnets with high magnetic transition temperatures.[24,38,39] CrSe and FeSe with nonlayered structures have been experimentally reported as 2D magnets[40,41]; however, the corresponding vdW layered structures have not been reported. Accessibility of materials in experiment is a crucial prerequisite for the final demonstration of magnetism. Exfoliation from bulk materials is one of the most common approaches for having access to 2D vdW materials. Most of the bulk phases of top 10 predictions have been prepared experimentally and deposited in ICSD or COD, namely, CrSe, FeSe, MnS_2, VBr_3, $FeBr_2$, $CrGa_2Se_4$, ScO_2, and FeOCl. Despite the absence of bulk materials of VSeTe and VSeS in experiment, monolayer transition metal dichalcogenides with a Janus structure like VSeTe and VSeS have been able to be synthesized experimentally,[42–44] suggesting that at least to some extent, 2D VSeTe and VSeS are also experimentally accessible. Remarkably, FeOCl has recently been shown to be an air-stable, promising 2D vdW magnet,[45] which emphasizes the high reliability of our framework.

3 | CONCLUSION

In summary, we have established the framework of MI-driven materials discovery to accelerate experimental discovery of 2D vdW magnets. Historical validation of predictions demonstrates that with the help of the framework, 2D vdW magnets with high probability for experimental demonstration may be well captured from a wide corpus of materials science literature. Many predicted 2D vdW magnets based on the literature published before 2017 (the starting point for experimental exploration of 2D vdW magnets)[36,37] have been experimentally demonstrated in the years since. Another notable aspect of the framework for materials discovery is that its low computational cost, which makes it accessible to personal computers (see the Supporting Information for more details). Furthermore, to provide relevant candidate materials for experimental researchers, we identified 10 promising 2D vdW magnets for experimental demonstration based on the published literature before 2022. It is worth emphasizing that mechanistically, the framework should also be applicable for mining many additional functional materials from large amounts of literature data. The high growth in the number of the published papers year on year has been predictable, and hence the framework could have profound implications for the ongoing development of the field of 2D magnetism.

4 | EXPERIMENTAL SECTION

4.1 | Text processing

Through Elsevier's Scopus application programming interface (API),[46] we collected the abstracts and titles of about 4.89 million English-language materials science literatures since 1921. The reasons for selecting the abstracts and titles as input data are presented in the Supporting Information. Our framework aims to use the theory of MI to characterize the relations between materials and properties underlying massive literature data. We found almost no formulae of materials in the titles and abstracts of the literatures published before 1921. 1921 was therefore selected as the starting point. The corresponding abstracts and titles were segmented and tokenized via natural language toolkit.[47] The tokens that were recognized as valid chemical formulae using pymatgen[48] were standardized (e.g., MoS_2 was converted to Mo_1S_2). Meaningless punctuations and stop words ("that", "it", "is", etc.) were discarded from the corpus. In addition, materials science literatures contain a large number of phrases, such as "fuel cell" and "solar cell". Since the meaning of the phrase often cannot be regarded to be identical to those of its constituent words, it was considered separately during training. That is, these potential phrases detected via genism[49] were spliced into one word using "_" (e.g., "solar cell" was transformed into "solar_cell"). The workflow for text processing is shown in Figure S2.

4.2 | Model training for IS computation

The architecture of SGNS is schematically illustrated in Figure 2. The model is based on a neural network with a single hidden layer. The numbers of input and output neurons are in keeping with that of distinct words in corpus. In the hidden layer, the neuron number was set to 200, and no activation function was used. Each word in the corpus was assigned an index and was passed to the neural network as a one-hot vector. The goal of the PMI-based neural network is to predict the context words given a center word. For training, the window, negative sample, and batch size were set to 10, 1, and 10 000, respectively, and an initial learning rate of 0.01 was employed. After training, IS between two words was calculated by multiplying input and output weights.

4.2.1 | MIC computation

For two variables X and Y, their relation can be captured by the grid partition G (also called as an $x-$by$-y$ grid) on their scatterplot. Given a finite dataset $D=\{(x_i,\ y_i),\ i=1,\ ...,\ n\}$, where n is the size of the dataset and $x_i\in X, y_i\in Y, G$ can be obtained by partitioning the x_i values and y_i values of D into x bins and y bins, respectively. The maximal mutual information (MMI) of on $-D\,G$ is defined as follow:

$$I^*(D,x,y)=\max I\left(D|_G\right)$$

where $D|_G$ represents the distribution of pairs $(X,\ Y)$ on G and $I\left(D|_G\right)$ denotes the MI of on $-D\,G$. It is obvious that $D|_G$ is dependent on G; as a result, the MMI values with different grids can be normalized, and the normalized results can be recorded in a characteristic matrix $M(D)$. Accordingly, the entry $M(D)_{x,y}$ can be described as:

$$M(D)_{x,y}=\frac{I^*(D,x,y)}{\log\min\{x,y\}}$$

MIC refers to the largest normalized MI value in $M(D)$ and is calculated as:

$$MIC(D)=\max_{xy<B(n)} M(D)_{x,y}$$

where the function $B(n)$ determines the grid size. According to the recommendation from the previous work,[35] $B(n)$ was set to $n^{0.6}$ in this work. The MIC computation was performed following a two-step process: (i) for any materials M_i, IS values between it and two target applications were calculated, respectively, and recorded as a pair $(E_i,\ F_i)$; (ii) all the pairs were input into the minepy package,[50] and then the MIC value between two target applications was calculated.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (21825103, 22103026, and U21A2069) and the Fundamental Research Funds for the Central Universities (2020kfyXJJS050).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Menghao Wu https://orcid.org/0000-0002-1683-6449
Huiqiao Li https://orcid.org/0000-0001-8114-2542
Tianyou Zhai https://orcid.org/0000-0003-0985-4806

REFERENCES

1. Jo J, Calavalle F, Martin-Garcia B, et al. Exchange bias in molecule/Fe_3GeTe_2 van der Waals heterostructures via spinterface effects. *Adv Mater*. 2022;34(21):2200474.
2. Zhang H, Ni Z, Stevens CE, et al. Cavity-enhanced linear dichroism in a van der Waals antiferromagnet. *Nat Photon*. 2022;16(4):311-317.
3. Xu C, Li X, Chen P, Zhang Y, Xiang H, Bellaiche L. Assembling diverse skyrmionic phases in Fe_3GeTe_2 monolayers. *Adv Mater*. 2022;34(12):2107779.
4. Kurebayashi H, Garcia JH, Khan S, Sinova J, Roche S. Magnetism, symmetry and spin transport in van der Waals layered systems. *Nat Rev Phys*. 2022;4(3):150-166.
5. Shin I, Cho WJ, An ES, et al. Spin-orbit torque switching in an all-van der Waals heterostructure. *Adv Mater*. 2022;34(8):2101730.
6. Zhuo W, Lei B, Wu S, et al. Manipulating ferromagnetism in few-layered $Cr_2Ge_2Te_6$. *Adv Mater*. 2021;33(31):2008586.
7. Wang X, Cao J, Lu Z, et al. Spin-induced linear polarization of photoluminescence in antiferromagnetic van der Waals crystals. *Nat Mater*. 2021;20(7):964-970.
8. Ni Z, Haglund AV, Wang H, et al. Imaging the Néel vector switching in the monolayer antiferromagnet $MnPSe_3$ with strain-controlled Ising order. *Nat Nanotechnol*. 2021;16(7):782-787.
9. Hwangbo K, Zhang Q, Jiang Q, et al. Highly anisotropic excitons and multiple phonon bound states in a van der Waals antiferromagnetic insulator. *Nat Nanotechnol*. 2021;16(6):655-660.
10. Li B, Wan Z, Wang C, et al. Van der Waals epitaxial growth of air-stable $CrSe_2$ nanosheets with thickness-tunable magnetic order. *Nat Mater*. 2021;20(6):818-825.
11. Zhang K, Han S, Lee Y, et al. Gigantic current control of coercive field and magnetic memory based on nanometer-thin ferromagnetic van der Waals Fe_3GeTe_2. *Adv Mater*. 2021;33(4):2004110.
12. Wang Y, Wang C, Liang SJ, et al. Strain-sensitive magnetization reversal of a van der Waals magnet. *Adv Mater*. 2020;32(42):2004533.

13. Zhang L, Huang X, Dai H, et al. Proximity-coupling-induced significant enhancement of coercive field and Curie temperature in 2D van der Waals heterostructures. *Adv Mater*. 2020; 32(38):2002032.
14. Verzhbitskiy IA, Kurebayashi H, Cheng H, et al. Controlling the magnetic anisotropy in $Cr_2Ge_2Te_6$ by electrostatic gating. *Nat Electron*. 2020;3(8):460-465.
15. Ostwal V, Shen T, Appenzeller J. Efficient spin-orbit torque switching of the semiconducting van der Waals ferromagnet $Cr_2Ge_2Te_6$. *Adv Mater*. 2020;32(7):1906021.
16. Zhang W, Wong PKJ, Zhu R, Wee ATS. Van der Waals magnets: wonder building blocks for two-dimensional spintronics? *InfoMat*. 2019;1(4):479-495.
17. Wang QH, Bedoya-Pinto A, Blei M, et al. The magnetic genome of two-dimensional van der Waals materials. *ACS Nano*. 2022; 16(5):6960-7079.
18. Hossain M, Qin B, Li B, Duan X. Synthesis, characterization, properties and applications of two-dimensional magnetic materials. *Nano Today*. 2022;42:101338.
19. Zhang T, Grzeszczyk M, Li J, et al. Degradation chemistry and kinetic stabilization of magnetic CrI_3. *J Am Chem Soc*. 2022; 144(12):5295-5303.
20. Bian R, Li C, Liu Q, et al. Recent progress in the synthesis of novel two-dimensional van der Waals materials. *Natl Sci Rev*. 2022;9(5):nwab164.
21. Lu S, Zhou Q, Guo Y, Zhang Y, Wu Y, Wang J. Coupling a crystal graph multilayer descriptor to active learning for rapid discovery of 2D ferromagnetic semiconductors/half-metals/metals. *Adv Mater*. 2020;32(29):2002658.
22. Lu S, Zhou Q, Guo Y, Wang J. On-the-fly interpretable machine learning for rapid discovery of two-dimensional ferromagnets with high Curie temperature. *Chem*. 2022;8(3):769-783.
23. Wang B, Zhang X, Zhang Y, et al. Prediction of a two-dimensional high-T_C f-electron ferromagnetic semiconductor. *Mater Horiz*. 2020;7(6):1623-1630.
24. Zhang C, Nie Y, Sanvito S, Du A. First-principles prediction of a room-temperature ferromagnetic Janus VSSe monolayer with piezoelectricity, ferroelasticity, and large valley polarization. *Nano Lett*. 2019;19(2):1366-1370.
25. Torelli D, Moustafa H, Jacobsen KW, Olsen T. High-throughput computational screening for two-dimensional magnetic materials based on experimental databases of three-dimensional compounds. *npj Comput Mater*. 2020; 6(1):158.
26. Shannon CE. A mathematical theory of communication. *Bell Syst Tech J*. 1948;27(3):379-423.
27. Fennell KA, Vassiliadis D, Lam EYN, et al. Non-genetic determinants of malignant clonal fitness at single-cell resolution. *Nature*. 2022;601(7891):125-131.
28. Nieh EH, Schottdorf M, Freeman NW, et al. Geometry of abstract learned knowledge in the hippocampus. *Nature*. 2021; 595(7865):80-84.
29. Liu B, Hong A, Rieke F, Manookin MB. Predictive encoding of motion begins in the primate retina. *Nat Neurosci*. 2021;24(9): 1280-1291.
30. Hasan M, Paul S, Mourikis AI, Roy-Chowdhury AK. Context-aware query selection for active learning in event recognition. *IEEE Trans Pattern Anal Mach Intell*. 2020; 42(3):554-567.
31. Liu X, Yang C, You J, Kuo CCJ, Kumar BVKV. Mutual information regularized feature-level Frankenstein for discriminative recognition. *IEEE Trans Pattern Anal Mach Intell*. 2022;44(9): 5243-5260.
32. Ragonesi R, Volpi R, Cavazza J, Murino V. Learning unbiased representations via mutual information backpropagation. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops; 2021:2723-2732.
33. Harris ZS. Distributional structure. *WORD*. 1954;10(2–3): 146-162.
34. Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J. Distributed representations of words and phrases and their compositionality. In: NIPS; 2013: 3111-3119.
35. Reshef DN, Reshef YA, Finucane HK, et al. Detecting novel associations in large data sets. *Science*. 2011;334(6062):1518-1524.
36. Gong C, Li L, Li Z, et al. Discovery of intrinsic ferromagnetism in two-dimensional van der Waals crystals. *Nature*. 2017; 546(7657):265-269.
37. Huang B, Clark G, Navarro-Moratalla E, et al. Layer-dependent ferromagnetism in a van der Waals crystal down to the monolayer limit. *Nature*. 2017;546(7657):270-273.
38. Guan J, Huang C, Deng K, Kan E. First-principles prediction of room-temperature ferromagnetic semiconductor MnS_2 via isovalent alloying. *J Phys Chem C*. 2019;123(15):10114-10119.
39. Guan Z, Ni S. Predicted 2D ferromagnetic Janus VSeTe monolayer with high Curie temperature, large valley polarization and magnetic crystal anisotropy. *Nanoscale*. 2020;12(44): 22735-22742.
40. Zhang Y, Chu J, Yin L, et al. Ultrathin magnetic 2D single-crystal CrSe. *Adv Mater*. 2019;31(19):1900056.
41. Xu J, Li W, Zhang B, et al. Free-standing 2D non-van der Waals antiferromagnetic hexagonal FeSe semiconductor: halide-assisted chemical synthesis and Fe^{2+} related magnetic transitions. *Chem Sci*. 2022;13(1):203-209.
42. Lu AY, Zhu H, Xiao J, et al. Janus monolayers of transition metal dichalcogenides. *Nat Nanotechnol*. 2017;12(8):744-749.
43. Qin Y, Sayyad M, Montblanch ARP, et al. Reaching the excitonic limit in 2D Janus monolayers by in situ deterministic growth. *Adv Mater*. 2022;34(6):2106222.
44. Guo Y, Lin Y, Xie K, et al. Designing artificial two-dimensional landscapes via atomic-layer substitution. *Proc Natl Acad Sci U S A*. 2021;118(32):e2106124118.
45. Zeng Y, Gu P, Zhao Z, et al. 2D FeOCl: a highly in-plane anisotropic antiferromagnetic semiconductor synthesized via temperature-oscillation chemical vapor transport. *Adv Mater*. 2022;34(14):2108847.
46. Rose ME, Kitchin JR. Pybliometrics: scriptable bibliometrics using a Python interface to Scopus. *SoftwareX*. 2019;10:100263.
47. Wagner W. Steven Bird, Ewan Klein and Edward Loper: natural language processing with Python, analyzing text with the natural language toolkit. *Lang Resour Eval*. 2010;44(4):421-424.
48. Ong SP, Richards WD, Jain A, et al. Python materials genomics (pymatgen): a robust, open-source python library for materials analysis. *Comput Mater Sci*. 2013;68(2013):314-319.
49. Rehurek, R, Sojka, P. Software framework for topic modelling with large corpora. In: Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks; 2010.
50. Albanese D, Filosi M, Visintainer R, Riccadonna S, Jurman G, Furlanello C. minerva and minepy: a C engine for the MINE

suite and its R Python and MATLAB wrappers. *Bioinformatics*. 2013;29(3):407-408.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Song H, Zhao Y, Turner E, et al. Capturing 2D van der Waals magnets with high probability for experimental demonstration from materials science literature. *InfoMat*. 2023;5(4):e12397. doi:10.1002/inf2.12397