BIOPRESERVATION AND BIOBANKING
Volume 22, Number 4, 2024
Mary Ann Liebert, Inc.
DOI: 10.1089/bio.2023.0074

Open camera or QR reader and scan code to access this article and other resources online.

Update of the Minimum Information About BIobank Data Sharing (MIABIS) Core Terminology to the 3rd Version

Niina Eklund,[1] Cäcilia Engels,[2–4] Michael Neumann,[5] Andrzej Strug,[6] Esther van Enckevort,[7] Ronny Baber,[8] Margreet Bloemers,[9] Annelies Debucquoy,[10] Aad van der Lugt,[11] Heimo Müller,[12] Lauri Parkkonen,[13] Philip R. Quinlan,[14] Esmond Urwin,[14] Petr Holub,[15] Kaisa Silander,[1] and Gabriele Anton[16]

Introduction: The Minimum Information About BIobank Data Sharing (MIABIS) is a biobank-specific terminology enabling the sharing of biobank-related data for different purposes across a wide range of database implementations. After 4 years in use and with the first version of the individual-level MIABIS component *Sample*, *Sample donor,* and *Event*, it was necessary to revise the terminology, especially to include biobanks that work more in the data domain than with samples.

Materials & Methods: Nine use-cases representing different types of biobanks, studies, and networks participated in the development work. They represent types of data, specific sample types, or levels of organization that were not included earlier in MIABIS. To support our revision of the *Biobank* entity, we conducted a survey of European biobanks to chart the services they provide. An important stakeholder group for biobanks include researchers as the main users of biobanks. To be able to render MIABIS more researcher-friendly, we collected different sample/data requests to analyze the terminology adjustment needs in detail. During the update process, the Core terminology was iteratively reviewed by a large group of experts until a consensus was reached.

Results: With this update, MIABIS was adjusted to encompass data-driven biobanks and to include data collections, while also describing the services and capabilities biobanks offer to their users, besides the retrospective samples. The terminology was also extended to accommodate sample and data collections of nonhuman origin. Additionally, a set of organizational attributes was compiled to describe networks.

Discussion: The usability of MIABIS Core v3 was increased by extending it to cover more topics of the biobanking domain. Additionally, the focus was on a more general terminology and harmonization of attributes

[1]Finnish Institute for Health and Welfare, Helsinki, Finland.
[2]German Biobank Node (GBN), Charité – Universitätsmedizin Berlin, Berlin, Germany.
[3]German Cancer Consortium (DKTK), Heidelberg, Germany.
[4]Charité University Hospital Berlin, Berlin, Germany.
[5]University Hospital Würzburg, Würzburg, Germany.
[6]Department of Medical Laboratory Diagnostics, Medical University of Gdansk, Gdansk, Poland.
[7]University Medical Center Groningen, University of Groningen, Groningen, The Netherlands.
[8]Leipzig Medical Biobank, Leipzig, Germany and Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, University of Leipzig Medical Center, Leipzig, Germany.
[9]ZonMw Organisation for Health Research and Development, the Hague, The Netherlands.
[10]Belgian Cancer Registry, Brussels, Belgium.
[11]Erasmus Medical Center, Rotterdam, The Netherlands.
[12]Medical University of Graz, Graz, Austria.
[13]Department of Neuroscience and Biomedical Engineering, Aalto University, Espoo, Finland.
[14]University of Nottingham, Nottingham, United Kingdom.
[15]BBMRI-ERIC, Graz, Austria.
[16]Helmholtz Zentrum München, Neuherberg, Germany.

with the individual-level entities *Sample, Sample donor, and Event* to keep the overall terminology minimal. With this work, the internal semantics of the MIABIS terminology was improved.

Keywords: MIABIS, biobank, network, standardization, interoperability

Introduction

The Minimum Information About BIobank Data Sharing (MIABIS) is a biobank-specific terminology enabling the sharing of minimal biobank-related data for different purposes across a wide range of database implementations. The development of MIABIS was initiated in 2012 by the Biobanking and BioMolecular Resources Research Infrastructure of Sweden (BBMRI.se).[1] Since 2017, the Common Service IT (CS IT)[2] within the Biobanking and Biomolecular Resources Research Infrastructure—European Research Infrastructure Consortium (BBMRI-ERIC) has coordinated the further development of the MIABIS terminology. Originally, the main aim of MIABIS Core was to establish a means to describe administrative information about biobanks, which at the time was not systematically represented in any terminology. At present, MIABIS Core version 2[3] is used in several biobank registers and catalogs,[4,5] it is implemented as part of data models for biobank and research information management systems,[6] and it is integrated in other data models[7] and biobanking ontologies.[8] The MIABIS terminology has recently been complemented with the extension of an individual-level component to describe samples, sample donors, and related events.[9]

MIABIS Core v2, together with the MIABIS individual-level reference data model for *Sample*, *Sample donor*, and *Event*, render the terminology heavily sample-centric. Biobanks, however, can also be built around data generated from samples and/or obtained from donors and may not include biological samples at all. Imaging biobanks[7,10] are an example of such biobanks. In addition, many biobanks today offer mostly data to researchers, as their infrastructure already contains extensive omics-data derived from biosamples.[11] The biobanking community also includes banks collecting nonhuman samples relevant to public health,[12] such as microbes from the environment. Owing to the increasing demand for collaborative work and large datasets, the need for research networks combining the expertise of individual network members[13–15] has arisen.

Thus, there was a need to update MIABIS Core, which we addressed with the efforts and results described here.

Materials and Methods

We identified nine different use-cases representing different types of biobanks, studies, and networks operating with big data or specific sample/data types, not previously defined in MIABIS (see Supplementary Table S1). In addition, in 2020-2021, we collected 38 sample/data requests from BBMRI-ERIC national nodes (from Austria, Belgium, Finland, Germany, The Netherlands, and Poland) to make informed decisions on how the terminology should be adapted (summarized in Supplementary Table S2).

In addition, to describe the biobanks at a more practical level, we conducted a survey in September 2021 targeting over 500 biobanks in the BBMRI-ERIC network about the services that the biobanks offer to their clients or what they have included in their infrastructure, such as availability of sample storing services and facilities, possibility to provide access to additional data, possibility to provide analytical services to biobank samples, and possibility to provide cell culture services. Seven biobanks responded to the survey, and their answers were included in the *Biobank* entity revision.

The names and descriptions of the entities were evaluated to ensure their compatibility with the new domain extensions. Each attribute was then thoroughly reviewed, and suggested modifications were discussed by the working group until consensus was reached. The MIABIS working group consisted of experts in biobanking, data (such as antimicrobial data resistance data, omics, epidemiological and population-based lifestyle data, clinical information, imaging data), interoperability, and IT. The assembly of experts in the MIABIS working group included the authors of the article and individuals mentioned in the acknowledgments.

Results

Moving from MIABIS Core v2 to v3

In this work, we propose an update for the MIABIS Core v2 entities *Biobank*, *Sample collection,* and *Study,* as well as an update of the structure of the MIABIS Core by including the new entity *Network*, renaming the former entity *Study* to *Research resource*, and defining it as an optional entity in the Core. We will first show our main aims and describe the important aspects of the update process before presenting the new MIABIS Core version 3.0.

Aims and conduction of the MIABIS update

We pursued the following aims with the MIABIS update:

(1) Incorporate the concept of data-driven biobanks and data collections to support biobanks providing mostly data:
As accompanying data and sample-derived data are becoming ever more relevant in biobanking, the descriptions of *Collection* and *Research resource* entities were rephrased to widen their applicability to data derived from samples or donors. We also revised attributes, e.g., the "Data categories" attribute to better reflect data-driven biobanks.

(2) Synchronize MIABIS Core attributes and values with the MIABIS individual-level component *Sample*, *Sample donor,* and *Event*:
To achieve synchronization with the individual-level entities *Sample*, *Sample donor,* and *Event*, it was necessary to update some MIABIS Core attributes, e.g.,

"Storage temperature" and "Material type," which were subsequently renamed as "Sample type."

(3) Allow the description of nonhuman sample and data collections related to biomedical research:
To be able to include samples of nonhuman origin, new sample types were added to the "Sample type" attribute value list, such as "Specimen from environment or food," and in general, attribute descriptions were changed to be more generic and include nonhuman samples.

(4) Restructure of the MIABIS Core:
We included an entity on biomedical networks, which has become more and more important for biobanks and in biomedical research. Additionally, we renamed the entity *Study* to *Research resource* and made this entity an optional one.

Thus, the MIABIS Core 3.0 consists of the following three entities: *Network*, *Biobank,* and *Collection*. Together, they form a more natural combination of biobank data sharing, and the information provided by these entities needs to be already used in several catalog solutions, such as the BBMRI-ERIC Directory.[4] We made the *Research resource* optional, as this entity is not relevant in current biobank catalog solutions but may receive more attention in the future to reflect active and successfully conducted projects of biobanks and network infrastructures. The updated MIABIS structure is shown in Figure 1, which also includes perspectives of an additional component on *Dataset* entities.

(5) Define a set of attributes for services and capabilities the biobanks offer to their customers apart from providing retrospective samples.

(6) Adjust the MIABIS terminology to provide better search capabilities for available data and samples for the main stakeholder group of biobanks: researchers.

From the 38 researchers' queries, we were able to identify similar features and group them by their type (Supplementary Table S2). We ensured that the researchers' needs are included in the *Collection* entity update, especially in the "Dataset type" attribute. We included most of the needs, but some requirements were too specific and would be better represented in their own extension, such as clinical data.

(7) Review entity names and descriptions and MIABIS metadata for overall consistency.

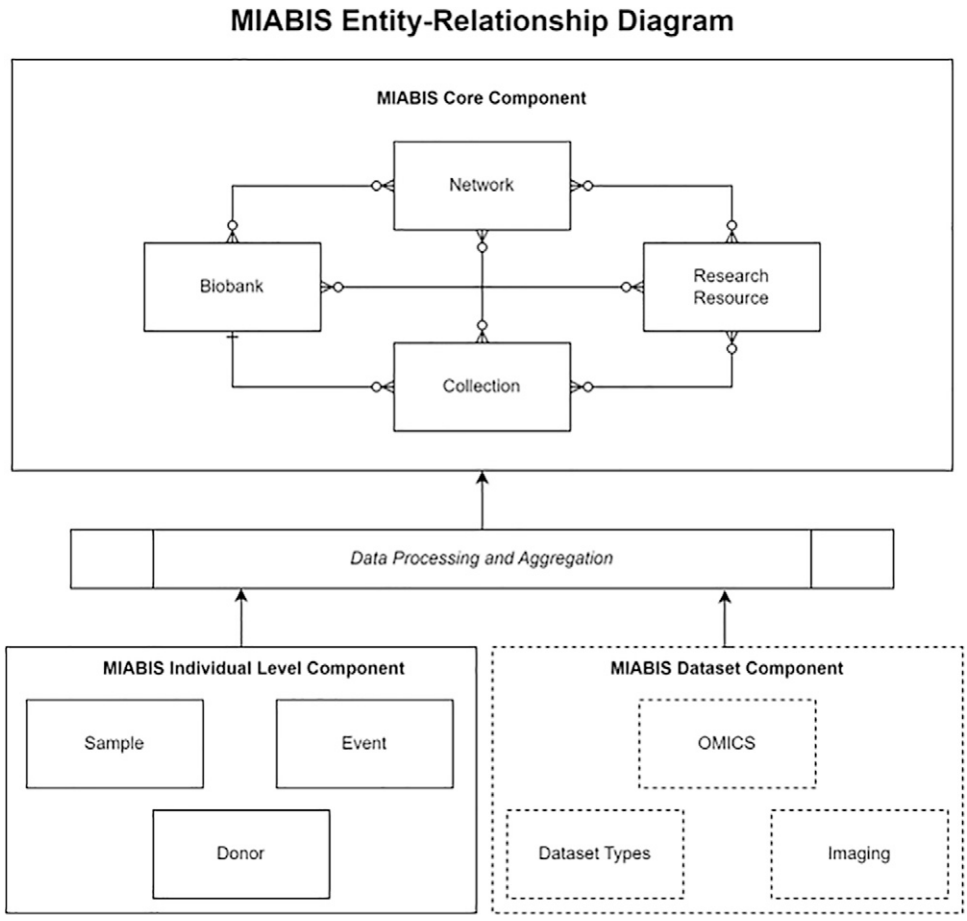

FIG. 1. The MIABIS terminology structure consists of aggregate-level and individual-level components. Together the *Biobank*, *Collection*, and *Network* form the Core, with the optional *Research resource* being maintained along the Core component. The individual-level component *Sample*, *Sample donor*, and *Event* are a basis to provide information from which some data elements of the *Collection* can be aggregated. Here, we already anticipate the next component on datasets, omics, and imaging, which will be on both aggregate and individual levels. This diagram uses crow's foot notation to depict the relationship between the Core entities. MIABIS, Minimum Information About BIobank Data Sharing.

One goal of this revision was to achieve a level of generalization, to make the entities and attributes as reusable and self-explanatory as possible. Therefore, entity names and descriptions were revised and updated as shown in the next chapter. We aimed to keep the entity descriptions short and informative but still generic to avoid the need for constant adaptation. This way, we will ensure the broad usability of the terminology in different settings.

We reviewed the descriptive structure and metadata of the MIABIS terminology. All the attribute descriptions were synchronized across the component, and the original description field from MIABIS v2 was split into two columns to accommodate descriptions and constraints. In addition, all the attribute codes were updated throughout the terminology so that each attribute has a unique and informative entity-specific code using the coding scheme described earlier.[9] As part of the metadata revision, the data types for allowed values were also revised to be more accurate by changing the "Text list" data type to "Enumerated values."

MIABIS Core 3.0 entities

The entities included in the MIABIS Core v3 are defined as follows:

- **Biobank**: A legal entity or part of one that performs biobanking (modified according to ISO 20387:2018).[16]
- **Collection**: A set of samples and/or data items collected and/or derived in a common context, stored under a common stewardship, and available for future research (modified according to Merino-Martinez *et al*, 2016).[3]
- **Research resource**: A set of samples and/or data items used and/or analyzed in a common context in past or current research. A *research resource* may combine material from multiple collections and biobanks (modified according to Merino-Martinez *et al*, 2016).[3]
- **Network**: A group of interconnected biobanks, collections, and/or research resources with defined common governance (modified according to National Cancer Institute Thesaurus (NCIT) definition [Ontobee: NCIT]).[17]

In addition, we adapted the definition of a sample to include samples of nonhuman origin.

- **Sample**: A sample is a portion or a quantity of biological material that is collected from a sample donor or the environment, or a digital representation of these, such as an image (modified according Eklund et al., 2020).[9]

We aimed to keep the entity descriptions short and informative, but still generic to avoid the need for constant adaptation. This way we will ensure the broad usability of the terminology in different settings.

Attribute lists for Core entities

When revising the attributes, we carefully considered backwards compatibility with previous MIABIS versions. However, some carefully considered breaks were accepted to ensure a more accurate representation of attributes.

Biobank entity

The existing attributes in the given *Biobank* entity, presented in Table 1, were found to be mostly adequate to describe organizational and administrative information about biobanks. Only "Contact information" was updated by merging it with the structured attribute "Principal investigator" from MIABIS Core v2 *Study* (structured attributes are presented in Supplementary Table S3A), as both attributes were very similar. In addition, based on the survey of biobanks, we added three new attributes for biobank capabilities and one for quality. The "Quality Management standard" attribute is the first quality attribute included in the MIABIS dataset and gives the opportunity to describe the quality accreditation or certification status of biobanks. The proposed list of values is minimal; however, we recommend that it be expanded to include local requirements when implemented.

Collection and Research resource entities

The *Collection* and *Research resource* attribute lists (presented in Tables 2 and 3, respectively) needed more extensive revision. The attribute lists of the two entities *Collection* and *Research resource* were merged, because they contained many similar attributes.

Regarding synchronization between attributes in the Core component and in the individual-level component, the "Detailed sample type" and "Sample storage temperature" attributes were the references for value list updates to the *Collection*'s "Storage temperature" and "Material type," subsequently renamed as "Sample type." In addition, when expanding the scope of MIABIS to include samples of nonhuman origin, we added new sample types to the list, such as "Specimen from environment or food," and made the attribute description more generic. Because the list of sample types in the *Sample* entity's "Detailed sample type" attribute is extensive, we aggregated the list in the mapping table (Table 4). For the "Storage temperature" attribute, we synchronized the value list with the same attribute in the *Sample* entity, where the "Sample storage temperature" values are based on Standard PREanalytical Code (SPREC) v3.[18] The value lists for the "Age low unit" and "Age high unit" attributes were also synchronized with the *Sample* entity.

The *Collection* entity has been extended from its previous version to include new attributes for "Dataset type," "Use and access conditions," "Publications," "Sample source," "Collection setting," and "Collection status." We replaced the original "Data categories" attribute—a mix of data types and sources—with the "Dataset type" attribute. "Dataset type" is a more refined attribute that is used to categorize available data. Its value list combines the MIABIS Core v2 "Data categories" value list, its revision in the *Sample donor* entity, and the requirements identified from use-cases collected sample/data requests. In addition, another attribute from the *Sample* entity, "Sample use restrictions," is introduced in the *Collection* entity. To make it more generic, we renamed the attribute to "Use & Access conditions" and revised the value list. We recognize that it may not always be possible to specify usage restrictions at the collection level. However, if these access conditions apply to the whole collection, this information is very valuable to researchers.

Table 1. Attribute Definitions in the Biobank Entity

Attribute code	*Mapping to Core v2*	*Attribute name*	*Value type*	*Description*	*Constraints*	*Cardinality*
MIABIS-BIOBANK-01	MIABIS-2.0–01	ID	Text	ID of the Biobank	The ID is technical and given by the implementation. Implementation will provide instructions on how to form the ID	1
MIABIS-BIOBANK-02	MIABIS-2.0–02	Acronym	Text	Short name in use for the Biobank, if applicable		0..1
MIABIS-BIOBANK-03	MIABIS-2.0–03	Name	Text	Name of the Biobank (preferably in English)		1
MIABIS-BIOBANK-04	MIABIS-2.0–04	URL	Text	Complete http-address for the Biobank		0..1
MIABIS-BIOBANK-05	MIABIS-2.0–05	Juristic person	Text	Juristic person, e.g., a university, concern, county council, etc., for the Biobank		1
MIABIS-BIOBANK-06	MIABIS-2.0–06	Country	Text	Two-letter code for the country of the Biobank	ISO-standard 3166 alpha2	1
MIABIS-BIOBANK-07	MIABIS-2.0–07	Contact information	Structured data	Contact information for the contact person/person responsible of the Biobank (Structured attribute)		1
MIABIS-BIOBANK-08	MIABIS-2.0–08	Description	Text	Description of the Biobank in English	Free text description of the Biobank. Recommendation max. 2000 char.	1
MIABIS-BIOBANK-09	NEW	Infrastructural capabilities	Enumerated values	The technical infrastructural capabilities that the Biobank can offer to the clients. It can be ≥1 of the following values: Sample storage, Data storage, Biosafety abilities		0..n
MIABIS-BIOBANK-10	NEW	Organizational capabilities	Enumerated values	The organizational capabilities and services that the Biobank can provide to support clients. It can be ≥1 of the following values: Recontact with donors, Facilitating clinical trials, Setting up prospective collections, Access to omics data, Access to laboratory analysis data, Access to donors' clinical		0..n

(*continued*)

Table 1. (Continued)

Attribute code	*Mapping to Core v2*	*Attribute name*	*Value type*	*Description*	*Constraints*	*Cardinality*
				data, Access to pathology archive, Access to radiology archive, Access to national medical registries, Other		
MIABIS-BIOBANK-11	NEW	Bioprocessing and analytical capabilities	Enumerated values	Bioprocessing and analytical services that the Biobank can offer to the clients. It can be ≥1 of the following values: Biochemical analyses, DNA analysis, Nucleic acid extraction, Genomics, Proteomics, Metabolomics, Histology, Cell-lines processing, Virology, Sample processing, Sample shipping, Sample quality control services, Other		0..n
MIABIS-BIOBANK-12	NEW	Quality management standard	Enumerated values	The standards that the Biobank is certified or accredited for. It can be ≥1 of the following values: ISO 20387, ISO 9001, Other	Attribute value list can be expanded by the implementation, and implementation needs to verify provided data	0..n

The table shows the entity-specific attribute code, terminology mapping to the previous Core version, attribute name, value types, attribute descriptions and value lists, any constraints on implementation in databases, and cardinalities on implementation.

Table 2. Attribute Definitions in the Collection Entity

Attribute code	*Mapping to Core v2*	*Attribute name*	*Allowed values*	*Description*	*Constraints*	*Cardinality*
MIABIS-COLLECTION-01	MIABIS-2.0–01	ID	Text	ID of the Collection	The ID is technical and given by the implementation. Implementation will provide instructions on how to form the ID	1
MIABIS-COLLECTION-02	MIABIS-2.0–02	Acronym	Text	Short name in use for the Collection. If applicable		0..1
MIABIS-COLLECTION-03	MIABIS-2.0–03	Name	Text	Name of the Collection (preferably in English)		1
MIABIS-COLLECTION-04	MIABIS-2.0–04	URL	Text	Complete http-address for the Collection		0..1
MIABIS-COLLECTION-05	MIABIS-2.0–07	Contact information	Structured data	Contact information for the contact person/person responsible of the Collection (Structured attribute)		1
MIABIS-COLLECTION-06	MIABIS-2.0–08	Description	Text	Description of the Collection in English	Free text description of the sample collection. Recommendation max. 2000 char.	1
MIABIS-COLLECTION-07	NEW	Sample source	Enumerated values	The source from which the samples were collected or isolated. Can be one of the following values: Human, Animal, Environment		0..1
MIABIS-COLLECTION-08	MIABIS-2.0–09	Sex	Enumerated values	The sex of the individuals in the Collection. Can be ≥1 of the following values: Male, Female, Unknown, Undifferentiated, Not applicable		1..n
MIABIS-COLLECTION-09	MIABIS-2.0–10	Age Low	Integer	Age of youngest sample donor at time of sample donation		0..1 (If applicable)
MIABIS-COLLECTION-10	MIABIS-2.0–11	Age High	Integer	Age of oldest sample donor at time of sample donation		0..1 (If applicable)
MIABIS-COLLECTION-11	MIABIS-2.0–12A	Age Low unit	Enumerated values	Unit defining Age Low. It can be one of the following values: Years, Months, Weeks, Days, Gestational weeks		0..1 (If applicable)
MIABIS-COLLECTION-12	MIABIS-2.0–12B	Age High unit	Enumerated values	Unit defining Age High. It can be one of the following values: Years, Months, Weeks, Days, Gestational weeks		0..1 (If applicable)

(*continued*)

Table 2. (Continued)

Attribute code	*Mapping to Core v2*	*Attribute name*	*Allowed values*	*Description*	*Constraints*	*Cardinality*
MIABIS-COLLECTION-13	MIABIS-2.0–13	Dataset type	Enumerated values	Types of datasets (groups of related data) obtained or otherwise derived from donors or their specimens. It can be ≥1 f the following values: Lifestyle dataset, Environmental dataset, Physiological dataset, Biochemical dataset, Clinical dataset, Psychological dataset, Genomic dataset, Proteomic dataset, Metabolomic dataset, Body (Radiological) image, Whole slide image, Photo image, Genealogical records, Other		0..n
MIABIS-COLLECTION-14	MIABIS-2.0–14	Sample type	Enumerated values	The biospecimen saved from a biological entity for propagation, e.g., testing, diagnostics, treatment or research purposes. It can be one of the following values: Blood, Buffy coat, Cancer cell lines, Digital sample, DNA, Entire body organ, Feces, Embryo or fetal tissue, Immortalized cell lines, Isolated microbes, Other body fluid, Plasma, Primary cells, Postmortem tissue, RNA, Saliva, Serum, Specimen from environment or food, Swab, Tissue (Frozen), Tissue (FFPE), Urine, Other		1
MIABIS-COLLECTION-15	MIABIS-2.0–15	Storage temperature	Enumerated values	The long-term storage temperature at which the sample is stored after preparation, based on SPREC v3.: RT (Room temperature), 2°C to 10°C, −18°C to −35°C, −60°C to −85°C, less than −135°C, Liquid nitrogen vapor-phase, Liquid nitrogen liquid-phase, Other		0..n

(*continued*)

Table 2. (Continued)

Attribute code	*Mapping to Core v2*	*Attribute name*	*Allowed values*	*Description*	*Constraints*	*Cardinality*
MIABIS-COLLECTION-16	MIABIS-2.0–17	Disease	Structured data	The disease or disease category of main interest in the Collection, if any. It can be multiple diseases or disease categories. (Structured attribute)		0..n
MIABIS-COLLECTION-17	NEW	Sample collection setting	Enumerated values	The context in which the sample collection was/is conducted. It can be ≥1 of the following values: Routine health care setting, Clinical trial, Research study, Public health/population based study, Museum and/or archeological collection, Environment, Unknown, Other		0..n
MIABIS-COLLECTION-18	MIABIS-2.0–16	Collection design	Enumerated values	The overall design of the collection that explains how the collection was/is built up. It can be ≥1 of the following values: Case–control, Cross-sectional, Longitudinal cohort, Twin-study, Quality control study, Population-based cohort, Disease-specific cohort, Birth cohort, Microbial collection (if applicable with resistance data), Reference collection, Rare disease collection, Other		0..n
MIABIS-COLLECTION-19	NEW	Use & Access conditions	Enumerated values	The conditions that may change the availability of the samples/data in the collection. It can be ≥1 of the following values: Commercial use, Collaboration, Specific research use, Genetic data use, Outside EU access, Xenograft, Other animal work, Other		0..n
MIABIS-COLLECTION-20	NEW	Collection status	Enumerated values	The state of the Collection functions: Active, Ended, Other		0..1
MIABIS-COLLECTION-21	MIABIS-2.0–20	Total number of subjects	Integer	Total number of subjects included in the Collection		0..1
MIABIS-COLLECTION-22	MIABIS-2.0–22					0..n

(*continued*)

TABLE 2. (CONTINUED)

Attribute code	*Mapping to Core v2*	*Attribute name*	*Allowed values*	*Description*	*Constraints*	*Cardinality*
		Inclusion criteria	Enumerated values	Information on type of parameters that determine which subjects will become Collection participants. Can be several values: Health status, Hospital patient, Use of medication, Gravidity, Age group, Familial status, Sex, Country of residence, Ethnic origin, Population representative sampling, Lifestyle/Exposure, Other		
MIABIS-COLLECTION-23	NEW	Publications	Text	List of key publications produced in the Collection (provide DOI's, if possible)		0..n

The table shows the entity-specific attribute code, terminology mapping to the previous Core version, attribute name, value types, attribute descriptions and value lists, any constraints it may have when implemented in databases, and cardinalities for implementation.

To support the representation of nonhuman samples, we introduced a new "Sample source" attribute to describe the source from which the sample was collected or isolated, e.g., human, animal, or environment. In addition, the value "Specimen from environment or food" was added to the "Sample type" attribute. The possibility to include nonhuman sample collections is making certain attributes optional when implemented (e.g., age) and new null values to attributes, e.g., "Not applicable" value was added to the "Sex" attribute. However, using these new null values results in a loss of information because the reason why the value is null is not known; therefore, the textual collection descriptions in the "Description" attribute become more important. Finally, we renamed the "Total number of participants" attribute to "Total number of subjects" to also allow the description of nonhuman sample collections, and we rephrased the definition of "Inclusion criteria" for the same reason.

To make the entities more generic and to support a variety of different types of sample/data collections, we renamed the "Collection type" attribute to "Collection design" and added a complementary "Sample collection setting" attribute to describe the primary setting for the sample/data collection. The value lists for both attributes were revised and new values added. New attributes include "Publications," as citations in publications are considered a way to convey confidence in the sample provider's credentials and track record.[19] In addition, we added "Collection status" to describe the activity and current state of the collection; this in particular can be important when searching for active collaborators for specific topics. The *Collection* entity also uses the structured attribute "Disease", now allowing for disease categories in addition to specific disease codes (see Supplementary Table S3B). Furthermore, we omitted the attribute "Total number of samples" from the merged *Collection* and *Research resource* attribute list, as it required more active updating by data providers than the other attributes.

Subsequently, after updating the *Collection* and *Research resource* attributes, we acknowledged the need to adjust the *Sample* and *Sample donor* attributes accordingly.

Network entity

Many of the use cases that influenced the MIABIS Core v3 update were research resources operating as networks rather than being stand-alone biobanks or sample/data collections (see Supplementary Table S1). This highlighted the importance of introducing a new dimension into the administrative information presented in MIABIS, and thus, the new entity, *Network*, was added into the Core component. We anticipate that a *Network* can contain not only biobanks but also *Collections* or other *Networks* directly. Thus, *Networks* can also form recursive structures. An example of such is the Belgian BBMRI network BBMRI.be, where the biobanks contain multiple collections, and which then can form networks.[20] Examples of Networks within Networks are BBMRI-ERIC Network consisting of National Node Networks,[21] and RD Connect, a global network for rare diseases, which includes an Italian Telethon Network of Genetic Biobanks in addition to biobanks hosting rare-disease collections.[22,23]

The *Network* attributes are based upon the attributes describing *Biobanks*, as outlined in Table 5. To describe

Table 3. Attribute Definitions in the Research Resource Entity

Attribute code	*Mapping to Core v2*	*Attribute name*	*Allowed values*	*Description*	*Constraints*	*Cardinality*
MIABIS-RESEARCHRE-SOURCE-01	MIABIS-2.0–01	ID	Text	ID of the Research resource	The ID is technical and given by the implementation. Implementation will provide instructions on how to form the ID	1
MIABIS-RESEARCHRE-SOURCE-02	MIABIS-2.0–02	Acronym	Text	Short name in use for the Research resource. If applicable		0..1
MIABIS-RESEARCHRE-SOURCE-03	MIABIS-2.0–03	Name	Text	Name of the Research resource (preferably in English)		1
MIABIS-RESEARCHRE-SOURCE-04	MIABIS-2.0–04	URL	Text	Complete http-address for the Research resource		0..1
MIABIS-RESEARCHRE-SOURCE-05	MIABIS-2.0–07	Contact information	Structured data	Contact information for the contact person/ person responsible of the Research resource (Structured attribute)		1..n
MIABIS-RESEARCHRE-SOURCE-06	MIABIS-2.0–08	Description	Text	Description of the Research resource in English	Free text description of the Research resource. Recommendation maximum 2,000 characters	0..1
MIABIS-RESEARCHRE-SOURCE-07	NEW	Sample source	Enumerated values	The source from which the samples were collected or isolated. It can be one of the following values: Human, Animal, Environment		0..1
MIABIS-RESEARCHRE-SOURCE-08	MIABIS-2.0–09	Sex	Enumerated values	The sex of the individuals in the Research resource. Can be ≥1 of the following values: Male, Female, Unknown, Undifferentiated, Not applicable		1..n
MIABIS-RESEARCHRE-SOURCE-09	MIABIS-2.0–10	Age Low	Integer	Age of youngest sample donor at the time of sample donation		0..1 (If applicable)
MIABIS-RESEARCHRE-SOURCE-10	MIABIS-2.0–11	Age High	Integer	Age of oldest sample donor at the time of sample donation		0..1 (If applicable)
MIABIS-RESEARCHRE-SOURCE-11	MIABIS-2.0–12A	Age Low Unit	Enumerated values	Unit defining Age Low. It can be one of the following values: Years, Months, Weeks, Days, Gestational weeks		0..1 (If applicable)
MIABIS-RESEARCHRE-SOURCE-12	MIABIS-2.0–12B	Age High Unit	Enumerated values	Unit defining Age High. It can be one of the following values: Years, Months, Weeks, Days, Gestational weeks		0..1 (If applicable)

(*continued*)

TABLE 3. (CONTINUED)

Attribute code	*Mapping to Core v2*	*Attribute name*	*Allowed values*	*Description*	*Constraints*	*Cardinality*
MIABIS-RESEARCHRE-SOURCE-13	MIABIS-2.0–13	Dataset type	Enumerated values	Types of datasets (groups of related data) obtained or otherwise derived from donors or their specimens. It can be ≥1 of the following values: Lifestyle dataset, Environmental dataset, Physiological dataset, Biochemical dataset, Clinical dataset, Psychological dataset, Genomic dataset, Proteomic dataset, Metabolomic dataset, Body (Radiological) image, Whole slide image, Photo image, Genealogical records, Other		0..n
MIABIS-RESEARCHRE-SOURCE-14	MIABIS-2.0–14	Sample type	Enumerated values	The biospecimen saved from a biological entity for propagation, e.g., testing, diagnostics, treatment, or research purposes. Can be one of the following values: Blood, Buffy coat, Cancer cell lines, Digital sample, DNA, Entire body organ, Feces, Embryo or fetal tissue, Immortalized cell lines, Isolated microbes, Other body fluid, Plasma, Primary cells, Postmortem tissue, RNA, Saliva, Serum, Specimen from environment or food, Swab, Tissue (Frozen), Tissue (FFPE), Urine, Other		1
MIABIS-RESEARCHRE-SOURCE-15	MIABIS-2.0–15	Storage temperature	Enumerated values	The long-term storage temperature at which the sample is stored after preparation, based on SPREC v3.: RT (Room temperature), 2°C to 10°C, −18°C to −35°C, −60°C to −85°C, less than −135°C, Liquid nitrogen vapor-phase, Liquid nitrogen liquid-phase, Other		0..n
MIABIS-RESEARCHRE-SOURCE-16	MIABIS-2.0–17	Disease	Structured data	The disease or disease category of main interest in the Research resource, if any. Can be multiple diseases or disease categories. (Structured attribute)		0..n
MIABIS-RESEARCHRE-SOURCE-17	NEW	Sample collection setting	Enumerated values	The context in which the sample collection was/is conducted. It can be ≥1 of the following values: Routine health care setting, Clinical trial, Research study, Public health/population based study, Museum and/or archeological collection, Environment, Unknown, Other		0..n

(*continued*)

TABLE 3. (CONTINUED)

Attribute code	*Mapping to Core v2*	*Attribute name*	*Allowed values*	*Description*	*Constraints*	*Cardinality*
MIABIS-RESEARCHRE-SOURCE-18	MIABIS-2.0–16	Research resource design	Enumerated values	The overall design of the collection that explains how the Research resource was/is built up. It can be ≥1 of the following values: Case–control, Cross-sectional, Longitudinal cohort, Twin-study, Quality control study, Population-based cohort, Disease-specific cohort, Birth cohort, Microbial collection (if applicable with resistance data), Reference collection, Rare disease collection, Other		0..n
MIABIS-RESEARCHRE-SOURCE-19	NEW	Use & Access conditions	Enumerated values	The conditions that may change the availability of the samples/data in the collection. It Can be ≥1 of the following values: Commercial use, Collaboration, Specific research use, Genetic data use, Outside EU access, Xenograft, Other animal work, Other		0..n
MIABIS-RESEARCHRE-SOURCE-20	NEW	Research resource status	Enumerated values	The state of the Research resource functions: Active, Ended, Other		0..1
MIABIS-RESEARCHRE-SOURCE-21	MIABIS-2.0–20	Total number of subjects	Integer	Total number of subjects included in the Research resource		0..1
MIABIS-RESEARCHRE-SOURCE-22	MIABIS-2.0–22	Inclusion criteria	Enumerated values	Information on type of parameters that determine which subjects will become Research resource participants. It can be several values: Health status, Hospital patient, Use of medication, Gravidity, Age group, Familial status, Sex, Country of residence, Ethnic origin, Population representative sampling, Lifestyle/Exposure, Other		0..n
MIABIS-RESEARCHRE-SOURCE-23	NEW	Publications	Text	List of key publications produced in the Research resource (provide DOIs, if possible)		0..n

The table shows the entity-specific attribute code, terminology mapping to the previous Core version, attribute name, value types, attribute descriptions and value lists, any constraints it may have when implemented in databases, and cardinalities for implementation

TABLE 4. MAPPING TABLE OF THE SAMPLE ENTITY'S "DETAILED SAMPLE TYPE" VALUES TO MORE AGGREGATED "SAMPLE TYPE" ATTRIBUTE VALUES

Detailed Sample type values	*Mapping to Aggregated Sample type value*
Amniotic fluid	**Other body fluid**
Ascites fluid	**Other body fluid**
Bile	**Other body fluid**
Body cavity fluid	**Other body fluid**
Bone	Tissue (frozen or FFPE)
Bone marrow aspirate	**Other body fluid**
Bone marrow plasma	**Other body fluid**
Bone marrow, whole	**Other body fluid**
Breast milk	**Other body fluid**
Bronchoalveolar lavage	**Other body fluid**
Buffy coat	**Buffy coat**
Cancer cell lines	**Cancer cell lines**
Cerebrospinal fluid	**Other body fluid**
Cord blood	**Other body fluid**
Dental pulp	**Other body fluid**
Digital sample	**Digital sample**
DNA	DNA
Embryo	**Embryo or fetal tissue**
Entire body organ	**Entire body organ**
Feces	Feces
Fetal tissue	**Embryo or fetal tissue**
Fibroblasts	**Primary cells**
Food specimen	**Specimen from environment or food**
Gas, exhaled (=breath)	Other
Gastric fluid	**Other body fluid**
Hair	Other
Immortalized cell lines	Immortalized cell lines
Isolated microbes	**Isolated microbes**
Menstrual blood	**Other body fluid**
Nail	Other
Nasal washing	**Other body fluid**
Pericardial fluid	**Other body fluid**
Peripheral blood mononuclear cells (PBMC)	**Primary cells**
Placenta	**Entire body organ**
Plasma	Plasma
Pleural fluid	**Other body fluid**
Primary cells	**Primary cells**
Post-mortem tissue	**Post-mortem tissue**
Proteins	Other
Red blood cells	**Primary cells**
RNA	RNA
Saliva	Saliva
Semen	**Other body fluid**
Serum	Serum
Specimen from environment	**Specimen from environment or food**
Sputum	**Other body fluid**
Stem cells and iPS cells	Immortalized cell lines
Swab	**Swab**
Sweat	**Other body fluid**
Synovial fluid	**Other body fluid**
Tears	**Other body fluid**
Teeth	**Entire body organ**
Tissue specimen	Tissue (frozen or FFPE)
Umbilical cord	Tissue (frozen or FFPE)
Urine	Urine
Urine sediment	Urine
Vitreous fluid	**Other body fluid**
Whole blood	Blood
Whole blood, dried (e.g. Guthrie cards)	Blood

All the new Aggregated values compared with MIABIS Core v2 are shown in bold.

Table 5. Attribute Definitions in the Network Entity

Attribute code	*Mapping to Core v2*	*Attribute name*	*Value types*	*Description*	*Constraints*	*Cardinality*
MIABIS-NETWORK-01	MIABIS-2.0–01	ID	Text	ID of the Network	The ID is technical and given by the implementation. Implementation will provide instructions on how to form the ID	1
MIABIS-NETWORK-02	MIABIS-2.0–02	Acronym	Text	Short name in use for the Network. If applicable		0..1
MIABIS-NETWORK-03	MIABIS-2.0–03	Name	Text	Name of the Network (preferably in English)		1
MIABIS-NETWORK-04	MIABIS-2.0–04	URL	Text	Complete http-address for the Network		0..1
MIABIS-NETWORK-05	MIABIS-2.0–05	Juristic person	Text	Juristic person, e.g., a university, concern, county council, etc., for the Network		1
MIABIS-NETWORK-06	MIABIS-2.0–06	Country	Text	Two-letter code for the country of the Network	ISO Standard 3166 alpha2. In Network component, multiple values are allowed	1..n
MIABIS-NETWORK-07	MIABIS-2.0–07	Contact Information	Structured data	Contact information for the contact person/ person responsible of the Network (Structured attribute)		1
MIABIS-NETWORK-08	MIABIS-2.0–08	Description	Text	Description of the Network in English	Free text description of the Network. Recommendation maximum 2,000 characters	1
MIABIS-NETWORK-09	NEW	Network status	Enumerated values	The state of the Network functions: Active, Ended, Other		1
MIABIS-NETWORK-10	NEW	Network members	Text	Names of the entities (organizations, biobanks, collections and/or research resources, etc.) involved in the Network.		0..n
MIABIS-NETWORK-11	NEW	Common collaboration topics	Enumerated values	Topics that the Network partners collaborate on. It can be ≥1 of the following values: Common charter, Common SOPs, Common data access policy, Common sample access, policy, Common MTA, Common image access policy, Common image MTA, Common representation, Common URL, Other.		0..n
MIABIS-NETWORK-12	NEW	Network type	Enumerated values	Type or main collaboration area of the Network. It can be one of the following values: BBMRI-ERIC National Node network, Biobank network, Collection network, Disease-specific network, Project network, Rare Disease network, Other		0..1

The table shows the entity-specific attribute code, terminology mapping to the previous Core version, attribute name, value types, attribute descriptions and value lists, any constraints it may have when implemented in databases, and cardinalities for implementation.

Networks more specifically, network collaboration attributes were added to describe the members of a *Network* and the type of collaboration taking place within the *Network*. In addition, the "Network status" attribute allows distinguishing between active and dormant networks, which can be important when trying to find collaborators in specific domains.

Discussion

In the current work, we have succeeded in synchronizing the existing Core attributes with the MIABIS individual-level component *Sample, Sample donor,* and *Event* to make the aggregate-level entities compatible with the individual-level ones. We also adapted and extended the existing Core component to meet the requirements of typical use-cases not only for human samples but also for animal or environmental samples as far as biobanking is concerned. At the same time, we added the new Core entity *Network* to cover an additional organizational level where biobanks often operate.

One of the main goals of this update was to adjust the terminology to better account for data-driven biobanks and digital samples. The working group discussions in this area were complex, but in this revision, we made a first effort to include digital samples and to better represent the different types of data stored in biobanks. It is expected that the dataset types will be expanded over time as biobank samples are used in research resources and projects and transformed into data with better-defined structures. Therefore, further work is needed to describe specific datasets in more detail, e.g., for omics and imaging. We will follow international activities in these areas, e.g., the European Genome-Phenome Archive (EGA),[24] to adhere to their respective data models. In addition, a generic component for datasets to reflect the numerous possibilities of sample use for biomedical data generation will be among the next efforts of the MIABIS working group. As with the current update of the MIABIS Core to version 3, based on the MIABIS individual-level component for *Sample*, *Sample donor,* and *Event*, a new Core update to version 4 may be required to accommodate the work on the dataset component. Focusing our future efforts on addressing the need to describe datasets in more detail will also further facilitate the move toward individual-level data integration in various database implementations such as federated search platforms.

With this work, we updated the semantic interoperability of the MIABIS terminology and improved its actual use and suitability, despite implementation challenges and barriers.

Acknowledgments

The authors thank the members of the MIABIS Core v3 Working Group Andrea Gutierrez, Andres Metspalu, Blazej Marciniak, Christoph Doellinger, Eline Vinke, Enrico Glerean, Heiko Ritzmann, Jens Haberman, Joanna Vella, Jörg Geiger, Kristjan Metsalu, Lisa Scheefhals, Luciano Milanesi, Marije van der Geest, Mario Schattschneider, Mia Liljeström, Michael Hummel, Morris Swertz, Noemi Deppenwiese, Riad Gacem, Ruben Haeck, Sara Casati, Tomas Snäckerström, and Zdenka Dudová for their contributions and support of this work. The authors are also grateful to all the reviewers who provided their valuable expert opinions at various stages of the work

Authorship Contribution Statement

N.E.: Conceptualization, Resources, Investigation, Data Curation, Writing—Original Draft. C.E.: Resources, Investigation, Data Curation, Writing—Review & Editing. M.N.: Investigation, Validation, Writing—Review & Editing. A.S.: Investigation, Validation, Visualization, Writing—Review & Editing. E.V.E.: Resources. R.B.: Resources, Writing—Review & Editing. M.B.: Resources. A.D.: Resources, Writing—Review & Editing. A.V.D.L.: Resources. H.M.: Resources. L.P.: Resources. P.R.Q.: Investigation, Resources. E.U.: Investigation, Resources. P.H.: Resources, Funding Acquisition, Writing—Review & Editing, Supervision. K.S.: Conceptualization, Resources, Writing—Review & Editing, Supervision. G.A.: Writing—Original Draft, Supervision.

Author Disclosure Statement

The authors declare no conflicts of interest.

Funding Statement

This work was supported by BBMRI-ERIC Common Service IT, EU Horizon 2020 supported projects EOSC-Life (grant agreement no. 824087), European Joint Programming for Rare Diseases (EJP RD) (grant agreement no. 825575), and A European-wide foundation to accelerate Data-driven Cancer Research (EOSC4Cancer) (grant agreement no. 101058427).

Supplementary Material

Supplementary Table S1
Supplementary Table S2
Supplementary Table S3

References

1. Norlin L, Fransson MN, Eriksson M, et al. A minimum data set for sharing biobank samples, information, and data: MIABIS. Biopreserv Biobank 2012;10(4):343–348.
2. BBMRI-ERIC. Common service IT. Graz, Austria; 2023. Available from: https://www.bbmri-eric.eu/bbmri-eric/common-service-it [Last accessed: February 20, 2023].
3. Merino-Martinez R, Norlin L, van Enckevort D, et al. Toward global biobank integration by implementation of the minimum information about biobank data sharing (MIABIS 2.0 Core). Biopreserv Biobank 2016;14(4):298–306.
4. Holub P, Swertz M, Reihs R, et al. BBMRI-ERIC directory: 515 biobanks with over 60 million biological samples. Biopreserv Biobank 2016;14(6):559–562.
5. Quinlan PR, Lawrence E, Pourabdolla A, et al. The UK Clinical Research Collaboration (UKCRC) tissue directory and coordination centre: The UK's Centre for facilitating the usage of human samples for medical research. Open J Bioresour 2017;4.
6. T'Joen V, Vaneeckhaute L, Priem S, et al. Rationalized development of a campus-wide cell line dataset for implementation in the biobank LIMS system at bioresource center Ghent. Front Med 2019;6:137.
7. Scapicchio C, Gabelloni M, Forte SM, et al. DICOM-MIABIS integration model for biobanks: A use case of the EU PRIMAGE project. Eur Radiol Exp 2021;5(1):20.

8. Brochhausen M, Zheng J, Birtwell D, et al. OBIB-a novel ontology for biobanking. J Biomed Semantics 2016;7:23.
9. Eklund N, Andrianarisoa NH, van Enckevort E, et al. Extending the minimum information about biobank data sharing terminology to describe samples, sample donors, and events. Biopreserv Biobank 2020;18(3):155–164.
10. European Society of Radiology (ESR). ESR position paper on imaging biobanks. Insights Imaging 2015;6(4):403–410.
11. Kinkorová J, Topolčan O. Biobanks in the era of big data: Objectives, challenges, perspectives, and innovations for predictive, preventive, and personalised medicine. Epma J 2020;11(3):333–341.
12. Antelo V, Salazar C, Martínez A, et al. First release of the Bacterial Biobank of the Urban Environment (BBUE). Microbiol Resour Announc 2018;7(16):e01201-18.
13. Unim B, Haverinen E, Mattei E, et al. Mapping European research networks providing health data: Results from the InfAct Joint Action on health information. Arch Public Health 2022;80(1):23.
14. Schüttler C, Prokosch HU, Hummel M, et al. The journey to establishing an IT-infrastructure within the German Biobank Alliance. PLoS One 2021;16(9):e0257632.
15. Schüttler C, Buschhüter N, Döllinger C, et al. Anforderungen an eine standortübergreifende Biobanken-IT-Infrastruktur. Pathologe 2018;39(4):289–296.
16. International Organization for Standardization. ISO 20387:2018 Biotechnology—Biobanking—General requirements for biobanking. Geneva: ISO; 2018.
17. Ontobee, NCI Thesaurus OBO Edition. Network definition. Ann Arbor, Michigan, USA; 2020. Available from: http://www.ontobee.org/ontology/NCIT?iri=http://purl.obolibrary.org/obo/NCIT_C61377 [Last accessed: February 28, 2020].
18. Betsou F, Bilbao R, Case J, et al. Standard PREanalytical Code Version 3.0. Biopreserv Biobank 2018;16(1):9–12.
19. Lawrence E, Sims J, Gander A, et al. The barriers and motivators to using human tissues for research: The views of UK-based biomedical researchers. Biopreserv Biobank 2020;18(4):266–273.
20. Biobanking and biomolecular resources research infrastructure of Belgium. Brussel, Belgium; 2023. Available from: www.bbmri.be [Last accessed: June 2, 2023].
21. Biobanking and Biomolecular resources Research Infrastructure—European Research Infrastructure Consortium. Graz, Austria; 2023. Available from: www.bbmri-eric.eu [Last accessed: June 2, 2023].
22. RD Connect. RD Connect Project. Barcelona, Spain; 2023. Available from: https://rd-connect.eu [Last accessed: June 2, 2023].
23. Fondazione. Telethon. Telethon Network of Genetic Biobanks. Italy; 2023. Available from: http://biobanknetwork.telethon.it [Last accessed: June 2, 2023].
24. Lappalainen I, Almeida-King J, Kumanduri V, et al. The European genome-phenome archive of human data consented for biomedical research. Nat Genet 2015;47(7):692–695.

Address correspondence to:
Niina Eklund
THL Biobank
Department of Knowledge Brokers
Finnish National Institute for Health and Welfare
PO Box 30, FI-00271 Helsinki
Finland

E-mail: niina.eklund@bbmri-eric.eu