Revolutionizing Materials Research with Quantum Computation: The Role of High-Throughput Experimental Materials Databases

[1]Anandakumar Haldorai and [2]Shrinand Anandakumar
[1]Center for Research and Development, Sri Eshwar College of Engineering, Coimbatore, Tamil Nadu, India.
[2]PSBB Millennium School, Coimbatore, Tamil Nadu, India.
[1]anandakumar.psgtech@gmail.com, [2]shrinand.psbb@gmail.com

Correspondence should be addressed to Anandakumar Haldorai : anandakumar.psgtech@gmail.com

Article Info
Journal of Computational Intelligence in Materials Science (https://anapub.co.ke/journals/jcims/jcims.html)
Doi: https://doi.org/10.53759/832X/JCIMS202402011.
Received 18 March 2024; Revised from 02 August 2024; Accepted 12 September 2024.
Available online 15 September 2024.

Abstract – The combination of quantum computing (QC) with artificial intelligence (AI) has the potential to significantly transform several industries and enhance quality of life. AI applications, such as autonomous vehicles and image recognition, significantly depend on supervised learning, which is a type of machine learning. Researchers are now examining the intersection of AI and multi-agent planning systems (MAPS) using quantum algorithms and game theory. Quantum machine learning and optimization techniques use the advantages of QC to accelerate training and optimization processes, drawing inspiration from quantum physics. With machine learning methods, it is now becoming commonplace in materials research to make predictions about the characteristics of inorganic solid-state materials. However, a significant issue lies in the limited availability of datasets. This article explores the structure and characteristics of the High Throughput Experimental Materials (HTEM) Database, which contains a vast amount of experimental data on inorganic materials obtained using high-throughput technologies. The HTEM Database provides a significant resource by offering comprehensive EDIM built using high-throughput exploration methodologies. The database offers a multitude of search and visualization functionalities and may be accessible using a web-based interface. Regular updates are performed on the HTEM DB to integrate new data and mitigate processing issues.

Keywords – Artificial Intelligence, Machine Learning, Quantum Materials, Quantum Computing, High Throughput Experimental Materials.

I. INTRODUCTION

Renowned quantum computer (QC) scientists, such as Lov Grover and Peter Shor, have brought up the subject of the distinctive characteristics that quantum computers offer in contrast to regular classical computers. Shor's approach provides a substantial increase in computational speed for the process of factorizing large numbers. This functionality has the potential to significantly impair the efficiency of the commonly used Rivest-Shamir-Adleman (RSA) public key encryption method [1], which is often utilized on the Internet. Chung, Yu, and Wong [2] displayed the capacity of a quantum database to provide a 100% improvement in search query performance. Bauer et al. [3] confirmed Feynman's initial conjecture on quantum simulation, thereby establishing a path towards the effective simulation of materials at the quantum scale.

Data is stored by quantum computer (QC) by using “qubits” or “quantum bits”. These qubits may be imagined of as a pair of quantum states (QS) associated with “artificial atoms” generated by quantum mechanics. This issue of MRS Bulletin discusses several techniques for producing these artificial atoms. The systems in issue include a range that extends from atomic ions to complex structures like Josephson junctions and quantum dots, which consist of billions of atoms. Under certain situations, these devices demonstrate behavior that is similar to atomic systems. The quantum mechanical description of these entities comprises their internal states, where the information of an N-qubit structure is stored in the N-body probability amplitude. Usually, the qubit is formed by selecting two QS from the several artificial atoms’ states, resulting in a quantum system with just two levels. More precisely, a stores data in the quantum-mechanical extents of the fundamental states of each of the N-body, two-level systems. These are a pair of intricate numerals. The memory of a QC is composed of N qubits, which may be precisely manipulated to experience certain physical interactions.

To process this information, qubits must interact with one other through a certain arrangement of classical control indication, such as electromagnetic fields. These signals link the qubits and perform quantum logic performance. This

quantum algorithm (QA) facilitates the advancement of all quantum data. The primary factor behind the remarkable potential of quantum computing, especially in specific tasks, is in the control of interactions among N qubits, resulting in the proliferation of 2N complex coefficients. However, the data kept in these qubits is crucial. The complex properties that characterize a qubit may be violated in two main ways. Initially, multiple qubits are obtained by extracting the lowest energy state and a higher energy level from a physical system. The deterioration of the qubit's state and the loss of quantum information happen as a result of the energy relaxation process towards the lowest energy state, which defines the duration of radiation for the system. Dephasing is a mechanism that may cause the loss of quantum data. The interaction between a qubit and its surrounding environment offers summation relaxation pathways that affect both the duration and phase of the qubit's state.

When the qubit socializes with other unconsidered two-level systems, such as residual environmental degrees of freedom, it will cause a seemingly random change in the qubit's phase, even if the amplitudes remain unchanged. It is essential to minimize the interactions between the qubit and its physical surroundings in quantum technologies. Over the last 15 years, the discipline of quantum computing has evolved from a niche and esoteric physics` branch to a challenging and rigorous frontier of emerging technology, engineering, and materials science [4]. While there are yet no QCs that are deemed "useful" and capable of scaling, there has been steady advancement in the creation of many potential candidate designs. The creation of a quantum computer has many hurdles that are closely related to topics that are particularly suitable for the area of materials science.

This article aims to justify the design and use of the High Throughput Experimental Materials (HTEM) Database. The fusion of artificial intelligence (AI) with quantum computing has the capacity to transform several industries and augment our standard of living. Machine learning methods, namely supervised learning, are crucial in many AI applications, such as image recognition and autonomous vehicles. The rapid expansion of data and ongoing progress in computer capabilities have significantly increased the potential of machine learning and artificial intelligence to transform several industries and improve our overall well-being. The rest of the article is organized as follows: The second section presents an overview of AI for quantum computing. In this section, we review AI and ML, different applications of quantum computing in AI, quantum ML algorithms, and quantum-inspired optimization algorithms. The third section focuses on related works on quantum computing. The fourth section focusses on making forecasts from known materials in experimental databases based on the application of machine learning. The fifth section reflects on the results and discussion of database infrastructure and content, and database explorations and application. Lastly, the sixth section presents a summary of the results.

II. OVERVIEW OF AI FOR QUANTUM COMPUTING

Overview of AI and ML

Supervised learning in machine learning refers to the process of instructing an algorithm by presenting it with a collection of labeled instances. This allows the program to forecast results for novel, unmarked data. The agent acquires knowledge via a process of experimentation, where it is rewarded for making favorable choices and penalized for making unfavorable ones [5]. Machine learning is essential in several AI applications, including image recognition and autonomous cars. With the exponential growth of data and the continuous advancement of computer power, the capacity of machine learning and artificial intelligence to revolutionize several sectors and enhance our quality of life grows more substantial.

Applications of Quantum Computing in AI

Quantum algorithms for learning

The intersection of quantum processing with AI has proved to be significant in the field of machine learning. Multiple scholarly articles, such as [6] and [7] are dedicated to the quantum extension of computational learning theory. Their objective is to discover QAs that surpass the efficiency of current classical methods in learning particular classical gadgets, like Boolean functions. This study is intricately connected to quantum intricacy theory.

Quantum algorithms for decision problems

Decision difficulties may often be expressed using decision trees (DT). Childs and Farhi [8] demonstrated that QAs using Hamiltonian evolution may solve choice issues depicted by a certain class of DT much quicker than conventional random walks, with an exponential speedup. However, this does not indicate any superiority of quantum computing over conventional computation for this specific category of issues, since they may also be efficiently addressed by alternative conventional techniques.

Quantum search

A significant portion of the first AI research focused on the exploration of search methodologies. This phenomenon may be attributed to two factors. Firstly, several artificial intelligence (AI) challenges can be simplified to a search process, such as data retrieval, planning, theorem proving, and scheduling. Secondly, computers possess the ability to do these tasks at a much higher speed compared to people. The Grover algorithm [9] demonstrates that quantum computers have the capability to outperform traditional computers in terms of speed. It is commonly anticipated that quantum computing would be extensively used in the field of artificial intelligence to effectively address a range of search-related issues. Quantum searching is anticipated to be among the first quantum computing approaches that would significantly impact AI. Moret-

Bonillo [10] provided a comprehensive analysis of the use of quantum search algorithms in the field of artificial intelligence. However, to this day, a decade later, there have been only a limited number of successful implementations of quantum searching in the field of artificial intelligence.

Quantum game theory

Game theory aims to mathematically represent a scenario in which individuals engage in interactions. The individuals participating in the game are referred to as players, their potential actions are known as moves, and a set of instructions that dictates the specific move to be executed in every given game scenario is termed a strategy. A strategy is a comprehensive plan of action that encompasses all potential possibilities within the established rules of the game. A pure strategy is a consistent choice of a certain move in a game, but a mixed strategy involves using a randomizing mechanism to pick among alternative moves. The utility of a game result to a player refers to a quantitative assessment of how desirable that outcome is for the player. A payout matrix assigns quantitative values to the players' utility for each possible result of the game. The participants are expected to optimize their utility within the prescribed rules of the game. Games characterized by players' options being immediately revealed are referred to as games of perfect data. These are the primary ones that are of relevance to us in this context.

A dominating approach is a strategy that consistently outperforms any other plan regardless of the actions taken by the opponent player(s). The Nash equilibrium (NE) has the utmost significance among the potential equilibria in game theory. It refers to a situation where no actor can increase their payout by unilaterally changing their strategy, because to the mix of methods being used. A Pareto optimum outcome is a situation in which no actor may increase their utility without causing a decrease in another person`s utility. Strategy A is considered evolutionarily stable versus strategy B if, for any adequately small and positive epsilon (ǫ), A outperforms B while playing against the mixed approach $(1 - ǫ)A + ǫB$. An evolutionarily stable strategy (ESS) is a strategy that remains stable and resistant to evolution when compared to all other strategies. The collection of all strategies that constitute Evolutionarily Stable Strategies (ESS) is a subtype of the NE of the game. A two-player zero-sum game is characterized by the participants having completely opposing interests. In other words, the total of the rewards for any outcome in the game is equal to zero. A saddle point in a game refers to an entry in the action matrix where the row player's payment is the lowest among all the entries in its row, and the column player's action is the highest among all the entries in its column.

The use of game theory in artificial intelligence is becoming prevalent, particularly in the context AI and multi-agent systems(MAS). In recent times, there have been proposals for quantum elongations of game theory in a series of works. For instance, Vijayakrishnan and Balakrishnan [11] established the concept of quantization in nonzero sum games involving two players, while Martínez-Martínez [12] offered quantum games involving more than two players. Duong et al. [13] posited that quantum game theory provides novel methodologies for addressing some challenges in artificial intelligence (AI). Additional possible applications of quantum processing in AI include (a) Representing knowledge via quantum juxtaposition and accelerating knowledge thinking through quantum parallelism. (b) Employing distributed quantum computing and quantum communication in MAS, specifically utilizing entanglement for cooperation.

Quantum Machine Learning Algorithms

The integration of QC with ML, forming a burgeoning field of research is referred to as Quantum machine learning (QML). It utilizes the benefits of quantum computing, such as parallel processing and quantum entanglement, to expedite the training and inference procedures of ML algorithms. The intricacy of ML tasks. Utilizing linear algebra for QML: The quantum logic operation, or qubits, applies a multiplication of the appropriate state vector by 2n*2n matrices. The pivotal method for transitioning from classical computing to quantum computing is the HHL (Harrow, Hassidim, and Lloyd) algorithm.

Quantum Boltzmann machine

Quantum Boltzmann machines are a specific kind of quantum neural network used to learn a collection of Hamiltonian parameters (wj) that may mimic an input state (ρtrain) by using a predetermined set of Hamiltonian terms (Hj). The quantum relative entropy (S) is used to quantify the accuracy of the approximation. It serves as an upper limit for the difference between the two states. Experimental methods may be used to determine the gradient of the relative entropy, and the Hamiltonian parameters are updated via gradient descent. Stoquastic Hamiltonians, characterized by real and non-positive off-diagonal matrix elements, are very compatible with this approach. However, there is yet no efficient classical counterpart for non-stoquastic Hamiltonians.

Quantum-Inspired Optimization Algorithms

To address optimization issues, classical algorithms known as quantum-inspired optimization algorithms use some of the ideas from quantum physics. These algorithms surpass conventional optimization strategies by effectively exploring search areas via the use of entanglement, disruption, and superposition ideas. Here are many optimization strategies that are influenced by principles from quantum mechanics: Quantum annealing is an optimization strategy that aims to identify the lowest energy state of a system by using quantum tunneling, which allows particles to pass through barriers they lack sufficient energy to overcome. Although often used for D-Wave quantum computers, this methodology may also be implemented on classical computers using traditional simulation approaches. Conventional optimization techniques, such as

Quantum-inspired Simulated Annealing, use thermal fluctuations to uncover the optimal solution. Quantum-inspired simulated annealing algorithms use principles of quantum mechanics to enhance the efficiency of search space exploration and discover solutions that are close to optimum.

Quantum-inspired optimization algorithms have shown superior efficiency compared to traditional optimization algorithms in handling some issues, especially those pertaining to optimization in complex systems. Nevertheless, it remains uncertain if they can surpass quantum algorithms in performance when executed on a quantum computer. Meanwhile, these algorithms and the knowledge gained from their development might contribute to the improvement of design suggestions for quantum hardware.

III. RELATED WORKS

In order to inspire the development of a data mechanism that gathers and enhances an experimental data stream for later utilization by sophisticated systems, Draxl and Scheffler [14] initially examine the requirements of data and experimental researchers (ER) in relation to their usual workflows, as illustrated in the materials science domain example depicted in **Fig 1**. This study delineates the shared and distinct data infrastructure needs of high-throughput experimental (HTE) data and materials researchers. The research community for HTE materials begins a study by developing a hypothesis, which is subsequently tested through experiments. The data undergo processing and analysis, and the findings are communicated via a peer-reviewed paper that examines the connections between structure, material production, composition, processing, performance, and characteristics. The materials-data researchers start their investigation by finding a collection of pertinent data. The dataset is then gathered and organized in order to facilitate filtering and analysis for the purpose of identifying relationships among the data, which ultimately provide the stated findings.

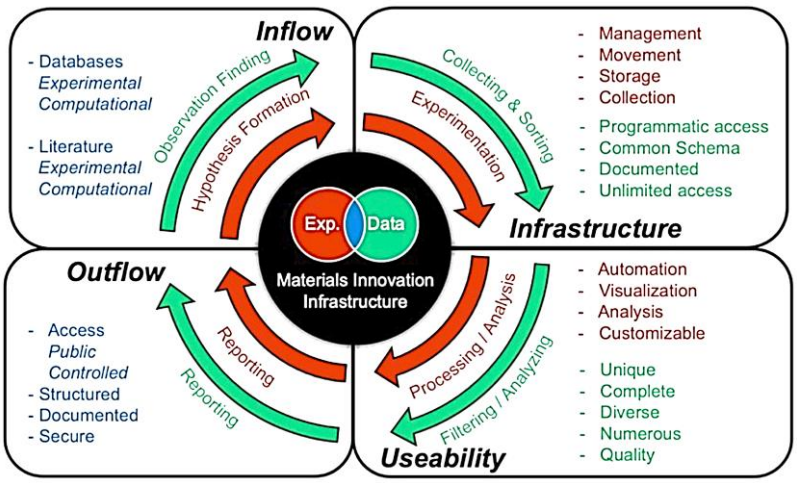

Fig 1. Requirements for managing data in the field of experimental materials research

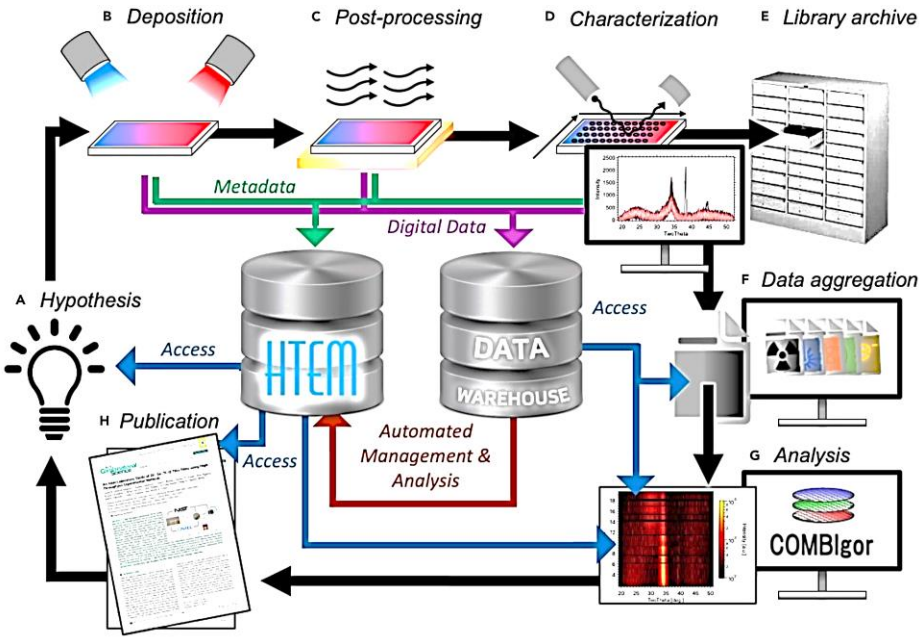

Fig 2. Methodologies and processes for conducting extensive materials research for experimentation and data analysis

According to Barseghian et al. [15], both the experimental and data processes have distinct necessities, but they may be combined into a unified workflow by categorizing the data demands as output, input, usability, or infrastructure, as seen in **Fig 1**. The experimental workflow necessitates the use of instruments for the acquisition, organization, and retention of recently created data, whereas the data workflow necessitates convenient retrieval of the stored data. In addition, the ER need tools for the analysis and extraction of insights from the information, while the information researchers want extensive, varied, and high-caliber datasets. Nevertheless, both experimental and data processes need access to a repository for new data and previously acquired data, resulting in a significant convergence in their outputs and inputs. The need for these overlapping needs led to the development of the RDI, which gathers, analyzes, and archives experimental metadata and data. Additionally, the HTEM-DB serves as a repository for experimental data and as a resource for data-driven investigations.

Fig 2 illustrates the integrated data and experimental process that is widely used by researchers in the field of components discovery at NREL. NREL has a significant amount of high-throughput experimentation (HTE) facilities and experience in thin-film components research (**Fig 2 A–2E**). Bhati, Nazeeruddin, and Maréchal [16] focus on the deposition and analysis of thin films. The films are typically deposited on square substrates measuring 50 × 50-mm (2 × 2”) and are characterized using a 4 × 11 sample mapping grid. These substrates and mapping grids are often used in various combinatorial spatially resolved characterization devices and thin-film deposition chambers at NREL. This experimental procedure at NREL has been compared to and evaluated against the methodologies used by other facilities. Other publications highlight the various types of chemical compositions (such as oxides, nitrides, chalcogenides, Li-containing materials, intermetallics) and characteristics (such as optoelectronic, electronic, piezoelectric, photoelectrochemical, thermochemical) to which these high-throughput experimentation (HTE) methods have been utilized.

The data (green) and experimental (red) research procedures include distinct infrastructure and usability prerequisites on the right side, while sharing common data intake and outflow requirements on the left side. The research procedures at NREL are integrated into a unified RDI and HTEM-DB. This allows for the use of current experimental data to generate novel insights in materials science using machine learning. The process commences with (A) formulating the experiment design, followed by the production (B) of material samples. These samples are then (C) subjected to treatment, (D) measured, and finally (E) stored in archives. The estimated data are gathered for the purpose of evaluation and are then published in a publication, where they provide information for future research. In order to collect metadata (green), measure data (purple), automate file management (red), and provide access (blue), data tools were developed at each step.

In [17], the sample libraries (SLs) mentioned in HTEM DB were prepared using PVD synthesis techniques, namely pulsed laser deposition (PLD) and co-sputtering, which were conducted in vacuum chambers. Both of these physical vapor deposition (PVD) procedures involve the conversion of solid substance precursors, often ceramic or metallic disks known as 'targets', into a plasma phase by the application of external energy stimuli. Sputtering utilizes electrified radio frequency power, while pulsed laser energy is used in pulsed laser deposition (PLD). The vapor that is produced subsequently condenses onto a smooth reinforcement, often circle or square plates made of silicon or glass, which are commonly known as 'substrates'. A thin film layer is created on the substrate's surfaces because of this condensation process. By targeting the substrate, combinatorial sputtering techniques allow purposeful shape differentiation of coatings, and maintain high throughput. Considering the pulsed laser deposition (PLD), the anticipated results are accomplished by proper synchronization of switches and rotational substrates. The aspect of sputtering is widely employed as an approach, which facilitates smooth transformation of newly established materials for the lab to large-scale production.

Research by Kuri et al. [18] reviewed sample libraries (SLs) in the HTEM database based on the application of various approaches such as X-ray diffraction, electrical diffraction, X-ray fluorescence, and optical sensing. The approaches of X-Ray Fluorescence (XRF) and X-Ray Diffraction (XRD) utilize x-rays to evaluate the crystalline nature, and thickness of thin film materials. Understanding these materials is fundamental to identify defects in the materials’ micro-structure at the interfaces and in an atomize samples. To effectively evaluate the resistance of leaves within a specific sample, a 4-point electric probe is employed to accurately quantify the output of energy. Similarly, optical spectroscopy utilizes light with the visible, infrared, and ultraviolet spectra to effectively quantify the transmittance and reflectance of materials. This is obtained by evaluating both electrical and optical data to compute the required figures. Using XRF, the samples’ thickness can then be determined. The optoelectrical characteristics, as assessed in this context, play a crucial role in identifying the most suitable material for a certain application.

Within the chosen substance structure, characterized by its specific chemical composition, the word "high-throughput" (HT) in the HTEM DB name signifies the increased rate of testing, as mentioned by Williams et al. [19]. The use of a slope enables the concurrent amalgamation of a vast quantity of chemicals, so achieving this objective. The shape of the material is often determined by the statistical probability of possessing the required attributes, together with prior experimental investigations and theoretical calculations. Choosing this option often leads to a decrease in throughput. To ascertain the essential features, we examine the precise technical prerequisites of the financial organizations that are backing the project. The HTEM DB contains instances of Co-based oxide material structures, such as Co-Ni-O, Co-Zn-Ni-O, Co-Zn-O, and Co-O, which have been identified by first-principles calculations. Research has been conducted on the potential use of these substances as transparent p-type emitters in the electrical interface activities of solar cells. The HTEM DB is a compilation of information from many sponsored research initiatives on specific structures systems, rather than an attempt to comprehensively include all potential merging of tables of elements. It is important to mention that HT studies expedite the

investigation of certain factors that can be easily tested using HT methods, like substrate temperature or chemical composition. Adjusting additional processing settings often results in further enhancement of the qualities of the materials.

According to Zakutayev et al. [20], using the previously described Zn-Ni-Co-O databank as an instance, the comprehensive analysis of structures reveals that the maximum specific conductance of 70 S/cm is seen in compositions that are rich in Ni and Co. Nevertheless, by refining the thin film accumulation technique and using post-accumulation annealing, it is possible to enhance the conductivity to 170 S/cm36. This demonstrates that the purpose of HTEM is to explore broad correlations and patterns throughout the data in order to pinpoint the most promising locations for processing parameters and chemical compositions in order to optimize them further.

This work introduces and describes the HTEM Database, which comprises substantial exploratory information for inanimate materials developed utilizing HT exploratory procedures. The database holds the promise of future research and development efforts.

IV. MAKING PREDICTIONS FROM KNOWN MATERIALS IN EXPERIMENTAL DATABASES USING ML

Machine learning (ML) is a sector within computer science that focuses on the development of algorithms capable of creating models based on existing data, identifying patterns and relationships within this data, and making predictions about information that is not yet accessible. The forecasts are based on data mining, which is the practice of uncovering sequences in large information via data techniques. ML techniques have lately achieved success in automating processes, interpreting natural language, and analyzing visual data. These successes have been possible due to the availability of extensive datasets that facilitate data-driven modeling.

These achievements also generated debates over the possibilities of 'Artificial Intelligence' in the field of research and 'The Fourth Paradigm' of information-driven scientific exploration. Applying AI to information-driven substances discovery is crucial in the field of materials research, since novel materials often serve as the foundation for significant advancements in contemporary technology. Advanced energy technologies have been facilitated by the utilization of gallium nitride in light-discharging rectifying valve, enabling enough solid-state lighting. Lithium-ion batteries' usage of intercalation materials has enabled the creation of electric vehicles. Furthermore, without the use of silicon material, contemporary computers would not have been possible to exist.

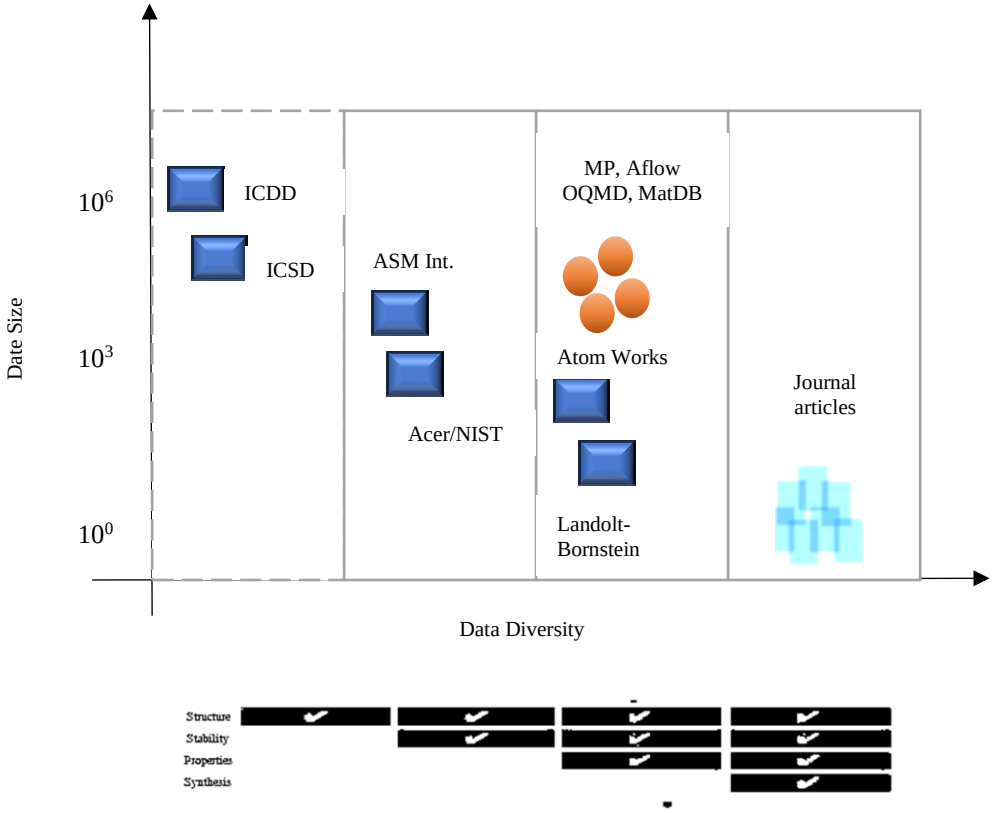

Fig 3. Relation between the variety and amount of materials data in present databases

Machine learning techniques have recently been used in computational materials science to forecast the structure, stability, and characteristics of inanimate solid state substances. These discoveries were made possible by advancements in feigning techniques at various length scales. The generated simulated substances information are maintained in continuously expanding publicly-accessible conceptional property databanks. Unlike calculations, the process of discovering materials

experimentally via ML is constrained by the scarcity of extensive and varied datasets (see **Fig 3**). The Inorganic Crystal Structure Database (ICSD), a vast collection of experimental data, consists of hundreds of thousands of entries. However, it lacks diversity since it only provides information on the structure and composition of the materials. The Landolt–Börnstein dataset [21] and the AtomWorks dataset [22] consist of hundreds to thousands of entries for various attributes. However, these datasets are not sufficiently big to train contemporary machine learning methods. Moreover, none of these databases provide synthesis data like pressure or temperature, which is essential for fabricating materials with desired characteristics.

In **Fig 3**, the combinatorial databanks, shown by red circles, are extensive and varied. Exploratory databanks, on the other hand, have restrictions in terms of their size and diversity, which restricts the applicability of machine learning techniques. ML in investigational materials research has mostly concentrated on implementing established methods that are appropriate for relatively small but intricate datasets. These datasets often consist of x-ray diffraction patterns, microscope pictures, or materials microstructure. A possible machine learning approach to generate extensive and varied materials datasets is using natural language processing techniques on research papers found in scientific publications.

Nevertheless, a significant proportion of journal articles in the field of experimental materials science is confined to the findings that writers subjectively deem as the most captivating, resulting in a substantial volume of unpublished 'dark information. In addition, the published papers often exhibit a bias towards favorable study findings, since the scientific community discourages the publishing of unsuccessful trials. The selective omission of unfavorable findings is a challenge for machine learning, since several algorithms need the inclusion of both positive and negative outcomes to facilitate effective training. In conclusion, a relatively tiny percentage of these journal article articles are connected to the related publishing of information, even while funding agencies and scientific journals are pushing for more public availability of research data.

In this article, we outline our advancements in developing the HTEM Database, which may be accessed in [23]. The HTEM DB is the first publicly accessible extensive compilation of investigational information for inanimate structures produced by HTE thin film methods. The HTEM DB currently comprises of data on the synthesis circumstances, optoelectronic property measurements, crystal structure, and chemical content, of the materials.

V. RESULTS AND DISCUSSION

Database infrastructure and content

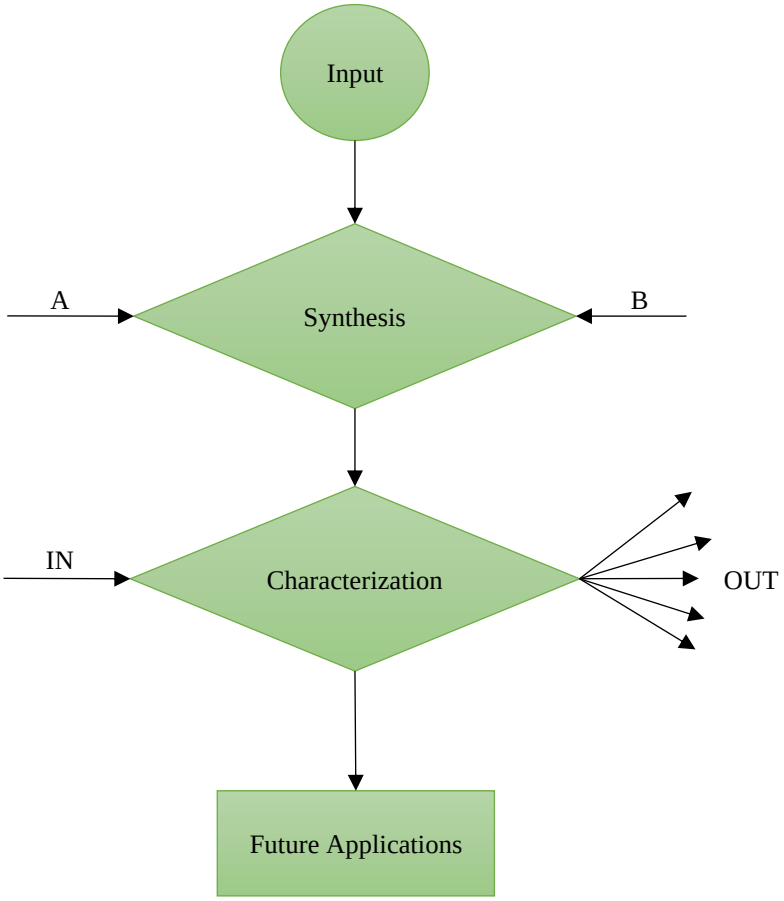

Fig 4. Characterization of each input in the library using spatially resolved characterisation techniques

The HTEM DB stores data on inorganic materials that have been produced as thin films, a configuration that is well-suited for conducting experiments on a large scale. Insulators, metals, and semiconductors make up the thin film SLs in

HTEM DB. These samples are generated using PVD procedures. Each sample in the library is characterized using spatially figured out characterization methods, as shown in **Fig 4**. The choice of materials for synthesis is often determined by computational predictions, previous experimental research, and particular target applications such as photoelectrochemical absorbers, solar cell materials, piezoelectrics, and transparent contacts. Materials with potential that are produced via the combination of spatially and synthesis detailed characterization are often refined through the use of conventional experimental techniques.

With the exception of a few intermetallic complexes, the majority of these components are found in compounds that are chalcogenide, nitride, or oxide. The Related Works section and specialized review articles give a concise overview of various high-throughput experimental approaches, with a specific emphasis on electrical and energy materials. This work does not mention other types of materials, like homogeneous catalysts or organic polymers, that may be studied using high-throughput experimentation. The reason for excluding them is because they are not presently integrated in the HTEM DB. The HTEM DB currently has 141,574 records of thin film inanimate substances as of 2018. The entries are organized into 4,356 sample libraries (SLs), which are spread out across about 100 distinct materials systems. **Fig 5** provides a visual representation, in the form of a bar chart, of the HTEM DB content. It specifically focuses on the 28 most frequently occurring metallic components.

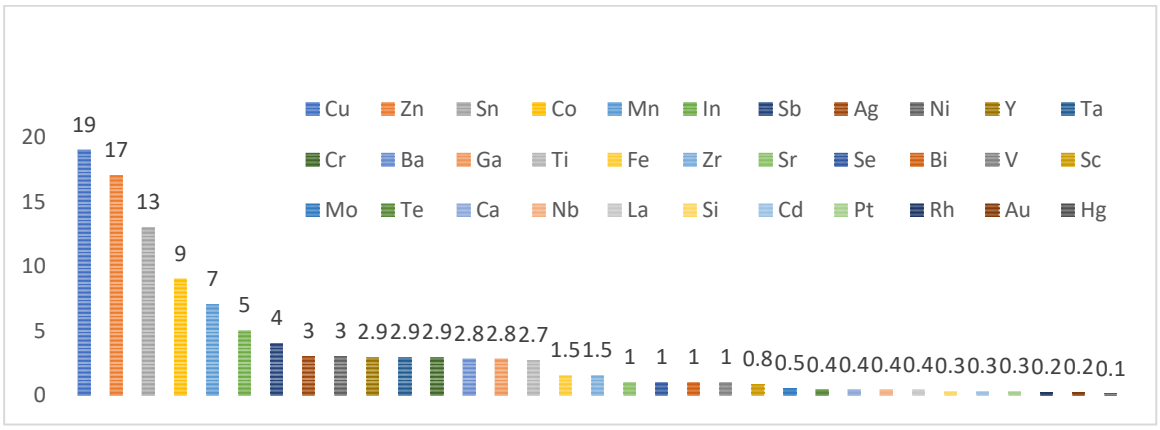

Fig 5. Identification of metallic components in the HTEM DB based on data from x-ray illumination analyses and accumulation precursor data

Most of these metallic structures are found in the form of compounds, with oxides accounting for 45%, chalcogenides for 30%, and nitrides for 20%. A small percentage, about 5%, also form inter-metallics. For these items, synthesis conditions including temperature (83,600), electrical conductivities (32,912), optical absorption spectra (55,352), composition and thickness (72,952), x-ray divergence patterns (100,848) are supplied. Of the sample entries, 80–90% have never been published before, and just around 10% have been reported in peer-reviewed writings.

At present, over half (50-60%) of the properties` measured are accessible to the public, while the other properties are owned by privately-funded organizations and continuing or completed public programs that still need data curation. The NREL Laboratory Information Management System (LIMS) facilitates substances research and supports the HTEM DB. The LIMS is tasked with the automated collection, categorization, and storage of synthesis information and measurement information in a data warehouse. The Extract, Transform, and Load (ETL) procedure consolidates specific information into a customized relational databank known as the HTEM DB. The HTEM DB may be accessible via a web-based interface (WBI) and other visualization and analytical tools using a quality-based application programming interface (API).

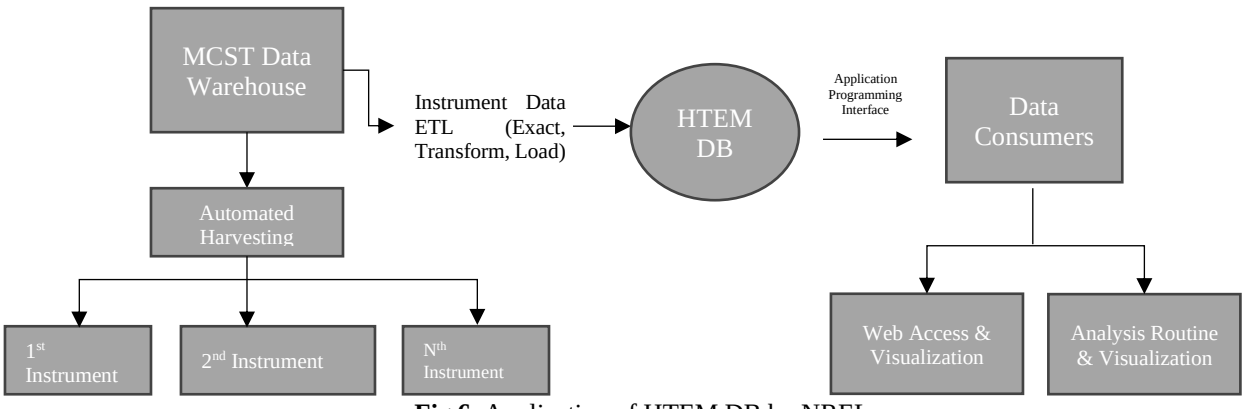

Fig 6. Application of HTEM DB by NREL

Fig 6 illustrates the use of HTEM DB, which is a unique Laboratory Information Management System (LIMS) built by NREL. This system has been created via extensive cooperation among computer scientists, materials researchers, programmers, and database architects over a significant period of time. Initially, the materials information is systematically collected from characterization and synthesis devices and stored in a data warehouse, which serves as a repository for metadata files and materials data. Subsequently, the ETL procedure harmonizes characterization and synthesis metadata and data, integrating them into the HTEM databank using an object-relational structure. An API is used to provide steady relation between client administration, such as the HTEM database, and statistical analysis tools, and online user interfaces.

The online user interface allows materials scientists worldwide, who lack access to specialized HT experimental equipment, to see a limited number of specific databanks at a time, even if they did not produce this information. For instance, the API in [24] provides computer experts with the ability to get a greater quantity of material databanks specifically for the purposes of ML and data mining. Both forms of access to the extensive and varied items information in HTEM DB have the potential to stimulate the ingenuity of researchers in unforeseen ways that are now challenging to predict. Additional technical information on the database architecture and the underlying system of managing data may be found in the Methods section.

Database Exploration and Application

Users primarily interact with HTEM DB via its specialized WBI located at htem.nrel.gov. Initially, the databank may be queried to get the example libraries of materials that include the desired components. Furthermore, the search results may be refined by applying other filters, such as completeness/data quality, synthesis conditions, and additional information. Ultimately, the refined search results may be shown and examined in an interactive manner, or saved for more comprehensive examination on the researcher's computer. To begin engaging with the HTEM DB, the first step is to do a search for the SLs pertaining to the desired substance. Customers could choose elements from the periodic table on the 'Search' page by using either the 'all' or 'any' search options. To obtain search results, the 'All' search option necessitates the existence of all designated items and maybe extra components. The 'any' search option allows for the retrieval of results as long as at least one of the chosen elements is present. Additional sophisticated search logic is organized for implementation in the future.

On the search page and the following pages, you may find comprehensive information on HTEM DB, including details about its purpose, statistics, and the application programming interface (API). Users are advised to constantly check the latest information on HTEM DB, as opposed to relying on this document, in order to stay informed about any changes or updates. After doing the search, the 'Filter' page displays a list of SLs that satisfy the search parameters. It also provides a sidebar for further refinement of the results. The SLs can be viewed in three different ways, each providing an increasing amount of information: compact view (including measured properties, data quality, database sample ID, and included elements), detailed view (including measurement/ synthesis date, sample number, deposition chamber, and the person responsible for generating the sample), and complete view (including synthesis parameters like power/ targets, flows/ gases, temperature/ substrate, time, and pressure). The introduction of the five-star data quality scale, with a 3-star rating for uncurated data, aimed to empower users to choose their preferred trade-off between data quantity and quality during analysis. All descriptors, including this one, may be used to categorize the search results or narrow them down by utilizing the sidebar located on the page`s left side. Furthermore, the user has the option to manually choose or deselect the sample library by using the checkboxes located on the page`s right side.

Once the filtering process is finished, users can utilize the 'Share Filter' option to share the search results with other users, creating a special web link. Additionally, the characteristics of the chosen SLs can be examined by pressing the 'Visualize' button. To prevent lengthy loading times, it is advisable to restrict the number of displayed SLs to fewer than 10-20 due to the abundance of data currently accessible. The product view on the 'Visualize' tab of the HTEM DB consists of three customizable charts per chosen SL. At the 'Visualize' page, you may specify the point size, y-axis, color scale, and x-axis for each of these three sets of plots using the drop-down choices. There are options to display both linear and logarithmic graphs. Right now, the scalar adjustable that can be schemed in this way are the sample thickness, chemical identify of each essential element, x/y correlates on the SL, and a variety of properties like electrical (standard deviation, conductivity, resistivity, sheet resistance, etc.), optical (mean visible transmittance, BG, etc.), and structural (picks` number in XRD pattern, etc.).

There are plans to develop multi-library plots, which will allow for the comparison of numerous sample libraries on a single plot, in the future. The Properties' page spectra view is organized similarly to the properties view. It displays plots that represent structural information, such as optical properties, and x-ray diffraction patterns, like absorption spectra. These plots are available for each of the selected SLs and can be displayed on either a logarithmic or linear scale. At now, the user is unable to set the axes of these plots. However, it is intended to include this capability in the future. Users may obtain thorough details about the SL from the clarify page by pressing the 'Library Details' button on the 'Properties' or 'Spectra' sections of the 'Visualize' page. A page that offers a brief synopsis of the sample library, complete with details, characteristics, and spectra, is displayed to users upon selecting the 'Library Summary' button.

Fig 7 displays an example of the spectra plots and property found in the HTEM DB for a specific sample library of Zn-Ni-Co-O. Panel (a) represents the thickness (size) and composition (color) of the sample at different positions (x and y) in the library. Panel (b) represents the direct bandgap (DBG) (size) and conductivity (color) of the sample at the same positions. Panel (c) displays a summary analysis figure that shows the relationship between the logarithm of conductivity and

composition. The color scale represents the DBG, while the point size represents the thickness. Panel (d) represents the x-ray diffraction patterns, whereas panel (e) represents the optical absorption spectra.

Fig 7 exhibits a segment of the library summary view pertaining to a specific Zn-Ni-Co-O SL. Based on the XRF data (see **Fig 7a**), there is a consistent linear variation in the zinc content throughout the sample library, as shown by the color scale. However, the thickness of the samples remains relatively constant across the library, as shown by the point size. The conductivity, as shown in **Fig 7b**, varies across the sample library, as shown by the color scale. However, the DBG stays rather consistent, as represented by the point size. By combining these two databanks (see **Fig 7c**), it is evident that the transmission, shown on a logarithmic axis, achieves its highest value of 4 S/cm at a Zn constitution of 33%. This relationship holds true independent of the direct band gap (represented by the color scale) and the density (represented by the point size). The optical absorption spectra and X-ray diffraction (XRD) patterns may be shown as well (**Fig 7d** and **Fig 7e**, respectively). This example demonstrates the use of the 'Visualize' tab to investigate and evaluate the information in HTEM DB.

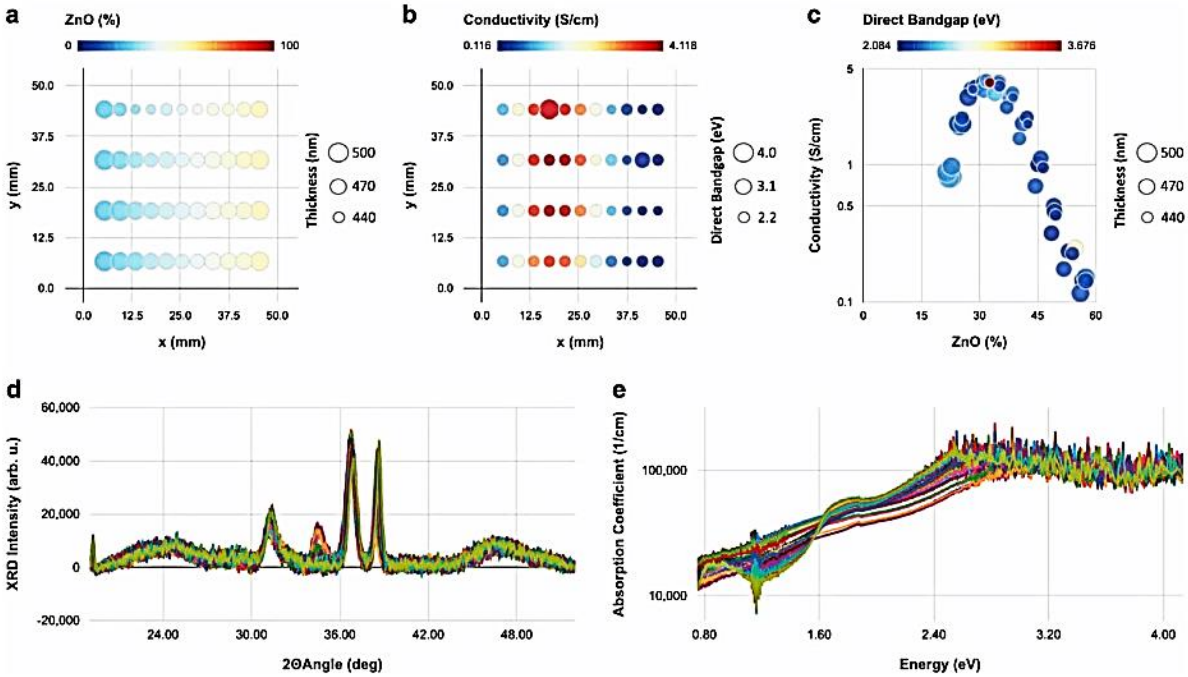

Fig 7. An illustration of the property and spectrum plots for a particular Zn-Ni-Co-O sample library that may be found in the HTEM DB

Additional categories of spectra and features will be included into the databank in the future. The data analysis capabilities of the HTEM DB online interface are restricted to generating basic spectra and property plots for a limited number of data collections, as previously mentioned. To conduct a more comprehensive study beyond the capabilities of the HTEM DB online interface, users have the option to download the datasets they are interested in [25]. They may then analyze these datasets using their preferred software programs such as Python, MatLab, Igor, Mathematica, and R, etc. The task may be achieved by using 'Download' buttons got on the spectra and property pages (with future implementation on the filter page as well), which save the information in a simple CSV format. To facilitate study of a greater quantity of information, it is recommended that users use a comprehensive and accessible API. This API provides data in many widely-used materials-specific formats, like Citrine's MIF format and NIST's MDCS format. It is advisable to use API data access instead of downloading the data to get the most up-to-date findings. The HTEM DB content is frequently updated to include extra data, address data processing errors, and curate data based on its consistency.

VI. CONCLUSIONS

Quantum computing and quantum algorithms have the capability to greatly accelerate the solution of certain choice problems that are represented by decision trees, resulting in exponential speedup. Quantum game theory and quantum communication may be used in multi-agent systems and distributed AI to augment cooperation and decision-making. Quantum machine learning integrates the advantages of quantum computing, such as simultaneous processing and quantum entanglement, with machine learning algorithms to accelerate training and inference methods. Quantum-inspired algorithms for search area exploration use superposition, interference, and entanglement more effectively than classical optimization approaches. Machine learning approaches have proven successful in the realm of materials research for data analysis and process automation. Nevertheless, the use of experimental machine learning techniques encounters difficulties in the realm of material identification owing to the scarcity of extensive and diverse datasets. The HTEM Database seeks to overcome this

limitation by providing a publicly accessible compilation of experimental data for inorganic compounds obtained via high-throughput practical methods.

The database architecture and content are constructed by conducting high-throughput testing and collecting data utilizing materials data infrastructure. The HTEM Database has comprehensive documentation on thin film inorganic materials, including detailed information on their crystal structure, chemical composition, synthesis circumstances, and optoelectronic property evaluations, among other aspects. The NREL LIMS streamlines the process of collecting, categorizing, and storing data, hence supporting chemicals research, and aiding the HTEM Database. Users may use a dedicated web interface to engage with the HTEM Database, which facilitates streamlined searching, filtering, and presentation of information. The HTEM Database is an invaluable resource for scholars interested in high-throughput experimental materials. The system provides enough data, simplifies searching and analysis, and creates opportunities for further study and growth. The database undergoes continual updates to address data processing difficulties and include new data. The HTEM Database plays a crucial role in developing materials research and the use of machine learning methods in these domains.

Data Availability

No data was used to support this study.

Conflicts of Interests

The author(s) declare(s) that they have no conflicts of interest.

Funding

No funding was received to assist with the preparation of this manuscript.

Competing Interests

There are no competing interests.

References

[1]. Y. Anshori, A. Y. E. Dodu, and D. M. P. Wedananta, "Implementasi Algoritma Kriptografi Rivest Shamir Adleman (RSA) pada Tanda Tangan Digital," Techno.COM Jurnal, vol. 18, no. 2, pp. 110–121, May 2019, doi: 10.33633/tc.v18i2.2166.

[2]. C. Y. Chung, H.-L. Yu, and K. P. Wong, "An advanced Quantum-Inspired evolutionary algorithm for unit commitment," IEEE Transactions on Power Systems, vol. 26, no. 2, pp. 847–854, May 2011, doi: 10.1109/tpwrs.2010.2059716.

[3]. B. Bauer, S. Bravyi, M. Motta, and G. K. Chan, "Quantum algorithms for quantum chemistry and quantum Materials science," Chemical Reviews, vol. 120, no. 22, pp. 12685–12717, Oct. 2020, doi: 10.1021/acs.chemrev.9b00829.

[4]. G. Ramsey and A. De Block, "The Dynamics of Science: Computational Frontiers in History and Philosophy of Science," University of Pittsburgh Press, Jan. 2020.

[5]. J. Brandts and G. Charness, "The strategy versus the direct-response method: a first survey of experimental comparisons," Experimental Economics, vol. 14, no. 3, pp. 375–398, Jan. 2011, doi: 10.1007/s10683-011-9272-x.

[6]. D. Ventura and T. Martinez, "A quantum Computational learning algorithm," arXiv (Cornell University), Jul. 1998, [Online]. Available: https://arxiv.org/pdf/quant-ph/9807052v1

[7]. S. Arunachalam, A. B. Grilo, T. Gur, I. C. Oliveira, and A. Sundaram, "Quantum learning algorithms imply circuit lower bounds," 2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS), Feb. 2022, doi: 10.1109/focs52979.2021.00062.

[8]. A. M. Childs and E. Farhi, "Quantum information processing in continuous time," 2004. [Online]. Available: https://dspace.mit.edu/bitstream/handle/1721.1/16663/56558690-MIT.pdf;sequence=2

[9]. A. Leider, S. Siddiqui, D. A. Sabol, and C. C. Tappert, "Quantum Computer search algorithms: Can we outperform the classical search algorithms?," in Advances in intelligent systems and computing, 2019, pp. 447–459. doi: 10.1007/978-3-030-32520-6_34.

[10]. V. Moret-Bonillo, "Emerging technologies in artificial intelligence: quantum rule-based systems," Progress in Artificial Intelligence, vol. 7, no. 2, pp. 155–166, Jan. 2018, doi: 10.1007/s13748-017-0140-6.

[11]. V. Vijayakrishnan and S. Balakrishnan, "Correspondence between quantization schemes for two-player nonzero-sum games and CNOT complexity," Quantum Information Processing, vol. 17, no. 5, Mar. 2018, doi: 10.1007/s11128-018-1870-5.

[12]. I. Martínez-Martínez, "A connection between quantum decision theory and quantum games: The Hamiltonian of Strategic Interaction," Journal of Mathematical Psychology, vol. 58, pp. 33–44, Jan. 2014, doi: 10.1016/j.jmp.2013.12.004.

[13]. T. Q. Duong, J. A. Ansere, B. Narottama, V. Sharma, O. A. Dobre, and O. A. Dobre, "Quantum-Inspired Machine Learning for 6G: fundamentals, security, resource allocations, challenges, and future research directions," IEEE Open Journal of Vehicular Technology, vol. 3, pp. 375–387, Jan. 2022, doi: 10.1109/ojvt.2022.3202876.

[14]. C. Draxl and M. Scheffler, "Big Data-Driven Materials Science and its FAIR data infrastructure," in Springer eBooks, 2020, pp. 49–73. doi: 10.1007/978-3-319-44677-6_104.

[15]. D. Barseghian et al., "Workflows and extensions to the Kepler scientific workflow system to support environmental sensor data access and analysis," Ecological Informatics, vol. 5, no. 1, pp. 42–50, Jan. 2010, doi: 10.1016/j.ecoinf.2009.08.008.

[16]. N. Bhati, M. K. Nazeeruddin, and F. Maréchal, "Critical analysis of decision variables for high-throughput experimentation (HTE) with perovskite solar cells," Solar Energy, vol. 262, p. 111810, Sep. 2023, doi: 10.1016/j.solener.2023.111810.

[17]. Y. Hao et al., "Dimeric Product of Peroxy Radical Self-Reaction Probed with VUV Photoionization Mass Spectrometry and Theoretical Calculations: The Case of C2H5OOC2H5," International Journal of Molecular Sciences, vol. 24, no. 4, p. 3731, Feb. 2023, doi: 10.3390/ijms24043731.

[18]. G. Kuri, H. Ramanantoanina, J. Bertsch, M. Martin, and I. Panas, "Chemical state and atomic scale environment of nickel in the corrosion layer of irradiated Zircaloy-2 at a burn-up around 45 MWd/kg," Corrosion Science, vol. 143, pp. 200–211, Oct. 2018, doi: 10.1016/j.corsci.2018.08.032.

[19]. A. J. Williams et al., "The CompTox Chemistry Dashboard: a community data resource for environmental chemistry," Journal of Cheminformatics, vol. 9, no. 1, Nov. 2017, doi: 10.1186/s13321-017-0247-6.

[20]. A. Zakutayev et al., "An open experimental database for exploring inorganic materials," Scientific Data, vol. 5, no. 1, Apr. 2018, doi: 10.1038/sdata.2018.53.

[21]. J. Winkelmann, "Introduction: Data extract from Landolt-Börnstein IV/15A: Diffusion of Gases in Gases, Liquids and their Mixtures," in Numerical data and functional relationships in science and technology, 2007, pp. 1–8. doi: 10.1007/978-3-540-49718-9_1.
[22]. H. Yoshihama and H. Kaneko, "Design of thermoelectric materials with high electrical conductivity, high Seebeck coefficient, and low thermal conductivity," Analytical Science Advances, vol. 2, no. 5–6, pp. 289–294, Dec. 2020, doi: 10.1002/ansa.202000114.
[23]. A. Belsky, M. Hellenbrandt, V. L. Karen, and P. Luksch, "New developments in the Inorganic Crystal Structure Database (ICSD): accessibility in support of materials research and design," Acta Crystallographica, vol. 58, no. 3, pp. 364–369, May 2002, doi: 10.1107/s0108768102006948.
[24]. I. Kavakiotis, O. Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas, and I. Chouvarda, "Machine learning and data mining methods in diabetes research," Computational and Structural Biotechnology Journal, vol. 15, pp. 104–116, Jan. 2017, doi: 10.1016/j.csbj.2016.12.005.
[25]. L. Tauxe et al., "PmagPy: Software package for paleomagnetic data analysis and a bridge to the Magnetics Information Consortium (MagIC) Database," Geochemistry Geophysics Geosystems, vol. 17, no. 6, pp. 2450–2463, Jun. 2016, doi: 10.1002/2016gc006307.