2023 Fall Allotrope Connect

Dates: Tuesday, October 31 – Thursday, November 2, 2023

Location: Benchling
10th floor
680 Folsom St
San Francisco, CA 94107
(Please proceed directly to the 10th floor to check-in for all days)

Allotrope Connect (Fall 2023)

Bringing together Allotrope members and the broader scientific community to discuss how we are delivering faster insights from new and existing data within organizations by improving standardization of data and its interpretation across laboratory and manufacturing operations.

With over 55 data models, and growing, Allotrope is leading the way in data standardization thereby allowing quicker access to insights within data.

2023 Fall Allotrope Connect Agenda

Day 1 – Foundation Board of Directors' Meeting and Technical Discussions

Time (PT)		Subject
Day 1: Tue, Oct 31		**Day 1: Open to Allotrope Foundation and APN Members** **Foundation Board of Directors Meeting and Hands-on Technical Discussions**
9:00	Meeting	**Board of Directors Meeting *(Invitation sent separately)***
9:00	Meeting	**Open discussion - Combined Technical Working Group meeting to discuss topics including model updates, new models, resolve issues, future developments, or other technical related topics.** ***Da Vinci conference room***

Day 2 - Public Session: Allotrope in Action

- **Welcome and Updates from Allotrope Foundation**
- **"Allotrope in Action" - Presentations from Allotrope Foundation and Allotrope Partner Network Members**
- **Implementation of Converters**
- **Welcome Reception and Dinner**

Time (PT)		Subject/Title	Speaker/ Moderator	Organization
Day 2: Wed, Nov 1		**Day 2 - Public Session**		
8:30-9:00		**Gatherings/networking** **Check-in at registration desk (10th Floor)**		
9:00	Presentation	**Welcome and short introduction**	**Corey Bakalarski**	**Allotrope Foundation**
9:15	Presentation	**Benchling welcoming address: "State of Tech in Biopharma", 2023 Report**	**Ashu Singhal**	**Benchling**
9:45	Presentation	**ASM at Scale Implementation – A Merck Update**	**Wes Schafer**	**Merck**
10:15		**Morning break**		
10:30	Presentation	**Facilitating ASM transformation, search, access, and usage via the Tetra Data Platform**	**Trevor Kent**	**TetraScience**
11:00	Discussion	**Bio-Rad roundtable discussion: "Modern Laboratory Workflow: Instrument Connectivity and Data Management in the Cloud"**	**Kashef Qaadri**	**Bio-Rad**
12:00		**Lunch break**		
1:00	Presentation	**Building digital documentary standards with Allotrope methodology**	**Jeff Shick**	**USP**
1:30	Presentation Demo	**Empowering Mass Spectrometry Through ZONTAL: Harnessing Allotrope Technologies for Vendor Agnostic Data Visualization and Drill Downs**	**Dennis Della Corte, Spencer Gardiner**	**Zontal**
2:00	Presentation	**Utilizing Allotrope models as an instrument data capture standard for data liquidity**	**Benson Lee**	**Ganymede**
2:30		**Afternoon break**		
2:45-4:15	Presentation Demo	**The power of an open source approach to data standardization and lab connectivity**	**Chris Severs, Joe Negri**	**Benchling**
4:15		**Closing Remarks**	**Corey Bakalarski**	**Allotrope Foundation**
5:30		**Welcome Reception (after workshop) and Dinner** ***Location to be announced***		

Day 3 - Allotrope Foundation Community Discussion

- **Strategic, business, and implementation topics**
- **Working Groups activities briefing**
- **Expanding Allotrope modeling domains**
- **Lessons learned and questions raised from implementing Allotrope**

Time (PT)		Subject/Title	Speaker/ Moderator	Organization
Day 3: Thu, Nov 2		**Day 3 – Open to Allotrope Foundation and APN Members**		
8:15-8:45		**Gatherings/networking Check-in at registration desk (10th Floor)**		
8:45		**Day 3 Opening Remarks - Day 2 reflections and day 3 intro**	**Corey Bakalarski**	**Allotrope Foundation**
9:00	Presentation	**Working Groups Readout**		**Allotrope Foundation**
	Presentation	**- Modeling Working Group Readout**	**Wes Schafer**	**Modeling WG**
	Presentation	**- Plate Reader Working Group Readout**	**Joe Negri**	**Plate Reader WG**
	Presentation	**- Chromatography WG Readout**	**Tim Lee**	**Chrom WG**
	Presentation	**- Mass Spec Working Group Readout**	**Sarah Robinson**	**Mass Spec WG**
9:30	Presentation	**Allotrope Business Model 3.0 Update**	**AF Executive Committee**	**Allotrope Foundation**
10:00-11:00		**Morning break. The building is conducting an evacuation drill on 11/2, 10am-11am. Plan for an outdoor hour break**		
Business track 11:00	Discussion	**Allotrope Business Model 3.0 (discussion and deep dive)**	**AF Executive Committee**	**Allotrope Foundation**
Business track 11:45	Discussion	**Allotrope Marketing Communications Strategy**	**AF Executive Committee**	**Allotrope Foundation**
Technical track 11:00-12:30	Discussion/Workshop (Conference room)	***"ASM for Flow Cytometry" Discussion/Workshop***	**Benson Lee, Shivani Ludwig**	**Ganymede, TetraScience**
12:30		**Lunch break**		
1:15	Discussion	**ASM converter development - Lessons Learned so far**	**Wes Schafer**	**Merck**
1:45	Discussion	**USP discussion session: *"Building a digital documentary standards with Allotrope methodology, process and tools"***	**Jeff Shick**	**USP**
2:15	Discussion	**Benchling discussion session: "ASM for Method"**	**Joe Negri**	**Benchling**
2:45		**Afternoon break**		
3:00	Discussion	**Hot topics and closing remarks**	**Corey Bakalarski**	**Allotrope Foundation**
3:45		**Meeting adjourned**		

Day 2 Presentations/Discussions/Demos

9:00am - Welcome and short introduction (Corey Bakalarski, Allotrope Foundation)

- **Corey E. Bakalarski**, Ph.D. Chairman of Allotrope Foundation. Senior Principal Bioinformatics Scientist, Computational Proteomics and Analytical Data Science Group Depts. of Microchemistry, Proteomics & Lipidomics and Research Informatics & Software Engineering, Genentech, Inc.

9:15am - Benchling welcoming address: "State of Tech in Biopharma", 2023 Report (Ashu Singhal, Benchling)

Benchling will be sharing their findings of the inaugural 2023 State of Tech in Biopharma report. In a new era defined by the interplay of biology, data, and AI, biopharma's tech strategy takes on critical importance. Benchling surveyed 300 experts from biopharma companies large and small — across R&D and IT — to understand current usage and impact of these technologies, in addition to the challenges companies face in fully implementing and adopting them. Based on these learnings, Benchling developed and recently launched Benchling Connect, an open platform that enables customers to automatically capture and maximize value from their instrument data. Built upon this platform, Benchling will also preview upcoming products that leverages LLM for data analysis and reporting.

- **Ashu Singhal** is co-founder of Benchling, pioneer of the R&D Cloud that powers the biotechnology industry. Since co-founding Benchling in 2012, Ashu has overseen the company's rapid growth. Today, more than 200,000 scientists at over 900 companies and 7,500 research institutions globally have adopted Benchling's R&D Cloud to make breakthrough discoveries and bring the next generation of medicines, food, and materials to market faster than ever before. As co-founder, Ashu is responsible for leading product development and strategy, helping scientists increase their productivity, collaborate more effectively, and reach new data-driven insights. Before Benchling, he co-founded a social media analytics company that was acquired by Twitter. He holds a B.S. in computer science and mathematics from the Massachusetts Institute of Technology.

9:45am - ASM at Scale Implementation – A Merck Update (Wes Schafer, Merck)

This is an update on a 2 ½ year project to "FAIRify" Merck's instrument data by implementing the Allotrope Simplified Model (ASM) at scale. It will convert data from over 6,000 instruments to ASM and leverages a Knowledge Graph to index and serve that data.

- **Wes Schafer**, Director, Business/Tech. Analysis, Develop and Make Product, Composition of Matter Product Line, MRL IT

10:30am - Facilitating ASM transformation, search, access, and usage via the Tetra Data Platform (Trevor Kent, TetraScience)

Last year, TetraScience co-presented with Merck on "FAIRify"ing its instrument data via the Allotrope Simple Model. A year later, TetraScience has facilitated the ingestion and ASM transformation of Liquid Chromatography (LC) data from the Empower, Chromeleon, and Chemstation Chromatography Data Systems (CDS), Fast Protein Liquid Chromatography (FPLC) data from Unicorn, and Gas Chromatography (GC) data from

Empower using TetraScience's Intermediate Data Schema (IDS). Lessons from this work have informed the "raw data to ASM" journey, which will be applied to the next phase of the project in plate readers and beyond. Additionally, we have outlined a path for programmatic ASM access via REST and SQL interfaces via Tetra Data Platform, a key component for enabling automated data workflows to other laboratory software, analytics, and AI.

- **Trevor Kent** serves as a Data Engineer for TetraScience where he has been working on an initiative to convert instrument data into Allotrope Simple Model at scale. Prior to joining TetraScience, he worked as a Data Scientist at Accenture with a particular focus in Network Science. He also earned a MS in Applied Mathematics and look forward to applying that skill set through direct usage of ASM's.

11:00am - Bio-Rad roundtable discussion: "Modern Laboratory Workflow: Instrument Connectivity and Data Management in the Cloud" (Kashef Qaadri, Bio-Rad

Join our roundtable discussion focusing on the synergies between cloud-based data management and laboratory instrument connectivity. In today's rapidly evolving scientific landscape, harnessing the power of the cloud and seamlessly integrating laboratory instruments are revolutionizing research. This session will delve into leveraging cloud platforms, for secure data storage, and optimizing laboratory workflows through enhanced instrument connectivity.

Key Discussion Points:

- Cloud-Based Data Management: Explore migrating laboratory data to the cloud, including scalability, accessibility, and cost-efficiency. Discuss data organization, storage, and retrieval in cloud environments.
- Instrument Connectivity: Explore instrument connectivity, IoT applications, and data standards. Discuss connected instruments may support data accuracy and streamline laboratory operations.
- Data Security and Compliance: Discuss concerns related to data security, privacy, and regulatory compliance in cloud-based environments. Explore encryption methods, compliance standards, and data governance practices specific to scientific research data.
- Collaborative Research: Discuss how cloud-based data management and instrument connectivity facilitate collaborative research efforts. Discuss case studies highlighting successful interdisciplinary collaborations and shared data initiatives.
- Automation and Integration: Discuss laboratory automation and integration of instruments with Laboratory Information Management Systems (LIMS). Discover how automation enhances productivity and reduces human errors.

- **Kashef Qaadri** is the Director of Global Marketing at Bio-Rad, where he oversees product management and product strategy.

1:00pm - Building digital documentary standards with Allotrope methodology (Jeff Shick, USP)

The pharma/biopharma companies are exploring advanced technical capabilities and implementing big data, AI, and digital technologies to accelerate production, distribution, smart facility and supply chain.

In 2022, FDA established an advanced manufacturing program, and the European Medicines Agency launched a Quality Innovation Expert Group. Their prime criteria are digitalization and technology adoption in manufacturing.

USP has independently evolved its digital reference standards and documentary standards to support these Pharma 4.0 initiatives. Our digital documentary standards product portfolio offers reusable, modulated, executable building blocks to rapidly create applications. It comes with the same scientific rigor of the quality trusted in the printed version of the compendium.

Our approach adopted the Allotrope methodology of data cubes, data models and ontologies components with underlying JSON data format.

- **Jeffery Shick,** R.Ph. Director, Translational Informatics, Digital & Innovation, USP. Jeff is a pharmacist and clinical informaticist from Washington, DC. For over 20 years he worked to implement and manage coding mechanisms that foster data granularity, application, exchange and portability of drug information databases and derive new products from existing data sets. Jeff is currently Director of the Translational Informatics division within the Digital & Innovation Office at the US Pharmacopeia. Translational Informatics aims to sustainably embed USP standards and solutions into the electronic systems and workflows of software used across the drug supply chain by delivering scientific content in the form of high-quality, machine-readable data and digital services. Translational Informatics is also a vehicle for advocacy, identifying areas where substandard digital offerings pose a public health risk, or where new digital standards / solutions would improve public health.

1:30pm - Empowering Mass Spectrometry Through ZONTAL: Harnessing Allotrope Technologies for Vendor Agnostic Data Visualization and Drill Downs (Dr. Dennis Della Corte and Spencer Gardiner, Zontal)

In the rapidly evolving landscape of mass spectrometry, the integration of innovative technologies has become paramount to enhance efficiency, accuracy, and analytical capabilities. This presentation delves into ZONTAL's usage of Allotrope Technologies, exploring how the platform revolutionizes mass spectrometry processes.

The objective of this presentation is to showcase how ZONTAL empowers researchers and analysts in the field of mass spectrometry. By enabling advanced data visualization, intuitive drill-down functionalities, and streamlined comparison across various vendors, the usage of Allotrope based technologies and other standards has significantly elevated the scientific and operational potential of mass spectrometry laboratories.

This presentation invites attendees to explore the transformative potential of ZONTAL and Allotrope technologies, emphasizing how this synergy has redefined the landscape of mass spectrometry. Through engaging discussions and real-world examples, participants will gain valuable insights into the practical applications of these technologies, empowering them to navigate the complexities of modern mass spectrometry with confidence and efficiency.

- **Prof. Dr. Dennis Della Corte** is Chief Science Officer at ZONTAL. He holds a PhD in Biophysics and MS Degrees in Biomedical Engineering and Medical Physics. His work focuses on data science and computational methods for drug discovery and protein design.
- **Spencer Gardiner** is a Data Scientist at ZONTAL. He holds a BSc in Chemical Engineering from Brigham Young University.

2:00pm - Utilizing Allotrope models as an instrument data capture standard for data liquidity (Benson Lee, Ganymede)

Data standardization and structuring projects have nuanced considerations due to what different stakeholders are looking for. This presentation discusses considerations/tradeoffs for implementation, advocating for rapid iteration / feedback cycle and stakeholder communication as key tenets to prioritize in implementation design.

- **Benson Lee**, Ganymede Founder (Implementations, Science, Eng). Before founding Ganymede, Benson was Head of Quantitative Markets at Affirm, where he built and oversaw quantitative research, capital markets deal structuring, and model risk management. He received BS and MEng degrees in Computer Science from Cornell.

2:45pm - The power of an open source approach to data standardization and lab connectivity (Chris Severs and Joe Negri, Benchling)

Every company conducting scientific research and development faces the same set of challenges when integrating lab instrument data into the scientific process. Lab instruments produce proprietary data, leading to unnecessary and costly obstacles to data governance, standardization, and connectivity. Scientists are forced to manually transfer data and manage customized data pipelines, preventing them from using data at scale, creating burdens on IT and automation teams, and introducing regulatory compliance risk.

To address these chronic data lifecycle challenges, Benchling has taken an open-source, ecosystem-based approach to data standardization and instrument connectivity, applying FAIR best practices. We've provided an open source library of instrument data converters (in ASM format partnering with Allotrope) to support this initiative.

In this presentation, we'll share:

- An overview of the open source approach and why it's powerful
- Key lessons learned from our partnership with the Allotrope working group on this initiative
- A demonstration of our open source contribution model in action

- **Chris Severs** leads the Automation and Analytics Engineering group at Benchling. Prior to Benchling, Chris worked on data science, ML, and applied research problems at eBay, PayPal, and Amazon over the past 10 years. Chris holds a PhD in Mathematics and did research as an academic mathematician before making a switch to industry.
- **Joe Negri** is a Product Manager for Automation and Analytics at Benchling. Joe worked as a research scientist for over 15 years before transitioning into product development. Prior to joining Benchling, Joe was a Product Manager at Agilent Technologies, and previously worked at the Broad Institute of Harvard and MIT, and the Dana-Farber Cancer Institute. Joe received his PhD in Biological and Biomedical Sciences from Harvard University, and bachelor's degree in neuroscience from Trinity College Dublin.

Day 3 Presentations/Discussions

8:45am - Opening Remarks. Day two reflections and day three introduction (Corey Bakalarski, Allotrope Foundation)

- **Corey E. Bakalarski**, Ph.D. Chairman of Allotrope Foundation. Senior Principal Bioinformatics Scientist, Computational Proteomics and Analytical Data Science Group Depts. of Microchemistry, Proteomics & Lipidomics and Research Informatics & Software Engineering, Genentech, Inc.

9:00am - Modeling Working Group Readout (Wes Schafer, Modeling WG)

- WG progress
- Modularization of the tabular models
- Upcoming modeling efforts

- **Wes Schafer**, Director, Business/Tech. Analysis, Develop and Make Product, Composition of Matter Product Line, MRL IT

9:00am – Plate Reader Working Group Readout (Joe Negri, Plate Reader WG)

- New WG,
- Common Plate Reader Model
- Objectives,
- Extend processed data in the model (SoftMax)

- **Joe Negri** is a Product Manager for Automation and Analytics at Benchling. Joe worked as a research scientist for over 15 years before transitioning into product development. Prior to joining Benchling, Joe was a Product Manager at Agilent Technologies, and previously worked at the Broad Institute of Harvard and MIT, and the Dana-Farber Cancer Institute. Joe received his PhD in Biological and Biomedical Sciences from Harvard University, and bachelor's degree in neuroscience from Trinity College Dublin.

9:00am - Chromatography Working Group Readout (Tim Leer, Chromatography WG)

- ASM Modularization
- Q2 and Q3 Release
- Community feedback on LC ASM
    - Common core and multiple detector support
- Prototyping Liquid Chromatography ASM Export t
- Chromatography Column model
- SMEs and APN participants

- **Tim Lee** is a senior software engineer for Shimadzu Scientific Instruments (SSI), US subsidiary of Shimadzu Corporation in Japan, headquartered in Columbia, Maryland. Tim joined SSI in 2014 and began participating in Allotrope in 2015. Starting with little scientific background in 2014, he enjoyed and continues to enjoy growing his knowledge along with the Allotrope community. Currently, He is assisting colleagues in Japan to integrate Allotrope into our software and keeping up-to-date with Allotrope.

9:00am – Mass Spectrometry Working Group Readout (Sarah Robinson, Mass Spectrometry WG)

- Walkup LC/MS Model status
- Walkup LC/MS review with instrument data
- MS WG activities
- New SME experience
- 2D Chromatogram from 3D data
- Processed data

- Generic MS tabular model
- Walkup LC/MS press release

- **Sarah Robinson** got her Ph.D. in Organic Chemistry studying Marine Natural Products with Phil Crews and did her postdoctoral studies in proteomics with Neil Kelleher. She current works at Genentech in Discovery Chemistry where she leads a group that performs structure elucidation and small molecule characterization.

9:30am – Allotrope Business Model 3.0 (Allotrope Foundation)

11:00am – Allotrope Business Model 3.0, Q&A and dive into topics (Allotrope Foundation)

11:45am – Allotrope Marketing Communications Strategy (Allotrope Foundation)

11:00am - "ASM for Flow Cytometry" Discussion/Workshop (Benson Lee, Ganymede. Additional SME: Shivani Ludwig, TetraScience)

- Brief discussion and intro on use case to be modeled, instruments and objectives
- Suggested Flow Cytometry model coverage
- Experiment planning step as a useful part of a complete ASM representation.
    -- Metadata and context of the run
    -- Handling recursive re-analysis
-Sample information, Instrument information, laser configuration, Instrument settings....
    -- Cell population gating
    -- Parameters and Measurements: Forward Scatter (FSC), Side Scatter (SSC), Fluorescence Channels, Compensation....
    -- Statistics and Analysis: Results data, process data, calculated data
- Integration with xMAP technology
- Exporting from data files/Original data files, such as FCS (Flow Cytometry Standard) files, which contain raw data and metadata
- Parsers that transform various flow instruments and data streams from planning and analysis.
- Suggested plan of action

Questions for discussion:

- Should we consider analytical software (like FlowJo) or instrument software as the interface?
    -- If we want to extract data from all sources, how can we ensure that we are able to have consistent representations from each piece of software, when approximations (bucketing) are not consistently performed between different analytical software? Relatedly, experimental setup data may not be readily parsable, which presents its own challenges on picking a common data shape.
- Should we include plate map design (for example, wells designated as FMO controls) as part of the standard? If so, how do we balance flexibility and consistency?
- How do we ensure that the Allotrope schema lines up well with pre-existing ISAC standards for flow cytometry event data and gating files?
- Should we be considering data at the event level, or at the aggregate statistics level?
- Should we consider a particular use case (i.e. - immunophenotyping) to frame out the data that is necessary/desirable to collect?
    -- How do we ensure alignment with already established schemas for cell viability assays / cell counters?
    -- How should we consider different use cases? For example, using a flow cytometer to sort cells would necessitate more careful collection of experimental setup, but this may be overkill for simpler experimental runs.

- **Benson Lee**, Ganymede Founder (Implementations, Science, Eng). Before founding Ganymede, Benson was Head of Quantitative Markets at Affirm, where he built and oversaw quantitative research, capital markets deal structuring, and model risk management. He received BS and MEng degrees in Computer Science from Cornell.

1:15pm - ASM converter development - Lessons Learned so far (Wes Schafer, Merck)

- **Wes Schafer**, Director, Business/Tech. Analysis, Develop and Make Product, Composition of Matter Product Line, MRL IT

1:45pm - USP discussion session: "Building a digital documentary standards with Allotrope methodology, process and tools" (Jeff Shick, USP)

- Current USP ontologies and the expected USP ontology domains expansions
    -- Alignments of the USP ontology with AFO (and BFO)
- USP expected models and domains to build a digital documentary standards
- Expectations on USP ontologies and data models merging to Allotrope AFO and ASM
- Compare Allotrope Governance process with USP digital documentary standards Governance process
    -- Tools and infrastructure from Allotrope that can help USP
- Suggestions and best practices on applying USP digital documentary standards to real world situations.

- **Jeffery Shick,** R.Ph. Director, Translational Informatics, Digital & Innovation, USP. Jeff is a pharmacist and clinical informaticist from Washington, DC. For over 20 years he worked to implement and manage coding mechanisms that foster data granularity, application, exchange and portability of drug information databases and derive new products from existing data sets. Jeff is currently Director of the Translational Informatics division within the Digital & Innovation Office at the US Pharmacopeia. Translational Informatics aims to sustainably embed USP standards and solutions into the electronic systems and workflows of software used across the drug supply chain by delivering scientific content in the form of high-quality, machine-readable data and digital services. Translational Informatics is also a vehicle for advocacy, identifying areas where substandard digital offerings pose a public health risk, or where new digital standards / solutions would improve public health.

2:15pm - Benchling discussion session: "ASM for Method (Joe Negri, Benchling)

Instrument methods are critical tools in analytical laboratories as they ensure that analytical instruments are used consistently and reliably to produce accurate and reproducible data. They are also important for regulatory compliance, as pharma is required to follow standardized methods to meet quality and safety standards. In a recent successful Pistoia PoC, ADM (full graph) was used to implement method exchange between different HPLC systems as well as different pharma labs. In the discussion, we will analyze the option to create a simple model to standardize an instrument method in a modular way .
Key parameters include:

-Analytical instrument type
-Sample preparation procedures
-Instrument settings (including calibration) to ensure that the instrument is operating within defined parameters.
-Data acquisition method that specifies how data should be collected
-Analytical technique that provides step-by-step instructions on how to perform the analysis.
-Calibration standards

-Quality control measures, to ensure data accuracy and precision.
-Optional data analysis: While not always a part of the instrument method itself, there may be instructions or references to data analysis procedures or software
-Optional method validation that may include parameters like specificity, accuracy, precision, linearity, and robustness.
-Documentation of the instrument method for traceability and regulatory compliance.

- **Joe Negri** is a Product Manager for Automation and Analytics at Benchling. Joe worked as a research scientist for over 15 years before transitioning into product development. Prior to joining Benchling, Joe was a Product Manager at Agilent Technologies, and previously worked at the Broad Institute of Harvard and MIT, and the Dana-Farber Cancer Institute. Joe received his PhD in Biological and Biomedical Sciences from Harvard University, and bachelor's degree in neuroscience from Trinity College Dublin.