Journal of Pharmaceutical and Biomedical Analysis 263 (2025) 116907

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.journals.elsevier.com/journal-of-pharmaceutical-and-biomedical-analysis

The Pistoia Alliance's methods database project: Instrument, chromatographic data system, and vendor-agnostic digital transfer of machine-readable high-performance liquid chromatography-ultraviolet methods using the allotrope data format

Pankaj Aggarwal [a], Azzedine Dabo [b], Cheng Sun [a], Vincent Antonucci [a], Wolfgang Colsman [c], Heiko Fessenmayr [d], Kenneth M. Wells [e], Juliet McComas [e], Gerhard Noelken [f], Birthe V. Nielsen [f,*]

[a] Merck & Co., Inc., Rahway, NJ, USA
[b] Medicine Development and Supply, Pharmaceutical Research and Development, GSK, Stevenage Hertfordshire, SG1 2NY, UK
[c] ZONTAL Inc, Plantation, FL, USA
[d] Agilent Technologies, Hewlett-Packard Str. 8, Waldbronn 76337, Germany
[e] Medicine Development and Supply, Pharmaceutical Research and Development, GSK, Collegeville, PA 19426, USA
[f] The Pistoia Alliance, 401 Edgewater Place, Suite 600, Wakefield, MA 01880, USA

ARTICLE INFO

Keywords:
Chromatographic Data System
HPLC
Allotrope Data Format
Interoperability
Precompetitive collaboration
FAIR

ABSTRACT

The Pistoia Alliance has successfully completed a pilot on the digital transfer of analytical High-Performance Liquid Chromatography (HPLC) methods and results between chromatography data systems (CDS) via a central data storage system using a standardized machine-readable data format, which transforms methods from paper documents to digital instructions. Critical method and result parameters in two example CDSs have been harmonized with a novel ontology and RDF-based graph data model created in this work. The authors will demonstrate how this new degree of data standardization can simplify data exchange day to day in the laboratory, and describe how the embedded semantics will position scientists to perform on-demand modern queries of data that has been automatically aggregated across vendor solutions. The developed solution was successfully tested at the analytical labs at Merck Sharp & Dohme LLC, a subsidiary of Merck & Co., Inc., Rahway, NJ, USA (MSD) and GSK, Stevenage, UK where there has been an effective transfer of HPLC information between different systems and sites to prove the concept and initial use cases. The work also begins to demonstrate the potential to realize many other use cases such as a series of critical improvements to current method transfer that eliminate the manual keying of data to reduce risk, steps, and error while improving overall flexibility.

1. Introduction

Driven by the need to improve efficiency and reduce operational costs and infrastructure, pharmaceutical companies have increased outsourcing to contract research organizations (CROs) and contract manufacturing organizations (CMOs) during the past decade, such that it is now standard practice for most companies. The global drug discovery outsourcing market was valued at USD 3.3 billion in 2021 and is expected to reach USD 12.7 billion by 2030 [1,2]. One aspect that has often been overlooked is the entry point into an outsourcing engagement, which usually is a document that uses either semi-templated or completely free text to describe the experiment or scientific procedure(s) to be followed. Accompanying analytical data from these methods often exist in a proprietary electronic format in the instrument control software of the commercial product used to acquire the data or are transcribed into a summary document containing only selected parameters judged to be most important by the author at the time the method is being transferred. These document representations of analytical methods often use inconsistent terminology and may contain unanticipated errors in documentation that were missed during review and approval stages, leading to unnecessary variability.

Similarly, sharing methods and associated data within a company or

* Corresponding author.
E-mail addresses: Methodshub@pistoiaalliance.org, birthe.nielsen@pistoiaalliance.org (B.V. Nielsen).

https://doi.org/10.1016/j.jpba.2025.116907
Received 16 September 2024; Received in revised form 16 April 2025; Accepted 16 April 2025
Available online 17 April 2025

between collaborators at different companies often employ the same types of semi-templated documents as above. Unavoidably, these practices lead to human interpretation errors, create time-intensive implementations, and reduce the reproducibility and efficiency of method exchange or transfers. One study estimates that more than 50 % (equivalent to USD $28 billion in the US alone) of preclinical research is not reproducible, and certainly some of this is the result of ambiguity or errors in reported method details [3]. The authors estimate that $2.2 M in annual savings are possible within a large pharmaceutical company (Analytical R&D department of approx. 500 analytical scientists) if a FAIR (Findable, Accessible, Interoperable, and Reusable) [4] representation of HPLC-UV methods could be implemented due to improved method searching, development, sharing, re-use, analytics, and error avoidance especially during pilot plant scale GMP processing of clinical materials. Therefore, there is a strong drive to create a consistent, digital representation of analytical methods that can be broadly applied to research and development as reliable, machine-readable instructions to instrument control software to reduce unintended "human error".

Prior efforts to create digital representations of methods have been worthy efforts, but do not yet address the many challenges and opportunities present in the modern analytical laboratory in the age of emerging data science. In their perspective, Bai and co-workers [5] conclude that the difficulty of achieving general interoperability remains an issue to be addressed, and reinforcing that a prerequisite condition towards digitalization, the absence of standardized data representations and exchange protocols, is seen as one of the critical challenges faced by the community [6]. Scientists need more than a consistent, templated representation of methods and analytical results in documents and databases, since these often require manual steps to assemble, which is time consuming, prone to errors, and requires additional cycles of review. Instead, scientists need a scalable approach to have instrument software output data in a common, semantically-enabled data format that enables greatly simplified visualizations, analyses, modeling, simulations, etc. This is needed not only to simplify *on-demand* analyses of 'like' data (e.g., across chromatography vendors), but also to accelerate insights across multiple data domains by connecting data from experimentation, samples, methods, instruments, etc, in one connected semantic graph. This work is a critical first step in showing that it is possible to digitally transfer chromatographic instrument instructions across different CDS platforms enabled by common semantics (ontologies) written into common and scalable data structures (data model) that are reusable across data domains.

1.1. Methods database project charter

The Pistoia Alliance is a not-for-profit, multi-company members' organization committed to lowering the barriers to innovation in life science R&D. It achieves this aim by improving the interoperability of R&D business processes through pre-competitive collaboration [7]. It draws its membership from pharmaceutical R&D and other life science R&D organizations, commercial information providers, technology companies, and other publicly funded research organizations. The Pistoia Alliance brings together key stakeholders to identify the root causes of R&D inefficiencies and then develops best practice recommendations and technology implementations to overcome common obstacles. Projects use a shared-risk funding model, where project members each contribute to fund projects and provide needed governance.

Method recapitulation across business units within pharmaceutical research companies was identified as a pre-competitive challenge common to most life science organizations. This challenge requires a collaborative, cross-industry resolution and seemed an ideal project for the Pistoia Alliance to charter, and HPLC-UV methods were selected as the target use case since these methods are ubiquitous in pharmaceutical laboratories. In early 2019, the Pistoia Alliance sponsor companies gathered to create the requirements and specifications for a desired digital, interoperable methods database. From this output, common HPLC-UV parameters were identified that would need to be standardized as part of the solution to meet the success criteria established for the selected HPLC-UV use case. It should be noted that this effort does not include an attempt to standardize data processing algorithms across CDS (e.g., integration parameters), only the representation of the instrument methods and processed results.

It was decided that the targeted technology solution would have to be powered by a semantically-enabled data model to provide consistent context and structure to method representations and link important concepts together to facilitate *on-demand* advanced data queries across experiments, samples, methods, and results parameters. Since HPLC-UV results are highly dependent on the simultaneous interaction of multiple method parameters, the representation of this data type should establish the durable relationship between parameters so they can be used to help interpret chromatographic results using standardized data queries of the format. The Pistoia Methods Database project selected the Allotrope Framework to create a standardized, vendor-independent, digital representation of HPLC-UV methods and the ZONTAL Life Science Data Platform as the solution to host the first example of a vendor-independent HPLC-UV methods repository to explore the possibilities of a fully FAIR ecosystem where methods and results data are truly findable, accessible, interoperable, and reusable by design without manual intervention.

1.2. The allotrope framework

Greater digitization of paper-based processes helps overcome the barriers of reproducibility and interoperability, but also sets the stage for machine learning by ensuring data is easily machine-readable from the point of creation [8]. The effort to digitalize a method description should start with the elements that can easily be parameterized and have significant impact on results. For example, in an HPLC experiment, the parameters typically provided to the CDS, sending instructions to the instrument, fall in this category. However, all method parameters are not critical to successfully reproduce a method as they may only have minor effects on results, so judged not worth standardizing. Still other parameters represented in a CDS may be specific to an instrument manufacturer's product, and therefore not possible to standardize across vendors. Employing input from subject matter experts from the sponsoring companies, 55 key HPLC-UV method parameters were selected for standardization into digital instructions as part of a graph representation of an HPLC method. The data standard selected to enable this project was the Allotrope Framework [9], which consists of multiple technology components that were applied:

- Allotrope Foundation Ontologies (AFO) is a Basic Formal Ontology (BFO)-inspired ontology [Ontologies | allotropefoundation] used to provide harmonized terminology, definitions, unique resource identifiers, and map key relationships across the 55 key HPLC-UV method concepts identified to facilitate advanced semantic searches and analytics. Use of ontologies is a critical component of the Allotrope Framework, making it the only data standard for instrument data to employ formal ontologies, which enables advanced semantic data queries.
- Allotrope Data Models (ADM) are the blueprints for how to semantically represent entities using AFO terms using the Resource Description Framework (RDF), the W3C graph model for the semantic web and a library of modeling patterns that are re-used across different domains to consistently structure data (e.g., many of the same modeling patterns are reused to model other instruments in the Allotrope catalogue, demonstrating the scalability of the approach and simplifying downstream data analyses). The constraints enforced by ADM through validation ensure the above 55 parameters will be used consistently to structure HPLC-UV methods from all sources. The relationships between these terms are now codified and

can be used in advanced data queries of not only the terms and result values, but also the relationships between them.

- Allotrope Data Format (ADF) is a high-performance binary data format based upon Hierarchical Data Format version 5 (HDF5), an open-source file format used for large, complex, or heterogeneous data with a file directory-like structure. This product consistently formats all ADMs into three sections – a data description for semantic data, a data cube for n-dimensional results, and a data package for additional arbitrary content such as images, videos, or data files. Use of this format consistently allows for automated verification that data files conform to the data standard, ensuring technologists that the files should perform as expected in downstream visualization and analysis software without any further modification. Knowing this to be true across data acquired from multiple vendor software products can reduce significant time spent on manual verification and translation. Details of the HPLC-UV data model created for this project and governed by the Allotrope Foundation are provided in the Materials and Methods section.

1.3. The ZONTAL methods database

The ZONTAL Methods Database is a module of the ZONTAL Life Science Data Platform, a GxP compliant cloud solution that provides scientific data management, workflow execution and data analytics to science-based companies [10]. ZONTAL allows for management and execution of analytical methods and analysis of raw data and results across a broad spectrum of analytical techniques. It automates the integration of instruments and laboratory informatics applications to eliminate manual data transcriptions and data movement, reduce operational cost, and increase data integrity, data quality and process efficiency. The ZONTAL platform natively uses the Allotrope Framework to contextualize and structure data into a consistent semantic data format, so it was a good solution option to build a first example of a universal methods database. Using the ZONTAL platform, this work explores what is possible when methods are stored centrally as common digital instructions regardless of where they were created, and then sent directly to a CDS to execute the method in the laboratory without manual intervention. The approach is directly scalable to methods and data from other instrument types as well beyond the HPLC-UV example in this work.

2. Experimental

2.1. Methods database approach & pilot

The Methods Database project is part of the Pistoia Alliance's Methods Hub, which is envisioned to be a digital platform where semantically interoperable analytical methods and supportive tools are available to the Pharmaceutical Industry. In 2021, after the development of the data model, a second phase of the Methods Database project began with the objective of demonstrating the ability to bidirectionally exchange HPLC-UV methods as digital instructions between a centralized methods database and two different CDS platforms, and subsequently use the exchanged methods to acquire data on HPLCs from different vendors. To achieve this goal and demonstrate interoperability, adapters needed to be built to transform the native format of HPLC-UV methods created in Empower and OpenLab™ CDS to the Allotrope LC-UV method data format, as depicted in Fig. 1. At the end of 2021, the pilot was expanded to include export of chromatographic results in Allotrope Data Format from the CDS to the ZONTAL Methods Database.

In this pilot study, ADF converters were constructed to enable digital transfer of HPLC-UV methods between chromatographic instruments and CDS across vendors. The pilot was intended to reimagine methods exchange and transfer, validate the possibility of centralized methods storage as executable instructions, and demonstrate a step towards greater cyber-resiliency if methods could be automatically recapitulated instead of being recreated. This would allow for a significant streamlining of current methods related workflows, Fig. 2, for example, methods transfer with a CRO.

Specifically, the pilot was designed to demonstrate the following capabilities via functioning software:

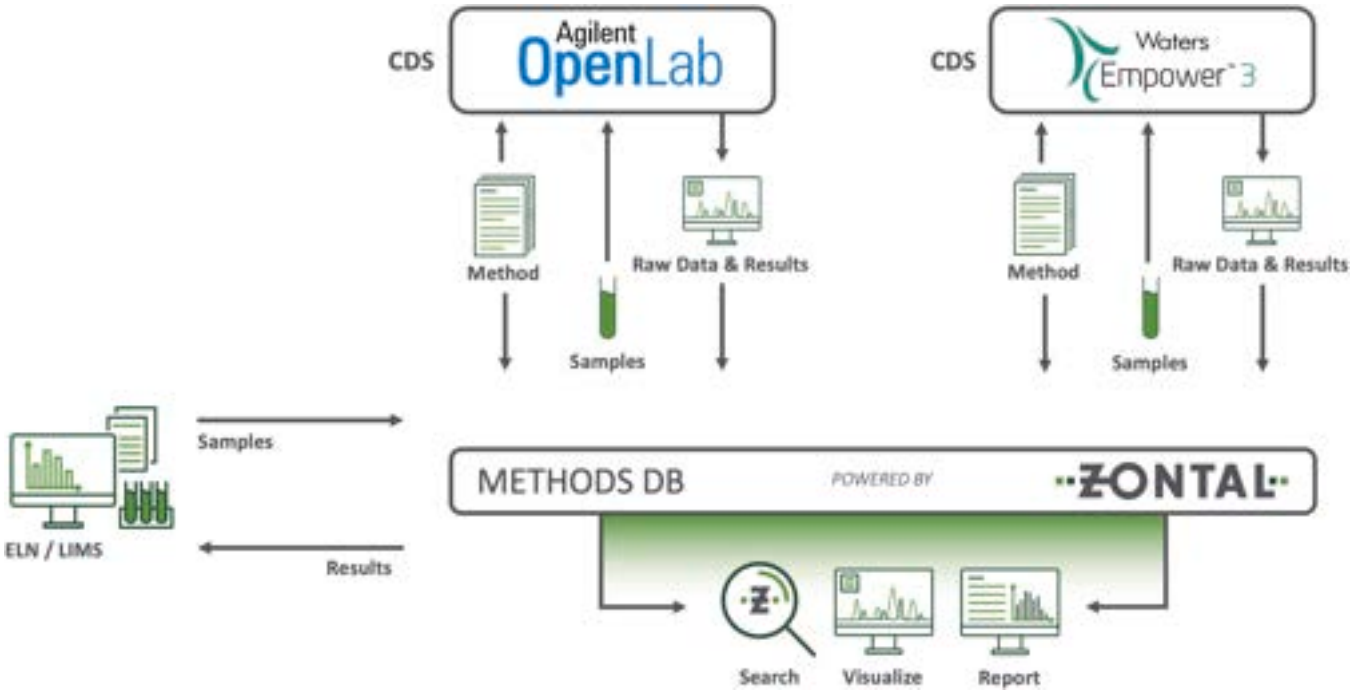

Fig. 1. Creation of CDS adapters to support method parameters via the cloud. After execution, method parameters/acquisition parameters in Water Empower™ CDS are converted from JSON to the ADF representation and exported to the Methods Database in ZONTAL space. The ADF of the desired method can be downloaded and imported through the OpenLab™ CDS interface, where the ADF adapters convert the acquisition parameters for automatic execution.

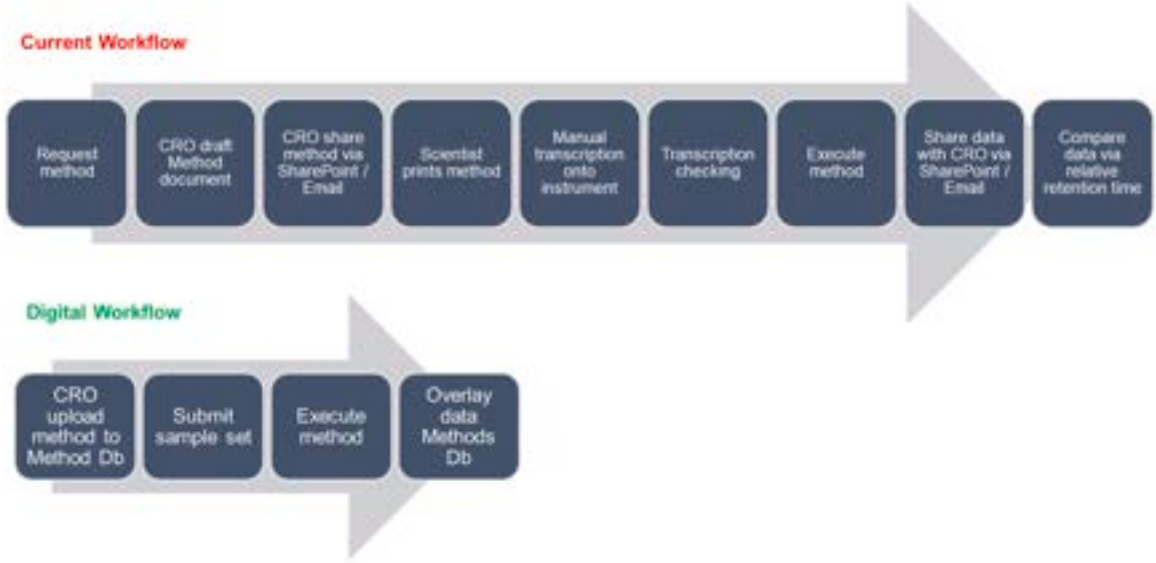

Fig. 2. A digital workflow reduces documentation effort and the chance of human transcription errors when method parameters are transferred across different instruments or between Chromatographic Data Systems.

- Digital import of HPLC-UV isocratic and gradient methods using OpenLab™ CDS and Empower™ CDS
- Digital transfer of method data between HPLC instruments (Agilent 1260 and 1290)
- Digital transfer of method data between Empower™ CDS and OpenLab™ CDS
- Sequence information via Electronic Notebook (ELN)
- On-demand visualization and analysis of method and result information using ZONTAL's analytics capabilities

To show interoperability across different CDS manufacturers, OpenLab™ CDS (Agilent Technologies) and Empower™ (Waters Corporation) were selected as these solutions are widely used in pharmaceutical development laboratories. The concept can be extended to cover other CDS manufacturers' software as well by similarly converting the native output to the same Allotrope Framework output.

2.2. The LC-UV data model

A graph HPLC-UV Class Model from the Allotrope Foundation was used to represent the HPLC experiment and the chromatographic results in this work. A second graph HPLC-UV Methods Class Model was created as part of the Methods Hub project and governed by the Allotrope Foundation. Allotrope Foundation Ontologies are publicly available at Ontologies | allotropefoundation. A schematic representation of these large class models is provided in Fig. 3, including a list of the classes that are part of each model to provide a sense of the scope addressed by each model. Individual terms included in each graph model are found in the Allotrope Ontology, but those included in the HPLC-UV Methods Class Model are also summarized in Appendix 1 for convenience. These are the core terms and graph data structure that must be used to represent chromatography output from any vendor's product if one wishes it to be compliant with the Allotrope data standard, which is programmatically verified using a schema validator. An additional >100 vendor-specific parameters (unique concepts not common across all vendor products) are also supported using a vendor-specific ontology for use cases that

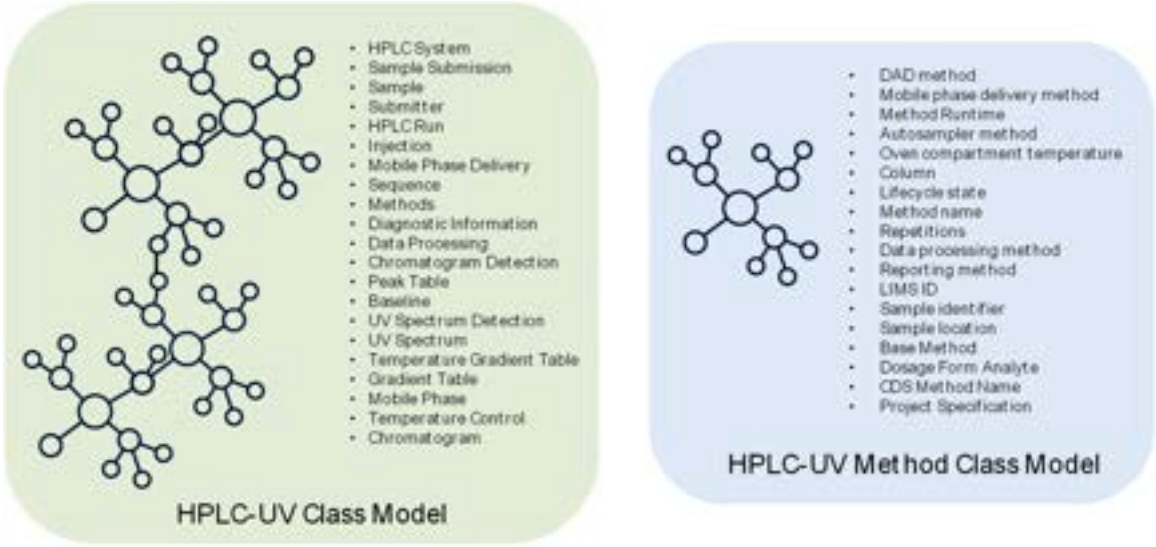

Fig. 3. Schematic representation of the HPLC-UV and HPLC-UV Method Class Models.

require this additional content for downstream queries or analytics, demonstrating the flexibility of the Allotrope data standard. The data model allows for procedural control and therefore mobile phase, or sample preparations were not included due to the complexity of these. An extension to this PoC could connect these in the semantic graph representation, however this was out of scope for this study.

2.3. CDS adapters

For an interoperable transfer between Empower™ CDS and OpenLab™ CDS, ADF adapters were built (Orbis Lab systems and Agilent Technologies). The Orbis Empower™ CDS Adapter was deployed to each Empower Citrix Server to allow direct access and control of the CDS. The Orbis Gateway establishes a fast communication to the Orbis Empower Adapter using TCP and provides a single point of contact to interact with all Empower Citrix Servers in the member's network. The configuration of the Empower™ CDS integration is managed using the Orbis Lab Data Platform, a SaaS cloud application available as a VPC and optionally on-premises, which will provide an intermediate storage layer for the in-transit data between Empower™ CDS and the ZONTAL Methods Database. The Orbis Gateway integrates with ZONTAL using RESTful web services. ZONTAL converts between the JSON format used by the Orbis Gateway and the ADF format using the Allotrope Data Model.

The Agilent OpenLab™ CDS Acquisition Method Export / Import Adapter has been embedded as a Plug-in into the OpenLab™ CDS Acquisition Client [12]. The Method Export Adapter creates an ADF file with a semantic web graph using AFO and Agilent vendor-specific taxonomy terms. Afterward, the ADF file is automatically uploaded into ZONTAL. Import is conveniently done directly from the OpenLab™ CDS Acquisition Client using the interface by selecting the ADF file to import, reading the semantic graph, and automatically setting the parameters in the target OpenLab™ CDS method.

The requirements to build these capabilities into your own software product can be obtained from the authors upon request to help facilitate the pre-competitive effort to create a FAIR data ecosystem.

2.4. LC-UV method & instrumentation - MSD

The ACQUITY UPLC/UV start up standard kit solution 2 (4.00 g/mL ± 1.0 % of 2-acetylfuran, acetanilide, acetophenone, propiophenone, butylparaben, benzophenone and valerophenone) purchased from Waters Corporation (Milford, MA, USA) was used as the test mixture for the analysis and 50:50 water:methanol was injected as blanks. All experiments were conducted using an Agilent 1290 Ultra-High-Performance Liquid Chromatography (UHPLC) system with a diode array detector (DAD). Separation was achieved using Waters ACQUITY UPLC BEH C18 Column (2.1 mm×50 mm, 1.7 μm) from Waters Corporation (Milford, MA, USA). The column temperature was controlled at 40°C throughout the experiments. A binary gradient containing 10 % methanol in 90 % water was used as mobile phase A and acetonitrile as mobile phase B, Table 1. The seal wash used was 10 % isopropanol mixed with 90 % water, and the needle wash was a 50:50 water:acetonitrile mixture. The injection volume was set to 5 μl with a mobile phase flow rate of 0.5 mL/min. The autosampler temperature was kept at room temperature, and data was collected at 254 nm with reference (360, 100 nm) and with reference off.

Table 1
Mobile phase binary gradient for analysis of theWaters start-up standard kit solution 2.

Time (min)	Mobile phase A (%)	Mobile phase B (%)
0.00	90	10
0.75	90	10
3.00	5	95
3.10	5	95
3.30	90	10
5.00	90	10

2.5. LC-UV method & instrumentation - GSK

An in-house GSK standard mix solution (uracil, theophyline, 2-acetylfuran, m-cresol, acetopheneone, propiophenone, benzofuran, butyrophenone, valerophenone, hexanophenone, and octanophenone) was used as the test mixture for the analysis and 50:50 water:acetonitrile was injected as blanks. All experiments were conducted using an Agilent 1290 UHPLC with a DAD. The chromatography was achieved using C18 Column (3.0 mm×50 mm, 2.6 μm) from Phenomenex (Torrance, CA, USA). The column temperature was controlled at 40°C throughout the experiments. A binary gradient containing TFA in water was used as mobile phase A and TFA in acetonitrile was used as mobile phase B as detailed in Table 2. Mobile phase A was composed of TFA in water and TFA in acetonitrile as mobile phase B. The injection volume was set to 1 μl with a mobile phase flow rate of 1.0 mL/min. The auto sampler temperature was kept at room temperature and data was collected at 220 nm with reference off.

3. Results and discussion

The manuscript presents the first successful transfer of instrument parameters from a proprietary format to a vendor-neutral standard format using the Allotrope Foundation Ontologies. This consensus data schema, based on semantic principles, was created to consolidate many existing vendor output formats onto a single format (context, structure, language) that will make it easier to aggregate data for further analysis (ML / AI). Method parameters covered in the Methods Database were successfully exported from the CDS (OpenLab™ CDS or Empower™ CDS) into a base method stored in ZONTAL. The method was successfully exported back into the CDS. Fig. 4 shows identical instructions for the binary pump, and Fig. 5 shows those for a quad pump. All 55 acquisition parameters in the data model were successfully transferred into and subsequently exported from the Methods Database. The objective of this pilot study was not to capture analytical data, but rather to demonstrate the feasibility of digitally transferring methods (HPLC acquisition parameters) between various vendor software controlling systems and hardware using the ADF data model. This objective was successfully achieved. However, scientists also expressed interest in viewing chromatographic results that could be correlated with the methods. Consequently, we expanded the scope to incorporate the visualization of chromatographic data. The ability to visualize the chromatographic data with ease will minimize downtime and establish faster if methods meet the suitability criteria, as methods can often be successfully analysed on different hardware, with some conditions being impacted by the difference in instrumentation. Below outlines that using the Methods Hub cloud platform, the chromatographic results could be compared with ease to identify any differences.

3.1. Overlay of MSD and GSK data (OpenLab™ CDS Data)

The 5 min linear gradient instrument method was drafted on Agilent 1290 binary UPLC at MSD connected to OpenLab™ CDS and the data was collected in the same stand-alone CDS. Thereafter, the instrument method and collected data were converted and transferred to ZONTAL in

Table 2
Mobile phase binary gradient used for the analysis of the GSK test mix standard solution.

Time (min)	Mobile phase A (%)	Mobile phase B (%)
0.00	100	0
8.00	5	95
8.01	100	0
9.50	100	0

Fig. 4. Method viewed in OpenLab™ CDS versus ZONTAL showing identical data.

Fig. 5. Method viewed in Empower™ CDS versus ZONTAL showing identical data (quad pump to binary pump system).

the ADF format. The method instruction file was downloaded by the scientist at GSK and an instrument method specific to the attached Agilent 1290 binary UPLC instrument was created electronically using the procedure outlined above. The scientist at GSK executed the method using the instrument method and thereafter uploaded the method as well as collected data onto ZONTAL after conversion into ADF format (all files are available as supplementary information).

The overlaid chromatograms collected on the same Agilent 1290

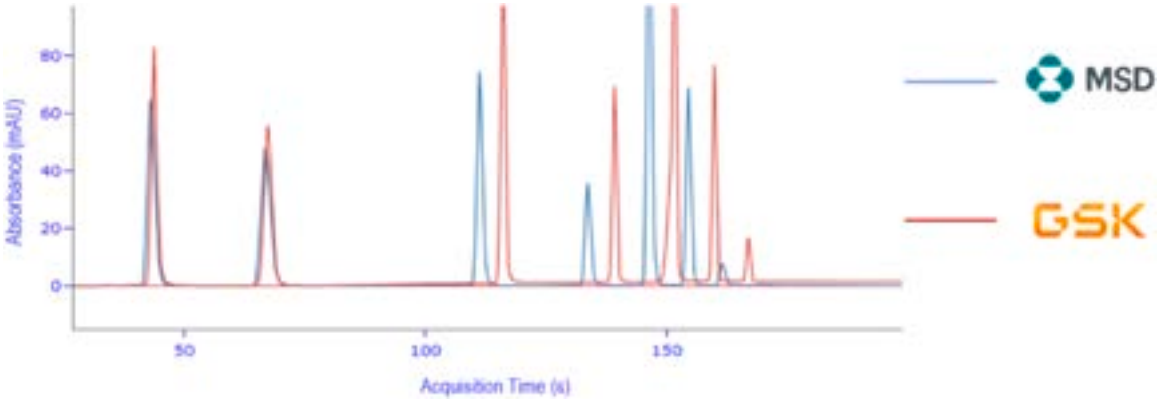

Fig. 6. Import (GSK) and export (MSD) of method parameters between two pharmaceutical companies with results acquired on different Agilent 1290 binary UPLC instruments connected to OpenLab™ CDS. The method instructions were sent via ZONTAL.

binary hardware and Agilent software, shown in Fig. 6, and the comparison metrics around chromatography parameters (retention time, peak tailing and resolution) in Table 3 confirm the successful transfer of the instrument method from one pharmaceutical company to another. The availability of chromatograms and metadata around method parameters and instrument configurations at one place allowed the scientist to conduct instantaneous data analytics. The retention times of the peaks on the GSK instrument were delayed by an average of ~ 0.2 min, which could be attributed to differences in instrument dwell volume. The retention time differences between the data collected on two instruments would have passed the classical system suitability criteria established for every HPLC method [11]. However, having the capability to do this type of data analytics in a quick turnaround time enables the identification of root cause analysis, which would be of significant use in establishing method controls through method lifecycle management and thereby reducing the Out-Of-Specification (OOS) results or support continuous manufacturing operations in the pharmaceutical industry.

The GSK's non-GMP 8-minute generic HPLC method was electronically transferred onto Empower™ CDS and Agilent 1290 instrument using ZONTAL. The method was exported from OpenLab™ CDS and stored on ZONTAL, and electronically sent to the Empower™ CDS through Orbis Gateway. Additionally, the GSK's non-GMP 8-minute was written on a 1260 instrument and transferred onto the 1290 instrument using Empower CDS through Orbis Gateway, Fig. 7.

For a digital transfer of method parameters, a data standard was created applying the Allotrope data modeling patterns. Method parameters can be represented by information objects that carry a literal value or quantity values, including a unit. For example, a parameter of injection volume can specify the volume of a liquid sample that should be injected during an HPLC run and has a numeric value and a unit of microliter. The Allotrope data modeling patterns provide a standardized approach to contextualizing such information. Information objects of parameters are attached to a more detailed graph structure that enriches the individual parameters with additional information. The graph provides accurate context to the method parameters, allowing applications and systems to process these specifications correctly and enable us to precisely distinguish, for example, between specified and actual values. The information of a method is represented by a semantic graph structure that enables processing of information based on its context. For the scientists to work with their analytical methods, they need a well-structured, human-readable version. This representation needs to provide all the method parameters in a scientifically appropriate context and make it easy to compare methods and follow the lifecycle of methods through the development process. Exported method instructions can be viewed in ZONTAL (Fig. 4, Fig. 5). Successful export and import of method instructions were achieved, as evidenced by the overlay of the Empower™ CDS and OpenLab™ CDS data, which was acquired through ZONTAL. Instrument instructions were successfully exported in a machine-readable format from one pharmaceutical company to another without any need for human transcription or manual keying. While downloading the method instructions to a new instrument, any deviations from the minimum required configurations for e.g., pressure limits, pump settings, and solvent compressibility, were prompted as messages for the user to accept or reject. The decision of the user was electronically captured in the audit trail for QC purposes. This feature enabled the electronic transfer of methods between different hardware configurations, such as binary to quaternary or software configurations, resulting from firmware updates on a specific vendor instrument. The results in Figs. 6 and 7 show good repeatability and reproducibility across platforms with different hardware and software configurations. A slight change in peak height was observed, which could be due to different standards and mobile phase batches used for the experiments. Also, the difference in retention time of the peaks, which is well within the acceptable system suitability criteria (Table 3), could be assigned to different column lots or a difference in dwell volume of both instruments.

The integration with ELN or LIMS systems is possible to allow for closed-loop experimentation, where manual data transfers are completely eliminated from the process. A sample list is created in the ELN and submitted to ZONTAL for execution. ZONTAL then forwards the sample list and the selected instrument method to the CDS adapter. Once results are returned by the CDS adapter, ZONTAL ingests the results into the ELN system to automate the experiment write-up, Fig. 8.

Table 3
Comparison of chromatographic parameters between runs at different companies.

Suitability Parameter	Result	Acceptance Criteria
Retention Time Difference[a]	1.35 %	≤ 2 %
Peak Tailing[b]	1.2 (MSD); 1.0 (GSK)	< 1.5
Resolution[c]	3.7 (MSD); 4.6 (GSK)	> 1.5

[a] Average difference in retention time for all 7 peaks
[b] Average of Peak Tailing value for all Peaks
[c] Resolution between peak 6 and 7

4. Conclusions

Pre-competitively, we performed HPLC-UV methods data modeling and applied it to build a specific example solution of a vendor-neutral Methods Hub to seamless exchange methods as digital instructions with different control software/instruments. HPLC method parameters are specific to both the CDS and the HPLC Instrument vendor and model, and have not been easily accessible in the past. This will be the first time method parameters and raw data are available in an interoperable data format across a variety of software and hardware systems commonly used in the industry: In this study, 55 key HPLC instrument instructions were successfully transferred in a digital format between two pharmaceutical company sites and across different control systems (CDSs) via a cloud-based application. Methods are often executed on similar chromatographic hardware which are not necessarily controlled by similar software. Using the common data format, it was possible to convert the instrument instructions across vendor systems. Method parameters in a digital executable format reduce the cost of method lifecycle management between functions. All Instrument control systems will be able to read and write from the ADF Methods file and allow seamless exchange between instrument hardware, depending on the preferences and needs of the individual laboratory. It also enables more accessible access to methods and data for QA/QC and regulatory purposes, as well as further data investigations, data re-use, and system changes made easier with data from different sources in one consistent data structure. To make the Methods Database an integral part of the enterprise infrastructure of every analytical Lab, flexible adapters to every ELN system will allow seamless exchange of method and result data. The use of standardized ADF format across vendors should also simplify the design of these interfaces since the method and result data structure and context will be consistent. Additionally, with the use of semantic standards to represent data, advanced on-demand data queries are possible across commercial software products, helping to move us all one step closer to realizing a FAIR data ecosystem to drive insights.

Standards and interoperability have become more important in recent years. Pre-competitive collaboration can result in the development of industry standards and interoperable solutions. Standardization facilitates compatibility and integration between different technologies, systems, and platforms, promoting seamless data exchange and workflow interoperability across the industry. This, in turn, fosters innovation by reducing barriers to adoption. A key strategic priority of the Pistoia Alliance is to Deliver Data Driven Value. With this pilot, this Pistoia Alliance project has demonstrated the benefits of precompetitive, cross-company collaborations, where joint funding led to the development of a semantically-enabled data model to allow for digital,

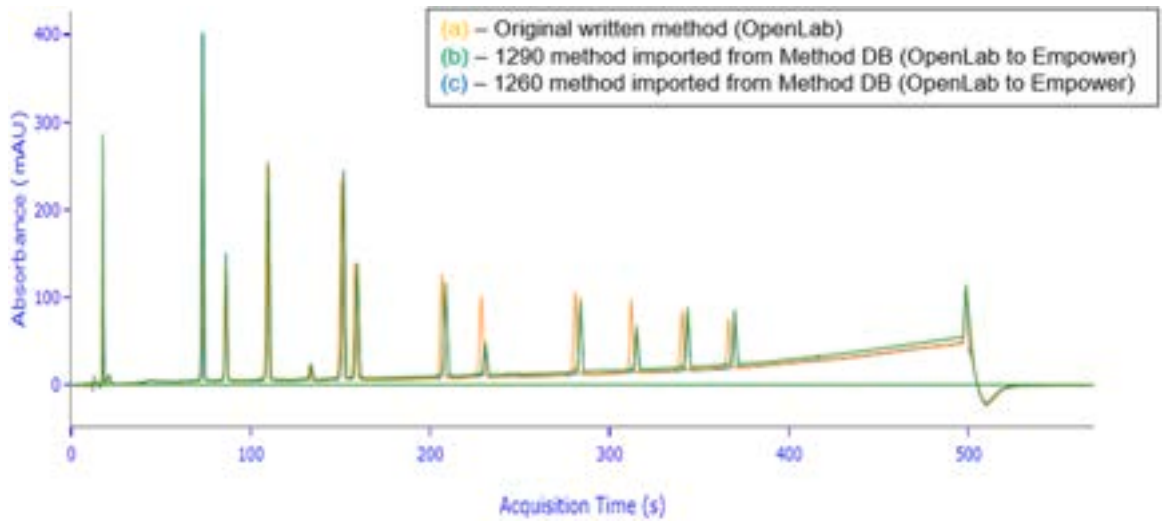

Fig. 7. HPLC chromatogram overlay of non-GMP 8-minute generic method (1290 Instrument): (a) original written method on 1290 using OpenLab™ CDS, (b) 1290 method imported from ZONTAL on Empower CDS and (c) 1260 method imported from ZONTAL on Empower™ CDS.

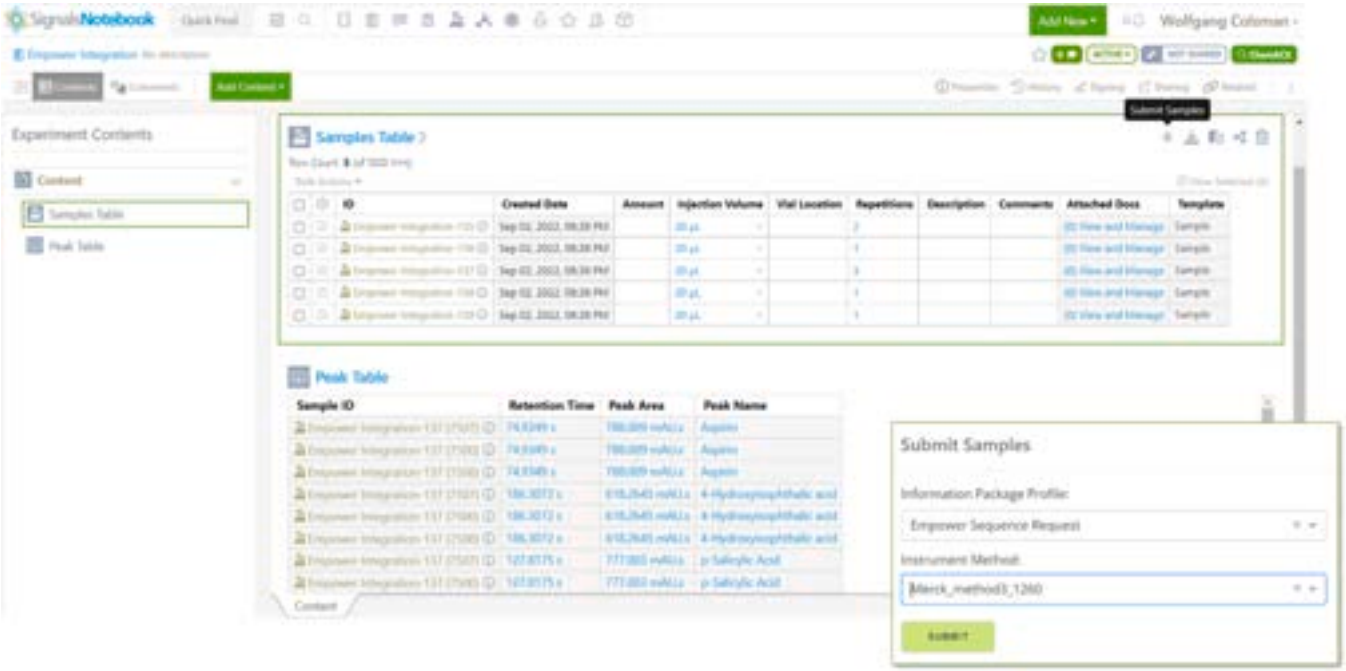

Fig. 8. Integration with ELN or LIMS systems allows for closed-loop experimentation.

interoperable exchange of information, with the goal of increasing method reproducibility, and support a fully FAIR ecosystem where methods and results data are truly findable, accessible, interoperable, and reusable by design without manual intervention.

CRediT authorship contribution statement

Noelken Gerhard: Writing – original draft, Project administration, Methodology, Conceptualization. **McComas Juliet:** Writing – review & editing, Methodology. **Nielsen Birthe Vejby:** Writing – review & editing, Writing – original draft, Project administration, Conceptualization. **Dabo Azzedine:** Writing – original draft, Investigation, Formal analysis, Data curation. **Aggarwal Pankaj:** Writing – original draft, Investigation, Formal analysis, Data curation. **Antonucci Vincent:** Writing – review & editing, Writing – original draft, Methodology, Funding acquisition, Conceptualization. **Sun Cheng:** Data curation. **Fessenmayr Heiko:** Writing – original draft, Software, Methodology, Conceptualization. **Colsman Wolfgang:** Writing – original draft, Visualization, Software, Methodology. **Wells Kenneth M.:** Funding acquisition, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors wish to acknowledge the Methods Database project team and the Steering Committee; Pistoia Alliance; the team at Orbis Lab systems (part of Astrix), Aidan Finnerty, Marco Pellegrini, Michael

Gannon, Luca Teodori; The Allotrope Foundation; and OSTHUS.

Appendix 1. AFO terms in the methods database HPLC-UV model

#	Parameter	#	Parameter
1	sequence table / repetition	29	instrument method / detector method / data acquisition time
2	sequence table / injection volume	30	instrument method / detector method / detector / channel / bandwidth
3	sequence table / sample type	31	instrument method / detector method / detector / channel / index
4	sequence / sample name	32	instrument method / detector method / detector / channel / wavelength
5	sequence table / sample ID	33	instrument method / detector method / detector / detector type
6	sequence table / vial position	34	instrument method / detector method / flow cell pathlength
7	sequence table / instrument method	35	instrument method / detector method / wavelength scan range
8	sequence table / action type	36	instrument method / detector method / wavelength scan step
9	sequence table / processing method	37	instrument method / autosampler method / autosampler temperature
10	sequence table / reporting method	38	instrument method / autosampler method / injection volume
11	sequence table / base instrument method name	39	instrument method / autosampler method / washing solvent
12	sequence table / base processing method name	40	instrument method / column method / column / chemistry
13	sequence table / CDS project name	41	instrument method / column method / column / description
14	sequence table / LIMS ID	42	instrument method / column method / column / inner diameter
15	sequence / project name	43	instrument method / column method / column / length
16	sequence table / run time	44	instrument method / column method / column / particle size
17	instrument method / pump method / pump / maximum flow gradient	45	instrument method / column method / column / temperature
18	instrument method / pump method / pump / pressure limit	46	instrument method / column method / column / vendor
19	instrument method / pump method / pump / pump type	47	drug product / dosage description
20	instrument method / pump method / pump / run time	48	material / identifier
21	instrument method / pump method / solvent name	49	material / role
22	instrument method / pump method / time events / flow	50	method / analyte
23	instrument method / pump method / time events / gradient curve	51	method / effective date
24	instrument method / pump method / time events / index	52	method / lifecycle status
25	instrument method / pump method / time events / solvent%	53	method / sample matrix
26	instrument method / pump method / time events / time	54	product name
27	instrument method / instrument type	55	project identifier
28	instrument method / detector method / acquisition rate		

Appendix B. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jpba.2025.116907.

References

[1] Grand View Research, Drug discovery outsourcing market size, share & trends analysis report by drug type (small, large molecules), therapeutic area (respiratory system, immunomodulation), workflow, and segment forecasts, 2022–2030, Grand View Research, April 2022 (Report ID: 6277979).

[2] Transparency Market Research, Drug discovery outsourcing market (service type: biological services, chemical services, lead identification & screening, lead optimization, and others; drug type: small molecules and large molecules; therapeutic area: respiratory system, anti-infective, oncology, cardiovascular, gastrointestinal, central nervous system, immunomodulation, and others)—global industry analysis, size, share, growth, trends, and forecast, 2020–2030, Transparency Market Research, June 2020 (Report No.: TMRGL1841).

[3] L.P. Freedman, I.M. Cockburn, T.S. Simcoe, The economics of reproducibility in preclinical research, PLoS Biol. 13 (2015) e1002165, https://doi.org/10.1371/journal.pbio.1002165.

[4] M.D. Wilkinson, M. Dumontier, I.J. Aalbersberg, et al., The FAIR guiding principles for scientific data management and stewardship, Sci. Data 3 (2016) 160018, https://doi.org/10.1038/sdata.2016.18.

[5] J. Bai, L. Cao, S. Mosbach, et al., From platform to knowledge graph: evolution of laboratory automation, JACS Au 2 (2022) 292–309, https://doi.org/10.1021/jacsau.1c00438.

[6] G. Schneider, Automating drug discovery, Nat. Rev. Drug Discov. 17 (2018) 97–113, https://doi.org/10.1038/nrd.2017.232.

[7] Pistoia Alliance, Available from: https://www.pistoiaalliance.org/ (accessed 15 September 2023).

[8] P.N. Robinson, M.A. Haendel, Ontologies, knowledge representation and machine learning for translational research: recent contributions, Yearb. Med. Inf. 29 (2020) 159–162, https://doi.org/10.1055/s-0040-1701991.

[9] Allotrope Foundation, Available from: ⟨https://www.allotrope.org⟩ (accessed 15 September 2023).

[10] ZONTAL, Available from: ⟨https://www.zontal.io⟩ (accessed 15 September 2023).

[11] International Conference on Harmonization (ICH), Validation of analytical procedures: text and methodology Q2(R1), November 2005. Available from: ⟨https://www.ich.org⟩ (accessed 15 September 2023).

[12] Agilent Technologies, OpenLAB CDS requirements guide. Available from: ⟨https://www.agilent.com/cs/library/usermanuals/public/OpenLABCDSRequirements.pdf⟩ (accessed 15 September 2023).