BioMADE's Bioindustrial Manufacturing Data Perspective

- Andrew Eichenbaum, PhD
- Sr. Director of Data Science - BioMADE

- February 8, 2024
- 2024 Industrial Digital Ecosystem Summit

Credit: Novozymes

Outline

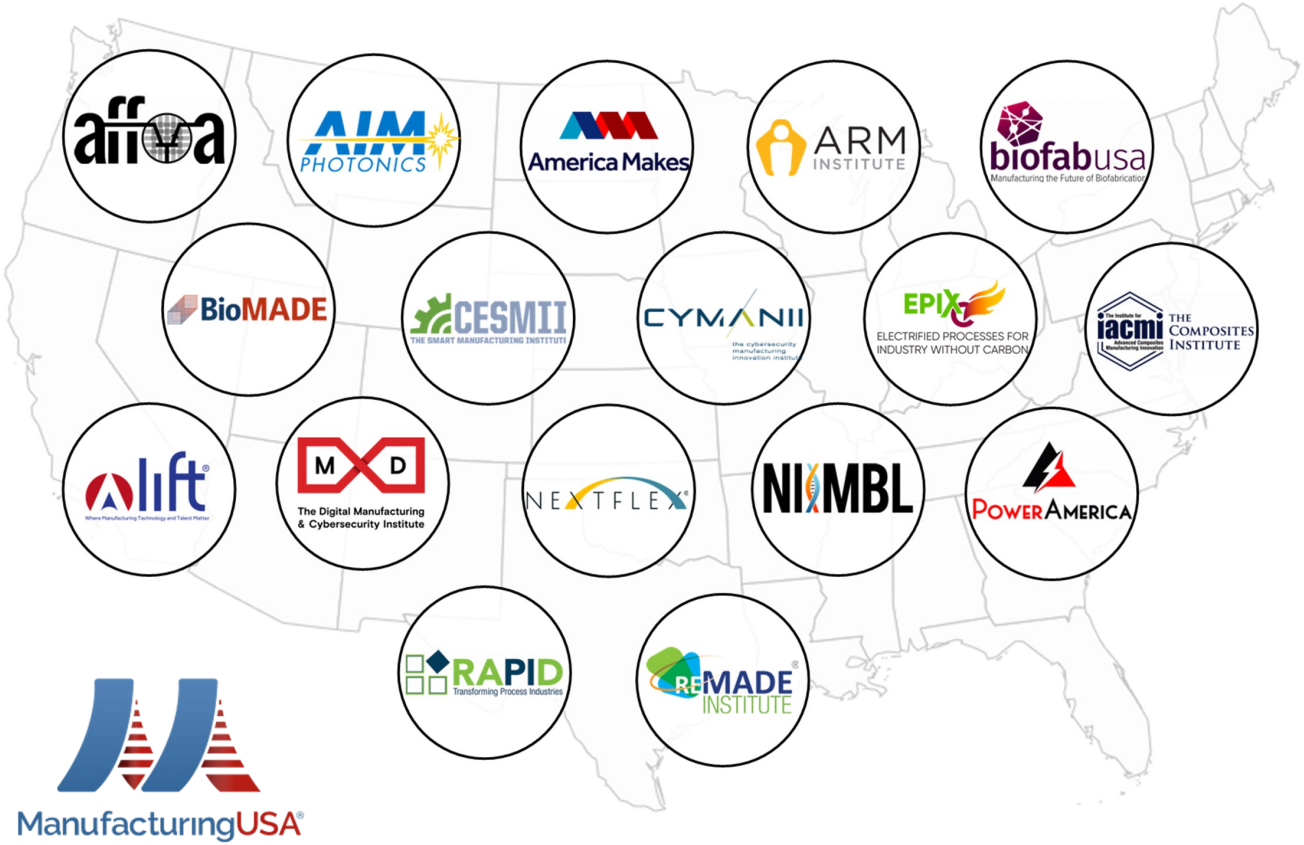
- Intro to BioMADE and Bioindustrial Manufacturing
- Efficiency through Advanced Process Modeling
- Bioindustrial Manufacturing Facilities Modeling
- Canonical Bioindustrial Data Model As a Path to Standards

BioMADE: the Bioindustrial Manufacturing Innovation Institute

About Us

BioMADE launched in 2021 and is an independent non-profit, public-private partnership initiated by the U.S. Department of Defense. In partnership with our members, we are securing America's future through biomanufacturing innovation, education, and collaboration by:

- Creating a dynamic member ecosystem
- Providing funding opportunities for members
- Increasing access to U.S. domestic scale up infrastructure



What is Bioindustrial Manufacturing?

Bioindustrial manufacturing uses living organisms such as bacteria, yeast, and algae to make new products or replacements for current products that are more sustainable and environmentally friendly than current processes

By harnessing the power of biology, bioindustrial manufacturing can make myriad products that Americans use every day

Bioindustrial manufacturing is key part of the bioeconomy, which could have an economic impact of up to $4 trillion annually within the next 10–20 years

Applications

- Novel and performance-driven chemicals, materials, catalysts, sensors, probiotics, and more
- Compounds that go into footwear, ink, and engine coolant
- Fibers that become coffee capsules, diapers, cups, and electronics
- Skincare products
- Growable concrete and on-site production of fuels, lubricants, and other critical materials

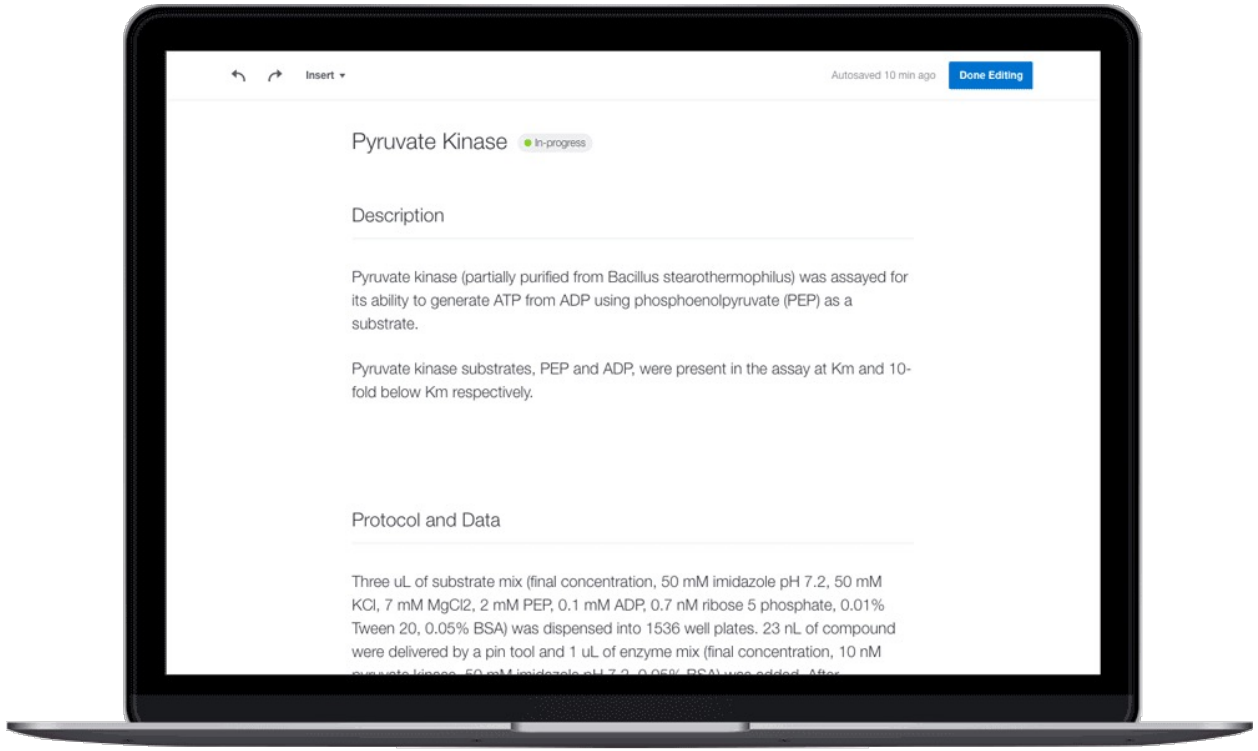
What Can Bioindustrial Manufacturing Create?

Applications

- Carbon-negative chemicals that can be used for water treatment, concrete, fertilizers, and detergents
- PFAS alternatives and bio-based fire-resistant composite materials
- Bioplastics and durable fibers
- Chemicals used to make compostable tote bags, coffee capsules, and food packaging
- Growable cement and alternative natural rubber to make tires
- Proteins, probiotics, fragrances, and skincare products

Efficiency through Advanced Process Modeling

Current State – Documents & Digital Lab Notebooks

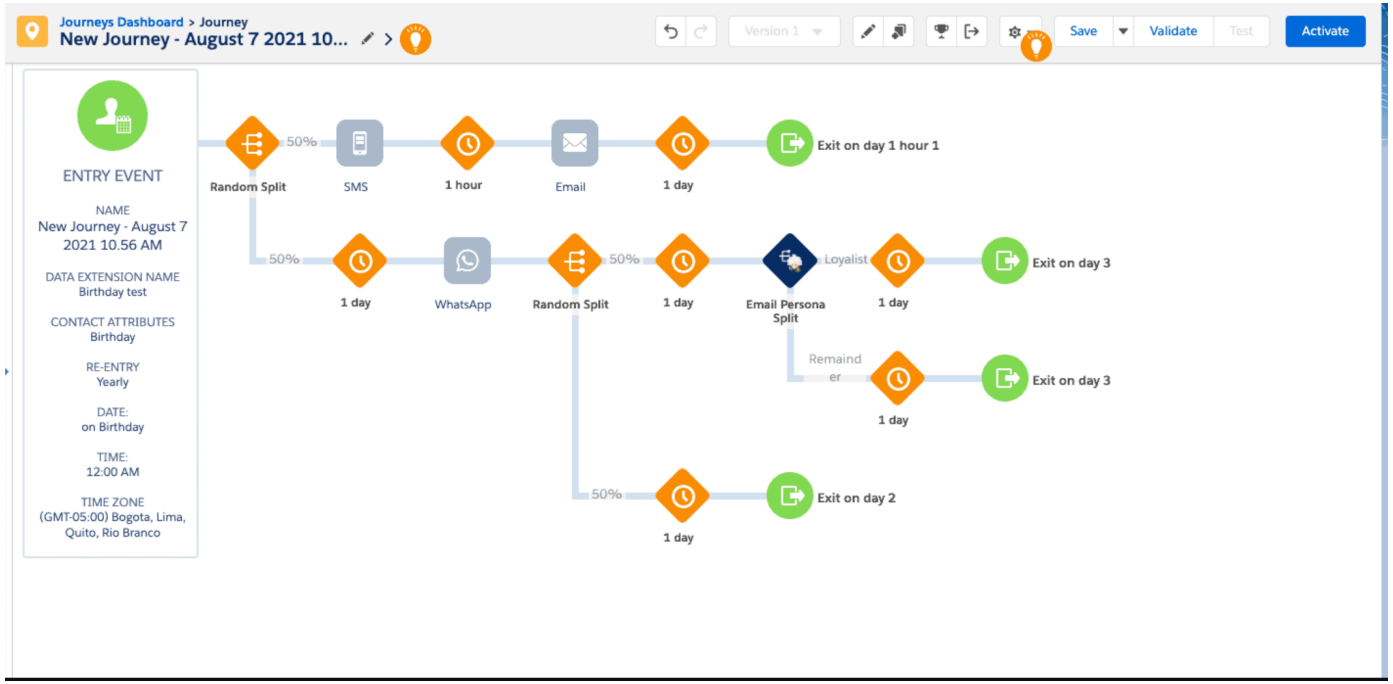


CDD Vault Electronic Lab Notebook

- Current state of the art in Electronic Lab Notebooks is a set of form driven documents
- Adds efficiency and clarity to lab notebooks but provides no insight into protocol/process changes
- Users are left with a set of digital files that must be manually searched and reviewed to provide insights

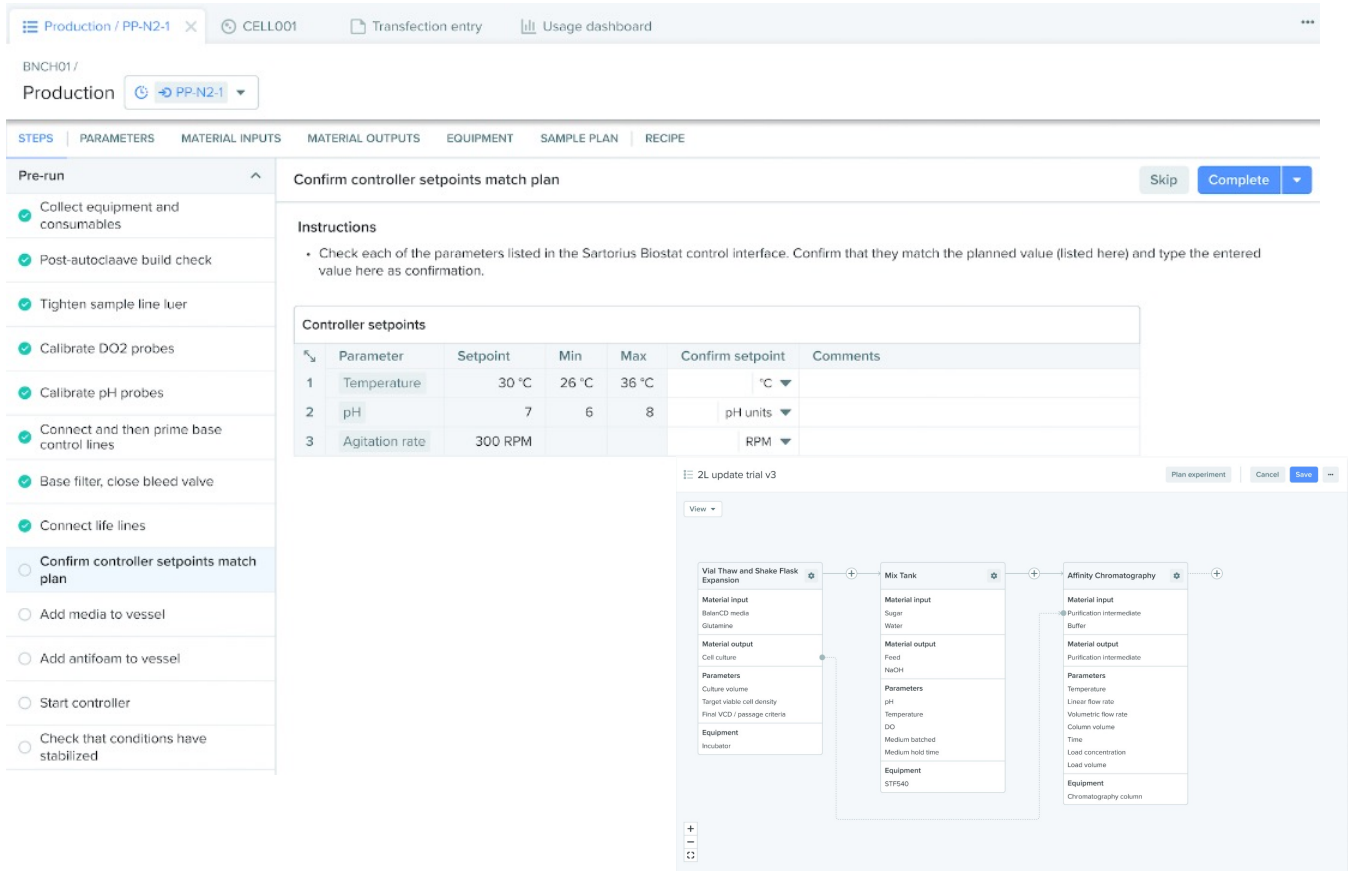
Future State – An Example

- Salesforce journey mapping allows Marketers to create and dynamically manage customer experiences
- Ability to then review an individual's journey at any point in time
- Why can there not be similar functionality for bioprocesses?



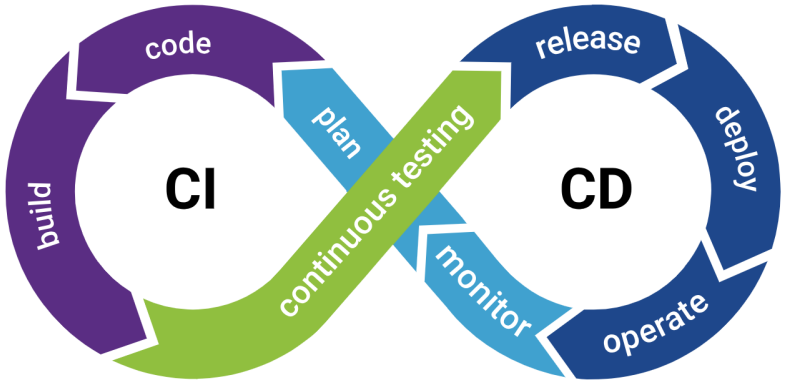
Future State – Process Workflow Model

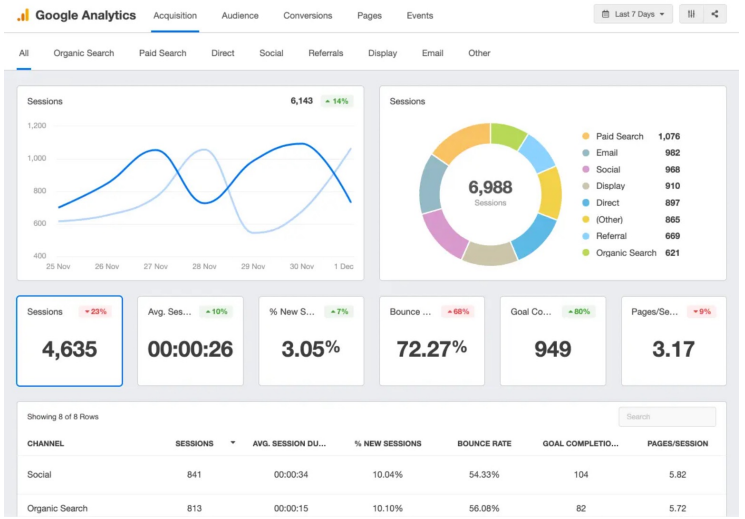
- Currently there are simple process models to guide development
- No way to easily understand the changes that have happened
- No way to link results to a set of process changes or issues at runtime



Benchling Bioprocess

Future State – Tech's Current State





With the introduction of Constant Integration and Constant Deployment tools with real time feedback tech companies know within hours if not minutes how changes have affected their users.

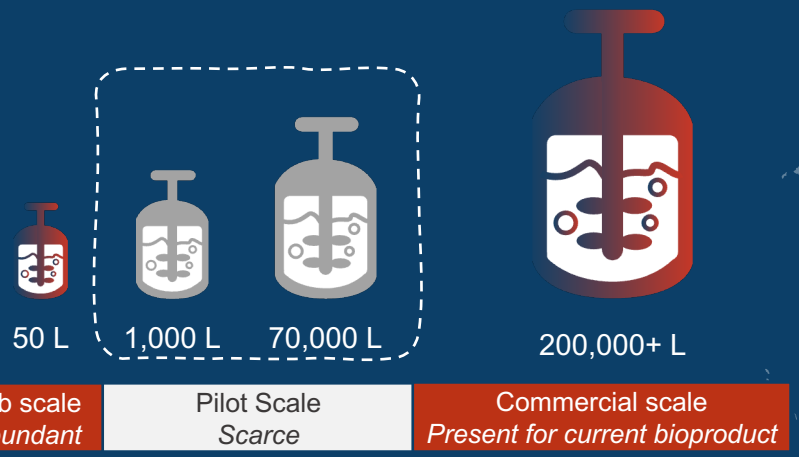
The goal is to have similar functionality such that a lab manager can ask of a digital platform, "What has changed in the past 4 runs, and how has that changed the outcome?"

Bioindustrial Manufacturing Facilities Modeling

The Domestic Scale-Up Infrastructure Gap

Gap is filled by going overseas

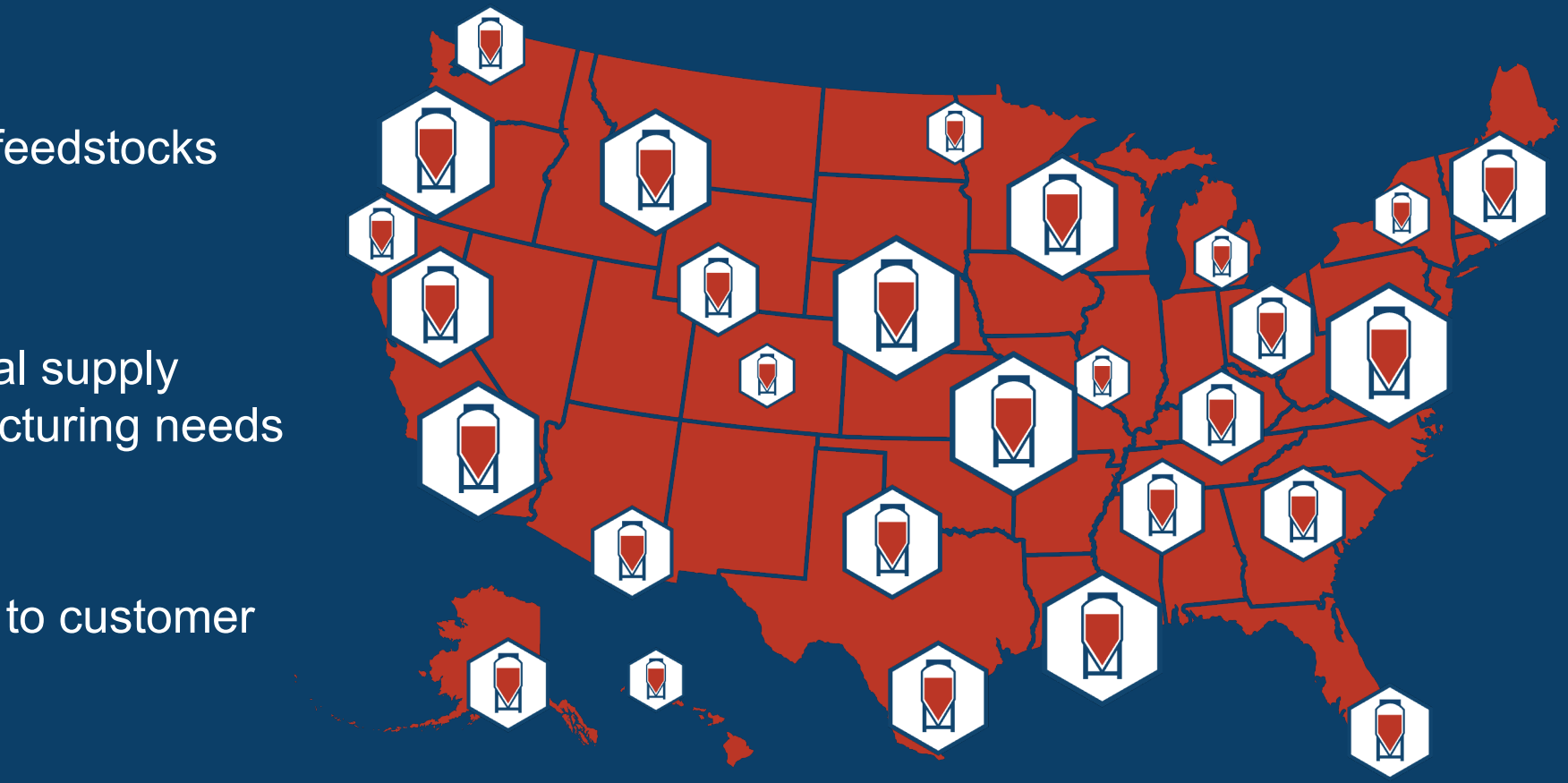
Lack of domestic scale-up facilities forces U.S. innovators to look elsewhere to scale or manufacture their biotechnology innovations



An Existing Network of National Pilot Facilities

Diversify and Specialize Infrastructure for a Robust and Resilient Network of Manufacturing Capabilities

- Leverage regional feedstocks
- Couple with regional supply chains and manufacturing needs
- Proximity locations to customer markets for testing



The Ontology of Bioindustrial Manufacturing CapEx

- There is a large range of CapEx process systems for bioindustrial manufacturing
  - Analytical Systems
  - Batch Vessel or Continuous Reaction Systems
  - Drying and/or Centrifugation
  - Extraction and/or Distillation
- Each of these systems provide a unique set of capabilities that when combined define a unique workflow
- The goal is to build an ontology of standard approaches and their associated systems so that we can map overall capability of a facility

Bioindustrial Manufacture Capabilities

High Level Bioindustrial Step

	Seed Train	Fermentation	Separation	Purification
Purpose	Scale up initial bio-source to production scale	Run bioprocess	Separate produced material from slurry	Refine produced material to final state
Equipment	1-50 liter reactor, assay equipment	10K+ liter reactor	Centrifuge	Filter or membrane
Non-Obvious Equipment	Clean room or hood	Assay equipment, feedstock management	Dryer	Storage

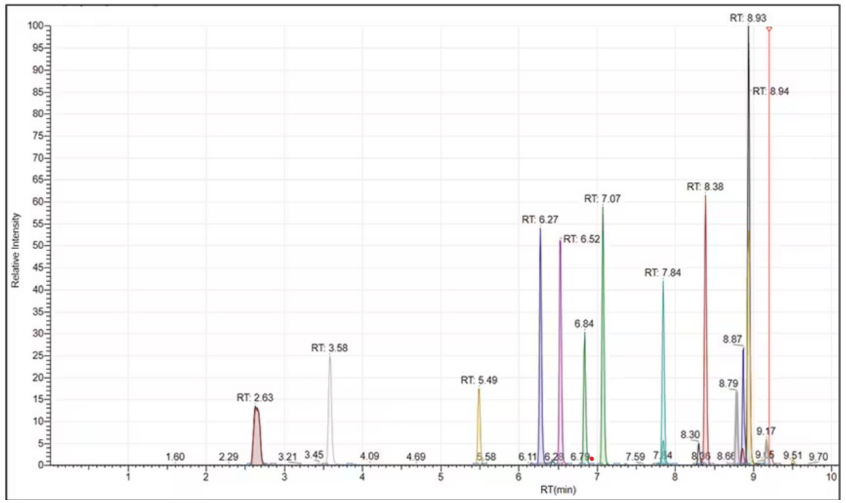
Without a full mapping of all equipment to its range of purposes, modeling will be non-optimal

Canonical Bioindustrial Data Model As a Path to Standards

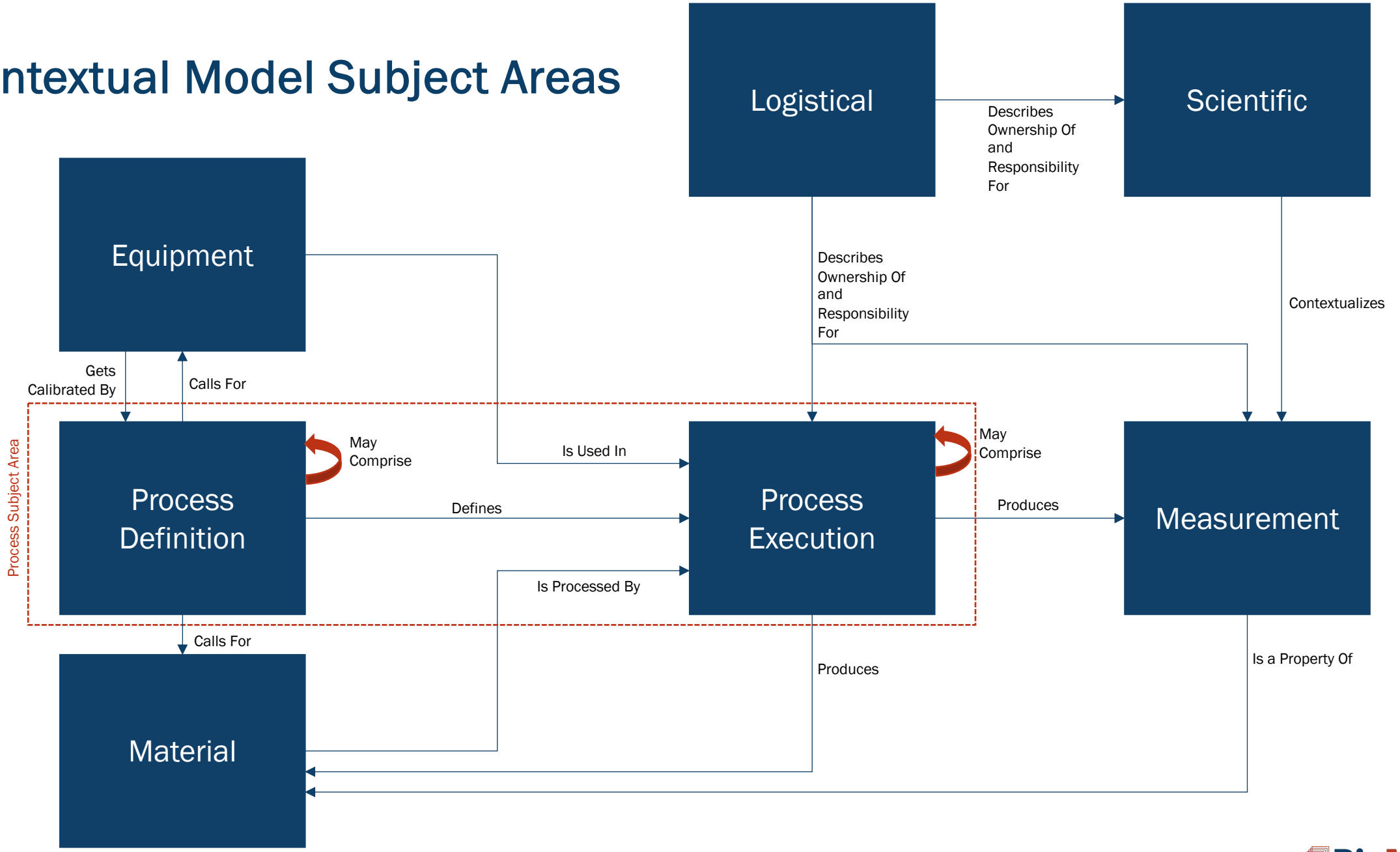
Credit: Ginkgo Bioworks

No Standard Data Formats for Bioindustrial Systems

- There is a large range of CapEx systems for Bioindustrial Manufacturing
  - Analytical Systems, Batch Vessel or Continuous Reaction Systems…
- For a single type of system there are multiple data formats
  - HPLC – High Performance Liquid Chromatography
    - Standard spectral output as a function of time
    - For each brand (model, software release) there are different output formats – binary, excel, flat…
      - Different formats provide different levels of data specificity
- Need to define a canonical, normalized model



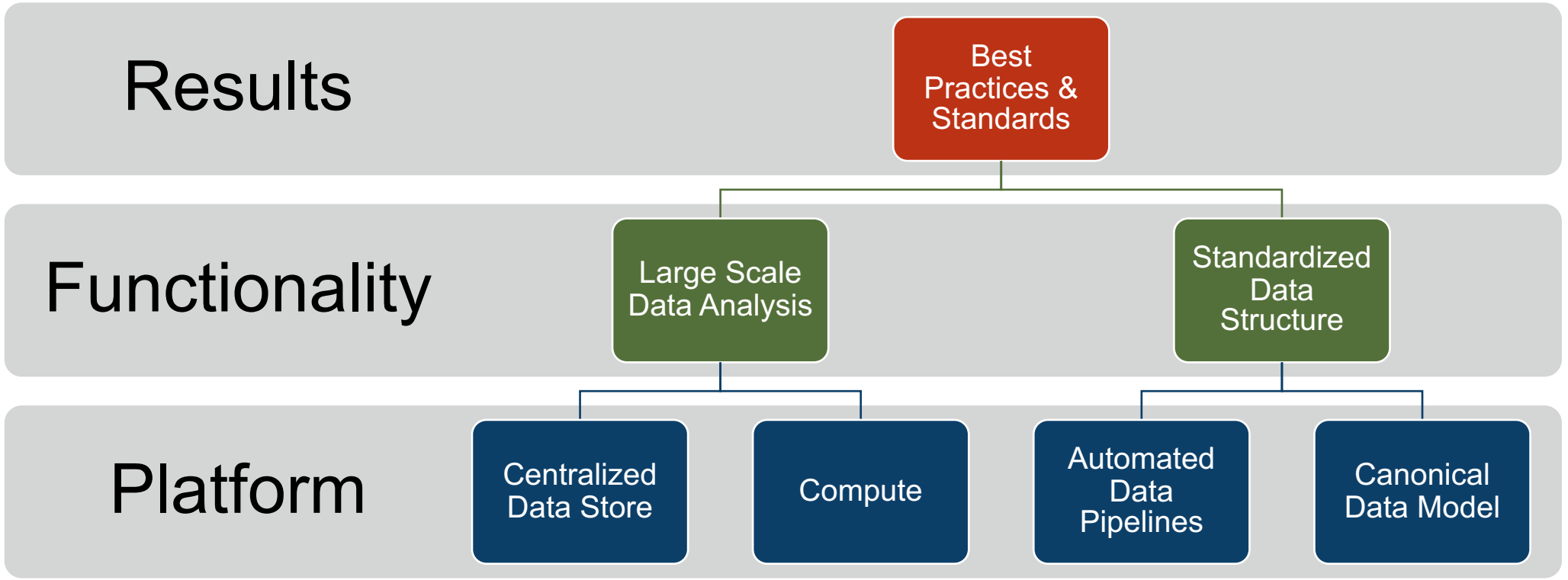
Contextual Model Subject Areas



Long Term Vision for Bioindustrial Data Platform

- **Data Storage:** A single centralized and secure store for all data
- **Data Acquisition:** Near real time data ingestion with fast availability and low overhead including automated data gathering options
- **Data Tools:** Native shareable analytic, dashboarding, and data science tools
- **APIs/Interfaces:** A well-documented and complete set of externally facing APIs and interfaces for integration purposes
- **Expert Data Store:** A centralized, managed, and trusted data store of information needed when performing analysis
- **Intuitive Interfaces:** Multiple intuitive interfaces for ease of interacting and sharing data
- **Templating:** Data and process standards/templates for ease of understanding and knowledge transfer

Canonical Data Models Lead to Standards



› BioMADE's goal is to create a platform where data can be anonymized across horizontal and vertical views to extract best practices and standards

Open Questions

- Multiple data models and ontologies have been created for Biotech and Bioindustrial processes, why is there no large-scale adoption?
- What would drive acceptance and usage of a bio-focused ontology and data model?
- What are the primary interfaces needed with a bio-focused data model?
- What other derived information would be useful to the bioindustrial community?
- What are the unknown unknowns?