FINAL REPORT

Project title	Laboratory Notebook and Database (LabNDB)
Project leader	Prof. Markus Krüger
Project team members	Jakob Harden, Lukas Briendl, Matthias Lenz, Mathias Eisner, Barbara Schmid
Reporting period	01.07.2021 – 31.12.2021
Institute	Institute of Technology and Testing of Construction Materials
Authors	Jakob Harden

1. Goals and results

We managed to achieve all of the proposed objectives. Some of them more elaborate and richer than expected, some of them in the form of a working prototype (ready and open for further development).

Table 1: Proposed objectives and achieved results, compared

No.	Proposed objectives	Achieved results
1	workflows to gather data from laboratory processes	data structure (backend) is well defined, user-interface and administrative interface finished, basic set of laboratory tests integrated, feature allows for simple integration of additional/new laboratory tests
2	sample and material database	data structure (backend) is well defined, user-interface and administrative interface finished, feature is open for future enhancements
3	role based access permissions to the data	feature implemented and tested, predefined configuration sets for different use-cases need to be implemented
4	data export to CyVerse/iRODS (to support the research and publication process)	finished gathering required informations for the implementation, integrated into solution, initial tests performed
5	IT-infrastructure, deployment of the solution	server setup completed, solution deployed in 3 different stages (testing, staging, production), solution is available inside the network of TU-Graz, backup of database and uploaded files is configured and working, test version supplied for FELMI and IMAT
6	data analysis tools	Integrated in test data tables in the sense of automatically calculated/updated fields for test results

Highlights:

We could manage to create the backend data structure during the previous reporting period in a way that allowed us to continue the development in this reporting period without any major changes or the need for a redesign. This saved us a lot of time and gave rise to some additional features and details we did not plan to implement in the beginning (e.g. device and section booking calendar).

The interest of other stakeholders in our solution (FELMI, IMAT).

Last but not least the self-commitment of the project team members to this project: B. Schmid, M. Eisner, P. Knabl, S. Sherifova, J. Harden. The success of this project is mainly owed to this circumstance and allowed us to achieve more results than we expected.

2. Work packages, milestones and current progress

2.1 Overview tables

Table 1: Work packages

WP No.	Work package title	Stage of completion (in %)	Scheduled date		Current date	
			Start	End	Start	End
1	Start phase	100	18.01.21	26.02.21	13.01.21	31.03.21
2	Implementation phase I	100	01.03.21	28.05.21	01.04.21	30.06.21
3	Implementation phase II	100	31.05.21	27.08.21	01.07.21	03.09.21
4	Final implementation phase	100	30.08.21	02.10.21	06.09.21	29.10.21
5	Final project phase	100	05.10.21	17.12.21	02.11.21	17.12.21

Table 2: Milestones

Milestone No.	Milestone Title	Scheduled date	Current date	Milestone achieved on
1	Project start	15.01.21	13.01.21	13.01.21
2	Implementation start	01.03.21	01.04.21	01.04.21
3	Evaluation of implementation phase I	31.05.21	01.07.21	02.07.21
4	Evaluation of implementation phase II	30.08.21	03.09.21	15.10.21
5	Evaluation of final implementation phase	05.10.21	26.11.21	26.11.21
6	Project end	17.12.21	17.12.21	17.12.21

2.2 Description of the work carried out during the reporting period

Work package: Implementation phase II

In this phase the major task was to implement the user-interface for all required features. Some enhancements to the data structure (database) were made as we needed them for additional features. We also enriched the solution with some features we did not plan in the beginning.

Work carried out during this phase:

- Dashboards. To lower the entrance level for users new to this solution we organized the access to all features in dashboards. They allow for a simple and quick overview and easy access.
- Context help. To enable the user for fast and simple access to information how certain features are to be used a context related help was added (display help text along with the feature). This should avoid the necessity of reading user manuals stored elsewhere. By now a working prototype is implemented that is enhanced in the ongoing development.
- Administrative interface. Along with all features explained below an administrative user-interface was implemented. It allows to manage the data of the entire database. Changes and modifications can be made that are not intended to be handled on the web frontend by the user and are therefore hidden away from the user-interface. This feature will be used by the database manager (a person with elevated skills in data management and database structures).
- Background features. In several places additional information is needed (e.g. the supplier who delivered a material, funders, customers, projects, user profiles, ...). This additional features are managed in the administrative user-interface only and deliver support all other features of the solution.
- Authentication and authorization. The first place to get access to the solution is the login site. Login credentials are provided by the database manager after user identification. The user profile is used

to identify data created by certain users. It also allows for role based authorization and to grant certain database access rights (insert, modify and delete data).

- Process management. The process management feature is the central point to gather and to handle laboratory measurement data.
  - A process holds the following information: which person started or ended the process; date and time when the process was started and ended; the related project; whether it is an accredited process or not (allows for accredited test settings); all laboratory tests performed in the course of this process (process activities); context related help
  - Each process activity consists of: the sample used to perform this laboratory test; a set of measurement data; the person who performed this activity; date and time when the laboratory test was performed; related documents, images, rawdata (can be uploaded to the server); context related help
  - Recurring processes (including a predefined set of process activities) can be created from process templates. This allows for simply re-use of common process settings.
- Material management. This feature provides for the management of all consumable materials used in the laboratory. Along with each material the following informations are stored: inventory (where it is coming from, supplier information); storage (know where your materials are, including a storage history); general material properties; type specific material properties (different properties for solid, liquid, aeriform, etc. materials, 'modular data structure'); related documents and images (can be uploaded to the server); material usage statistics (storage amount, how much material is left)
- Sample management. Samples are the very basis for all laboratory tests. Samples can also be divided into subsamples. Each sample holds the following informations: sample source (from supplier, customer, funder, consumable material, material mix, 'modular data structure'); sample origin (artificial or biological, 'modular data structure'); storage (know where your samples are, including a storage history); history of laboratory tests performed on this sample (a brief summary of the test results is shown); related documents and images (can be uploaded to the server); related parent sample and related subsamples are listed; context related help
- Laboratory facility management. This management tool is needed as a basis for material and sample storage and the devices used for laboratory tests (process activities). It has it's own dashboard that allows for simple access to: laboratory areas (a set of rooms); laboratory rooms; laboratory room sections (workbenches); test devices; booking events for room sections and test devices
- Export to Cyverse/iRODS. The export facility allows to export (almost) all tables in the database to various file formats (e.g. CVS, JSON, netCDF4, ...). Once the data file and the metadata (DataCite standard) is created this files can be transferred to a predefined CyVerse/iRODS server (using iCommands). The metadata associated to a data file can be updated before the export is done. Account credentials of a CyVerse server have to be provided when the data is exported. Requirements regarding the DSGVO are considered in this place.
- File management. In many places of the solution (process activities, materials, samples) files can be uploaded to the server. The file management allows for simple access to these files. The feature also is also capable of file filtering and file download.
- Database statistics. This feature is provided as a first prototype for future releases. It is used to aggregate database statistics to support the laboratory management process (e.g. how often was a certain test performed during the last year or by a person, …). This is especially usefull for quality management and accredited testing but also provides general information about the available data. The statistical informations are plotted as bar or pie charts.

Work package: Final implementation phase

When we evaluated implementation phase II we found out that all objectives of the proposal are satisfied and went ahead to the deployment of the solution and enhancements needed for future development.

Work carried out during this phase:

- Redesign of the booking calendar. Here we decided to add a graphical interface to the booking events (Django-Scheduler) and we also joined all kind of booking events into one datastructure. This allows for re-using the feature in many different places (device booking, room-section booking, etc.). This gives rise to a more open and reusable data structure and simplifies the further development.
- Deployment of the solution. We also implemented deployment scripts used to deploy the solution on a web-server in a convinient way. Now we are able to deploy the solution in three different stages -

"Testing", "Staging", "Production". The "Testing" stage is used during development to check the impact of new features and/or changes made to the code. "Staging" is used to check whether the solution is ready for the postgres DBMS or not and to perform initial tests with real data. This stage also allows for minor changes in source code and data structure. The "Production" stage is used to run the solution with real data in production mode. Running in this stage the solution can actually be used by laboratory staff and scientific staff.

- Gathering information for future development. In the beginning of November the project manager (J. Harden) had a talk with with the laboratory manager of IMBT/TVFA (M. Lenz). Outcomes of this talk: good overview of the provided tools (dashboards), simple to use, three additional process activities (often used laboratory tests) need to be implemented.
- At 26. Nov. 2021 the project team held a project management workshop. The purpose of this workshop was an elaborate code- and feature review and to plan for further development of the solution. In the course of this workshop the evaluation of the final implementation phase was completed.

Work package: Final project phase

In this phase the solution was tested and documented. We also provided test instances of LabNDB for other stakeholders and finished off the necessary reports.

- Stable user-interface and code documentation. After the implementation phases we had to test and settle the solution. The labeling in the user-interface was made consistent. Typing errors were removed and the code documentation was updated. We also redesigned some code sections to make future development more simple.
- Providing test versions. On request we provided two "Testing" deployments for FELMI and IMAT. Now these two institutes can review the solution and provide us with information and suggestions for the ongoing development.
- File parser. We implemented a file parser for the "Messphysik" measurement data file format. This file format is used very often in our lab's but also in many other lab's of this university. This file parser is providing the data in a way (pandas dataframe) that can easily be re-used in the user-interface to plot measurement data and is a very good basis for future data analysis tools.
- Create the final report for submission to the RDM Marketplace.

Changes: At the beginning we thought that an already existing feature of the Django web-framework regarding the integration of SI-units for measurement data fields would satisfy our requirements. At the end we found out that this feature has some unhandy limitations in usage and configuration. So we decided to implement the integration of SI-units on our own. Our own implementation now provides for simple usage and configuration of measurement data fields, it is well documented and open for future development.

The first design of the laboratory room section and test device booking consisted of a simple definition of the booked time range and some checks to avoid double bookings. First tests turned out that this implementation does not provide a good overview over existing bookings and the availability of test devices and room sections. So we decided to go ahead with a graphical solution that makes this kind of information much more visible. It takes more effort regarding the implementation and leads to a more complex deployment but leads to much better usability of the solution in the end.

3. Project team, cooperation and sustainability

Since the last report the project team did not change.

Cooperations: We continued our cooperation with the Institute of Biomechanics (P. Knabl, S. Sherifova). In general this cooperation turned out to be very beneficial. We shared our informations about the requirements of the solution, our experiences in software development and knowledge on the tools we used. We also helped each other to solve the issues that came up. Here we also want to emphasize the cooperation between the developers (M. Eisner, B. Schmid, P. Knabl) who really commited themselve to this project and did the work toghether without any major interactions with the project management. It is planned that this cooperation will be continued beyond the limits of this project piloting phase. The first step towards this ongoing cooperation was a project management workshop on 26.11.2021 to plan for the future development of the solution. Furthermore there is an upcoming cooperation with the Institute of Electron Microscopy and Nanoanalysis – FELMI. They showed much interest in our solution and we are looking forward to continue the development with their expiriences and suggestions they will share with us.

Dissemination: A talk was given by J. Harden on 08.11.2021 to present the solution to the RDM Marketplace pilot team from the Institute of Electron Microscopy and Nanoanalysis - FELMI (W. Grogger, A. Zankel, R. Schmidt, J. Gurker). They were pretty surprised about the completeness of the solution and the rich featured user-interface. Together we found out that LabNDB would be a good starting point for the future activities regarding RDM on their institute. Main outcomes of the talk: we provide a test instance of the solution for FELMI and we keep contact to share experiences needed for the future development of LabNDB and best practices in RDM in general.

Furthermore a presentation of the solution at the Institute of Technology and Testing of Construction Materials is planned.

Sustainability: The development of the solution will continue on our institute beyond the limits of this RDM Marketplace pilot project. Therefore we employed our project developers (M. Eisner, B. Schmid) until end of the first quarter of 2022. Further development is continued in cooperation with the Institute of Biomechanics and FELMI. The goal is to develop a stable version of this solution for production use within the next year.

4. Project costs

The costs are related to the fee of the two developers (M. Eisner, B. Schmid) employed for this project. Other contributions are institute-inkind and therefore not part of the costs shown below. No money was spent on software for the reason that we only use open-source-software. At the beginning of this project we planned to spend money on tablet computers and server infrastructure. This was not necessary for the reason that we are able to run the solution on existing measurement PC’s and office workstations. The server infrastructure is provided by a virtual machine available on the institute’s application server.

At this point we can state that there are no deviations in the proposed cost plan.

Table 3: Project costs

Funded	Spent
22.500,00	22.500,00

In this place we also want to mention the institute-inkind that is not considered in the project costs. The inkind consists of project management, software development and other informational contributions to this project. Estimated amount of inkind to the end of this reporting period: approx. € 23.000,00. Estimated total project cost (funded and inkind): approx. € 55.500,00.