EuroSun2026

TAXONOMY-BASED MAPPING OF EXISTING TES DATABASES: COVERAGE GAPS AND INTEROPERABILITY OPPORTUNITIES

Summary

A systematic mapping and characterization of the TES database landscape is proposed using a taxonomy. The framework is organized by data-generation pathway and level of abstraction, and coupled with a hierarchical structure that maps internal data dependencies and data-flow logic. Based on a survey of 30 databases, diversity across key analytical dimensions – namely scope, methodology, data types, provenance representation, and quality assurance – is quantified, and recurring trends are identified. The findings are further disaggregated by thermal energy storage methods, revealing well-populated and underrepresented knowledge areas across sensible, latent, and thermochemical heat storage technologies. The work aims to foster discussion on more effective cross-database interoperability and to accelerate collaboration within the TES research community.

Keywords: Thermal storage materials, database taxonomy, knowledge mapping, interoperability

Field of application for the findings: Thermal Energy Storage

Addressed audience: Oral presentation

1. Introduction

Over the last decades, thermal energy storage (TES) has gained increasing prominence as a key technology for decarbonizing the energy system. This has accelerated research and innovation efforts, resulting in the generation of large volumes of data. However, despite the growing number of TES databases and repositories, the overall structure of the available data landscape remains only partially characterized, particularly with respect to coverage across database (DB) classes, gaps in the representation of specific knowledge domains, and the potential interoperability among existing resources. Systematic mapping of database coverage across TES technologies can provide a basis for identifying which knowledge domains, data types, and methodological approaches have been most extensively developed and curated, while also revealing areas that remain fragmented or underrepresented. When further resolved by storage method – such as sensible heat storage (e.g., water), latent heat storage with phase change materials (PCMs), and thermochemical storage including sorption technologies and thermochemical materials (TCMs) – the analysis may also reveal opportunities for transferring knowledge and methodological approaches. In this regard, the current TES data space is inherently multidimensional, spanning multiple levels of abstraction and descriptor types. It ranges from "material" properties – such as structural attributes available, e.g., in IZA Database of Zeolite Structures *(International Zeolite Association, 1996)*; chemical thermodynamics and kinetics parameters reported, among others, in NIST/ARPA-E Database of Novel and Emerging Adsorbent Materials (ISODB) *(National Institute of Standards and Technology, 2016)*; and thermophysical properties, for example in the NIST Chemistry WebBook (SRD 69) *(National Institute of Standards and Technology, 1996)* – to "TES system" metrics, as in the Thermal Energy Storage Technologies Dataset *(ECHO Project, 2024)*, including thermal efficiency, system lifetime, together with other techno-economic metrics. In addition, currently available TES data are generated through heterogeneous, often non-standardized methodological pathways, including experimental campaigns, correlation-derived approaches, and physics-based modelling. This methodological heterogeneity propagates directly to the database level, where differences in definitions, testing conditions, calibration protocols, and uncertainty quantification practices often limit comparability and hinder data reuse.

This contribution surveys 30 TES databases that are, to the best of the authors' knowledge, among the most relevant and widely used in research projects. It applies a taxonomy framework to map database coverage across storage technologies, knowledge domains, and database classes, thereby revealing both well-populated and underrepresented areas of the TES data landscape. On this basis, the study also identifies interoperability opportunities among existing resources, particularly where complementary coverage, shared descriptors, or partially aligned data structures could support stronger cross-database collaboration. To illustrate the practical implications of the proposed framework, two DBs – the IEA Energy Storage TCP Task 48 "Thermal Materials Database" and the AIT's "slPCMlib solid-liquid Phase Change Materials library" – are discussed in detail.

2. Taxonomy and Domain Mapping of TES Databases

A preliminary taxonomy framework is proposed to classify TES databases according to the primary mode of data generation: "*Fundamental*" if derived from first-principles governing equations and constitutive relations; "*Simulation*" if generated by physics-based computational models that numerically solve mathematical/physical formulations to predict system properties and behavior across operating conditions; "*Correlation-based*" if learned from experimental and/or theory-informed datasets via correlation/regression/ML; "*Experimental*" if obtained under controlled laboratory testing or measurements protocols; "*Application-oriented*" if generated from field performance evidence, potentially including indicators such as technology readiness level (TRL) or economic metrics relevant to a specific market-ready technology. Furthermore, database classes are hierarchically distributed across the following abstraction levels: "*L0: Fundamental*" (e.g., first-principles-derived properties or representations), "*L1: Material*" (e.g., PCM), "*L2: Component*" (e.g., PCM integrated in tank), "*L3: System*" (e.g., PCM storage tank integrated in solar thermal system). Moreover, material-level data can be further classified according to the provenance of the underlying domain knowledge into *"Structural"* (e.g., crystallinity, pore structure, morphology), *"Chemical"* (e.g., chemical composition, corrosion stability), and "*Thermophysical"* (e.g., specific heat capacity, thermal conductivity, density) categories. Notably, database classes may span different levels of abstraction; for instance, simulation data may be generated at the material level (e.g., via MD simulations), as well as at the component level (e.g., via CFD-based simulations of component-scale heat transfer), and at the system level (e.g., via multiphysics simulations of overall system operation). The resulting matrix highlights potential interconnections among different storage technologies, as well as opportunities for knowledge and methodological transfer between them. The presentation will provide a visual mapping of these relationships, further complemented by representative examples of typical descriptors and their underlying data-generation pathways, resolved by storage method, i.e., sensible, latent, and thermochemical heat storage.

The domain mapping of the reviewed 30 TES databases according to the proposed taxonomy reveals that, allowing for non-exclusive assignment across classes and abstraction levels (i.e., databases may belong to multiple categories), the distribution is clearly dominated by material-level DBs (27 in total). In particular, thermophysical properties DBs represent the largest share (12 experimental, 10 correlation-based, 2 simulation), followed by chemical DBs (6 experimental, 8 correlation-based, 5 simulation) and structural DBs (6 experimental, 4 correlation-based, 3 simulation). This indicates a strong emphasis on material characterization rather than component- or system-level integration, and suggests that these resources are comparatively mature and largely self-contained, rather than dependent on other database types. Across these material datasets, experimental and correlation-based databases are most prevalent, denoting higher reliance on measured data and data-driven modeling over purely fundamental or simulation-based approaches. Simulation databases are found only at the material level, suggesting limited data aggregation and standardization at higher abstraction levels. Moreover, component and system levels (L2–L3) are underrepresented (4 in total), confirming a gap in shared, aggregated datasets addressing integrated TES components and end-to-end systems. Finally, fundamental DBs are only sparsely represented as standalone resources (2 in total). Instead, fundamental knowledge appears to be embedded indirectly within correlation-based and simulation databases. When resolved by storage technology, the mapping reveals additional distinct patterns. Thermochemical storage exhibits the broadest distribution across database classes and is the only technology represented by standalone fundamental databases, reflecting the greater complexity and broader knowledge representation of the underlying physical and chemical principles. Notably, application-oriented databases at system level are observed only for sensible heat storage, consistent with the higher maturity, broader deployment, and greater availability of operational and in-situ performance data for these TES technologies. Latent heat storage DBs are concentrated primarily at the material thermophysical level, especially in experimental and correlation-based databases, while remaining weakly represented at both the fundamental and higher abstraction levels. This pattern aligns with the literature, which has primarily focused on calorimetric characterization of PCMs, whereas more recent studies have started addressing this gap through more fundamental thermodynamic approaches, including entropy-based analysis *(Mehling, 2025)*.

3. IEA Energy Storage TCP Task 48 Thermal Materials Database

ES TCP Task 48 aims to accelerate research and development of TES materials, and advances the field by developing robust measurement guidelines for key material properties, collecting new and innovative TES materials alongside their synthesis and production pathways, and establishing a "Thermal Materials Database" to support technology development and deployment. The DB (available under: https://thermalmaterials.org/) can be classified as "Experimental" at "Material" level (L1). Designed to deliver high-quality, standards-compliant thermophysical data for PCMs and TCMs, the DB aligns with established calorimetry guidelines *(Gschwander et al., 2016)* and reports on measurement uncertainty for the published measurement values. Database users can consult, download, and compare measurement datasets via interactive plotting and according to the FAIR principles. Covered properties include, for example: enthalpy, viscosity, and thermal conductivity (PCM), as well as gas isotherms/isobars, and density (TCM). Users are encouraged to contribute to the database expansion via a structured data submission workflow, including a step-by-step upload guide. All submissions undergo mandatory Task-member quality review prior to publication.

4. AIT solid-liquid Phase Change Materials library slPCMlib

The solid-liquid Phase Change Materials library slPCMlib (available under: https://slpcmlib.ait.ac.at/) can be classified as "Correlation-based" database at "Material" level (L1) for latent heat storage materials. It contains data, correlations and numerical code for the computation of so-called "effective" thermal material properties of solid/liquid PCMs showing a non-isothermal phase transition behavior. Supported simulation software comprises Python, Matlab, Modelica/Dymola, and Ansys Fluent, for which the computer code can be directly exported *(Giordano et al., 2025)*. The database supports interactive plotting, data search, and as additional feature numerical code generation for user-defined PCMs not contained in slPCMlib.

5. Conclusions

This study provides a first systematic characterization of the TES data landscape, revealing coverage gaps and opportunities for knowledge and methodological transfer across TES technology domains, while resolving the emerging patterns by storage method. The proposed taxonomy framework is intended to stimulate discussion on how interoperability could be strengthened across different databases, as exemplified by the IEA ES TCP Task 48 "Thermal Materials Database" and AIT's "slPCMlib". Further aspects and implementation details will be addressed in the corresponding presentation.

6. References

ECHO Project (2024) Thermal Energy Storage Technologies Dataset. Available at: https://lookerstudio.google.com/reporting/1dd662a1-938f-4dd8-91a9-ef70ccdd72ab/ (Accessed: March 8, 2026).

Giordano, F. et al. (2025) "Statistical analysis of heat transfer with non-isothermal melting in 116 commercial PCMs using the apparent heat capacity method in ANSYS Fluent," Journal of Energy Storage, 140, p. 118793. Available at: https://doi.org/10.1016/j.est.2025.118793.

Gschwander, S. et al. (2016) "Standardization of PCM Characterization via DSC." Available at: http://dx.doi.org/10.18462/iir.pcm.2016.0008.

International Zeolite Association (1996) IZA Database of Zeolite Structures. Available at: https://www.iza-structure.org/databases/ (Accessed: March 8, 2026).

Mehling, H. (2025) "Review of the Use of Entropy to Understand the Thermodynamics of Pure-Substance PCMs," Entropy, 27(11), p. 1130. Available at: https://doi.org/10.3390/e27111130.

National Institute of Standards and Technology (1996) NIST Chemistry WebBook (SRD 69). Available at: https://webbook.nist.gov/chemistry/ (Accessed: March 8, 2026).

National Institute of Standards and Technology (2016) NIST/ARPA-E Database of Novel and Emerging Adsorbent Materials (ISODB). Available at: https://adsorption.nist.gov/isodb/ (Accessed: March 8, 2026).