
PGD du projet "Bioimagerie - Plateforme MICA"

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "Horizon 2020 FAIR DMP (anglais)" fourni par Commission européenne.

Plan Details

Plan title	PGD du projet "Bioimagerie - Plateforme MICA"
Version	First version
Fields of science and technology (from OECD classification)	Biological sciences (Natural sciences)
Language	fra
Creation date	2019-02-01
Last modification date	2022-06-03
Identifiant	2389
Identifiant type	DOI
License	Etalab Open License 2.0

Project Details

Project title	Bioimagerie - Plateforme MICA
Abstract	<p>Le projet Mutimage UCA soutenu par l' Académie 4 de l'Idex UCA JEDI, est de proposer à l'ensemble de la communauté en sciences de la vie de l'Université Côte d'Azur (UCA) un outil cadrant les processus de gestion des images et donnant un accès aux ressources de façon complètement délocalisée. Cet outil conçu pour le partage des images permettra aussi à la communauté locale de bénéficier plus largement des outils d'analyse d'image développés sur la plateforme MICA. Le projet s'appuie sur une base de données OMERO installée et gérée pour la partie hardware en partenariat avec l'Institut Français de BioInformatique.</p>

Funding	<ul style="list-style-type: none">• Université Côte d'Azur : MutImage_UCA
----------------	---

Start date	2019-01-01
-------------------	------------

Produits de recherche :

1. Images de la biologie et océanologie (Image)

Contributeurs

Nom	Affiliation	Rôles
BEKKOUCHE Faisal	CNRS	
Frédéric BRAU	CNRS	<ul style="list-style-type: none"> • DMP manager • Personne contact pour les données • Project coordinator

Droits d'auteur :

Le(s) créateur(s) de ce plan accepte(nt) que tout ou partie de texte de ce plan soit réutilisé et personnalisé si nécessaire pour un autre plan. Vous n'avez pas besoin de citer le(s) créateur(s) en tant que source. L'utilisation de toute partie de texte de ce plan n'implique pas que le(s) créateur(s) soutien(nen)t ou aient une quelconque relation avec votre projet ou votre soumission.

PGD du projet "Bioimagerie - Plateforme MICA"

1. Data summary

L'objectif de cette base de données est de proposer aux équipes de recherche en sciences de la vie de l'Université Côte d'Azur et EMBRC France un outil pour gérer leurs images dans le cadre de leurs projets de recherche. Le périmètre de cette base de données se limite aux entités sus-nommées et leurs collaborateurs externes académiques nationaux et internationaux dans le cadre de collaborations scientifiques. Elle est destinée à accueillir des images scientifiques uniquement.

Cet outil a pour but de faciliter la gestion, le partage et l'analyse des images. Par ailleurs cette base de données servira de support pour organiser, coordonner, structurer et mutualiser les ressources et expertises en analyse d'images quantitative en biologie.

Mode d'obtention : données déposées par des utilisateurs d'appareils d'acquisition d'images, principalement générées par des microscopes. Ces images sont destinées à être partagées de façon libre entre collaborateurs identifiés d'un même groupe ou avec des collaborateurs externes identifiés.

Origine des données : images expérimentales correspondant à des observations microscopiques, récolte d'échantillons marins, de photos sous-marines...

Type de données : images multidimensionnelles : 2D, 3D, associant des informations spectrales et/ou temporelles, images pyramidales histologiques de tissus, images à haut contenu. Avec la possibilité de lier aux images des fichiers d'information ou d'analyse de différents types (tableaux graphiques, graphes, documents texte, présentations, vidéos)

Nature : images de biologie cellulaire, d'histologie, organismes vivants, tissus, organes

Périmètre thématique : Images scientifiques, biologie cellulaire, biologie du développement, neurosciences, biologie marine. Cette base de données n'a pas vocation à accueillir des images sensibles (santé, humain), sous réserve qu'elles soient anonymisées et liées à un projet de recherche ou d'enseignement.

Format de données : Images propriétaires de microscopie (.ism, .lif, .oif, .oib, .czi, .stk, .mvd2...) ou en TIFF Open Microscopy Environment Bioformat (OME TIFF) : TIFF associant des métadonnées au format XML.

Réutilisation : Les images sont susceptibles d'être réutilisées pour des analyses d'images par des logiciels tiers via des passerelles permettant de déposer automatiquement les résultats dans la base de donnée ou par des téléchargements, analyses déportées...

Volumétrie de stockage : la capacité originale d'accueil des données sur le serveur de l'Institut Français de Bioinformatique est de 15To. Cette capacité pourra être étendue.

Utilité des données : images scientifiques de projets en cours pour discussions et échanges quotidiens au sein des équipes de recherche et/ou avec des collaborateurs externes.

2. FAIR data

Standard utilisé :

Le format de Fichier utilisé par la base de donnée OMERO est le format standard Open Microscopy Environment OME-XML. C'est un format de fichier indépendant du fournisseur pour les données d'images biologiques, incluant les métadonnées courantes de la microscopie photonique, importées lors de la conversion (système, objectifs, échelles, canaux, détecteurs, réglages...). Il peut être utilisé comme format de fichier de données à part entière, ou comme moyen d'encoder des métadonnées dans un fichier TIFF ou BigTIFF (pour lequel il existe la spécification OME-TIFF). La norme est maintenue par l'Open Microscopy Environment Consortium et a été mise à jour pour la dernière fois en juin 2012. Seuls les images dont les formats (150) sont compatibles avec le standard OME pourront être ajoutés dans la base de données.

Conventions de nomage : aucune convention n'est imposée à la base de données. Pour des questions éventuelles de migration des données et pour faciliter leur administration par les équipes de recherche et les spécialistes en analyse d'images, il sera préférable de privilégier des noms de fichiers courts (<50 caractères) sans espaces ni caractères spéciaux, avec des "_". Par exemple "Date_Nom de l'expérience_Conditions_Numerotation"

Métadonnées complémentaires : Les utilisateurs pourront associer des métadonnées complémentaires lors de l'importation des images dans la base avec l'outil Omero Insight ou directement via l'accès en ligne. Ces métadonnées seront liées aux paramètres de prise de vue, aux caractéristiques des échantillons et de leur préparation. Elles seront propres aux groupes constitués et seront intégrées aux images dans le format OME-XML.

"Findable" : les projets, datasets, images, figures, fichiers attachés, etc. sont associées à un identifiant unique (ID) et une URL ou URI unique. Des mots clefs ("tags" ou "key values") peuvent être ajoutées aux images par l'utilisateur à l'importation ou lors de la consultation. Ces identifiants et les mots clefs associés peuvent être utilisés pour accéder aux données grâce à des outils de recherche inclus dans la base.

Accessibilité

Méthodes et logiciels d'accès

Chaque utilisateur identifié auprès de son référent peut avoir un accès à la base de données à l'adresse : <https://bioimage.france-bioinformatique.fr/omero-mica/webclient/login/> :

* via un compte utilisateur avec un login unique et mot de passe,

* ou à un groupe public sans restriction d'accès avec des restrictions d'écriture. Ce groupe pourra accueillir des images libres de droit ou sous license Etalab 2.0 compatible avec Creative Commons (CC-BY), sans faire office de plateforme de publication.

Les données sont disponibles via l' [interface web](#) sus-nommée ou un client Java (Omero Insight ou Importer) pouvant s'installer localement sur les postes de travail individuels. La base de données est également accessible via une interface de programmation (API d'Omero), mais également au travers d'outils d'analyse, logiciels tiers, et déconvolution (Huygens Remote Manager, logiciels d'analyse d'image, de programmation).

L'ensemble des modalités d'accès sont disponibles sur les sites web <http://mica.unice.fr> : des liens pour accéder à l'interface web d'accès ou vers les pages de l'équipe Omero (<https://www.openmicroscopy.org/>), les distributions logicielles, les documentations associées, FAQ...

La base de données comporte :

-d'une part des groupes privés nominatifs par équipe de recherche auxquels sont associés des comptes individuels nominatifs. Chaque équipe a un groupe contenant les comptes des membres de l'équipe et un groupe de collaboration pouvant contenir les comptes des membres de l'équipe et des comptes de collaborateurs extérieurs au périmètre (laboratoires en sciences de la vie de l'Université Côte d'Azur et EMBRC France) de la base. A leur demande 1 à 2 groupes de collaborations supplémentaires pourront être créés. A l'intérieur de ces groupes les différents comptes inclus auront accès à l'ensemble des images et données du groupe. Pour accéder aux comptes une authentification individuelle est nécessaire.

Les référents sont regroupés dans un groupe dédié. Ils créent/suppriment les groupes, les comptes utilisateurs.

Les personnels des services d'imagerie de la plateforme MICA et des plateforme d'EMBRC France posséderont des comptes et seront ajoutés aux comptes collaborations des équipes en fonction de leur soutien à ces équipes. Il n'y aura pas de groupe dédié à ces services.

Protection des données

L'objectif de la base étant d'offrir un espace de gestion de projets de recherche en cours, les groupes d'équipes sont privés pour protéger les travaux de recherche confidentiels des équipes préalables aux publications.

Interopérabilité : Omero permet :

* un import direct (grâce à Omero Insight) des images depuis des postes d'acquisition d'images ou stations personnelles dans des formats propriétaires (150 formats, voir rubrique "Findable") qui sont lisibles dans la base grâce au standard OME.

* Via l'interface web : le partage, la visualisation, l'édition, l'analyse et la modification des images pour créer des figures sans modification des données brutes quelque soit le poste de travail, le navigateur web ou le système d'exploitation.

* Un export des images brutes ou modifiées, de leurs métadonnées, des figures, aux formats d'origine ou OME-TIFF depuis/vers des logiciels propriétaires (Matlab) ou libres (ImageJ, ICY, Fiji, Cell Profiler, Orbit, R, Python...) au travers de passerelles logicielles (plugins).

* Il est possible que les images d'un même dataset proviennent d'un même échantillon, d'une même acquisition sur cet échantillon. Dans certains cas de formats de fichiers constructeur plusieurs séries d'images peuvent être liées à un même fichier sans pouvoir défaire ce lien. D'autre part, la création de figures (avec OMERO.Figure) contenant des images de différents projets et datasets engendre de fait un lien entre ces images.

* Via les arborescences et mots-clefs il est possible d'avoir des liens vers d'autres jeux données : génomique, protéomique, etc. Il n'y pas de protocole défini pour l'instant à moins d'utiliser 1) la rubrique "Image details" d'OMERO pour ajouter des informations ou des liens URL, 2) des fichiers attachés. Ces deux entrées peuvent être liées au différents niveaux d'arborescence d'OMERO (projets, datasets, images).

Il n'y pas de standard imposé d'appellation des images ou des mots clefs, chaque groupe et compte étant responsable de ses nomages et mots clefs, tout en étant vigilants à une convention de nomage a minima facilitant les exports et analyses (voir rubrique "Findable"). Dans un esprit d'amélioration continue des outils d'ajout de métadonnées contextuelles (conditions expérimentales et instrumentales, modèles biologiques) seront proposés dans le futur, en adéquation avec l'ontologie proposée dans la base de données "Image Data Resource" (IDR) éventuellement sous le format ISA Commons. Il serait possible d'utiliser le plugin MapR qui permet d'utiliser des attributs liés au datasets et aux projets à condition de mettre en place un ontologie commune.

Réutilisable : les images disponibles dans la partie publique de la base de données (enseignement et formations) sont disponibles sous license Etalab 2.0, ce qui sous-entend de mentionner la paternité de l'«Information» : sa source (au moins le nom du «Concédant») et la date de dernière mise à jour de l'«Information» réutilisée.

Les images déposées dans la base données par les équipes dans leurs groupes respectifs ont une durée de vie limitée à 5 ans (avec prolongation possible de 1 an supplémentaire sur demande). Au delà de cette période les équipes seront invitées à déposer les images ayant conduit à valorisation dans la base de donnée "Image Data Resource" (IDR) sous license Creative Commons (CC-BY, compatible avec Etalab 2.0)

3. Allocation of resources

Administration

La base de données est hébergée au centre de calcul de l'IN2P3 à Lyon et gérée par l'Institut Français de BioInformatique (IFB). L'IFB assure la maintenance logicielle, matérielle (mises à jour, montée en capacité de stockage et de calcul, sécurité) et la sauvegarde des données (réplication et rétention sur 90 jours). Cette administration se fait en lien avec les porteurs de projet pour UCA et EMBRC France. Elle garantit une sauvegarde, un accès permanent et sécurisé aux ressources.

Localement, sur les différents sites UCA et EMBRC France des référents par site administrent les groupes et les comptes utilisateurs et sont coordonnés par les pilotes du projet. Ils apportent une assistance par site et au travers d'un alias de support commun qui regroupe les référents, pilotes et l'administrateur : omero-support@listes.univ-cotedazur.fr

Les administrateurs et référents communiquent avec l'ensemble des utilisateurs avec l'alias : omero-users@listes.univ-cotedazur.fr pour les informer des mises à jour, des interruptions de service, des formations...

Gestion des groupes et comptes

La base de données comporte :

- d'une part des groupes privés nominatifs par équipe de recherche auxquels sont associés des comptes individuels nominatifs. Chaque équipe aura donc un groupe contenant les comptes des membres de l'équipe et un groupe de collaboration pouvant contenir les comptes des membres de l'équipe et des comptes de collaborateurs extérieurs au périmètre (laboratoires en sciences de la vie de l'Université Côte d'Azur et EMBRC France) de la base. A leur demande 1 à 2 groupes de collaborations supplémentaires pourront être créés. A l'intérieur de ces groupes les différents comptes inclus auront accès à l'ensemble des images et données du groupe. Pour accéder aux comptes une authentification individuelle est nécessaire.

Les référents sont regroupés dans un groupe dédié. Ils créent/suppriment les groupes, les comptes utilisateurs.

Les personnels des services d'imagerie de la plateforme MICA et des plateformes d'EMBRC France posséderont des comptes et seront ajoutés aux comptes collaborations des équipes en fonction de leur soutien à ces équipes. Il n'y aura pas de groupe dédié à ces services.

- d'autre part un groupe public sans restriction d'accès avec des restrictions d'écriture. Ce groupe pourra accueillir des images libres de droit ou sous licence Etalab 2.0, compatible avec Creative Commons (CC-BY), sans faire office de plateforme de publication.

Gestion des mots de passe

Les référents de chaque institut doivent choisir un mot de passe « robuste », c'est-à-dire composé d'au moins dix caractères alphanumériques, d'une majuscule et un ou plusieurs caractères spéciaux. L'utilisateur devra modifier le mot de passe lors de la première connexion avec un niveau de sécurité au moins équivalent.

La base de données OMERO présente actuellement une méthode d'authentification avec un niveau de sécurité faible. En effet, l'utilisateur a la possibilité de choisir un mot de passe sans aucune restriction. Dans ce contexte, afin de renforcer la sécurité de la base de données, nous envisageons de nous tourner, si c'est possible, vers une fédération d'identité s'appuyant sur les technologies SSO (« Single Sign-on » désignant les technologies qui permettent aux utilisateurs de mutualiser la phase d'authentification entre plusieurs services). Ceci aussi pour faciliter le travail des utilisateurs sur plusieurs systèmes, tout en réduisant la charge administrative liée à la gestion de l'accès à ces systèmes.

Temporalité des comptes :

- Les comptes des utilisateurs contractuels ayant quitté les équipes (doctorants, stagiaires, post-doctorants) seront désactivés.

Préalablement les images seront détruites ou réaffectées au chef d'équipe du groupe concerné et resteront accessibles pour ce groupe.

Cette mise à jour manuelle des comptes sera effectuée deux fois par an, en septembre et janvier. Les administrateurs de la base enverront un message aux référents qui feront remonter les départs d'utilisateurs de leurs laboratoires. En cas de changement de propriétaire des images un formulaire en ligne est disponible pour spécifier le compte de départ, d'arrivée et les projets concernés. Ceci pour permettre à l'administrateur (F. Bekkouche) d'effectuer l'ensemble des changements de propriétaires répertoriés en une seule fois.

Les comptes des utilisateurs supprimés seront dénomés "A_utiliser_numero", désactivés et affectés au groupe "User désactivés"

- Dans le futur, une procédure de curation automatique des comptes est envisagée suivie d'une vérification annuelle. Les référents seront informés de la liste des comptes inusités et valideront le changement de propriétaire des images, voire leur suppression.

Durée de conservation des données : Cette base de données ayant vocation à être utilisée pour gérer des projets, elle ne constitue pas une solution de stockage pérenne et les images seront conservées pour une durée maximale de 5 ans. Les utilisateurs seront avertis par courriel pour supprimer les projets ou images dépassant la durée maximale prévue. Une prolongation pourra être accordée exceptionnellement pour une durée maximale de 1 an. Dans l'esprit FAIR les utilisateurs de la base de données seront invités à déposer les images ayant conduit à valorisation dans la base de données "Image Data Resource" (IDR) sous licence Creative Commons (CC-BY).

Politique de coûts

L'accès aux ressources matérielles et environnantes de cette base est régi par l'établissement d'une convention entre Université Côte d'Azur et l'Institut Français de Bioinformatique qui définit les services accessibles, la durée, la volumétrie, les coûts humains...

4. Data security

Une sauvegarde automatique est effectuée 2 fois par jour à midi et à minuit avec une rétention de 90 jours.

Un certificat est installé sur le serveur UCA, le certificat active le cadenas et le protocole « https » dans les navigateurs, afin d'assurer une connexion sécurisée entre le serveur web et le navigateur.

5. Ethical aspects

La base de donnée héberge des informations à caractère personnel pour y avoir accès : nom, prénom, adresse électronique, laboratoire et équipe d'appartenance. Seuls les référents et administrateurs qui gèrent les comptes ont accès à ces informations.

L'accès aux données dans les groupes d'équipes est régie par les équipes et les référents qui peuvent limiter l'accès.

Le dépôt d'images sur la partie publique est sous la responsabilité de l'utilisateur qui les y dépose.

Cette base de données n'a pas vocation à accueillir des images sensibles (santé, humain), sous réserve qu'elles soient anonymisées et liées à un projet de recherche ou d'enseignement et que cela réponde au cadre légal en vigueur : <https://www.cnil.fr/fr/recherche-medicale-quel-est-le-cadre-legal>

6. Other

Cas particulier d'utilisateurs hors périmètre dans le cadre de prestations externes :

Les utilisateurs externes des services d'imagerie de la plateforme MICA ou EMBRC France ou autres services ou appareils producteurs d'images disposeront d'un groupe temporaire ayant une dénomination rattachée au laboratoire d'où sont issues les images. Ce groupe sera créé temporairement pour une durée d'1 mois et une volumétrie maximale de 100Go. Au delà, si l'utilisateur souhaite utiliser la base de donnée sur une période plus grande une tarification sera appliquée sur a base du temps d'utilisation et pour une volumétrie adaptée.