



Yareta  
Specifications



- Introduction .....3**
- 1 Standards, protocols and registries.....4**
- 2 Technologies and open-source applications.....5**
- 3 Compliance .....6**
  - Compliance with the Swiss Law ..... 6
  - Compliance with FAIR principles ..... 6
- 4 Infrastructure .....7**
  - Open-source ..... 7
  - Universal..... 7
  - Modular architecture based on RESTful APIs..... 7
  - Deployment options..... 7
  - Redundancy and geographic separation..... 7
  - Multiple archival storage systems..... 7
  - Data security ..... 7
  - Policy-based data management system..... 8
  - Angular user interface..... 8
  - Multi-language support..... 8
- 5 Functionalities .....9**
  - SWITCHaai authentication ..... 9
  - ORCID integration..... 9
  - Preservation Space..... 9
  - Role-based access control ..... 9
  - Submission policy management..... 9
  - Preservation policy management ..... 10
  - Dissemination policy management [upcoming feature] ..... 10
  - Pre-ingest workflow preserving folder hierarchy and data structure ..... 10
  - Automated file assessment..... 10
  - Rich metadata ..... 10
  - Date of publication and optional embargo period..... 11
  - Data access management ..... 11
  - Licenses management..... 11
  - Persistent identifier (DOI) ..... 11
  - Metadata harvesting via OAI-PMH..... 11
  - IIIF support ..... 11
  - Finding data..... 11
  - Integration..... 12
  - Tags ..... 12
  - Preview functionality ..... 12
  - Resources ..... 12
  - Customizable functionalities ..... 13

## INTRODUCTION

Yareta is a local implementation of the DLCM solution, dedicated to long-term preservation and publishing of research data in the Canton of Geneva. Open-source and compliant with the OAIS reference model (ISO 14721) and international standards, the solution is designed to support the acquisition, validation, ingest, storage, management, preservation, and dissemination of all types of research data, in compliance with Swiss funder requirements (SNSF) and FAIR principles.

Below are some of the key features of the solution. *Customizable functionalities are identified in italic in the text.*

Version 1.0

Prepared by: Lydie Echernier, Research Advisor and Service Manager for Yareta, e-Research Team, UNIGE  
David Gold, Support for Yareta, e-Research Team, UNIGE

Validated by: Pierre-Yves Burgi, Head of the Scientific Information and Research Department (RISE), IT Division, UNIGE, and Representative for Yareta  
Hugues Cazeaux, Head of the e-Research Team, UNIGE, and Solution Architect for Yareta

Creation date: May 19, 2020

Last revision date: October 28, 2022

Publication date: October 28, 2022

## 1 STANDARDS, PROTOCOLS AND REGISTRIES

[ISO 14721](#): Open Archival Information System (OAIS) reference model

[FAIR principles](#): set of guiding principles in order to make data findable, accessible, interoperable and reusable

[OAI-PMH](#): Open Archives Initiative Protocol for Metadata Harvesting

[DataCite open descriptive metadata schema](#)

[METS](#): Metadata Encoding & Transmission Standard (Library of Congress)

[XML](#) and [JSON-based](#) file system metadata

[PREMIS](#): Preservation Metadata Maintenance Activity (Library of Congress)

[LOCKSS](#): Lots of Copies Keep Stuff Safe (Stanford University)

[BagIt](#): hierarchical file packaging format (Library of Congress)

[PUID scheme](#): file format validation (UK National Archives)

[re3data classification of subject areas](#) (German Research Foundation)

[List of research domains and disciplines](#) (SNSF)

## 2 TECHNOLOGIES AND OPEN-SOURCE APPLICATIONS

[SWITCHaai](#) authentication

[SWITCHengines](#) long-term storage service

[DOI](#) (Digital Object Identifier)

[ORCID](#) (Open Researcher and Contributor ID)

[Creative Commons licenses](#)

[Java](#) programming language and [Spring Boot](#)

[Docker](#): open-source containerization platform

[Kubernetes](#): open-source container-orchestration system

[Angular](#): open-source JavaScript framework

[PHP](#): open source scripting language

[RESTful APIs](#): web services

[IIIF](#) (Image Interoperability Framework)

[S3](#) (Amazon Simple Storage Service): object storage through a web service interface

Tape file system: files storage on magnetic tape

[FITS](#): file format identification (Harvard)

[Elasticsearch](#): open source search and analytics engine

[ClamAV](#): open source antivirus engine

[JSmol](#): open-source Java viewer for chemical structures in 3D

[Jenkins](#): open source automation server (deployment and test)

[MySQL](#) and [MariaDB](#): open-source relational database

[ActiveMQ](#): messaging server

| Feature                                       | Benefit   |
|---|---|
| <p><b>COMPLIANCE WITH THE SWISS LAW</b></p>   | <p>Yareta solely relies on existing data centers in Switzerland for archival storage, making it suitable to preserve your sensitive data according to the legal provisions (<a href="#">Federal Act on Data Protection</a>).</p>  |
| <p><b>COMPLIANCE WITH FAIR PRINCIPLES</b></p> | <p>Yareta follows <a href="#">FAIR principles</a> and is eligible for funding for data upload for researchers under a <a href="#">SNSF grant</a>.</p> <p><b>Findable</b></p> <p>Yareta registers your descriptive metadata with <a href="#">DataCite</a> [F2, F4] to assign a unique and persistent identifier (<a href="#">DOI</a>) to your datasets [F1] and provides discoverability over all hosted descriptive metadata [F4], which include the respective DOIs [F3].</p> <p><b>Accessible</b></p> <p>Yareta offers all descriptive metadata for harvesting and, if license conditions allow, datasets for download without authentication [A1.2] through open standard [A1.1] protocols (<a href="#">OAI-PMH</a> and <a href="#">RESTful APIs</a>) [A1]. Data is retained according to its respective retention policy, and descriptive metadata is maintained after disposal and data destruction. Yareta is developed at the University of Geneva and powered by <a href="#">DLCM</a>, mandated by <a href="#">swissuniversities</a>. The University of Geneva is the guarantor of the sustainability of the solution. [A2]</p> <p><b>Interoperable</b></p> <p>Yareta provides descriptive and administrative metadata in XML format, and custom metadata in both <a href="#">XML</a> and <a href="#">JSON</a> formats [I1] using internationally-recognized open vocabularies from various authoritative sources such as <a href="#">DataCite</a>, the <a href="#">UK National Archives</a> and the <a href="#">Library of Congress</a> [I2]. In order to improve data linking, datasets in YARETA can be annotated with custom rich metadata mapping to community standards and ontologies (such as <a href="#">DublinCore</a>), and combined in collections interlinking datasets and with qualified referencing in the metadata schema (using the <code>relatedIdentifier</code> property) [I3].</p> <p><b>Reusable</b></p> <p>Yareta requires a mandatory set of descriptive metadata from the <a href="#">DataCite metadata scheme</a> and records administrative metadata (<a href="#">METS</a>) for the management of data files [R1.3]. It allows customization with additional metadata, including 'Recommended' and 'Optional' DataCite metadata properties as well as optional custom rich metadata [R1], validated against their respective scheme [R1.3]. The descriptive metadata of all datasets are released openly (i.e. with a <a href="#">Public</a> access level) and establish a direct link between the dataset and the license (Yareta provides standard <a href="#">Creative Commons licenses</a> by default and allows for custom alternative licenses) [R1.1]. Descriptive metadata include authorship details of the <a href="#">Organizational Unit</a> hosting the dataset and of its <a href="#">Contributors</a>, including a link to their <a href="#">ORCID</a> unique researcher identifier, when provided [R1.2].</p> |

## 4 INFRASTRUCTURE

| Feature   | Benefit   |
|---|---|
| <b>OPEN-SOURCE</b>                                | Yareta is 100% built on top of open-source technologies and non-proprietary protocols to keep the solution flexible, scalable and open to future developments and upgrades. The DLCM solution powering Yareta is also open-source. This means that you may use the infrastructure that best fits your institutional needs.  |
| <b>UNIVERSAL</b>                                  | Yareta maintains an agnostic view towards file formats and object types, which enables it to preserve essentially any type of file. It is intended to be generic to the broadest range of research datasets. Yareta does not aim to replace discipline- or community-specific repositories, but to complement their use by offering a solution compliant with international standards applied to information sciences on preservation, accessibility and reuse of data.   |
| <b>MODULAR ARCHITECTURE BASED ON RESTFUL APIS</b> | The modular architecture of the solution makes it easy to add functionalities to Yareta without affecting the core. Developments can take place independently across the various modules of the solution, and RESTful APIs provide scalability and large-scale performance for the distributed modules and integrated systems. <i>The <a href="#">well-documented RESTful APIs</a> are suitable for an easy <a href="#">integration of Yareta with your preferred applications and services</a>, and ensure that your data never get locked in a single system.</i>   |
| <b>DEPLOYMENT OPTIONS</b>                         | <p>Yareta offers durable and cost-effective long-term preservation in Switzerland. <i>Yareta proposes a complete deployment flexibility: you can choose to use the Yareta web portal or to deploy Yareta with many different forms and configurations, in a cloud of your choice, on-premise on your own infrastructure, or a mix of both.</i></p> <p>Yareta is a cloud-native application compatible with on-premise, Docker or Kubernetes deployment, operating on a group of two or more host machines (nodes) that have their own technology, copy of the datasets and backups. <i>The <a href="#">number of nodes and their geographical distribution</a> can be specified.</i> Multiple copies replicated across multiple locations mitigate the risk of data corruption and loss.</p> <p>As an <a href="#">API-based solution</a>, Yareta has many advantages. Thanks to its modular infrastructure, Yareta enables institutions to streamline critical data processes, enforce policies and monitor performance, with a significant spare on development and maintenance costs.</p> |
| <b>REDUNDANCY AND GEOGRAPHIC SEPARATION</b>       | Yareta follows good practices by maintaining multiple copies in geographically distant storage locations/systems to protect your data against disaster, hardware or software failure, and human error. <i>To maximize the durability and availability of your data, you can define a custom <a href="#">redundancy-based retention policy and data location</a>, by specifying the number of instances you want to be retained and your preferred data center locations.</i>  |
| <b>MULTIPLE ARCHIVAL STORAGE SYSTEMS</b>          | Yareta operates on a multitier and distributed architecture to preserve your data on file-based systems and across different types of storage infrastructures. Yareta integrates natively with Amazon S3 compatible storage systems, such as SWITCH's long-term storage solution. Big volumes of data can be managed and tiered across magnetic data tapes. <i>You can select the most relevant system for your preservation plan depending on your requirements, budget and strategy.</i>  |
| <b>DATA SECURITY</b>                              | To ensure that the service is reliable, Yareta uses different processes to confirm the authenticity and to independently verify the integrity of the data.  |

| Feature   | Benefit   |
|---|---|
| <ul style="list-style-type: none"> <li>• PREMIS preservation metadata</li> <li>• Checksums</li> <li>• Audit trails</li> </ul> | <p>Preservation metadata (PREMIS) create a data-centric trust chain for submission to the repository, for archival storage and for dissemination from the repository. Files are processed on upload to generate <a href="#">key assessment information</a> and <a href="#">metadata</a> for file management and long-term preservation planning and, for each file, a checksum is calculated. Checksums are used to ensure the file's integrity is being maintained in archival storage on a regular schedule. Preservation planning tasks are monitored with administrative dashboards and notification systems, and audit trails offer a back-end view of system use.</p>   |
| <p><b>POLICY-BASED DATA MANAGEMENT SYSTEM</b></p>   | <p>The definition of repository trustworthiness rests on a list of tasks specified by the <a href="#">OAIS reference model</a> and enforced using computer actionable policies. These policies specify the tasks required to manage submission of data (SIP), creation of archives (AIP) and dissemination of datasets (DIP). <i>Yareta invite users to define some properties involved in these <a href="#">submission</a>, <a href="#">preservation</a> and <a href="#">dissemination</a> policies, enabling researchers to carry on as usual with the workflow that is right for them.</i></p>   |
| <p><b>ANGULAR USER INTERFACE</b></p>  | <p>Yareta uses Angular (an open-source JavaScript framework) to create a better and more efficient user experience. The Angular framework allows the Yareta development team to remain focused on the core logic of the DLCM solution while offering a smooth yet powerful portal, rich in content, fast to load and easy to use, with a modern interface design that comes with <i>both light and dark themes</i>. Yareta integrates Formly, a (JSON powered) dynamic library used to design Angular Reactive Forms. <i>It allows for the customization of an <a href="#">expanded DataCite descriptive metadata</a> form to match your needs and facilitate a most efficient identification of your datasets.</i></p> |
| <p><b>MULTI-LANGUAGE SUPPORT</b></p>  | <p><i>Yareta supports datasets in multiple languages, and its user interface is already available in English and French.</i></p>  |



## 5 FUNCTIONALITIES

| Feature  | Benefit   |
|--|---|
| <b>SWITCHAAI AUTHENTICATION</b>  | Access to Yareta is specifically intended to anyone with ties to academic and research institutions in the Geneva canton and their collaborating partners. A SWITCHaai authentication is required before accessing any other functionality in Yareta, and all user actions are logged in an audit trail for accountability. Yareta is opened to SWITCHedu-ID, allowing collaboration with non-academic national and international partners. Access permissions can be defined both at the <a href="#">Organizational Unit</a> level (control over five fine-grained <a href="#">user roles</a> ) and at the individual dataset level (choice of <a href="#">Closed</a> , <a href="#">Restricted</a> or <a href="#">Public</a> release of your data).  |
| <b>ORCID INTEGRATION</b>   | Yareta offers an ORCID <a href="#">integration</a> for the purpose of name disambiguation. By linking your unique researcher identifier to your personal profile on Yareta, you ensure that your research data activities and outputs are correctly attributed to you.  |
| <b>PRESERVATION SPACE</b> <ul style="list-style-type: none"> <li>• <b>Organisational Units</b></li> <li>• <b>Contributors</b></li> <li>• <b>Request notifications</b></li> </ul> | <p>Your Preservation Space is your private workspace on Yareta. It is characterized by a series of Organizational Units (research teams, research projects, departments, laboratories, etc.), a group of Contributors (members of Organizational Units) and a set of notifications (requests sent and requests received).</p> <p>In your Preservation Space, you can list all Organizational Units to consult their <a href="#">Public</a> archives, access the ongoing deposits of the Organizational Units you are a member of, and request membership to other Organizational Units to consult their <a href="#">non-public</a> archives. You can also list all the Contributors to your Organizational Units and their deposits. <i>In addition, you can modify the Organizational Units you are a <a href="#">Manager</a> of, and control who can edit deposits within them by assigning each Contributor a specific <a href="#">role</a>.</i></p> <p><i>Yareta notifies you of peoples' requests to share your <a href="#">non-public</a> archives. You can grant people a <a href="#">Visitor</a> access to your Organizational Units so they can consult all of your <a href="#">Restricted</a> archives, and choose to accept or refuse individual request to access your <a href="#">Closed</a> archives on a case-by-case basis.</i></p> |
| <b>ROLE-BASED ACCESS CONTROL</b>   | Yareta uses a role-based access control to grant appropriate permissions to users. All <a href="#">Contributors</a> access <a href="#">Organizational Units</a> with authenticated user accounts, and are assigned roles such as Manager, Steward, Approver, Creator or Visitor. <i>Yareta delegates responsibilities for managing an <a href="#">Organizational Unit</a> to the Manager, who defines the different policies governing the use of the unit and the access to the data it contains. <a href="#">Contributors</a> to an <a href="#">Organizational Unit</a> share access to the datasets and can perform certain actions in the deposits. For more information about the actions that each role can perform, see <a href="#">a detailed explanation of the roles</a>.</i>   |
| <b>SUBMISSION POLICY MANAGEMENT</b>  | <i>Yareta allows you to add a validation step to your workflow. The submission policy sets the workflow that controls the validation of your deposits. It defines how <a href="#">Contributors</a> can collaborate to publish a dataset, and more specifically who can approve a dataset or not. Choosing "Deposit with approval" as your submission policy enables an <a href="#">Approver</a> or a data <a href="#">Steward</a> (such as a domain expert, archivist, librarian, etc.) to review datasets sent to validation and to approve, reject or to send them back to edition for modification. With the "Deposit without approval" policy, <a href="#">Contributors</a> who are assigned <a href="#">Creator</a> role in your <a href="#">Organizational Unit</a> can also decide to validate datasets themselves.</i>  |

| Feature  | Benefit   |
|--|---|
| <p><b>PRESERVATION POLICY MANAGEMENT</b></p>   | <p>The data preservation policy defines how long Yareta retains your data for. Within the framework of good scientific practice, Yareta proposes the recommended retention periods of 10 years for research data and 15 years for medical devices, but <i>you may set your own retention policy for your <a href="#">Organizational Unit</a></i>. When the retention period ends, Yareta deletes the data from the archival storage. Your <a href="#">descriptive metadata</a> is maintained, so your <a href="#">DOI</a> remains a permanent mean of identifying your dataset.</p>   |
| <p><b>DISSEMINATION POLICY MANAGEMENT [UPCOMING FEATURE]</b></p> <ul style="list-style-type: none"> <li>• <b>Web publishing</b></li> <li>• <b>IIIF dissemination policy</b></li> </ul>   | <p><i>Yareta offers an independent set of dissemination policies based on protocols and interoperability standards to provide a timely and seamless portability of your data to different information systems, devices and applications. By default, Yareta supports support web publishing and deployment of your datasets to custom servers and websites. An independent <a href="#">IIIF</a> dissemination policy allows for the publishing of specific image data files to IIIF-based servers and tools like for instance <a href="#">Bodmer Lab catalog system</a>.</i></p> <p><i>The development of customized dissemination policies can provide to you with greater manageability to improve the visibility and use of your datasets across all systems.</i></p>  |
| <p><b>PRE-INGEST WORKFLOW PRESERVING FOLDER HIERARCHY AND DATA STRUCTURE</b></p>   | <p>Yareta includes a pre-ingest module that can automatically acquire, unpack and ingest data at scale, for a seamless deposit of datasets from a variety of systems. The pre-ingest module can extract, list and test compressed data form a ZIP file, while preserving your folder hierarchy and data structure in the solution.</p>  |
| <p><b>AUTOMATED FILE ASSESSMENT</b></p> <ul style="list-style-type: none"> <li>• <b>Virus scanning</b></li> <li>• <b>File format identification</b></li> <li>• <b>Data compliance level assessment</b></li> <li>• <b>File filtering</b></li> </ul> | <p>Yareta brings together file format identification (FITS) with data compliance level assessment, using structured format policy registries provided by the internationally-recognized UK National Archives (PUID scheme) and Library of Congress (Recommended Formats Statement), in order to encourage the use of open standards. <i>A fine-grained filtering is applied so you can validate, ignore or exclude files based on their status after automatic assessment and virus scanning (ClamAV).</i></p>  |
| <p><b>RICH METATDATA</b></p> <ul style="list-style-type: none"> <li>• <b>DataCite descriptive metadata</b></li> <li>• <b>METS administrative metadata</b></li> <li>• <b>XML and JSON custom metadata</b></li> </ul>                                | <p>Allowing for rich metadata, Yareta makes sure your information is findable and identifiable in the future.</p> <p>Yareta integrates the DataCite metadata scheme to support dataset citation, discovery and persistence. It requires a minimum set of descriptive metadata during the submission process so that a <a href="#">persistent digital object identifier (DOI)</a> can be generated and assigned to each dataset. <i>For a more accurate and consistent identification of your datasets, you can very easily customize the form to your own needs with chosen fields from the pool of DataCite metadata properties.</i> Yareta provides discovery services over all descriptive metadata for citation and retrieval purposes.</p> <p>Administrative metadata (METS) is collected automatically during the deposit, allowing for internal tracking and management purposes. It supports Yareta’s specific preservation and dissemination functions by providing information to help understand data files, to <a href="#">secure their fixity and integrity</a>, and to enable their access and use by the designated community.</p> <p><i>Yareta gives you the ability to associate custom metadata to your dataset to describe your data. You can define a metadata scheme in a format that is unique and meaningful to your research field, such as XML or JSON. YARETA validates both the type (format) of the file and its content (structure) against the metadata scheme.</i></p> |

| Feature  | Benefit  |
|--|--|
| <b>DATE OF PUBLICATION AND OPTIONAL EMBARGO PERIOD</b>   | <i>Yareta allows you to determine the date of publication if your dataset is not to be published immediately. An optional embargo period can also be specified together with a more exclusive access level. Your data will only be accessible after the embargo has expired. The <a href="#">descriptive metadata</a> remain discoverable, to ensure the findability of your dataset.</i>  |
| <b>DATA ACCESS MANAGEMENT</b>  | Yareta makes your data and metadata accessible in a secure and controlled way, so you can comply with legal requirements. <i>Flexible access management offers you some control over who can access your datasets and how.</i> When Public sharing of your research data might be inappropriate – when anonymization cannot be carried out effectively or in case of commercial agreement – you may wish to use the Restricted or Closed access levels, which ensure that research data will not be published openly. Only members of your <a href="#">Organizational Units</a> will be able to see your Restricted datasets, and they will be required to submit a request so you can grant them access to your Closed datasets. The <a href="#">DataCite metadata</a> , describing the dataset, remain available to ensure its findability, but the actual data is only made available to authorized users upon demand.                                      |
| <b>LICENSES MANAGEMENT</b>   | As a best practice, Yareta encourages you to apply a license to your datasets to retain your copyrights and to control the attribution, use, distribution and modification of your data, by informing others about what you allow them to do with your data. Yareta offers a default choice of open licenses. <i>You may ask for additional widely-used open licenses to be included, and additional custom, modified or non-standard terms can be created on demand.</i><br><br>In respect of the <a href="#">FAIR principles</a> , a license must be specified for each dataset released under a <a href="#">Public</a> access with Yareta. Without a choice, a default CC-BY (Creative Commons Attribution 4.0) license will apply. <i>You can also define your preferred default license in your <a href="#">Organizational Unit</a>. [upcoming feature]</i>   |
| <b>PERSISTENT IDENTIFIER (DOI)</b>   | YARETA provides metadata registration with DataCite's DOI (Digital Object Identifier) to make data sharing FAIR, by ensuring that your dataset is uniquely and permanently identifiable and citable in scientific publications. Each dataset published with Yareta receives a DOI at the time of publication. <i>If required, DOIs can be reserved for a dataset upon deposit, before actually making it publicly available, allowing you to quote your data in any articles being prepared.</i>   |
| <b>METADATA HARVESTING VIA OAI-PMH</b>   | Each dataset published with Yareta is automatically <a href="#">indexed with DataCite</a> and offered for harvesting via the OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) standardized protocol to ensure maximum dissemination and findability of your research data.  |
| <b>IIIF SUPPORT</b>  | Yareta is able to deliver images and metadata according to the IIIF specifications, the International Image Interoperability Framework for publishing and sharing content.   |
| <b>FINDING DATA</b><br><ul style="list-style-type: none"> <li>• Archives browsing</li> <li>• Keyword searching</li> <li>• Filtering</li> </ul> | Yareta provides discovery services over all hosted <a href="#">descriptive metadata</a> on Yareta. To find datasets, you can browse through <a href="#">Organizational Unit</a> and dataset titles, or you can enter one or more keywords in the search box. When connected to your Yareta account, a filter gives you additional search options to narrow your search by <a href="#">Public</a> datasets (including downloadable data), or Private ones, i.e. datasets with <a href="#">Restricted</a> or <a href="#">Closed</a> access, including <a href="#">descriptive metadata</a> only, for which you will be able to request access to.<br><br>You can use the most common types of search expressions and Boolean operators to refine your search: <ul style="list-style-type: none"> <li>• AND: Finds results where both terms next to AND are present (a search without any operator includes an implicit AND between the terms entered)</li> </ul> |

| Feature   | Benefit  |
|---|--|
|   | <ul style="list-style-type: none"> <li>• OR: Finds results containing either one or the other term entered</li> <li>• - [minus]: Finds results not including the term following the minus sign</li> <li>• " " [quotes]: Finds results containing the exact phrase within the quotes (punctuation is ignored)</li> <li>• * [asterisk]: Finds results containing terms starting with the letter(s) followed by the asterisk (returns all published archives if used alone)</li> </ul>  |
| <p><b>INTEGRATION</b></p>   | <p>Yareta serves as an integrated solution with a central repository for long-term preservation and publishing of research data, which can be synchronized with many other satellite applications. The solution already supports different open-source and proprietary technologies. It provides <a href="#">RESTful APIs</a> in Python to easily integrate any instrument, application and data center you are using, so your data is up to date and completely interconnected with minimal manual effort.</p> <p><b>Examples of integration projects</b></p> <ul style="list-style-type: none"> <li>• Connect lab instruments, LIMS (Laboratory Information Management System) and ELN (Electronic Laboratory Notebook platform), such as SLIMS (Genohm) and openBIS (ETHZ), to automatically feed datasets with results, locations, properties and samples from your experiments.</li> <li>• Connect visualization and analysis tools, such as Tableau, to create dashboards and conduct analyses.</li> <li>• Integrate any other application to automatically pull in your data and keep it in sync, including custom or proprietary solutions. For example, Yareta integrates Jupyter Notebooks to allow you to write custom Python scripts to process your data.</li> <li>• Integrate with external data centers to back up your data onto your own <a href="#">Amazon S3 or other systems of record</a>.</li> </ul> <p><i>Yareta can be integrated to any online service or software that offers API functionality. Consult our <a href="#">Integration Guide</a> and <a href="#">RESTful API Guide</a> to know how to connect Yareta to your research environment.</i></p> |
| <p><b>TAGS</b></p> <ul style="list-style-type: none"> <li>• Research domains</li> <li>• Keywords</li> </ul>                                   | <p>Yareta uses tags to categorize and filter your data. There are two types of tags available: research domain tags and keyword tags. Research domain tags can be used to categorize your <a href="#">Organizational Unit</a>'s work based on Swiss and international registries, such as the SNSF research domains and disciplines and the classification of subject areas of the German Research Foundation used by re3data. Keyword tags can be helpful to specify the content and purpose of your datasets in relation to a specific topic or research area in complement to the title and short data description.</p>   |
| <p><b>PREVIEW FUNCTIONALITY</b></p>   | <p>Yareta includes a universal visualization tool for previewing a large selection of data types (documents, images, audio, video, etc.) to repository users, including for instance 2D and dynamic 3D displays of molecular structure files using JSmol.</p>  |
| <p><b>RESOURCES</b></p> <ul style="list-style-type: none"> <li>• Tooltips</li> <li>• User guide</li> <li>• Technical documentation</li> </ul> | <p>Yareta has created a set of resources to help you understand its various capabilities. You can learn to use the solution in various ways: by accessing the user guide and tooltips within the application, by reading DLCM technical documentation on the DLCM Sandbox (<a href="#">Integration Guide</a>, <a href="#">RESTful API Guide</a> and <a href="#">DLCM Tools Guide</a>), and by exploring the resources on the <a href="#">Yareta web page</a>. Tooltips (onscreen labels) provide helpful information about some options when you position the pointer over them. If you need additional information about a feature, you can access the user guide at any time from within the application.</p>  |

| Feature                                    | Benefit   |
|--|---|
| <p><b>CUSTOMIZABLE FUNCTIONALITIES</b></p> | <p>Yarea offers a high degree of freedom with functionalities to fit your usual work environment and personal preferences. (Customizable functionalities are identified in <i>italic</i> in the text.)</p> <p>Customizable Yarea functionalities:</p> <p>Your profile</p> <ul style="list-style-type: none"> <li>• <i>Link your ORCID</i> to your personal profile to get your datasets correctly attributed to you</li> <li>• Choose between <i>light and dark modes</i></li> </ul> <p>Your Preservation Space</p> <ul style="list-style-type: none"> <li>• Keep a complete view over <i>Organizational Units</i>, their <i>Contributors</i> and deposits</li> <li>• Manage <i>request notifications</i> (to request, grant or deny access)</li> <li>• Have responsibilities delegated to you: as a <i>Manager</i>, control the <i>Organizational Unit</i>'s policies, its <i>Members</i> and their access level (<i>roles</i>)</li> <li>• Implement your own workflow by choosing to enable an optional validation step in your <i>submission policy</i></li> <li>• Define your <i>preservation policy</i> to meet your funder or institutional required timeframe for retention, preservation and access</li> <li>• Opt for the <i>dissemination policy</i> that suits your needs (web or <i>image-based servers publishing</i>)</li> <li>• Include your <i>discipline and keywords</i> to improve the visibility of your datasets in your research community</li> <li>• Define a default <i>license</i> for your <i>Organizational Unit</i></li> </ul> <p>Your deposits</p> <ul style="list-style-type: none"> <li>• <i>Reserve a DOI</i> ahead of the publication of your dataset</li> <li>• Choose to validate, ignore or exclude files based on their status after <i>automatic assessment</i> and virus scanning</li> <li>• Choose your <i>date of publication</i> and apply an optional embargo</li> <li>• Define who can access your data and how (<i>Public</i>, <i>Restricted</i> or <i>Closed</i>)</li> <li>• Choose <i>licenses</i> from a selection of widely-used open licenses and ask for additional custom, modified or non-standard terms</li> <li>• Choose a retention period for your datasets among the <i>preservation policies</i> set at the <i>Organizational Unit</i> level</li> <li>• Add or edit <i>metadata</i>, by expanding DataCite descriptive metadata form (optional properties) and customizing your metadata according to a chosen scheme (XML or JSON)</li> <li>• Connect Yarea to your research environment through <i>RESTful APIs</i> (third party <i>integrations</i>)</li> </ul> |

This text is made available under a CC-BY-SA licence.

Photographic credit: UNIGE, Jacques Erard, 2018 <https://phototheque.unige.ch/unige:121641>

Contact:

e-Research  
Université de Genève  
Division du système et des technologies de l'information et de la communication / IT Division  
24 rue du Général-Dufour  
1211 Genève 4  
[eresearch@unige.ch](mailto:eresearch@unige.ch)