Battery Data Management

How data management can accelerate battery development

Improving EV Affordability: The Role of Battery Data Management in Development

Let's face it, before the cost of an electric vehicle (EV) is not (significantly) lower than a car with an internal combustion engine (ICE), mass adoption will be tough. When anticipating future developments - this means in the battery field looking at Asian developments - the path to cost reduction is plausible and near. Industry leaders such as BYD, Changan, and Hyundai are already reshaping market expectations by offering electric cars at price points below €10,000. Drawing on insights from a recent Gartner study, the horizon where battery electric vehicles (BEVs) become more cost-effective to produce than a comparable ICE vehicle is not as distant as once thought. By 2027, innovative manufacturing techniques are anticipated to lower production costs significantly.

However, despite these encouraging signs, a sobering reality check reminds us that the journey toward affordable electric transportation is far from over. The strong economic contrast remains evident as EVs, on average, still cost 40% more to produce than their ICE counterparts. This disparity underscores the challenging, yet crucial, path forward in achieving the economic and environmental promises of electric mobility.

In the pursuit of making electric vehicle (EV) production economically sustainable, a pivotal shift is required, not in incremental adjustments, but in fundamentally transforming the product development process.

At the Munich Motor Show 2023, Kai Grünitz, responsible for technical development at Volkswagen, speaks about cutting their development process from 54 months to 36 months:

"The question is how we reduce test cycles and stay within the quality. [...] That's the key to success. It's not done yet but we know what to do." - *Kai Grünitz*

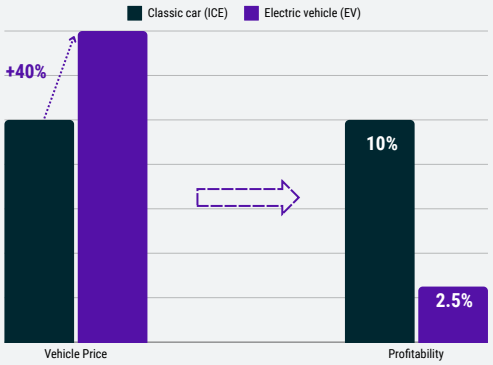
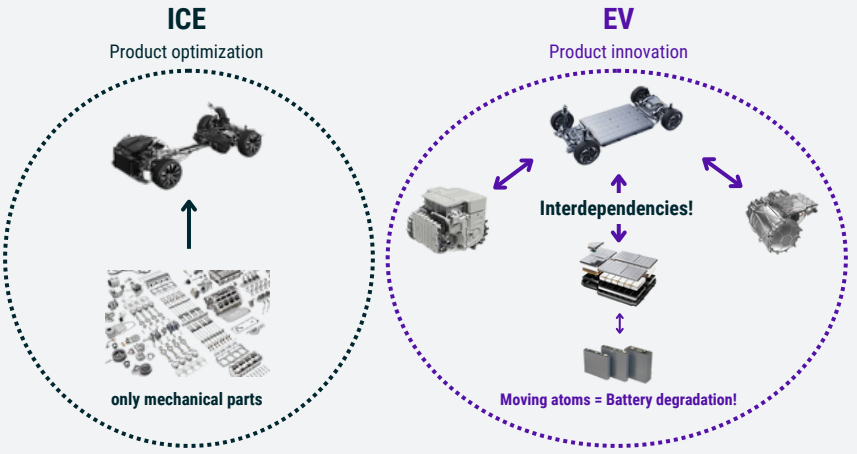
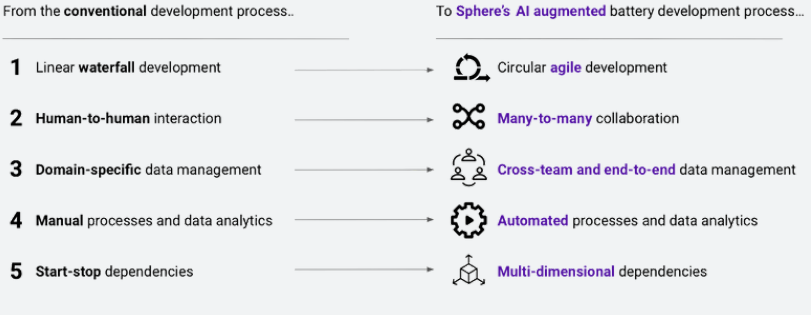


Figure 1: *Cost comparison graph of classic cars vs. electric vehicles.*



Currently, the profitability of EVs appears to be lower due to the **intricate development process** of the battery components involved and the strong **interdependencies** among them.

The barrier of mass EV adoption can be overcome by **reducing the costs of battery development while also trying to speed up the process**.



Today's methodologies in vehicle design and manufacturing are a representation of precision, optimized for developing Internal Combustion Engine (ICE) vehicles with their tens of thousands of components—a truly remarkable undertaking. Automakers have mastered the art of predicting how each component behaves, ensuring reliability and performance. One key thing has changed though. With the inclusion of batteries in the BOM of a car, we now have moving atoms in play. Current operating models are not designed to deal with this added complexity and dynamics. Moreover, the battery technology sector is in a constant state of motion, with new chemistries, designs, and engineering solutions emerging at a pace that traditional automotive development cycles can't keep up with. The typical 5-7 year development cycle of an ICE vehicle is proving to be inadequate for the fast-evolving EV market.

To stay competitive and achieve a sustainable model for EV production, it's clear that a revolution in product development practices is necessary. Embracing agility, fostering innovation, and adapting to the rapid advancements in battery technology are no longer options but imperatives for the future of the automotive industry.

Transforming the entire development process spans across multiple areas: organizational structures, culture, systems, and processes are just a few capabilities that have to change. In this discussion, we want to focus on one key area that is pushing this holistic transition: transforming how we manage battery data. You might be asking, "What sets battery data management apart from, say, managing the data for turbochargers?" Or "What role does AI have to play in all of this?" And importantly, "How can we mix traditional data management concepts with new methods to unlock new value?" These are some of the questions we are trying to untangle in this paper.

The specifics of battery data management in the battery development process

The importance of solid data management has been discussed in the industry for decades. We are not trying to reinvent the wheel but to tailor existing concepts to the specifics of battery data. We see three main characteristics that make battery data hard and very important to manage properly.

Data volume and variety

As one of the first steps in the battery development process, an automotive or power tool manufacturer alike needs to select the right cell technology and supplier for the defined technical specifications. As part of this strategic benchmarking project, a lot of cells need to be tested. The proper analysis of one cell alone, and we speak from firsthand experience, leads to gigabytes of performance data. Now let's extrapolate this to all subsequent tests that will have to be performed once a cell has been selected: detailed charging tests, application-specific aging tests, thermal- and pressure tests, module tests, pack tests, and many more. An enormous volume of data needs to be managed, analyzed, and synchronized between different data customers downstream.

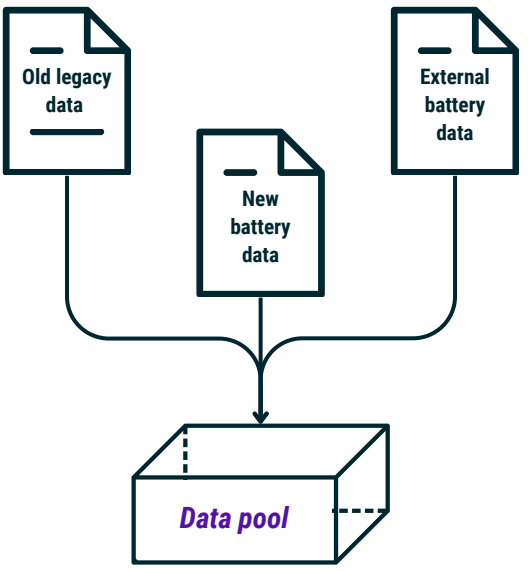


Figure 2: *Current situation - How to integrate different data sources?*

In addition, to the challenging amount of data to manage, multiple sources of this data add further complexity. OEMs work with multiple test facilities (internal/external) and development partners to generate these different data points. The quality of these data files needs to be validated, the structure has to be harmonized and ideally put in context with tests performed in-house.

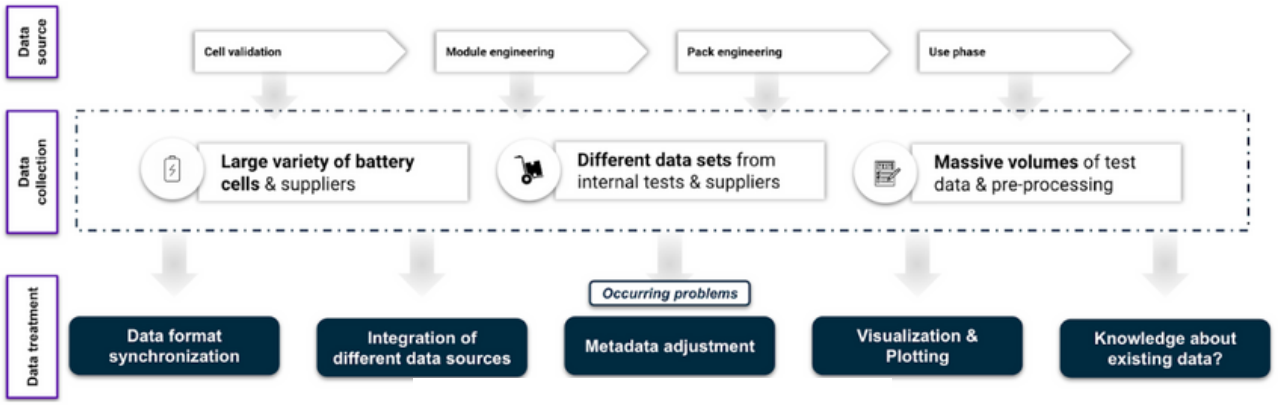


***Figure 3**: Occurring problems during test data management.*

Dependencies along the development process

Whereas an ICE engine covers more components than a battery, a battery development project presents significantly higher degrees of dependencies within the different lifecycle stages. Almost every process step and team is dependent on the understanding of the cell performance and behavior. How does a battery cell swell under certain conditions and to what special requirements does this lead on a module level? How do the thermal management and temperature gradients within the pack influence aging?

To truly understand and simulate these dependencies, data points from a single cell need to be trackable up to pack level and vice versa. Therefore, managing the lifecycle of a battery is complex and, when done correctly, leads to significant efficiencies in the product development process.

Prerequisites for effective BMS

The Battery Management System (BMS) monitors and controls the health, performance, and safety of a battery pack. The BMS and its hosted algorithms are strongly tailored to the composition of the pack, the module, and the individual cell. This software layer represents an opportunity for a competitive edge against other systems by simply maximizing the output of a given battery pack.

Not surprisingly, the quality of a BMS is highly dependent on the quality of the data that is used to parametrize its algorithms. Hence the BMS teams are one of the important data customers for cell test data.

The rise of AI and the need for new data architectures

In the introduction, we made the case for the urgent need to transform the battery development process to enable faster and, therefore, cheaper production of batteries. We are convinced that innovative AI models (known through the application of large language models (LLM) in the context of ChatGPT), can and will play a very important role in enabling this transformation. Under the hood of these models, one key innovation is the development of very powerful, and now also scalable, transformer models. These large neural networks have the potential to support the acceleration of battery development through two features. First, they will significantly accelerate the transition from large datasets to insights, forming the basis for decision-making. Secondly, they will augment engineers to contextualize and amend these insights to draw new conclusions. Further information about the architecture and innovation of this technology can be found in our white paper: "A new form of Artificial Intelligence".

Without going into too much technical detail, we want to make the point that for any industry it will be hard to argue that these AI models won't add significant value by efficiently managing processes and automatizing specific tasks/applications. However, to make use of this toolkit, the data architectures need to be rethought. Which, unfortunately, is a hard undertaking without any shortcuts. Now why is this a unique opportunity for battery data?

In the midst of radically changing their battery development processes, companies must also reconsider their operating models from the ground up. This has to include adapting their data architecture to be compatible with artificial intelligence (AI). This task is comparatively less daunting for organizations already in the process of comprehensive transformation than for those focused solely on implementing isolated AI projects without reevaluating their foundational operating models.

In the subsequent section, we will outline in a bit more detail, how exactly data management principles should change to make sure they can integrate with AI tools.

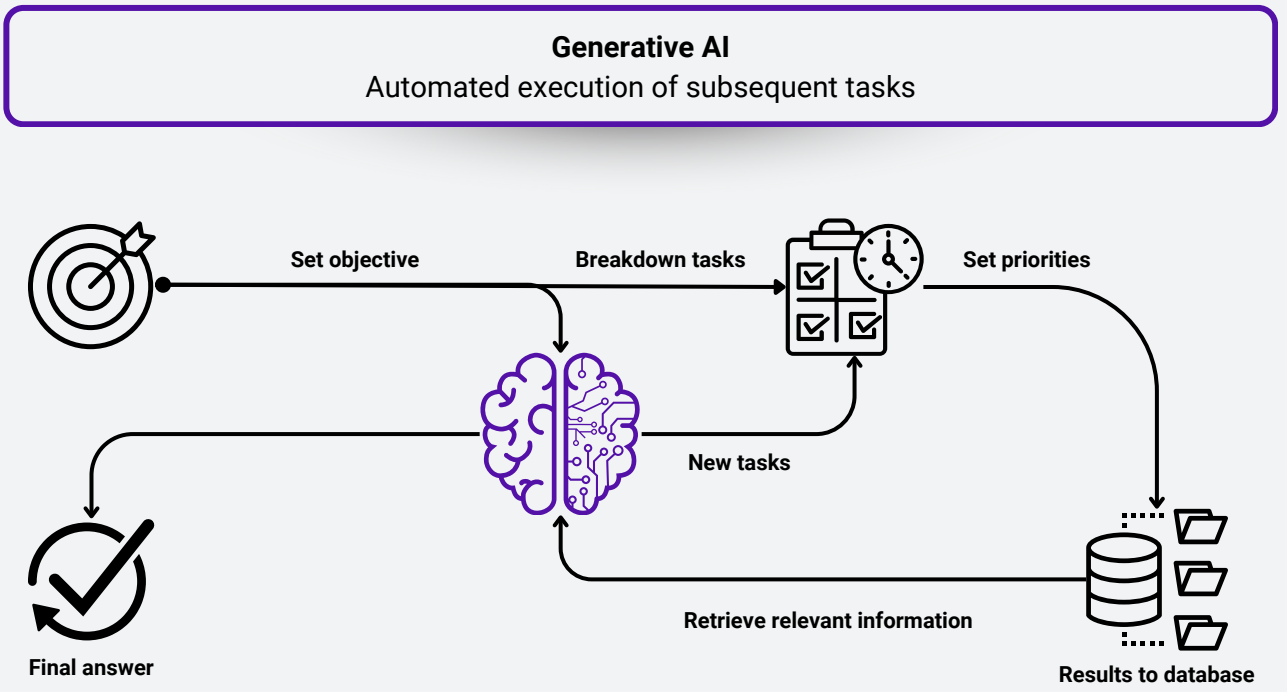


Figure 4: *Connected AI Ecosystem.*

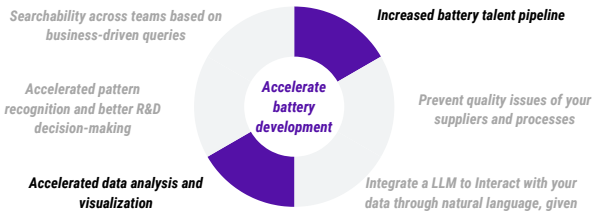
What needs to be considered when managing battery data?

Battery data, by its nature, presents a unique set of challenges and opportunities. Characterized by its high volume, diverse variety, intricate lifecycle dependencies, and significant competitive value, it sits at the intersection of high complexity and high potential. With the opportunity to deploy value-adding AI tools, we propose five essential concepts to guide the development of your data management strategies, ensuring you unlock the full potential of your battery data.

(1) Use business-friendly analytics tools

Finding highly skilled battery talent poses a significant challenge in itself; unnecessarily complicating the scenario by requiring advanced coding skills only serves to tighten the already small talent pool. In the fast-paced battery development sector, adopting easy-to-use data analytics and business intelligence software is essential - let's focus on innovation rather than coding challenges. This strategy not only accelerates development but also democratizes it, a critical advantage in a highly competitive market.

We believe that large language models (LLMs) will pivot the focus in data analytics towards the business domain, reducing the emphasis on complex data engineering tasks. Ideally, data analysis should start with a business question, avoiding the pitfalls of complex coding or time-consuming visualization tasks. For battery engineers, this shift can significantly streamline their work, making the process quicker and more focused on delivering actionable insights.

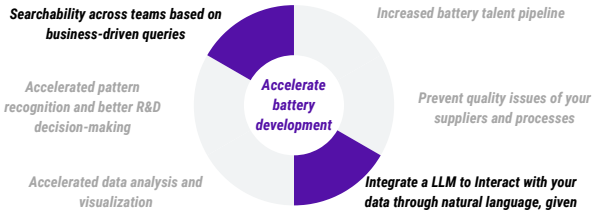


(2) Manage your metadata

Metadata is data about data. In the realm of battery data, this means for example: the type of test performed, the nominal capacity of the cell, the test equipment used, or simply the time and date of the test. Metadata operates largely behind the scenes, serving as the scaffolding that supports data architecture rather than the star performer. It's the backstage crew of the data show: essential but out of the spotlight.

For the majority who interact with data at a surface level, the indirect benefits of metadata do not immediately register in terms of direct application or impact, making its value less apparent.

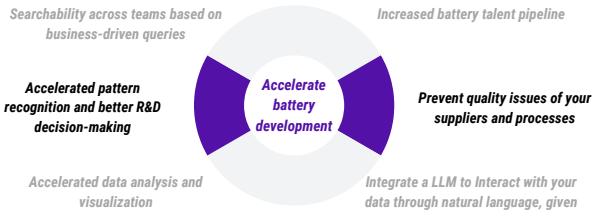
We believe that for battery data, however, metadata is particularly important, because of its ability to connect and contextualize data frames for use cases that are not particularly clear at the point of data generation.



(3) Standardize your data quality checks

Managing the data quality of data generated by your internal teams is one thing. But when you are relying on the input of different external data sources, which is naturally the case in a battery development process, diligent data quality management is key to getting the most out of your data.

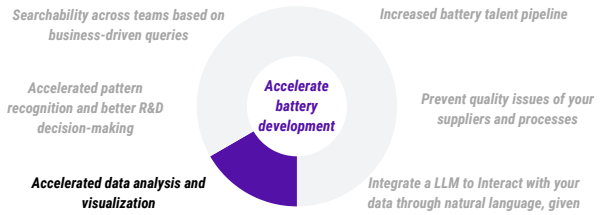
The good thing is that, in most cases, a battery development team knows exactly what a "gold standard" data set should look like. Let's make use of that! It truly is a low-hanging fruit to include the most important data quality checks as part of your data ingestion process.



(4) Automatize data treatment and visualization

After an ideally automated data quality check - to transform raw data to insights and KPIs that are used by a downstream development team - a lot of different activities need to be performed. Feature engineering, data transformation, data visualization, and interpretation. We conducted a survey with 10 different battery cell testing projects and the results are stunning, yet not surprising. A test engineer, on average, spends 2 hours alone on developing graphs (in mostly Excel) from test raw data for a single cell. Time that could be better spent on more critical tasks, especially since, in the battery development process, every day is crucial.

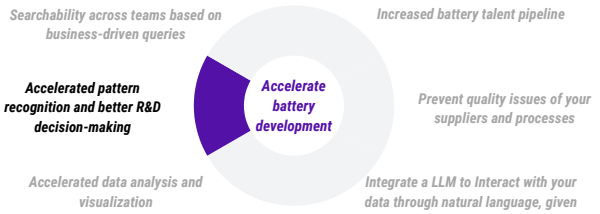
Data analysis and visualization tasks have to be automated and managed in a way so that, upon user request (ideally through an LLM - see concept 1), they can transform your raw data into business intelligence. Now the question is, how to build all data treatment algorithms and scripts in a scalable way? Luckily, based on our experience, 80% of the user queries can be covered by no more than 50 data treatment scripts. However, the underlying parameters of every general script can differ for every individual test. That's why it is important to write these scripts flexibly so that the parameters for each test can be tailored by the (non-techie) end user through a config file or even a fancy UI: reach out if you are interested in getting access to our library of generic battery data transformation scripts.



(5) Allow your engineers to tag data

The key distinction in the development process between an ICE car (or corded power tool) and an EV (or cordless power tool) lies in the end product's reliance on the unpredictable nature of electrochemical reactions. How does this affect data management? The relationship between two data points becomes much less straightforward, potentially influenced by hundreds of factors. Therefore, it's crucial for your development process to evolve with each new piece of data. Enabling your development team to annotate data with comments, questions, or observations about their findings can be immensely valuable. For example, a team member might note, "Comment: the swelling was more pronounced on the anode side." This approach empowers your team to contribute to a collective understanding and continuous improvement process.

These additional, retrospectively, and manually added metadata can become immensely valuable when trying to identify patterns (for your AI) and correlations.



Bringing it all together - our vision of battery data management

We have developed a battery data management operating model in which your battery engineer can initiate data inquiries in natural language about a specific challenge or question (Concept 1). This smart architecture, powered by LLMs, would then pinpoint the ideal data set (Concept 2), ensure its quality (Concept 3), choose the appropriate data treatment method (Concept 4), and generate not only the pertinent graph but also propose potential technical hypotheses (Concept 5).
This streamlined process promises to enhance decision-making and accelerate battery development

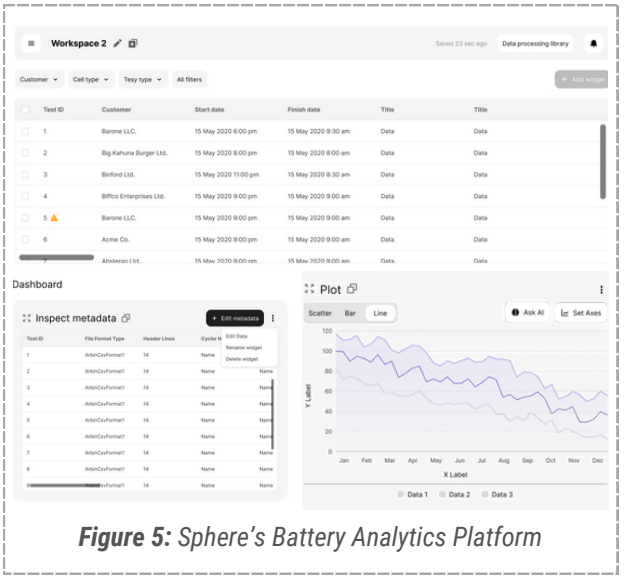


Figure 5: *Sphere's Battery Analytics Platform*

Sphere's approach

At Sphere, we have developed a data analytics platform that is based on the data management concepts introduced in this white paper with the sole purpose of accelerating battery development.

Through our advanced battery cell testing center, we have built a solution to make us faster in transforming raw battery data into insights. Proven by dozens of battery development projects and terabytes of battery cell data, we now want to share our knowledge and tools with the market: from battery data strategies, and battery fine-tuned LLMs, to the implementation of automated data quality and visualization algorithms. We are very keen to get your input into our solution and discuss how we can collaborate.

Use cases of how improved data management can accelerate your product development

Improved R&D based on field data

Faster cell selection

Integration of third parties

Contact us to find out more:

info@sphere-energy.eu

About the authors

Luca Scherrer
Co-Founder of Sphere Energy

Dr. Lukas Lutz
Co-Founder of Sphere Energy