



Yareta

Spécifications



Introduction	3
1 Normes, protocoles et registres	4
2 Technologies et applications open source	5
3 Conformité	6
Respect de la loi suisse.....	6
Conformité aux principes du FAIR.....	6
4 Infrastructure	8
Open-source.....	8
Universel.....	8
Architecture modulaire basée sur des API RESTful.....	8
Options de déploiement.....	8
Redondance et séparation géographique.....	8
Systèmes de stockage d'archives multiples.....	8
Sécurité des données.....	9
Système de gestion des données basé sur des politiques.....	9
Interface utilisateur Angular.....	9
Support multilingue.....	9
5 fonctionnalités	10
Authentification de SWITCHaai.....	10
Intégration de l'ORCID.....	10
Espace de préservation.....	10
Contrôle d'accès basé sur les rôles.....	10
Gestion de la politique de soumission.....	10
Gestion de la politique de préservation.....	11
Gestion de la politique de diffusion [fonctionnalité à venir].....	11
Flux de travail pré-ingestion préservant la hiérarchie des dossiers et la structure des données.....	11
Évaluation automatisée des dossiers.....	11
Des métadonnées riches.....	11
Date de publication et période d'embargo facultative.....	12
Gestion de l'accès aux données.....	12
Gestion des licences.....	12
Identifiant persistant (DOI).....	13
Récolte de métadonnées via OAI-PMH.....	13
Soutien de l'IIF.....	13
Trouver des données.....	13
Intégration.....	13
Tags.....	14
Fonctionnalité de prévisualisation.....	14
Ressources.....	14
Fonctionnalités personnalisables.....	14

INTRODUCTION

Yareta est une implémentation locale de la solution DLCM, dédiée à la préservation à long terme et à la publication des données de recherche dans le canton de Genève. Open source et conforme au modèle de référence OAIS (ISO 14721) et aux normes internationales, la solution est conçue pour soutenir l'acquisition, la validation, l'ingestion, le stockage, la gestion, la préservation et la diffusion de tous les types de données de recherche, en conformité avec les exigences des financeurs suisses (FNS) et les principes FAIR.

Vous trouverez ci-dessous quelques-unes des principales caractéristiques de la solution. *Les fonctionnalités personnalisables sont identifiées en italique dans le texte.*

Version 1.0

Préparé par : Lydie Echernier, conseillère en recherche et responsable du service Yareta, équipe e-Research, UNIGE
David Gold, support pour Yareta, équipe e-Research, UNIGE

Validé par : Pierre-Yves Burgi, Chef du Département Recherche et Information Scientifique (RISE), Division IT, UNIGE, et représentant de Yareta
Hugues Cazeaux, Chef de l'équipe e-Research, UNIGE, et architecte de solution pour Yareta

Date de création : 19 mai 2020

Date de la dernière révision : November 14, 2022

Date de publication : November 14, 2022

[ISO 14721](#) : modèle de référence du Système Ouvert d'Archivage d'Information (OAIS)

[Principes FAIR](#) : ensemble de principes directeurs visant à rendre les données trouvables, accessibles, interopérables et réutilisables.

[OAI-PMH](#) : Protocole de l'Open Archives Initiative pour la collecte de métadonnées.

[Schéma ouvert de métadonnées descriptives DataCite](#)

[METS](#) : Metadata Encoding & Transmission Standard (Bibliothèque du Congrès)

Métadonnées des systèmes de fichiers [basées sur XML](#) et [JSON](#)

[PREMIS](#) : Preservation Metadata Maintenance Activity (Bibliothèque du Congrès)

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

[BagIt](#) : format d'emballage de fichiers hiérarchiques (Library of Congress)

[Schéma PUID](#) : validation du format de fichier (UK National Archives)

[Classification re3data des domaines thématiques](#) (Fondation allemande pour la recherche)

[Liste des domaines de recherche et des disciplines](#) (FNS)

2 TECHNOLOGIES ET APPLICATIONS OPEN SOURCE

Authentification [SWITCHaai](#)

Service de stockage à long terme [SWITCHengines](#)

[DOI](#) (Digital Object Identifier)

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

[Licences Creative Commons](#)

Langage de programmation [Java](#) et [Spring Boot](#)

[Docker](#) : plateforme de conteneurisation open source

[Kubernetes](#) : système d'orchestration de conteneurs open source.

[Angular](#) : framework JavaScript open source

[PHP](#) : langage de script open source

[API RESTful](#) : services web

[IIIF](#) (Cadre d'interopérabilité des images)

[S3](#) (Amazon Simple Storage Service) : stockage d'objets via une interface de service web

Système de fichiers sur bande : stockage des fichiers sur bande magnétique

[FITS](#) : identification du format de fichier (Harvard)

[Elasticsearch](#) : moteur de recherche et d'analyse open source

[ClamAV](#) : moteur antivirus open source

[JSmol](#) : visualisateur Java open source de structures chimiques en 3D

[Jenkins](#) : serveur d'automatisation open source (déploiement et test)

[MySQL](#) et [MariaDB](#) : bases de données relationnelles open source

[ActiveMQ](#) : serveur de messagerie

Fonctionnalité	Prestations
RESPECT DE LA LOI SUISSE	Yareta s'appuie uniquement sur des centres de données existants en Suisse pour le stockage d'archives, ce qui permet de préserver vos données sensibles conformément aux dispositions légales (Loi Fédérale sur la Protection des Données).
CONFORMITÉ AUX PRINCIPES DU FAIR	<p>Yareta suit les principes FAIR et peut bénéficier d'un financement pour le téléversement de données pour les chercheurs dans le cadre d'une subvention du FNS.</p> <p>Trouvable Yareta enregistre vos métadonnées descriptives avec DataCite [F2, F4] afin d'attribuer un identifiant unique et permanent (DOI) à vos ensembles de données [F1] et permet de découvrir toutes les métadonnées descriptives hébergées [F4], qui comprennent les DOI respectifs [F3].</p> <p>Accessible Yareta propose toutes les métadonnées descriptives pour le harvesting et, si les conditions de licence le permettent, les ensembles de données pour le téléchargement sans authentification [A1.2] par le biais de protocoles standard ouverts [A1.1] (OAI-PMH et API RESTful) [A1]. Les données sont conservées conformément à leur politique de conservation respective, et les métadonnées descriptives sont conservées après la purge des données. Yareta est développé à l'Université de Genève et alimenté par DLCM, mandaté par swissuniversities. L'Université de Genève est le garant de la pérennité de la solution. [A2]</p> <p>Interopérable Yareta fournit des métadonnées descriptives et administratives au format XML, ainsi que des métadonnées personnalisées aux formats XML et JSON [I1] utilisant des vocabulaires ouverts reconnus internationalement et provenant de diverses sources reconnues telles que DataCite, les Archives nationales britanniques et la Bibliothèque du Congrès [I2]. Afin d'améliorer la mise en relation des données, les ensembles de données dans Yareta peuvent être annotés avec des métadonnées riches personnalisées correspondant aux normes et aux ontologies de la communauté (telles que DublinCore), et combinés dans des collections reliant les ensembles de données entre eux et avec un référencement qualifié dans le schéma de métadonnées (en utilisant la propriété <code>relatedIdentifier</code>) [I3].</p> <p>Réutilisable Yareta requiert un ensemble obligatoire de métadonnées descriptives provenant du schéma de métadonnées DataCite et des métadonnées administratives des dépôts (METS) pour la gestion des fichiers de données [R1.3]. Il permet la personnalisation avec des métadonnées supplémentaires, y compris des propriétés de métadonnées DataCite "recommandées" et "facultatives", ainsi que des métadonnées riches personnalisées facultatives [R1], validées par leur schéma respectif [R1.3]. Les métadonnées descriptives de tous les ensembles de données sont publiées ouvertement (c'est-à-dire avec un niveau d'accès public) et établissent un lien direct entre l'ensemble de données et la licence (Yareta fournit des licences Creative Commons standard par défaut et permet des licences alternatives personnalisées) [R1.1]. Les métadonnées</p>

descriptives comprennent les détails relatifs à l'[unité organisationnelle](#) qui héberge le jeu de données et de ses [contributeurs et contributrices](#), y compris un lien vers leur identifiant unique de chercheur/euse [ORCID](#), lorsqu'il est fourni [\[R1.2\]](#).

Fonctionnalité	Prestations
OPEN SOURCE	Yareta est entièrement conçu à partir de technologies open source et de protocoles libres, afin que la solution reste flexible, évolutive et ouverte aux développements et mises à jour futurs. La solution DLCM qui alimente Yareta est également open source. Cela signifie que vous pouvez utiliser l'infrastructure qui correspond le mieux à vos besoins institutionnels.
UNIVERSEL	Yareta a une vision agnostique des formats de fichiers et des types d'objets, ce qui lui permet de préserver essentiellement tout type de fichier. Il est conçu pour être générique à la plus large gamme d'ensembles de données de recherche. Yareta ne vise pas à remplacer les repositories spécifiques à une discipline ou à une communauté, mais à compléter leur utilisation en offrant une solution conforme aux normes internationales appliquées aux sciences de l'information sur la préservation, l'accessibilité et la réutilisation des données.
ARCHITECTURE MODULAIRE BASÉE SUR DES API RESTFUL	L'architecture modulaire de la solution permet d'ajouter facilement des fonctionnalités à Yareta sans affecter le noyau. Les développements peuvent se faire indépendamment dans les différents modules de la solution, et les API RESTful assurent l'évolutivité et les performances à grande échelle des modules distribués et des systèmes intégrés. <i>Les API RESTful bien documentées permettent une intégration facile de Yareta avec vos applications et services préférés, et garantissent que vos données ne seront jamais enfermées dans un seul système.</i>
OPTIONS DE DÉPLOIEMENT	<p>Yareta offre une préservation à long terme durable et rentable en Suisse. <i>Yareta propose une flexibilité de déploiement complète : vous pouvez choisir d'utiliser le portail web Yareta ou de déployer Yareta avec de nombreuses formes et configurations différentes, dans un cloud de votre choix, sur site sur votre propre infrastructure, ou un mélange des deux.</i></p> <p>Yareta est une application cloud-native compatible avec un déploiement sur site, Docker ou Kubernetes, fonctionnant sur un groupe de deux ou plusieurs machines hôtes (nœuds) qui possèdent leur propre technologie, une copie des ensembles de données et des sauvegardes. <i>Le nombre de nœuds et leur répartition géographique peuvent être spécifiés.</i> Les copies multiples répliquées sur plusieurs sites atténuent le risque de corruption et de perte de données.</p> <p>En tant que solution basée sur une API, Yareta présente de nombreux avantages. Grâce à son infrastructure modulaire, Yareta permet aux institutions de rationaliser les processus de données critiques, d'appliquer des politiques et de surveiller les performances, tout en réduisant considérablement les coûts de développement et de maintenance.</p>
REDONDANCE ET SÉPARATION GÉOGRAPHIQUE	Yareta suit les bonnes pratiques en conservant plusieurs copies dans des emplacements/systèmes de stockage géographiquement distants afin de protéger vos données contre les catastrophes, les défaillances matérielles ou logicielles et les erreurs humaines. <i>Pour maximiser la durabilité et la disponibilité de vos données, vous pouvez définir une politique de conservation personnalisée basée sur la redondance et l'emplacement des données, en spécifiant le nombre d'instances que vous souhaitez conserver et les emplacements de vos centres de données préférés.</i>
SYSTÈMES DE STOCKAGE D'ARCHIVES MULTIPLES	Yareta fonctionne sur une architecture multi-niveaux et distribuée pour préserver vos données sur des systèmes basés sur des fichiers et sur différents types d'infrastructures de stockage. Yareta s'intègre nativement aux systèmes
• Système de fichiers	

Fonctionnalité	Prestations
<ul style="list-style-type: none"> • Protocole S3 • Bandes de données 	de stockage compatibles Amazon S3, comme la solution de stockage à long terme de SWITCH. Les gros volumes de données peuvent être gérés et hiérarchisés sur des bandes de données magnétiques. <i>Vous pouvez sélectionner le système le plus pertinent pour votre plan de préservation en fonction de vos exigences, de votre budget et de votre stratégie.</i>
SÉCURITÉ DES DONNÉES <ul style="list-style-type: none"> • Métadonnées de préservation PREMIS • Checksums • Traces d'audit 	Pour garantir la fiabilité du service, Yareta utilise différents processus pour confirmer l'authenticité et vérifier de manière indépendante l'intégrité des données. Les métadonnées de préservation (PREMIS) créent une chaîne de confiance centrée sur les données pour leur soumission au repository, pour leur stockage dans les archives et pour leur diffusion à partir du repository. Les fichiers sont traités lors du téléchargement afin de générer des <u>informations d'évaluation clés</u> et des <u>métadonnées</u> pour la gestion des fichiers et la planification de la préservation à long terme et, pour chaque fichier, un checksum est calculé. Les checksums sont utilisés pour s'assurer que l'intégrité du fichier est maintenue dans le stockage d'archives selon un calendrier régulier. Les tâches de planification de la préservation sont suivies à l'aide de tableaux de bord administratifs et de systèmes de notification, et des traces d'audit offrent une vue en amont de l'utilisation du système.
SYSTÈME DE GESTION DES DONNÉES BASÉ SUR DES POLITIQUES	La définition de la fiabilité d'un dépôt repose sur une liste de tâches spécifiées par le <u>modèle de référence OAIS</u> et appliquées à l'aide de politiques informatiques exploitables. Ces politiques spécifient les tâches nécessaires pour gérer la soumission des données (SIP), la création d'archives (AIP) et la diffusion des jeux de données (DIP). <i>Yareta invite les utilisateurs et les utilisatrices à définir certaines propriétés impliquées dans ces politiques de <u>soumission</u>, de <u>conservation</u> et de <u>diffusion</u>, permettant ainsi aux chercheurs de poursuivre comme d'habitude le flux de travail qui leur convient.</i>
INTERFACE UTILISATEUR ANGULAR	Yareta utilise Angular (un framework JavaScript open source) pour créer une expérience utilisateur meilleure et plus efficace. Le framework Angular permet à l'équipe de développement de Yareta de rester concentrée sur la logique de base de la solution DLCM tout en offrant un portail fluide mais puissant, riche en contenu, rapide à charger et facile à utiliser, avec un design d'interface moderne qui propose <i>des thèmes clair et sombre</i> . Yareta intègre Formly, une bibliothèque dynamique (alimentée par JSON) utilisée pour concevoir des formulaires réactifs Angular. <i>Elle permet de personnaliser un formulaire de <u>métadonnées descriptives DataCite étendu</u> pour répondre à vos besoins et faciliter une identification plus efficace de vos ensembles de données.</i>
SUPPORT MULTILINGUE	<i>Yareta prend en charge des ensembles de données en plusieurs langues, et son interface utilisateur est déjà disponible en anglais et en français.</i>

Fonctionnalité	Prestations
AUTHENTIFICATION DE SWITCHAAI	L'accès à Yareta est spécifiquement destiné à toute personne ayant des liens avec les institutions universitaires et de recherche du canton de Genève et leurs partenaires de collaboration. Une authentification SWITCHAAI est requise avant d'accéder à toute autre fonctionnalité de Yareta, et toutes les actions des utilisateurs et utilisatrices sont enregistrées dans une trace d'audit pour la responsabilité. Yareta est ouvert à SWITCHedu-ID, permettant la collaboration avec des partenaires nationaux et internationaux non universitaires. Les autorisations d'accès peuvent être définies à la fois au niveau de l' <u>unité organisationnelle</u> (contrôle de cinq <u>rôles d'utilisateur</u> très précis) et au niveau de l'ensemble de données individuel (choix de la publication <u>fermée</u> , <u>restreinte</u> ou <u>publique</u> de vos données).
INTÉGRATION DE L'ORCID	Yareta propose une <u>intégration</u> ORCID à des fins de désambiguïsation des noms. En associant votre identifiant unique de recherche à votre profil personnel sur Yareta, vous vous assurez que les activités et les résultats de vos données de recherche vous sont correctement attribués.
ESPACE DE PRÉSERVATION <ul style="list-style-type: none"> • Unités organisationnelles • Contributeurs/trices • Demander des notifications 	<p>Votre espace de préservation est votre espace de travail privé sur Yareta. Il est caractérisé par une série d'unités organisationnelles (équipes de recherche, projets de recherche, départements, laboratoires, etc.), un groupe de contributeurs (membres des unités organisationnelles) et un ensemble de notifications (demandes envoyées et demandes reçues).</p> <p>Dans votre Espace de préservation, vous pouvez répertorier toutes les Unités organisationnelles pour consulter leurs archives <u>publiques</u>, accéder aux dépôts en cours des Unités organisationnelles dont vous êtes membre, et demander l'adhésion d'autres Unités organisationnelles pour consulter leurs archives <u>non publiques</u>. Vous pouvez également répertorier tous les contributeurs de vos Unités organisationnelles et leurs dépôts. En outre, <i>vous pouvez modifier les unités organisationnelles dont vous êtes le/la <u>Manager</u> et contrôler qui peut modifier les dépôts en leur sein en attribuant un <u>rôle</u> spécifique à chaque contributeur et chaque contributrice.</i></p> <p><i>Yareta vous notifie les demandes des personnes souhaitant accéder à vos archives <u>non publiques</u>. Vous pouvez accorder aux personnes un accès <u>visiteur</u> à vos unités organisationnelles afin qu'elles puissent consulter toutes vos archives <u>restreintes</u>, et choisir d'accepter ou de refuser les demandes individuelles d'accès à vos archives <u>fermées</u> au cas par cas.</i></p>
CONTRÔLE D'ACCÈS BASÉ SUR LES RÔLES	Yareta utilise un contrôle d'accès basé sur les rôles pour accorder les permissions appropriées. Tous les <u>contributeurs et toutes les contributrices</u> accèdent aux <u>unités organisationnelles</u> à l'aide de comptes authentifiés, et se voient attribuer des rôles tels que Manager, Steward, Approver, Creator ou Visitor. <i>Yareta délègue les responsabilités de gestion d'une <u>unité organisationnelle</u> au Manager, qui définit les différentes politiques régissant l'utilisation de l'unité et l'accès aux données qu'elle contient.</i> Les <u>membres</u> d'une <u>unité organisationnelle</u> partagent l'accès aux jeux de données et peuvent effectuer certaines actions dans les dépôts. Pour plus d'informations sur les actions que chaque rôle peut effectuer, voir <u>une explication détaillée des rôles</u> .
GESTION DE LA POLITIQUE DE SOUMISSION	<i>Yareta vous permet d'ajouter une étape de validation à votre flux de travail.</i> La politique de soumission définit le flux de travail qui contrôle la validation de vos dépôts. Elle définit comment les <u>membres</u> peuvent collaborer pour publier un

Fonctionnalité	Prestations
	<p>jeu de données, et plus précisément qui peut approuver ou non un jeu de données. Choisir "Dépôt avec approbation" comme politique de soumission permet à un Approver ou à un Steward (expert du domaine, archiviste, bibliothécaire, etc.) d'examiner les jeux de données envoyés à la validation et de les approuver, de les rejeter ou de les renvoyer à l'édition pour modification. Avec la politique "Dépôt sans approbation", les membres auxquels sont attribué le rôle de Creator dans votre unité organisationnelle peuvent également décider de valider eux-mêmes les jeux de données.</p>
<p>GESTION DE LA POLITIQUE DE PRÉSERVATION</p>	<p>La politique de conservation des données définit la durée de conservation de vos données par Yareta. Dans le cadre des bonnes pratiques scientifiques, Yareta propose les périodes de conservation recommandées de 10 ans pour les données de recherche et de 15 ans pour les dispositifs médicaux, mais <i>vous pouvez définir votre propre politique de conservation pour votre unité organisationnelle</i>. Lorsque la période de conservation prend fin, Yareta supprime les données du stockage d'archives. Vos métadonnées descriptives sont conservées, de sorte que votre DOI reste un moyen permanent d'identifier votre ensemble de données.</p>
<p>GESTION DE LA POLITIQUE DE DIFFUSION [FONCTIONNALITÉ À VENIR].</p> <ul style="list-style-type: none"> • Publication sur le Web • Politique de diffusion de l'IIIF 	<p><i>Yareta offre un ensemble indépendant de politiques de diffusion basées sur des protocoles et des normes d'interopérabilité pour assurer une portabilité rapide et transparente de vos données vers différents systèmes d'information, dispositifs et applications. Par défaut, Yareta prend en charge la publication sur le web et le déploiement de vos ensembles de données vers des serveurs et des sites web personnalisés. Une politique de diffusion IIIF indépendante permet la publication de fichiers de données d'images spécifiques vers des serveurs et des outils basés sur IIIF, comme par exemple le système de catalogue du laboratoire Bodmer.</i></p> <p><i>L'élaboration de politiques de diffusion personnalisées peut vous offrir une plus grande facilité de gestion pour améliorer la visibilité et l'utilisation de vos ensembles de données dans tous les systèmes.</i></p>
<p>FLUX DE TRAVAIL PRÉ-INGESTION PRÉSERVANT LA HIÉRARCHIE DES DOSSIERS ET LA STRUCTURE DES DONNÉES.</p>	<p>Yareta comprend un module de pré-ingestion qui peut acquérir, décompresser et ingérer automatiquement des données à l'échelle, pour un dépôt transparent des ensembles de données à partir d'une variété de systèmes. Le module de pré-ingestion peut extraire, lister et tester des données compressées à partir d'un fichier ZIP, tout en préservant la hiérarchie des dossiers et la structure des données dans la solution.</p>
<p>ÉVALUATION AUTOMATISÉE DES DOSSIERS</p> <ul style="list-style-type: none"> • Analyse de virus • Identification du format de fichier • Évaluation du niveau de conformité des données • Filtrage des fichiers 	<p>Yareta associe l'identification du format de fichier (FITS) à l'évaluation du niveau de conformité des données, en utilisant des registres structurés de politiques de format fournis par les Archives nationales britanniques (schéma PUID) et la Bibliothèque du Congrès (Recommended Formats Statement), reconnu au niveau international, afin d'encourager l'utilisation de normes ouvertes. <i>Un filtrage à grain fin est appliqué afin que vous puissiez valider, ignorer ou exclure des fichiers en fonction de leur statut après évaluation automatique et analyse antivirus (ClamAV).</i></p>
<p>DES MÉTADONNÉES RICHES</p> <ul style="list-style-type: none"> • Métadonnées descriptives DataCite • Métadonnées administratives METS 	<p>Grâce à la richesse des métadonnées, Yareta garantit que vos informations pourront être trouvées et identifiées à l'avenir.</p> <p>Yareta intègre le schéma de métadonnées DataCite pour faciliter la citation, la découverte et la persistance des ensembles de données. Il exige un ensemble minimal de métadonnées descriptives au cours du processus de soumission afin qu'un identifiant d'objet numérique (DOI) persistant puisse être généré et</p>

Fonctionnalité	Prestations
<ul style="list-style-type: none"> • Métadonnées personnalisées XML et JSON 	<p>attribué à chaque ensemble de données. <i>Pour une identification plus précise et plus cohérente de vos ensembles de données, vous pouvez très facilement adapter le formulaire à vos propres besoins avec des champs choisis dans le pool de propriétés de métadonnées de DataCite.</i> Yareta fournit des services de découverte sur toutes les métadonnées descriptives à des fins de citation et de recherche.</p> <p>Les métadonnées administratives (METS) sont collectées automatiquement lors du dépôt, permettant un suivi interne et des objectifs de gestion. Elles soutiennent les fonctions spécifiques de préservation et de diffusion de Yareta en fournissant des informations qui aident à comprendre les fichiers de données, à garantir leur fixité et leur intégrité, et à permettre leur accès et leur utilisation par la communauté désignée.</p> <p><i>Yareta vous donne la possibilité d'associer des métadonnées personnalisées à votre ensemble de données pour décrire vos données. Vous pouvez définir un schéma de métadonnées dans un format unique et significatif pour votre domaine de recherche, tel que XML ou JSON. Yareta valide à la fois le type (format) du fichier et son contenu (structure) par rapport au schéma de métadonnées.</i></p>
<p>DATE DE PUBLICATION ET PÉRIODE D'EMBARGO FACULTATIVE</p>	<p><i>Yareta vous permet de déterminer la date de publication si votre jeu de données ne doit pas être publié immédiatement. Une période d'embargo facultative peut également être spécifiée, ainsi qu'un niveau d'accès plus exclusif. Vos données ne seront accessibles qu'après l'expiration de l'embargo. Les métadonnées descriptives restent consultables, afin de garantir la facilité de recherche de votre ensemble de données.</i></p>
<p>GESTION DE L'ACCÈS AUX DONNÉES</p>	<p>Yareta rend vos données et métadonnées accessibles de manière sécurisée et contrôlée, afin que vous puissiez vous conformer aux exigences légales. <i>La gestion flexible des accès vous offre un certain contrôle sur qui peut accéder à vos ensembles de données et comment.</i> Lorsque le partage public de vos données de recherche n'est pas approprié - lorsque l'anonymisation ne peut pas être effectuée efficacement ou en cas d'accord commercial - vous pouvez utiliser les niveaux d'accès Restreint ou Fermé, qui garantissent que les données de recherche ne seront pas publiées ouvertement. Seuls les membres de votre unité organisationnelle pourront voir vos jeux de données restreints, et ils devront soumettre une demande pour que vous puissiez leur accorder l'accès à vos jeux de données fermés. Les métadonnées DataCite, qui décrivent l'ensemble de données, restent disponibles afin d'en assurer la recherche, mais les données proprement dites ne sont mises à la disposition des utilisateurs autorisés que sur demande.</p>
<p>GESTION DES LICENCES</p>	<p>En tant que meilleure pratique, Yareta vous encourage à appliquer une licence à vos ensembles de données afin de conserver vos droits d'auteur et de contrôler l'attribution, l'utilisation, la distribution et la modification de vos données, en informant les autres de ce que vous leur permettez de faire avec vos données. Yareta propose un choix par défaut de licences ouvertes. <i>Vous pouvez demander l'inclusion d'autres licences ouvertes largement utilisées, et des conditions supplémentaires personnalisées, modifiées ou non standard peuvent être créées sur demande.</i></p> <p>Dans le respect des principes FAIR, une licence doit être spécifiée pour chaque jeu de données diffusé en accès public avec Yareta. Sans choix, une licence par défaut CC-BY (Creative Commons Attribution 4.0) s'appliquera. <i>Vous pouvez</i></p>

Fonctionnalité	Prestations
	<i>également définir votre licence par défaut préférée dans votre unité organisationnelle. [fonctionnalité à venir]</i>
IDENTIFIANT PERSISTANT (DOI)	Yareta fournit l'enregistrement des métadonnées avec le DOI (Digital Object Identifier) de DataCite pour rendre le partage des données FAIR, en garantissant que votre jeu de données est identifiable de manière unique et permanente et qu'il peut être cité dans les publications scientifiques. Chaque jeu de données publié avec Yareta reçoit un DOI au moment de la publication. <i>Si nécessaire, des DOI peuvent être réservés pour un jeu de données au moment du dépôt, avant de le mettre à la disposition du public, ce qui vous permet de citer vos données dans tout article en préparation.</i>
RÉCOLTE DE MÉTADONNÉES VIA OAI-PMH	Chaque ensemble de données publié avec Yareta est automatiquement indexé avec DataCite et proposé pour harvesting via le protocole standardisé OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) afin de garantir une diffusion et une trouvabilité maximales de vos données de recherche.
SOUTIEN DE L'IIIF	Yareta est capable de fournir des images et des métadonnées conformément aux spécifications IIIF, le cadre international d'interopérabilité des images pour la publication et le partage de contenu.
TROUVER DES DONNÉES <ul style="list-style-type: none"> • Navigation dans les archives • Recherche de mots-clés • Filtrage 	<p>Yareta fournit des services de découverte sur toutes les métadonnées descriptives hébergées. Pour trouver des jeux de données, vous pouvez parcourir les noms des unités organisationnelles et des ensembles de données, ou saisir un ou plusieurs mots clés dans le champ de recherche. Lorsque vous êtes connecté à votre compte Yareta, un filtre vous offre des options de recherche supplémentaires pour restreindre votre recherche aux jeux de données publics (y compris les données téléchargeables), ou privés, c'est-à-dire les jeux de données à accès restreint ou fermé, comprenant uniquement les métadonnées descriptives, pour lesquels vous pourrez demander un accès.</p> <p>Vous pouvez utiliser les types d'expressions de recherche et les opérateurs booléens les plus courants pour affiner votre recherche :</p> <ul style="list-style-type: none"> • AND : trouve les résultats où les deux termes à côté de AND sont présents (une recherche sans opérateur comprend un AND implicite entre les termes saisis). • OR : recherche les résultats contenant l'un ou l'autre des termes saisis. • - [moins] : Recherche les résultats ne comprenant pas le terme qui suit le signe "moins". • " " [guillemets] : Recherche les résultats contenant l'expression exacte entre les guillemets (la ponctuation est ignorée). • * [astérisque] : Recherche les résultats contenant les termes commençant par la ou les lettres suivies de l'astérisque (renvoie toutes les archives publiées si utilisé seul).
INTÉGRATION	<p>Yareta sert de solution intégrée avec un repository central pour la conservation et la publication à long terme des données de recherche, qui peut être synchronisé avec de nombreuses autres applications satellitaires. La solution supporte déjà différentes technologies propriétaires et open source. Elle fournit des API RESTful en Python pour intégrer facilement tous les instruments, applications et centres de données que vous utilisez, afin que vos données soient à jour et complètement interconnectées avec un minimum d'effort manuel.</p> <p>Exemples de projets d'intégration</p> <ul style="list-style-type: none"> • Connectez les instruments de laboratoire, les LIMS (Laboratory Information Management System) et les ELN (Electronic Laboratory Notebook platform), tels que SLIMS (Genohm) et openBIS (ETHZ), pour alimenter

Fonctionnalité	Prestations
	<p>automatiquement les ensembles de données avec les résultats, les emplacements, les propriétés et les échantillons de vos expériences.</p> <ul style="list-style-type: none"> • Connectez des outils de visualisation et d'analyse, tels que Tableau, pour créer des tableaux de bord et effectuer des analyses. • Intégrez toute autre application pour extraire automatiquement vos données et les garder synchronisées, y compris les solutions personnalisées ou propriétaires. Par exemple, Yareta intègre Jupyter Notebooks pour vous permettre d'écrire des scripts Python personnalisés pour traiter vos données. • Intégrez des centres de données externes pour sauvegarder vos données sur votre propre Amazon S3 ou d'autres systèmes d'enregistrement. <p><i>Yareta peut être intégré à tout service en ligne ou logiciel offrant une fonctionnalité API. Consultez notre Guide d'intégration et notre Guide de l'API RESTful pour savoir comment connecter Yareta à votre environnement de recherche.</i></p>
<p>TAGS</p> <ul style="list-style-type: none"> • Domaines de recherche • Mots clés 	<p>Yareta utilise des balises pour catégoriser et filtrer vos données. Il existe deux types de balises : les balises de domaine de recherche et les balises de mots-clés. Les balises de domaine de recherche peuvent être utilisées pour catégoriser le travail de votre unité organisationnelle sur la base de registres suisses et internationaux, tels que les domaines et disciplines de recherche du FNS et la classification des domaines de la Fondation allemande pour la recherche utilisée par re3data. Les balises de mots-clés peuvent être utiles pour préciser le contenu et l'objectif de vos jeux de données par rapport à un sujet ou un domaine de recherche spécifique, en complément du titre et de la brève description des données.</p>
<p>FONCTIONNALITÉ DE PRÉVISUALISATION</p>	<p>Yareta comprend un outil de visualisation universel permettant de présenter aux utilisateurs et aux utilisatrices du repository une grande sélection de types de données (documents, images, audio, vidéo, etc.), y compris par exemple des affichages 2D et 3D dynamiques de fichiers de structure moléculaire utilisant JSmol.</p>
<p>RESSOURCES</p> <ul style="list-style-type: none"> • Infobulles • Guide de l'utilisateur • Documentation technique 	<p>Yareta a créé un ensemble de ressources pour vous aider à comprendre ses différentes fonctionnalités. Vous pouvez apprendre à utiliser la solution de différentes manières : en accédant au guide de l'utilisateur et aux infobulles dans l'application, en lisant la documentation technique du DLCM sur la Sandbox DLCM (Guide d'intégration, Guide de l'API RESTful et Guide des outils DLCM), et en explorant les ressources sur la page web de Yareta. Les infobulles (étiquettes à l'écran) fournissent des informations utiles sur certaines options lorsque vous placez le pointeur dessus. Si vous avez besoin d'informations supplémentaires sur une fonctionnalité, vous pouvez accéder au guide de l'utilisateur à tout moment depuis l'application.</p>
<p>FONCTIONNALITÉS PERSONNALISABLES</p>	<p>Yareta offre un haut degré de liberté avec des fonctionnalités adaptées à votre environnement de travail habituel et à vos préférences personnelles. (Les fonctionnalités personnalisables sont identifiées en italique dans le texte).</p> <p>Fonctionnalités personnalisables de Yareta :</p> <p>Votre profil</p> <ul style="list-style-type: none"> • Reliez votre ORCID à votre profil personnel pour que vos ensembles de données vous soient correctement attribués. • Choisissez entre les modes clair et foncé <p>Votre espace de préservation</p>

Fonctionnalité	Prestations
	<ul style="list-style-type: none"> ● Gardez une vue complète sur les unités organisationnelles, leurs membres et leurs dépôts. ● Gérer les notifications de demande (pour demander, accorder ou refuser l'accès) ● Avoir des responsabilités qui vous sont déléguées : en tant que Manager, vous contrôlez les politiques de l'unité organisationnelle, ses membres et leur niveau d'accès (rôles). ● Mettez en œuvre votre propre flux de travail en choisissant d'activer une étape de validation facultative dans votre politique de soumission. ● Définissez votre politique de préservation pour respecter les délais de conservation, de préservation et d'accès exigés par votre bailleur de fonds ou votre institution. ● Optez pour la politique de diffusion qui convient à vos besoins (publication sur le web ou sur des serveurs d'images) ● Incluez votre discipline et des mots-clés pour améliorer la visibilité de vos ensembles de données dans votre communauté de recherche. ● Définir une licence par défaut pour votre unité organisationnelle <p>Vos dépôts</p> <ul style="list-style-type: none"> ● Réservez un DOI avant la publication de votre ensemble de données. ● Choisissez de valider, d'ignorer ou d'exclure des fichiers en fonction de leur statut après évaluation automatique et analyse antivirus. ● Choisissez votre date de publication et appliquez un embargo facultatif. ● Définissez qui peut accéder à vos données et comment (public, restreint ou fermé). ● Choisissez des licences parmi une sélection de licences ouvertes largement utilisées et demandez des conditions supplémentaires personnalisées, modifiées ou non standard. ● Choisissez une période de conservation pour vos jeux de données parmi les politiques de conservation définies au niveau de l'unité organisationnelle. ● Ajouter ou modifier des métadonnées, en développant le formulaire de métadonnées descriptives de DataCite (propriétés optionnelles) et en personnalisant vos métadonnées selon un schéma choisi (XML ou JSON) ● Connectez Yareta à votre environnement de recherche par le biais d'API RESTful (intégrations tierces).

Ce texte est mis à disposition sous la licence CC-BY-SA.

Crédit photographique : UNIGE, Jacques Erard, 2018 <https://phototheque.unige.ch/unige:121641>

Contact :

Recherche en ligne
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