

ADOPT BBMRI-ERIC GRANT AGREEMENT NO 676550

DELIVERABLE REPORT

Deliverable no	D4.4
Deliverable Title	Reports on the access procedure
Contractual delivery month	M42 (March 2019)
Responsible Partner	INSERM
Author(s)	Petr Holub, Saher Maqsood, Romyana Proynova, Robert Reihs, David van Enckevort, Michael Hummel, Erik Steinfelder

REPORTS ON THE ACCESS PROCEDURE

Executive Summary

This report summarizes efforts of BBMRI-ERIC to implement Harmonized Access Procedure (delivered previously as Deliverable D4.1). Once created, the Harmonized Access Procedure was extended into the BBMRI-ERIC 's Access Policy, which was approved and adopted by BBMRI-ERIC Assembly of Members.

IT-assisted implementation of the Harmonized Access Procedure and subsequent Access Policy was delivered by developing and implementing the BBMRI-ERIC Negotiator. The BBMRI-ERIC Negotiator is an IT tool designed to support and monitor access requests, as a direct interface between researcher and biobank.

The BBMRI-ERIC Negotiator works by interfacing findability services, most notably the BBMRI-ERIC Directory (also created under ADOPT, Deliverable D3.1), which allows users to identify biobanks and/or collections that might have relevant biological material or data (or, in certain cases, services) for them. After identifying those candidate biobanks/collections in the BBMRI-ERIC Directory, the user moves to the BBMRI-ERIC Negotiator and submits requests for availability of those for a particular purpose. The system allows also free-text specification of complex requirements, as requests observed in the biobanking world in the recent years have been getting very detailed and complex (e.g., dependencies of timing of treatments and their outcomes).

In addition, the BBMRI-ERIC Negotiator saves the user (researcher) valuable time. The tool has been designed to support group communication—i.e., the user can identify multiple candidate biobanks/collections and submit a single request to all of them. The communication can be both group-based (so that all the request recipients can see it—e.g., to clarify the request or better define purpose) or private (so that only the given biobank/collection and the requester can see the communication). This functionality is designed to increase the efficiency of the requesting process as well as to provide better documentation and/or organisation of requests and responses.



While browsing the BBMRI-ERIC Directory is open to the public with no login, user identification is required to start the access negotiation process. So that both parties in the negotiations have a reasonable guarantee of their counterparts' identities, the BBMRI-ERIC Negotiator interfaces to the BBMRI-ERIC Authentication and Authorization Infrastructure (AAI). BBMRI-ERIC AAI employs eduGAIN as the primary federated source of identity information, but also provides a "hostel" functionality for users whose home institutions do not participate in eduGAIN.

The system has been implemented and put into production at <https://negotiator.bbmri-eric.eu/> and can also be reached from the BBMRI-ERIC Directory once search results appear. In addition, an online helpdesk has been set up.

In terms of the state of the art, a few systems exist that cover a part of the pipeline necessary for users to interact with biobanks. However, all of these systems assume that the target resources are a priori known to the user. In the case of large-scale biobanks featuring a mixture of structured and unstructured personal data enjoying data protection, the user simply cannot be aware of all available resources. The BBMRI-ERIC Negotiator, in contrast, actually identifies these resources even in the presence of only aggregate descriptors. Nonetheless there is potential for integration with a few existing tools for enhanced control of the access procedure.

As a sustainability model, the BBMRI-ERIC Negotiator has been adopted by BBMRI-ERIC Common Service IT in order to have resources for further development. As such, it has been evaluated for usability and a roadmap of additional functionality has been developed, based on user feedback. This includes support for return of data, tools for BBMRI-ERIC National/Organisational Nodes and biobank network representatives to monitor performance of their biobanks, as well as optimization of the user interfaces to ease the addition of new requests. The team developing the BBMRI-ERIC Locator is also separating the BBMRI-ERIC Negotiator into two components: project registration and request issuing.

The report concludes with a summary of performance indicators related to access, where it is shown that the percentage of collections registered in the BBMRI-ERIC Negotiator has remained fairly constant. This is because, based on results of user feedback and testing in the pilot phase, the BBMRI-ERIC Negotiator service was further developed and its integration with BBMRI-ERIC Directory made more visible to the requesters in 2018; in this time period, BBMRI-ERIC did not do any promotion of the service. However, after piloting was completed, the number of requests significantly increased at the beginning of 2019.



COPYRIGHT NOTICE



This work by Parties of the ADOPT BBMRI-ERIC Consortium is licensed under a Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>). This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No. 676550.

DOCUMENT LOG

Issue	Date	Comment	Author
D4.4rev1	2019-03-14	Initial version	Petr Holub
D4.4rev2	2019-09-17	Revised version	Petr Holub, Caitlin Ahern



Glossary

AAI An infrastructure comprising one or more services used to authenticate the user and provide sufficient information for applications to make authorization decisions.

AP BBMRI-ERIC Policy for Access to and Sharing of Biological Samples and Data. Published at http://www.bbmri-eric.eu/wp-content/uploads/AoM_10_8_Access-Policy_FINAL.pdf.

Assembly of Members Assembly of representatives of the member countries of BBMRI-ERIC

BBMRI-ERIC Biobanking and BioMolecular resources Research Infrastructure - European Research Infrastructure Consortium.

BBMRI-ERIC AAI An instance of AAI operated by BBMRI-ERIC to allow consistent authentication and authorization mechanisms across all BBMRI-ERIC services.

BBMRI-ERIC Directory A service provided by BBMRI-ERIC to enable basic findability of biobanks and their collections of samples/data. <https://directory.bbmri-eric.eu/>

BBMRI-ERIC Locator A service co-developed by BBMRI-ERIC and German Biobank Alliance to enable advanced findability of biobanks and their collections based on sample-level and donor-level data. <https://http://search.germanbiobanknode.de/>

BBMRI-ERIC Negotiator A service provided by BBMRI-ERIC to allow effective communication between researchers, requesting samples/data or other services, and biobanks. Subject of this ADOPT Deliverable D4.4. <https://negotiator.bbmri-eric.eu/>

FAIR Findable, Accessible, Interoperable, and Reusable – guiding principles for scientific data management and stewardship [8].

Member EU Member States, third countries as well as intergovernmental organisations may become Members of BBMRI-ERIC

MSC Message Sequence Chart

MTA A contract between the requester and the biobank specifying conditions under which the biological material and/or data is transferred from the biobank to a recipient. Data-only transfer agreement is sometimes called a Data Transfer Agreement (DTA).

National/Organisational Node A National Node or an Organisational Node as defined in the Statutes of BBMRI-ERIC.

Partner Biobank Biobanks participating in BBMRI-ERIC infrastructure as a part of National/Organisational Node that has BBMRI-ERIC Partner Charter signed. Note that biobanks remain in control to ultimately grant/deny access to potential users.

RRI Responsible Research and Innovation



Contents

Background	6
1 Access Policy of BBMRI-ERIC	7
2 Implementation BBMRI-ERIC Access Procedures	8
2.1 Design considerations	8
2.2 Architecture of BBMRI-ERIC Negotiator	11
2.3 Implementation of BBMRI-ERIC Negotiator	17
3 Comparison to State of the Art	19
4 Further Development and Adoption of BBMRI-ERIC Negotiator beyond ADOPT	20
4.1 Further Development of the BBMRI-ERIC Negotiator	20
4.2 Recommendations for BBMRI-ERIC and its National/Organisational Nodes	20
5 Statistics of Access	21
Bibliography	23



Background

This Deliverable describes efforts of BBMRI-ERIC on developing and implementing Access Policy of BBMRI-ERIC across biobanks that are a Partner Biobanks of BBMRI-ERIC as a European Research Infrastructure. It follows the previous conceptualizations of access principles developed as a part of BBMRI Preparatory Phase¹ and follows development of the Harmonized Access Procedure, delivered as Deliverable D4.1 of ADOPT project [1]. This deliverable starts from extending the Harmonized Access Procedure into AP and implementation of the access procedure into a tool called BBMRI-ERIC Negotiator. This tool is used since 2019 as a tool to inquire availability of samples and/or data from BBMRI-ERIC Partner Biobanks.

The original contractual dates were in M24 and M36, but these have been postponed as a part of the amendment, because at that time the implementation of the Access Policy was ongoing and no accesses have been given yet. See also the development of the performance indicators shown in Section 5.

¹ Grant Agreement 212111, 2008–2011.



1 Access Policy of BBMRI-ERIC

The Harmonized Access Procedure, delivered as Deliverable D4.1 of ADOPT project [1], became a starting point for developing AP (technically called “BBMRI-ERIC Policy for Access to and Sharing of Biological Samples and Data”). While the Harmonized Access Procedure focused on the actual process of requesting samples/data and role of different entities (researchers, biobanks, BBMRI-ERIC) in the process, the Access Policy adds information on legal basis and provides also governing ethical principles (Section 4 of the AP) and procedures (Section 5 of the AP). The actual access procedure (Section 6 of the AP) remains then same as defined in the original Harmonized Access Procedure.

The Access Policy was approved by BBMRI-ERIC Assembly of Members on November 8, 2017 and subsequently published on BBMRI-ERIC website.²

² http://www.bbmri-eric.eu/wp-content/uploads/AoM_10_8_Access-Policy_FINAL.pdf



2 Implementation BBMRI-ERIC Access Procedures

As anticipated already in the Harmonized Access Procedure Deliverable, the access procedure was supposed to be IT-assisted, with focus on enabling requests involving small number as well as large number of biobanks.

ADOPT project was the training ground to develop Access-IT tool fostering a FAIR Access. This meant the progressive harmonization of shared codes for the biobanked samples as fundamental condition. Meanwhile shaping the BBMRI-ERIC Negotiator as a learning-tool introduced a critical, practical Responsible Research and Innovation (RRI) milestone, de facto enriching, enhancing the resource, tracking the interaction between biobank and researchers, systematically considering request and returning results as evolving, key information directly related to sample/data.

From the high-level perspective, the access pipeline can be visualized as shown in Figures 1 and 2. As the first step, the requester needs to identify a set of potentially relevant biobanks to further inquire. This is done based on search using BBMRI-ERIC Directory or BBMRI-ERIC Locator services using structured search criteria. Once potentially relevant biobanks are identified, the requester selects to move to the BBMRI-ERIC Negotiator, at which point she needs to reveal the research project or research project proposal and she can also specify additional requirements on the samples/data (as a free text, hence supporting also very complex specifications). At the end of the process in the BBMRI-ERIC Negotiator, the user gets availability information from the biobanks and proceeds with them individually. The Access Procedure also works with the concept of return of data, where the researcher should offer the data she generates back to the biobank to enrich the value of the resource. This step is conceptually envisaged to be also supported by the BBMRI-ERIC Negotiator, but not yet practically implemented. After the researcher obtains availability information, she proceeds directly with the biobanks she chooses, signing MTA and have samples/data delivered.

The structured search options using BBMRI-ERIC Directory and BBMRI-ERIC Locator differ primarily in the depth of data that can be queried. While the BBMRI-ERIC Directory works on the level of highly aggregate data only and has relatively low participation barrier for any biobank, hence assuming all relevant biobanks having been registered in it, the BBMRI-ERIC Locator works on querying donor-level and sample-level data harmonized to a common data model. BBMRI-ERIC Locator features federated querying architecture, where participating biobanks run their own instance of a Connector, which communicates with the Locator, processes obtained requests and returns statistics (counts) of matching donors and samples. Architecture and implementation of BBMRI-ERIC Locator is subject to ADOPT Deliverable D3.5. Because of much higher demands on participating biobanks, both in running a Connector component but namely because of need to obtain in-depth structured data on donors and samples upfront and harmonize it to the target data model, not all the biobanks are anticipated to participate in the BBMRI-ERIC Locator.

2.1 Design considerations

The following design principles have been defined based on real-world considerations:



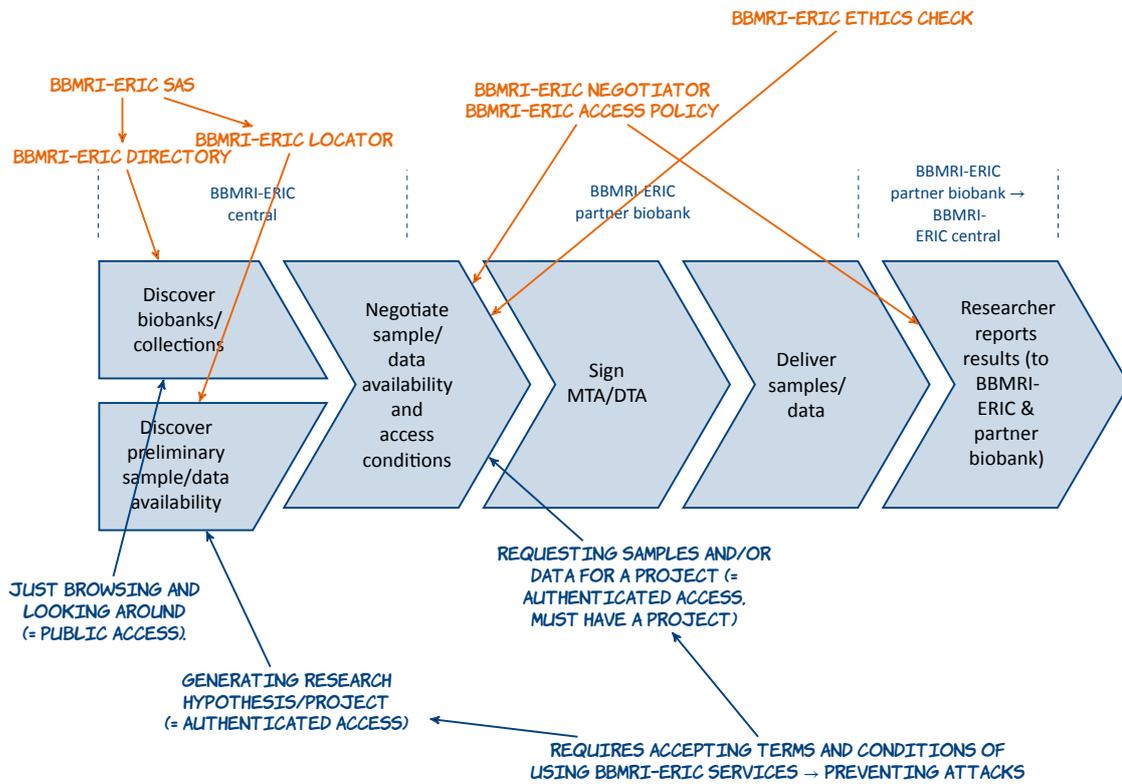


Figure 1: High-level overview of access pipeline.

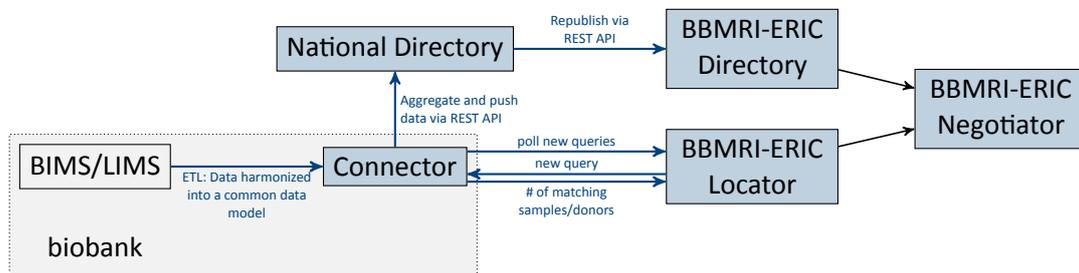


Figure 2: High-level perspective on IT components implementing the access pipeline. Thick blue arrows are subject to standardization in BBMRI-ERIC Interoperability Forum and/or MIABIS.



D-1 System must work in presence of approximative aggregate data in the source data catalogs, without access to sample-level or donor-level information.

Extensive mining of structured data from source systems, namely from clinical hospital information systems relying heavily on unstructured data maintained for diagnostic/therapeutic purposes, is extremely resource demanding; particularly so if high accuracy and reliability of the extracted data is needed.

Hence the designed system must be able to tolerate absence of detailed data, has to be able to provide access even to mixture of biobank having in-depth structured data readily available as well as those biobanks that have only approximate aggregate descriptors available about collections of their samples/data.

D-2 Approximative data in source data catalogs should approximate upper bound of what is available.

In case of approximating/estimating aggregate descriptors, one must decide if the approximation of the data is done toward minimizing advertisements (lower bound = being conservative) or maximizing advertisements (upper bound = being optimistic) in the source data catalogs. The design principle says that the upper bound should be used in the source catalogs. The reason for this is based on several observations: (a) if the biobanks advertise only lower bound, the uncertain part of the collections are rendered invisible for any access requests and can't be clarified in the subsequent steps of access negotiation; (b) aggregate descriptors, like the descriptors based on MIABIS 2.0 Core in the BBMRI-ERIC Directory, may not allow true "AND queries" on the level of donors or samples³ [4, 6] and thus the resulting data structure naturally approaches the upper bound; (c) even if the right combination of samples with the requested parameters is available, it may render inaccessible for other reason, e.g., quality not being sufficient or informed consent not being compatible with the purpose for which the samples/data are requested; hence this filtering out of potentially available collections or samples might occur anyway.

D-3 System must be able to take advantage of sample-level or donor-level data in the source data catalogs.

The sample-level catalogs such as RD-Connect Sample Catalogue⁴ [3] or federated querying systems having access to sample-level or donor-level data inside the biobanks, such as BBMRI-ERIC Locator, provide much more accurate estimates of sample or donor availability. Hence these should be able to interface to the access support systems, and ideally the requester should be able to prioritize results from them.

D-4 The whole access pipeline must be able to cope with the mixture of structured and unstructured search criteria.

While the structured search criteria is the basis for identifying candidate biobanks in the initial step, the unstructured data often contain very critical information that is hard to be expressed in a structured way: e.g., intended analytic methods for samples, or complex requirements on patients phenotype combined with clinical treatment regimes and outcomes.

Another example of unstructured criteria is description of the project/purpose, to which the informed consent or other legal basis for processing personal data must be mapped in the source

³ Absence of true "AND queries" stems from the flat composition of attributes describing Collections [of samples and/or data]. E.g., collection with disease set = C19, C20 and material type set = Blood, DNA does not mean, there are all possible combinations of diseases and material available. Hence a query for C19 and Blood may identify a biobank, which does not have exactly this combination, despite the fact that they *must* have samples from patients with C19 and they *must* have Blood samples.

⁴ <https://samples.rd-connect.eu/>



data, in order to make sure the purpose and the legal basis are compatible. There are efforts to make access conditions at least partially structured and effectively machine readable as as ADA-M [9] and Data Use Ontology⁵ (DUO) [2], there are many subtle variations that are hard to express using those.

D-5 System must allow negotiating with large number of biobanks at the same time, while minimizing burden on both requesting researchers and bioankers.

Because of low sensitivity of search using aggregate descriptors in catalogs like BBMRI-ERIC Directory, the search can result in large number of candidate biobanks and their collections. Therefore it is important that the process poses least possible burden on dealing with many resources both on requesters and biobankers.

D-6 System must minimize adoption barrier as possible for both researchers and biobankers.

This requirement equally applies to academic users, users coming from health care and users coming from industry. For biobankers, it also means that adding their collections to the access ecosystem in the simplest possible mode (via aggregated information only) should be also low-barrier process.

D-7 System must guarantee authenticity of requests and responses to requests.

This requires that both the requests and the reactions of biobanks come from authenticated users, albeit the authentication process and namely the initial registration process might increase the barrier of adoption.

D-8 System should generate positive feedback loop (incentives) for improving the data quality in the source catalogs.

The biobanks are typically hesitant to invest a lot of effort in providing “perfect” (accurate and very detailed) data available upfront in the catalogs. The access process should provide them with additional incentives: if they provide too conservative data, they are likely to be omitted from the search results of candidate biobanks; if they exceed the upper bound substantially, they might get a lot of irrelevant requests. The earlier situation is self-modulating, as most biobanks have to demonstrate utilization of their resources to their funders. The latter situation is partially self-motivating for the biobanks as receiving too many irrelevant requests increases their workload on filtering those. But it can be also visible in the metrics of the biobank performance (too high rejection to acceptance ratio, possibly also too high time to react to the request).

D-9 The actual decision on sample/data release and delivery of these to the requester is up to direct communication between the requester and the biobank.

This principle follows the fact that biobanks enjoy rather high level of independence within the federated architecture of BBMRI-ERIC and this principle is also codified in AP of BBMRI-ERIC.

2.2 Architecture of BBMRI-ERIC Negotiator

The main purpose of the BBMRI-ERIC Negotiator is to implement steps 1–2 of the AP: to help the researcher to identify biobanks that have samples and/or data relevant for the request.

⁵ <https://ega-archive.org/data-use-conditions>



The BBMRI-ERIC Negotiator is a relatively simple centralized service implementing a web-based communication system with a specific workflow of communication. It allows three modes of communication:

- (a) researcher to file a request;
- (b) biobank representatives to communicate to the researcher about the request so that the communication is visible also to the other biobanks participating in the same request; i
- (c) biobank representatives to communicate confidentially with the requester sensitive information such as availability information and access conditions (cost of access, etc.).

Based on the MIABIS 2.0 Core and Directory data models, the basic unit used for identifying relevant resources available in biobanks to be negotiated with, are *collections* for several reasons. While the biobank entity in MIABIS represents the institutional envelope, the collection entity represents collection of samples/data. Hence when searching for samples and/or data, it is the collections that are identified and it should be known to biobank representatives, why they were identified as potential candidates for the given request. Collections can also form hierarchies, hence allowing biobanks to refine aggregate descriptors of available samples/data, allowing for more specific search results. Second, in the BBMRI-ERIC Directory data model, the collections might have separate contact information, overriding the contact of the hosting biobank; this is because some biobanks act as number of individual collections hosted under the responsibility of the same institutional biobank. However, *we use term "biobanker" also for collection representatives in this paper for clarity and consistency reasons.*

Request workflow. This access process will consists of the following steps from the requester's perspective, as also shown graphically in the Figures 3 and 7:

1. **Registration and authentication of the user** in order to ensure that biobanks have a trusted communication partners, ideally also with trusted institutional affiliations (for liability reasons). This is implemented by interfacing to the BBMRI-ERIC AAI as described in the paragraph below.
2. **Selecting candidate collections in the source system** that might have samples and/or data for the particular request, using search in BBMRI-ERIC Directory or BBMRI-ERIC Locator (or possibly other external catalogs such RD-Connect Finder or Catalogue [3]). This is achieved using structured search criteria, which are dependent on the particular source system – typically it allows searching for diseases, material types, gender/age groups, etc. These search criteria are then serialized into a URL that can be used by the BBMRI-ERIC Negotiator to call the system back, when allowing the user to change the structured search criteria, as well as simple textual representation to be shown for the user in the BBMRI-ERIC Negotiator as a part of the request. At the end of this step, the user is transferred to the BBMRI-ERIC Negotiator.
3. **File the request**, combining structured search criteria (from the previous step) with unstructured data about requirements on the samples/data. Possibility to introduce any free-text description of the request stems from practical experience that the requests are becoming increasingly complex (e.g., complex inclusion criteria for research participants contributing their samples/data, requirements on particular treatment types and results of treatments) and introducing



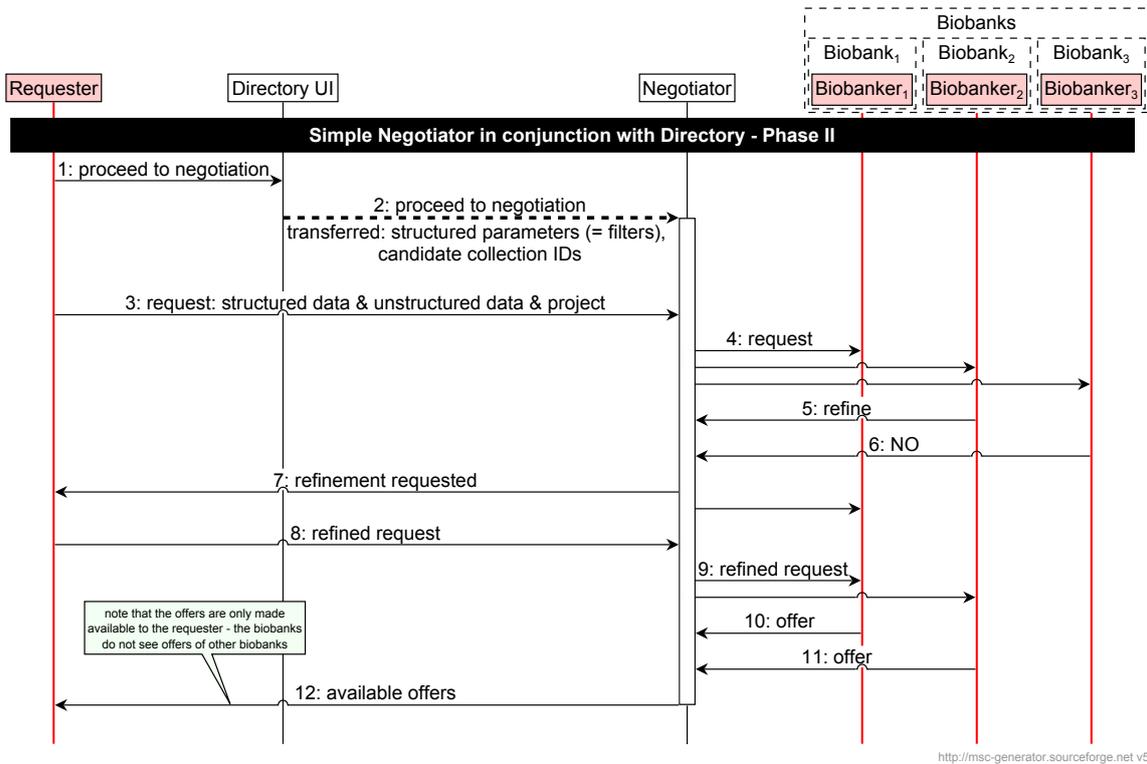


Figure 3: Overview of user interaction with BBMRI-ERIC Negotiator captured using a MSC. Specific example of implementation is shown in Figure 7.



domain-specific languages to query those properties in structured manner would impair usability of the system for people who are not experts in IT. The request also contains information about the project and if ethics vote is available; this information is necessary for biobanks to decide if human biological material and/or data can be used for a particular purpose and if they need to perform their own internal ethics vote.

4. **Notification to biobankers** for the new request. Notifications are sent via emails to all the registered representatives of the matching collections for the given request. This uses contact information obtained from the BBMRI-ERIC AAI.
5. **Biobankers review the request**, possibly asking for additional information from the researchers if needed in order to make decision on what samples/data is available and can be meaningfully used for the given purpose. There are two ways how the biobankers can interact with the requester: group communication (“public communication” is used as the term) visible to all the biobanks and collection representatives within the same request, or private biobank–requester communication. The advantage of the group communication is that if there is any generic clarification needed for the request, it is sufficient if one biobank asks and the requester refines the request accordingly. This can substantially speed up the process of negotiation.
6. **Refine the request** based on the feedback from biobanks to the point that biobanks are able to decide if they have relevant samples/data.
7. **Obtain availability information** from the biobanks: what samples and/or data are available for the given purpose and under what conditions (joint project, fee based on cost-recovery, etc.). This is typically communicated via the private biobank–requester communication in the Negotiator.

Based on the obtained availability information, the user can then directly interact with one or more biobanks s/he selects—this involves getting into a contractual relation with them by signing MTAs. This direct interaction follows Design Principle **D-9**.

Detailed technical communications schema of the BBMRI-ERIC Negotiator with other components of BBMRI-ERIC IT ecosystem is shown in Figure 4. We discuss these in the individual paragraphs below.

Authentication and authorization. The BBMRI-ERIC Negotiator relies on BBMRI-ERIC AAI to perform authentication and provide sufficient information for authorization decisions. Prior to filing a request or accessing existing requests in the BBMRI-ERIC Negotiator, all users must be authenticated, thus complying with Design Principle **D-7**. BBMRI-ERIC AAI relies on eduGAIN as the primary federated source of identity information, but also provides a hostel functionality for users whose home institutions are not participating in eduGAIN (typical for users coming from health care or industrial R&D), hence also complying with Design Principle **D-6**. BBMRI-ERIC AAI contains information about institutional affiliation of the users in case that they authenticate via their home organization identity provider. It also uses groups (and resulting user attributes) to maintain information on which user represents which collections; pushing this information into the BBMRI-ERIC Negotiator allows to make a decision of a user is “only” a researcher, a biobank/collection representative, or combination of both.



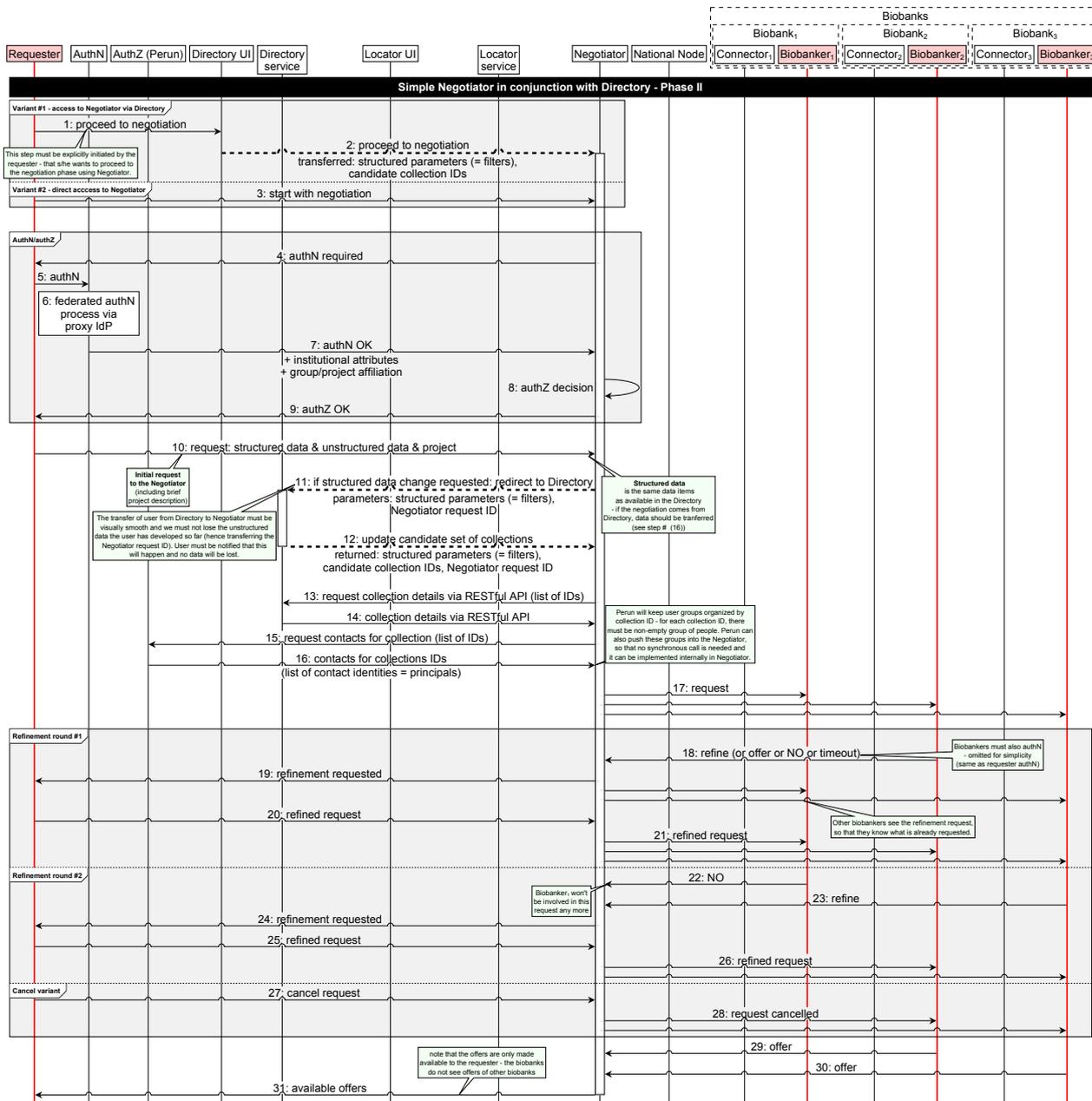


Figure 4: Technical overview of interaction between BBMRI-ERIC Directory, Authentication and Authorization Infrastructure (marked as AuthN and AuthZ respectively), and BBMRI-ERIC Negotiator captured using a MSC.



Integration with BBMRI-ERIC Directory. BBMRI-ERIC Negotiator has been integrated with BBMRI-ERIC Directory since the beginning as the main source of information on available collections of samples/data. Both BBMRI-ERIC Negotiator and BBMRI-ERIC AAI retrieve list of collections and their representatives from the BBMRI-ERIC Directory.

Structured search criteria can be used with any attribute of Collections and it is possible to transition to the BBMRI-ERIC Negotiator from both the default simple user interface as well as the advanced user interface of the BBMRI-ERIC Directory. The screenshot of the BBMRI-ERIC Directory including the button to transfer to the BBMRI-ERIC Negotiator is shown in the Figure 5.

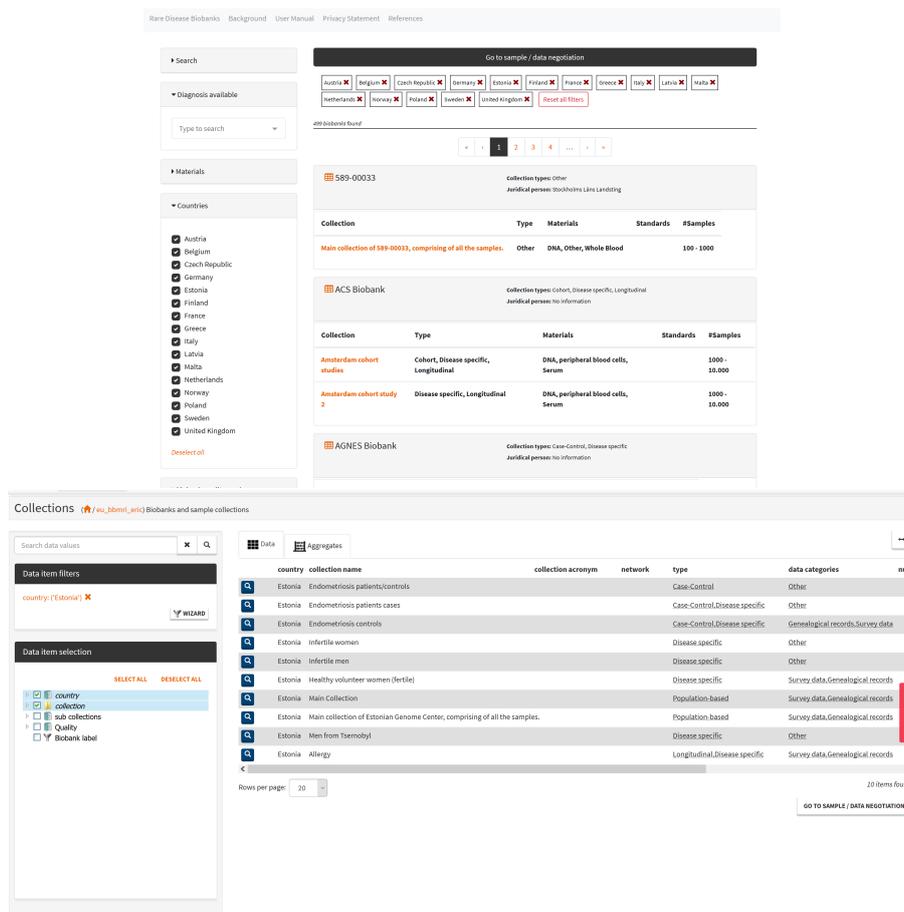


Figure 5: BBMRI-ERIC Directory user interfaces (simple above, advanced below) showing the buttons to go to the BBMRI-ERIC Negotiator (for simple it's on the top, for advanced it is at the bottom).

2.3 Implementation of BBMRI-ERIC Negotiator

The Negotiator has been implemented using Java technology with common frameworks and is available as open-source in a public source-code repository.⁶ The actual service is operated at <https://negotiator.bbmri-eric.eu/>. The service is also reachable from within BBMRI-ERIC Directory, allowing to identify candidate biobanks first and then proceed to the BBMRI-ERIC Negotiator (after which login is required to complete subsequent steps). The service has its helpdesk line available via Request Tracker operated by BBMRI-ERIC at <negotiator@helpdesk.bbmri-eric.eu>.

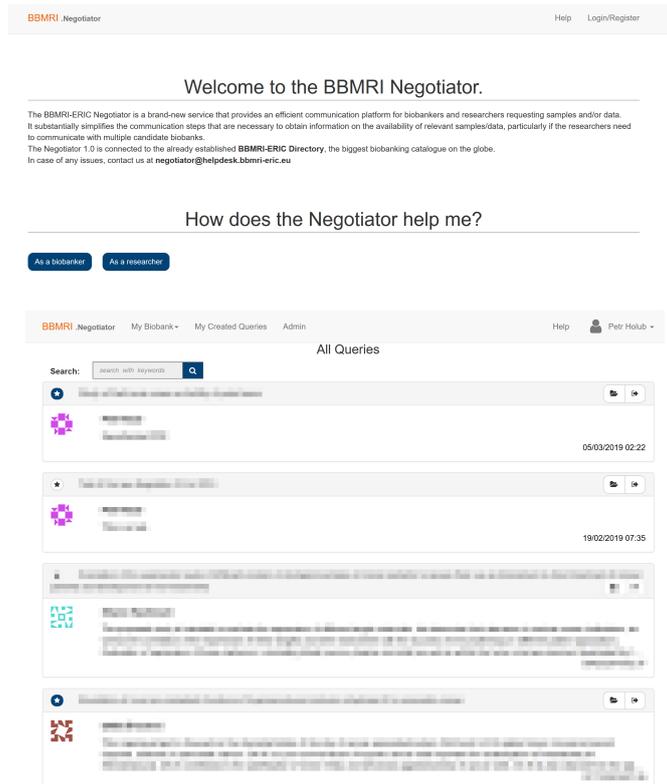


Figure 6: Showcase of the BBMRI-ERIC Negotiator user interface. Note that the actual requests have been blurred in order to stick to the confidentiality required by the AP.

⁶ <https://bitbucket.org/medicalinformatics/>



Search: C00-C75 - Malignant neoplasms, stated or presumed to be primary, of specified sites, except of lymphoid, haematopoietic and related tissue X
 C00-C97 - Malignant neoplasms X C18 - Malignant neoplasm of colon X Certified by accredited body X Reset all filters

11 biobanks found

REQUEST SAMPLES

1 2 >

BioBanca Istituzionale Bari
 ISO-9001 ✓
 Collection types: Disease specific, Hospital, Sample collection
 Juridical person: IRCCS Istituto Tumori "Giovanni Paolo II" Bari

Collection	Type	Materials	Standards	#Samples
Breast Cancer Serum with Mammographic Characteristics	Disease specific, Hospital	Serum		1500
	Hospital	Tissue (frozen)		
Breast Cancer pathological and non affected tissues with 5 to 10 years follow up from 3000 patients	Disease specific	Tissue (frozen)		7500
Gastro Intestinal Cancer Pathological and non affected tissues from 2000 to 2013 with 5 to 10 years follow up from 1600 patients	Disease specific, Hospital, Sample collection	Tissue (frozen)		3660
Gynecological Cancer Pathological and non affected tissues from 2000 to 2013 with 5 to 10 years follow up from 567 patients	Disease specific, Hospital, Sample collection	Tissue (frozen)		1219
Gastro Intestinal Cancer with synchronous multiple samples	Disease specific, Hospital, Sample collection	Plasma, Serum, Tissue (frozen)		908
Gynecological Cancer with synchronous multiple samples	Disease specific, Hospital, Sample collection	Plasma, Serum, Tissue (frozen)		167

Biobank Graz
 ISO-9001 ✓
 Collection types: Population-based, Disease specific
 Juridical person: Medical University of Graz

Collection	Type	Materials	Standards	#Samples
Colorectal Cancer Collection	Population-based, Disease specific	Tissue (frozen), Serum, Tissue (paraffin preserved)		169815
Fluids Collection	Population-based, Disease specific	Serum, Whole Blood, Other, Urine, Plasma		1173482

Biobanque Hôpital Erasme-ULB
 ISO-15189 ✓
 Collection types: Case-Control, Disease specific, Longitudinal, Quality control, Twin-study
 Juridical person: Hôpital Erasme

Query details

On this page you can provide the details for your query.
 First define the search parameters for the samples you are looking for. Click the "Edit Search Parameters" button to change them. If you like to add additional search queries, for the same catalogue or an other catalogue use the "Add new Search Query" button to be redirected to one of the catalogues.
 Then fill out the form below, giving your query a title and describing both the project you are working on and adding more detailed information for this request in particular.
 If you have additional information in PDF files available, you can attach them to this query.

BBMRI_Negotiator My Biobank - My Created Queries Admin Help

BBMRI-ERIC Directory selected disease types are C18 - Malignant neoplasm of colon,C00-C75 - Malignant neoplasms, stated or presumed to be primary, of specified sites, except of lymphoid, haematopoietic and related tissue,C00-C97 - Malignant neoplasms 22 Edit Search Parameters

Add new Search Query

* Title

* Project description

* Request description

Ethics Note

Attachments (only PDF files allowed) Browse Upload

Add query

Figure 7: User interface of the system when filing a new request into the Negotiator. The first step is selection of relevant biobanks in the Directory based on structured data search (in this case third-party accredited biobanks hosting C.18 material/data), followed by actual entry of the request into the Negotiator.



3 Comparison to State of the Art

There are at several systems implementing parts of access pipeline suitable for biobanks and data repositories. One is REMS [5], Resource Entitlement Management System developed by CSC Finland,⁷ which focuses on the workflow of approval process of data or sample release. It allows to implement complex processes based on the source data or sample sets, whose approval is required to proceed into the next step of the decision process. Another system is Podium⁸ [7], developed by The Hyve⁹ and BBMRI.nl, with primary focus on sample release approval and delivery process, including tracking the actual sample shipment. Yet another system is ARIA¹⁰ developed by Instruct Research Infrastructure, with primary focus on management of service or technology on-site access visits and remote access to a service or a technology.

All of these systems come from one fundamentally different design principle, differing from Design Principle **D-1**: they have assumption that the set of resources (samples, data sets, or services/technologies) are a priori known and access to a subset of these resources is requested. This is, however, not realistic in the near- to mid-term future (if ever at all) in case of large-scale biobanks featuring mixture of structured and unstructured personal data enjoying data protection. Browsing of sample catalogs is indeed available in specific cases (e.g., in rare disease domain as witnessed by RD Sample Catalogue or TELETHON access system), but it is not generally available and even these catalogs provide only limited depth of information that can be used as search criteria. BBMRI-ERIC Negotiator, however, focuses on actually identifying these resources even in the presence of only aggregate descriptors available; it allows biobanks to dynamically create sample/data sets based on the mixture of requested structured and unstructured search criteria, and provide availability information on this set to the requester. Because of the Access Policy, the actual approval of release, signing MTA and delivery of samples/data is left to the biobanks.

However, systems like Podium could be connected to the BBMRI-ERIC Negotiator conceptually, if the biobanks are interested in having a common system for approval of sample/data release and tracking of delivery. There are even tentative discussions between BBMRI-ERIC and BBMRI.nl on integrating these tools. Similarly, BBMRI-ERIC Negotiator works together with ARIA for supporting trans Research Infrastructure requests within CORBEL Project and it is expected to continue and further optimize this process in EOSC-Life project. In these requests, the BBMRI-ERIC Negotiator can be used either prior to filing a request into ARIA, allowing to identify specific biobanks/collections for further collaboration, but in some cases the request filed into ARIA actually assumes the use of BBMRI-ERIC Negotiator as a next step once the request is approved; the latter approach poses a inherent risk that after the approval of the project, the user may discover that there are no biobanks/collections to provide samples/data for the project.

⁷ <https://www.csc.fi/>

⁸ <https://podium.bbMRI.nl/>

⁹ <https://thehyve.nl/>

¹⁰ <https://www.structuralbiology.eu/help/about-aria>



4 Further Development and Adoption of BBMRI-ERIC Negotiator beyond ADOPT

After being developed within this project, the BBMRI-ERIC Negotiator has become a key component of BBMRI-ERIC functionality as a European research infrastructure, mission of which is to facilitate access to the quality-defined biological material and data.

4.1 Further Development of the BBMRI-ERIC Negotiator

After the end of the ADOPT funding for the BBMRI-ERIC Negotiator, further development has been picked up by BBMRI-ERIC Common Service IT. As such, it has been evaluated for usability and also based on the users' feedback, there is a roadmap of additional functionality developed. This includes support for return of data, tools for BBMRI-ERIC National Nodes and biobank network representatives to monitor performance of their biobanks, as well as optimization of the user interfaces to make it more efficient to insert new requests. The team developing BBMRI-ERIC Locator is also implementing separation of the BBMRI-ERIC Negotiator into two components: project registration and request issuing. Thus it would be more effective in the situation with multiple requests per project, where the requester (or a team) files a project and then without duplication of information can submit many requests. Recently it has been extended with the support for having multiple different source catalogs as a part of the RD-Connect project, so that it can be integrated with the two main RD-Connect cataloging tools: Registry & Biobank Finder¹¹ and Sample Catalogue¹² [3].

4.2 Recommendations for BBMRI-ERIC and its National/Organisational Nodes

At the end of the ADOPT project, 113 representatives were registered in the Negotiator to represent 271 collections (16.2% of collections available in the BBMRI-ERIC Directory). The main goal of BBMRI-ERIC is to keep onboarding additional representatives, so that there is sufficient number of service providers actively responding to the requests filed by the requesters into the BBMRI-ERIC Negotiator. BBMRI-ERIC sees it as one of the main roles of its National/Organisational Nodes, to support adoption of the core IT tools nationally. BBMRI-ERIC has a commitment to support this adoption process by helping to develop training materials and to keep developing the platform based on the received user feedback. Both BBMRI-ERIC and its National/Organisational Nodes shall work on promoting the adoption of the tools by the requesters; this should be namely done by public relations aimed at relevant communities (scientific conferences and meetings in the fields utilizing biological material and/or data from the biobanks), advertisements in the scientific media, as well as use of social media and other supportive means.

¹¹ <http://catalogue.rd-connect.eu/>

¹² <https://samples.rd-connect.eu/>



5 Statistics of Access

Selected key performance indicators related to the performance of the performance of BBMRI-ERIC Negotiator and related tools has been collected in the pilot phase in 2018 as well as after moving into production in 2019 are shown in Table 2.

Month	Collections in Directory	Biobanks in Directory	Registered collections in Negotiator	Number of new requests	Total number of requests
Jan-18	1,429	531	16.10%	0	0
Feb-18	1,429	531	19.40%	1	1
Mar-18	1,435	531	20.10%	0	1
Apr-18	1,435	531	20.30%	1	2
May-18	1,480	542	21.10%	2	4
Jun-18	1,488	545	21.00%	3	7
Jul-18	1,488	545	23.50%	0	7
Aug-18	1,497	546	24.50%	0	7
Sep-18	1,543	550	23.80%	0	7
Oct-18	1,538	551	23.30%	0	7
Nov-18	1,517	554	23.70%	2	9
Dec-18	1,524	558	23.60%	0	9
Jan-19	1,524	555	23.60%	3	12
Feb-19	1,665	578	22.80%	8	20

Table 2: Development of key performance indicators on access over the time.

Collections in Directory is a number of collections that are available in the BBMRI-ERIC Directory.

Discussion of results: One can see a gradual increase in the number of collections, which corresponds to two effects: number of new resources being added (which can be also observed by a number of biobanks increasing) and having more refined aggregate descriptors of available samples/data in the BBMRI-ERIC Directory.

Biobanks in Directory is a number of biobanks represented in the BBMRI-ERIC Directory

Discussion of results: While the number is gradually increasing in the 2018–2019, one can see two counteracting effects: new resources are being added, which increases the number, and multiple smaller biobanks or standalone collections being aggregated into bigger institutional biobanks. The latter effect can result in drop of the number, while the actual service provided is typically improved, as the resulting bigger biobank has also bigger pool of resources to perform its services.

Registered collections in Negotiator is a percentage of collections, which have at least one representative registered in the BBMRI-ERIC AAI – and thus this information being also available in the BBMRI-ERIC Negotiator.



Discussion of results: After the initial on-boarding period in the piloting in the Q1–Q2/2018, one can see almost no increase in Q3–Q4. The reason is that based on results of user feedback and testing in the pilot phase, the BBMRI-ERIC Negotiator service was further developed and its integration with BBMRI-ERIC Directory made more visible to the requesters, and BBMRI-ERIC did not do any promotion of the service in this period. New activity started in Q1/2019, after updated BBMRI-ERIC Negotiator and BBMRI-ERIC Directory were deployed.

Number of new requests/Total number of requests indicates the total number of real researchers requests, excluding any test requests issued by either requesters or developers.

Discussion of results: For the same reasons discussed in the previous performance indicator, after piloting in 2018, the actual move of the service into production resulted in very significant growth of requests issued in 2019.



Bibliography

- [1] G. Dagher, P. Holub, M. Hummel, A. Jalanko, O. Törnwall, K. Silander, M. Lavitrano, K. Zatloukal, M. Hansson, M. T. Mayrhofer, M. Mendy, P. Quinlan, and I. Schlünder. *Harmonised Access Procedure to Samples and Data*. ADOPT BBMRI-ERIC Deliverable D4.1. Sept. 2016. DOI: 10.5281/zenodo.823013. URL: <https://doi.org/10.5281/zenodo.823013>.
- [2] S. O. Dyke, A. A. Philippakis, J. R. De Argila, D. N. Paltoo, E. S. Luetkemeier, B. M. Knoppers, A. J. Brookes, J. D. Spalding, M. Thompson, M. Roos, et al. "Consent codes: upholding standard data use conditions". In: *PLoS genetics* 12.1 (2016), e1005772.
- [3] S. Gainotti, P. Torreri, C. M. Wang, R. Reihs, H. Mueller, E. Heslop, M. Roos, D. M. Badowska, F. Paulis, Y. Kodra, et al. "The RD-Connect Registry & Biobank Finder: a tool for sharing aggregated data and metadata among rare disease researchers". In: *European Journal of Human Genetics* 26.5 (2018), p. 631.
- [4] P. Holub, M. Swertz, R. Reihs, D. van Enckevort, H. Müller, and J.-E. Litton. "BBMRI-ERIC directory: 515 biobanks with over 60 million biological samples". In: *Biopreservation and biobanking* 14.6 (2016), pp. 559–562.
- [5] M. Linden, T. Nyrönen, and I. Lappalainen. "Resource entitlement management system". In: (2013). URL: https://www.immagic.com/eLibrary/SOURCE/TERNA_NL/T130606C/LINDEN.docx.
- [6] R. Merino-Martinez, L. Norlin, D. van Enckevort, G. Anton, S. Schuffenhauer, K. Silander, L. Mook, P. Holub, R. Bild, M. Swertz, et al. "Toward global biobank integration by implementation of the minimum information about biobank data sharing (MIABIS 2.0 Core)". In: *Biopreservation and biobanking* 14.4 (2016), pp. 298–306.
- [7] A. Moors, S. Schuur, J. Lynch, D. Ellisor, J. Ness, J. Trevillian, R. Pugh, W. Schleif, S. McCall, M. Datto, et al. "The International Society for Biological and Environmental Repositories Presents Abstracts from Its Annual Meeting Thinking BIG in TEXAS: Seizing BIG Opportunities in Biobanking Through Data, Collaboration, and Innovation May 20–24, 2018 Dallas, Texas". In: *Biopreservation and Biobanking* 16.3 (2018), A–1.
- [8] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, et al. "The FAIR Guiding Principles for scientific data management and stewardship". In: *Scientific data* 3 (2016).
- [9] J. P. Woolley, E. Kirby, J. Leslie, F. Jeanson, M. N. Cabili, G. Rushton, J. G. Hazard, V. Ladas, C. D. Veal, S. J. Gibson, et al. "Responsible sharing of biomedical data and biospecimens via the "Automatable Discovery and Access Matrix"(ADA-M)". In: *NPJ Genomic Medicine* 3.1 (2018), p. 17.

