OPEN U

BLOOM

European hub for online & blended learning, mobility & networking

Co-funded by the Erasmus+ Programme of the European Union

BLOOM hub Technical Specifications

September 2022

BLOOM hub Technical Specifications

This report is the deliverable of D2.6 of the OpenU project
By J. Posel, Freie Universität Berlin

Co-funded by the Erasmus+ Programme of the European Union

Disclaimer: Funded by the European Commission in the framework of the Erasmus+ co-funded project OpenU (606692-EPP-1-2018-2-FR-EPPKAC-PI-POLICY).
The European Commission support for the production of this publication does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Table of contents

1. Overview

The OpenU project's Work Package 2 assembled partners in the design and implementation of this European Hub, named the BLOOM Hub. For this European digital infrastructure aimed at both higher education institutions, professionals, and students alike, we have chosen a modular approach that provides for the necessary components to tackle the needs and requirements addressed within the European Higher Education Area.

Specifically, we sought to provide a user-friendly infrastructure that features sharing spaces for ideas and resources to promote learning offers, co-create and co-deliver innovative pedagogies and learning opportunities. It enables linking existing local repositories like course catalogues or course content of universities, aggregating information, and thus making it accessible to users from different backgrounds, supported by a taxonomy and an efficient search engine. As well, more geared at the cooperation of institutions both bilaterally or within the European University Alliances in regards to organizing and supporting common endeavours like joint programmes or joint diplomas, the Hub is providing a front- and backend for administrative support based on needs not covered by local IT systems, which most often are historically designed and geared towards a single-institution approach, requiring incorporation and management of external people, for example Erasmus exchange students, by applying a handling like local users.

Based on the user needs analysis and the review of existing solutions, as detailed in deliverable D2.4, we have chosen Sakai as the learning management system (LMS) module for the Community Portal and the Site Browser component of the BLOOM Hub. Sakai is an open-source educational software platform designed to support teaching, research, and collaboration. It features an outstanding support for learning tools interoperability (LTI), an education technology specification by the IMS Global Learning Consortium allowing the integration of existing content or tools in the context of a course or project site, thus allowing access to and integration of, for example, existing courses in a locally hosted LMS for students of partner institutions without the need to grand access to local systems or having to manage local accounts.

2. Technical specifications

The BLOOM Hub is designed to be modular, and its technical specifications are dependent on the different software components the Hub is combined of, namely the public facing Hub website, the Community Portal with the Site Browser, and the curriculum/program administration interface.

2.1 The public facing website

The BLOOM Hub public facing website, serving as host of the Pedagogical Guidelines as well as an overview and entry point towards the other Hub modules, is being powered by the WordPress content management system (CMS) in its current release version 6.1.1. It is being hosted on a virtual machine running the current release version 11 of Debian Linux, which is the anticipated next/future long-term support (LTS) release that will receive updates, especially security updates, roughly up to June 2026. The virtual machine running the public facing website has been assigned four CPUs, 16 Gb of RAM and 250 Gb of storage. For the web server, the Apache Software Foundations Apache HTTP Server was chosen, as well as MariaDB for the relational database management system (DBMS), a community-developed free and open-source fork of the popular MySQL DBMS.

2.2 The Community Portal and the Site Browser

The BLOOM Hub Community Portal and Site Browser are powered by Sakai, a free and open-source learning management system (LMS) platform that allows both for a feature-rich on site experience to support teaching, research and collaboration, as well as integrating third-party systems, for example content shared by universities and higher education institutions within the realm of European university alliances or to a broader public, by means of learning tools interoperability (LTI), a specification developed by the IMS Global Learning Consortium. Sakai consists of web applications based on the Spring Framework, an application framework for building enterprise Java application using the Java programming platform/language. The Spring framework covers relevant aspects of modern application development, like aspect-oriented programming (AOP), core containers and data access and integration.

For the technical requirements, the systems serving the Community Portal and the Site browser powered by Sakai requires the Apache Software Foundations Apache HTTP Server as well as Apache project tomcat web containers for Jakarta servlets, server pages and web sockets, allowing us to run Java web applications. Again, for the relational database management system (DBMS), MariaDB has been chosen.

We are operating the servers in an industry-standard division between productive hosts and testbed/development/quality assurance hosts. Currently, its sizing is three servers with each 4 CPUs, 16 Gb RAM and 400 Gb fiber channel storage.

2.3 The curriculum/program administration interface

The BLOOM Hub curriculum/program administration interface draws heavily from the MyCampus development at Freie Universität Berlins Department of Mathematics and Computer Sciences. As it is a monolithic Java Spring-Boot web application, as opposed to a Spring Framework application, it features autoconfiguration which allows applications to be initialized with pre-set dependencies and integrated plugins, an opinionated approach to adding and configuring said dependencies to reduce the possibility of configuration errors, and the possibility to run the application without a supporting infrastructure.

Within the BLOOM Hub context, the curriculum/program administration interface runs within the requirements of the tomcat instances used for Sakai (cf. chapter 2.2). Thus, for providing access to the interface, additional tomcats are launched on the Community Portal servers and made available to authorized users.

3. Back-end setup and authentication

3.1 Back-end setup

In order to get a first version of the BLOOM Hub released, we have opted to set up a test bed system by creating a so-called branch of the MyCampus repository dubbed BLOOM. Within this branch, all development regarding the functionalities as well as the implementation of the BLOOM Hub is merged.

The systems providing the BLOOM Hub in an industry-standard division between productive hosts and testbed/development/quality assurance hosts have been set up at Freie Universität Berlin. We have opted for virtual machines on a VMWare cluster system hosted on site, in order to benefit from high

availability[1], easy resource allocation and the ability to quickly and automatically being able to scale up additional host machines should the load or the number of concurrent users warrant it. As well, the datacentre is equipped with best practice precautions regarding (physical) access control, emergency power supply as well as backup capacity on a tape library in a separate, secluded fireproof facility.

The typical worker node/host is set up with the operating system Debian Linux, the Apache project web server as load balancer and for serving static resources, as well as several Apache project tomcat web containers for Jakarta servlets, server pages and web sockets, allowing to run Java web applications. The database engine used is MariaDB, a free and open-source fork of the MySQL relational database management system. The node/host setup, both regarding automation of deployment and updating, as well as software packaging and configuration file management, is managed centrally using the Python-based open-source software stack Salt[2].

3.2 Authentication through EduGAIN

An important part of handling portals and web services geared towards the higher education community is authentication and authorization, as the mere administration of user accounts is not sufficient in most contexts, especially when bridging between institutions and tapping into existing ecosystems. Having the need to provide for verified attributes, be it name, e-mail address or affiliation to an institution, when connecting LMS and opening up course content and their related exams, we have relied on the EduGAIN infrastructure, which is an international interfederation service operated by GÉANT, the pan-European data network for the research and education community. EduGAIN connects both identity providers (IdPs) and service providers (SPs) within the national education and research networks (NRENs), allowing trusted access to services by authenticating users at their home organization, allowing the transfer of required attributes towards the specified service provider. As of December 2022, 80 federations and national research and education networks participate in EduGAIN, allowing users of over 5.200 identity providers at institutions' access to more than 3.600 service providers[3]. After having met all the requirements regarding the EduGAIN code of conduct as well as per the REFEDS Research and Scholarship entity category, the BLOOK Hub is listed as a service provider within the EduGAIN network since March 22nd, 2022.

Having met these first steps in staging the infrastructure, the public release of the first version of the BLOOM Hub was introduced to the project partners and with the interested public on the occasion of the High Level Authority and Consortium Meeting in Paris, France in March 2022 (cf. Fig. 1).

[1] „Best Practices for VMware vSphere® High Availability Clusters", https://docs.vmware.com/en/VMware-vSphere/8.0/vsphere-availability/GUID-39731BEC-EB0C-48C9-813B-CAF9DE884FD5.html, May 31st, 2019. Retrieved December 14th, 2022.

[2] „Welcome to Salt Project ", https://saltproject.io, December 14th, 2022. Retrieved December 14th, 2022.

[3] Federations in eduGAIN, https://technical.edugain.org/, December 16th, 2022. Retrieved December 16th, 2022.

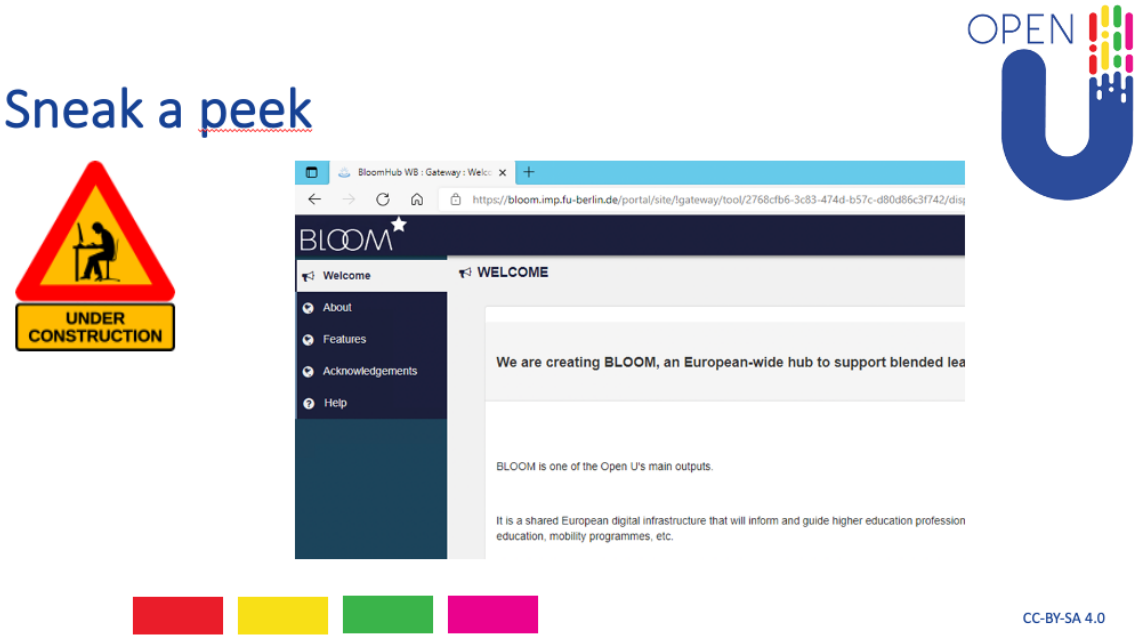

Fig. 1: Slide of the presentation at the Consortium Meeting in Paris, France on March 25th, 2022, introducing the live demonstration of the first release of the BLOOM Hub.

3.3 Shortcomings and possible avenues for additional authentication methods

While we are convinced that the eduGAIN interfederation is, as outlined, the method of choice in regard to authentication and authorization within the higher education community, bridging institutions and service providers securely together, we understand that not all participants and stakeholders within the European higher education area are yet connected to eduGAIN.

Within the Open U project, this was the case for certain public ministries partners, or partners that are organized separately for example as associations of universities. Acknowledging the various requirements in order to join eduGAIN, mainly differing in the applicable policies of the national research and education network where the affected partner is located, we have provided for a separate signup process that allows for manual registration and, depending on the choice of the operating body, confirmation in order to allow individuals to access the BLOOM Hub if their institution has no access to eduGAIN.

This workflow can be tailored towards allowing self-registration based on e-mail domains, allowing the authorisation of specific domains, or by allowing registration and confirmation entirely based on conventional user-submitted information and loop-based confirmation via e-mail. Currently, this self-registration workflow is implemented for all Open U project partners which are not integrated with eduGAIN.

Additionally, during the implementation of the eduGAIN based authentication, we found that even if the home organization is connected to the eduGAIN federation and per se supports authentication through said federation, the implementation of the agreed standards within the federation differs based on local interpretations regarding legal requirements, mainly in the application of the EU GDPR. Even though we published relevant documents like a privacy statement, the code of conduct compliance statement and our inclusion within the REFEDS Research and Scholarship Entity Category, we often had to bilaterally communicate and ask for whitelisting of our service provider in order to enable users from

organizations to access the BLOOM Hub. This sometimes was a matter of hours and quick e-mail exchanges, but could sometimes last for weeks involving data protection officers in both organizations.

As well, the underlying the Security Assertion Markup Language (SAML) profile used for Shibboleth single sign on authentication at most sites, and especially local implementations in regard to the home organizations Identity Provider (IdP) displays a shortcoming which made it sometimes very difficult to get a consistent mapping of the required attributes transmitted from the home organizations Identity Provider (IdP) towards the BLOOM Hub operating as a Service Provider (SP) within eduGAIN.

We currently request the following attributes:

displayName (SAML:2.0)	urn:oid:2.16.840.1.113730.3.1.241	(required)
eduPersonPrincipalName (SAML:2.0)	urn:oid:1.3.6.1.4.1.5923.1.1.1.6	(required)
eduPersonScopedAffiliation (SAML:2.0)	urn:oid:1.3.6.1.4.1.5923.1.1.1.9	(required)
eduPersonTargetedID (SAML:2.0)	urn:oid:1.3.6.1.4.1.5923.1.1.1.10	(required)
mail (SAML:2.0)	urn:oid:0.9.2342.19200300.100.1.3	(required)

Specifically, the SAML metadata are specified only with a binary yes/no scheme when requesting and processing attribute lists. A combination and processing using Boolean operators is currently not possible. But with regards to the concept of privacy by design mandated by the EU GDPR, we would have needed only one identifier, which is mapped to be the local username within the Hub. This would have translated to:

```
<AND>
  mail
  displayName
  <OR>
    pairwise-id
    eduPersonTargetedId
    eduPersonPrincipalName
  </OR>
</AND>
```

As this is currently not possible within Shibboleth, we had to settle for a broader attribute request within the eduGAIN metadata. This is the former request:

pairwise-id	-> required
eduPersonTargetedID	-> optional
displayName	-> required
mail	-> required

We require a persistent ID in order to consistently map user accounts within the BLOOM Hub to the corresponding users authenticated through the home organization IdPs. While the concept of such persistent IDs being “persistent, revocable, non-reassignable, opaque, targeted, non-global identifier for identifying the subject in a SAML assertion”[4] have been introduced with the SAML v2.0 specification in 2005, unfortunately there is still a number of home organizations that have not yet implemented the privacy friendly pairwise ID, and this left their users stranded when trying to access the BLOOM Hub. As well, many of the contacted home organizations were not able to specify a time

[4] „NameIdentifiers“, https://shibboleth.atlassian.net/wiki/spaces/CONCEPT/pages/928645231/NameIdentifiers, June 4th, 2020, retrieved December 14th, 2022.

limit in which their IdP would be able to provide a pairwise ID, or the scheduled time frames where on a long-term scale.

The broader request translates to:

pairwise-id -> required
eduPersonTargetedID -> required
eduPersonPrincipalName -> required
displayName -> required
mail -> required

This allows for a broader, easier access for users within the eduGAIN federation, as most participating home organizations can supply either the pairwise ID, an eduPersonTargetedID or an eduPersonPrincipalName, thus alleviating the need to individually help users and contacting their home organizations.

Nevertheless, this required a local implementation of mapping between external identities and local assigned usernames and profiles, as the transmitted information from the home organizations IdP could change when their local configuration is changed or enhanced, rendering the BLOOM Hub users profile stale, and leaving him with a fresh, new profile, which needs manual adjusting and correction. This happens for example when the home organizations IdP finally supports either the pairwise ID or the Targeted ID, which it then prefers to release as user attribute over the Principal Name. Users having accessed the BLOOM Hub and worked with it using an eduPersonPrincipalName as the identifier are then finding themselves in a fresh, empty profile and user account on the Hub, seemingly having lost their content and settings until the link is (manually) adjusted.

To ease the possible impact on the user experience and to circumvent these SAML implementation shortcomings, we are actively encouraging the adoption of the pairwise ID within the AAI community of the Deutsches Forschungsnetz e.V., our national research and educational network (NREN), which itself is also connected with the European NRENs and with the eduGAIN community.

Also, we envision additional ways and means of authenticating to the BLOOM Hub platform, which would also alleviate the outlined shortcomings. Specifically, we envision to implement authentication and authorization through verifiable student IDs, in two settings following up on the Open U projects outcome, one together with KU Leuven and Università di Bologna within the EBSI network, and one together with the ERUA European university alliance by implementing their ERUA ID, a self-sovereign ID.

Project Number: 606692-EPP-1-2018-2-FR-EPPKA3-PI -POLICY

Project Duration: 47 months

Start date: 20-02-2019

End date: 19-12-2022

Coordination: Université Paris 1 Panthéon-Sorbonne

Co-funded by the
Erasmus+ Programme
of the European Union