Vision & Scope

Contents

Problem statement

In research, tracking and managing experimental data is crucial for comprehensively analysing results and making decisions within the experiments. Current methods for tracking and recording experimental data in BioSphere Innovations Ltd. laboratories are not efficient. The information is stored in a mix of analog and disconnected digital formats - on paper, in isolated files and a digital sample registry. The fragmentation of data makes it difficult to trace experimental progress, or analyze outcomes.

A unified system is needed to solve the problem with data management.

Vision statement

Through Laboratory Experiment Management System (LEMS) researchers can securely upload, store, organize and retrieve data about experiments done in BioSphere Innovations Ltd. laboratories.

Unlike current data storing systems, LEMS will facilitate data manipulations. An API interface will support transfering the data in dashboards, reports, and broader analytics, which is helpful to spot trends and opportunities within the data.

The system will seamlessly integrate into workflows and encourage decision making and collaboration in the company.

Alternatives and Competition

The targeted system will compete with some strong laboratory experiment management solutions currently available in the market. Available alternatives for stakeholders include purchasing ready-made competitor software available in the market, developing a custom in-house solution, or maintaining existing analog/digital fragmented systems. This section addresses some known competitor products and their strengths/weaknesses as perceived by users. This analysis holds strategic importance for product positioning and identifying differentiating features.

Benchling

A powerful platform used in molecular biology and genetics studies. Its UI and UX, along with team collaboration features, make a difference in the market. However, it has limitations in chemical experiment management and its high cost restricts accessibility.

LabArchives

A solution known for its success in digitalizing experiment logs. Being cloud-based allows users to access it from anywhere. However, the system's customizability is limited. It is known that it cannot meet technological expectations such as advanced data analysis and API integration.

LabWare

A product known for having an extensive feature set. This allows it to be integrated and managed across different lab processes. However, the system's complexity on the UX side keeps the learning process lengthy. Additionally, its installation cost is high.

Dotmatics

It is preferred due to its powerful data analysis tools. However, it has limited usage areas in biological experiment processes and does not meet modern design expectations in terms of user interface.

LabCollector

It has a modular structure. It is easy to install and fast to operate. However, it may be inadequate in terms of performance and data management in large-scale projects; additionally, the lack of advanced reporting tools makes detailed analyses at the decision-maker level difficult.

Business requirements/goals/success criteria

Business Requirements

Functional Requirements:

- **Data Management:** The system should allow researchers to register new samples, assign them to experiments, and thack their lifecycle. The system must centrally record and track all experimental data (related to chemical and biological samples), replacing existing fragmented analog and digital data entry methods.

- **Sample-Experiment Association:** The system must ensure seamless association and tracking of specific samples with their corresponding experiments.
- **Workflow Integration:** The system must support a structured and traceable experiment lifecycle workflow that aligns with real-world laboratory practices. The workflow includes:
  - **Initiation**: Experiments are initiated based on schemas provided by designated personnel (e.g. Team Lead).
  - **Preparation**: Researchers select target species and chemicals from internal databases and prepare for treatment.
  - **Treatment & Sample Creation**: Samples are treated with various chemical concentrations across multiple time points. New samples are created and must be registerable directly within LEMS.
  - **Data Collection**: Researchers extract biological material, perform analysis (e.g., LC-MS), and attach result files to experiments.
  - **QA Review & Completion**: Once all data is collected and validated, and all samples are registered, the QA team reviews the experiment. Upon approval, the experiment is marked as completed.
  - **Additional Experiment States**:
    - **Paused**: Temporarily locked to prevent changes.
    - **Reassigned**: Can be transferred to another researcher.
    - **Failed**: If errors (e.g., incorrect chemical usage) occur.
- **Progress Tracking:** The system must facilitate clear tracking of experimental progress (e.g., experiment status, phase tracking).
- **Collaboration Support:** LEMS must provide shared access to experimental data among global research teams, enhancing collaboration.
- **Data Analysis and Insights:** The system must enable comprehensive analysis of experimental results to derive meaningful insights (e.g., report generation, filtering, comparison capabilities).
- **API:** The system should provide a read-only API designed specifically to enable secure, structured data retrieval for reporting and analytics systems. This API will allow external platforms (e.g., BI dashboards, Tableu etc.) to extract experimental records, sample metadata, and workflow statuses.

    - Only authorized systems with valid access tokens can access the API.
    - API responses must be structured in standardized formats (e.g., JSON, CSV, XML) to support automated parsing and integration.
    - The API must support filtering such as by date, user, experiment ID to enable targeted data extraction
    - All API requests must be logged in the system audit trail, including with the information of requester, timestamp and dataset accessed.
    - Export endpoints must be limited with experiment data, sample details, and activity logs. No write operations are allowed via API.
    - System may offer bulk data export for offline reporting for limited roles such as Department Manager, Admin and QA roles.
- **Reporting and Analytics Integration:** The system must support broader reporting needs, including dashboards, reports, and high-level analytics. Customized dashboards should reflect user roles and responsibilities.
- **Protocol Management (Future Scope):** The system must include a centralized protocol management module that allows storing, updating, and distributing standardized experimental procedures, ensuring consistency across labs.
- **Audit Trail:** The system must have a comprehensive audit trail to ensure traceability,accountability and compliance with regulatory standards.
    - All user actions which are create, update, and access data within the system must be auto logged.
    - Logged events must include timestamp, user id, role, action name (updated sample, accessed report etc), affected experiment id, or sample id (basically entity id).
    - Audit log must support filtering by user, date range, action type and entity id.
    - Audit log must support export in PDF and CSV formats.
    - Audit logs must be read-only to ensure data integrity.
    - Only authorized roles (such as QA Offices and Admins) can view and export audit logs.

Non-Functional Requirements

- **Security:** Must meet strict security standards of BioSphere Innovations Ltd. for processing sensitive chemical and biological research data.
- **Usability:** Should provide a user-friendly interface and experience, ensuring high adoption rates among researchers and laboratory teams.
- **Integration:** The system must integrate with the existing digital sample recording system and other legacy systems, support different time zones and potential language factors, and comply with regulatory requirements in the UK, EU (Lithuania), and India.
- **Data Quality:** The system should reduce manual data entry errors by 80% and ensure 99% data consistency with reporting systems, thereby maximizing data quality and reliability.
- **Sustainability & Maintenance:** The system must be suitable for long-term maintenance and development while sustainably meeting performance goals.

Goals

- **Modernizing Data Management:** Transitioning from fragmented, manual record-keeping to a unified digital platform.
- **Enhancing Data Quality:** Ensuring accuracy and traceability in experimental data to minimize errors.
- **Strengthening Global Collaboration:** Enabling real-time data access and sharing across international teams.
- **Accelerating Decision-Making:** Supporting rapid trend analysis and strategic decisions with integrated data.
- **Optimizing Workflows and Security:** Streamlining laboratory processes while ensuring data integrity and security.
- **Ensuring Standardization Across Labs:** Providing a centralized protocol management system to enforce consistent experimental procedures across international research sites.

Success Criteria

Performance and Functionality

- Experiment data must be retrievable via the API within a maximum of 5 seconds from user entry.
- The system must ensure 99% data integrity for experiment records.
- At least 90% of users should log in regularly and actively use the system.

Data Security & Compliance

- The LEMS platform must be 100% compliant with General Data Protection Regulation(GDPR) and relevant regional data protection regulations.
- Multi-factor authentication (MFA) must be implemented to prevent unauthorized access.
- The risk of data loss must be reduced to zero, with a full system backup performed every 24 hours.

User Experience and Process Optimization

- User satisfaction should exceed 85% within the first 6 months.
- The training process should be completed within 3 hours, enabling new users to quickly adapt to the system interface.
- The time required for data sharing between laboratories should be reduced by 50% compared to current processes.

Continuous Improvement and Risk Management

- System downtime must be managed to remain under 5 hours per year.
- The feedback mechanism should evaluate user requests within two weeks and implement appropriate changes in the system.
- New system updates must be delivered with minimal disruption to user operations.

Constraints

Existing Infrastructure Integration: LEMS should integrate with the existing digital sample registration system and potentially other legacy systems, avoiding a full replacement approach as much as possible.

Global Operations and Time Zones: The system design must account for the geographic distribution of laboratories in the UK, Lithuania, and India, considering different time zones and potential language factors.

Regulatory Diversity: LEMS must comply with potentially different legal requirements regarding data privacy and scientific data integrity in the UK, the EU (Lithuania), and India.

Data Migration: Existing historical experimental data (both analog and digital) should be assessed for migration to the new system, which can be a complex and time-consuming process.

Resource Availability: The project timeline and scope will be influenced by the availability of BioSphere Innovations Ltd.'s internal personnel and subject matter experts (SMEs).

Security Requirements: The system must meet BioSphere Innovations Ltd.'s strict security standards for handling sensitive chemical and biological research data.

Budget Constraints: The project must be delivered within a defined budget, which will impact technology choices and implementation scope.

Data Storage Limitations: Storage systems must accommodate not only structured data but also large file uploads (e.g., images, PDFs), which may require scalable storage solutions.

Protocol Migration Complexity: Migrating and digitizing legacy protocol documentation may require SME validation and structured document management design.

Product Overview

Business Rules

- **Status Flow in Experiments**
  - An experiment must have a lifecycle such as: *Planned → In Progress → Completed (it could be successful, failed or disabled)*
    - Every action should be logged with timestamps.
    - Roles such as the author of the experiment, *Team Lead, QA managers* can mark an experiment Completed.
- **Data Validation**
  - Fields such as sample ID, experiment ID, and such fields must be mandatory to save.
  - Some parameter values must fall within the accepted lab-specific range or require justification.
- **File Upload Criteria**
  - Only specific file types are allowed (.pdf, .csv, .xlsx, .jpg).
  - Each file must be associated with either an experiment, a sample, or a protocol entry.
- **User Permissions**
  - Users must act within their predefined roles. For example:
    - A “Technician” cannot edit experiment protocols.
    - A “Data Analyst” cannot register new samples.
- **API Security and Storage**
  - External systems can retrieve experiment data only through authenticated API calls using an access token.
  - Each API request is logged with metadata for audit purposes.

- **Backup System**
  - Daily system backup is mandatory.
  - Data must be retained for a minimum of 5 years in accordance with GDPR and scientific regulatory standards.

Assumptions and Dependencies

1. The majority of users (~85%) will be accessing the system using modern web browsers (e.g., Chrome, Firefox, Safari). The remaining ~15% may use older supported browsers, alternative browsers (e.g., Opera, Edge), or interact with the system via specific integrations or potential mobile/desktop clients in the future.Researchers interacting with the system will have basic digital literacy and receive training during onboarding.
2. Development and testing efforts will primarily focus on the majority group (~85%) to ensure optimal performance and usability, while ensuring basic functionality remains accessible for the remaining users.
3. Laboratories will continue using the existing digital sample registry until full LEMS integration is rolled out.
4. Regulatory compliance requirements (GDPR, local scientific data laws) will remain unchanged during deployment.
5. Stakeholders will be available for validation of experimental workflows and protocol standardization before MVP release.
6. The infrastructure (network speed, storage, compute) at each lab is sufficient to support centralized cloud-based architecture.

Product Features

The high-level capabilities of the system are listed in here. It is designed to provide essential services to users according to various roles.

Sample & Result Tracking

The system should enable researchers to accurately track chemical and biological samples throughout their lifecycle from receive to disposal and correlate them with experimental results. This includes the ability to record detailed information about each sample and link them to specific experiments and their outcomes.

Experiment Logs System

Current digital sample registry will be integrated in the system to support experiment logs. Besides, ability to create new samples and experiments will be essential part of the system

Role Based Access

The system maintains security and control via role-based access. Users are allowed to access, view or modify relevant data and features according to their permissions or roles. This ensures data security and integrity.

Dashboards

The system should provide customized dashboards for specific roles. Each stakeholder has immediate access to the most relevant data and metrics for their roles. Dashboards will support real-time visibility and facilitate quick decision-making.

- **Researchers** will view experiment progress, assigned sample statuses, and pending tasks related to their active experiments.
- **Lab Managers** will monitor workflow bottlenecks, lab-wide sample throughput, team workload, and experiment distribution.
- **QA Officers** will access compliance-related metrics such as audit logs, validation statuses, and protocol usage consistency.
- **Administrators** will see system usage metrics, login statistics, API activity, and overall operational health indicators.

Each dashboard will display KPIs, summaries, and alerts relevant to the user's scope of responsibility, enabling efficient tracking of both scientific and operational workflows.

Experiment Workflow

The system must implement a clearly defined experiment lifecycle with specific states and transitions:

- **Lifecycle States**:
  - *Planned*: Created by the root designer or team lead.
  - *In Progress*: Researcher begins chemical treatment and data collection.
  - *Paused*: Temporarily disabled for review or due to blocking conditions.
  - *Reassigned*: Can be handed off to another researcher.
  - *Failed*: Marked as unsuccessful due to procedural or chemical errors.
  - *Completed*: All samples registered, data uploaded, QA-reviewed.
- **System Capabilities**:
  - Experiments must be linkable to samples, species, and chemical data.
  - Integration with the legacy sample registry must allow fetching existing sample data and registering new ones from within LEMS.

- Researchers should be able to attach analysis results and supporting files directly to experiments.
- Only authorized roles (e.g., QA Officer) should have the ability to finalize experiments.

These capabilities will ensure that all experiment-related activities are digitally traceable, align with real lab processes, and comply with regulatory and internal review requirements.

API Data Exports

The system should provide APIs which are secure and efficient extraction of experimental data for external systems. It may help the customer for reporting and more detailed analysis. Also system should allow to export as PDF,CSV and Excel formats for recorded experimental data, reports and logs. This feature facilitates offline analysis, and sharing.

Centralized Protocol Management System (Not Included to MVP)

The system should have a centralized section to store, manage and distribute the standardized experimental protocols. This ensures experimental procedure consistency across labs in different countries and encourages the best practice implementation.

Audit Trail

The system should capture all critical events (e.g., data creation, update, data access). Each log entry must include:

- Timestamp
- User ID and Role
- Action Type
- Entity Affected

The audit trail must be filterable (by user, date, entity), exportable (CSV/PDF), and visible only to QA Officers and Admins to support regulatory compliance (e.g., GDPR, GLP).

File Upload and Management

The system should provide a secure file upload mechanism for users to upload files like analysis results, images, legal documents, supporting documents, experiments, samples etc. This ensures all the related information is consolidated in one place.