Engineering Biology Metrics and Technical Standards for the Global Bioeconomy

Americas Workshop Report

Contents

Introduction

Historically, standards are often developed to remedy failures to interoperate. A particularly disastrous example is the Great Baltimore Fire of 1904: the various fire hoses from neighboring towns brought to fight the fire were incompatible with the local fire hydrants, resulting in unusable equipment and a more devastating blaze.[1] The need for standardized fire hose couplings was resolved with a 1905 standard, set by the National Fire Protection Association and other organizations for fire hydrants and hose couplings.[2]

To be impactful, standards must address a pressing need and be widely agreed upon, accepted, and adopted. In short, standards must be useful. They are needed at points of transfer, including data, IP, and material transfer, to ensure efficiency and efficacy, and to reduce risks when developing technologies with partners. As such, standards can alleviate common pain points to accelerate progress in a given field. Metrics and measurement science facilitate decision making by describing properties using quantitative data, and are crucial to supporting standards. That said, not every problem has a standards- or metrics-based solution. This report reflects that standards identification is complementary to and deeply intertwined with other supporting actions, including technological advancements, policy updates, and developing educational resources.

As advancements made in the field of engineering biology galvanize the development of a robust and diverse bioeconomy, standards setting becomes imperative for the success and growth of the biotech industry. Because these new biotechnologies and bio-based processes must integrate and compete with established ones, such as chemicals production in the petrochemical industry, preemptively identifying and establishing critical standards and metrics will help to accelerate the development of the bioeconomy and ensure product safety and efficacy.

Workshop Summary

The Engineering Biology Metrics and Technical Standards for the Global Bioeconomy project aims to identify community and stakeholder driven scientific, technical, operational, and semantics standards to enable and drive scale-up capabilities, improve reproducibility across batches and geographies, and enhance the performance of microbial factories and bio-products. Ultimately, these metrics and standards will enable global trade and markets for bio-based and bio-produced products. To achieve this, the project leadership Task Force hosted three workshops, held in different locations around the world, to convene representatives from industry, government, and academia to discuss barriers, opportunities, and potential for consensus on technical standards and metrics for engineering biology. These activities will lay the groundwork to establish open voluntary standards for engineering biology that enable the rapid growth and success of the bioeconomy.

The Engineering Biology Metrics and Technical Standards Workshop for the Americas region, hosted by the Engineering Biology Research Consortium (EBRC) and the National Institute of Standards and

---

[1] 1904 Baltimore Fire: 80 Blocks Burned And Lessons Learned. *History Daily* https://historydaily.org/1904-baltimore-fire.

[2] Seck, M. D. & Evans, D. D. *Major U.S. cities using national standard fire hydrants, one century after the great Baltimore fire*. NIST IR 7158 https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7158.pdf (2004) doi:10.6028/NIST.IR.7158.

Technology (NIST), included an evening reception at the Cosmos Club in Washington D.C. on June 7th, 2023, followed by two days of technical programming at the Institute for Bioscience and Biotechnology Research (IBBR) in Rockville, Maryland on June 8th and 9th. The first day of technical discussions featured plenary talks and panel discussions, and encouraged a high level of audience discussion and participation. The discussion topics and ongoing feedback from the participants informed the breakout topics on the second day, which were used to delve into details about specific metrics- and standards-related topics.

The Americas workshop welcomed participants from four countries in North and South America, though participants were predominantly U.S.-based. Thirty-eight organizations were represented, including six U.S. federal agencies; technical experts were present from 33 of the organizations. In total, some 60 attendees from industry, academia, and the U.S. federal government engaged in passionate and thoughtful dialogue about the standards and metrics that are needed in their industries and fields.

The objectives for the Americas Workshop were the following:

1. Survey the current ecosystem of engineering biology metrics and standards.
2. Identify metrics and standards needed to promote innovation and commercialization in a global bioeconomy.
3. Identify developments, technical and otherwise, required to achieve the metrics and standards needed.

An abridged agenda for the Americas Workshop is included below.

8 June 2023	
8:30 AM	**Welcome to Day 1** *India Hook-Barnard (EBRC) and Andrea Hodgson (Schmidt Futures)* Dr. Hook Barnard and Dr. Hodgson will provide an overview and objectives of the workshop.
9:00 AM	**Developing Metrics and Setting Standards** *Paul Freemont (Imperial College) and Plenary Discussion* Prof. Freemont will present key definitions for the workshop, describe current, past and failed efforts, and the purpose for the current effort.
9:30 AM	**The Current State of Standards and Metrics** *Sheng Lin-Gibson (NIST) and Plenary Discussion* Dr. Lin-Gibson will present the current state of engineering biology metrology, metrics, and standards and NIST efforts.
10:00 AM	Break
10:30 AM	**Informal Standards are Barriers to Using Non-Model Organisms** - Sarah Richardson (MicroByre)
10:50 AM	**Panel 1: Upstream Processing and Feedstocks** *Moderator: Jane Romantseva* *Panelists:* Aaron Schaller (MeliBio), Jonathan Jacobs (ATCC), Marilene Pavan (LanzaTech), Sarah Richardson (MicroByre), Swami Srinivas (Ginkgo)

12:00 PM	Lunch
1:00 PM	**Downstream Process Development for Precision Fermentation** - Stan Herrmann (Amyris)
1:20 PM	**Panel 2: Downstream Processing and Scale Up** *Moderator: Emily Aurand* *Panelists:* Elizabeth Onderko (Capra Biosciences), Sean Hunt (Solugen), Stan Herrmann (Amyris), Steve Evans (BioMADE), Vikramaditya Yadav (UBC)
2:30 PM	Break
3:00 PM	**Standards and Benchmarks for Automated Experimentation** - Pete Kelly (Align to Innovate)
3:20 PM	**Panel 3: Process Development and Data** *Moderator: Cynthia Ni* *Panelists:* Dave Vance (BU DAMP Lab), Emiley Eloe-Fadrosh (LBL), Nathan Hillson (LBL), Pete Kelly (Align to Innovate)
5:00 PM	Adjourn

9 June 2023	
9:00 AM	Welcome to Day 2: Overview and Objectives
9:30 AM	Breakout Session 1 A. Standards and metrics for **engineer*ed* biology as the product** (e.g., T-cells, crops) B. Feedstocks: components, consistency, and sustainability C. Standards and metrics for **engineer*ing* biology as the process** (e.g., organism, enzyme, strain platforms for biomanufacturing)
10:45 AM	Break
11:15 AM	Breakout Session 2 A. Standards that support regulations and biosecurity B. Translating and coordinating with existing standards and benchmarks C. Metrology: tools, platforms, and equipment
12:30 PM	Lunch
1:45 PM	Breakout Session 3 A. Best practices for data sharing and platform interoperability B. Safety, sourcing, traceability, public perception C. Connecting capabilities and competencies of CMOs for scale up and DSP
3:00 PM	Plenary Discussion and Workshop Summary
4:00 PM	Adjourn

Detailed notes were taken during the technical discussions at the workshop, and were carefully read, summarized, and organized following the workshop. Though the plenary talks, panel discussions, and breakout sessions spanned a variety of topics, common themes emerged from all sessions. These themes became the recommendations in this report; they were arranged into broad categories which are the sections herein. Workshop participants were invited to edit and provide feedback on the report as it was drafted, to ensure that the arrangement of the themes and supporting details maintained fidelity with the workshop discussions.

Report Overview

As one participant said, "it's helpful to get everything in the soup" as we attempt to forecast what standards and metrics will be most beneficial to accelerating the bioeconomy. A plethora of opportunities, obstacles, and needs were discussed during the workshop, and many of those are summarized in this report. The report includes high level takeaways that are important to consider in metrics and standards development activities; the remainder of the report is divided into four sections of recommendations: (1) Metrics, (2) Short-Term Technical Standards, (3) Long-Term Technical Standards, and (4) Operational Best Practices. The recommendations address common needs and pain points related to engineering biology in industry. The technical standards recommendations are ordered from short-term, discrete developments, to long-term, continuous progress and are indicative of how long it will likely take to resolve and accomplish the recommendations, rather than a desired or prescribed timeline. These time spans are reflective of the perspectives of the workshop participants; achieving any of them sooner rather than later is likely to confer benefits to the bioeconomy. This workshop and report are meant to be a first step in standards and metrics development efforts in the Americas region; the needs identified from all three regions will be compared to find areas of agreement and consensus. The contents of this workshop and resulting report should be further explored by experts and stakeholders across industry to determine specific and actionable next steps.

High Level Takeaways

The context of the bioeconomy is important to consider as engineering biology standards and metrics are being proposed and developed. Some broad takeaways to keep in mind include the following:

- Redundant efforts to establish standards that already exist in adjacent industries should be avoided; existing standards should be modified to meet the needs of biomanufacturing, bioprocesses, and bioproducts, where applicable. Notably, novel capabilities that biological systems lend should be considered, so that established standards do not hinder the usage of these unique properties. Standards should be revisited regularly to confirm that they remain applicable to current technologies.
- As biomanufacturing processes move toward scale up, they may converge on, or interact with, established fields, such as chemicals or materials production. In these cases, standards and metrics should focus on the unique properties of the biological systems, while ensuring their interoperability with other technologies.
- Informal standards arising from established practices can hinder innovation if not addressed. For example, vendors of cell cultivation equipment frequently design to the needs of the largest market segments, and often support only a limited range of temperatures, agitation, gas control, and other functionalities. Meanwhile, few vendors cater to the distinct demands of non-model organisms, which represent a smaller, but important, market segment. Standards formed around

the most utilized biological systems should take care to not inhibit advancements in novel systems.

- Standards set for infrastructure are the most difficult to abide by, especially in an international context, due to high cost and limited flexibility. Application- and outcome-based standards could be more easily adopted around the world.
- Standards development, including the standards recommendations contained in this report, will be most beneficial to startups and small companies. Larger companies, with the resources to develop and maintain their own measurements and standards, will likely continue to keep their resources private for a competitive advantage and security reasons. Ultimately, establishing useful standards and metrics will support the development of diverse products and processes within a robust bioeconomy, as they can help to level the playing field between companies of varying sizes and resources.

Metrics

In this workshop and report, metrics are defined as *the measurements made towards assessing the (technical, economic, social, etc.) viability of a product or process*. The measurement needs that were identified can be binned into three categories: measurements within a process, metrics to compare processes, and biocontainment metrics. Some of the recommended metrics are for technologies to enable discrete measurements, while others may require the thoughtful combination of multiple measurements to assess a broader concept, such as sustainability or biocontainment.

Measurements Within a Process

Advancing measurement capabilities for biological processes will improve predictability and enable better decision making during scale up and production runs. There is an objective to not only measure the desired outcome (e.g., productivity, titer, yield), but also the contextual and environmental conditions that select for, or influence, the outcome (e.g., temperature, inhibitors, etc). In some countries, the cost of labor will be lower than the cost of instrumentation, so manual process measurements may be more economical and advantageous compared to in-line, automated measurements. Generally, the confidence in measurements, including accuracy and precision, should be improved for both manual and automated measurements. Enabling comparisons between manual and automated measurements will also be important as the field progresses globally.

The characteristics needed for metrology and measurement technologies to be adopted in industry are as follows:

- Low cost of instrumentation, adoption, and integration;
- High level of reliability, reproducibility, and ruggedness;
- Ability to interface with process control software;
- Flexibility to support different products or processes; and,
- Preferably continuous, in-line, and/or wireless.[3]

[3] Wireless instrumentation is a relatively new technology. There are security concerns that arise with the use of wireless systems in facilities that must be considered and addressed before implementation. The U.S. Cybersecurity and Infrastructure Security Agency (CISA) provides resources on this topic.

Specific process variables that are important to quantify and need improved measurement technologies include the following:

- Biomass or cell density. Currently, there is no commercial, in-line sensor for cell density. Optical density is used as a proxy for small-scale laboratory experiments, but is challenging for high cell density conditions, such as those found in bioreactors. Other measurement probes, such as spectroscopy- or radio-frequency-based probes, are used for large-scale applications. Continued innovation, benchmarking, and confidence-building around cell density quantification measurements are needed that embody the characteristics listed above, and can handle the complexities of cell culture media.
- Product, impurity, inhibitor, and feedstock concentrations. Being able to measure these values, especially in the fermentation broth, without having to perform a full purification, would decrease the time required to assess scale up and production runs in biomanufacturing.
- Internal conditions within a bioreactor. There is a general need for instrumentation that would allow for continuous monitoring and data collection for important process parameters, especially those that would enable feedback control loops for more automated runs. These instruments should be functional in a wide array of bioreactor designs.

Metrics to Compare Processes

Metrics that characterize the carbon footprint, (specifically biomass) ability, or other aspects of performance of products and processes within the bioeconomy, and of conventional, non-biological processes, would enable quantifiable comparisons between processes, as well as motivate innovation for new biological processes. Currently, evaluations such as life cycle analysis (LCA) attempt to capture some of these characteristics; however, variations in approaches, boundary selection, and data quality yield vastly different results. This highlights the need for standards and best practices in systemic and impact analyses (discussed below), in addition to metrics development.

Biocontainment Metrics

An open discussion in the field is what to measure to quantify biocontainment. Most industrial processes are designed for the organism to be physically contained in a structure, such as a bioreactor or other piece of equipment. Loss of such physical containment presents a risk of unintended release. However, there are some processes that necessitate deliberate release, where the organism is applied to a crop field or other uncontained environment. The levels of acceptable risk with regards to containment of modified organisms varies greatly across institutions, organizations, and governments. Even the language around biocontainment is highly subjective: the terms “outside of a structure,” “outside of containment,” and “deliberate release” are used across various U.S. government agencies to describe a loss or lack of biocontainment. Thus, the topic of biocontainment is a complex technical, semantic, policy, and social issue.

Metrics to quantify biocontainment are lacking and vary between countries. Biocontainment characteristics of an engineered organism might include the frequency of horizontal gene transfer (HGT), plasmid loss, distance found from the application area, and how long the product can exist outside of its intended location. Many of these characteristics are already components of risk assessments required by state and federal regulating agencies, such as the Environmental Protection Agency (EPA) and Department of Agriculture (USDA), in the U.S. The way these factors are measured is not uniform and

could potentially be standardized, contributing to streamlined research and development activities and assessment and regulatory processes. Furthermore, different metrics may be appropriate for different application contexts and environments. These metrics could also serve as the basis for international coordination.

In addition to the technical metrics and standards that are needed for the bioeconomy, stakeholders are not in agreement on the extent of biocontainment methods that should be required. Some want stringent biocontainment for every application and process; others want biocontainment proportionate to the risk associated with release of the organism or system. These are important policy and regulatory considerations that can be facilitated by common metrics for biocontainment.

Short-Term Technical Standards

There were a number of topics for which the most discrete and detailed discussions took place, suggesting that they are common and pressing considerations for individuals and companies. The clarity around these recommendations indicates that they can be developed in the short-term. These recommendations comprise standards surrounding bioeconomy lexicon and language, data, protocols, and automation and liquid handling.

Lexicon and Language

A standard lexicon should be developed to enable clear communication across the bioeconomy and with the public. There is currently a lack of public, consistent terminology across sectors to describe methods, tools, and resources;[4] even the term "bioeconomy" lacks a consistent definition for experts, policymakers, and the public. Standard language would facilitate greater understanding between technology transfer partners, and with regulatory bodies. NIST has recently begun to compile a Bioeconomy Lexicon (https://www.nist.gov/bioscience/nist-bioeconomy-lexicon), though it is still sparsely populated. The lexicon should include both technical and descriptive terms. Technically-detailed language is needed to specify operations. For example, the term "anaerobic" is used to describe organisms that can (facultative anaerobe) or must (obligate anaerobe) grow in the absence of oxygen. Additionally, processes can also be described as anaerobic, when oxygen is transiently or perpetually absent. Standardizing more detailed, functional descriptions would allow for better understanding of the organism and process in use.

The lexicon could serve as a helpful tool for the general public to better understand the bioeconomy. The term "GMO" (genetically modified organism), despite its ubiquitous use, is quite controversial in the public sphere. Though the scientific community has generated its own definitions, GMO remains a general use term without a specified definition across diverse audiences.[5] The GMO label could come with additional information about what the "M" (modifications) achieve, such as increased product yield, reduced impurities or byproduct generation, or decreased energy consumption. Perhaps standard definitions and increased transparency could help with public perception.

[4] The American Society for Testing and Materials (ASTM) provides E1705-15: Standard Terminology Relating to Biotechnology and E3072-22a: Standard Terminology for Industrial Biotechnology and Synthetic Biology.

[5] ASTM E3214-19: Standard Classification for Industrial Microorganisms provides definitions for engineered organisms, though GMO is not a broadly well-understood term.

In addition to the lexicon, the language of the bioeconomy should be developed in a broader sense. A bioeconomy ontology and schema could enable fluency and quicker understanding of the properties of what is described and how it relates to other entities. These language structures can support data interpretation and integration, and would be useful for artificial intelligence (AI) applications. This type of language development work can build off the shared lexicon, and will be an ongoing, longer-term effort.

Data

Technological advancements and scale up are underpinned by data exchange. There are no widely adopted standards for data sharing. Because individual organizations collect and curate data subjectively, it is challenging to understand and use data generated from an external source. Furthermore, without incentives and secure mechanisms to collect and share useful, compatible data, efforts in troubleshooting or scale up are often duplicated across organizations.

Standards around data structure, collection, sharing, and security will accelerate technology transfer and communal problem solving across the bioeconomy. Specifically, the following standards should be developed:

1. Data structure and format. A standard structure and format for data collection and sharing would enable better readability and more accurate understanding of the process or technology represented. Formats that are broadly accessible by a variety of software are preferred. These standards should specify: i) data labeling (e.g. column headers) as informed by the standard lexicon described above; ii) unit determination and descriptions; iii) mandatory and optional data bins or categories; and iv) how to indicate technical or biological replicates; among other aspects of data structure. These data standards can also help to ensure high quality data is being collected and shared. Additionally, data structure and format standards would be beneficial for the pooling of disparate data sets and databases, such as for multi-omic analyses. These standards would allow for the integration of genomics, metabolomics, proteomics, and other -omics data for a more complete understanding of an organism than is possible from the individual data sets alone. Associated metadata (described below) for disparate data sets will also be critical to their integration.
2. Metadata collection. Data without context of experimental conditions is meaningless, especially for bioprocesses and biological systems. Specific contextual metadata collection should be standardized, and include properties such as the host or source organism, pre-growth conditions, fermentor design, temperature, agitation, software used, and instrument type and settings.
3. Data sharing. Standard data sharing interfaces between organizations, such as equipment manufacturers, researchers, companies, scale up and production facilities, and government agencies, could enable smoother technology handoffs. Having these standards in place would allow startups to format their data for the interface from the outset. Furthermore, if the metadata standards are incorporated as well, there is a higher likelihood of successful process transfer during scale up. Easier data sharing through standardization can also accelerate innovation and solutions to common scale up, downstream processing, and other industrial challenges. Standards and best practices around sharing negative data would further strengthen collective problem-solving efforts and reduce resources being wasted on redundant, siloed efforts; beyond

standards, incentivization for negative data sharing is critical for its practice in research and industry. Data sharing comes with data security concerns; these are discussed below.

4. Data security. For data pooling initiatives to achieve industry-wide problem solving of common challenges, standards and best practices should be set to ensure protection of company intellectual property (IP) and identity, as well as to ensure the quality and integrity of the data. IP protection tools, such as one-year embargo periods on data and anonymization at the data sharing interface, can help to secure shared data. These measures should be balanced with data traceability.

Full context for process and experimental data, including negative data, will also enable better data curation for automation, AI, and machine learning (ML) applications. Currently, data sets are small and poorly annotated. Standardized formatting, metadata annotation, and sharing would allow small data sets to be pooled more effectively in order to leverage ML for advancements in the bioeconomy.

Data standards could have the added benefit of facilitating data sharing globally. Language and cultural differences can create barriers to successful data sharing, or at least slow the process. The data standards enumerated above would establish a shared baseline and ease international data transfer.

There is a question of who will implement these data standards to generate useful, controlled, open data sets that will be critical for many of the standards and supporting technical developments outlined in this report. Technical universities in the European Union contribute heavily to these types of data. However, the U.S. lacks the incentives, both in federal funding and in publishing options that are recognized for tenure, for academic researchers and institutions to serve the same role. Like standards, quality, open data sets have the ability to accelerate bioeconomy development, and could be especially useful for startups and small companies.

Protocols

To achieve production readiness at commercial levels, protocols must evolve for larger scales, and diverse equipment and facilities; often, this involves moving from human-run, bench-scale experiments to large-scale fermentors and purification steps with some level of automation. The accuracy of automated processes often depends on the quality and readability of the protocol used. Standardized protocol language and documentation should be established to streamline the transfer of processes from humans to machines, from laboratories to scale up and production facilities, and across geographies. The Laboratory Open Protocol Language (LabOP, https://bioprotocols.github.io/labop/) is a community-based initiative that publishes open specifications for laboratory protocols. Though LabOP and other open source tools are not used by large companies, who invest in in-house, secure systems, development of protocol standards could be very useful for startups and small companies. The existing protocol standards specified by LabOP and others should be considered when developing new or supplementary standards. The data standards described above will lend nicely to standard protocols, as information such as original equipment model, settings, conditions, and software could inform appropriate transfer to new equipment. Key aspects to incorporate into protocol standards include the following:

1. Quantitative descriptions of process steps to improve readability, reproducibility, and interoperability. Current common protocol steps say things like “mix well” and “centrifuge

briefly." These are difficult to transfer from human to machine, and often will be executed differently between humans as well. These steps should instead state clear, quantitative instructions, such as "mix by inverting 10 times" and "centrifuge at 10,000 g for 30 seconds," while including other requirements such as temperature, fixed angle rotor or swinging bucket specifications, among others. Other additional details that could be included are discussed below.

2. Additional and auxiliary information that is critical to the protocol, but not reflected in the steps. This can include:
    a. Equipment and instrument type, settings, and software, along with modifications that allow for the utilization of or transition between different tools. Making protocols instrument agnostic, if possible, is preferable.
    b. Precision requirements and tolerance of steps. Some steps have a much larger tolerance in accuracy and precision than others. A participant shared an anecdote of a functional, manual process that failed when it was automated. After troubleshooting, they determined that the automated process was better calibrated, and it began to work when the ~20% miscalibration of manual pipetting operations were accounted for.
    c. Underlying motivations and explanations for process conditions; these conditions should be specified as quantitatively as possible. For example, cold temperature requirements for a series of steps. Perhaps in a lab setting, using ice is the easiest way to keep components cold, but the underlying need could be to maintain a temperature of 4°C or below. In a different facility, there may be better ways to adhere to this temperature condition, so this auxiliary information could help to avoid unnecessary time and energy expenditures.
    d. Liquid class parameters and compatible materials (e.g. pipette tips) when using liquids with atypical characteristics.
3. A minimum requirement of necessary and sufficient information. While the above recommendations can help to specify necessary information in protocols, standards that prescribe over-specifying (unnecessary) information would be tedious and unlikely to be adopted. Giraldo et al. developed a guideline to achieve this for research protocols in the life sciences;[6] their work, and other protocol guideline efforts, could be adapted for industrial engineering biology protocols.

Developing rigorous protocols, including analytical protocols to monitor performance, could help to ensure a process is running correctly and ameliorate uncertainties when scaling a biological system. While broadly useful, these would be especially important for transfers between people and organizations.

Another initiative for protocol standardization is a community-supported effort to convert individualized operating procedures into standard operating procedures (SOPs). Building off the data and protocol standards above, AI could be implemented to generate SOPs from various organizations' protocols for common process steps. These SOPs could serve as a baseline from which to develop a protocol for a new process, or as a point of reference to improve existing protocols. Whether the SOPs should be

[6] Giraldo, O., Garcia, A. & Corcho, O. A guideline for reporting experimental protocols in life sciences. *PeerJ* **6**, e4795 (2018).

generated for the most common equipment, or be equipment-agnostic and solely outcome based, is a question that needs to be considered by the community.

Automation and Liquid Handling

Automation can improve process and batch reproducibility and enable data collection more quickly and in larger quantities. However, automation is currently expensive and can limit flexibility of process development. As technology advances to incorporate more automation in research, scale up, and commercial processes, advancements in modularity and interoperability are needed to piece together the steps of a full process. Standards for how robots and automated steps should interact will improve their implementation.

Technical advancements in liquid handling automation, specifically, will be readily integrated into process development activities in the short-term. Reference materials and benchmarking assays are needed for different liquid classes in order to accurately and reproducibly handle liquids with different physical properties (such as viscosity, vapor pressure, and surface tension) and to identify material compatibility (for example, for pipette tips and labware). Standardization around liquid handling representation could help users select the appropriate instrument and materials for their application, and should include:

- Working volume range;
- Precision and accuracy by working volume;
- Types of liquids or liquid classes that can be handled;
- Precision and accuracy by liquid type;
- Pipette tips supported by the liquid handling system;
- Cost of equipment; and,
- Compatible systems, softwares, and APIs for interfacing with the liquid handler.

While these automation and liquid handling recommendations will accelerate industrial research and scale up, translating these standards, reference materials, and benchmarks to large-scale production facilities will be crucial for these developments to be impactful at the commercial scale.

Long-Term Technical Standards

The recommendations in this section are more complex, and may build upon the metrics and short-term standards discussed above. These recommendations may require technical innovations, incentivization, or other co-requisites to achieve the standardization described. Thus, they will likely take longer to develop and will require more information gathering, especially as engineering biology technologies advance. As such, these standards, for process representations and modeling, feedstock diversity and specifications, and software, will likely have to be developed and revised continuously over time.

Process Representations and Modeling

Standard representations of processes can enable process readability, communication, and transfer. Block flow diagrams (BFDs), process flow diagrams (PFDs), and piping and instrumentation diagrams (P&IDs) are common process representations used by engineers in the chemicals industry and in some large biotech companies with their own production facilities. These representations are critical to reliably and efficiently describe processes and production facilities, and are necessary to perform an advanced

techno-economic analysis (TEA). Thus, they should be adapted and implemented for processes containing biological unit operations, and adopted by companies across the bioeconomy.

A standard process model for biological processes would be beneficial for process development and scale up. There are a number of established modeling programs for chemical processes; however, they are not all accessible on different computer operating systems, which limits their use. While biomanufacturing packages and programs have been developed,[7] they lack relevant physical property data sets to fully model biological processes, which are often aqueous and contain salts, proteins, and other impurities. These data sets are needed to support the computational methods underpinning the modeling programs. Additionally, developing bioprocess packages for existing, widely-used programs would allow bioprocess developers to leverage existing expertise in modeling software.

For both recommendations on process representations and modeling, these standards would allow interoperability with existing industries and competing platforms, as well as better depictions of processes that comprise both biological and chemical unit operations.

Feedstock Diversity and Specifications

There is potential in bio-based technologies – and the desire from companies – to use complex, biomass feedstocks in biomanufacturing processes. While there are technical advancements needed in this area, standards and metrics development can accelerate the incorporation of these complex feedstocks, including industrial, agricultural, and municipal wastes, into processes in shorter time scales. Efforts should be made to continue developing and publicizing existing resources[8] to assess the availability and commercial viability of biomass feedstocks. Further standards developments should include the following:

- Standard characterization and specification sheets for feedstocks, including details such as:
    - Identity and source;
    - Quantity, such as total volume, mass, etc.;
    - Composition, including, for example, carbon content, lignin content, types of components (e.g., hydrocarbons, sugars, sulfur compounds), components ratios, etc.;
    - Energy density;
    - Seasonality and/or long-term availability;[9]
    - Storage conditions;
    - Preprocessing conditions, if applicable; and,
    - A measure of sustainability or circularity, e.g. carbon index.

[7] Shanklin, T., Roper, K., Yegneswaran, P. K. & Marten, M. R. Selection of bioprocess simulation software for industrial applications. *Biotechnol. Bioeng.* **72**, 483–489 (2001).

[8] For example, Ecostrat (https://ecostrat.com/standards/) has developed a risk rating system for biomass feedstocks which considers availability, inventory, and quality among other factors; the Bioenergy Feedstock Library at Idaho National Laboratory (https://bioenergylibrary.inl.gov/Home/Home.aspx) is a database and sample repository for a variety of bioenergy feedstocks.

[9] The real, physical availability of a feedstock should be reported, rather than the theoretical. A Feedstock Readiness Level evaluation (https://www.caafi.org/tools/Feedstock_Readiness_Level.html) was developed for sustainable aviation fuels to assess the legitimacy of a feedstock by determining if there is a complete supply chain in place.

- A map of potential feedstock resources, that are specifically suitable for bioprocessing and include as much of the above characterization data as possible.[10]
- Third party composition certifications. This should include a bilateral impurities-reporting function, wherein feedstock users can report undocumented impurities or compositional elements while protecting their identity and IP.

The characterization data sets recommended above for feedstocks are also currently lacking for critical components and chemicals used in bioprocess development. The distinct characteristics and specifications needed to describe each component may vary. For example, physical property data in relevant solvents is highly important for process chemicals. Standardizing these types of characterization data for process inputs would help to bridge information gaps that currently exist during process development.

Software

Standards are needed to improve the usability of software for data collection, processing, and annotation, and to improve software interfacing with equipment, automation, and more. Currently, software is written for the most common use cases, with few options for users who are developing a unique application. Additionally, software developers often do not share the application programming interface (API), which connects the software to the instrument. Open source APIs and other process software would make these tools more useful for broader use-cases. Establishing standards or best practices for software customization and interoperability could expedite process development. Many companies perform tasks in-house that are not interoperable across platforms, such as developing reference databases or writing software to support flexible workflows. A coordinated approach to this could consolidate individual, redundant efforts.[11]

Operational Best Practices

Codifying best practices around common, challenging activities in the bioeconomy can further accelerate development and complement the technical standards described above. These best practices can also enable smoother interactions between entities, for example a startup company and a regulatory agency, as the expectations from each party will already be outlined. Best practices should be developed for scale up, pilot facility and contract manufacturing organization (CMO) selection, regulatory processes, and systemic and impact analyses of products and processes. Finally, there is a note about training and education around standards.

Scale Up Toolkit

Individual biomanufacturing startups often encounter similar obstacles as they mature their technologies. Rather than a trial-and-error approach, or through hiring a consultant to advise a company on rudimentary considerations, a baseline toolkit should be developed to accelerate the scale up process.

---

[10] There are existing efforts to map biomass feedstocks through the National Renewable Energy Laboratory (https://www.nrel.gov/gis/biomass.html) and the BioP2P Network (https://biomanufacturing.net/directory/?layers=feedstocks).

[11] There are existing efforts to develop a common language to program liquid handlers (PyLabRobot: https://github.com/PyLabRobot/pylabrobot) and open source instrument protocols (OpenTrons: https://opentrons.com/), both using Python.

The toolkit should emphasize prudent, systems-level examination of the entire process being developed, and encourage startups to consider factors such as the following:

- Employing good TEA modeling early (and often) can direct companies towards scale up problems that are the most impactful to address. Business risk, including market dynamics, market uptake, competitor reactions, government policy or incentive changes, and interest rates, must be mitigated alongside technical risks, to ensure a product will be successful on the market.
- Whole-process, systems thinking early in process development helps to ensure that resources are utilized appropriately. For example, downstream processing can present a significant cost, so the technical and financial planning of the entire process should begin in early stages.
- "The friend of every process is good process statistics," stated one participant. Quality process data can lend to statistics for process control math; the resulting process control is a powerful tool to inform scale up and to monitor and maintain performance.

These considerations leverage many of the metrics and standards recommended in this report, and can help to increase the likelihood of success in scale up and commercialization efforts.

Given that each company may have a unique application, it would be best to develop a community-verified **Scale Up Checklist** of considerations for different application sectors. Some examples for biomanufacturing include:

- Determining process constraints that may prevent or inhibit scale up. For example, if a biomanufacturing process requires antibiotics; this is a "no-go" for scale up, due to cost and potential development of resistance.
- Analyzing the target feedstock for scalability to the desired level.
- Understanding licensing, IP, and patent arrangements of the biological components acquired from an external entity. Genetic elements, plasmids, and strains from commercial kits or depositories often prohibit commercial use or further distribution from the original purchaser.

Though these points may seem simple, they are common stumbling blocks, and substantial resources could be saved if a scale up checklist could be consulted. The checklist should also include how to choose and prepare to work with scale up and production facilities; the Data Sharing standards (above) and Pilot Facility and CMO Selection best practices (below) would lend well to this. Finally, the toolkit should include guidelines on how to approach the regulatory process. Standardizing the regulatory workflow will be an important co-requisite to this resource and is discussed more below.

<u>Pilot Facility and CMO Selection</u>

Choosing a suitable pilot facility to scale up a new process, and subsequently a contract manufacturing organization (CMO) for production, are important but challenging steps towards commercialization. Common practice is for companies to continue to use the same scale up and production facilities due to familiarity and previous successful partnerships. However, this leaves a void for new startups or drastically different endeavors from a company's past technologies. Establishing a standardized, digital representation of the capabilities offered by pilot facilities and CMOs would aid companies in selecting the most appropriate partners for their application; it could also help the transition from pilot scale (~20,000 L) to CMO scale (150,000-1,000,000 L). Additionally, clear representations of these facilities would allow for a better understanding of how an existing process would need to be adjusted to be transferred to the next step in scale up. Some important attributes to include in the representation are:

- Available on-site production capacities;
- Unit operations, equipment, instrumentation, and software lists;
- Output data format;
- Feedstocks and media formulation specifications;
- Ambient environmental conditions, including temperature, humidity, elevation, etc.;
- Workforce and personnel;
- Demonstrated success with different processes;
- Schedule and availability of capacity;
- Physical and cyber security practices;
- Cost structure;
- QA/QC and troubleshooting support; and,
- Any prerequisites for working with the facility.

These types of representations could be integrated into a public resource platform such as www.biomanufacturing.net to enable comparison between facilities.

Pilot-scale unit operations in the U.S. are often split between multiple locations, though the best way to validate a process is by running it continuously at one site. How to best validate performance at pilot facilities and CMOs warrants further investigation and development. Developing robust validation methods for scale up facilities, in various configurations, could increase the likelihood of success when the process is transferred to a continuous, production-scale CMO. Additional benchmarking or accreditation could help to improve the appraisal process while selecting pilot facilities and CMOs; how to benchmark a bioprocess, or develop a benchmarking standard, is an open question.

Regulatory Pathways and Assessments

Navigating the regulatory landscape is a major hurdle that companies must overcome. For some startups, a failed effort to acquire regulatory approval can deplete their investment funding. A clear, streamlined process for getting regulatory approval, backed by standards and technical assessments, is needed to accelerate commercialization. The following hurdles must be addressed to achieve this:

- It is unclear what is required to approach agencies for regulatory approval.
- There isn't a clear metric, standard, or pathway to getting regulatory approval for novel organisms. Different regulatory agencies have different exempt or non-exempt organisms lists with varying degrees of clarity on how to add to them.[12]
- Startups are sometimes not aware of Confidential Business Information protections to their IP when verifying their strains or other novel biological systems.
- Bioidentical products to products that are already approved are not recognized if they are made from a different process or source organism. For example, there was an anecdote from the workshop of a bioidentical, dietary collagen product, made by fermentation, that was denied regulatory approval because the source of the collagen protein was an organism that is not customarily consumed in the U.S. There is no guidance on how to assess or regulate trace differences between bioidentical products. A standard approach in regulatory pathways that

[12] In the U.S., the Food and Drug Administration uses the Generally Recognized as Safe (GRAS) list, the Environmental Protection Agency uses the Tier 1 Exempt organisms list, and the Department of Agriculture has a negative list of organisms presenting plant and animal pathogenicity.

acknowledges approved bioidentical and biosimilar products will help companies capitalize on the immense diversity of bio-based products that engineering biology can bring to market.

Systemic and Impact Analyses

Standard assessment approaches, such as life-cycle analysis (LCA) and techno-economic analysis (TEA), need to be established that take into account the systemic impact of a product or process. For example, considerations such as carbon intensity of processes, renewability and long-term availability of feedstocks, or circularity of materials are not commonly considered in these types of analyses. Systemic analyses developed for bio-products and -processes should build on existing standards, such as International Organization for Standardization (ISO) 14001, which helps to manage environmental impacts, and the ISO 14040 series on LCA frameworks.[13] Other factors such as social metrics – for example, economic justice, community partnerships, or sourcing local materials – could be incorporated as well; development of social metrics should be done in collaboration with experts trained in relevant fields. Working from common boundary definitions and validated, quality data sets are necessary to make comparisons between different systems. This type of standard analysis would be a beneficial tool to compare bioproducts and bioprocesses to existing chemical alternatives as well. Using these standard analyses, a database of results for existing products and ingredients on the market, and for different unit operations, could facilitate streamlined comparisons and easier process evaluations.

Training

The need for training and education on standards and metrics came up repeatedly throughout the workshop. Currently, there is no education on industrial standards in biotechnology- or biomanufacturing-related training programs. This is not a standards-based recommendation, rather a note that training *in* standards can enable broad understanding and implementation of standards across the bioeconomy. Industry standards and metrics can be disseminated through early-, mid-, or late-career training. Stand-alone teaching material could be developed around standards that support the bioeconomy. Alternatively, education on standards could be integrated into relevant existing curricula. For example, standardized data formats, metadata collection, and protocol language could be integrated into college lab courses and the scientific discovery process; data security and facilities representation standards would be apt for technical workforce training programs.

Conclusion

This report makes recommendations on the metrics and standards developments needed to accelerate the bioeconomy in the Americas, as informed by the technical experts that attended the Americas Workshop. The recommendations herein are intended to make technology transfer, scale up, and commercialization more efficient for engineering biology-based products and processes.

This project is an initial step to compile the technical standards and metrics needs from regional technical experts, and represents a snapshot of the bioeconomy. The long-term trajectory of standards and metrics development must continue in consultation with industry experts and adapt to technological advancements. Even the initial standards that are developed based on recommendations from this

[13] ISO 14001:2015 specifies environmental management systems (https://www.iso.org/standard/60857.html); ISO 14040:2006 describes LCA frameworks (https://www.iso.org/standard/37456.html).

project should incorporate more detailed input from a broad range of industry, government, policy, and research stakeholders. These continued collaborations are crucial to ensure that resulting standards and metrics will be useful and widely adopted to accelerate the development of a global bioeconomy.