

TECHNICAL SPECIFICATION
Open tender
“Cloud-based software development, including a database, analytics, and other functionality”
(Identification number LMT/BRAVO/2026-14)

1. This specification describes the software product to be developed, the purpose of which is to collect medical information and data from sensor devices, to validate, store, organise and process this data in conjunction with other data (linearly and using AI algorithms) and organising access to external systems, providing a cloud-based platform (developing a new platform or adapting an existing one) as a service for conducting medical research related to breath analysis.
2. The tenderer shall submit a Technical Proposal in accordance with the open tender announced by AS “LabMedTech” for the procurement of “Cloud-based software development comprising a database, analyses and other functionalities” (identification number LMT/BRAVO/2026-14) the requirements of the Tender Documents and Technical Specifications, with the Technical Proposal providing a detailed description of the fulfilment, provision and implementation of all requirements, including all activities related to development, delivery and warranty technical support, which will result in the realisation of the subject matter of the procurement. The description must be sufficiently detailed to enable the Contracting Authority to be satisfied as to the Tenderer’s ability to fulfil all the requirements set out in the Tender Documents and the Technical Specifications.

TECHNICAL REQUIREMENTS / SPECIFICATION:

Development/Delivery/Warranty technical support*	Quantity	Unit	Requirements
1. Base platform for a cloud-based solution for conducting medical research	1	Platform	● Location and compliance: The platform must be located within the European Union (EU) and must strictly comply with the GDPR and other applicable EU and Latvian legislation on the protection of health data. All data, including backups, logs and test environments, must be stored and processed exclusively within the EU/EEA. ● Data processing and performance: The platform must be capable of receiving data, performing real-time or near-real-time data analysis, and carrying out quality control with anomaly detection. It must ensure drift compensation, normalisation of signal data in line with environmental changes, and the harmonisation of multiple instruments. ● Architecture: The solution must be implemented as a multi-tenant web application with a modular architecture and container orchestration. All services must be containerised and stored in a centralised container registry. The architecture must be flexible and extensible so that the Client can add new data processing steps, algorithms and models without modifying the Platform’s source code. ● Data storage: Key transaction and configuration data must be stored in a managed relational database. Large volumes of sensor and log data must be stored in a scalable object store, whilst a specialised time-series database must be used for time-series queries. ● Security and encryption: All data traffic must be encrypted. For external data traffic, TLS termination

			must be used on the load balancer or API gateway. For data that has not been accessed for a long time (at rest) in databases, object stores and attached storage, managed encryption with a key management system must be used. ● Access control and monitoring: A centralised identity management system with role-based access control must be used for authentication and authorisation. The platform must implement appropriate activity logging, metrics and system monitoring, using centralised auditing and security posture monitoring tools to detect incidents and threats. ● Portability: The platform must be made available to the client upon completion of the project and must be designed in such a way that it can also be deployed on the Client's server (as a containerised solution). The Client must be able to independently use data for model training and to integrate their own data processing algorithms and models developed in the <i>Python</i> environment, as well as to deploy, configure and replace them independently without the Supplier's involvement and without modifying the Platform's source code.
1.1. Secure, automated data exchange	1	Module	● Encryption and security: Data transmission functionality must be provided with mandatory encryption in transit (at least AES-256 and TLS 1.3 standards) and encryption at rest. ● Data exchange mechanisms: The module must expose RESTful or equivalent application programming interface (API) endpoints. Event-driven integration with asynchronous delivery and message buffering must be provided to ensure reliable disconnection. The RESTful API must allow respiratory analysis instruments to upload new measurement data. ● File transfer: Standard protocols (e.g. SFTP) must be supported for the secure transfer of files from medical devices or partners. ● Management: API keys, certificates and accreditation data must be stored in a secure secrets management system. Access control must be integrated with the API gateway. All access and configuration changes must be logged in audit logs and linked to an automated threat detection system in accordance with healthcare and GDPR requirements. The supplier must provide incident response plans in accordance with the GDPR, including notification of the contracting authority of a personal data breach within a timeframe that enables compliance with the requirements of Article 33 of the GDPR.
1.2. Data pre-processing, scaling, transformation, database	1	Module	● Signal processing logic: The module must implement pre-processing of time series and other data types for the analysis of data from breath analysis instruments. Data from breath analysis instruments contain measurement parameters, auxiliary data (e.g. environmental and instrument conditions), as well as sensor response data from multiple sensors. The signal processing pipeline must be configurable and extensible so that new instruments with varying data structures can be added at a later stage. The signal processing pipeline must include the determination of the measurement period and exhalation phase, visualisation of the sensor baseline, normalisation and calibration (to compensate for fluctuations in temperature, humidity

			and pressure, as well as sensor ageing). ● Error detection: Detection of erroneous signals (e.g. noise, excessively high peaks) and feature extraction (delta normalisation, exhalation characteristics) must be ensured. ● Data pipelines: Pre-processing must be implemented using scalable data transformation tools operating within a data lake with a centralised data catalogue. ● Instrument calibration data: must be stored in the system and appropriately integrated into the data stream. Calibration data may include reference measurements. A dedicated calibration workflow must be implemented to derive calibration parameters from calibration measurements, thereby minimising the impact of drifts, sensor ageing, cross-talk and device variations. ● Data storage formats: Processed and historical data must be stored in an object store in optimised, encrypted formats (e.g. CSV or JSON). A relational database must be used for structured queries, whilst a time-series database must be used for time series. ● Data Governance: Access to data and the evolution of schemas must be strictly controlled by the data governance layer, ensuring that users see only authorised data. Automated detection of sensitive personal data must be implemented in both raw and pre-processed layers.
1.3. Data analysis and modelling environment	1	Module	● Analytical functionality: The module must provide an environment in which the executed and maintained mathematical, statistical and machine learning (ML)-based algorithms (including neural networks) developed or requested by the Client, in order to generate forecasts and determine group affiliation with regard to diseases. Tools for analysing features and their impact must be provided. The scientific content, selection and validation of algorithms and models are determined by the Client; the Supplier does not develop research algorithms, but provides the means for their integration, training, execution and deployment. ● Machine learning (ML) platform: A unified, managed ML platform for data exploration and collaboration must be used for the connection, training and execution of models. Managed training processes must be provided using algorithms developed in the Python environment supplied by the Client or customised containers, without any changes to the Platform's source code. The Supplier shall provide a documented algorithm integration interface and at least two reference examples. ● MLOps and management: Experiment tracking, model version management (registry), and analysis of model explainability and bias for medical purposes must be implemented. Automated ML pipelines must be used for repeatable workflows. ● Infrastructure monitoring: All operations, compliance/control evidence and metrics must be automatically routed to a centralised logging and audit system. ● Integration of external models and large language models (LLMs): The platform must provide a configurable interface that allows the Client to independently connect, replace and disconnect external or

			locally deployed AI models, including large language models (LLMs), by specifying the service endpoint, authentication credentials and usage restrictions, without modifying the Platform's source code. A register of connected models and services, version and configuration management, usage tracking, and the ability to specify which data may be transferred to external services must be provided, whilst preventing the transfer of personal data outside the EU/EEA without the Customer's authorisation.
1.4. Deployment of the model for use by researchers and healthcare staff	1	Module	● Deployment architecture: Models developed or requested by the Client must be capable of being deployed, hosted and used in a scalable environment managed via an API gateway. A standardised model ingestion format (e.g. containerised model, ONNX or Python artefact) must be provided, along with model versioning and the ability for the Client to deploy a new model version independently without the Supplier's involvement. Access must be compliant with GDPR requirements. ● Network security and integration: Network access to model endpoints must be restricted using virtual private networks, security groups and private links, ensuring a secure connection to hospital and research networks. ● Monitoring: The system must monitor performance metrics (latency, error rate, throughput) in real time, as well as detect data and model drift. All API calls must be logged with full traceability in audit logs.
1.5. Interface for data browsing and research	1	Module	● User interface (UI): Researchers and clinicians must be provided with a secure web application that is globally accessible with low latency and protected by a web application firewall (WAF). ● Functionality: The interface must enable data browsing, report generation, result visualisation and data aggregation based on group membership and measurements. Business intelligence (BI) or interactive analytics dashboards must be integrated for information presentation. ● Different user types require different user interfaces: User types include administrators (who can perform backups, updates and user management), clinicians (doctors, nurses) with access to data on their own patients and the ability to upload new measurements, and researchers with the ability to access datasets and configure/implement data analysis workflows. ● Access and control: A centralised solution with role- and attribute-based access control (RBAC/ABAC) must be used for authentication and authorisation. All user interface actions, particularly data export and access, must be logged and retained for audit purposes.
1.6. Platform module for patient metadata collection and consent management	1	Module	● Database and security: Patient metadata and consent records must be stored in an isolated database with strict network controls and encryption. ● Architecture: Data entry and editing logic must be implemented as microservices behind an API gateway. ● Privacy and compliance: Immutable audit logs must be created for all data subject rights operations (access, rectification, erasure) and changes to consent. The system must provide an interface through which

			an individual gives active consent to data processing operations. All personal data across all datasets must be stored and processed in such a way as to ensure protection in accordance with the GDPR and applicable EU health data protection regulations.
1.7. Adapting the platform to handle new data types (GC and MOX integration)	1	Updated platform	• Scalability: The platform must be capable of assimilating additional data modalities, such as gas chromatography (GC) and metal oxide (MOX) sensor readings. GC data ingestion, peak detection and correlation must be ensured. • Data pipelines: New data processing and import logic must be implemented as automated data processing jobs that operate on data stored in the object store and are managed via the central data catalogue. • Big data analytics: Time-series features and peaks must be stored in relational or time-series databases. Large-scale multimodal feature sets must be stored in an optimised format, enabling high-performance SQL queries using analytical data warehouses or query tools. • Multimodal modelling: To combine features obtained from GC and MOX sensors, monomodal and multimodal models must utilise a unified MLOps lifecycle (training, experiments, pipelines and registry) within a centralised ML platform. • Visual analysis: Integrated visual analysis of chromatograms, GOS profiles and sensor signatures must be accessible via an API and displayed in secure, embedded dashboards. All access to both data modalities must be centrally audited and protected.

2. Integral elements of a cloud-based platform for conducting clinical trials

2.1. Examples of data and device interfaces

MOX sensor data	GC data	Patient metadata	Technical formats The specific formats supported must be specified:
• device identifier; • measurement date and time; • sensor channel identifier; • output signal value; • temperature; • humidity; • pressure; • flow rate; • measurement phase; • calibration information;	• example of a chromatogram file; • time axis; • signal intensity; • peak identifier; • retention time; • peak area; • peak height; • compound name, if known; • instrument identifier;	• age group; • gender, if required for the study; • diagnosis group; • study group; • medication and comorbidities fields; • consent status; • study visit identifier.	• JSON; • CSV; • Parquet; • XML, if required; • SFTP; • REST API; • OpenAPI specification; • maximum file size; • encoding; • time zones and date format;

error and status codes.	calibration and quality control data.		mandatory and optional fields.
---	---	--	--

2.2. Mandatory and future functionality*

A. Mandatory functionality to be delivered	B. Mandatory functionality to be developed, but to be specified during the project	C. Optional or future functionality
- user and rights management; - data retrieval via API and SFTP; - data validation; - data storage; - audit logs; - management of patient metadata and consent; - data browsing and export; - basic statistical analysis; - integration of Python algorithms, the Client's models and external AI services (including LLM); - deployment of the system within the EU; - containerisation and source code handover;	- signal normalisation; - calibration algorithms; - anomaly detection; - exhalation phase detection; - instrument harmonisation; - model training and deployment workflows; - ML model registry; - model interpretability;	- automatic identification of complete GC data connections; - multimodal clinical diagnostic models; - integration with external hospital systems; - automated clinical decision support functionality; - additional data modalities.

* These requirements may be implemented as future development work outside the scope of this procurement.

3. Acceptance criteria for each module*

No.	Module	Acceptance criterion – the module is accepted if:
1	The base platform	- is deployed in the specified environment; - it operates in a containerised environment; - the architecture is documented; - source code and CI/CD are available; - the system passes load testing; - all data is located within the EU.
2	Data exchange module	The API complies with the OpenAPI specification;

		• data is accepted and validated; • incorrect data is rejected with a clear error message; • message retry is supported; • SFTP is supported; • access operations are audited.
3	Data pre-processing	• the system processes specified sample files; • normalisation is carried out; • defined anomalies are identified; • the specified features are calculated; • the result is reproducible.
4	ML and model environment module	• it is possible to create and run a training process; • it is possible to register the model version; • it is possible to repeat the experiment; • model metrics are available; • the model can be deployed to an API endpoint; • the model explainability tool is operational; it is possible to connect and deploy a model prepared by the Client, as well as to connect an external AI or LLM service without modifying the source code
5	User interface	• only authorised functions are available for each role; • data browsing is possible; • report generation is possible; • data export is possible; • all actions are audited; • usability tests have been carried out.
6	Patient Consent Module	• consent can be given, withdrawn and renewed; • all changes are auditable; • withdrawal of consent affects data processing; • it is possible to comply with a data subject's request; • sensitive data is segregated.

7	GC/MOX module	• GC and MOX samples provided by the Client are imported; • peaks are identified; • the data is displayed graphically; • it is possible to link GC and MOX features; • it is possible to use the data in a single ML process.
---	----------------------	---

***Model acceptance criteria**

The Supplier must ensure that the Platform allows for the connection, training, comparison and deployment of at least three methodologically distinct model approaches provided by the Client, the performance of cross-validation, the calculation of predefined metrics and the attainment of reproducible results. The scientific content and selection of models shall be provided by the Client; the clinical suitability of the final model shall be assessed by the Client's scientific and clinical team.

4. Third-party components

Before using a component, the tenderer shall specify its licence, costs and restrictions on use. Components that restrict the Client's rights to modify, maintain or transfer the system to another maintainer may not be used, unless the Client has given its prior written approval.

5. Vendor lock-in

The solution must be deployable on the Client's infrastructure or in another service provider's cloud environment without significant software re-engineering.

6. Project phases and milestones

Phase	Time period	Deliverables
Phase 1 — Clarification of requirements and architecture	3–4 weeks.	• detailed requirements; • data models; • architecture document; • security architecture; • project plan; • risk register; • acceptance test plan.
Phase 2 — base platform	6–8 weeks.	• infrastructure; • user management; • databases; • API; • SFTP; • auditing;

		• basic user interface.
Stage 3 — data pre-processing	6–8 weeks.	• data pipelines; • validation; • normalisation; • calibration; • error detection; • data catalogue.
Stage 4 — ML and analytics	6–8 weeks.	• Environment for integrating ML and external models (including LLMs); • experiment tracking; • model registry; • training and deployment workflows; • dashboards and reports.
Phase 5 — GC/MOX integration	4–6 weeks.	• GC import; • Peak identification; • MOX data processing; • multimodal data structures; • visualisation.
Phase 6 — acceptance and implementation	3–4 weeks.	• security check; • performance testing; • user testing; • documentation; • training; • commissioning.

7. Separation of platform development from research and clinical validation

Subject of the procurement A — software platform

The supplier is responsible for:	The contracting authority shall provide:	Research algorithms (the Client's responsibility)
• infrastructure; • data reception; • databases; • data management;	• study design; • clinical expertise; • labelled data; • explanation of data quality;	• selection of signal processing methods; • feature engineering; • calibration models; • disease classification models;

• user interfaces; • ML platform; • Deployment of the client's models; • auditing; • security; • Technical integration of GC/MOX data.	• definitions of diagnoses and groups; • acceptable model quality metrics; • clinical validation; • a decision on the use of the model for medical purposes.	• comparison of models; • statistical validation
---	---	---

Name and surname of the applicant or their authorised representative _____
 Position _____

The date of signing of this document is the date of the attached secure electronic signature and its time stamp.

The tender (including this document) shall be signed by the tenderer's representative using a secure electronic signature and a time stamp.