



# Allotrope Foundation Quarterly Update 2024/06

Dear Allotrope Community,

We have continued our progress this quarter and improved or expanded the AFO and ASM with updates to share. Please note that access to links may require access to GitLab or other Allotrope Community resources. More details for access [here](#).

## Table of Content

Welcoming New Community Members .....	2
TU-Dortmund University, a New APN Member .....	2
Allotrope Foundation Ontology & Data Models (AFO/ASM) .....	2
AFO Updates .....	2
AFO Term Dictionary .....	3
Allotrope Foundation Simple Models (ASM) .....	3
ASM Directory .....	3
Modularization .....	3
ASM Updates .....	3
ASM Training Materials and Working with the ASM .....	4
ASM and ADM Modeling and Support .....	4
Tooling, Testing, QA and Automation Pipeline .....	4
PURL (AFO server) Improvements and Stability .....	5
Further enhancements to the ASM automated QA using the CI pipelines .....	5
General Maintenance .....	5
Working Group Updates .....	5
Modeling: (Notes: here) .....	5
Chromatography: (Notes: here) .....	6
Plate Reader: (Notes: here) .....	6
Mass Spectrometry: (Notes: here) .....	6
Flow Cytometry: (Notes: here) .....	7
Allotrope in the News .....	7
Projects within the Allotrope Community .....	8
Sample of Projects with the Allotrope Framework .....	8
Instrument Data Converters to ASM (Open Source) .....	9
Allotrope Publications .....	9
AF Community and Events .....	10
2024 Spring Allotrope Connect Workshop .....	10
Allotrope Data Framework Onboarding Guide .....	11
Allotrope YouTube Channel .....	11
Looking Forward .....	11

---

## *Welcoming New Community Members*

---

### *TU-Dortmund University, a New APN Member*

We would like to welcome a new member to the APN: [TU-Dortmund University, Faculty of Biochemical and Chemical Engineering](#)

“The Faculty of Biochemical and Chemical Engineering has existed at TU Dortmund University for over 50 years and is one of the largest and most successful in Germany in the field of process engineering. The faculty is committed to the development and improvement of safe, resource- and environmentally compatible processes and products for the chemical and biotechnological industries.”

Faculty members published a few research articles where the AFO was exercised. Initial article is available at [Link](#). Faculty members are also behind the NFDI4Cat initiative to secure the digital future of catalysis: Digital catalysis, the German National Research Data Infrastructure (NFDI) <https://nfdi4cat.org/en/about-us/>

---

## *Allotrope Foundation Ontology & Data Models (AFO/ASM)*

---

Modeling teams have continued working to align on proposals to expand the domain coverage of the AFO and ASM. Easily access files located on Client Connect [here](#) and more granular technical details available on GitLab, <https://gitlab.com/allotrope>.

See [www.allotrope.org/product-releases](http://www.allotrope.org/product-releases) for a full and updated list of available models.

### *AFO Updates*

Following the new and updated set of ASM released this quarter, a new AFO release is published. Please note that QUDT 1.0 is no longer merged into the Allotrope Merged Ontology Suite.

The Allotrope Merged Ontology Suite release is available on:

- BioPortal, the repository of biomedical ontologies published by the National Center for Biomedical Ontology at Stanford University: <https://bioportal.bioontology.org/ontologies/AFO>
- OLS4, the Ontology Lookup Service repository for biomedical ontologies published by the European Bioinformatics Institute: <https://www.ebi.ac.uk/ols4/ontologies/af0>
- Ontobee, Ontologies data server published by the University of Michigan Medical School: <https://ontobee.org/ontology/AFO> (Ontobee generates the AFO list of terms in an Excel spreadsheet as well as Tab Separated Values file)
- Client Connect: [here](#)
- Gitlab: <https://gitlab.com/allotrope/af0/-/tree/master/af0>
- Allotrope PURL sever: <http://purl.allotrope.org/> (listed under AFO>MERGED)
- JFrog Artifactory: <https://allotrope.jfrog.io/ui/repos/tree/General/AFO-release-public>



- Allotrope website: <https://www.allotrope.org/ontologies>

### *AFO Term Dictionary*

Allotrope Term Dictionary is available in both .xlsx (Excel) and .csv (Comma Separated Values) format and can be downloaded from the

- Allotrope website at: <https://www.allotrope.org/ontologies>
- Client Connect: [here](#)
- JFrog Artifactory: <https://allotrope.allotrope.jfrog.io/ui/repos/tree/General/AFO-dictionary-release-public>

---

## *Allotrope Foundation Simple Models (ASM)*

---

### *ASM Directory*

ASM Directory for the applicable sample JSON and JSON Schema files per technique is available for convenient viewing of file content using a browser. The directory contains links to the latest sample files and embedded (i.e., standalone) schema for all ASMs in REC (Recommended) status. Please refer to the directory on the public repository: <https://gitlab.com/allotrope-public/asm/-/blob/main/README.md#allotrope-simple-model-directory>

### *Modularization*

JSON Schemas allow for modularization and factoring out commonly used rules by utilizing references to other JSON schema files. The simple model schemas make use of this modular approach. The ASM Schema is defined using:

- Technique specific schema: a JSON Schema that contains the domain specific rules. It references the core declarations instead of each technique defining its own.
- Core schema: a JSON Schema that contains reusable, domain independent rules. The core schema defines value types for all possible values that may be used in tabular models.
- Other reusable schemas: Cube, Hierarchy, Manifest, Units, other future extensions

Having the basic rules factored out in a core and other schemas allows for later extensions without changing each technique specific schema. It ensures consistent writing and querying regardless of whether it's a single contained instrument or a modular stack with multiple detectors, pumps, or anything else. Motivation of the modular pattern is to drive consistent data structures across techniques, enabling data from different models to work seamlessly together.

### *ASM Updates*

New and updated Allotrope Simple Models release this quarter are available on:

- Client Connect: [here](#)
- GitLab: <https://gitlab.com/allotrope/adm/-/tree/master/>
- JFrog Artifactory: <https://allotrope.jfrog.io/ui/repos/tree/General/ADM-release-public>

Here is the list of the new and updated set of ASMs released this quarter.\*

ASM Model	Type	Maturity	Path
NMR	Tabular	REC	New
Electrophoresis	Tabular	REC	New
pH	Tabular	REC	Update
Balance	Tabular	REC	Update
Osmolality	Tabular	REC	Update
Automated reactors	Tabular	CR	Update
Plate readers kinetics (Absorbance, Fluorescence, Luminescence)	Tabular	REC	Update
Chromatography column	Tabular	CR	New
Core updates	Tabular	REC	Update

\* To find out how to access the related model's artifacts on GitLab:

<https://gitlab.com/allotrope/adm/-/wikis/Summary-Table-of-the-Governed-ADM-and-ASM-Techniques-Artifacts>

### *ASM Training Materials and Working with the ASM*

ASM training material is available on Allotrope public repository at the following locations:

- Brief introduction to ASM: <https://www.allotrope.org/allotrope-simple-model>
- ASM Primer: <https://gitlab.com/allotrope-public/asm-primer/-/wikis/home>
- ASM Jupyter Notebook Demo: <https://gitlab.com/allotrope-public/asm-jupyter-demo> It is a step-by-step example file for working with ASM files in a Jupyter Notebook. It was also tested with Google Colab.

### *ASM and ADM Modeling and Support*

ASM related support tickets can be opened at the ADM project (<https://gitlab.com/allotrope/adm/-/issues>).

The Product Team can generate ADM specific artifacts (SHACL and its deployment using ADF) by request.

The latest updated set of ASM models is available on Gitlab. New and updated models will be released in conjunction with the release of new tabular models. Adopters can generate example results of tabularized data based on the JSON ASM format.

In cases where there is no tabular model for a chosen instrumentation type or technique, the product team is available to support the drafting of a new tabular model and the Modeling Working Group is ready to review and govern drafted models.

---

***Tooling, Testing, QA and Automation Pipeline***

---



### *PURL (AFO server) Improvements and Stability*

The DevOps team completed the tasks to improve PURL (AFO server) stability:

- Resolve Fuseki (SPARQL server) database disk space consumption with every update.
- Update Fuseki (SPARQL server) to latest 5.0.0 version.
- Create server alerts using AWS Cloud Watch

### *Secured Connection with PURL (AFO server)*

Secured HTTP was enabled on PURL (AFO server): <https://purl.allotrope.org/>

### *Further enhancements to the ASM automated QA using the CI pipelines*

Further enhancements are anticipated during the coming quarter on the automated ASM QA tools using the CI pipelines:

- Add format checking to ASM. <https://gitlab.com/allotrope-open-source/allotrope-devops/-/issues/232>. It validates formats in ASM JSON files using Python JSON schema validation tool <https://python-jsonschema.readthedocs.io/en/latest/validate/#validating-formats>
- Check ASM/ADM manifests import existing AFO files: New checks to the manifest validator, JSON scripts, to ensure the exitance of the referenced .ttl files from the AFO. <https://gitlab.com/allotrope-open-source/allotrope-devops/-/issues/238>

### *General Maintenance*

- Maintenance work on the GitLab CI pipelines was done to resolve JFrog changes on how Java jar files are fetched.

We would like to thank Karin Colman from the PharmaLex engineering team for her dedication and commitment to improve the overall tooling, testing, QA and automation pipeline.

---

## **Working Group Updates**

---

Please note that the working groups meetings are recorded to improve access and transparency for those unable to attend or for the folks that are just interested in what's going on. To sign up for any working group, go to: [www.allotrope.org/working-groups](http://www.allotrope.org/working-groups)

### *Modeling: (Notes: [here](#))*

ASM modularization work continues across the different working groups. The Common Hierarchy Schemas is a collection of “Lego” like, reusable building blocks to create consistent hierarchical structures across the different models.

The Modeling WG was working on the following new models:

- NMR (REC)
- Electrophoresis (REC)



To support the development of the NMR model, the team was joined by Dimitris Argyropoulos, NMR Business Manager, ACD/Labs, Oxford, England

The following new modular documents/schemas were developed:

- Error document
- Data Systems Document
- Custom Information document
- Electronic Signature Document
- Electronic Project Record
- Analysis Sequence Document
- Data Region Document

*Chromatography: (Notes: [here](#))*

The Chromatography Working Group was working on the development of a new Chromatography Column model that was released this quarter. The work extends the current Chromatography Column model in the general LC model

The following new modular documents/schemas were developed:

- Substrate document
- Stationary Phase document
- Capability Document

The working group was joined by Dr. Dwight Stoll from the Department of Chemistry at Gustavus Adolphus College. Dr. Dwight maintains the Column Selectivity Database

<https://www.hplccolumns.org/about/index.php>

*Plate Reader: (Notes: [here](#))*

As of this year, the Plate Reader WG meets bi-weekly from 9am to 10am EST on Monday.

The team was working on Kinetic Data Cube (sample data is available on Client Connect [here](#)) for Kinetic reads (the reading or measurement of kinetic data in a plate reader). In plate readers, a kinetics read involves taking measurements at multiple time points to observe changes in a sample over time.

A new model update of Kinetics (Absorbance, Fluorescence, and Luminescence) was released this quarter.

On the horizon is further work on the model to include:

- Optical imaging
- Spectrum scan
- Area scans (use in immunology/biology)

*Mass Spectrometry: (Notes: [here](#))*

We would like to thank Graham McGibbon (ACD/Labs) for stepping forward and continuing to lead the MS working group.

- Quantitation MS tabular model in progress. It is the quantitation of compounds (with generic reusable structure where possible) using MS.
- Genentech shared reported MS data
- During this quarter, Lablicate demonstrated the ongoing work on the ADF/ASM plug-ins for OpenChrom with MS data.

### *Flow Cytometry: (Notes: [here](#))*

The Flow Cytometry Working Group was established this year following the 2023 Fall Connect Workshop technical meetings. The WG meets bi-weekly from 11am to 12pm EST on Friday, and it is chaired by Joe Negri from Benchling. Community members are welcome to join.

The working group is discussing the scope of the initial model and reuse of existing standards. Existing ones are Flow Cytometry specific, but they are not cross cutting techniques like ASM. A skeleton for an ASM Flow model was suggested and presented by Joe Negri from Benchling. Some of the members will provide sample Flow Cytometry data. The team created a set of open questions to be addressed such as:

- Acquiring data across multiple channels
- Referencing terms to an .FCS file and AFO mapping
- Describing gates and populations
- Other Flow Cytometry data standards or other ontologies standards to be considered

We would like to thank Joe Negri (Benchling) for his dedication, commitment, and leadership of the Flow Cytometry working group.

---

### *Allotrope in the News*

---

For the latest list of “Allotrope in the News”, please visit our website at:

<https://www.allotrope.org/allotrope-in-the-news>

Here is the listed recent news:

- True Instrument Integration Requires More Standardization: [Link](#)
- Unlocking the automation potential via open standards: [Link](#)
- Natural Language Processing-Based Extension of the Allotrope Foundation Ontology: [Link](#)
- Lessons from Shape Therapeutics: How to harness AI for R&D success: [Link](#)
- Pharmaceutical CMC ISA-88 Process Ontology: [Link-1](#) [Link-2](#)
- Pharma 4.0 Market: Global Industry Assessment & Forecast: [Link](#)
- Resetting R&D for AI: Biotech’s Path to AI-Ready Data & AI-Enabled Scientists: [Link](#)
- AFO and IDMP-O; showcasing the federation of ontologies: [Link](#)
- Using Chat GPT for automatic asset onboarding for Industrial IoT projects: [Link](#)
- Combining standardized lab instrument communication with standardized data formats is now a reality: [Link](#)

- Digital Evolution: Novo Nordisk's Shift to Ontology-Based Data Management: [Link](#)
- Bringing OPC-UA and Allotrope ASM together to enable real automation in the analytical lab: [Link](#)
- Unleashing lab connectivity presentation at the 2024 Spring Allotrope Connect was live at LADShack: [Link](#)
- Lablicate shared their latest advancements using Allotrope Foundation data models: [Link](#)
- ZONTAL's Summary of the 2024 Spring Allotrope Connect: [Link-1](#), [Link-2](#)
- Growing value of data standardization: Allotrope Foundation Connect Workshop Proceedings: [Link](#)
- How to jumpstart innovation in R&D? Biotech must solve the instrument connectivity challenge: [Link](#)
- An Introduction to Scientific Data Standards: [Link](#)
- Exciting news for the Allotrope community! The 2024/Q1 release is here, featuring new Spectrophotometry and Multi Analyte Profiling models, and much more!: [Link](#)
- Join the Allotrope Foundation on an Exciting Journey to Standardize Flow Cytometry Data Analysis!: [Link](#)

---

### *Projects within the Allotrope Community*

---

#### *Sample of Projects with the Allotrope Framework*

- Shimadzu Supports the Allotrope Data Format (ADF) with LabSolutions:
  - LabSolutions is Shimadzu's software platform designed to manage data from Shimadzu's various analytical instruments, including chromatography, mass spectrometry, spectroscopy, and other laboratory equipment.
  - "LabSolutions can be output in various formats, and it also supports output to the data format specified by the Allotrope Foundation"
  - More information is available at <https://www.an.shimadzu.co.jp/products/software-informatics/software-option/system-linkage/features.html>. (Please use the translation function of your browser)
- [Lablicate](#) is working on Allotrope ASM/ADF plugin for OpenChrom. The project was presented and demonstrated during the recent 2024 Spring Connect.
- Merck is populating a Knowledge Graph with ASM instance data. The project was presented during the recent 2024 Spring Connect. Merck is using the \$asm.pattern Keyword within the ASM Schema to populate the Knowledge Graph and NOT using any ADM/SHACL/RDF equivalent. The project presentation is available [here](#), view Python notebook [here](#), download Python notebook [here](#).
- Agilent demonstrated at the recent 2024 Spring Connect a proof of concept using the ASM as the Information Modelling layer in OPC-UA. The project presentation is available on the Allotrope YouTube channel [here](#)



### *Instrument Data Converters to ASM (Open Source)*

- **Benchling** is growing its platform for lab instrument connectivity and data management, Benchling Connect. With Connect, Benchling confronts industry-wide challenges with proliferation of proprietary instrument data models and vendor lock-in by mapping all instrument output to the Allotrope Simple Model (ASM) and making the converter codes open source and freely available on GitHub.
  - Benchling's blog: <https://www.benchling.com/blog/benchling-connect>
  - "allotropy", open source, Python library repository on GitHub: <https://github.com/Benchling-Open-Source/allotropy>
  - "allotropy", Python package for converting instrument data into ASM is on PyPI: <https://pypi.org/project/allotropy/>
  - Published the Benchling Connect - Adapter specific guides: [here](#)
- **IFP Energies Nouvelles (IFPEN)**, a French public research, innovation and training organization in the fields of energy, transport and the environment is developing a set of open-source converters to several ASM models and several instruments. The project is managed and developed by Maxime Visconte, Industrial and lab IT manager at IFPEN.
  - More information can be found on IFPEN's GitHub repository: <https://github.com/ifpen>
  - A set of utilities, shared between all the ASM converters were pushed to the Maven central repository: <https://central.sonatype.com/artifact/fr.ifpen.allotropeconverters/ASMUtils/1.0>
  - A presentation of this project is available on our YouTube channel: [here](#)

---

### *Allotrope Publications*

---

We have published 2-page summaries and updated the introductory presentation:

- **Allotrope Introductory Slide Deck**: can be downloaded from [here](#)
- **Allotrope Models & Domains**: can be downloaded from [here](#)
- **Allotrope Data Strategies**: can be downloaded from [here](#)

---

## *AF Community and Events*

---

### *2024 Spring Allotrope Connect Workshop*

The 2024 Spring Allotrope Connect Workshop took place at BASF Digital Solutions in Ludwigshafen, Germany. The workshop brought together over 110 in-person and online participants, and was very informative and focused on expanding opportunities for Allotrope in the era of Generative AI including:

- Generative AI Play in Lab Connectivity, Data Integration, and Automation
- Ontology Federation, Implementation, Extensions with NLP, and Role for AI&ML
- The Role of Data Standards in the Era of Generative AI

Here are some highlights from the discussions and presentations:

- Generative AI can hold substantial potential in streamlining lab integration projects through standardized interfaces like OPC-UA and data models such as the Allotrope Simple Model (ASM).
- Agilent demonstrated the future of an interoperable digital lab by integrating Liquid Chromatography ASM with OPC-UA, showcasing improved workflows and data accessibility.
- Microsoft highlighted automated asset onboarding in Industrial IoT projects using Generative AI with OPC-UA standardized interfaces, featuring OpenAI's capability to generate OPC-UA integration artifacts.
- Successful AI and ML applications in labs depend on standardized interfaces and data formats, as emphasized by BASF, which discussed integrating instruments with Laboratory Information Management Systems (LIMS) to reduce errors and data gaps.
- BASF also advocated for open-source data standards, ensuring high-quality, flexible data integration that is vendor-agnostic.
- Zontal presented a platform leveraging Generative AI for automated report generation and data connectivity, enhancing digital lab transformation.
- Andrew Chen demonstrated the use of Large Language Models (LLMs) to automate converting proprietary instrument formats to ASM, reducing development time and resources while improving testing through extensive test case generation.
- MSD showcased the construction of knowledge graphs from ASM files, enhancing advanced analytics and Generative AI capabilities.
- Benchling and IFP Energies Nouvelles (IFPEN) emphasized the importance of open-source standardized data formats in accelerating AI and ML applications.
- J&J illustrated how ontologies are the secret sauce to structured data that could enable "Digital Pharma 4.0" by making data FAIR (Findable, Accessible, Interoperable, and Reusable) .
- TU Dortmund highlighted NLP-based tools for dynamic ontology extension.
- Accurids and Amgen discussed ontology federation for semantic interoperability and compliance across the pharma value chain.
- Overall, data standards are crucial for advancing Generative AI technologies, ensuring data consistency and compatibility, and facilitating collaborative efforts.



- The AF Connect breakout discussions underscored the importance of data standards in AI/ML initiatives and the potential of Generative AI to accelerate their development and implementation by analyzing historical data, structuring unstructured data, and promoting data standards adoption.

Conclusion: The 2024 Spring Allotrope Connect Workshop emphasized the promising role of Generative AI and data standards in advancing lab connectivity, data integration, and automation. The adoption of ontologies, standardized interfaces, data models, and knowledge graphs was emphasized as key to unlocking the potential of AI/ML applications across various domains.

For more detailed information please refer to:

- 2024 Spring Allotrope Connect Workshop event page [here](#)
- Detailed program and list of speakers [here](#)
- Detailed presentations abstracts and speaker's bio can be downloaded [here](#)
- Playlist of Day-2 presentations recording on the Allotrope YouTube Channel [here](#)

### *Allotrope Data Framework Onboarding Guide*

The Allotrope Onboarding Guide wiki page was updated. Please refer to the following link: [Allotrope Data Framework Onboarding Guide](#)

### *Allotrope YouTube Channel*

Our YouTube channel has new a handle: <https://www.youtube.com/@allotropefoundation>. The Allotrope YouTube Channel hosts a technical playlist as well as the Allotrope Connect public presentations from 2017 and 2020 to the latest 2024 Spring Connect event.

The YouTube Channel videos are organized by playlists at: <https://www.youtube.com/@allotropefoundation/playlists>.

---

## *Looking Forward*

---

The Allotrope Product Team is looking forward to another productive 3<sup>rd</sup> quarter of 2024. We are looking to develop additional improvements to meet the evolving needs of our community.

Please contact us for any questions at [product\\_team@allotrope.org](mailto:product_team@allotrope.org).

Sincerely,

Allotrope Product Team