



DAMSS

DATA ANALYSIS
METHODS FOR SOFTWARE
SYSTEMS



15th Conference on

DATA ANALYSIS METHODS for Software Systems

November 28–30, 2024

Druskininkai, Lithuania, Hotel “Europa Royale”

<https://www.mii.lt/DAMSS>

Co-Chairmen:

Prof. **Gintautas Dzemyda** (Vilnius University, Lithuanian Academy of Sciences)

Dr. **Saulius Maskeliūnas** (Lithuanian Computer Society)

Programme Committee:

Dr. **Jolita Bernatavičienė** (Lithuania)

Prof. **Juris Borzovs** (Latvia)

Prof. **Janis Grundspenkis** (Latvia)

Prof. **Janusz Kacprzyk** (Poland)

Prof. **Ignacy Kaliszewski** (Poland)

Prof. **Bożena Kostek** (Poland)

Prof. **Tomas Krilavičius** (Lithuania)

Prof. **Olga Kurasova** (Lithuania)

Assoc. Prof. **Tatiana Tchemisova** (Portugal)

Assoc. Prof. **Gintautas Tamulevičius** (Lithuania)

Prof. **Julius Žilinskas** (Lithuania)

Organizing Committee:

Dr. **Jolita Bernatavičienė**

Prof. **Olga Kurasova**

Assoc. Prof. **Viktor Medvedev**

Laima Paliulionienė

Assoc. Prof. **Martynas Sabaliauskas**

Prof. **Povilas Treigys**

Contacts:

Dr. Jolita Bernatavičienė

jolita.bernatavicienne@mif.vu.lt

Tel. (+370 5) 2109 315

Prof. Olga Kurasova

olga.kurasova@mif.vu.lt

Copyright © 2024 Authors. Published by Vilnius University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

<https://doi.org/10.15388/DAMSS.15.2024>

ISBN 978-609-07-1112-5 (digital PDF)

© Vilnius University, 2024

Information System Architecture of the Lithuanian National Biobanking Infrastructure

Justas Trinkūnas¹, Roma Puronaitė^{2,3}, Laurent Jacotot⁴,
Mike Woodward⁴, Mindaugas Morkūnas²

¹ Vilnius Gediminas Technical University

² Vilnius University Hospital Santaros klinikos

³ Institute of Data Science and Digital Technologies
Vilnius University

⁴ Modul-Bio

Justas.Trinkunas@santa.lt

In 2019, the Human Biological Resource Center project was started to create a modern national biobanking infrastructure in Lithuania and join the BBMRI-ERIC. Project objectives were formulated as follows: to create a National Human Biological Resources Center (HBRC) with a unified and standardized management system for collecting, processing, storing, and managing biological samples and related health information. The Lithuanian National Biobank collaborates with various organizations and institutions to fulfil its mission of supporting health research and public health initiatives. Vilnius University Hospital Santaros Klinikos (VULSK) as the main implementer of the IT project, together with its partners National Center of Pathology (VPC), The National Cancer Institute (NCI), Vilnius University (VU), Innovative Medicine Center (IMC), Lithuanian University of Health Sciences (LSMU), Hospital of Lithuanian University of Health Sciences Kauno Klinikos (LSMUL KK), signed the HBRC Joint Activity Agreement and Lithuania joined the BBMRI-ERIC network as the full member in 2024.

With the rapid advance of multi-omics-based assays and medical decision-support artificial intelligence tools, cutting-edge biomedical research striving to deploy personalized medicine in clinical practice increasingly depends on high-quality, well-annotated samples and datasets. Following the broad participants' consent, the samples in the storage can be linked with clinical information to form biospecimen col-

lections and further serve the development of personalized medicine strategies. However, inconsistent operational procedures cause collections to be scattered across Lithuanian biobanks, making samples and data inaccessible and slowing down request processing.

To achieve this goal, we designed and implemented biobanking platform reflecting the biobanking process in the Lithuanian biobank network. The architecture has three main parts: the hospital infrastructure (VULSK HIS, NCI HIS, VPC HIS), the biobank laboratories software (Modul-Bio's MBioLIMS BioBanking Software) for biospecimen management, and the information technology infrastructure integrating biobank networking operations (biobank participant registry, HIS integrations, The State Data Agency integration).

MBioLIMS BioBanking software manages the complete life-cycle from reception to distribution of biospecimens and associated data. Its multi-site functionality facilitates the connection of different biobanks within the same network. This allows national biobanks to work in the same system autonomously but also share nomenclature and standardized processes. It is possible for a supervising site to view and search all biobanks' data in the cluster.

As the feasibility study of the implemented system, we report here a case study of a VULSK biobank with 12,300 patients enrolled by June 1st 2024. At this date, the biobank database contained more than 101,500 inpatient, 823,200 outpatient, and 6,200 emergency encounters (including historical) and more than 120,000 DICOM images, mostly Ultrasound (US), Computed Radiography (CR), and Computed Tomography (CT) modalities. Patients had cancer and blood diseases (TOP 3 according to ICD-10-AM cancer and blood disease groups, D70-D77 - 33.2%, C81-C96 - 32.2%, D60-D64 - 29.1%), infections (TOP 3 ICD-10-AM infection groups, U00-U49 - 29.2%, A30-A49 - 18.9%, B95-B97 - 8.2%), other diseases (TOP 3 ICD-10-AM other diseases groups, E70-E89 - 38.2%, I10-I15 - 35.0%, J09-J18 - 24.7%).