3rd NFDI4Energy Conference: Defining Energy Research Data Management Together
Proceedings to the 3rd NFDI4Energy Conference - Full Papers
https://doi.org/10.52825/ocp.v9i.3317

Published: 23 Mar. 2026

d-post

Device-Side Data Capture for FAIR Laboratories

James Fitz[1,*], Alexander Krause[1], Kashfia Mahin[1], and Carsten Schilde[1]

[1]Technische Universität Braunschweig, Institut für Partikeltechnik, Germany

*Correspondence: james.fitz@tu-braunschweig.de

Abstract. Laboratory research often depends on heterogeneous, frequently legacy instruments that export results by “dropping files” onto Windows workstations, creating provenance and metadata gaps before data reaches electronic lab notebooks and research data management (RDM) platforms. As a remedy, we present *data-post* (*d-post*), a workstation-side ingestion layer that monitors these file drops, validates lightweight human-readable naming at the moment of emission, assembles stable record folders with captured provenance, and synchronises records to an RDM platform (e.g., Kadi4Mat) through a narrow sync connector boundary.

d-post minimises coupling by relying on filesystem events and modular device plugins that handle device-specific export patterns, stability windows, and lightweight processing. Deployed at the Battery LabFactory Braunschweig on a particle size analyser, a universal testing machine, and a scanning electron microscope, it supports workflows such as grouping related artefacts, threading series via sample stems, and managing large or composite exports with appropriate stability policies and idempotent staging. These mechanisms aim to reduce operator burden and support sample-centric traceability across research, production, and characterisation chains.

Ongoing work targets broader platform support, richer metadata and ontology alignment, and tighter ELN/RDM integration, with the longer-term aim of a central raw-data repository and lab-scale knowledge graph to support discovery, reuse, and incremental automation. The *d-post* source code and configuration examples are available at https://github.com/altshiftj/d-post.

Keywords: FAIR Data, Research Data Management, Provenance, Laboratory Instruments, Data Ingestion, Filesystem Monitoring, Electronic Lab Notebooks, Battery Manufacturing

1. Introduction

University laboratories are critical infrastructures for energy-systems research and education. Yet experimental data and metadata are produced across a heterogeneous mix of instruments and practices, making reliable end-to-end provenance capture difficult. Over time, this yields results that are hard to find or reproduce and can contribute to the erosion of institutional knowledge as personnel change [1], [2], [3]. In order to prevent this knowledge loss, community guidance such as the FAIR principles emphasises machine-actionable metadata, persistent identifiers, and interoperable formats as prerequisites for findable, accessible, interoperable, and reusable research data [4], [5].

Downstream systems such as Electronic Lab Notebooks (ELNs), Laboratory Information Management Systems (LIMS), and other research data management (RDM) platforms support curation and publication workflows [6], [7]. In practice, however, platforms such as Kadi4Mat and Zenodo often receive data well after it is generated, relying on ad hoc manual steps or user-specific scripts [8]. This temporal and procedural gap between file creation and structured ingestion creates opportunities for metadata loss, naming inconsistencies, and broken traceability, particularly when staff turnover erodes local conventions.

Closing the ingestion gap requires standardising data ingress across modern and legacy instruments that expose proprietary or bespoke interfaces [9]. Standards such as OPC UA and SiLA2 offer typed, discoverable device interfaces [10], [11], [12], yet many instruments still behave as “drop-file” emitters. Workstation-side ingestion must therefore tolerate brittle exports while enforcing operator-facing rules for naming, raw formats, and contextual metadata within and across institutions.

These challenges are pronounced in interdisciplinary energy- and manufacturing-oriented research [13], [14], where long, heterogeneous process chains combine manual and automated data acquisition [15], [16]. Systems must also link process parameters with intermediate and final product properties, adding spatial and temporal complexity [17], [18]. Battery-cell manufacturing exemplifies these needs, and the Battery LabFactory Braunschweig (BLB) is such a representative setting where instrument-specific file drops and operator-dependent naming persist alongside emerging digital infrastructure efforts [15], [19], [20].

With these challenges in mind, we derive design requirements for a lightweight ingestion layer that operates adjacent to raw file emission while remaining compatible with research data environments.

- **R1 Non-invasive ingestion** Use filesystem events for low coupling.
- **R2 Pluggable device adapters** Provide modular processors to normalise heterogeneous outputs.
- **R3 Metadata enhancement & provenance** Capture contextual naming, device information, and transformation lineage.
- **R4 Interoperability hooks** Enable timely synchronisation/export into ELN/LIMS/RDM platforms.
- **R5 Operational resilience** Keep deployment simple, exposing clear logs/metrics to mitigate knowledge-loss risks.
- **R6 Extensibility & evolution** Leave headroom for new devices, richer traceability, quality gates, and metadata extraction.

To meet these requirements, we present *d-post*, which bridges raw file emission and downstream RDM platforms by capturing at emission, enforcing standardised naming, preserving native artefacts, and synchronising provenance-linked records [4], [11]. Our contributions include a plugin-based architecture that encapsulates device and workstation heterogeneity while keeping deployment lightweight, a practical scheme for naming validation and record construction at emission time, and interoperable synchronisation hooks. Finally, we present an initial deployment at the BLB that illustrates end-to-end operation and how naming enforcement and stability policies can improve operational reliability and user experience.

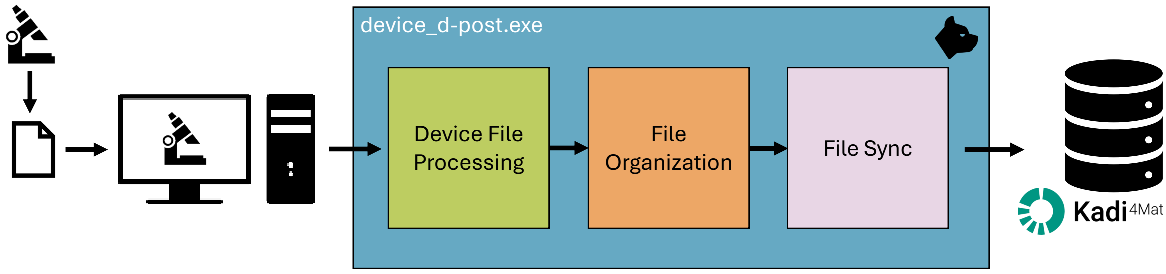


Figure 1. *Conceptual overview:* d-post *transforms raw instrument files into structured, provenance-bearing records on the workstation and synchronises them to an RDM platform (e.g., Kadi4Mat). The generic architecture supports heterogeneous devices while maintaining consistent capture and processing, followed by synchronisation to downstream platforms.*

2. Related Work

Efforts to automate data ingestion span from lightweight stream watchers to fully autonomous lab orchestrators. Complex research in chemistry, biology, and materials science, as well as mission-critical data capture and interpretation in healthcare, have encouraged such automation.

In modern healthcare settings, pipelines continuously mirror operational records to analytics backends and replace manual extraction with configurable flows, improving clinical monitoring, alert systems, and patient processing [21]. Materials science communities adopt similar concepts: CRADLE offers a distributed, data-centric environment that integrates heterogeneous instruments, storage, and analytics so workflows can be ingested, curated, and analysed at scale [22]. Automation-heavy pipelines described by Bai et al. demonstrate knowledge graph-based laboratory platforms that couple instruments, ontologies, and agents to steer experiments while integrating provenance capture directly within closed-loop optimisation routines [23], [24].

While this established science has informed this work, these environments often treat data pipelines as abstracted flows and automated analytics. The ultimate goal of laboratory automation is shared; however, the current focus of *d-post* is to provide a scientist-centric approach that encourages FAIR awareness and data sharing practices organically from the bottom up. Data originates at the device, the device is driven by the scientist, and the scientist is the ultimate consumer and steward of their data. This software aims to aid and inform scientists as they create data, facilitating research data management practices within the laboratory.

3. Methodology

d-post is written in Python and designed to run as a persistent background service on device workstations. The codebase is packaged into standalone executables, eliminating the need for operator-managed environments or dependencies. Each workstation-specific build embeds configuration via PC plugins that declare active device plugins and override default paths or naming rules as needed. Architecturally, *d-post* is designed to generalise beyond a single instrument class or a single RDM platform by separating workstation-side ingestion from platform synchronisation.

At ingestion time, it captures lightweight provenance at emission and mirrors record folders and metadata to downstream RDM platforms via connector-based synchronisation [4], [11]. Although our initial deployment targets Kadi4Mat, platform integration is concentrated in a narrow connector boundary: create or resolve a platform record, attach the local record folder, and write metadata/provenance fields. This keeps device-specific capture logic independent from platform-specific APIs and allows the same workstation behaviour to be reused across multiple platforms.

3.1 Runtime Architecture

d-post runs as a workstation-side service that monitors a configured upload folder and processes file-creation events in a single-threaded loop to keep user interactions responsive [25], [26]. Workstation configuration specifies watch/destination paths, naming rules, and the enabled device plugins. Background services expose structured logs and Prometheus metrics, complemented by a lightweight observability web UI.

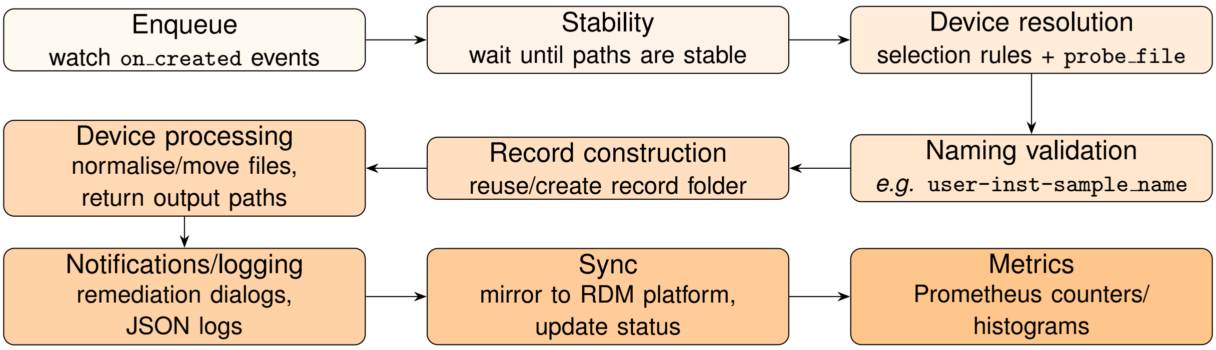


Figure 2. *Processing pipeline:* d-post *processes filesystem events into stable record folders by waiting for file stability, resolving the responsible device plugin, enforcing naming, constructing records, applying device-specific transformations, and synchronising records and provenance to a downstream RDM platform via a configurable sync connector while emitting logs and metrics.*

Instrument-specific behaviour is encapsulated in device plugins that implement a common processor interface, keeping the core pipeline generic while allowing independent evolution of device handling. Each plugin defines selection rules and any preprocessing or content probing needed to normalise outputs before appending them to the record folder. Workstation configuration binds enabled plugins to file selectors, batching rules, and metadata (e.g., identifiers, tags), ensuring only relevant devices are monitored.

Platform synchronisation is delegated to a sync connector that encapsulates the target RDM platform API for record attribution and organisation (e.g., ownership, collections, tags). This keeps connectors interchangeable: additional ELN/LIMS/RDM platform connectors can be implemented without changes to device plugins.

3.2 User Experience

d-post is designed to run unobtrusively as a background service, with user interactions relegated to exceptional cases, such as naming convention violations, where modal dialogs guide users to correct issues. To keep capture low-friction while maintaining consistency, *d-post* validates a lightweight filename schema (e.g., *user–inst–sample_name*, where `inst` denotes the user's institute abbreviation) and prompts operators to remediate violations when needed (Figure 3).

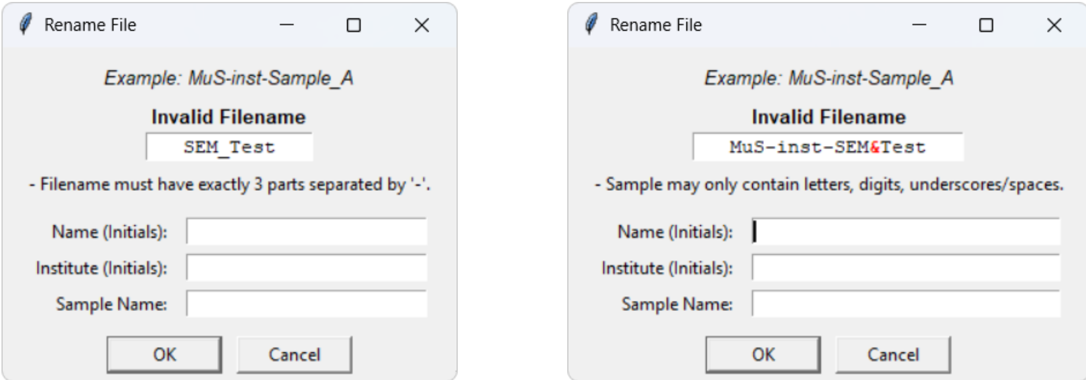


Figure 3. *Naming convention enforcement during data ingestion: prompts shown for improper formatting (left) and invalid characters (right). Comprehensive logging captures all actions and errors in structured JSON format, facilitating troubleshooting and audit trails.*

Upon proper naming and successful processing, the data is mapped to the user based on the provided filename, and the record is synchronised to the connected RDM platform. Permissions within the platform are assigned to the identified user, the generating device, and a collection of records aggregating all raw data generated by that user from all devices within the lab (Figure 4). Data backups are maintained on the respective device workstations as well.

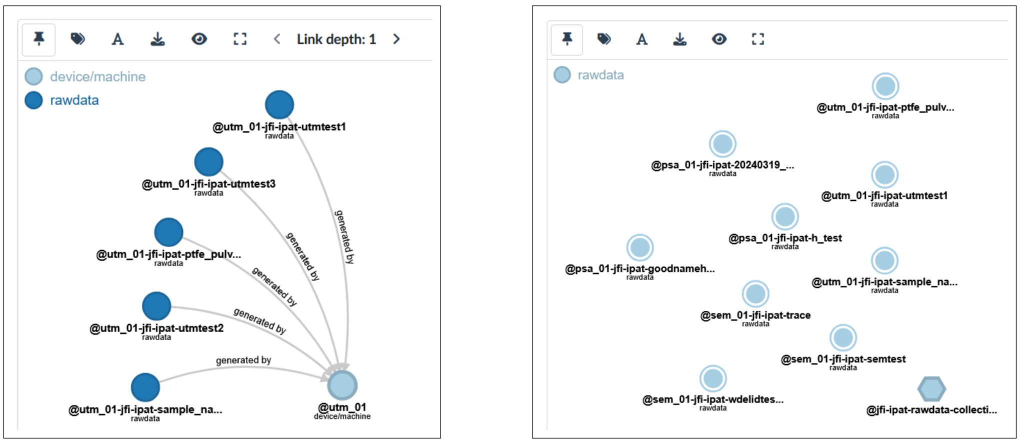
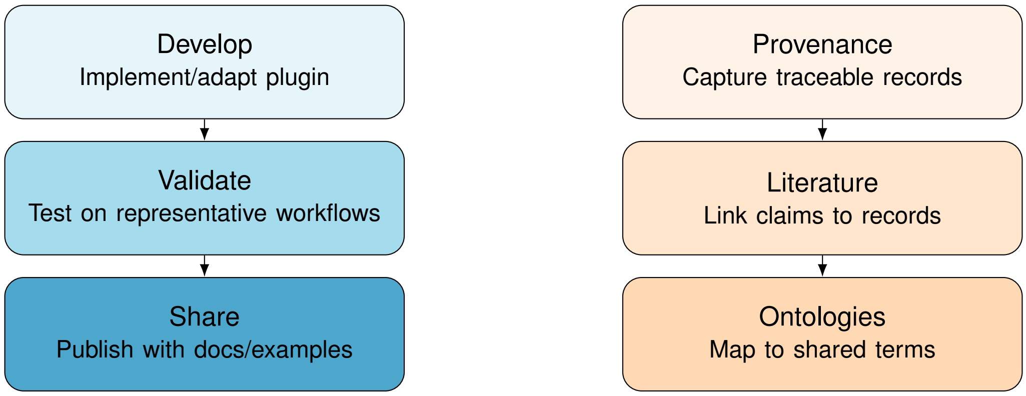


Figure 4. *Integration with Kadi4Mat: device links (left) and complete raw data collection (right) as managed by* d-post, *preserving provenance and metadata while providing an aggregate data overview for users. Here, each node is a set of files (i.e., a record) uploaded from the respective device.*

Device-specific example applications are provided in Appendix A.

4. Discussion

Building on the modular runtime architecture described in Section 3, we discuss practical scaling and near-term evolution of *d-post* along two threads: (i) widening instrument coverage through community-maintained device plugins and (ii) progressively enriching captured provenance so records can later align with shared semantics and support literature-facing traceability. Figure 5 summarises these two threads.



(a) *Community device/plugin adoption.*

(b) *Provenance–literature–ontology alignment linkage.*

Figure 5. *Two complementary evolution loops for* d-post*: a community contribution cycle that scales instrument coverage while maintaining robust, deterministic capture (left), and a progressive enrichment pathway in which captured provenance supports literature-facing traceability and later ontology alignment (right).*

4.1 Scalability Through Community Device Plugins

Long-term scale depends on widening instrument coverage beyond what a single laboratory team can maintain. *d-post* therefore follows a plugin model in which each device adapter encapsulates only the minimal device-specific logic (selection rules, preprocessing, and final placement), while the core application provides shared concerns such as naming validation, staging, record construction, logging, and sync orchestration.

A practical scaling path is community-driven: a public plugin template, contribution guidelines, fixture-based tests with representative raw files, and review criteria that check deterministic behaviour and provenance completeness. This governance model distributes maintenance effort while keeping quality controls central.

4.2 Future Work

The next phase focuses on scale and depth, pairing broader instrument coverage with richer context capture to strengthen traceability and interoperability. We aim to expand the catalogue of device plugins so *d-post* spans more lab topologies, and to embed richer in-stream metadata capture so contextual fields travel with every record and ELN upload. In parallel, we plan to retrofit legacy data with the same naming and record-construction rules to standardise historical drops.

To ensure that increased instrument coverage and richer context capture translate into reusable records, we view semantic interoperability as a valuable next step: making workstation-captured records reusable across platforms and future-proof for later integration. Rather than requiring fully specified ontologies at acquisition time, we aim for a lightweight provenance contract captured at emission that can be progressively enriched and mapped to shared terms downstream [4], [5], aligning with ontology-based traceability proposals for battery manufacturing data acquisition [16], [19].

In *d-post*, this contract acts as a *provenance spine* captured at ingestion: sample identifier, device/plugin identity, event time, record folder, and an agent/context signal inferred from naming and workstation placement. Optional links (e.g., method/protocol references, process-chain relations, derived artefacts) can be added later without increasing the operator burden at acquisition time.

With stable record identifiers (e.g., repository-backed dataset bundles) and consistent provenance, manuscripts could reference the records underlying a figure/table or key numerical claim, enabling auditable linkage from scientific claims directly to evidence. Beyond auditability, we see potential in using provenance-to-literature linkages to strengthen ontology alignment: publication context can provide an additional signal to guide mapping to existing terms and to prioritise enrichment where it matters for interpretation. Once consistent provenance exists, it can be aligned later to richer domain ontologies and broader knowledge-graph architectures without revisiting the original raw files [23], [24], [27], [28].

5. Conclusion

d-post shows that a modest ingestion layer can sit beside legacy instrumentation and assist in the production of data records that are immediately useful to downstream RDM platforms. By combining naming validation, provenance-preserving record folders, and platform synchronisation, *d-post* narrows the long-standing gap between "drop-file" practices and FAIR-minded curation without demanding protocol-level adoption from instrument vendors.

Device plugins isolate the difficult aspects of device-specific raw data normalisation into manageable modular units while higher-level services handle cross-cutting concerns. This separation of concerns keeps the core application stable while allowing incremental growth of device coverage.

Case-driven integrations at BLB further illustrate the payoff; detailed device walkthroughs are provided in Appendix A. The adoption of a particle size analyser, universal testing machine, and scanning electron microscope within the BLB suggests that upfront effort to standardise device analytical tests and enforce naming conventions can facilitate more reliable data capture and simplify research workflows for scientists. As *d-post* continues to expand, it aims to enhance both the efficiency of scientific processes and the quality of research data management.

Data availability statement

No new research data were generated or analysed in this study. Example files and instrument outputs processed by *d-post* originate from routine laboratory operations and may contain third-party, sensitive, or proprietary information; therefore, these data are not publicly available. Access to such data may be considered on reasonable request to the corresponding author, subject to institutional approvals and any applicable confidentiality, ethics, and data-protection constraints.

Underlying and related material

The *d-post* source code and configuration examples are available at https://github.com/altshiftj/d-post.

Author contributions

The authors' contributions are listed according to the CRediT taxonomy:

- **Conceptualization:** James Fitz, Alexander Krause, Carsten Schilde
- **Funding acquisition:** Carsten Schilde
- **Methodology:** James Fitz, Alexander Krause, Kashfia Mahin
- **Software:** James Fitz, Alexander Krause, Kashfia Mahin
- **Validation:** James Fitz, Alexander Krause, Kashfia Mahin
- **Writing – original draft:** James Fitz
- **Writing – review & editing:** James Fitz, Carsten Schilde

Competing interests

The authors declare no competing interests.

Funding

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under the Heisenberg Programme "Digital methods for complex systems in process and production engineering" (grant no. 500382045).

Acknowledgements

The authors thank colleagues and laboratory staff at the Battery LabFactory Braunschweig for feedback during deployment and for supporting instrument-side testing. The authors also acknowledge the use of large language model (LLM) assistance for drafting and editing text and for supporting software development; all outputs were reviewed and validated, and remain the responsibility of the authors.

A. Device-Specific Example Applications

To ground the architecture in practice, this appendix (i) recalls the end-to-end processing pipeline that turns filesystem events into stable record folders (Figure 2), (ii) summarises common workstation/device topologies observed in drop-file laboratories (Figure 6), and (iii) presents three concrete instrument integrations and their capture policies.

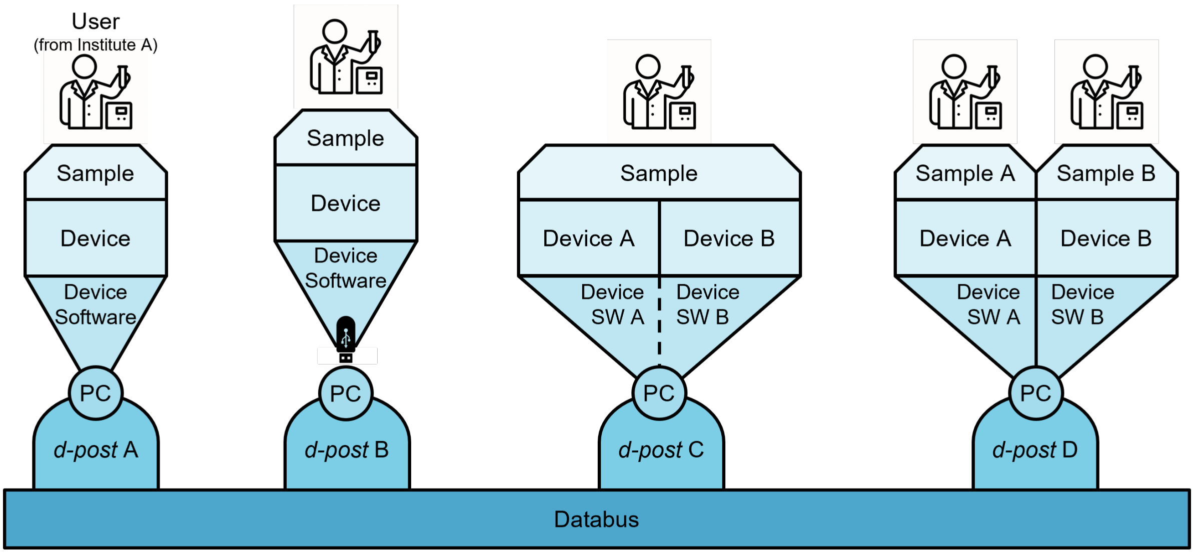


Figure 6. *Observed workstation/device configuration patterns in drop-file laboratories: multi-instrument, multi-sample data flows converge via device software and workstation plugins into a unified databus. The diagram highlights where lightweight integration work is needed to make heterogeneous device software emit predictable drops, enabling consistent capture and scalable deployment across devices, samples, and institutions.*

These patterns motivate the case studies below, which show where integration effort lives (instrument export configuration vs device plugin policy) and how heterogeneous drops can be made repeatable.

We illustrate this approach on three instruments at the BLB: a laser diffraction particle size analyser (PSA), a universal testing machine (UTM), and a scanning electron microscope (SEM). Each case starts from a concrete data-management problem and shows how *d-post* turns loose file streams into stable, sample-centric research data records.

A.1 Laser Diffraction Particle Sizing (Horiba Partica LA-960)

Laser diffraction particle sizing provides important insights into the size distribution of powders and slurries used in battery manufacture at the BLB [17], [29], [30]. The Horiba Partica LA-960 in use at the lab writes a native `.ngb` and an exported `.csv` per measurement, with data managed by bespoke software that leaves users repeatedly navigating within the device UI and file dialogs. Adopting the Partica LA-960 requires standardising file naming and configuring the manufacturer's software to automatically emit all files in a single folder for *d-post* ingestion. A flowchart depicting device-specific processing is shown in Figure 7.

Implementing *d-post* at the Horiba PSA can reduce manual handling and make capture more consistent for operators. By exposing file naming within the device software and configuring automatic saving and export to a single monitored folder,

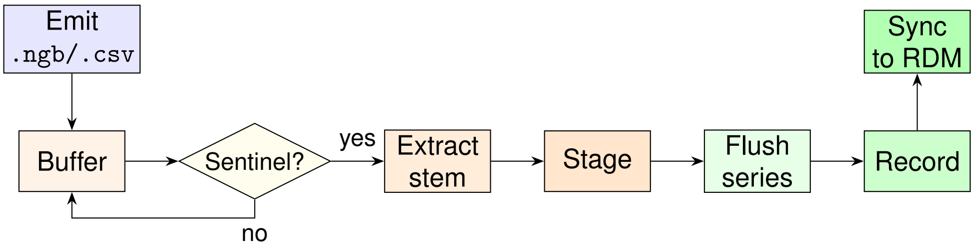


Figure 7. *Horiba PSA batch processing: d-post buffers each native `.ngb` until a matching `.csv` appears, pairing both as a single measurement. A `.csv` arriving first acts as a sentinel, triggering sample name extraction and series flush upon the final `.ngb` arrival. Measurement pairs are staged and, once complete, organised into record folders and synchronised to the RDM platform.*

measurement-time file handling is reduced. Researchers can focus on sample preparation and measurement rather than file-management while *d-post* validates naming, groups artefacts into record folders, and synchronises records to the RDM platform with captured provenance.

A.2 Universal Mechanical Testing (Zwick/Roell UTM)

Universal mechanical tests at the BLB help scientists explore the adhesive strength of materials to electrodes, as well as conductive behaviour of electrodes and materials under pressure [31], [32], [33]. When measuring, the Zwick/Roell Z020 UTM writes a native `.zs2` file per run and exports a single Excel workbook (`.xlsx`) at the end of testing, containing the per-run data and the statistical summary. As with the particle sizer, the device is constrained via its software to centralise naming and file sequencing so that *d-post* receives orderly, predictable drops. Figure 8 depicts the device-specific processing flow.

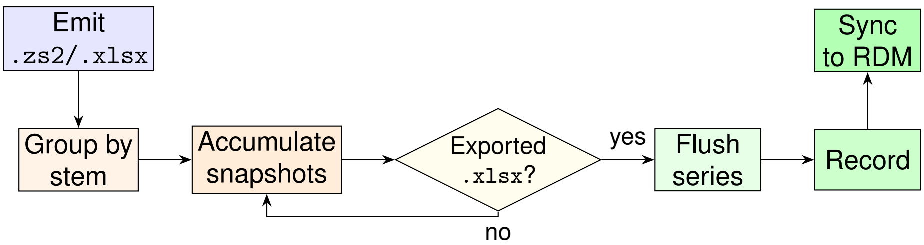


Figure 8. *Zwick UTM series processing: d-post groups native `.zs2` outputs by stem, accumulates run snapshots, and flushes the series when the exported Excel workbook (`.xlsx`) appears. Complete records are synchronised to the RDM platform.*

At the UTM, *d-post* can improve data consistency and reduce operator overhead. Automated saving and exporting to a single location, together with incorporating naming conventions within the device software, reduces manual file handling. This can save time and reduce the risk of data loss due to misnaming or misplacement.

A.3 Electron Microscopy (Phenom XL2 SEM)

In battery manufacture and research, scanning electron microscopy enables detailed visual analysis of electrodes and material deposition as well as elemental composition [34]. The Phenom XL2 SEM at the BLB emits `.tif` images and element identification (ELID) export directories that mix descriptors, images, and spectroscopy results. *d-post* adoption of the device requires little beyond enforcement on a sample naming field, except that ELID exports are written incrementally with temporary files appearing and vanishing unpredictably, requiring a robust stability policy before ingestion (Figure 9).

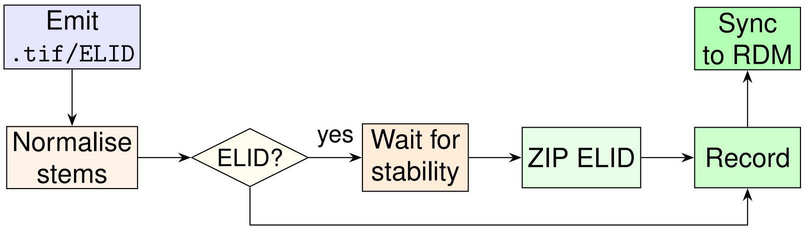


Figure 9. *Phenom SEM processing: d-post normalises image stems, applies stability windows, and archives ELID exports. `.tif` files bypass archiving and are added directly to records.*

At the Phenom SEM, the same pattern applies: naming enforcement and automated file handling reduce manual data-management steps, allowing researchers to focus on microscopy tasks.

A.4 Cross-Cutting Observations

Across these three integrations, the device-specific portion of the capture policy is small but essential: batch-completion triggers, stability windows for incremental or composite exports, and repeat-safe staging and finalisation so that repeated filesystem events converge to a deterministic record update. Within this envelope, *d-post* applies a consistent workstation-side pattern—deferred finalisation based on reliable triggers, idempotent staging of coherent bundles, preservation of native binaries, and sample-centric scoping—to support consistent, audit-ready records with low operator overhead and a reusable template for future device integrations.

References

[1] H. Collins, *Tacit and explicit knowledge*, eng. Chicago, Ill: University of Chicago Press, 2013, ISBN: 978-0-226-11382-1 978-0-226-00421-1.

[2] N. Galan, "Knowledge loss induced by organizational member turnover: A review of empirical literature, synthesis and future research directions (Part I)," en, *The Learning Organization*, vol. 30, no. 2, pp. 117–136, Apr. 2023, ISSN: 0969-6474, 0969-6474. DOI: 10.1108/TLO-09-2022-0107. Accessed: Nov. 7, 2025. [Online]. Available: http://www.emerald.com/tlo/article/30/2/117-136/379749.

[3] N. Galan, "Knowledge loss induced by organizational member turnover: A review of empirical literature, synthesis and future research directions (Part II)," en, *The Learning Organization*, vol. 30, no. 2, pp. 137–161, Apr. 2023, ISSN: 0969-6474, 0969-6474. DOI: 10.1108/TLO-09-2022-0108. Accessed: Nov. 7, 2025. [Online]. Available: http://www.emerald.com/tlo/article/30/2/137-161/379778.

[4] M. D. Wilkinson et al., "The FAIR Guiding Principles for scientific data management and stewardship," en, *Scientific Data*, vol. 3, no. 1, p. 160 018, Mar. 2016, ISSN: 2052-4463. DOI: 10.1038/sdata.2016.18. Accessed: Nov. 4, 2025. [Online]. Available: https://www.nature.com/articles/sdata201618.

[5] J. Leipzig, D. Nüst, C. T. Hoyt, K. Ram, and J. Greenberg, "The role of metadata in reproducible computational research," en, *Patterns*, vol. 2, no. 9, p. 100 322, Sep. 2021, ISSN: 26663899. DOI: 10.1016/j.patter.2021.100322. Accessed: Nov. 7, 2025. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2666389921001707.

[6] N. CARPi, A. Minges, and M. Piel, "eLabFTW: An open source laboratory notebook for research labs," *The Journal of Open Source Software*, vol. 2, no. 12, p. 146, Apr. 2017, ISSN: 2475-9066. DOI: 10.21105/joss.00146. Accessed: Dec. 14, 2025. [Online]. Available: http://joss.theoj.org/papers/10.21105/joss.00146.

[7] N. Brandt et al., "Kadi4Mat: A Research Data Infrastructure for Materials Science," en, *Data Science Journal*, vol. 20, p. 8, Feb. 2021, ISSN: 1683-1470. DOI: 10.5334/dsj-2021-008. Accessed: Nov. 3, 2025. [Online]. Available: http://datascience.codata.org/articles/10.5334/dsj-2021-008/.

[8] CERN, "Zenodo," CERN, Tech. Rep., 2013. [Online]. Available: https://zenodo.org.

[9] M. Porr et al., "Bringing IoT to the Lab: SiLA2 and Open-Source-Powered Gateway Module for Integrating Legacy Devices into the Digital Laboratory," en, *HardwareX*, vol. 8, e00118, Oct. 2020, ISSN: 24680672. DOI: 10.1016/j.ohx.2020.e00118. Accessed: Nov. 5, 2025. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2468067220300274.

[10] OPC Foundation, "OPC Unified Architecture Specification," OPC Foundation, Tech. Rep., 2023, Accessed: 2025-11-25. [Online]. Available: https://reference.opcfoundation.org/.

[11] M. Porr et al., "Implementing a digital infrastructure for the lab using a central laboratory server and the SiLA2 communication standard," en, *Engineering in Life Sciences*, vol. 21, no. 3-4, pp. 208–219, Mar. 2021, ISSN: 1618-0240, 1618-2863. DOI: 10.1002/elsc.202000053. Accessed: Nov. 5, 2025. [Online]. Available: https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/elsc.202000053.

[12] D. Juchli, "SiLA 2: The Next Generation Lab Automation Standard," en, in *Smart Biolabs of the Future*, S. Beutel and F. Lenk, Eds., vol. 182, Series Title: Advances in Biochemical Engineering/Biotechnology, Cham: Springer International Publishing, 2022, pp. 147–174, ISBN: 978-3-031-12581-2 978-3-031-12582-9. DOI: 10.1007/10_2022_204. Accessed: Nov. 5, 2025. [Online]. Available: https://link.springer.com/10.1007/10_2022_204.

[13] A. Turetskyy, "Data Analytics in Battery Production Systems," ISBN: 9783802783685, Ph.D. dissertation, Oct. 2022. DOI: 10.24355/dbbs.084-202209091126-0.

[14] G. Ventura Silva, M. Thomitzek, T. Abraham, and C. Herrmann, "Simulation-based Assessment of Energy Demand and Costs Associated with Production Scrap in the Battery Production," *SNE Simulation Notes Europe*, vol. 32, no. 3, pp. 143–150, 2022, ISSN:

23060271. DOI: 10.11128/sne.32.tn.10614. Accessed: Jan. 17, 2024. [Online]. Available: https://www.sne-journal.org/10614.

[15] A. Turetskyy, S. Thiede, M. Thomitzek, N. von Drachenfels, T. Pape, and C. Herrmann, "Toward Data-Driven Applications in Lithium-Ion Battery Cell Manufacturing," en, *Energy Technology*, vol. 8, no. 2, p. 1 900 136, 2020, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ente.201900136, ISSN: 2194-4296. DOI: 10 . 1002 / ente . 201900136. Accessed: Jan. 8, 2024. [Online]. Available: https : / / onlinelibrary.wiley.com/doi/abs/10.1002/ente.201900136.

[16] J. Wessel, A. Turetskyy, O. Wojahn, C. Herrmann, and S. Thiede, "Tracking and Tracing for Data Mining Application in the Lithium-ion Battery Production," en, *Procedia CIRP*, vol. 93, pp. 162–167, 2020, ISSN: 22128271. DOI: 10.1016/j.procir.2020.03.071. Accessed: Jan. 12, 2024. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2212827120306910.

[17] H. Dreger, M. Huelsebrock, L. Froboese, and A. Kwade, "Method Development for Quality Control of Suspensions for Lithium-Ion Battery Electrodes," en, *Industrial & Engineering Chemistry Research*, vol. 56, no. 9, pp. 2466–2474, Mar. 2017, ISSN: 0888-5885, 1520-5045. DOI: 10.1021/acs.iecr.6b02103. Accessed: Dec. 2, 2025. [Online]. Available: https://pubs.acs.org/doi/10.1021/acs.iecr.6b02103.

[18] M.-A. Filz, S. Gellrich, A. Turetskyy, J. Wessel, C. Herrmann, and S. Thiede, "Virtual Quality Gates in Manufacturing Systems: Framework, Implementation and Potential," en, *Journal of Manufacturing and Materials Processing*, vol. 4, no. 4, p. 106, Dec. 2020, Number: 4 Publisher: Multidisciplinary Digital Publishing Institute, ISSN: 2504-4494. DOI: 10.3390/jmmp4040106. Accessed: Jan. 12, 2024. [Online]. Available: https://www.mdpi.com/2504-4494/4/4/106.

[19] J. Wessel, A. Turetskyy, O. Wojahn, T. Abraham, and C. Herrmann, "Ontology-based Traceability System for Interoperable Data Acquisition in Battery Cell Manufacturing," en, *Procedia CIRP*, vol. 104, pp. 1215–1220, 2021, ISSN: 22128271. DOI: 10.1016/j.procir.2021.11.204. Accessed: Jan. 12, 2024. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2212827121011021.

[20] A. Naumann, S. Süß, M. Mennenga, and C. Herrmann, "Towards an integrated control system for a scrap-free circular production of lithium-ion batteries," en, *Procedia CIRP*, vol. 120, pp. 297–302, 2023, ISSN: 22128271. DOI: 10.1016/j.procir.2023.08.053. Accessed: Nov. 7, 2025. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2212827123007242.

[21] M. Al-Agil et al., "Enhancing clinical data retrieval with Smart Watchers: A NiFi-based ETL pipeline for Elasticsearch queries," en, *BMC Medical Informatics and Decision Making*, vol. 24, no. 1, p. 255, Sep. 2024, ISSN: 1472-6947. DOI: 10.1186/s12911-024-02633-w. Accessed: Nov. 24, 2025. [Online]. Available: https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-024-02633-w.

[22] T. G. Ciardi et al., "Materials data science using CRADLE: A distributed, data-centric approach," en, *MRS Communications*, vol. 14, no. 4, pp. 601–611, Jul. 2024, ISSN: 2159-6867. DOI: 10.1557/s43579-024-00616-6. Accessed: Nov. 3, 2025. [Online]. Available: https://link.springer.com/10.1557/s43579-024-00616-6.

[23] J. Bai, L. Cao, S. Mosbach, J. Akroyd, A. A. Lapkin, and M. Kraft, "From Platform to Knowledge Graph: Evolution of Laboratory Automation," en, *JACS Au*, vol. 2, no. 2, pp. 292–309, Feb. 2022, ISSN: 2691-3704, 2691-3704. DOI: 10.1021/jacsau.1c00438. Accessed: Nov. 14, 2025. [Online]. Available: https://pubs.acs.org/doi/10.1021/jacsau.1c00438.

[24] J. Bai et al., "A dynamic knowledge graph approach to distributed self-driving laboratories," en, *Nature Communications*, vol. 15, no. 1, p. 462, Jan. 2024, ISSN: 2041-1723. DOI: 10.1038/s41467-023-44599-9. Accessed: Nov. 14, 2025. [Online]. Available: https://www.nature.com/articles/s41467-023-44599-9.

[25] Watchdog Developers, *Watchdog: Filesystem events monitoring*, https://pypi.org/project/watchdog/, Accessed: 2025-11-25.

[26] A. K. Paul, R. Chard, K. Chard, S. Tuecke, A. R. Butt, and I. Foster, "FSMonitor: Scalable File System Monitoring for Arbitrary Storage Systems," in *2019 IEEE International Conference on Cluster Computing (CLUSTER)*, Albuquerque, NM, USA: IEEE, Sep. 2019, pp. 1–11, ISBN: 978-1-7281-4734-5. DOI: 10.1109/CLUSTER.2019.8891045. Accessed: Dec. 1, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/8891045/.

[27] J. Bai et al., "Twa: The World Avatar Python package for dynamic knowledge graphs and its application in reticular chemistry," en, *Digital Discovery*, vol. 4, no. 8, pp. 2123–2135, 2025, ISSN: 2635-098X. DOI: 10.1039/D5DD00069F. Accessed: Nov. 5, 2025. [Online]. Available: https://xlink.rsc.org/?DOI=D5DD00069F.

[28] A. Wein et al., "Developing Interoperable Ontologies for Research Data Management in the Energy Domain," in *Joint Ontology Workshops (JOWO 2024), 14th International Conference on Formal Ontology in Information Systems (FOIS 2024)*, Enschede, Netherlands, Jul. 2024.

[29] M. Weber, J. K. Mayer, and A. Kwade, "The Carbon Black Dispersion Index DI_{CB} : A Novel Approach Describing the Dispersion Progress of Carbon Black Containing Battery Slurries," en, *Energy Technology*, vol. 11, no. 5, p. 2 201 299, May 2023, ISSN: 2194-4288, 2194-4296. DOI: 10.1002/ente.202201299. Accessed: Dec. 2, 2025. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/ente.202201299.

[30] T. Grenda, H. Meer, J. Keilhofer, R. Daub, and A. Kwade, "Impact of Dissolver Setup on the Performance of Nickel-Rich Active Material Cathodes for Lithium Ion Batteries," en, *Energy Technology*, vol. 13, no. 11, p. 2 500 331, Nov. 2025, ISSN: 2194-4288, 2194-4296. DOI: 10.1002/ente.202500331. Accessed: Dec. 2, 2025. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/ente.202500331.

[31] G. Lenze, F. Röder, H. Bockholt, W. Haselrieder, A. Kwade, and U. Krewer, "Simulation-Supported Analysis of Calendering Impacts on the Performance of Lithium-Ion-Batteries," en, *Journal of The Electrochemical Society*, vol. 164, no. 6, A1223–A1233, 2017, ISSN: 0013-4651, 1945-7111. DOI: 10.1149/2.1141706jes. Accessed: Dec. 14, 2025. [Online]. Available: https://iopscience.iop.org/article/10.1149/2.1141706jes.

[32] J. K. Mayer et al., "Influence of the Carbon Black Dispersing Process on the Microstructure and Performance of Li-Ion Battery Cathodes," en, *Energy Technology*, vol. 8, no. 2, p. 1 900 161, Feb. 2020, ISSN: 2194-4288, 2194-4296. DOI: 10.1002/ente.201900161. Accessed: Dec. 2, 2025. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/ente.201900161.

[33] M. Bredekamp, L. Gottschalk, M. Peter, and A. Kwade, "Structuring Electrodes for Lithium-Ion Batteries: A Novel Material Loss-Free Process Using Liquid Injection," en, *Energy Technology*, vol. 12, no. 8, p. 2 400 256, Aug. 2024, ISSN: 2194-4288, 2194-4296. DOI: 10.1002/ente.202400256. Accessed: Dec. 14, 2025. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/ente.202400256.

[34] R. Moschner et al., "Impact of the sulfurized polyacrylonitrile cathode microstructure on the electrochemical performance of lithium–sulfur batteries," *Advanced Science*, vol. 12, no. 15, p. 2 415 436, Apr. 2025. DOI: 10.1002/advs.202415436.