Allotrope Foundation Release Notes 2021/12

Dear Allotrope Community,

We have continued our progress this quarter and improved or expanded the AFO, ADM, ASM, and ADF further in the following areas 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

Applied Graphs

We would like to welcome Applied Graph www.appliedgraphs.com. Here is Applied Graphs introduction message to the community: "Applied Graphs helps innovative customers build and deploy knowledge graphs. With a particular focus on life sciences and health care, we bring deep expertise in the relevant technology, vocabularies, and software development practices that are core to deploying production-ready semantic graphs. We believe the semantic data paradigm has the power to reshape how data is created, consumed, shared, and acted upon. It is our mission to guide and support you in this work.

Applied Graphs is grateful to be a member of the Allotrope Partner Network. We believe that the Foundation is designing and building a set of information structures and tools that will enable both rapid growth of knowledge graphs and a massively improved data exchange paradigm. Our focus is simply on semantic graphs, and over the course of many years, we have never seen a project so captivating, rigorous, and ready for use. Thanks to everyone who is making this happen. It is a unique achievement with a great future."

University of Cambridge

We would like to welcome the University of Cambridge – School of technology, the department of Chemical Engineering and Biotechnology https://www.tech.cam.ac.uk/. The global research team is working on a program called "CARES", The Cambridge Centre for Advanced Research and Education (CARES) in Singapore". Some of the CARES research programs can be viewed at https://www.cares.cam.ac.uk/about-us/.

Texas Tech University

We would like to welcome Texas Tech University (TTU) – Data Intensive Computing Laboratory (DISCL) https://discl.cs.ttu.edu/doku.php?id=home. The research team at TTU is working on data management of self-describing files. Please check the publications from the DISCL research team at https://discl.cs.ttu.edu/doku.php?id=publications.

Allotrope Data Format (ADF):

We would like to thank the dedicated HDF Group engineering team for the 2021 support: Aleksandar Jelenak and Brian Sawicki.

ADF Library Version 1.5.4RF Released

The ADF Library version 1.5.4RF was recently released by the HDF group support team.

- The team was working on improved .Net C# NuGet package versioning. Initially a pre-release of the C# NuGet package 1.5.4 was issued and tested by a community member. The new release of the new C# NuGet package resolves the product versioning issues ADF #111
- Following the above improvement, the GitLab CI pipeline for the C# product release was updated and completed including:
    - Produce ZIP archives from ADF C# NuGet packages for release package
    - Automatically generate the required assembly version information during build stage
    - Automatically verify assembly version information in built ADF DLLs
- Similarly, a pre-release of the java package 1.5.4 was issued tested by a community member. The package is fixing the use of wrong IRI for hash codes (ADF #101)
- In addition, a bug of a supported special character in the Data Package file and folder name was fixed (ADF#122)

ADF Library version 1.5.4RF is available for download via:

- Client Connect: here
- JFrog Artifactory:
    - Java Library: https://allotrope.jfrog.io/ui/repos/tree/Properties/libs-release-internal
    - C# Library: https://allotrope.jfrog.io/ui/repos/tree/Properties/libs-release-internal-net

Support tickets can be opened at the ADF repo (here). Please select the “new bug” template to describe the issue and attach any supporting artifact.

Assessment of Any ADF Library Vulnerability to the Recent Log4J Exploit

Aleksander Jelenak from the HDF Group support team reported his assessment about the recent Log4i exploit (https://logging.apache.org/log4j/2.x/security.html) and any possible ADF Library vulnerability. The assessment is reported on the following ADF ticket https://gitlab.com/allotrope/adf/-/issues/124.

Following this assessment, we conclude that the ADF Library is not vulnerable to the Log4j exploit (CVE-2021-44228)

Allotrope Simplified Model (ASM):

We would like to thank the dedicated and talented Osthus engineering team for this project: Thomas Weber, Gerhard Wickler, Harsh Thakker, and Jennifer Aschenbrenner.

New Set of Allotrope Simple Data Model Released

26 new ASM Models - "Fresh from the oven": The product team is excited to officially publish the second set of the new JSON based simplified models (ASM) as part of the Q4 release. It includes the release of a Liquid-Chromatography ASM model. Altogether, we have already published 31 ASM models. The project is within its final phases. The product team will make the final few models available officially to the community in the coming weeks, as soon as they pass the compatibility testing with the related ADM.

The ASM Transformation Tool Project

Following the ASM 2021 RFP the product team together with a professional software development team from OSTHUS is working diligently on the final execution phases of this project. The ASM is a simple text representation of the Allotrope tabular models using JSON syntax. It is an extension to the existing ADM in its Tabular and Aggregation structure. Compatibility between the ASM data representation and the ADM data representation is achieved by transforming the syntactic representation of an ADM into an ASM simplified format, while maintaining data and semantic consistency.

The project included several milestones and deliverables including:

- ASM JSON Schema definition
- ADM to ASM transformation tool
- Tool testing including ASM models generation
- ADM to ASM transformation wiki
- ASM Primer wiki

Working with the Allotrope Simple Model

The second set of ASM models is available. The few remaining models are coming soon. Adopters can generate example results of tabularized data based on the new JSON ASM format and may get access to the project wiki.

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 WG is ready to review drafted models.

Allotrope Data Modeling (ADM):

Tensile ADM Release

The Modeling WG, with the initiative and SME support by Dow, was working on the development of a Tensile Model (Tabular). Tensile testing, also known as tension testing, is a fundamental materials science and engineering test in which a sample is subjected to a controlled tension until failure (https://en.wikipedia.org/wiki/Tensile_testing). This Tensile tabular model is released with a REC maturity status

Updates to Tabular Models

During the development of the ADM to ASM transformation methodology and tool, several bugs were found in several tabular models. Bug fixes were issued accordingly.

Intact Mass ADM Promoted

Last quarter, the Mass Spectrometry WG led by Adam Kimball (Applied Graphs) and Graham McGibbon (ACD/Labs) was focusing its effort to create a new graph model for Intact Mass analysis, which is the assessment of a protein's total molecular weight by mass spectrometry. Test data can be found here. This quarter the Intact Mass graph model is promoted with a REC maturity status.

Allotrope Foundation Ontology & Data Models (AFO/ADM/ASM)

Modeling teams have continued working to align on proposals to expand the domain coverage of the AFO, ADM and ASM. Easily access files located on Client Connect here and more granular technical details available on GitLab, here.

See www.allotrope.org/product-releases for full list of available models updated on December, 2021.

AFO Updates

The Allotrope Merged Ontology Suite release is available on:

- BioPortal, the repository of biomedical ontologies: https://bioportal.bioontology.org/ontologies/AFO
- Client Connect: here
- Gitlab: https://gitlab.com/allotrope/afo/-/tree/master/afo
- Allotrope PURL sever: http://purl.allotrope.org/ (listed under AFO>MERRGED)
- JFrog Artifactory: https://allotrope.jfrog.io/ui/repos/tree/General/AFO-release-public
- Allotrope website: https://www.allotrope.org/ontologies

ADM/ASM Updates

The Allotrope Data Models and Allotrope Simple Model Suite release is available with a restricted access to the community members 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

ADM Model Creation or Updates

New or Updated Data Models in Recommended (REC) status below:

Model	Type	Maturity	Path	GitLab Links
Tensile Test	Tabular	REC	New	Diagram Shape
MS Intact Mass	Graph	REC	CR>REC	Diagram Shape

ASM Model Creation or Updates

Here is the list of the set of ASMs released in this quarter.

ASM Model	Type	Maturity	Path	GitLab Links
balance	Tabular	REC	ASM released in 2021/09	JSON schema Sample file
bga	Tabular	REC	New ASM	JSON schema Sample file
bulk-density	Tabular	REC	New ASM	JSON schema Sample file
cell-counting	Tabular	REC	New ASM	JSON schema Sample file
cell-culture-analyzer	Tabular	REC	New ASM	JSON schema Sample file
conductivity	Tabular	REC	ASM released in 2021/09	JSON schema Sample file
disintegration	Tabular	REC	New ASM	JSON schema Sample file
dsc	Tabular	REC	New ASM	JSON schema Sample file
dvs	Tabular	REC	New ASM	JSON schema Sample file
fluorescence	Tabular	REC	New ASM	JSON schema Sample file
ftir	Tabular	REC	New ASM	JSON schema Sample file
light-obscuration	Tabular	REC	New ASM	JSON schema Sample file
liquid-chromatography	Tabular	REC	New ASM	JSON schema Sample file
loss-on-drying	Tabular	REC	New ASM	JSON schema Sample file
luminescence	Tabular	REC	New ASM	JSON schema

				Sample file
nephelometry	Tabular	REC	New ASM	JSON schema Sample file
optical-microscopy	Tabular	CR	New ASM	JSON schema Sample file
osmolality	Tabular	REC	ASM released in 2021/09	JSON schema Sample file
oven-kf	Tabular	REC	New ASM	JSON schema Sample file
pcr	Tabular	REC	New ASM	JSON schema Sample file
ph	Tabular	REC	New ASM	JSON schema Sample file
psd	Tabular	REC	New ASM	JSON schema Sample file
raman	Tabular	REC	New ASM	JSON schema Sample file
sem	Tabular	CR	New ASM	JSON schema Sample file
specific-rotation	Tabular	REC	ASM released in 2021/09	JSON schema Sample file
surface-area-analysis	Tabular	REC	New ASM	JSON schema Sample file
tablet-hardness	Tabular	REC	New ASM	JSON schema Sample file
thermogravimetric-analysis	Tabular	REC	New ASM	JSON schema Sample file
titration	Tabular	REC	New ASM	JSON schema Sample file
ultraviolet-absorbance	Tabular	REC	New ASM	JSON schema Sample file
x-ray-powder-diffraction	Tabular	REC	ASM released in 2021/09	JSON schema Sample file

Term Mapping Database

Most recent Allotrope Terminology Mapping (ATM) project Proof of Concept can be demonstrated upon request to the Product Team. The PoC demonstrates the mapping of vendor specific instrument to the *Cell-Counting* and *qPCR* Tabular models.

The term mapping project is Allotrope's collating instrument term mappings (to the standardized AFO) that already exist in the Community and housing them in a single central database. The Term Mapping database will:

- Support Allotrope Partners and Members by reducing the tedious efforts often required by teams to map terms
- Provide direct value to the entire Community by reducing the time needed to harmonize data

Allotrope Product Team is coordinating with Members and Allotrope's Partner Network to gather mapping information to ensure consistency of mapping across vendors. Several vendor instrumentation manuals are already available on Client Connect. The artifacts are categorized by the Allotrope ADM. Access to this repository is restricted and may be available upon request.

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

Modeling: (Notes: https://highq.in/5hj7qpacex)

- Progress noted in AFO/ADM section
- The Modeling WG was focusing on the development of a Tensile Model (Tabular).
- A tentative instrument priority list for modeling in 202 was provided to the Modelling Working Group by George Van Den Driessche from Biogen.

Chromatography: (Notes: https://highq.in/1xgh2gjurj)

- Following the Q3 tabular modeling effort, the LC-UV ASM model is released
- The Chromatography working group members are evaluating possible the next modeling efforts. Some of the possible direction include: Linking column information, assaying or diagnostic data capture, and compound identification (The compound identification modelling is a cross technique effort). Community members are welcome to join and express their priorities.
- In addition, the Chromatograph working Group may be looking on additional tabular modeling in the Chromatography domain and the related ASM.

MS: (Notes: https://highq.in/6igryjhpus)

We would like to thank Adam Kimball and Graham McGibbon for stepping forward and leading the MS working group as of May 2021.

- The team was working on promoting the base model use case of measuring the intact mass of proteins to a recommended status.
- The model was revised to handle different types of peaks and their relationships to various mass spectra (processed mass spectrum, deconvoluted mass spectrum, and base mass spectrum).

- The working group was analyzing the SPARQL queries to an ADF data instance file of the intact mass model (such as Deconvoluted Mass Spectra for a chromatographic peak etc.), possible model extensions and discussed the possibility of a derivative into a tabular model.

Semantic: (Notes: https://highq.in/6ihpr1kx5x)

- Continued discussions related to improving design and functionality.

Tooling, Testing, QA and Automation Pipeline

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

Raptor was added to the AFO QA Pipelines to improve BioPortal deployment

Following an analysis of the merged AFO that was done by the BioPortal support team at Stanford University, an additional tool called Raptor was added to the AFO QA pipeline. Raptor provides a set of parsers and serializers that generate Resource Description Framework (RDF) triples by parsing syntaxes or serialize the triples into a syntax. For more information, please refer to https://librdf.org/raptor/

Raptor was added as a job to the AFO pipeline, specifically to the "Check" stage. The output is to a log. To view the AFO pipeline, please refer to https://gitlab.com/allotrope/afo/-/ci/editor?tab=1

BioPortal examines (nightly) our PURL server and auto pulling the content in case of an updated release.

Automating AFO Deployment to Arifactory

The AFO pipelines are automating the deployment of the AFO to the JFrog Artifactory. Please refer to the "Deploy" stage of the AFO pipeline https://gitlab.com/allotrope/afo/-/ci/editor?tab=1. To view and downloaded AFO deployed packages on JFrog Artifactory, please refer to https://allotrope.jfrog.io/ui/repos/tree/General/AFO-release-public

Migration of the IRI Minting tool (Domain ID generator)

The IRI Minting Tool (domain ID generator) was migrated to Allotrope AWS infrastructure. The source code was loaded to the Allotrope GitLab opensource repository and can be found at https://gitlab.com/allotrope-open-source/ontology-qa-tools/-/tree/master/artificial-id-generator

In addition, Allotrope specific production Minting tool server documentation such as counter IDs, AWS installation, admin, and troubleshooting is located under the Allotrope-Internal GitLab group repository https://gitlab.com/allotrope-internal. Access to this repository is restricted and may be available upon request.

Migration of the Allotrope PURL server repository

The Allotrope PURL server source code repository and the documentation wiki including the installation instructions was migrated to the Allotrope-Internal GitLab group repository https://gitlab.com/allotrope-internal. Access to this repository is restricted and may be available upon request.

The Online Allotrope SHACL Validator is Offline

Given the availability of public opensource tools for SHACL validation, the online Allotrope SHACL Validator is taken offline. The tool and the source code are available on GitLab and Client Connect:

- GitLab: https://gitlab.com/allotrope-open-source/shacl-validator
- ClientConnect: https://highq.in/80gu82c6zo

A good opensource online tool such as the *SHACL Play!* from Sparna can be found at:

- Github: https://github.com/sparna-git/shacl-play
- Online: https://shacl-play.sparna.fr/play/

It is a free online RDF data validation with SHACL. The *SHACL Play!* embeds TopBraid SHACL API from TopQuadrant.

AF Community

Allotrope YouTube Channel

The Allotrope YouTube Channel hosts the Allotrope Connect public presentations from 2020 to the latest fall 2021 Connect event. Public presentations from 2017 Connect event are available as well. The YouTube Channel is available here and organized by playlists here.

Community Website

A reminder that the Allotrope Foundation Community website is an effort to improve and centralize the non-public knowledge and communication with the community. The GitLab-hosted website can be accessed here (http://community.allotrope.org). Note this requires GitLab access.

Some of the pages to look for:

- Allotrope Foundation Data Models (ADMs): https://community.allotrope.org/resources/reference/semantic/allotrope_data_models/ The page includes information about the structure of ADM repository under the section "*ADMS artifacts repository*"
- ADF Application Programming Interface (API): https://community.allotrope.org/resources/reference/adf/api/
- ADF Specification: https://docs.allotrope.org/
- Allotrope Framework Semantic Style Guide: https://community.allotrope.org/resources/reference/semantic/style_guide/afo_style_guide/
- Allotrope Foundation Data Model and Ontology Governance Process: https://community.allotrope.org/resources/reference/semantic/governance/afo_adm_governance_process/

(If you would like access to GitLab, please fill out our onboarding questionnaire, here – It's a Google Form, so if your firewall denies access to Google drive links, please contact amnon.ptashek@allotrope.org).

Looking Forward - 2022

The Product Team is working on its resources and development planning for FY 2022.

- As part of the planning, we need your feedback. It would be very helpful to get an estimated list of new models that your company plans to develop with or submit to Allotrope governance in 2022.
- George Van Den Driessche from Biogen already provided to the Modelling Working Group a tentative instrument priority list for modeling in 2022 and we are looking into a very productive year.
- In addition, the Product Team is working on a plan to enhance our tooling and automation pipelines for QA and release of the AFO, ADM and ASM products. QA tools are integrated into the GitLab CI pipelines and run automatically as new content is developed. It optimizes the operational SDLC process and making it more efficient as well as error-proof with a high Return of Investment.

The Allotrope Product Team is looking forward to another very productive year in 2022 and to develop additional improvements to meet the evolving needs of our community.

Please contact us for any questions at product_team@allotrope.org.

Wishing the entire Allotrope community, a Happy New Year of 2022,

Allotrope Product Team