From the Rock Floor to the Cloud: A Systematic Survey of State-of-the-Art NLP in Battery Life Cycle

Tosin Adewumi[k*] **Martin Karlsson,**[b] **Marcus Liwicki,**[k] **Mikael Sjödahl,**[c] **Lama Alkhaled,**[k] **Rihab Gargouri,**[k] **Nudrat Habib,**[k] **and Franz Hennies**[d]

Keywords: Keywords: *Battery*, *TLP*, *NLP*, *Survey*

Abstract: We present a comprehensive systematic survey of the application of natural language processing (NLP) along the entire battery life cycle, instead of one stage or method, and introduce a novel technical language processing (TLP) framework for the EU's proposed digital battery passport (DBP) and other general battery predictions. We follow the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method and employ three reputable databases or search engines, including Google Scholar, Institute of Electrical and Electronics Engineers Xplore (IEEE Xplore), and Scopus. Consequently, we assessed 274 scientific papers before the critical review of the final 66 relevant papers. We publicly provide artifacts of the review for validation and reproducibility. The findings show that new NLP tasks are emerging in the battery domain, which facilitate materials discovery and other stages of the life cycle. Notwithstanding, challenges remain, such as the lack of standard benchmarks. Our proposed TLP framework, which incorporates agentic AI and optimized prompts, will be apt for tackling some of the challenges.

1. Introduction

In the fast-advancing field of natural language processing (NLP) and its many growing application areas, including battery production, new tasks are evolving. A battery is an energy storage system (ESS) that is capable of supplying electrical energy and is usually made of electrochemical substances (Mitali et al., 2022; Winter & Brodd, 2004). Meanwhile, NLP is a field with methods and tools for machines to analyze, understand and, possibly, generate natural human language (Eisenstein, 2019; Liddy, 2001). The battery production life cycle, as in many other fields, involves humans who communicate in natural language and store records of some of the communication about the entire complex value chain. Both successful and unsuccessful attempts along the value chain are usually documented in structured or unstructured reports. Indeed, the increasing storage demand because of sustainability is leading to

[k]**Machine Learning Group, Luleå University of Technology (LTU), Sweden**
[b]**Research Center for Advanced Battery Technology, LTU, Sweden.**
[c]**Luleå University of Technology (LTU), Sweden.**
[d]**Lund University (LU), Sweden.**
[*]**Email: tosin.adewumi@ltu.se**

increased research and documentation involving the popular Li-ion battery (LiB), among others, with liquid electrolytes, known for their conductivity, to possible replacement by non-flammable solid-state electrolytes (SSE) (Mahbub et al., 2020; Shon & Min, 2023). The documentations create a wealth of information and may include technical manuals, production logs, patents, scientific literature, standards, quality reports, maintenance records, digital battery passport (DBP), and more. Therefore, NLP and technical language processing (TLP), which addresses a specific domain's technical terms (Brundage et al., 2021; Dima et al., 2021; Löwenmark et al., 2023), have the capacity to lead to dramatic speed-up in battery innovation, including materials discovery, leading to shorter innovation time. This is more so that ESSs are very important for a sustainable future since renewable energy systems (RESs) depend on natural resources for generation, which are limited in quantity or are seasonal (e.g. sunshine or wind) (Mitali et al., 2022).

Many factors are important for the improvement of battery production and the understanding of battery behavior. These include both qualitative and quantitative information on material composition, environmental conditions, understanding of the chemical reactions, and mass transport (Lyonnard et al., 2025). Experimental methods, which are usually effective, have their drawbacks. For example, they can be resource-intensive, time-consuming, and involve trial-and-error studies (Zuo et al., 2025; Mahbub et al., 2020). Therefore, researchers have observed that incorporating NLP in parts of the battery production pipeline can facilitate relevant parts of the value chain, especially as there are thousands of articles on battery research from which to mine (Mahbub et al., 2020). Consequently, as a motivation, this study has the **main objective of a comprehensive survey of the application of NLP along the entire value chain of a battery life cycle, from material sourcing to repurposing or end of life**. The research question we address is what are the NLP tasks along the entire value chain of a battery life cycle? To address the question, we followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method (Page et al., 2021) for systematic survey of the literature between 2017 and 2025 using 3 reputable databases or search engines: Google Scholar, Institute of Electrical and Electronics Engineers Xplore (IEEE Xplore), and Scopus. As a result, we assessed 274 papers before our final list of 66 papers for our findings, where we (1) map the field, (2) compare methods, (3) identify challenges, and (4) introduce a novel TLP framework for battery-related predictions.

Contributions

This work provides a suitable entry point for newcomers to the field of NLP applications in the battery domain and an overview of the intersection of NLP and battery science. Our main contributions are there-fold and address gaps in the existing literature. They include:

1. We provide, to the best of our knowledge, the first comprehensive systematic survey of SotA NLP across the whole value chain of a battery life cycle.
2. We curate and synthesize important comparative information about the results, models, data, and existing challenges from many different studies.
3. We introduce a comprehensive yet simple and novel TLP framework for digital battery passport (DBP) and general battery predictions, which combines the capabilities of artificial intelligence (AI) agents, large language models (LLMs), and optimized hard and soft prompts.

The rest of this paper is structured as follows. In the next section (Section 2), we discuss related work that have surveyed NLP applications in the battery domain, the basics of NLP and its traditional tasks, and the battery life cycle, from material sourcing to recycling or repurposing. In Section 3, we describe the methodology, including the search databases, criteria, and the PRISMA method and chart. In Section 4, we present our findings in a clear and structured way, comparing and describing some important works across the value chain. In Section 5, we describe ongoing challenges and introduce our novel TLP framework for battery predictions. Finally, in Section 6, we conclude with the main insights from the study and possible future work.

2. Background

This section is intended to provide background information on related work that have surveyed NLP applications in the battery domain and the gaps they have. It provides the basics of NLP for battery engineers, who may be new to the field, and 11 of its traditional tasks in relation to the battery domain. The section also discusses the battery life cycle stages, which involve about 17 stages, from material sourcing to recycling or repurposing.

2.1. Related Work

There have been several attempts at surveying NLP in the battery domain. They are, however, limited to specific parts of the battery life cycle. For example, the work by Jiang et al. (2025) reviewed NLP application in materials science, with a focus on information extraction (IE) and materials discovery. They observed that within materials science, NLP is still in the early stages but the trend towards the development of material-specific pretrained language models is growing. Zuo et al. (2025) surveyed the applications of LLMs, addressing two questions of what LLMs offer to support battery-related tasks and how more effective LLMs may be developed? In contrast, our work offers a more comprehensive survey by including discriminative models. In the review by Pievaste et al. (2025), they highlighted macro-scale property prediction for the properties and performance of bulk materials in real-world applications, including multi-physics performance of a battery, as an example. The survey centered on machine learning (ML) methods for materials science, including traditional algorithms and deep learning architectures. Pei et al. (2025) reviewed the literature, examining how NLP facilitates the understanding and design of novel materials in materials science. The focus of their review was on the limitations of LLMs, the creation of a materials discovery pipeline, and the potential of generative pretrained transformer (GPT) models to synthesize existing knowledge for sustainable materials. Osaro et al. (2025) surveyed materials discovery and chemical synthesis, observing that AI-driven frameworks use existing compounds to predict new battery electrode materials with optimized properties for conductivity and stability. Kim et al. (2025) reviewed R-based energy forecasting for different energy sources instead of the more common Python language, including a brief mention of batteries. Lin & Chou (2025) conducted a systematic review of AI in patent analysis, discussing works that used principal component analysis (PCA) and random forest (RF).

Earlier works, such as Singh et al. (2022), surveyed deep learning approaches in NLP for battery materials. They identified classification, IE, summarization and materials discovery as some of the tasks or applications of NLP in the battery domain. In their perspective paper, Zhao et al. (2024) discussed the potential for insights on battery research from LLMs, with particular attention on fast charging and ChatGPT. Clark et al. (2022) focused on the state of ontology development, emphasizing the need for such in battery research and production. Lee et al. (2023) conducted a short review on NLP for materials discovery. They observed that the number of scientific papers from different databases that are used vary significantly across different research studies, lacking standardization, and sometimes involved the use of optical character recognition (OCR) for older documents. A review of LiB state-of-charge (SoC) estimation using deep learning was carried out by Liu et al. (2022), where they classified the methods into two: structured adjustment and unstructured improvement. They observed that LiB is currently the dominant battery type and identified the usual data acquisition process for training and verification datasets during a hypothetical driving cycle. Their review was limited to the SoC and some deep neural networks, especially recurrent neural networks (RNNs). A comprehensive survey along the entire value chain, such as our work, can draw lessons from one stage to benefit other stages of the value chain.

2.2. NLP Basics

NLP has advanced considerably over the decades from simple rule-based methods for machine translation (MT) tasks to deep neural network (NN) and LLMs for general tasks. The general NLP pipeline for solving a task involves the iterative steps of (1) data acquisition, (2) preprocessing (e.g. removal of unwanted characters), (3) tokenization (e.g. splitting into words), (4) model selection, (5) model training, (6) validation and hyper-parameter tuning, (7) testing, and (8) deployment (Lane & Dyshel, 2025; Srinivasa-Desikan, 2018). The process of model training with NNs involves an important step of embedding creation, where the tokenized data is transformed into numerical representation (or embeddings) of a lookup table in relatively small dimension (Lane & Dyshel, 2025). The NN model training is based on minimizing a loss function (e.g. cross-entropy loss) to maximize a given utility function by comparing the model predictions to ground truth labels, in supervised learning, and backpropagating to iterate the procedure until a convenient validation loss is obtained so that the model can generalize to unseen data at test time. This pipeline is based on splitting the data into 3 parts - training, validation and test splits. The procedure is called unsupervised learning when there are no labels to compare predictions with and it must only learn patterns within the data. These two paradigms form the ends of the learning spectrum with variants in between, including reinforcement learning (RL). LLMs undergo self-supervised learning as pretraining before undergoing different types of post-training and deployment for inference.

Models have become more complex and deep from the shallow, non-contextual networks of the past, such as Word2Vec (Mikolov et al., 2013). Important concepts from past models can be found in recent models (e.g. embeddings in the early layers of deep models). Recent state-of-the-art (SotA) models, like LLMs, are based on the SotA Transformer architecture, which has both an encoder and decoder (Vaswani et al., 2017). For efficiency, researchers decouple the encoder for discriminative language understanding tasks (e.g. text classification (TC)) while the decoder is used for language generation (or generative) tasks (e.g. summarization) though it is also suitable for discriminative tasks. Examples of discriminative models include bidirectional encoder representations from transformers (BERT) and its many variants (Devlin et al., 2019; Liu et al., 2019; He et al., 2021) while examples of generative models include the GPT series (Brown et al., 2020; OpenAI et al., 2024). Some of these models are used in battery-related research and predictions, as will be discussed later. For example, GPT4 has been used in chemical named entity recognition (NER) and other tasks by Lee et al. (2024a). Qwen2.5-7B and Gemma2-9B are LLMs used in Zheng et al. (2025). Below, we identify some traditional tasks of NLP, i.e. those before the advent of LLMs, where the first 6 are categorized as natural language understanding (NLU) tasks and the remaining as natural language generation (NLG) tasks, except the last one, which is an unsupervised learning task. NLU tasks are usually evaluated with metrics like accuracy or F1 score and NLG tasks with n-gram-based metrics like ROUGE (Lin, 2004; Gehrmann et al., 2021) or semantic-based models like BERTScore (Zhang et al., 2020). Noteworthy that despite the improvements in NLP, challenges, such as hallucinations and biases, still exist in the field (Adewumi et al., 2025a; Pettersson et al., 2024).

Some Traditional NLP Tasks

1. TC: It is the general activity of automatically labeling natural language texts with thematic categories from a predefined set (Sebastiani, 2002). A variety of standard definitions of TC tasks exists in NLP. Some examples are sentiment analysis (SA), author attribution, news classification, and NER (Gasparetto et al., 2022; Habib et al., 2025). Each of these sub-tasks have many datasets available and they are widely used in research. In battery research, TC is used to automatically select battery related documents from a large collection, making literature screening and dataset building faster. For example, BatteryBERT was fine-tuned on labeled abstracts to classify whether a paper is battery-related, enabling more focused materials data mining (Huang & Cole, 2022a).

2. NER: It seeks to identify and classify mentions of specific entities often referred to as rigid designators into predefined semantic categories such as chemical compounds, persons, locations, organizations, and others (Adelani et al., 2021). NER enables structured extraction of key information from unstructured text, supporting tasks like search, summarization, and knowledge graph construction. Commonly used datasets include CoNLL-2003, OntoNotes 5.0, and W-NUT (Hu et al., 2024; Tjong Kim Sang & De Meulder, 2003; Derczynski et al., 2017). NER can be used to extract chemical and material names from battery literature to support the automatic generation of structured datasets or to identify materials, synthesis descriptions, and phase labels that are relevant to battery research from large collections of scientific abstracts (Huang & Cole, 2022b; Weston et al., 2019).
3. Part-of-Speech (PoS) tagging: It is the process of automatically assigning each word in a sentence a grammatical category such as noun, verb, adjective, or adverb, based on its definition and context within the sentence (Toutanova et al., 2003). It is a core NLP task that helps models to understand sentence structure, supporting downstream applications like parsing, NER, and IE. Some of the datasets include part of the Penn Treebank and AfricaPOS (Dione et al., 2023; Marcus et al., 1993). While PoS tagging improves contextual understanding, it can struggle with word ambiguity, domain-specific terms, and noisy technical text (Manning, 2011).
4. IE: It refers to the automatic identification and organization of specific types of information, like entities, relationships, and events from unstructured or semi-structured text, transforming natural language into structured, machine-readable data (Grishman, 1997). IE can help battery research and production by extracting key data, like capacity, voltage, and efficiency, from text and turning them into structured formats. This makes it easier to build battery databases, compare performances, and spot useful trends for material and design choices (Huang & Cole, 2022b). An example dataset includes POLYIE (Cheung et al., 2024).
5. SA: It is a TC task of automatically detecting and classifying emotions or opinions in text. This is typically positive and negative or may contain neutral (Maas et al., 2011). It plays a key role in applications such as customer feedback analysis, social media monitoring, and product review mining. It enables users to extract information from textual data automatically to support fast decisions about customer satisfaction or product improvement, even within the battery value chain (Wankhade et al., 2022). Example datasets include Yelp, internet movie database (IMDb), and movie review datasets (Minaee et al., 2021; Maas et al., 2011).
6. Question Answering (QA): It is a task involving systems that can understand a natural language question and automatically return an accurate and relevant answer, either from a given context (Extractive QA) or based on previously learned knowledge (Abdel-Nabi et al., 2023). This task is particularly useful in education (Adewumi et al., 2025b; Pettersson et al., 2024). QA systems allow users to retrieve precise information quickly without reading full documents, even in the battery domain, making them highly useful for time-sensitive tasks and large-scale knowledge access. Examples of datasets include Stanford Question Answering Dataset (SQuAD) 2.0 and Conversational Question Answering Challenge (CoQA) (Yatskar, 2019).
7. Summarization: It aims to produce a concise and coherent summary of a document or documents that captures the main ideas of the source text while reducing its length (El-Kassas et al., 2021). Text summaries of battery documents can lead to more efficient workflows. Like with many generative tasks that have different objective metrics to evaluate the tasks, human evaluation is usually the gold standard, though subjective. Datasets include BOOKSUM, ArXiv, and PubMed (Kryscinski et al., 2022; Cohan et al., 2018).
8. MT: It is a process of automatic conversion of text from one natural language to another (Jiang & Lu, 2020). Since this helps in allowing people to access content in different languages without human intervention, even in cross-border partnerships in the battery domain, it is useful in communication and multilingual applications. While MT can speed up information access, it's challenging when it concerns complex languages, idioms, and other linguistic issues (Wang et al., 2022; Adewumi et al., 2022b). Datasets include those from the Workshop on Statistical Machine Translation (WMT) and FLORES (Guzmán et al., 2019; Bojar et al., 2014).

9. Automatic speech recognition (ASR): This is also known as speech-to-text (STT) and is the process of converting spoken language into written text using computational models (Kheddar et al., 2024). ASR is useful in applications like voice assistants and transcription services as it helps users interact with machines more naturally. Challenges include dealing with background noise, accents, and overlapping speech (Benzeghiba et al., 2007). Datasets include LibriSpeech and Google-Sc, among others (Kheddar et al., 2024).
10. Dialog Generation: It is the single or multi-turn conversations generated by models communicating with humans or another model (Adewumi et al., 2022a). The task can be divided into task-oriented or open-domain dialog generation. LLMs are used in open-domain dialog generation, including materials dialog, and are being adapted in task-oriented dialog systems also (e.g. flight bookings) (Lee et al., 2024b). Phenomenal improvements in dialog coherence and other features have been witnessed in the past few years in open-domain dialog generation. However, some limitations like hallucinations still exist (Adewumi et al., 2024c). A dataset example includes the Ubuntu dialog dataset (Adewumi et al., 2022a).
11. Topic modeling (TM): It learns patterns in the data in an unsupervised manner to discover latent topics and infer topic proportions of documents (Wu et al., 2024). It can be used to support other tasks for better results. Two common challenges with topic modeling are trivial topics, which are based on uninformative words, and repetitive topics, where synonyms exist. An example of a dataset for the task is 20newsgroup (Wu et al., 2024).

2.3. Battery Life Cycle

The increase in RESs and electric vehicles (EVs) has made the battery production industry one of the fastest-growing industries (Nekahi et al., 2024). The LiB has attracted particular interest because of its high energy, high power density, long cycle life, and potential across electronic mobility and stationary storage. Battery production is a complex process that involves many steps, from material sourcing, mining and refining raw materials to assembling cells together into finished battery packs (Xiao et al., 2025; Örüm Aydin et al., 2023; Nekahi et al., 2025). During every step of the production process, specific physical and chemical processes must be followed, supported by quality assurance and environmental controls. These processes generate large volumes of textual data throughout the entire life cycle, which can be leveraged through NLP for process analysis, compliance, and optimization (Lee et al., 2025; Jiang et al., 2025). Figure 1 depicts the stages of the battery life cycle. We categorize and discuss each of the stages briefly below.

2.3.1. Material sourcing & Mixing

Material sourcing is the foundation of the entire production chain. Cobalt, nickel, manganese, graphite, and lithium are key raw materials, each with unique electrochemical properties. Mining and refining operations are required to achieve the purity levels needed for batteries (Xiao et al., 2025). For example, lithium comes from hard rock mineral deposits (e.g. spodumene, $LiAl(SiO_3)_2$) and brines pools (e.g. in South America). The refining process, such as acid roasting and precipitation, then follows to produce lithium carbonate or hydroxide. Material sourcing significantly impacts both environmental outcomes and production costs (Degen et al., 2023). It also generates extensive documentation, such as mining reports, safety sheets, and procurement contracts. These sources of documentation coupled with the wealth of information on the internet make NLP tasks like materials discovery, question answering (QA), and information retrieval (IR) through deep analysis of the available information an important component of this stage, and possibly the first, to guide informed decisions before **material mixing**, where conductive additives and polymer binders are homogenously dispersed in a solvent (usually water or N-methyl-2-pyrrolidone) to create a homogenous slurry (Lee et al., 2025).

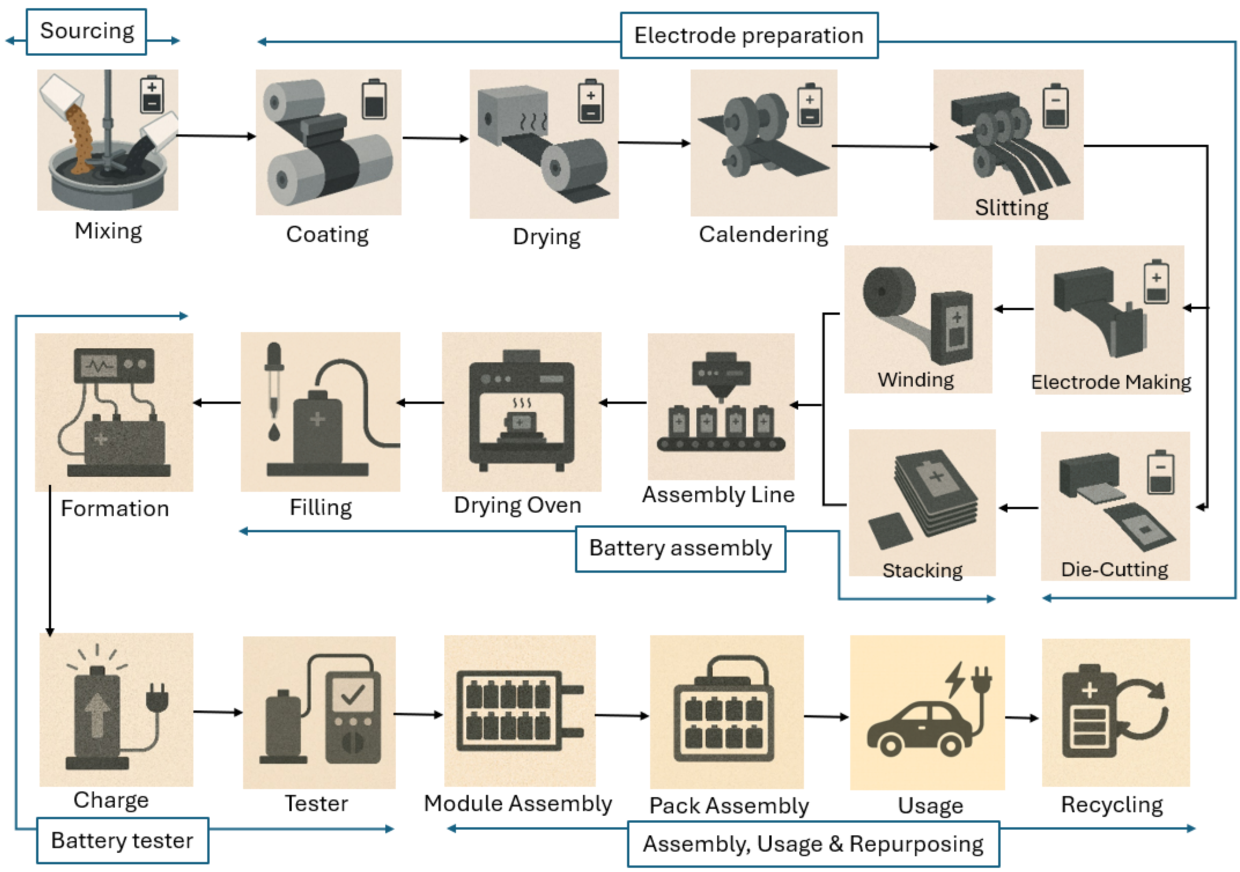

Figure 1. Battery life cycle (Images are ChatGPT-generated).

2.3.2. Electrode preparation (front stage process)

Electrode preparation, which converts refined powder into active electrodes, is the most costly and time-consuming stage of the manufacturing process (Nekahi et al., 2024; Örüm Aydin et al., 2023). Two main approaches exist to process the electrode fabrication step: solvent-based wet processing (which currently dominates) and solvent-free dry processing. Dry processing is gaining attention due to its lower energy use and environmentally friendly process. The main sub-steps are: **coating**, in which the slurry is applied onto a metallic foil (aluminum or copper) while maintaining precise control of the thickness, **drying**, which removes solvent and forms a solid film, representing one of the most energy intensive steps in the process, **calendering**, where the electrodes are compressed between hard rollers to improve contact and mechanical integrity, **slitting**, where wide rolls are cut into narrow electrode strips, **electrode making and die-cutting**, where electrode making is for producing both cathode and anode sheets and die-cutting is for trimming the coated foils into specific geometries, such as cylindrical, prismatic, or pouch shapes (Degen et al., 2023). Each sub-step requires optimization to balance electrochemical performance with mechanical robustness. The process documentations, including mixing ratios, coating logs, temperature and humidity records, and inspection reports, serve as rich textual sources for NLP-driven process optimization and predictive quality control (Liu et al., 2021). At the laboratory scale, various coating methods are available, including spray coating, spin coating, dip coating, comma-bar coating, ink-jet printing, electrophoretic deposition, doctor blading, and slot-die coating (Degen et al., 2023).

2.3.3. *Battery Assembly*

In the assembly stage, electrodes and separators are combined into complete cells (Attia et al., 2025). During **winding and stacking**, depending on the type, electrodes can be wound (for cylindrical cells) or stacked (for prismatic or pouch cells). The anode, separator, and cathode layers are rolled into a tight 'jelly roll' to ensure uniform ion transport. Stacking arranges sheets in alternating layers, enabling high packing density and consistent performance (Degen et al., 2023). **Assembly line** ensures electrode tabs are welded to terminal leads, and the assembled cell goes through the **drying oven** to remove residual moisture, that is, the solvent and water from the cathodes and anodes, respectively. Finally, **electrolyte filling** introduces a lithium salt solution into the porous structure in a controlled atmosphere. Each step produces operational and safety logs that can be analyzed using NLP to detect anomalies or optimize production sequences.

2.3.4. *Battery tester*

This stage initiates the cell's electrochemical activity through controlled charge and discharge cycles, which generate the solid electrolyte interphase (SEI) and cathode electrolyte interphase (CEI). These layers are critical for cell stability and safety (Schomburg et al., 2024; Weng et al., 2023). **Formation** generally occurs over several days under constant current, constant voltage (CCCV) conditions. After formation, cells undergo a series of test, including impedance measurement, self-discharge assessment, and safety evaluations, to verify performance consistency and compliance with quality standards (Liu et al., 2021). Generally, at the **charge** stage, the charging method determines battery capacity (Huang & Cole, 2020). Methods include CCCV and the cutoff voltages. The **tester** is used to qualify and verify the performance of the battery components to ensure they meet electrical and safety standards. Logs generated during the formation and testing stages are often stored in digital manufacturing systems, which again can serve NLP application purposes to improve yield and enable predictive maintenance.

2.3.5. *Assembly, Usage & Repurposing*

Validated cells are combined into modules in a **module assembly**. A typical battery module consists of several cells connected in series or parallel. Multiple modules are then combined in the **pack assembly**, which, like the module assembly, is equipped with thermal management and electronic control systems (Liu et al., 2021). In addition, the pack contains a battery management system (BMS), cooling components, and safety devices. Once assembled, the pack undergoes conditioning cycles to verify the functionality of sensors, wiring, and control electronics prior to deployment for **usage**. The safe use of LiBs relies on the BMS, which manages several parameters, including SoC of the batteries, for efficient function during the life of the batteries (Bian et al., 2024). The BMS also continuously monitors voltage and temperature to ensure safety and efficiency. The SoC indicates the capacity left in a LiB, thereby providing information against overcharging or undue discharging. Methods of estimating SoC may be divided into 4 categories: look-up table, Ampere-hour counting, model-based, and data-driven. The data-driven method has been gaining attention in recent times (Bian et al., 2024). **Recycling and repurposing** forms the final stage of the battery life cycle. Valuable materials such as lithium, nickel, cobalt, and copper are recovered through recycling and reintroduced into the supply chain to support a circular battery economy. Recycling processes and life cycle assessments produce a wide range of textual data, from dismantling instructions and recovery reports to environmental impact studies and compliance records (Lee et al., 2025). Batteries that cannot be recycled come to their end of life and must be disposed off properly or repurposed.

3. Methodology

We followed the PRISMA method (Page et al., 2021) for systematic reviews for a thorough and balanced survey to achieve the stated objective of this work. Figure 2 presents the PRISMA flow diagram for the review. It follows the rigorous and auditable general guidelines recommended for conducting a systematic literature review (Page et al., 2021; Adewumi et al., 2024a). We searched 3 reputable databases or search engines with relevant (inclusion) terms and focused on search results in English.

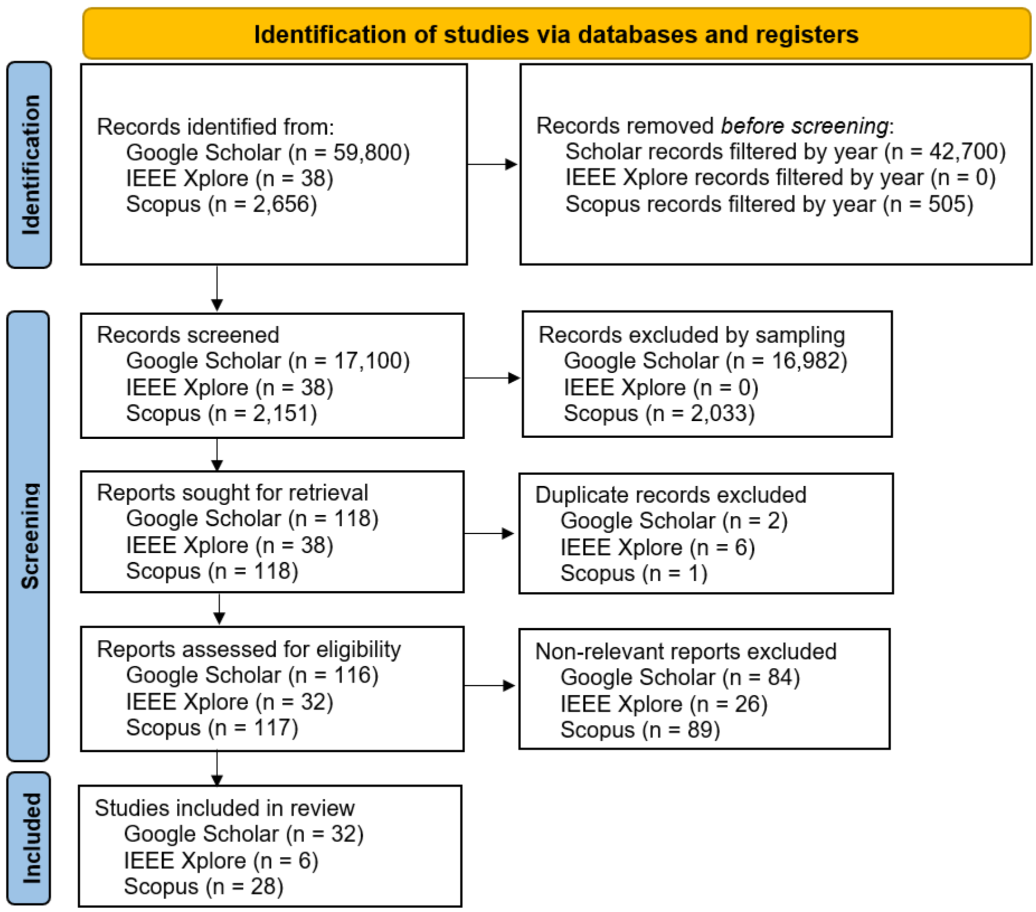

Figure 2. PRISMA flow diagram for the systematic review.

3.1. The Databases

The databases (or search engines) are Google Scholar, IEEE Xplore, and Scopus. They are suitable because they index the major publishers or databases. Each database complements the other. For example, IEEE Xplore and Scopus serve as additional validation for Google Scholar, which sometimes includes archived (or non-pair-reviewed) papers while Scholar has a wider coverage of results compared to the other two.

3.2. Search Criteria

The search terms (or keywords) we settled for are 3: 'text', 'language processing', and 'battery'. These are based on the inclusion criterion of the objective of the study. This is because NLP commonly involves written text, 'language processing' applies to natural language processing, technical language processing, and materials language processing, and battery applies to any stage of the battery life cycle. The search terms were combined using the AND operator on all the 3 databases as "text AND 'language processing' AND battery". We also attempted 4 combined terms: 'text', 'language', 'processing', and 'battery' but realized this gave a lot more non-relevant results than our earlier terms. Also, we chose not to include many other terms to avoid losing relevant articles since the objective was to cover the entire life cycle. Hence, we settled for our initial 3 search terms.

3.3. The PRISMA method

As depicted in Figure 2, the number of search results at every stage of the screening, including the year filter, are provided. Our choice of year filter (or cut off) from 2025 back to 2017 was because 2017 marked a pivotal period in NLP with the introduction of the Transformer architecture (Vaswani et al., 2017; Huang & Cole, 2022b). To decide the number of papers to review, we used sampling as a heuristic, though we're not randomly sampling from a population since the most relevant papers appear at the top or foremost pages. Given a 95% confidence level, 9% margin of error, and the biggest population of 17,100 from Google Scholar, we arrive at 118 as the sample size. Hence we used this value for Scopus also and reviewed all the 38 papers from IEEE Xplore since it returned a much smaller number in results. For the same reason we did not filter IEEE Xplore by year. Screening the titles and abstracts revealed papers that were not relevant. In situations where it was not immediately clear if a paper was relevant, we performed a quick automatic search for our keywords (especially 'battery') in the body of the paper. Furthermore, we observed that for some non-relevant papers, especially health-inclined ones, the term "battery of ..." or similar appears, where battery is used to mean something different, e.g. a series of things. For validation and reproducibility of our work, the search result links[1], downloadable Scholar list,[2] and the list of non-relevant papers are in the appendix.

4. Findings: NLP in the Battery Life Cycle

After the careful review of all the relevant papers from the search results, we categorized each paper (identified as DatabasePaperNumber) based on the NLP tasks they address, resulting in Table 1 while we synthesized some of the SotA results, where applicable, into Table 2. In Table 1, there are 16 tasks and about 8 are new or different from the 11 identified in Section 2.2. Some of the papers address multiple tasks. In Table 2, SotA results and models are compared for different datasets per task, where applicable. Different metrics are reported, depending on the task, and the category of each model is also identified. Some of the tasks identified in Table 1 have no results reported in Table 2 because the relevant paper in the earlier table is either a review paper or does not report a relevant SotA result. Hence, 11 main tasks are reported in Table 2 and the category of models are deep learning, classical, LLM, and knowledge graph (KG). In the following subsections, we discuss some details of the relevant papers based on their contributions.

[1] which are time-varying and sensitive to ranking algorithms
[2] https://drive.google.com/file/d/1iXNTIU7eLt-nSUteZnrl2o3Wq-LueMhL/view?usp=sharing
[3] From different journals
[4] Unclear from the paper.
[5] DSC - sample descriptors category

Table 1. Distribution of research papers across tasks. Review$_r$ and Position$_p$ papers are subscripted.

No.	NLP task	Google Scholar	IEEE Xplore	Scopus
1	TC	G2$_r$, G7, G8, G23, G25, G31,	I6	S5, S29
2	QA	G3$_p$, G4, G7, G8, G23		
3	Materials discovery	G2$_r$, G9$_r$, G10, G17, G22, G37$_r$, G47$_r$	I6	S2, S3$_r$, S7, S43$_r$, S46$_r$, S73$_r$, S78$_r$, S79, S81, S82, S100$_p$
4	IE	G1, G4, G7, G10, G11, G12, G19, G31, G38, G56, G57, G100		S1, S5
5	NER	G1, G4, G7, G8, G10, G23, G25, G49, G100	I19	S1
6	Summarization	G25		
7	Abbreviation detection	G4, G31		
8	SA	G75	I18, I27	S13, S18, S19, S53, S55, S67
9	Policy model consistency (PMC)			S49
10	Relational triple extraction (RTE)			S88
11	SoC estimation	G13, G40$_r$	I1	
12	Energy prediction		I26	S91$_r$
13	Recipe extraction and retrieval (RER)	G1, G5, G6		
14	IR	G2$_r$, G6, G28		S12, S80
15	TM	G1, G18, G89	I6	S1, S53, S80, S90, S92, S97
16	Ontology development	G15$_r$		
	Non-relevant papers	G14, G16, G20, G21, G24, G26, G27, G29, G30, G32, G33, G34, G35, G36, G39, G41, G42, G43, G44, G45, G46, G48, G50, G51, G52, G53, G54, G55, G58, G59, G60, G61, G62, G63, G64, G65, G66, G67, G68, G69=G71, G70=G72, G73, G74, G76, G77, G78, G79, G80, G81, G82, G83, G84, G85, G86, G87, G88, G90, G91, G92, G93, G94, G95, G96, G97, G98, G99, G101, G102, G103, G104, G105, G106, G107, G108, G109, G110, G111, G112, G113, G114, G115, G116, G117, G118	I2, I3, I4, I5, I7, I8, I9, I10, I11, I12, I13, I14, I15, I16, I17, I20, I21, I22, I23, I24, I25, I28, I29, I30, I31, I32=133=134 =135=136 =137=138	S4, S6, S8, S9, S10, S11, S14, S15, S16, S17, S20, S21, S22, S23, S24, S25, S26, S27, S28, S30, S31, S32, S33, S34, S35, S36, S37, S38, S39, S40, S41, S42, S44, S45, S47, S50, S51, S52, S54, S56, S57, S58, S59, S60, S61, S62, S63, S64, S65, S66, S68, S69, S70, S71, S72, S74, S75, S76, S77, S83, S84=S63, S85, S86, S87, S89, S93, S94, S95, S96, S98, S99, S101, S102, S103, S104, S105, S106, S107, S108, S109, S110, S111, S112, S113, S114, S115, S116, S117, S118

4.1. Material Discovery and Knowledge Mining

Apparently, much of the work in the literature and research has gone in the area of material discovery and knowledge mining (such as IE and NER), as shown in Table 1 in terms of the number of papers. He & Zhang (2021) conducted unsupervised learning of the literature by using Word2Vec for prediction of solar-chargeable battery materials in materials discovery. They calculated the cosine similarities (for relevance) between different materials and certain words like "photo-rechargeable". Choi & Lee (2024) studied in-context learning with LLMs in what they call materials language processing (MLP), which may be considered a subset of TLP, since battery terms are specific materials terms (Huang & Cole, 2022b). BatteryBERT was introduced by Huang & Cole (2022a) for the tasks of text classification and extractive QA by employing continued pretraining from the origianl BERT weights by Devlin et al. (2019) before finetuning. A toolkit that is based on the BatteryBERT is BatteryDataExtractor, which was used for the extraction of chemical data (Huang & Cole, 2022b). Its implementation is, however, limited to two turns of QA and does not follow a natural conversation style. Lee et al. (2024a) went further and introduced text-to-battery recipe for recipe

Table 2. Synthesis of some SotA results. Specific metrics or models are identified with superscripts.

No.	NLP task	Data	**Metric score** (F1 /Accuracy^{a} /RMSE^{rm} /ROUGE-1^{RG} /Cosine^{cs}) %	**Model** (Deep learning /Classical^{c} /LLM^{l} /Knowledge graphs^{kg})
1	TC	Battery abstracts	94.47^{a} (Huang & Cole, 2022a)	BatterySciBERT
		Battery papers[3]	96.6^{a} (Choi & Lee, 2024)	GPT3^{l} finetuned
		Battery papers[1]	97^{a} (Zheng et al., 2025)	GPT4^{l} finetuned
2	QA	SQuADv1.1 dev	89.16 (Huang & Cole, 2022a)	BatteryBERT-cased
		Battery device QA	88.21 (Choi & Lee, 2024)	GPT3^{l} finetuned
3	Materials discovery	SpringerLink & ScienceDirect	57^{cs} (He & Zhang, 2021)	Word2Vec^{c}
4	IE	*Unknown* articles	89.58 (Gou et al., 2024)	*Unkown*
		MKG	88.27 (Ye et al., 2024)	Darwin^{l} (ER)
		NLP4SIB	84.1 (Munjal et al., 2023)	SciBERT
		Auto-generated battery materials	67.98 (Huang & Cole, 2020)	*Unknown*
		Battery papers[1]	85 (Zheng et al., 2025)	Gemma2-9B^{l}
5	NER	Cell-Assembly	94.61 (Lee et al., 2024a)	BatteryBERT
		CHEMDNER[4]	95.98 (Huang & Cole, 2022b)	BatteryOnlyBERT-uncased
		Solid-state materials	95.1[5] (Choi & Lee, 2024)	GPT3^{l} finetuned
		MatScholar	87 (Choudhary & Kelley, 2023)	XLNet
		MS-Mention	91.47 (O'Gorman et al., 2021)	SciBERT
		MS-Mention	91.85 (O'Gorman et al., 2021)	SciBERT+SOFC
		MKG	92.96 (Ye et al., 2024)	Darwin^{l} (ER)
		manuals	93.99 (Ren et al., 2025)	CE-RTJE
6	Summarization	arXiv-cond-mat	46.5^{RG} Choudhary & Kelley (2023)	T5
7	Abbreviation detection	PLOS[1]	95.16 (Huang & Cole, 2022b)	BatteryOnlyBERT-cased
		Unkown articles	92.71 (Gou et al., 2024)	*Unkown*
8	SA	BYD NEVs	91.46 (Xu et al., 2025b)	MA-DKGCN^{kg}
		AutoHome	81 (Na et al.)	ALBERT
		X (formerly Twitter)	91.9 (Maghsoudi et al., 2025)	PyABSA
9	Policy model consistency (PMC)	Chinese policy documents	7.46 (Liang et al.)	ChatGPT^{l}
10	Relational triple extraction (RTE)	manuals	95.6 (Ren et al., 2025)	CE-RTJE^{kg}
11	SoC estimation	Panasonic-18650PF	1.89^{rm}, 2.21^{rm}, 2.02^{rm}, 2.42^{rm} (HWFET, LA92, UDDS, US06) at -20°C (Bian et al., 2024)	ChatGLM-6B^{l}
		LG-18650HG2	1.89^{rm}, 1.83^{rm}, 1.69^{rm}, 1.48^{rm} at -20°C (Bian et al., 2024)	ChatGLM-6B^{l}
		Samsung-18650-20R	1.91^{rm}, 2.75^{rm}, 2.72^{rm}, 1.92^{rm} at 0°C (Bian et al., 2024)	ChatGLM-6B^{l}
		A123-18650	1.69^{rm}, 2.46^{rm}, 1.86^{rm} (DST, FUDS, US06 at -10°C) (Bian et al., 2024)	ChatGLM-6B^{l}

extraction from the scientific literature. Recipes contain both materials and instructions about the materials. **A major challenge in materials discovery is the lack of a representative metric** or evaluation framework for evaluating successful discoveries without the need for expensive experimentation for validation. In addition, IE is a challenging task in the domain because it requires multiple critical variables for materials selection. To address this challenge, Nie et al. (2022) proposed a semantic knowledge graph (KG) dedicated to Li-ion battery (LiB) cathodes and featuring a dual-attention component that refines word embeddings. Their implementation is based on the bidirectional long short-term memory (BiLSTM). This is, however, known for limitations like less capacity for long-time dependencies and reliance on serial processing Graves et al. (2005), especially compared to the SotA Transformer architecture.

The material sodium is gaining attention because it is abundant and cheap, leading to Sodium-ion battery (SiB), which has similar working principles to LiB. Gou et al. (2024) performed IE for cathode materials of SiB. Their approach combined a number of other related tasks (e.g. chemical NER and TC) to improve performance on IE. Also, in the chemical NER contribution by O'Gorman et al. (2021) they annotated labels over a new corpus of 595 synthesis procedures and made the data (MS-Mentions) publicly available. Bai et al. (2025) also used NLP to guide the screening of single-atom catalysts for the Na-S batteries. In the work by Zheng et al. (2025), they introduced cognition-enhanced instruction framework (CEIF), where a teacher model provided feedback, prompt refinement, and optimized training data to guide the learning process of student models. Park et al. (2025) introduced Chemeleon, which was designed to generate chemical compositions and crystal structures from both textual descriptions and 3-dimensional structural data. However, the vastness of the possible combinations of chemical composition makes comprehensive exploration time-consuming and computationally demanding. Regarding explainability, an important transparency concept, Xu et al. (2025a) used Shapley additive explanations (SHAP) as explainable AI while employing biterm topic modelling for the analysis of patents related to LiB research. SHAP, however, like other posthoc explainability methods does not provide intuitive explanations for humans like textual explanations chain-of-thought (CoT) reasoning from LLMs.

4.2. Battery Production, Maintenance & Sustainability

Interestingly, several deep learning methods have been explored in the diagnostics of batteries, including SoC estimation (Bian et al., 2024). Adequate monitoring, diagnosis and maintenance ensure the prevention of battery overcharge, undue discharge, and explosion, thereby extending battery life (Liu et al., 2022). Bian et al. (2024) explored SoC estimation for LiB using LLMs. The task can be challenging due to harsh temperatures and dynamic operations, which are just a couple of factors out of several others that challenge reliable estimates within a module. In their work, they proposed hard prompt generator to translate LiB data into instruction and answer text while a soft prompt encodes task-specific information of various LiBs into a set of independent vectors. Since datasets are important for data-driven battery property predictions and estimates, some work focused on this. Huang & Cole (2020) introduced an automatically-generated materials dataset from the literature by extracting battery materials and 5 functional properties using ChemDataExtractor, which is a toolkit with NLP techniques for materials science. The extracted functional properties are capacity, conductivity, Coulombic efficiency, energy density, and voltage. Shon & Min (2023) extracted the ionic conductivities of SSE from scientific literature to create a materials database. Meanwhile, El-Bousiydy et al. (2021) noted that the lack of standard benchmark and how electrode and cell properties are reported in the domain are some of the challenges in LiB research giving rise to issues of lack of reproducibility among others.

Furthermore, Oprea & Bâra (2025) rightly identified that there are environmental concerns over the production of batteries, their use and disposal while Ren et al. (2025) addressed the disassembly of retired electric vehicle battery packs in their work by proposing disassembly-oriented knowledge graph. Deep learning methods face challenges such as poor generalization and robustness, especially when moving from LiB to LiB (Bian et al., 2024). This is complicated at harsh sub-zero ambient temperatures when SoC drops drastically with fluctuations due to chemical reactions in the LiBs. Besides, to use language modeling for SoC estimation, there is the need for models to learn measurements in numerical format, though this is a typical challenge (Jiang et al., 2025).

5. Discussion

The application of NLP in the battery domain is growing. However, challenges exist, as already highlighted. Below, we further highlight some of the key challenges and subsequently discuss details of our proposed novel TLP framework that is capable of addressing some of the challenges.

5.1. Challenges

Missing in Table 2 is agentic AI for any of the tasks. This is clearly a limitation in existing attempts, which our framework addresses. Also missing is any task related to the new EU DBP or digital product passport (DPP). These are in addition to some of the earlier identified gaps and limitations. Among some of the challenges in battery research are the lack of standard battery benchmark data, proprietary data, vocabulary standards, and interoperability of tools (Clark et al., 2022). Different authors used different self-created datasets. As a result, the SotA comparison made by Choi & Lee (2024) in their work may be considered unfair because their results were compared with other works which evaluated different datasets. Also, it would have been helpful if many of the authors gave unique names to their datasets for easy identification. Many of these shortcomings have negatively affected transparency in the field.

5.2. TLP framework for DBP and battery predictions

AI transparency provides an avenue for responsible human oversight (Adewumi et al., 2025a). The new EU DBP, which is an electronic record that may be stored in the cloud and is part of the DPP, provides the required transparency along the entire battery life cycle and promotes a circular economy (Popowicz et al., 2025; Rizos & Urban, 2024). The passport contains battery-specific texts, numeric values and dates. The initiative is similar to the concepts of data card, data statement, or model card in NLP (Pushkarna et al., 2022; Bender & Friedman, 2018; Mitchell et al., 2019). Data cards or statements provide summaries of ML data in a structured way with explanations of the processes and rationale behind the data. The DBP also aligns with the Findability, Accessibility, Interoperability, and Reusability (FAIR) guiding principles for scientific data management, which formulates a guideline for those who want to enhance the reusability of their data, and provides machines with the ability to automatically find and use such data (Wilkinson et al., 2016). It does not include personal data of the stakeholders, who include end-users, manufacturers, recyclers, and regulatory bodies, among others (Berger et al., 2022).

Just as it is beneficial for LLMs to have detailed prompts or instructions for improved performance (Adewumi et al., 2024b), it is beneficial to provide as much detail as necessary in the digital passport for the benefit of AI systems, besides the stakeholders. Consequently, the information, including those at relevant points in the value chain, can be transformed for TLP for BMS modeling and predictions since a BMS is responsible for monitoring and managing a battery pack for safe and optimal performance (Shen & Gao, 2019). The DBP will provide information on material origin, composition, chemical substances, carbon footprint, capacity, hazardous substances, state-of-health (SoH), battery status (e.g. reused), number of charging and discharging cycles, performance, recycling and disposal aspects, among others. BMSs typically provide information on the SoH and the expected battery lifetime (Berger et al., 2022). Accurately estimating battery state in extreme temperatures is one of the challenges in the field and the TLP framework aims to address that.

We propose a comprehensive but simple TLP framework for battery predictions, including various tasks identified in Table 2. Figure 3 depicts the framework. The information provided by the DBP or BMS will be useful within the framework. It involves a battery-agnostic model context protocol (MCP) AI agent that can connect to external tools that have information of the DBP or BMS and provides soft prompts (continuous feature vectors learned by prompt-tuning), which will be combined with optimized hard prompts (plain text inputs enhanced with gradient-based optimization) (Wen et al., 2023). The combination will then be supplied as input to a capable multimodal LLM for relevant TLP task predictions, e.g. SoC estimation or IR. In addition, the framework has a component capable of being a standalone part (indicated in the purple box of Fig. 3), which can perform tasks related to traditional datasets. The

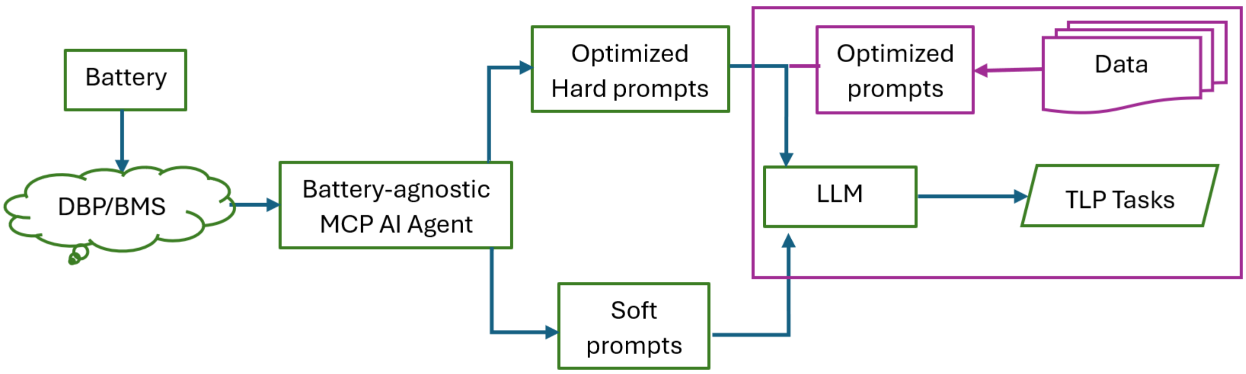

Figure 3. TLP framework for DBP and battery-related predictions.

strategy of combining both hard and soft prompts (or prompt parameters) can be beneficial by providing complementary perspectives as detailed and precise input to the LLM. The benefit of using LLMs for the tasks (instead of traditional methods) is that they have emergent properties and the capability to perform coding and other ML tasks that can support predictions. One disadvantage with dense LLMs is their size but using a mixture of experts (MoE) can be useful for mitigating this because not all the parameters need to be activated during inference.

6. Conclusion

The application of NLP along the entire battery life cycle is growing, with increasing research, reports, logs and other types of documents, as observed from the papers reviewed in this systematic survey. This work presented a comprehensive survey in the field, describing some traditional NLP tasks and emerging ones, as they relate to the battery domain. We showed, with the battery life cycle, that NLP can play important parts at different stages, from materials sourcing to recycling and repurposing. Our findings revealed that challenges still exist, especially with standardizing the benchmarks, creating the right evaluation framework or metric for materials discovery, and employing even more modern NLP methods, such as agentic AI. The EU DBP is another innovation not addressed in any of the papers surveyed. Hence, we introduced a novel TLP framework for battery predictions to address some of the challenges. With the planned DBP in the near future, the TLP framework will have even more data for driving better predictions. The inclusion of multimodal AI (featuring text, image, and sensor data) will also provide improved performance along the battery value chain.

Acknowledgement

This work is partly supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP), funded by Knut and Alice Wallenberg Foundations.

A. Search links

Google scholar[2]:
https://scholar.google.com/scholar?q=text+AND+%22language+processing%22+AND+battery

&hl=en&as_sdt=0%2C5&as_ylo=2017&as_yhi=2025

IEEE Xplore:

https://ieeexplore.ieee.org/search/searchresult.jsp?action=search&newsearch=true&matchBoolean=true&queryText=(%22All%20Metadata%22:text)%20AND%20(%22All%20Metadata%22:%22language%20processing%22)%20AND%20(%22All%20Metadata%22:battery)&highlight=true&returnFacets=ALL&returnType=SEARCH&matchPubs=true&rowsPerPage=50&pageNumber=1

Scopus:
https://www.scopus.com/results/results.uri?sort=plf-f&src=s&sid=87cba225c81851a1f3193513dccf661b&sot=a&sdt=a&sl=80&s=text+AND+%22language+processing%22+AND+battery+AND+PUBYEAR+%3E+2016+AND+PUBYEAR+%3C+2026&origin=savedSearchNewOnly&txGid=f0e09e9bcf4ff3781509c6543c2c2898&sessionSearchId=87cba225c81851a1f3193513dccf661b&limit=10

List of non-relevant papers

Below, we identify the non-relevant papers in the search results by listing DatabasePaperNumber (and the focus of the paper).

Google scholar G14 (about electrical equipment malfunction not battery), G16 (about consumer sentiments), G20 (neurocognitive assessment batteries - a set of clinical equipment), G21 (materials of inorganic glasses), G24 (Encryption in Mobile Ad Hoc Network), G26 (Quaternion algebra), G27 (Skateboard Monitoring Device), G29 (technology lexical database), G30 (requirement management in engineering), G32 (vehicle diagnostics using free-text customer service reports), G33 (Crowdfunding Videos), G34 (Android-Based Text Extraction), G35 (Materials Applications), G36 (in-text citation analysis), G39 (Analysis of Software Industry), G41 (neural databases), G42 (L2 Listening Proficiency), G43 (ensemble learning), G44 (Video summarization), G45 (Corporate Sustainability Reports), G46 (systems to detect significant future business changes), G48 (Unstructured Data from Medical Reports), G50 (Detection of Cognitive Decline), G51 (sentiment analysis (SA) for movie reviews), G52 (Digital shop floor management), G53 (Plastic Waste Recycling), G54 (Aviation Safety Reports), G55 (ADVANCES IN Computer Science), G58 (proceedings document), G59 (corrosion-resistant alloy), G60 (Monitoring Alzheimer's Disease), G61 (Fault Diagnosis of Signal Equipment), G62 (Language impairment in adults), G63 (clinical neuropsychology), G64 (properties of alloys), G65 (design research), G66 (therapist facilitative interpersonal skills), G67 (Loneliness in Older Adults), G68 (exploring GPT-3), G69=G71 (mild cognitive impairment), G70=G72 (Semantic Web), G73 (fMRI Dataset), G74 (sarcasm), G76 (product reviews), G77 (Personality and Psychological Distress), G78 (general language processing), G79 (Customer Satisfaction), G80 (Cognitive Functions), G81 (Clinical Data Generation), G82 (Psycholinguistic Assessments), G83 (opportunity discovery), G84 (IoT for aquaculture), G85 (Human-Machine Interaction), G86 (Arabic Sentiment Analysis), G87 (Symptom Documentation), G88 (Software Test Case Generation), G90 (extractive text summarization), G91 (cognitive plausibility), G92 (extraction from polymer literature), G93 (morphemic boundaries), G94 (Arabic NLP), G95 (Arabic Text Categorization), G96 (personnel selection), G97 (Generative AI), G98 (organisational culture), G99 (Energy Districts), G101 (Complex Engineered Systems), G102 (Legal Informatics), G103 (Customers' Sentiment Analysis), G104 (Measurement Extraction), G105 (IoT Based Voice Assistant), G106 (deep learning for NLP), G107 (Geolocation Context), G108 (Eye movements), G109 (Grid Monitoring), G110 (neuroimaging), G111 (psychological constructs), G112 (Examination System), G113 (Depression Disorder), G114 (Sentiment Classification), G115 (RAMS Information for Metro Vehicles), G116 (Neurodevelopmental Disorders:), G117 (Readers with Autism), G118 (Aspect Sentiment Analysis).

IEEE Xplore I2 (Supply Chain Information), I3 (Short Text Clustering), I4 (general natural language processing), I5 (Aphasia Speech), I7 (Computer Crimes Safety of IOT Networks), I8 (Software Energy Consumption), I9 (edge devices), I10 (Sentiment Polarity Computation), I11 (Comparative Sentiment Analysis), I12 (News Articles Summarization), I13 (Correcting Typing Errors), I14 (sentiments around aspects), I15 (Spinal Cord Stimulator), I16 (Unveiling Comment Insights), I17 (geotemporal visualization of Twitter analysis), I20 (Cross-Lingual Topic Model), I21 (opinion mining), I22 (Opinion Mining via Intrinsic and Extrinsic Domain Relevance), I23 (Aspect-Based Opinion Mining), I24 (Hospital Assisting Multitasking System), I25 (Agriculture Monitoring), I28 (fuzzy domain sentiment ontology), I29 (Mobile Cloud Support), I30 (general Generative AI), I31 (Speech Signal Quality Assessment), I32=I33=134=135=136=137=138 (Practical Guide to Machine Learning, NLP, and Generative AI),

Scopus S4 (Arabic text summarization), S6 (medical text classification), S8 (general priors-augmented retrieval priors-augmented retrieval Industrial applications of LLMs), S9 (digital matching methods), S10 (properties of crystalline materials), S11 (stress detection), S14 (hydrogen supply), S15 (perovskite solar devices), S16 (dementia), S17 (ophthalmic diseases), S20 (metal-organic frameworks), S21 (spin-orbit torque materials), S22 (TAP rule security detection), S23 (sentence representation of GitHub), S24 (technologies in aviation), S25 (electrocatalytic hydrogen evolution), S26 (speech-in-noise hearing tests), S27 (cognitive impairment), S28 (clinical review generation), S30 (patients with stroke), S31 (schizophrenia, bipolar disorder, and depression), S32 (patients with schizophrenia), S33 (EEG data for differential diagnosis), S34 (mechanical parts), S35 (Alzheimer's disease), S36 (imidazolium-based ionic liquids), S37 (collaborative robots), S38 (drug-target binding affinity), S39 (Williams syndrome), S40 (thermal CO2 hydrogenation), S41 (Vocal Course), S42 (subsidy policy on new energy vehicles), S44 (immunogenicity), S45 (dyslexia intervention), S47 (Aircraft EWIS safety risk), S48 (waste management policy), S50 (computational literature review), S51 (Cost Outcome Pathway), S52 (media discourse), S54 (molecular representation), S56 (open-source Agile practices), S57 (Fiber-Reinforced Composites), S58 (Photovoltaic forecasting), S59 (polymer processing database), S60 (metal-organic frameworks), S61 (electronic health records), S62 (Malware detection), S63 (Mongolian Emotional Speech Synthesis), S64 (hydropower projects), S65 (human–robot collaboration), S66 (multidimensional information extraction), S68 (electricity load forecasting), S69 (priors-augmented retrieval), S70 (atypical power system failures), S71 (hyperbolic soliton families), S72 (electroretinogram signal generation), S74 (category frameworks), S75 (Process Systems Engineering), S76 (Neural tracking of continuous speech), S77 (Emotion Vocabulary Recognition), S83 (tropical flood susceptibility), S84 (Emotional Speech Synthesis), S85 (fuzzy technology trajectories), S86 (AI-autonomous robotics), S87 (bigrams and promising patents), S89 (IT Professional Skills), S93 (Power Sample Feature Migration), S94 (Electric Vehicle Charging), S95 (disentanglement, selection, and reaggregation), S96 (fake review detection), S98 (competence-based HRM), S99 (Task Crowdsourcing), S101 (speech in noise), S102 (consumers' personalized preferences), S103 (Transcranial Direct Current Stimulation), S104 (Hard-of-Hearing Children), S105 (supercapacitor researches), S106 (Management of Skin Conditions), S107 (Approaches for 3D-Printing), S108 (blood pressure estimation), S109 (Digital Health Interventions), S110 (Tools for Scientific Review Writing),

References

Abdel-Nabi, H, Awajan, A & Ali, MZ (2023), 'Deep learning-based question answering: a survey,' *Knowledge and Information Systems*, **65**(4), pp. 1399–1485.

Adelani, DI, Abbott, J, Neubig, G, D'souza, D, Kreutzer, J, Lignos, C, Palen-Michel, C, Buzaaba, H, Rijhwani, S, Ruder, S et al. (2021), 'Masakhaner: Named entity recognition for african languages,' *Transactions of the Association for Computational Linguistics*, **9**, pp. 1116–1131.

Adewumi, T, Alkhaled, L, Gurung, N, van Boven, G & Pagliai, I (2024a), 'Fairness and bias in multimodal ai: A survey,' *arXiv preprint arXiv:2406.19097*.

Adewumi, T, Alkhaled, L, Imbert, F, Han, H, Habib, N & Löwenmark, K (2025a), 'Ai must not be fully autonomous,' *arXiv preprint arXiv:2507.23330*.

Adewumi, T, Habib, N, Alkhaled, L & Barney, E (2024b), 'Instruction makes a difference,' in Sfikas, G & Retsinas, G (eds.), *Document Analysis Systems*, Springer Nature Switzerland, Cham, pp. 71–88.

Adewumi, T, Habib, N, Alkhaled, L & Barney, E (2024c), 'On the limitations of large language models (llms): False attribution,' *arXiv preprint arXiv:2404.04631*.

Adewumi, T, Liwicki, F & Liwicki, M (2022a), 'State-of-the-art in open-domain conversational ai: A survey,' *Information*, **13**(6), doi:10.3390/info13060298.

Adewumi, T, Liwicki, FS, Liwicki, M, Gardelli, V, Alkhaled, L & Mokayed, H (2025b), 'Findings of mega: Maths explanation with llms using the socratic method for active learning,' *arXiv preprint arXiv:2507.12079*.

Adewumi, T, Vadoodi, R, Tripathy, A, Nikolaido, K, Liwicki, F & Liwicki, M (2022b), 'Potential idiomatic expression (PIE)-English: Corpus for classes of idioms,' in Calzolari, N, Béchet, F, Blache, P, Choukri, K, Cieri, C, Declerck, T, Goggi, S, Isahara, H, Maegaard, B, Mariani, J, Mazo, H, Odijk, J & Piperidis, S (eds.), *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, European Language Resources Association, Marseille, France, pp. 689–696.

Attia, PM, Moch, E & Herring, PK (2025), 'Challenges and opportunities for high-quality battery production at scale,' *Nature Communications*, **16**(1), p. 611.

Bai, R, Yao, Y, Lin, Q, Wu, L, Li, Z, Wang, H, Ma, M, Mu, D, Hu, L, Yang, H et al. (2025), 'Preferable single-atom catalysts enabled by natural language processing for high energy density na-s batteries,' *Nature Communications*, **16**(1), p. 5827.

Bender, EM & Friedman, B (2018), 'Data statements for natural language processing: Toward mitigating system bias and enabling better science,' *Transactions of the Association for Computational Linguistics*, **6**, pp. 587–604.

Benzeghiba, M, De Mori, R, Deroo, O, Dupont, S, Erbes, T, Jouvet, D, Fissore, L, Laface, P, Mertins, A, Ris, C et al. (2007), 'Automatic speech recognition and speech variability: A review,' *Speech communication*, **49**(10-11), pp. 763–786.

Berger, K, Schöggl, JP & Baumgartner, RJ (2022), 'Digital battery passports to enable circular and sustainable value chains: Conceptualization and use cases,' *Journal of Cleaner Production*, **353**, p. 131492.

Bian, C, Duan, Z, Hao, Y, Yang, S & Feng, J (2024), 'Exploring large language model for generic and robust state-of-charge estimation of li-ion batteries: A mixed prompt learning method,' *Energy*, **302**, p. 131856.

Bojar, O, Buck, C, Federmann, C, Haddow, B, Koehn, P, Leveling, J, Monz, C, Pecina, P, Post, M, Saint-Amand, H et al. (2014), 'Findings of the 2014 workshop on statistical machine translation,' in *Proceedings of the ninth workshop on statistical machine translation*, pp. 12–58.

Brown, T, Mann, B, Ryder, N, Subbiah, M, Kaplan, JD, Dhariwal, P, Neelakantan, A, Shyam, P, Sastry, G, Askell, A et al. (2020), 'Language models are few-shot learners,' *Advances in neural information processing systems*, **33**, pp. 1877–1901.

Brundage, MP, Sexton, T, Hodkiewicz, M, Dima, A & Lukens, S (2021), 'Technical language processing: Unlocking maintenance knowledge,' *Manufacturing Letters*, **27**, pp. 42–46.

Cheung, J, Zhuang, Y, Li, Y, Shetty, P, Zhao, W, Grampurohit, S, Ramprasad, R & Zhang, C (2024), 'POLYIE: A dataset of information extraction from polymer material scientific literature,' in Duh, K, Gomez, H & Bethard, S (eds.), *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, Association for Computational Linguistics, Mexico City, Mexico, pp. 2370–2385, doi:10.18653/v1/2024.naacl-long.131.

Choi, J & Lee, B (2024), 'Accelerating materials language processing with large language models,' *Communications Materials*, **5**(1), p. 13.

Choudhary, K & Kelley, ML (2023), 'Chemnlp: a natural language-processing-based library for materials chemistry text data,' *The Journal of Physical Chemistry C*, **127**(35), pp. 17545–17555.

Clark, S, Bleken, FL, Stier, S, Flores, E, Andersen, CW, Marcinek, M, Szczesna-Chrzan, A, Gaberscek, M, Palacin, MR, Uhrin, M et al. (2022), 'Toward a unified description of battery data,' *Advanced Energy Materials*, **12**(17), p. 2102702.

Cohan, A, Dernoncourt, F, Kim, DS, Bui, T, Kim, S, Chang, W & Goharian, N (2018), 'A discourse-aware attention model for abstractive summarization of long documents,' in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, Association for Computational Linguistics, New Orleans, Louisiana, pp. 615–621, doi: 10.18653/v1/N18-2097.

Degen, F, Winter, M, Bendig, D & Tübke, J (2023), 'Energy consumption of current and future production of lithium-ion and post lithium-ion battery cells,' *Nature energy*, **8**(11), pp. 1284–1295.

Derczynski, L, Nichols, E, van Erp, M & Limsopatham, N (2017), 'Results of the WNUT2017 shared task on novel and emerging entity recognition,' in *Proceedings of the 3rd Workshop on Noisy User-generated Text*, Association for Computational Linguistics, Copenhagen, Denmark, pp. 140–147, doi:10.18653/v1/W17-4418.

Devlin, J, Chang, MW, Lee, K & Toutanova, K (2019), 'Bert: Pre-training of deep bidirectional transformers for language understanding,' in *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pp. 4171–4186.

Dima, A, Lukens, S, Hodkiewicz, M, Sexton, T & Brundage, MP (2021), 'Adapting natural language processing for technical text,' *Applied AI Letters*, **2**(3), p. e33.

Dione, CMB, Adelani, DI, Nabende, P, Alabi, J, Sindane, T, Buzaaba, H, Muhammad, SH, Emezue, CC, Ogayo, P, Aremu, A, Gitau, C, Mbaye, D, Mukiibi, J, Sibanda, B, Dossou, BFP, Bukula, A, Mabuya, R, Tapo, AA, Munkoh-Buabeng, E, Memdjokam Koagne, V, Ouoba Kabore, F, Taylor, A, Kalipe, G, Macucwa, T, Marivate, V, Gwadabe, T, Elvis, MT, Onyenwe, I, Atindogbe, G, Adelani, T, Akinade, I, Samuel, O, Nahimana, M, Musabeyezu, T, Niyomutabazi, E, Chimhenga, E, Gotosa, K, Mizha, P, Agbolo, A, Traore, S, Uchechukwu, C, Yusuf, A, Abdullahi, M & Klakow, D (2023), 'MasakhaPOS: Part-of-speech tagging for typologically diverse African languages,' in Rogers, A, Boyd-Graber, J & Okazaki, N (eds.), *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Association for Computational Linguistics, Toronto, Canada, pp. 10883–10900, doi:10.18653/v1/2023.acl-long.609.

Eisenstein, J (2019), *Introduction to natural language processing*, MIT press.

El-Bousiydy, H, Lombardo, T, Primo, EN, Duquesnoy, M, Morcrette, M, Johansson, P, Simon, P, Grimaud, A & Franco, AA (2021), 'What can text mining tell us about lithium-ion battery researchers' habits?' *Batteries & Supercaps*, **4**(5), pp. 758–766.

El-Kassas, WS, Salama, CR, Rafea, AA & Mohamed, HK (2021), 'Automatic text summarization: A comprehensive survey,' *Expert systems with applications*, **165**, p. 113679.

Gasparetto, A, Marcuzzo, M, Zangari, A & Albarelli, A (2022), 'A survey on text classification algorithms: From text to predictions,' *Information*, **13**(2), p. 83.

Gehrmann, S, Adewumi, T, Aggarwal, K, Ammanamanchi, PS, Aremu, A, Bosselut, A, Chandu, KR, Clinciu, MA, Das, D, Dhole, K, Du, W, Durmus, E, Dušek, O, Emezue, CC, Gangal, V, Garbacea, C, Hashimoto, T, Hou, Y, Jernite, Y, Jhamtani, H, Ji, Y, Jolly, S, Kale, M, Kumar, D, Ladhak, F, Madaan, A, Maddela, M, Mahajan, K, Mahamood, S, Majumder, BP, Martins, PH, McMillan-Major, A, Mille, S, van Miltenburg, E, Nadeem, M, Narayan, S, Nikolaev, V, Niyongabo Rubungo, A, Osei, S, Parikh, A, Perez-Beltrachini, L, Rao, NR, Raunak, V, Rodriguez, JD, Santhanam, S, Sedoc, J, Sellam, T, Shaikh, S, Shimorina, A, Sobrevilla Cabezudo, MA, Strobelt, H, Subramani, N, Xu, W, Yang, D, Yerukola, A & Zhou, J (2021), 'The GEM benchmark: Natural language generation, its evaluation and metrics,' in Bosselut, A, Durmus, E, Gangal, VP, Gehrmann, S, Jernite, Y, Perez-Beltrachini, L, Shaikh, S & Xu, W (eds.), *Proceedings of the First Workshop on Natural Language Generation, Evaluation, and Metrics (GEM)*, Association for Computational Linguistics, Online, pp. 96–120, doi:10.18653/v1/2021.gem-1.10.

Gou, Y, Zhang, Y, Zhu, J & Shu, Y (2024), 'A document-level information extraction pipeline for layered cathode materials for sodium-ion batteries,' *Scientific Data*, **11**(1), p. 372.

Graves, A, Fernández, S & Schmidhuber, J (2005), 'Bidirectional lstm networks for improved phoneme classification and recognition,' in *International conference on artificial neural networks*, Springer, pp. 799–804.

Grishman, R (1997), 'Information extraction: Techniques and challenges,' in *International summer school on information extraction*, Springer, pp. 10–27.

Guzmán, F, Chen, PJ, Ott, M, Pino, J, Lample, G, Koehn, P, Chaudhary, V & Ranzato, M (2019), 'The FLORES evaluation datasets for low-resource machine translation: Nepali–English and Sinhala–English,' in Inui, K, Jiang, J, Ng, V & Wan, X (eds.), *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Association for Computational Linguistics, Hong Kong, China, pp. 6098–6111, doi:10.18653/v1/D19-1632.

Habib, N, Adewumi, T, Liwicki, M & Barney, E (2025), 'Trends and challenges in authorship analysis: A review of ml, dl, and llm approaches,' *arXiv preprint arXiv:2505.15422*.

He, M & Zhang, L (2021), 'Prediction of solar-chargeable battery materials: A text-mining and first-principles investigation,' *International Journal of Energy Research*, **45**(10), pp. 15521–15533.

He, P, Gao, J & Chen, W (2021), 'Debertav3: Improving deberta using electra-style pre-training with gradient-disentangled embedding sharing,' *arXiv preprint arXiv:2111.09543*.

Hu, Z, Hou, W & Liu, X (2024), 'Deep learning for named entity recognition: a survey,' *Neural Computing and Applications*, **36**(16), pp. 8995–9022.

Huang, S & Cole, JM (2020), 'A database of battery materials auto-generated using chemdataextractor,' *Scientific Data*, **7**(1), p. 260.

Huang, S & Cole, JM (2022a), 'Batterybert: A pretrained language model for battery database enhancement,' *Journal of chemical information and modeling*, **62**(24), pp. 6365–6377.

Huang, S & Cole, JM (2022b), 'Batterydataextractor: battery-aware text-mining software embedded with bert models,' *Chemical Science*, **13**(39), pp. 11487–11495.

Jiang, K & Lu, X (2020), 'Natural language processing and its applications in machine translation: a diachronic review,' in *2020 IEEE 3rd international conference of safe production and informatization (IICSPI)*, IEEE, pp. 210–214.

Jiang, X, Wang, W, Tian, S, Wang, H, Lookman, T & Su, Y (2025), 'Applications of natural language processing and large language models in materials discovery,' *npj Computational Materials*, **11**(1), p. 79.

Kheddar, H, Hemis, M & Himeur, Y (2024), 'Automatic speech recognition using advanced deep learning approaches: A survey,' *Information fusion*, **109**, p. 102422.

Kim, M, Kim, H & Moon, J (2025), 'Beginner-friendly review of research on r-based energy forecasting: Insights from text mining.' *Electronics (2079-9292)*, **14**(17).

Kryscinski, W, Rajani, N, Agarwal, D, Xiong, C & Radev, D (2022), 'BOOKSUM: A collection of datasets for long-form narrative summarization,' in Goldberg, Y, Kozareva, Z & Zhang, Y (eds.), *Findings of the Association for Computational Linguistics: EMNLP 2022*, Association for Computational Linguistics, Abu Dhabi, United Arab Emirates, pp. 6536–6558, doi:10.18653/v1/2022.findings-emnlp.488.

Lane, H & Dyshel, M (2025), *Natural language processing in action*, Simon and Schuster.

Lee, D, Choi, J, Mizuseki, H & Lee, B (2024a), 'Text-to-battery recipe: A language modeling-based protocol for automatic battery recipe extraction and retrieval,' *arXiv preprint arXiv:2407.15459*.

Lee, D, Mizuseki, H, Choi, J et al. (2025), 'Building an end-to-end battery recipe knowledge base via transformer-based text mining,' *Communications Materials*, **6**, p. 100, doi:10.1038/s43246-025-00579-1.

Lee, JH, Lee, M & Min, K (2023), 'Natural language processing techniques for advancing materials discovery: a short review,' *International Journal of Precision Engineering and Manufacturing-Green Technology*, **10**(5), pp. 1337–1349.

Lee, JH, Shin, D & Hwang, Y (2024b), 'Investigating the capabilities of large language model-based task-oriented dialogue chatbots from a learner's perspective,' *System*, **127**, p. 103538.

Liang, J, Wang, Y, Li, W & Wang, W (????), 'Quantitative evaluation of china's energy storage policies: A chatgpt-based pmc index modelling approach,' *Available at SSRN 5235882*.

Liddy, ED (2001), 'Natural language processing,' *In Encyclopedia of Library and Information Science, 2nd Ed. NY.*

Lin, CY (2004), 'ROUGE: A package for automatic evaluation of summaries,' in *Text Summarization Branches Out*, Association for Computational Linguistics, Barcelona, Spain, pp. 74–81.

Lin, TY & Chou, LC (2025), 'A systematic review of artificial intelligence applications and methodological advances in patent analysis,' *World Patent Information*, **82**, p. 102383.

Liu, K, Hu, X, Zhou, H, Tong, L, Widanage, WD & Marco, J (2021), 'Feature analyses and modeling of lithium-ion battery manufacturing based on random forest classification,' *IEEE/ASME Transactions on Mechatronics*, **26**(6), pp. 2944–2955.

Liu, Y, He, Y, Bian, H, Guo, W & Zhang, X (2022), 'A review of lithium-ion battery state of charge estimation based on deep learning: Directions for improvement and future trends,' *Journal of Energy Storage*, **52**, p. 104664.

Liu, Y, Ott, M, Goyal, N, Du, J, Joshi, M, Chen, D, Levy, O, Lewis, M, Zettlemoyer, L & Stoyanov, V (2019), 'Roberta: A robustly optimized bert pretraining approach,' *arXiv preprint arXiv:1907.11692*.

Löwenmark, K, Taal, C, Vurgaft, A, Nivre, J, Liwicki, M & Sandin, F (2023), 'Labelling of annotated condition monitoring data through technical language processing,' in *15th Annual Conference of the Prognostics and Health Management Society, PHM 2023. Salt Lake City, USA. 28 October 2023 through 2 November 2023*, Prognostics and Health Management Society, vol. 15.

Lyonnard, S, Biscari, C, Bozzini, B, Casas-Cabanas, M, Calisto, BM, Fransson, M, Graceffa, R, Hennies, F, Hinrichsen, B, Karlsson, M et al. (2025), 'Building a community lightsource meta-infrastructure to accelerate battery innovation in europe,' *Journal of Physics: Energy*, **7**(3), p. 031001.

Maas, AL, Daly, RE, Pham, PT, Huang, D, Ng, AY & Potts, C (2011), 'Learning word vectors for sentiment analysis,' in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, Association for Computational Linguistics, Portland, Oregon, USA, pp. 142–150.

Maghsoudi, M, Mohammadi, N & Bakhtiari, M (2025), 'A novel approach to customer segmentation for product development on social media data: Integrating aspect-based sentiment analysis and text mining,' *Knowledge-Based Systems*, p. 114269.

Mahbub, R, Huang, K, Jensen, Z, Hood, ZD, Rupp, JL & Olivetti, EA (2020), 'Text mining for processing conditions of solid-state battery electrolytes,' *Electrochemistry Communications*, **121**, p. 106860.

Manning, CD (2011), 'Part-of-speech tagging from 97% to 100%: is it time for some linguistics?' in *International conference on intelligent text processing and computational linguistics*, Springer, pp. 171–189.

Marcus, MP, Santorini, B & Marcinkiewicz, MA (1993), 'Building a large annotated corpus of English: The Penn Treebank,' *Computational Linguistics*, **19**(2), pp. 313–330.

Mikolov, T, Chen, K, Corrado, G & Dean, J (2013), 'Efficient estimation of word representations in vector space,' *arXiv preprint arXiv:1301.3781*.

Minaee, S, Kalchbrenner, N, Cambria, E, Nikzad, N, Chenaghlu, M & Gao, J (2021), 'Deep learning–based text classification: a comprehensive review,' *ACM computing surveys (CSUR)*, **54**(3), pp. 1–40.

Mitali, J, Dhinakaran, S & Mohamad, A (2022), 'Energy storage systems: A review,' *Energy Storage and Saving*, **1**(3), pp. 166–216.

Mitchell, M, Wu, S, Zaldivar, A, Barnes, P, Vasserman, L, Hutchinson, B, Spitzer, E, Raji, ID & Gebru, T (2019), 'Model cards for model reporting,' in *Proceedings of the conference on fairness, accountability, and transparency*, pp. 220–229.

Munjal, M, Prein, T, Venugopal, V, Huang, KJ & Olivetti, E (2023), 'Extracting a database of challenges and mitigation strategies for sodium-ion battery development,' in *AI for Accelerated Materials Design-NeurIPS 2023 Workshop*.

Na, J, Long, R, Chen, H, Wang, X, Yang, S, Sun, Q & Huang, Z (????), 'Decoding dissatisfaction drivers and sentiment trends in bev demand: An integrated lda and deep learning analysis,' *Available at SSRN 5174549*.

Nekahi, A, Dorri, M, Rezaei, M, Bouguern, MD, Madikere Raghunatha Reddy, AK, Li, X, Deng, S & Zaghib, K (2024), 'Comparative issues of metal-ion batteries toward sustainable energy storage: Lithium vs. sodium,' *Batteries*, **10**(8), p. 279.

Nekahi, A, Feyzi, E, Srivastava, M, Yeganehdoust, F, Reddy, AKMR & Zaghib, K (2025), 'Advanced lithium-ion battery process manufacturing equipment for gigafactories: Past, present, and future perspectives,' *IScience*, **28**(7).

Nie, Z, Zheng, S, Liu, Y, Chen, Z, Li, S, Lei, K & Pan, F (2022), 'Automating materials exploration with a semantic knowledge graph for li-ion battery cathodes,' *Advanced Functional Materials*, **32**(26), p. 2201437.

OpenAI, :, Hurst, A, Lerer, A, Goucher, AP, Perelman, A, Ramesh, A, Clark, A, Ostrow, A, Welihinda, A, Hayes, A, Radford, A, Mądry, A, Baker-Whitcomb, A, Beutel, A, Borzunov, A, Carney, A & more (2024), 'Gpt-4o system card,' .

Oprea, SV & Bâra, A (2025), 'Unveiling the nexus between energy storage and electricity markets in academic publications. a data-driven analysis of emerging trends and market dynamics using nlp, sentiment analysis and probabilistic modeling,' *Journal of Energy Storage*, **106**, p. 114917.

Örüm Aydin, A, Zajonz, F, Günther, T, Dermenci, KB, Berecibar, M & Urrutia, L (2023), 'Lithium-ion battery manufacturing: Industrial view on processing challenges, possible solutions and recent advances,' *Batteries*, **9**(11), p. 555.

Osaro, E, Karpinski, N, Alornyo, S & Ighalo, JO (2025), 'Artificial intelligence in materials science and chemistry: Past, present and future trajectories,' *Materials Today Chemistry*, **49**, p. 103115.

O'Gorman, T, Jensen, Z, Mysore, S, Huang, K, Mahbub, R, Olivetti, E & McCallum, A (2021), 'Ms-mentions: consistently annotating entity mentions in materials science procedural text,' in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pp. 1337–1352.

Page, MJ, McKenzie, JE, Bossuyt, PM, Boutron, I, Hoffmann, TC, Mulrow, CD, Shamseer, L, Tetzlaff, JM, Akl, EA, Brennan, SE et al. (2021), 'The prisma 2020 statement: an updated guideline for reporting systematic reviews,' *bmj*, **372**.

Park, H, Onwuli, A & Walsh, A (2025), 'Exploration of crystal chemical space using text-guided generative artificial intelligence,' *Nature Communications*, **16**(1), p. 4379.

Pei, Z, Yin, J & Zhang, J (2025), 'Language models for materials discovery and sustainability: Progress, challenges, and opportunities,' *Progress in Materials Science*, p. 101495.

Pettersson, J, Hult, E, Eriksson, T & Adewumi, T (2024), 'Generative ai and teachers–for us or against us? a case study,' in *Proceedings of the 14th Scandinavian Conference on Artificial Intelligence SCAI*.

Pievaste, I, Belouettar, S, Mercuri, F, Fantuzzi, N, Dehghani, H, Izadi, R, Ibrahim, H, Lengiewicz, J, Belouettar-Mathis, M, Bendine, K et al. (2025), 'Artificial intelligence in materials science and engineering: Current landscape, key challenges, and future trajectories,' *Composite Structures*, p. 119419.

Popowicz, M, Pohlmann, A, Schöggl, JP & Baumgartner, RJ (2025), 'Digital product passports as information providers for consumers—the case of digital battery passports,' *Business Strategy and the Environment*.

Pushkarna, M, Zaldivar, A & Kjartansson, O (2022), 'Data cards: Purposeful and transparent dataset documentation for responsible ai,' in *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, pp. 1776–1826.

Ren, Y, Wu, J, Zhuang, C, Sun, X, Guo, H, Wu, J, Chen, Y & Liu, J (2025), 'Automated disassembly-oriented knowledge graph construction for retired battery packs using a candidate entity-based relational triple joint extraction method,' *Advanced Engineering Informatics*, **67**, p. 103525.

Rizos, V & Urban, P (2024), 'Implementing the eu digital battery passport: Opportunities and challenges for battery circularity,' *CEPS Papers*, (42379).

Schomburg, F, Heidrich, B, Wennemar, S, Drees, R, Roth, T, Kurrat, M, Heimes, H, Jossen, A, Winter, M, Cheong, JY et al. (2024), 'Lithium-ion battery cell formation: status and future directions towards a knowledge-based process design,' *Energy & Environmental Science*, **17**(8), pp. 2686–2733.

Sebastiani, F (2002), 'Machine learning in automated text categorization,' *ACM computing surveys (CSUR)*, **34**(1), pp. 1–47.

Shen, M & Gao, Q (2019), 'A review on battery management system from the modeling efforts to its multiapplication and integration,' *International Journal of Energy Research*, **43**(10), pp. 5042–5075.

Shon, YJ & Min, K (2023), 'Extracting chemical information from scientific literature using text mining: building an ionic conductivity database for solid-state electrolytes,' *ACS omega*, **8**(20), pp. 18122–18127.

Singh, G, Mittal, N & Chouhan, SS (2022), 'A systematic review of deep learning approaches for natural language processing in battery materials domain,' *IETE Technical Review*, **39**(5), pp. 1046–1057.

Srinivasa-Desikan, B (2018), *Natural Language Processing and Computational Linguistics: A practical guide to text analysis with Python, Gensim, spaCy, and Keras*, Packt Publishing Ltd.

Tjong Kim Sang, EF & De Meulder, F (2003), 'Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition,' in *Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003*, pp. 142–147.

Toutanova, K, Klein, D, Manning, CD & Singer, Y (2003), 'Feature-rich part-of-speech tagging with a cyclic dependency network,' in *Proceedings of the 2003 human language technology conference of the north american chapter of the association for computational linguistics*, pp. 252–259.

Vaswani, A, Shazeer, N, Parmar, N, Uszkoreit, J, Jones, L, Gomez, AN, Kaiser, Ł & Polosukhin, I (2017), 'Attention is all you need,' *Advances in neural information processing systems*, **30**.

Wang, H, Wu, H, He, Z, Huang, L & Church, KW (2022), 'Progress in machine translation,' *Engineering*, **18**, pp. 143–153.

Wankhade, M, Rao, ACS & Kulkarni, C (2022), 'A survey on sentiment analysis methods, applications, and challenges,' *Artificial Intelligence Review*, **55**(7), pp. 5731–5780.

Wen, Y, Jain, N, Kirchenbauer, J, Goldblum, M, Geiping, J & Goldstein, T (2023), 'Hard prompts made easy: Gradient-based discrete optimization for prompt tuning and discovery,' in Oh, A, Naumann, T, Globerson, A, Saenko, K, Hardt, M & Levine, S (eds.), *Advances in Neural Information Processing Systems*, Curran Associates, Inc., vol. 36, pp. 51008–51025.

Weng, A, Olide, E, Kovalchuk, I, Siegel, JB & Stefanopoulou, A (2023), 'Modeling battery formation: Boosted sei growth, multi-species reactions, and irreversible expansion,' *Journal of The Electrochemical Society*, **170**(9), p. 090523.

Weston, L, Tshitoyan, V, Dagdelen, J, Kononova, O, Trewartha, A, Persson, KA, Ceder, G & Jain, A (2019), 'Named entity recognition and normalization applied to large-scale information extraction from the materials science literature,' *Journal of chemical information and modeling*, **59**(9), pp. 3692–3702.

Wilkinson, MD, Dumontier, M, Aalbersberg, IJ, Appleton, G, Axton, M, Baak, A, Blomberg, N, Boiten, JW, da Silva Santos, LB, Bourne, PE et al. (2016), 'The fair guiding principles for scientific data management and stewardship,' *Scientific data*, **3**(1), pp. 1–9.

Winter, M & Brodd, RJ (2004), 'What are batteries, fuel cells, and supercapacitors?' *Chemical reviews*, **104**(10), pp. 4245–4270.

Wu, X, Nguyen, T & Luu, AT (2024), 'A survey on neural topic models: methods, applications, and challenges,' *Artificial Intelligence Review*, **57**(2), p. 18.

Xiao, J, Cao, X, Gridley, B, Golden, W, Ji, Y, Johnson, S, Lu, D, Lin, F, Liu, J, Liu, Y et al. (2025), 'From mining to manufacturing: Scientific challenges and opportunities behind battery production,' *Chemical Reviews*.

Xu, Y, Zhang, X & Xu, Y (2025a), 'Technology opportunity discovery based on biterm topic model and explainable artificial intelligence model,' in *2025 8th International Conference on Artificial Intelligence and Big Data (ICAIBD)*, IEEE, pp. 819–824.

Xu, Z, Wu, Y, Tang, L & Gui, S (2025b), 'From user-generated content to quality improvement: A multi-granularity analysis of customer satisfaction and attention in new energy vehicles using deep learning,' *Computers in Industry*, **173**, p. 104380.

Yatskar, M (2019), 'A qualitative comparison of CoQA, SQuAD 2.0 and QuAC,' in Burstein, J, Doran, C & Solorio, T (eds.), *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, Association for Computational Linguistics, Minneapolis, Minnesota, pp. 2318–2323, doi:10.18653/v1/N19-1241.

Ye, Y, Ren, J, Wang, S, Wan, Y, Razzak, I, Hoex, B, Wang, H, Xie, T & Zhang, W (2024), 'Construction and application of materials knowledge graph in multidisciplinary materials science via large language model,' *Advances in Neural Information Processing Systems*, **37**, pp. 56878–56897.

Zhang, T, Kishore, V, Wu, F, Weinberger, KQ & Artzi, Y (2020), 'Bertscore: Evaluating text generation with bert,' in *International Conference on Learning Representations*.

Zhao, S, Chen, S, Zhou, J, Li, C, Tang, T, Harris, SJ, Liu, Y, Wan, J & Li, X (2024), 'Potential to transform words to watts with large language models in battery research,' *Cell Reports Physical Science*, **5**(3).

Zheng, Y, Li, S, Liu, Z, Wang, Y, Guo, D, Wang, Y, Han, X & Ouyang, M (2025), 'Cognition-enhanced instruction framework: Accelerating structured battery knowledge extraction with low-parameter models,' *Journal of Energy Chemistry*.

Zuo, W, Zheng, H, He, T, Vishwanath, V, Chan, MK, Stevens, RL, Amine, K & Xu, GL (2025), 'Large language models for batteries,' *Joule*, **9**(8).