

Déposer ses données sur Recherche Data Gouv

CÉCILE ARÈNES
15 FÉVRIER 2023



BIBLIOTHÈQUE
UNIVERSITAIRE



1

LE SITE RECHERCHE DATA GOUV



Un écosystème au service du partage et de l'ouverture des données de recherche

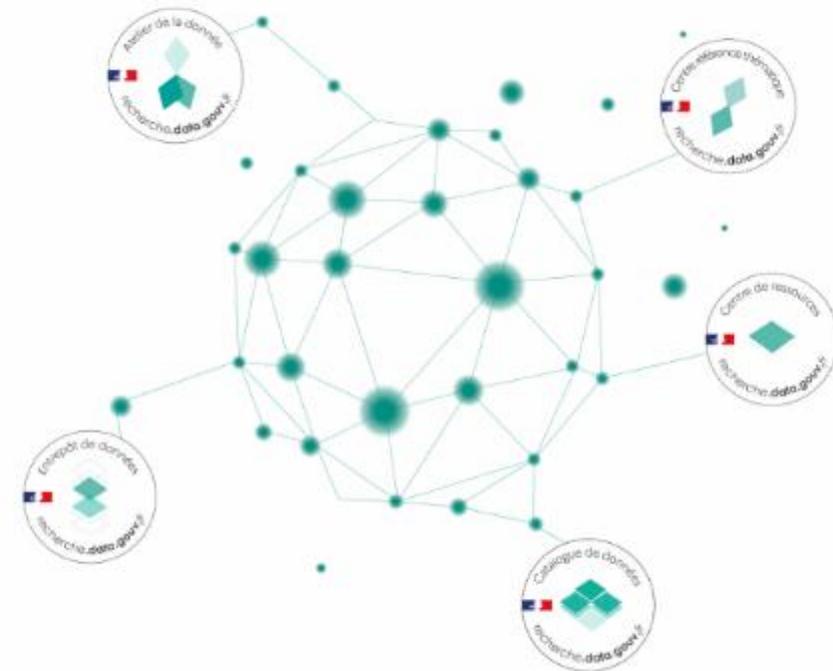
Plusieurs modules :

- [Des ateliers de la donnée](#)
- [Des centres de référence thématiques](#)
- [Des centres de ressources](#)

Deux modules pour déposer, publier et signaler des données :

- [Un entrepôt pour déposer et utiliser des données](#)
- Un catalogue pour rechercher les données publiées sur l'entrepôt ou sur des entrepôts externes

<https://recherche.data.gouv.fr/fr/page/a-propos-de-recherche-data-gouv>



Les ateliers de la donnée

Des services généralistes, de proximité :

- 5 ateliers labellisés à ce jour,
- 8 ateliers en cours de labellisation

Sorbonne Université n'est pas encore labellisé



Les centres de référence thématiques

Des centres d'expertise par domaine scientifique :

- Rôle dévolu aux infrastructures de recherche

Astronomie et astrophysique



Système Terre et Environnement



DATA
TERRA



Pôle National
de Données de Biodiversité

Sciences Humaines et Sociales



Huma-Num^{IR*}

Biologie-Santé

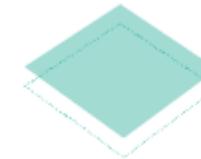


INSTITUT FRANÇAIS DE BIOINFORMATIQUE

Les centres de ressources

4 centres :

- Centre de ressources entrepôt-catalogue Recherche Data Gouv
- Centre de ressources DoRANum Recherche Data Gouv
- Centre de ressources compétences Recherche Data Gouv
- Centre de ressources OPIDoR Recherche Data Gouv



Centre de ressources
entrepôt-catalogue
Recherche Data Gouv



Centre de ressources
supports pédagogiques, e-formations
Recherche Data Gouv



Centre de ressources
compétences
Recherche Data Gouv



Centre de ressources
outils communs aux ateliers
Recherche Data Gouv

2

L'ENTREPÔT

Un entrepôt généraliste et pluri-disciplinaire

Recherche Data Gouv :

- Est dédié aux données associées aux publications
- Est destiné aux disciplines n'ayant pas encore d'entrepôts disciplinaires
- Est destiné aux projets pluri-disciplinaires
- N'a pas vocation à se substituer au paysage des entrepôts disciplinaires existants

Organisation de l'entrepôt

Un espace générique

<https://entrepot.recherche.data.gouv.fr/dataverse/espacegenerique>



Des espaces d'établissement

- Espace Sorbonne Université : <https://entrepot.recherche.data.gouv.fr/dataverse/sorbonne-univ>
- Possibilité de demander la création d'une collection pour une unité ou un projet



Organisation de l'entrepôt



BIBLIOTHÈQUE
UNIVERSITAIRE

Sur le site =>



Recherche Data Gouv (Recherche Data Gouv) Génération datapaper

Statistiques

341 923 téléchargements

Contact Partager

L'entrepôt pluridisciplinaire *Recherche Data Gouv* est une solution souveraine pour le partage et l'ouverture des données de recherche produites par les communautés qui ne disposent pas d'un entrepôt disciplinaire reconnu. Il est basé sur le logiciel libre *Dataverse*. Le dépôt des données doit se faire dans l'espace institutionnel attribué à l'établissement dont relève un des contributeurs. Un espace générique est dédié aux données produites par les établissements ne disposant pas d'espace dédié.

Retrouver les actualités et événements de la plateforme *Recherche Data Gouv*.

The *Recherche Data Gouv* multidisciplinary repository is a sovereign solution for sharing and opening up data produced by communities that do not have a recognized disciplinary repository. It is based on the *Dataverse* software. Data should be deposited in the space assigned to an institution which one of the contributor belongs to. A generic space is available for data produced by institutions which do not such a space yet themselves.

The latest news and events on the *Recherche Data Gouv* platform.



Chercher dans ce dataverse...

Chercher Recherche avancée

Dataverses (389)

Datasets (1 574)

Fichiers (22 606)

Catégorie du dataverse

Projet de recherche (143)

1 à 10 de 1 963 résultats

Trier

Lexique (Université de Lille)
26 août 2022 Université de Lille



Espace de dépôt des jeux de données liés aux articles publiés dans la revue Lexique : <https://lexique.univ-lille.fr/>

Principes FAIR, Findable



F1 : Les données et les métadonnées sont identifiées par un identifiant global unique et pérenne.

=> Fourni par l'entrepôt

F2 : Les métadonnées décrivant les données sont riches.

=> **Déposant**

F3 : Les données et les métadonnées sont enregistrées et indexées dans un dispositif permettant de les rechercher.

=> Fourni par l'entrepôt

F4 : Les métadonnées spécifient l'identifiant de la donnée.

=> Fourni par l'entrepôt

Principes FAIR, Accessible



A1 : Les données et les métadonnées sont accessibles par leur identifiant via un protocole de communication standardisé.

A1.1 : Le protocole utilisé est ouvert, libre et peut être implémenté de manière universelle.

A1.2 : Le protocole utilisé permet l'authentification et l'autorisation si besoin.

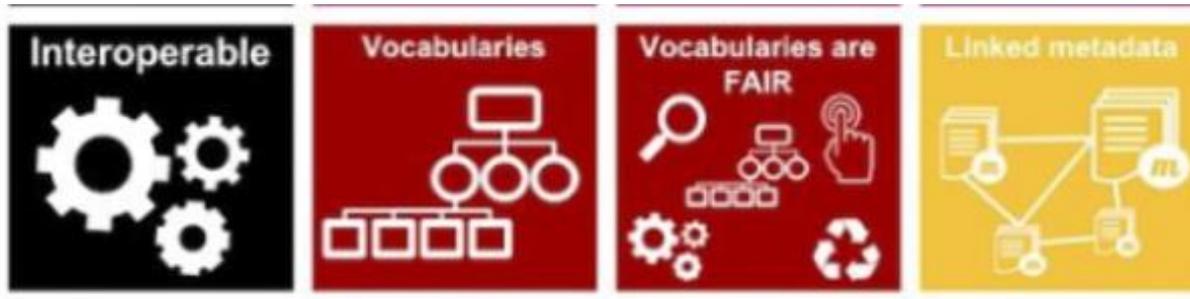
=> Fourni par l'entrepôt

A2 : Les métadonnées sont accessibles même quand les données ne le sont plus.

=> Fourni par l'entrepôt

Dépositant : si possible, données ouvertes, éventuellement sous conditions

Principes FAIR, Interopérable



I1 : Les données et les métadonnées utilisent un langage formel, accessible, partagé et largement applicable pour la représentation des connaissances.

=> [Déposant](#)

I2 : Les données et les métadonnées utilisent des vocabulaires qui respectent les principes FAIR.

=> [Déposant](#)

I3 : Les données et les métadonnées incluent des liens vers d'autres (méta)données.

=> Fourni par l'entrepôt

Principes FAIR, Réutilisable



R1 : Les données et les métadonnées ont des attributs multiples et pertinents.

R1.1 : Les données et les métadonnées sont mises à disposition selon une licence explicite et accessible.

=> [Déposant](#)

R1.2 : Les données et les métadonnées sont associées à leur provenance.

=> [Déposant](#)

R1.3 : Les données et les métadonnées correspondent aux standards des communautés indiquées.

=> [Déposant](#)

Les différentes briques de l'entrepôt

	Mitobese (INRAE) 14 nov. 2022 NuMeCan		Collection
	Widespread latitudinal asymmetry in marginal population performance 11 nov. 2022 - Data INRAE Pulido, Fernando; Castagneyrol, Bastien; Rodríguez-Sánchez, Francisco; Cáceres, Yónatan; Pardo, Adhar Johannes; Valladares, Fernando; Ehrlén, Johan; Jump, Alistair S.; Svenning, Jens-Christian; Hampe, Arnd latitudinal asymmetry in marginal population performance", https://doi.org/10.15454/STOJ93 , Recherche Data Gov, v.1, UNF:6:Xjcq1ZmSbtsZVAXeipy6Qw== [fileUNF]		Jeu de données
	sampleSpecies-rawdata.txt 14 nov. 2022 - Evolutionary analyses and expression patterns of TCP family genes in Ranunculales text/plain - 1.6 Ko - MD5: ebae04967adadb12dda489f83a30de55 name of the raw sequencing files associated to each assembly		Fichier

Qu'est-ce qu'un jeu de données ?

« Agrégation, sous une forme lisible, de données brutes ou dérivées présentant une certaine "unité", rassemblées pour former un ensemble cohérent. »

Gaillard R.- 2014. [De l'Open data à l'Open research data : quelle\(s\) politique\(s\) pour les données de recherche ?](#)

- **Ensemble cohérent, composé d'un ou plusieurs fichiers**
- **Décrit par des métadonnées**
- **Identifié par un DOI**
- **Comportant une licence (si possible ouverte), pour indiquer les possibilités de réutilisation**

Les métadonnées

Espace Générique

Recherche Data Gov > Espace Générique >

Fast and interactive positioning of proteins within membranes

Version 1.4



LANREZAC, André; LAURENT, Benoist; SANTUZ, Hubert; FÉREY, Nicolas; BAA DEN, Marc. 2022. "Fast and interactive positioning of proteins within membranes". <https://doi.org/10.57745/NSHWZ>, Recherche Data Gov, V1

Citer le dataset > Pour en apprendre davantage sur le sujet, consultez le document Data Citation Standards [en]

Modèles d'accès au dataset >

Contact Partager

Statistiques d'utilisation sur les datasets

47 consultations

3 téléchargements

0 citation

Description

This dataset contains the input, configuration and output files described in Lanrezac et al 2022 (in press). It also contains binaries of UnityMol, MDDriver and BioSpring, which are the programs used to in this interactive simulations. English (2022-09-30)

Subject

Chemistry, Medicine, Health and Life Sciences, Computer and Information Science, Physics

Mot-clé

Molecular dynamics, Transmembrane protein visualisation, Transmembrane protein analysis, Software engineering

Related Publication

Lanrezac et al 2022 (in press)

Fichiers Métadonnées Conditions Versions

1 fichier



impata_algorithms.zip
application/zip - 204.5 Mo - 7 oct. 2022 - 3 téléchargements
MD5: 3f6634657148bb1713c84cd6a6512df



Fichiers Métadonnées Conditions Versions

Exporter les métadonnées >

Citation Metadata

Identifiant pérenne du dataset

doi:10.57745/NSHWZ

Date de publication

2022-10-07

Title

Fast and interactive positioning of proteins within membranes

Contact

Utiliser le bouton de courriel ci-dessus pour joindre la personne-contact.

BAA DEN, Marc (CNRS)

Author

LANREZAC, André (UPC, CNRS, LBT, France) - ORCID: 0000-0003-2069-4161
LAURENT, Benoist (UPC, CNRS, LBT, France) - ORCID: 0000-0002-1714-2236
SANTUZ, Hubert (UPC, CNRS, LBT, France) - ORCID: 0000-0001-6149-9480
FÉREY, Nicolas (UPS, LIMSI, France) - ORCID: 0000-0002-7713-001X
BAA DEN, Marc (UPC, CNRS, LBT, France) - ORCID: 0000-0001-6472-0466

Producer

Centre National de la Recherche Scientifique (CNRS) <https://www.cnrs.fr/>

Production Date

2022

Distributor

Entrepôt-Catalogue Recherche Data Gov

Description

This dataset contains the input, configuration and output files described in Lanrezac et al 2022 (in press). It also contains binaries of UnityMol, MDDriver and BioSpring, which are the programs used to in this interactive simulations. English (2022-09-30)

Subject

Chemistry; Medicine, Health and Life Sciences; Computer and Information Science; Physics

Keyword

Molecular dynamics http://edamontology.org/operation_2476 (EDAM) <http://edamontology.org>
Transmembrane protein visualisation http://edamontology.org/operation_2241 (EDAM) <http://edamontology.org>
Transmembrane protein analysis http://edamontology.org/operation_0270 (EDAM) <http://edamontology.org>
Software engineering http://edamontology.org/topic_3372 (EDAM) <http://edamontology.org>

Kind of Data

Dataset; Model; Software

Data Origin

analysis data; simulation data; computer code

Software

UnityMol Version: 1.1.4
BioSpring Version: 111ed953

Related Publication

Lanrezac et al 2022 (in press) doi:

Grant Information

French State Initiative d'Excellence
Agence Nationale de la Recherche (ANR): "DYNAMO", ANR-11-LABX-0011
Agence Nationale de la Recherche (ANR): "CACSICE", ANR-11-EQPX-0008

Depositor

LAURENT, Benoist

Deposit Date

2022-09-30

Licences

Recherche Data Gov > Data INRAE > NuméCan > Mitobese >

Mitobese_COLON

Version 1.0



ROUDRY, Gaëlle; GUERRETTE, Thomas; IAN, Annali. 2022. "Mitobese_COLON", <https://doi.org/10.57745/CLIXA.GJ>, Recherche Data Gov, V1

Citer le dataset - Pour en apprendre davantage sur le sujet, consulter le document Data Citation Standards [en]

Modalités d'accès au dataset -

Contact Partager

Statistiques d'utilisation sur les datasets

?

0 consultation ?

0 téléchargement ?

0 citation ?

Description ?

16S sequence fasta file of the colonic microbiota data and associated metadata of the study by Guerrette et al where mice were fed two different obesogenic diets and control diets

Subject ?

Medicine, Health and Life Sciences

Fichiers

Métadonnées

Conditions

Versions

Conditions d'utilisation ^

Licence accordée ?

Les normes de la communauté Dataverse de même que les bonnes pratiques scientifiques exigent que toute source utilisée soit citée correctement. Veuillez utiliser la référence bibliographique ci-dessus générée par Dataverse.

Aucune licence n'a été sélectionnée pour ce dataset.

Conditions d'utilisation ?



LICENCE OUVERTE
OPEN LICENCE

Licence Ouverte / Open Licence Version 2.0 compatible CC BY

Recherche Data Gov > Data INRAE > Omics Dataverse > SPOMics >

Evolutionary analyses and expression patterns of TCP family genes in Ranunculales

Version 1.0



DAMERVAL Catherine; DEVEAUX Yves; CARMINE Claudot; LE GUILLOUX Martine; BRUNAUD Véronique; SOUBIGOU-TACONNAT Ludwine; CAIUS José; DELANNOY Etienne; NADOT Sophie; JABBOUR Florian; CONDE E SILVA Natalia, 2022. "Evolutionary analyses and expression patterns of TCP family genes in Ranunculales", <https://doi.org/10.57745/2G1VCP>, Recherche Data Gov, V1

Citer le dataset - Pour en apprendre davantage sur le sujet, consulter le document Data Citation Standards [en]

Modalités d'accès au dataset -

Contact Partager

Statistiques d'utilisation sur les datasets

?

0 consultation ?

0 téléchargement ?

0 citation ?

Description ?

this dataset contains the transcript and protein assemblies for *Aconitum napellus*, *Clematis stans*, *Ficaria verna*, *Heliborus orientalis* and *Nigella damascena* inferred from de novo assembly of Illumina RNA-seq reads. The rawsequencing data are available from CATdb under the accession ngs2015_16_ranunculaceae English

Subject ?

Earth and Environmental Sciences

Mot-clé ?

transcriptome, proteome

Fichiers

Métadonnées

Conditions

Versions

Conditions d'utilisation ^

Licence accordée ?

Les normes de la communauté Dataverse de même que les bonnes pratiques scientifiques exigent que toute source utilisée soit citée correctement. Veuillez utiliser la référence bibliographique ci-dessus générée par Dataverse.

Licence CC0 — « Transfert dans le domaine public »



Les versions

Protective ability of oligosaccharides to preserve lactic acid bacteria (LAB) during freezing, freeze-drying and storage

Version 1.1



Cenard, Stéphanie; Passot, Stéphanie; Ghibaudo, Florencia; Girardeau, Amélie; Romano, Nelson; Gomez-Zavaglia, Andrea; Fonseca, Fernanda, 2020, "Protective ability of oligosaccharides to preserve lactic acid bacteria (LAB) during freezing, freeze-drying and storage", <https://doi.org/10.15454/DOLRJO>, Recherche Data Gouv, V1

Citer le dataset ▾ Pour en apprendre davantage sur le sujet, consulter le document Data Citation Standards [en].

Modalités d'accès au dataset ▾

Contact

Partager

Statistiques d'utilisation sur les datasets



144 consultations

35 téléchargements

0 citation

Description

This study concerns the evaluation of the protective ability of oligosaccharides (FOS produced by enzymatic synthesis from sucrose) to preserve three lactic acid bacteria (*Lactobacillus delbrueckii* ssp *bulgaricus* CFL1, *Lactobacillus plantarum* WCFS1, and *Carnobacterium maltaromaticum* CNCM I-3298) during freezing, freeze-drying at pilot scale, and two weeks of storage of freeze-dried samples at 25°C. The dataset includes, for each of the three strains: i) the fermentation conditions (temperature, pH, and base consumption); ii) the stabilization conditions (freezing temperature and containers, freeze-drying recipe and parameters monitored during the cycle); iii) the quantification of the acidifying activity (using the CinAc system, AMS, Frepillon, France) and viability (using the plate count method) before and after freezing, after freeze-drying, and after storage.

Subject

Medicine, Health and Life Sciences; Agricultural Sciences

Mot-clé

Lactic acid bacteria, oligosaccharides, protective effect, freezing, freeze-drying

Fichiers

Métadonnées

Conditions

Versions

Dataset	Résumé	Collaborateurs	Publié
1.1	Métadonnées bibliographiques : Project Information (Modifié); Common Metadata : (1 Modifié); Voir les renseignements	Recherche Data Gouv Admin	9 févr. 2021
1.0	Il s'agit de la première version publiée.	Fernanda Fonseca	10 déc. 2020

La modération des jeux de données déposés

Sur l'espace générique de Recherche Data Gouv :

- Modération a priori (à anticiper)

Sur les espaces d'établissement, choix de l'institution :

- Modération a priori (SU) ou a posteriori

Générer un data paper



BIBLIOTHÈQUE
UNIVERSITAIRE

Sur RDG : <https://entrepot.recherche.data.gouv.fr/>

Par exemple : 10.57745/T5YEMM

Déposer et publier les jeux de données

Connexion avec son compte institutionnel



BIBLIOTHÈQUE
UNIVERSITAIRE

Recherche Data Gouv (Recherche Data Gouv) Génération datapaper

243 257 téléchargements

L'entrepôt pluridisciplinaire Recherche Data Gouv est une solution souveraine pour le partage et l'ouverture des données de recherche produites par les établissements qui ne disposent pas d'un entrepôt disciplinaire reconnu. Il est basé sur le logiciel libre Dataverse. Le dépôt des données doit se faire dans l'espace institutionnel attribué à l'établissement dont relève un des contributeurs. Un espace générique est dédié aux données de bac à sable.

The Recherche Data Gouv multidisciplinary repository is a sovereign solution for the sharing and opening of research data produced by institutions which do not have a disciplinary repository. It is based on the Dataverse software. Data should be deposited available for data produced by institutions which do not have a disciplinary repository. The latest news and events on the Recherche Data Gouv platform.

Se connecter

Vous pouvez vous connecter avec votre compte institutionnel afin de déposer vos jeux de données

SORBONNE UNIVERSITÉ

Bienvenue sur le portail d'authentification

Votre identifiant

Votre mot de passe

Se connecter

Se connecter

Connectez-vous ou inscrivez-vous avec votre compte institutionnel — En apprendre davantage. Vous quittez votre établissement ? Contactez le Support Recherche Data Gouv pour toute demande de support.

Votre établissement: Sorbonne Université Continuer

Autres options

Nom d'utilisateur/Courriel

S'inscrire pour obtenir un compte

- Paris Sorbonne Université (PELICAN)
- Pole de Vinci
- PSL - Université de Recherche Paris Sciences & Lettres
- PSL HEC - SAIT Lyon Saint-Herme
- Rectorat de la Martinique
- Rectorat de Strasbourg
- REMIAT Access Check
- Réseau Campus
- Sciences Po Bordeaux
- Sciences Po Grenoble
- Sciences Po Lille
- Sciences Po Lyon
- Sciences Po Paris
- Sciences Po Rennes
- SIRIM
- SIGMA Clermont
- SIRMA Business School
- Sorbonne Université

Note : Il est également possible créer un compte externe en cliquant le bouton [S'inscrire] mais ce compte n'aura pas de droit associé

Déposer et publier les jeux de données

Connexion avec son identifiant ORCID



Si vous possédez déjà un identifiant ORCID, vous pouvez vous connecter avec votre ORCID ID

Se connecter

Connectez-vous ou inscrivez-vous avec votre compte ORCID — [En apprendre davantage](#). Vous éprouvez des problèmes ? Veuillez contacter [Recherche Data Gouv](#) pour obtenir de l'aide.

ORCID est une initiative communautaire, ouverte et sans but lucratif visant à fournir un registre d'identifiants uniques de chercheurs ainsi qu'une méthode transparente pour lier les activités et produits de la recherche à ces identifiants. ORCID est unique dans sa capacité à atteindre les intervenants de toutes disciplines, secteurs de recherche et frontières nationales confondus ainsi que de coopérer avec d'autres systèmes d'identifiants. Pour en apprendre davantage visitez orcid.org/about.

Ce dépôt utilise votre ORCID ID pour l'authentification (vous n'avez donc pas besoin d'une autre combinaison nom d'utilisateur / mot de passe). Le fait que votre compte ORCID soit associé à vos datasets facilite également la recherche des datasets que vous avez publiés.

Se connecter

Connectez-vous ou inscrivez-vous avec votre compte institutionnel — [En apprendre davantage](#). Vous quittez votre établissement ? Contactez le [Support Recherche Data Gouv](#) pour toute demande de support.

Votre établissement

[Me permettre de saisir le nom de mon établissement](#)

Autres options

[S'inscrire pour obtenir un compte Dataverse.](#)

Projets en cours à SU



BIBLIOTHÈQUE
UNIVERSITAIRE

- Un projet de guichet unique autour des données de recherche porté par SACADO et la BSU
- Un espace institutionnel sur Recherche Data Gouv <https://entrepot.recherche.data.gouv.fr/dataverse/sorbonne-univ>

3

DÉMONSTRATION

Pour les exercices, utiliser le bac à sable de RDG



BIBLIOTHÈQUE
UNIVERSITAIRE

<https://demo.recherche.data.gouv.fr>

Source et remerciements

[Classe virtuelle Recherche Data Gouv](#) :

Déposer des données / Arnaud, Alexis (UGA); Cocard, Sylvie (INRAE); Jacquemot, Marie-Christine (Inist-CNRS) Sans, Damien (INRAE); Sowinski, Claire (Inist-CNRS); Szabo, Dimitri (INRAE); Tang, Cathy (INRAE), 2022.

Les supports sont envoyés à l'issue des classes virtuelles, il n'est pas nécessaire de s'inscrire. Prochaine session pour le dépôt de données : mardi 07 mars, de 10h à 12h

MERCI



BIBLIOTHÈQUE
UNIVERSITAIRE



Sauf mention contraire, cette présentation est mise
à disposition selon les termes de la Licence
Creative Commons Attribution 2.0 France.
Icônes : freepik