



Gestion des données au cours du projet de recherche

3 heures

Fred de Lamotte - Montpellier
<https://orcid.org/0000-0003-4234-1172>

Julien Seiler - Strasbourg
@julozi



Bonnes Pratiques

Un projet sur la durée (ré-intro)

La vie des données

Les principes FAIR

Stockage des données

Un environnement de travail sûr

Le nommage des fichiers

Les formats de fichiers

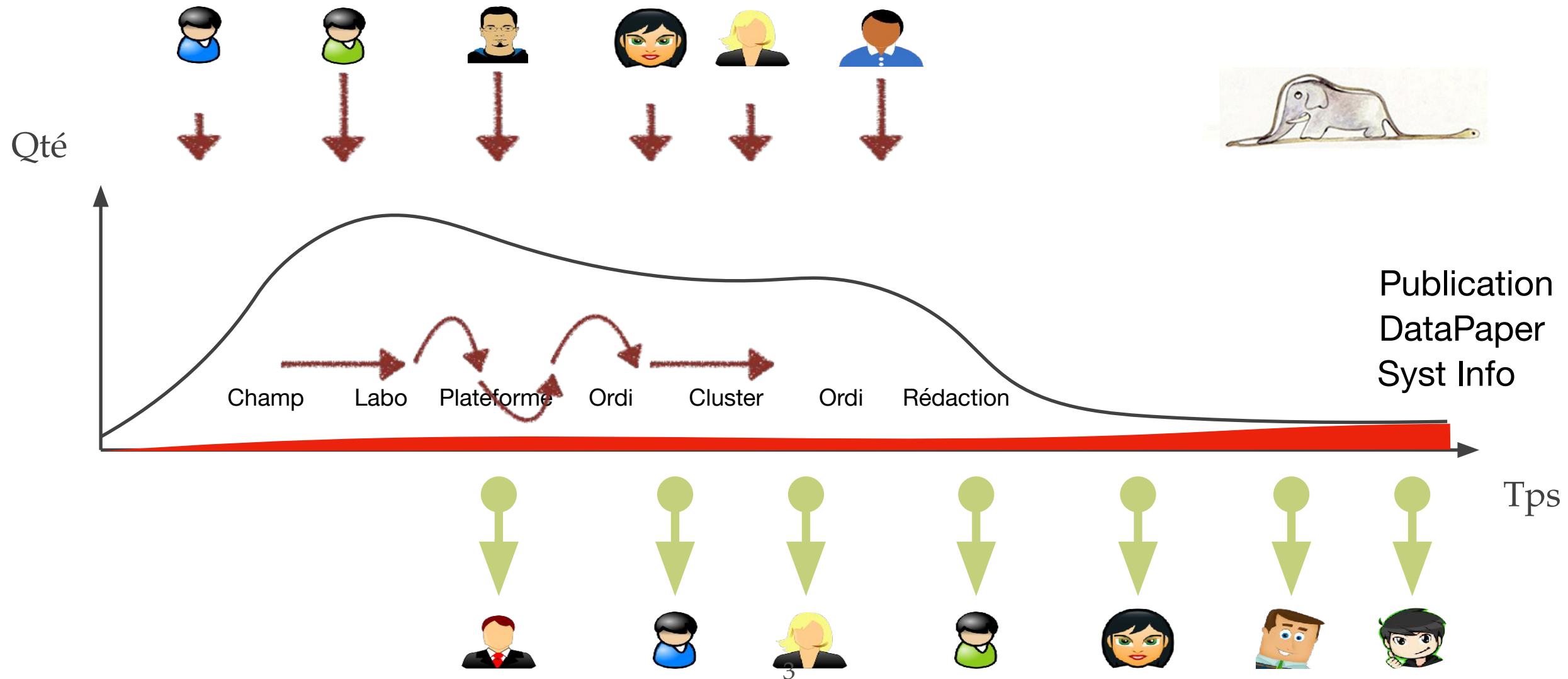
Organisation des données

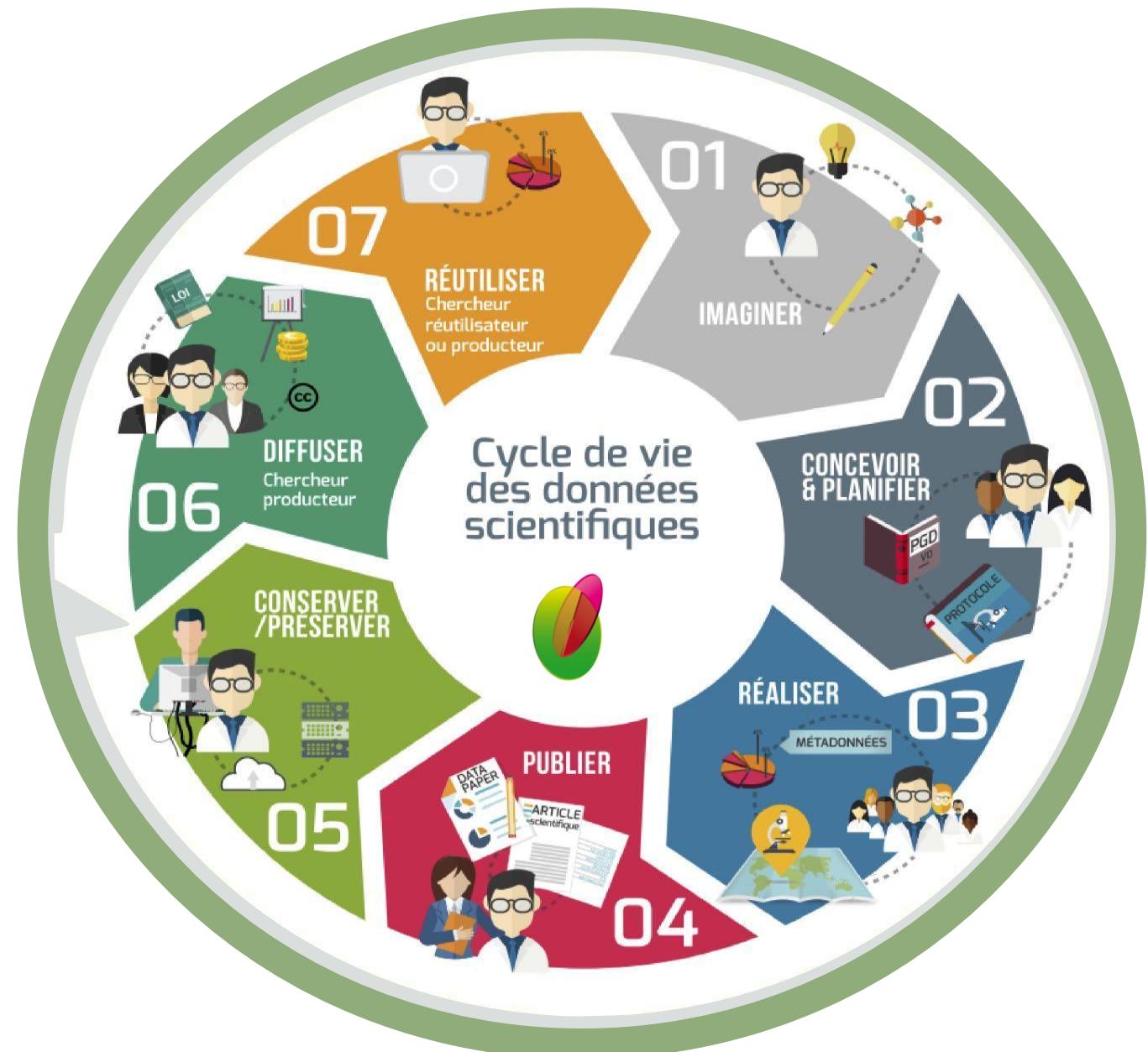
Protéger ses données

La suppression des données

Outils et solutions

Rappel : un projet sur la durée





Bonnes pratiques dans la gestion des données

Plusieurs personnes

Plusieurs techniques

Plusieurs lieux

Plusieurs années

Ne rien perdre

Pouvoir retrouver

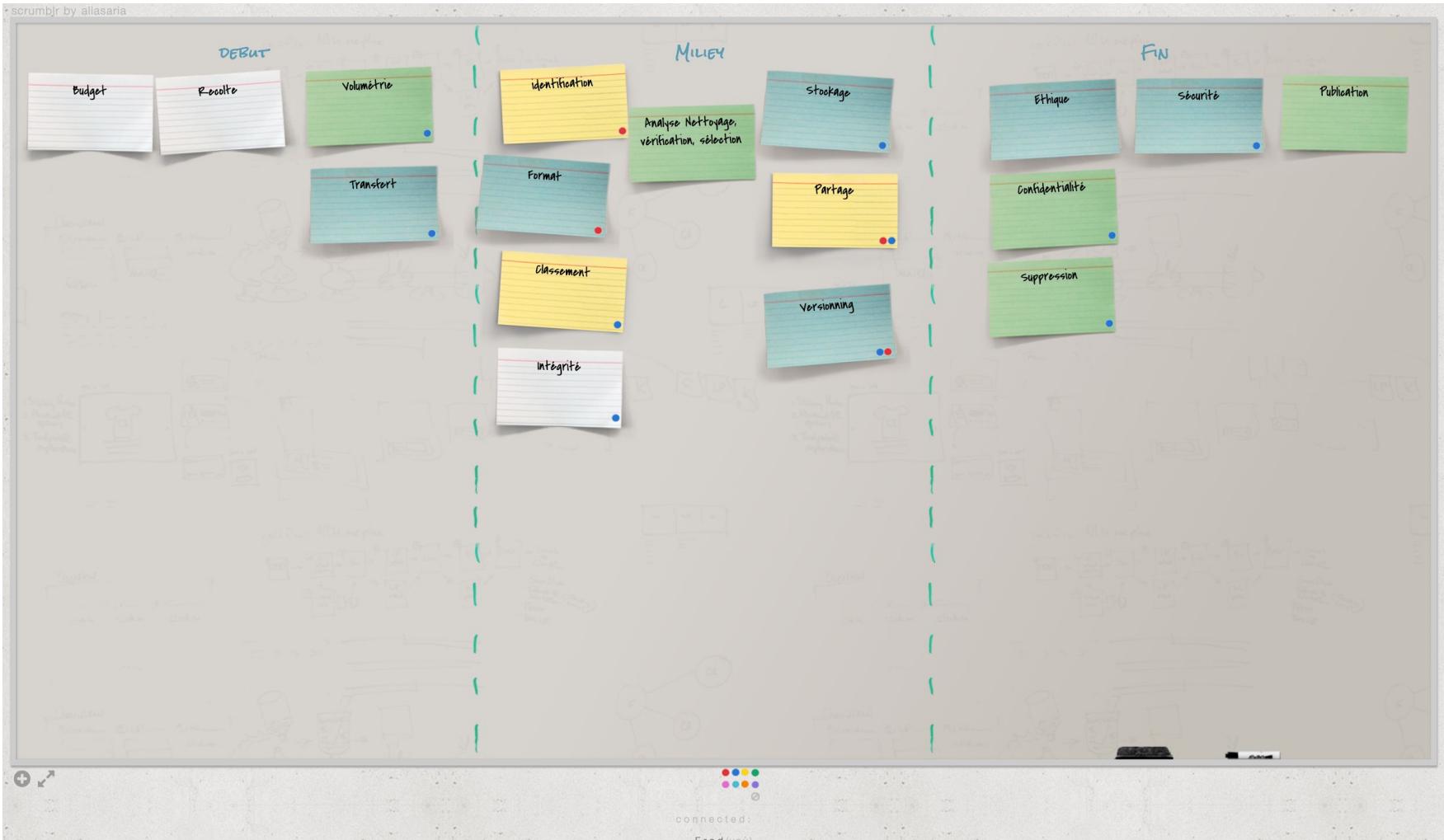
Pouvoir réanalyser

Pouvoir partager

La vie des données



La vie des données le long du projet



Un cas d'usage



Un cas d'usage

Tout au long de cette session nous allons vous présenter des bonnes pratiques et des outils en nous mettant en situation au travers d'un cas d'usage.

Pour ce faire, imaginons que nous sommes une équipe de recherche et que nous souhaitons démarrer un nouveau projet de recherche.

Ce projet, nécessitera de mener de nombreuses expérimentations et acquisition de données diverses.
Nous espérons également qu'il nous permettra de proposer quelques bons papiers.

Nous nous efforcerons également tout au long du projet de garder en tête les principes FAIR que nous souhaiterons notamment mettre en oeuvre au travers de la publication de données

Attention, cette formation contient du placement de produits :-)

Un cas d'usage

Contexte : les chercheurs/ingénieurs effectuent leur travail sur leur PC, mais avec obligation de sauvegarder sur un serveur.

Situation : ce matin, je m'aperçois que mon PC est inaccessible : visiblement le disque dur est mort.

À part acquérir un nouveau poste de travail, comment vais-je récupérer mon environnement logiciel & données ?

Un environnement de travail sûr



Un environnement de travail sûr

Comprendre l'environnement de travail que vous utilisez avant de démarrer votre projet :

Votre poste de travail :

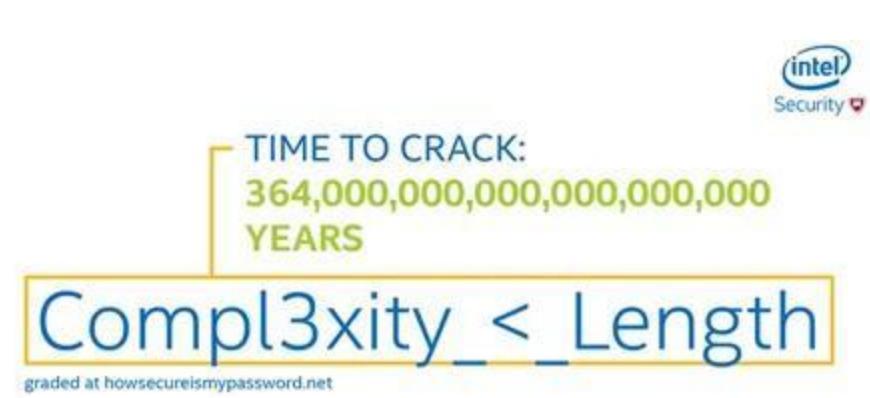
- Y'a-t-il des sauvegardes (stratégie 3-2-1) ?
 - **3** copies sur au moins **2** systèmes différents dont au moins **1** est distant = **0** inquiétude
Par exemple : stockage en RAID (copie locale) + sauvegarde sur un disque externe qui reste au labo
- Votre environnement est-il mis à jour régulièrement ?
- Disposez-vous d'un antivirus (à jour) ?
- Vos données sont-elles chiffrés (en cas de vol) ?

Vos solutions de stockage :

- Y'a-t-il des sauvegardes (stratégie 3-2-1) ?
- Est-ce que la pérennité est en phase avec vos besoins ?
- L'environnement est-il mis à jour régulièrement ?

Un environnement de travail sûr

Vos mots de passes (au pluriel)



Un environnement de travail sûr

Vos mots de passes (au pluriel)

- Utilisez-vous des mots de passe robustes ?

Type de mot de passe	Taille de clé équivalente	Force	Commentaire
Mot de passe de 8 caractères dans un alphabet de 70 symboles	49	Très faible	Taille usuelle
Mot de passe de 10 caractères dans un alphabet de 90 symboles	65	Faible	
Mot de passe de 12 caractères dans un alphabet de 90 symboles	78	Faible	Taille minimale recommandée par l'ANSSI pour des mots de passe ergonomiques ou utilisés de façon locale.
Mot de passe de 16 caractères dans un alphabet de 36 symboles	82	Moyen	Taille recommandée par l'ANSSI pour des mots de passe plus sûrs.
Mot de passe de 16 caractères dans un alphabet de 90 symboles	104	Fort	
Mot de passe de 20 caractères dans un alphabet de 90 symboles	130	Fort	Force équivalente à la plus petite taille de clé de l'algorithme de chiffrement standard AES (128 bits).

Exemple : N,cn'eplr.2IMcb! (16 caractères, alphabet de 90 symboles)

Un environnement de travail sûr

Vos mots de passes (au pluriel)

- Utilisez-vous un mot de passe différent pour chaque fournisseur de service ?
- Utilisez-vous un gestionnaire de mot de passe ?
 - BitWarden
- Renouvez-vous vos mots de passe régulièrement ?
- Utilisez-vous une procédure sécurisé pour communiquer un mot de passe à vos collègues ? (par exemple pastebin.com)

Optional Paste Settings

Syntax Highlighting:	None
Paste Expiration:	Burn after read
Paste Exposure:	Unlisted
Folder:	
Password	<input checked="" type="checkbox"/> Enabled iif5zL8zErFBehs6hfhjGr7djcbvhjre34v!
	<input checked="" type="checkbox"/> Burn after read NEW
Paste Name / Title:	The root password
Create New Paste	

Un gestionnaire de mot de passe : Bitwarden



Bitwarden est un service en ligne qui vous permet de créer un coffre fort dans lequel vous allez pouvoir enregistrer tous vos mots de passe.

OpenSource

et donc pérenne

Gratuit

*mais n'hésitez pas à payer la souscription
Premium pour soutenir le projet*

Accessible

*Application Mac, Windows, Linux, Web,
iPhone et Android*

1. Créer votre compte sur <https://bitwarden.com/>
2. Choisissez votre mot de passe maître (size matter)
3. Installer les applications sur vos appareils et les extensions de vos navigateurs
4. Enregistrer vos mots de passes dans votre coffre fort Bitwarden

En plus :

- Générateur de mot de passe robuste intégré
- Analyse de vos mots de passes et reporting
- Partage de mots de passe entre collègue

Notre espace de stockage



Un cas d'usage

Situation : après 7 mois d'attente, Sam Lee me transmet des données critiques par le biais d'un clé USB.

Q : quelle est ma démarche ?

Stockage des données

Fonction fondamentale : **la conservation des données**

Stockage :

- désigne des méthodes et des technologies permettant de conserver des données
- concerne tous les types de supports de stockage de masse (DD, Clé USB...) ou support de stockage dématérialisé (cloud)
- intègre des problématiques d'usage collaboratif : dépôt, partage.

Critères de sélection pour choisir un support de stockage :

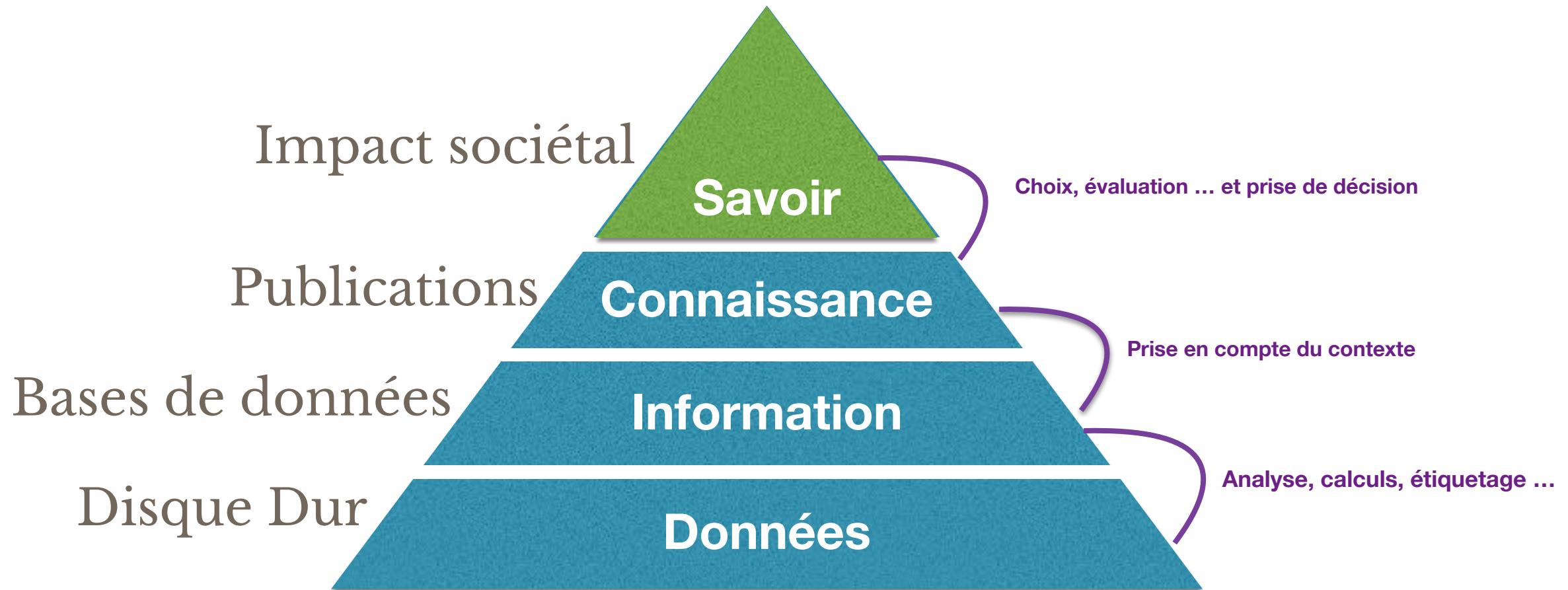
- la fréquence d'utilisation des données,
- les besoins en capacité de stockage (taille),
- la sécurité des données,
- la vitesse d'accès à la donnée
- la fiabilité et le coût du support

Stockage des données

Les besoins courants pour la gestion de données lors d'un projet de recherche

- Des espaces de stockages adaptés à vos données (données scientifiques, documents bureautiques, bases de données, code source)
- Des outils adaptés à la gestion des droits des collaborateurs
- Des solutions de publication et d'archivage des données

Data Pyramid



Stocker et sécuriser : quels compromis ?

Comparatif de systèmes de stockage des données

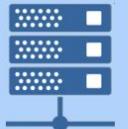
Support de stockage	Sécurité	Accès	Coût	Remarque d'utilisation
 Ordinateur professionnel	 Sujet au piratage informatique, aux détériorations et pannes	 Pas adapté au partage, nécessite l'utilisation d'un support externe ou d'Internet (mail, cloud...)	 Pas de coût supplémentaire ou coût peu important	- Pour un stockage temporaire - Nécessité de crypter les données confidentielles et sensibles
 Support externe	 - Sujet au vol, à la perte du support - Durée de vie limitée (dégradation du matériel)	 Facilement transportable, il permet de transférer les données vers un autre ordinateur	 Pas de coût supplémentaire ou coût peu important	- Pour un stockage temporaire - Nécessité de crypter ou de sécuriser physiquement les données confidentielles et sensibles
 Serveur institutionnel	 Stockage fiable, durable et sécurisé (contre le vol, le piratage, les incendies...)	 La connexion au serveur institutionnel ne facilite pas le travail avec des personnes extérieures	 Coût assez important mais pas forcément répercuté sur l'usager	- Pour un stockage plus pérenne - Adapté pour le stockage de données sensibles et des versions « stables » de vos données - Toutes les institutions ne proposent pas ce service
 Serveur Cloud	 On ne sait pas vraiment où sont stockées les données, ni ce qu'elles deviennent	 Permet un travail synchronisé avec toutes les personnes ayant été autorisées au partage	 Payant à partir d'une certaine limite de stockage	- Pour un partage avec des personnes externes à l'institution - Ne pas y mettre de données sensibles ou confidentielles - Pas de contrôle sur la procédure de sauvegarde des données

Tableau tiré de <http://doranum.fr/le-stockage-des-donnees/>

Comprenez l'infrastructure que vous utilisez

Performance vs Sécurité

- Une infrastructure de calcul nécessite une solution de stockage **performante** :
 - accès massivement parallèle aux données
 - disques rapides
- Pour gagner en performance, on désactive les mécanismes de sécurité :
 - Moins voire pas de snapshots
 - Pas de réPLICATION
 - Pas de sauvegarde
- Pour gagner en sécurité, on réduit la performance
- A capacité identique, le coût d'une infrastructure performante et d'une infrastructure sécurisé est le même



Comprenez l'infrastructure que vous utilisez

- Infrastructure de calcul ne rime pas toujours avec infrastructure de stockage

 Security  Search

IFB Core Cluster Documentation

Quick start guide

Logging in

Job submission (Slurm) >

Software environment >

Data >

Tutorials >

Cluster description

Security

Backup

There is no backup for the main storage.

Some snapshots are available to protect against deletion by error but only one by day and for days.

All servers and services are deployed using Ansible (and configurations are under revision control).

Main infrastructure services are backed up.

Charte d'utilisation ROMEO

Conditions d'accès et règles de bon usage des ressources ROMEO

Version 2017/12

Créé en 2002, le Centre de Calcul Régional ROMEO accompagne les chercheurs de la région dans leurs activités numériques. La description complète des ressources et de leur utilisation est décrite sur <http://romeo.univ-reims.fr>

La présente demande, d'ouverture ou de maintien de compte sera étudiée et validée par le comité scientifique du centre de calcul et mis en œuvre par le personnel ROMEO.

L'utilisateur s'engage, sous risque de fermeture de son compte sans préavis, à :

- consulter, corriger et améliorer les informations contenues sur le site pour toute question
- consulter les *notes de maintenance* sur le site web et sur les messages d'accueil des machines
- ne pas utiliser la machine comme espace de stockage ou de sauvegarde
- ne pas utiliser la machine comme passerelle depuis l'extérieur vers le réseau de l'URCA
- maintenir à jour ses coordonnées dans la rubrique *mon compte* du site web
- mettre à jour les projets dont il est responsable ou membre ainsi que la liste de ses publications dans la rubrique « *mon compte* » du site web
- mentionner l'utilisation de ROMEO sur vos communication :
 - *Ce travail a été réalisé avec le concours du Centre de Calcul Régional ROMEO*
 - *This work was partially supported by the French HPC Center ROMEO*
- prendre toute mesure afin d'empêcher l'utilisation de compte par des tiers (ne pas divulguer son mot de passe, choisir un mot de passe suffisamment complexe)
- participer aux événements organisés par le Centre de Calcul
- lire son mail régulièrement et répondre aux demandes venant du Centre de Calcul
- de manière générale, se conformer aux règles d'utilisations (batch, utilisation des scratchs, ...) disponibles dans la rubrique *techno-centre* du site web
- libérer les espaces scratchs après leur utilisation
- communiquer avec l'équipe technique à l'adresse romeo@univ-reims.fr
- utiliser le site de support pour toute demande d'intervention <https://romeo.univ-reims.fr/ticket>
- participer à la diffusion des résultats scientifiques (posters, vidéos, ...)
- respecter les aspects légaux liés aux logiciels
- ne pas utiliser les ressources du centre à des fins criminelles, de violation ou tentative

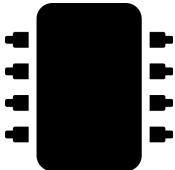
24

Le NNCR IFB

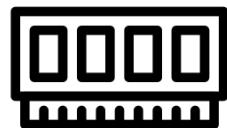
National Network of Computing Resources

Une offre de service **cloud** et **cluster** couvrant l'ensemble du territoire Français

Le cluster national IFB



4300 coeurs



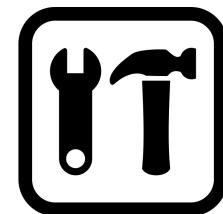
20 To RAM



2 Po



Une communauté
d'entraide



Plus de 400
outils



SSH
Jupyter
RStudio
Galaxy 25

Le NNCR IFB

Cluster	Localisation du Data center	Coeurs	RAM (Go)	Stockage (To)
IFB Core	IDRIS - Orsay	5 042	26 542	2 000
Genotoul	Toulouse	6 128	34 304	3 000
ABiMS	Roscoff	2 608	10 600	2 500
GenOuest	Rennes	1 824	7 500	2 300
Migale	Jouy en Josas	1 084	7 000	350
BiRD	Nantes	560	4 000	500

Le cluster national de l'IFB

Allocation des espaces de stockage par projet :

- 250 Go par projet extensible sur demande argumentée
- Un projet peut être accessible à plusieurs utilisateurs
- Un utilisateur peut demander plusieurs espaces projet
- Pas de sauvegarde

Bientôt disponible :

- Mise à disposition d'un espace scratch avec un quota plus important pour des besoins ponctuel (suppression automatique des fichiers les plus anciens)
- Sauvegarde des espaces projets

Demander un espace projet IFB cluster

<https://my.cluster.france-bioinformatique.fr>



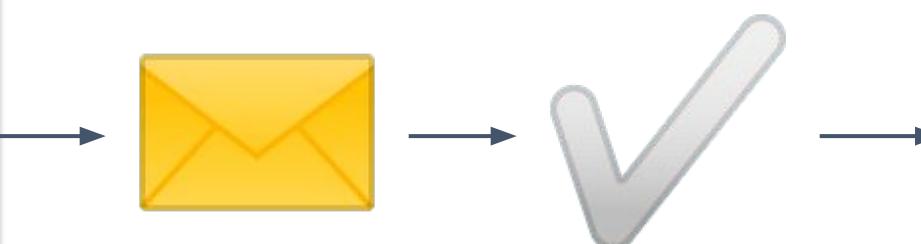
Sign in to access your account

User id
Password
Sign in

Lost your password? [RESET](#)

or

Create an account



Request a new project creation

Name **(required)**

Avoid generic name, team name, technology name or your name. Please, choose a project name that matches your cluster research project. If you treat several projects, it is quite possible for you to request more project spaces.

Size (GB)

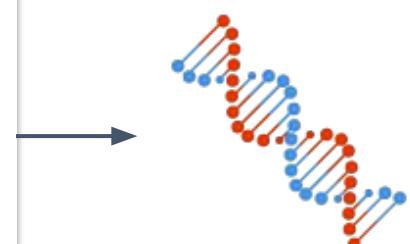
Optional, for information only

Financing

Optional, for information only

Description

Ask Admin **Cancel**



Créer un compte

Vérifier son adresse email

Attendre la validation de son compte

Demander un espace projet

Commencer à travailler

Accès à l'espace projet



FileZilla



Terminal



MobaXterm



SSH/SFTP

core.cluster.france-bioinformatique.fr

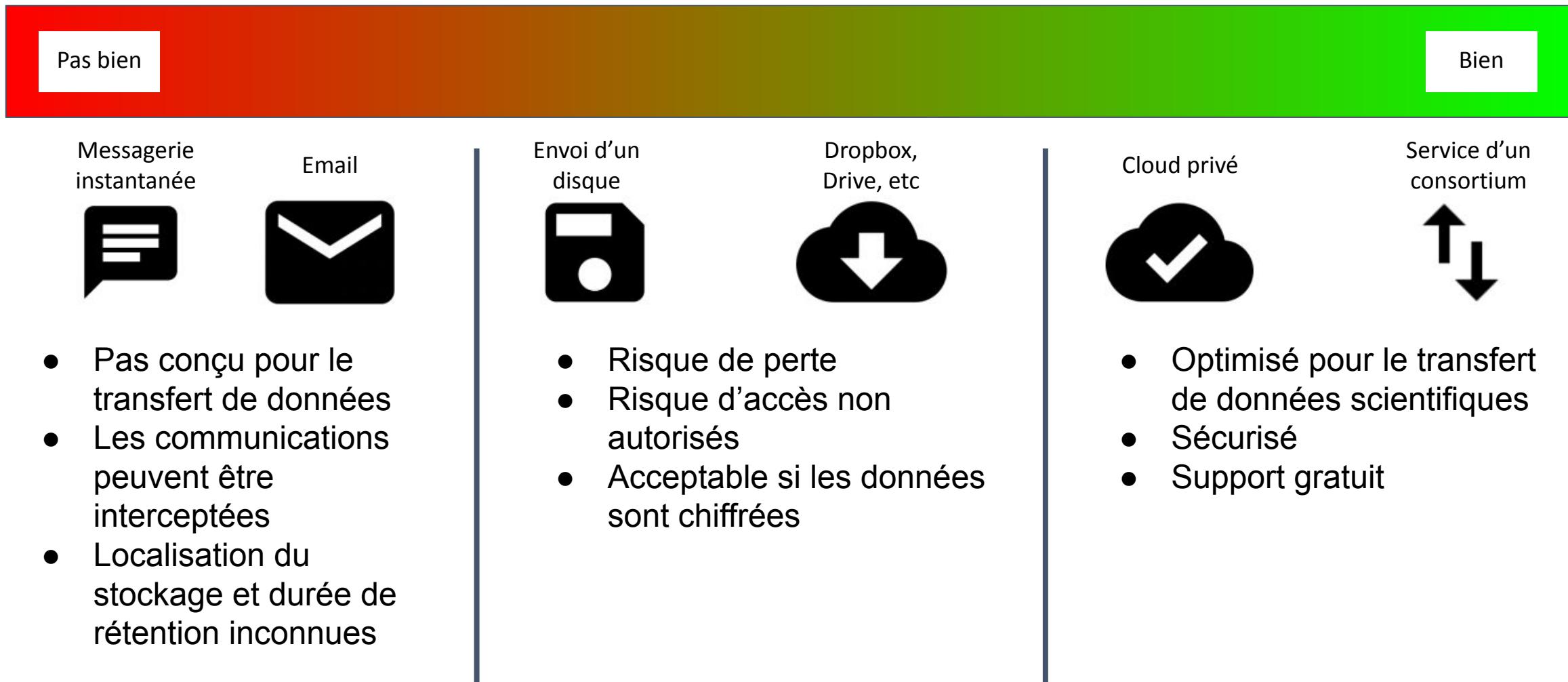


RStudio

.cluster.france-bioinformatique.fr

Transfert de vos données de recherche

Comment transmettre vos données ?



Transfert de vos données de recherche

Démonstration :

- A l'aide de son terminal : la copie via SSH avec scp
- A l'aide d'un client sFTP : FileZilla
- A l'aide de son navigateur : JupyterHub

La vitesse de transfert dépend de :

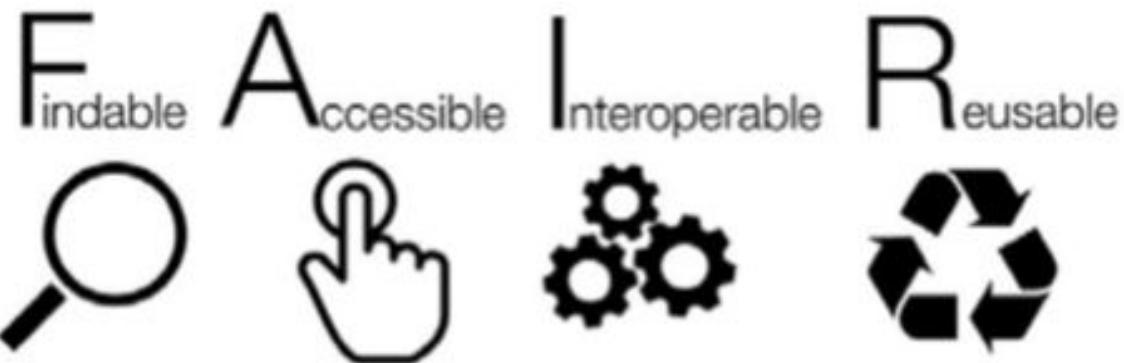
- L'outil utilisé
- L'infrastructure source et destination
- Le réseau
- La granularité des données

*N'oubliez pas le
chiffrement !!!*

Un cas d'usage

Situation : C'est parti pour un projet de biologie intégrative sur 3 ans, au programme acquisitions de nombreux types de données (imagerie, séquençage, phénotypage) et analyses intensives.

Q : Expliquez votre approche de nommage et d'organisation des fichiers (le nom des fichiers doit obligatoirement comprendre au moins la date)



'How FAIR are your data?' checklist, CC-BY by Sarah Jones & Marjan Grootveld, [EUDAT](#). Image CC-BY-SA by [SangyaPundir](#)

Findable

It should be possible for others to discover your data. Rich metadata should be available online in a searchable resource, and the data should be assigned a persistent identifier.

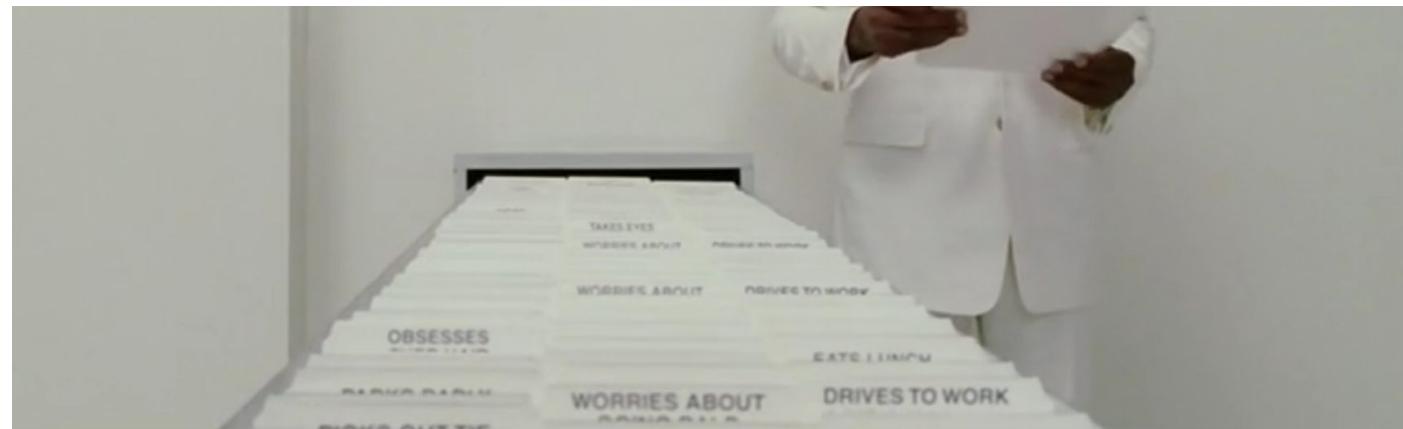
- A persistent identifier is assigned to your data
- There are rich metadata, describing your data
- The metadata are online in a searchable resource e.g. a catalogue or data repository
- The metadata record specifies the persistent identifier

Interoperable

Data and metadata should conform to recognised formats and standards to allow them to be combined and exchanged.

- Data is provided in commonly understood and preferably open formats
- The metadata provided follows relevant standards
- Controlled vocabularies, keywords, thesauri or ontologies are used where possible
- Qualified references and links are provided to other related data

Le nommage des fichiers



COMMENT

Open Access



CrossMark

Gene name errors are widespread in the scientific literature

Mark Ziemann¹, Yotam Eren^{1,2} and Assam El-Osta^{1,3*}

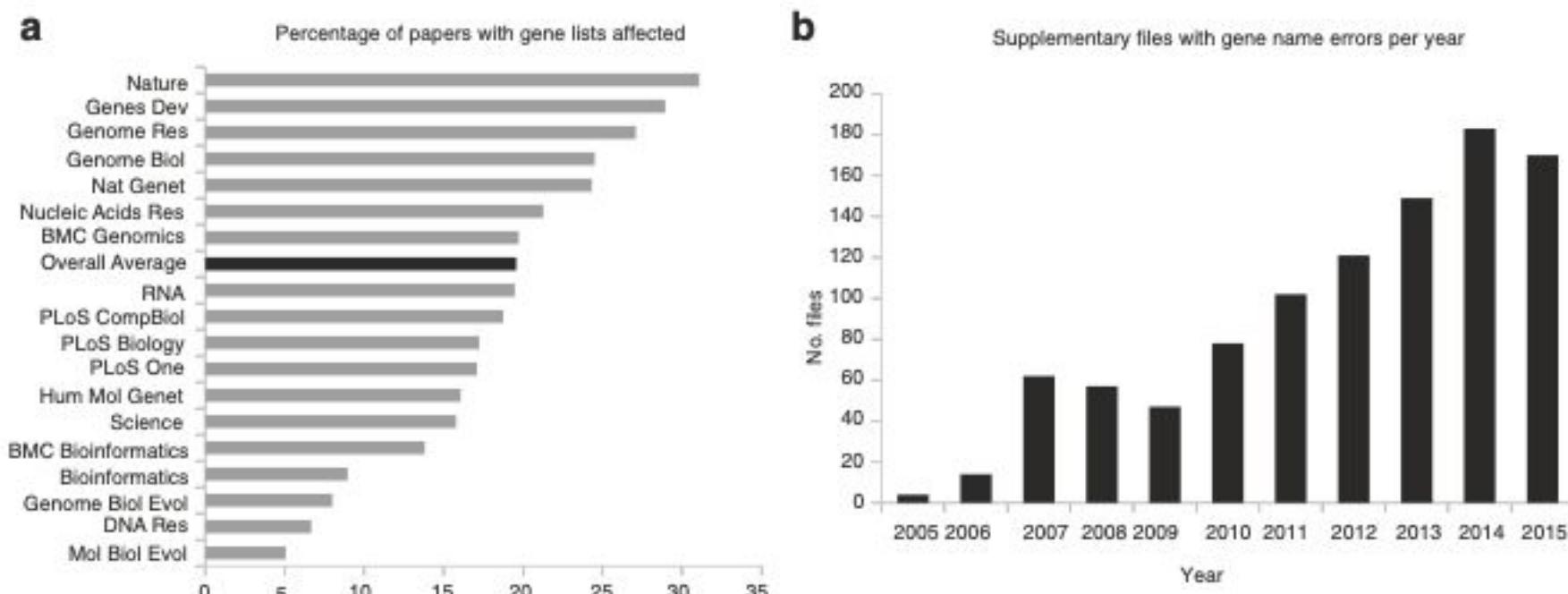


Fig. 1 Prevalence of gene name errors in supplementary Excel files. **a** Percentage of published papers with supplementary gene lists in Excel files affected by gene name errors. **b** Increase in gene name errors by year

1) DONNER UN NOM BREF ET EXPLICITE (*SI VOUS N'ÊTES PAS UNE MACHINE*)

Le nom d'un fichier doit être le plus court possible et doit refléter son contenu. Il peut contenir son **sujet** (rapport d'activité, résultats d'analyse, bilan, etc.), sa **date**, sa **version** (si le document est destiné à évoluer).

Ce qui signifie qu'il est inutile d'écrire des mots vides comme : "le", "un", "des", "et", etc. Ils n'apportent pas plus de sens pour comprendre de quoi le document retourne.

Pour réduire le nombre de caractères, utilisez des **abréviations** facilement reconnaissables (CR pour compte rendu par exemple).



Reunion



20150407_CR



20150407_Minutes



20150407_OJ

Dans le dossier réunion, on retrouve :

un document de compte rendu (CR)
un document sur les minutes de la réunion et
un document sur l'ordre du jour (OJ).

Le mot réunion n'apparaît pas dans la dénomination des fichiers car il est déjà précisé dans le nom du dossier.

2) NE PAS METTRE D'ESPACE NI DE CARACTÈRES SPÉCIAUX (*SURTOUT SI VOUS ÊTES UNE MACHINE*)

Il vaut mieux ne pas mettre d'espace entre les mots ni de caractères spéciaux (accents et cédille y compris). Pour séparer les mots, vous pouvez mettre une **majuscule** ou un **underscore “_”**.



Règles dénomination fichiers X



ReglesDenominationFichiers ✓

Cette règle permet d'éviter des problèmes d'affichage voire d'accès aux documents.

En effet, il se peut que d'un système d'exploitation à un autre (Windows vers Mac par exemple), le fichier ne soit pas reconnu.

Il vaut mieux donc bannir tout ce qui n'est pas alphanumérique : & / + > : ? % () etc.

3) INDICER LES DATES AU BON FORMAT (QUE VOUS SOYEZ OU NON UNE MACHINE)

Les dates doivent s'écrire au **format AAAAMMJJ** (année, mois, jour).

Commencer par l'année permet de **maintenir les documents dans un ordre chronologique.**



20150405_CR



20160310_CR



20160515_CR

Le document le plus récent se retrouvera toujours en fin de liste.

4) PLACER L'ÉLÉMENT IMPORTANT EN PREMIER

Il est préférable de bien choisir l'ordre des éléments dans le nom des fichiers. Mettez **en premier l'élément qui vous est le plus pertinent** pour pouvoir retrouver facilement votre document.



20150911_CR



20150911_Minutes



20151210_CR



20151210_Minutes

Ici par exemple, il sera plus facile de retrouver un compte rendu en fonction de sa date. L'élément le plus important du document est sa date, c'est donc elle qui apparaît en premier.

5) INDICER LES VERSIONS DES DOCUMENTS

Certains documents peuvent être amenés à évoluer, avoir plusieurs versions avant une version finale. Dans ces cas, il faut préciser la version dans le nom du fichier.



Pour faire simple, ne rédiger que la lettre V suivi du numéro de la version, en deux unités.

Si un numéro de version ne convient pas, vous pouvez le remplacer par une autre convention de nommage, par exemple VP pour Version Provisoire, VF pour Version Finale (si vous êtes joueur 😊).

1) DONNER UN NOM BREF ET EXPLICITE (SI VOUS N'ÊTES PAS UNE MACHINE)

MAIS ÇA SE DISCUTE

MES IDENTIFIANTS DOIVENT-ILS ÊTRE OPAQUES OU SIGNIFIANTS ?

On parle d'identifiant **signifiant** quand la chaîne de caractères qui constitue l'identifiant est formée à partir des métadonnées qui décrivent la ressource qu'on identifie.

Au contraire, un identifiant **opaque** est constitué d'une suite de lettres et de chiffres sans rapport avec le contenu de la ressource.

La question du choix entre identifiants signifiants ou opaques est liée à l'utilisation qui est faite de l'identifiant.

Si l'identifiant doit être lu ou manipulé par un utilisateur humain, on aura tendance à vouloir le rendre signifiant.

En revanche, si l'identifiant est principalement manipulé par des machines, la longueur de l'identifiant et son opacité importent peu.

<http://d.opencalais.com/genericHasher-1/c1008719-bca3-3c48-966a-192869743423.html>,

Ou

<http://dbpedia.org/resource/Manhattan>

Il faut être conscient qu'embarquer de la sémantique dans l'identifiant expose à être confronté aux questions d'ambiguïté de la langue. Il sera éventuellement nécessaire de faire porter par l'identifiant des éléments de désambiguïsation pour en garantir l'unicité.

Par exemple, il est nécessaire d'ajouter un élément de désambiguïsation dans l'identifiant pour distinguer l'objet « bidon » (<http://exemple.com/categorie/Bidon>) et la commune « Bidon » (<http://exemple.com/commune/Bidon>).

On voit le type d'ambiguïté que peuvent porter des termes comme « avocat » (homonymie) ou « opéra » (polysémie).

Par ailleurs, les possibles évolutions de la langue induisent un risque pour la pérennité.

Par exemple, le concept qui serait ainsi identifié : « <http://exemple.com/categorie/biodiesel> » devient d'une certaine façon obsolète, dès lors que le journal officiel du 22 juillet 2007 a préconisé l'utilisation de la forme française « biogazole » de préférence à « biodiesel ». Dans ce cas, ou bien l'identifiant du concept s'écarte du libellé du concept, ce qui réduit l'intérêt d'avoir un identifiant signifiant, ou bien on modifie l'identifiant du concept pour l'adapter à l'état de la langue, ce qui remet bien sûr en cause la pérennité de cet identifiant.

Recommandations

Critères	Niveaux d'exigence
Se conformer au plan de nommage de l'institution ou du domaine ou le cas échéant déterminer un plan de nommage spécifiant les éléments de la numérisation à doter d'un identifiant	Obligatoire
Garantir l'unicité des identifiants	Obligatoire
Veiller à la non-ré attribution des identifiants de la ressource concernée	Obligatoire
En cas de suppression ou de dépublication, en renseigner les circonstances et le cas échéant pointer vers une ressource de substitution	Recommandé
Publier sa politique de gestion des identifiants	Recommandé
Donner des identifiants non signifiants pour éviter toute ambiguïté liée aux langues naturelles	Recommandé
Dans le cas de l'utilisation d'identifiants pérennes, prévoir des URI HTTP [c'est-à-dire que les URI donnant directement accès à l'usager, via un navigateur, à une vue de la ressource].	Recommandé

Eléments	Règle	Exemple
Sujet	Obligatoire Il s'agit du sujet principal traité au sein du document. Utiliser des noms communs, écrits en lettres minuscules non accentuées.	projet formation évaluation
Séparateur	Les espaces sont interdits. Utiliser l'underscore (touche 8 du clavier) pour remplacer les espaces	« _ »
Type de document	Facultatif Qualifie la nature du document. Toute abréviation sera en lettres majuscules.	(CR) compte rendu (OJ) ordre du jour
Date	Obligatoire Date de création du document, date de l'événement. Format à l'américaine : AAAAMMMJJ. Nommage d'une période : utilisation d'un séparateur «_» ou « - ».	20180122 201608 2010 201501_07 ou 201501-07
Version du document	Obligatoire Distingue les différentes versions d'un document, signalées par un « V » majuscule suivi de deux chiffres ; version provisoire (VP) et la version finale (VF), version validée (VV). Un nouveau document créé à partir d'une version finale doit être sauvegardé sous un nouveau nom de manière à ne pas écraser la version précédente.	CR_CFU_V0.0 CR_CFU_V0.1 CR_CFU_VP, VF ou VV
Extension	Obligatoire L'extension est ajoutée automatiquement par le système et n'apparaît peut-être pas sur vos écrans.	.txt (fichier texte) .doc (fichier Word) .xls (fichier Excel)

Format de fichier



Un cas d'usage

Situation : vous devez traiter un fichier avec un format ‘propriétaire’, c'est à dire qui nécessite un logiciel non gratuit pour lire le fichier. Votre institution n'a aucune licence pour ce logiciel, et ne projette pas d'en acquérir.

Q : quelles sont les solutions possibles ?

Deux grandes catégories de formats : **textuels** et **binaires**.

Enjeu pour la préservation et l'exploitation des données

Formats « textuels »

- Suite d'octets représentant des caractères imprimables et affichables à l'écran
- Peuvent être lus dans un éditeur de texte
- Mais souvent besoin d'un logiciel spécifique pour interpréter la structure interne, matérialisée par certains caractères, et en donner une représentation informatique exploitable

Ex. de format textuel : HTML

Contenu lisible dans un éditeur texte :

```
<html>
<head></head>
<body>
<p>Bonjour <span style='color:red'>tout le monde</span></p></body>
</html>
```

Mais « interprétable » par un logiciel dédié (navigateur web) :



Bonjour tout le monde

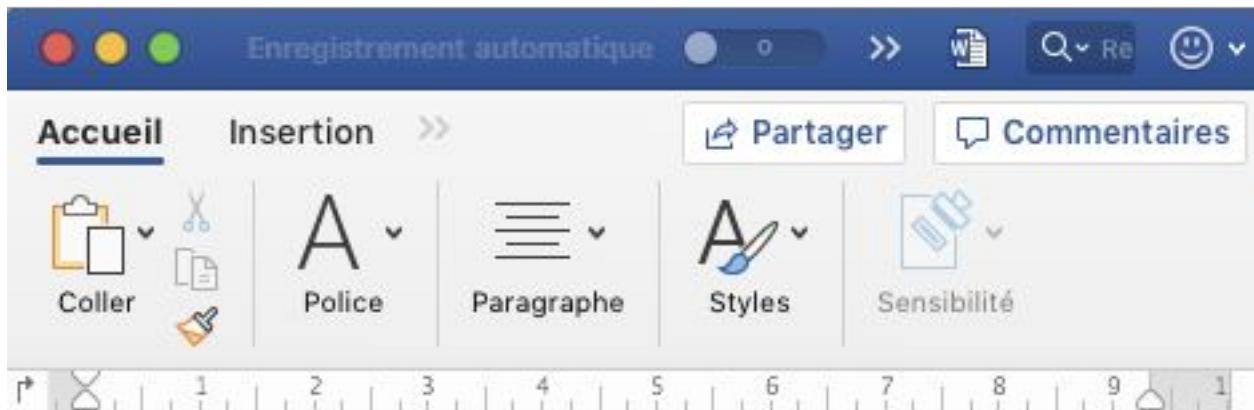
Caractères ordinaires + caractères ayant une valeur spéciales : < > /, etc.

Mots ayant des valeurs spéciales en HTML (« balises ») si encadrés par < > ou </>: <body>, </body>, etc...

Ex. de format textuel : RTF (texte structuré) Contenu lisible dans un éditeur texte

```
{\rtf1\adeflang1025\ansi\ansicpg1252\uc1\adef0\def0\stshfbch37\stshf  
och37\stshfhich37\stshfb0\deflang1036\deflangfe1036\themelang1036\theme  
langfe0\themelangcs0{\fonttbl{\f0\fbidi \froman\fcharset0\fprq2{\*\panos  
e 02020603050405020304}Times New Roman;}{\f34\fbidi \froman\fcharset0\fp  
rq2{\*\panose 02040503050406030204}Cambria Math;}  
  
\xmlMargin0\mrMargin0\mdefJc1\mwrapIndent1440\mintLim0\mnaryLim1}{\info{\  
author Mathieu Saby}{\operator Mathieu Saby}{\creatim\yr2018\mo6\dy10\hr  
13\min44}{\revtim\yr2018\mo6\dy10\hr13\min44}{\version2}{\edmins1}{\nofp  
ages1}{\nofwords3}{\nofchars19}  
  
\fs24\lang1036\langfe1033\loch\af37\hich\af37\dbch\af37\cgrid\langnp1036  
\langfnp1033 {\rtlch\fcs1 \af0 \ltrch\fcs0 \insrsid16651434 \hich\af37\  
dbch\af37\loch\f37 Bonjour }{\rtlch\fcs1 \af0 \ltrch\fcs0 \cf6\insrsid16  
651434\charrsid16651434  
  
\hich\af37\dbch\af37\loch\f37 tout le monde{\rtlch\fcs1 \af0 \ltrch\fcs  
0 \insrsid16651434
```

Mais uniquement interprétable avec Word, Libre office ou autre traitement de texte



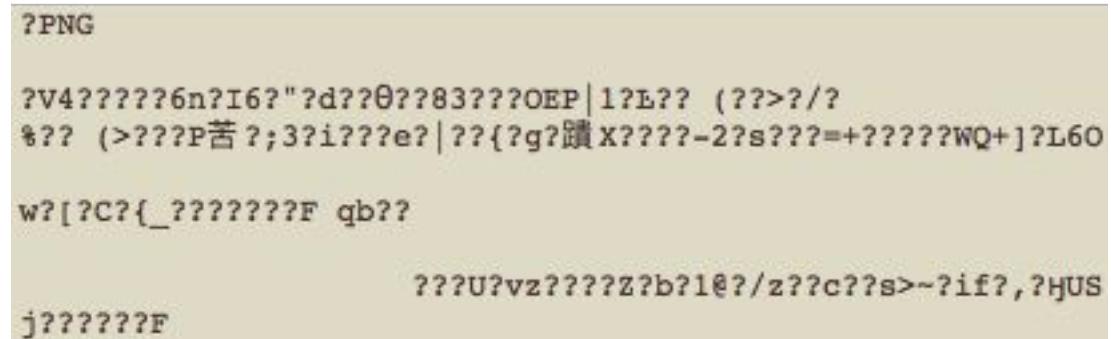
Bonjour tout le monde

Formats « binaires »

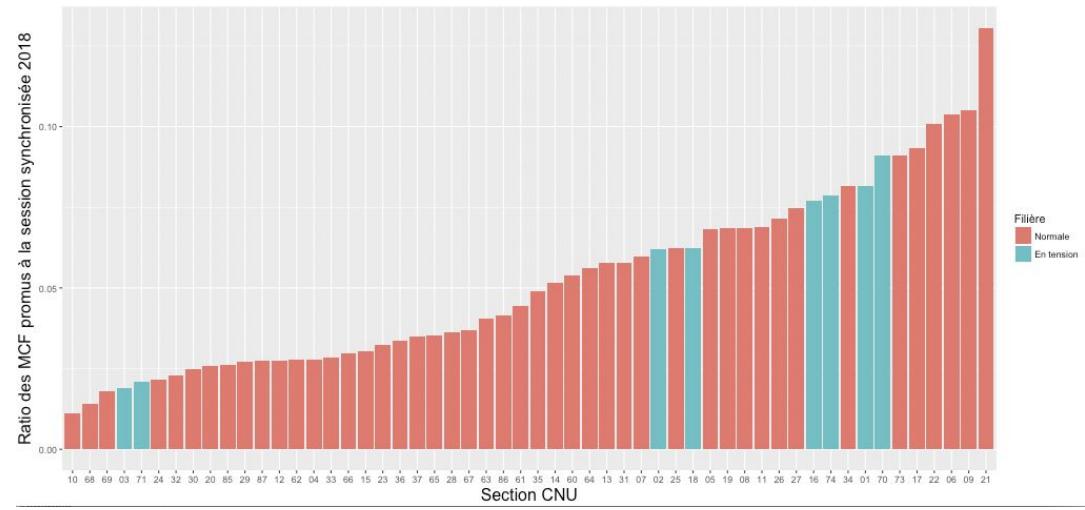
- Suite d'octets non interprétables comme des caractères imprimables ou affichables
- Structure interne opaque
- Besoin de logiciel spécifique pour les lire et les interpréter

Ex. de format binaire : PNG (image)

Contenu illisible dans un éditeur texte (à part «?PNG » au début)



Uniquement lisible et interprétable avec une visionneuse d'images.



Formats d'archives : ensemble de fichiers (binaires ou textuels) regroupés dans un fichier unique, compressé (ZIP, 7Z, RAR, TAR.GZ...) ou non (TAR)

Ex : le format DOCX (Word récent) est en fait une archive contenant plusieurs fichiers textuels (XML). Il faut changer l'extension en .ZIP pour lister le contenu.

```
[Content_Types].xml
rels
docProps
|   app.xml
|   core.xml
word
|   rels
|   document.xml.rels
|   document.xml
|   fontTable.xml
|   settings.xml
|   styles.xml
|   theme
|   theme1.xml
webSettings.xml
```

```
?xml version="1.0" encoding="UTF-8" standalone="yes"?
<w:document xmlns:wpc="http://schemas.microsoft.com/office/word/2010/wordprocessingCa
nvas" xmlns:cx="http://schemas.microsoft.com/office/drawing/2014/chartex" xmlns:cx1="
http://schemas.microsoft.com/office/drawing/2015/9/8/chartex" xmlns:cx2="http://schema
s.microsoft.com/office/drawing/2015/10/21/chartex" xmlns:cx3="http://schemas.microso
ft.com/office/drawing/2016/5/9/chartex" xmlns:cx4="http://schemas.microsoft.com/ffic
e/drawing/2016/5/10/chartex" xmlns:cx5="http://schemas.microsoft.com/office/drawing/2
016/5/11/chartex" xmlns:cx6="http://schemas.microsoft.com/office/drawing/2016/5/12/ch
artex" xmlns:cx7="http://schemas.microsoft.com/office/drawing/2016/5/13/chartex" xmln
s:cx8="http://schemas.microsoft.com/office/drawing/2016/5/14/chartex" xmlns:mc="http:
//schemas.openxmlformats.org/markup-compatibility/2006" xmlns:aink="http://schemas.mi
crosoft.com/office/drawing/2016/ink" xmlns:am3d="http://schemas.microsoft.com/office/
drawing/2017/model3d" xmlns:o="urn:schemas-microsoft-com:office:office" xmlns:r="http
://schemas.openxmlformats.org/officeDocument/2006/relationships" xmlns:m="http://sche
mas.openxmlformats.org/officeDocument/2006/math" xmlns:v="urn:schemas-microsoft-com:v
ml" xmlns:wp14="http://schemas.microsoft.com/office/word/2010/wordprocessingDrawing"
xmlns:wp="http://schemas.openxmlformats.org/drawingml/2006/wordprocessingDrawing" xml
ns:w10="urn:schemas-microsoft-com:office:word" xmlns:w="http://schemas.openxmlformats
.org/wordprocessingml/2006/main" xmlns:w14="http://schemas.microsoft.com/office/word/
2010/wordml" xmlns:w15="http://schemas.microsoft.com/office/word/2012/wordml" xmlns:w
16cid="http://schemas.microsoft.com/office/word/2016/wordml/cid" xmlns:w16se="http://
schemas.microsoft.com/office/word/2015/wordml/symex" xmlns:wpg="http://schemas.micro
soft.com/office/word/2010/wordprocessingGroup" xmlns:wpi="http://schemas.microsoft.com
/office/word/2010/wordprocessingInk" xmlns:wne="http://schemas.microsoft.com/office/w
ord/2006/wordml" xmlns:wps="http://schemas.microsoft.com/office/word/2010/wordprocess
ingShape" mc:Ignorable="w14 w15 w16se w16cid wp14"><w:body><w:p w:rsidR="00FE14AA" w:
rsidRDefault="00FE14AA"><w:r><w:t xml:space="preserve">Bonjour </w:t></w:r><w:r w:rsi
dRPr="00FE14AA"><w:rPr><w:color w:val="FF0000"/></w:rPr><w:t xml:space="preserve">tou
t le </w:t></w:r><w:bookmarkStart w:id="0" w:name="_GoBack"/><w:bookmarkEnd w:id="0"/
><w:r w:rsidRPr="00FE14AA"><w:rPr><w:color w:val="FF0000"/></w:rPr><w:t>monde</w:t></
w:r><w:p><w:sectPr w:rsidR="00FE14AA" w:rsidSect="00F22F18"><w:pgSz w:w="11900" w:h=
"16840"/><w:pgMar w:top="1417" w:right="1417" w:bottom="1417" w:left="1417" w:header=
"708" w:footer="708" w:gutter="0"/><w:cols w:space="708"/><w:docGrid w:linePitch="360
"/></w:sectPr></w:body></w:document>
```

Quelles conséquences pour vous ?
Et pour ceux qui arriveront plus tard ?

Formats et logiciel ?

Allez à : <https://scrumblr.ethibox.fr/format>

et listez les formats que vous connaissez dans les colonnes idoines

Les logiciels nécessaires pour traiter les formats cités sur scrumblr :

Fonctionnent-ils en ligne ou après installation sur un ordinateur ?

Fonctionnent-ils avec un système d'exploitation particulier (Windows, Mac, Linux) ?

Sont-ils liés à un type d'ordinateur ou à un instrument particulier (ex : microscope) ?

Sont-ils gratuits ou payants ? Qui paye ?

S'ils n'existaient plus ou si vous n'y avez plus accès, pourriez-vous continuer à travailler ?

L'éditeur du logiciel (ou la communauté) est-il en bonne santé ?

? Que proposez-vous pour garantir la pérennité de l'accès à vos données ?

Recommandations sur le format des fichiers

Privilégiez les formats ouverts afin de faciliter le partage des données

Format ouvert

Spécifications publiques et gratuites

Aucune restriction légale pour l'utiliser

Format indépendant du logiciel utilisé qui assure l'interopérabilité des données

Maintenu par une organisation à but non lucratif

Format fermé

Spécifications non publiques

Des restrictions légales s'opposent à son utilisation (droit d'auteur, copyright, brevet)

Format lisible qu'avec un logiciel particulier

Format propriétaire

Recommandations sur le format des fichiers

Type	Format conseillé	Format non conseillé
Document texte	PDF, TXT, ODT	MS Word, RTF
Feuille de calcul	ODS, CSV	MS Excel, PDF, OOXML
Base de données	SQL, SIARD, DB tables (.CSV)	MS Access, dBase (.dbf), HDF5
Données statistiques	SPSS Portable, STATA, XML, CSV, TXT	SAS et R
Images	JPEG, TIFF, PNG	DICOM
Audio	BWF, MXF, Matroska (.mka), FLAC, OPUS	<u>WAVE</u> , <u>MP3</u> , <u>AAC</u> , <u>AIFF</u> , <u>OGG</u>
Video	MXF, MKV	MPEG-4, MPEG-2, AVI, QuickTime (.mov, .qt)
Information géographique	GML, MIF/MID	ESRI Shapefiles, MapInfo, KML
Images géoréférencées	GeoTIFF (.tif, .tiff)	TIFF World File
Raster	ASCII GRID (.asc, .txt)	ESRI GRID

File formats for digital content: Probability for full long-term preservation

Content type	High	Medium	Low
Text	<ul style="list-style-type: none"> • Plain text (encoding: USASCII, UTF-8, UTF-16 with BOM) • XML (includes XSD/XSL/XHTML, etc.; with included or accessible schema) • PDF/A-1 (ISO 19005-1) (*.pdf) 	<ul style="list-style-type: none"> • Cascading Style Sheets (*.css) • DTD (*.dtd) • Plain text (ISO 8859-1 encoding) • PDF (*.pdf) (embedded fonts) • Rich Text Format 1.x (*.rtf) • HTML (include a DOCTYPE declaration) • SGML (*.sgml) • Open Office (*.sxw/*.odt) • OOXML (ISO/IEC DIS 29500) (*.docx) • Microsoft Word 2007 or newer (*.docx) 	<ul style="list-style-type: none"> • PDF (*.pdf) (encrypted) • Microsoft Word 2003 or older (*.doc) • WordPerfect (*.wpd) • DVI (*.dvi) • All other text formats not listed
Raster image	<ul style="list-style-type: none"> • TIFF (uncompressed) • JPEG2000 (lossless) (*.jp2) • PNG (*.png) 	<ul style="list-style-type: none"> • BMP (*.bmp) • JPEG/JFIF (*.jpg) • JPEG2000 (lossy) (*.jp2) • TIFF (compressed) • GIF (*.gif) • Digital Negative DNG (*.dng) 	<ul style="list-style-type: none"> • MrSID (*.sid) • TIFF (in Planar format) • FlashPix (*.fpx) • PhotoShop (*.psd) • RAW • JPEG 2000 Part 2 (*.jpf, *.jpx) • All other raster image formats not listed
Vector graphics	<ul style="list-style-type: none"> • SVG (no Java script binding) (*.svg) 	<ul style="list-style-type: none"> • Computer Graphic Metafile (CGM, WebCGM) (*.cgm) 	<ul style="list-style-type: none"> • Encapsulated Postscript (EPS) • Macromedia Flash (*.swf) • All other vector image formats not listed
Audio	<ul style="list-style-type: none"> • AIFF (96kHz 16bit PCM) (*.aif, *.aiff) • WAV (96kHz 24bit PCM) (*.wav) 	<ul style="list-style-type: none"> • SUN Audio (uncompressed) (*.au) • Standard MIDI (*.mid, *.midi) • Ogg Vorbis (*.ogg) • Free Lossless Audio Codec (*.flac) • Advance Audio Coding (*.mp4, *.m4a, *.aac) • MP3 (MPEG-1/2, Layer 3) (*.mp3) 	<ul style="list-style-type: none"> • AIFC (compressed) (*.aifc) • NeXT SND (*.snd) • RealNetworks 'Real Audio' (*.ra, *.rm, *.ram) • Windows Media Audio (*.wma) • Protected AAC (*.m4p) • WAV (compressed) (*.wav) • All other audio formats not listed
Video	<ul style="list-style-type: none"> • Motion JPEG 2000 (ISO/IEC 15444-4)??.mj2) • AVI (uncompressed/native, motion JPEG) (*.avi) • QuickTime Movie (uncompressed/native, motion JPEG) (*.mov) 	<ul style="list-style-type: none"> • Ogg Theora (*.ogg) • MPEG-1, MPEG-2 (*.mpg, *.mpeg, wrapped in AVI, MOV) • MPEG-4 (H.263, H.264) (*.mp4, wrapped in AVI, MOV) 	<ul style="list-style-type: none"> • AVI (others) (*.avi) • QuickTime Movie (others) (*.mov) • RealNetworks 'Real Video' (*.rv) • Windows Media Video (*.wmv) • All other video formats not listed



Recommandations sur le format des fichiers

Définition légale du **format ouvert** en France (loi no 2004-575 du 21 juin 2004) :

On entend par standard ouvert tout protocole de communication, d'interconnexion ou d'échange et tout format de données **interopérable** et dont les **spécifications techniques sont publiques et sans restriction d'accès ni de mise en œuvre**.

-> format bien documenté et utilisable sans demander d'autorisation

Ex : DOCX, ODT, XLSX, ODS, XML ; DOC et XLS depuis 2006

Format **fermés (ou propriétaires)** : format non ou mal documenté et/ou dont l'utilisation est payante ou soumise (au moins en théorie) à autorisation

Ex : PSD, JPEG, logiciels SPSS, STATA, SAS, NVIVO, Altas.ti

En pratique, on peut souvent travailler avec un format fermé populaire et le **convertir** en format ouvert. **Mais il faut vérifier si la conversion altère les informations, et prendre des mesures de compensation si nécessaire.**

Ex : la conversion XLSX -> CSV perd les mises en forme.

Formats standardisés

La documentation d'un format peut devenir une norme officielle nationale ou internationale ou un standard de facto.

Ex :

PDF/A1 est une version standardisée (ISO 19005) du format PDF. Les autres versions de PDF ne sont pas standardisées

Les formats Libre office (ODS, ODT...) sont standardisés (ISO/IEC 26300)

Le format XML est standardisé par une « recommandation » du W3C (équivaut à une norme)

Le format CSV est décrit dans la RFC 4180 de l'IETF, mais n'est pas réellement standardisé (la RFC est un document indicatif), plusieurs versions existent

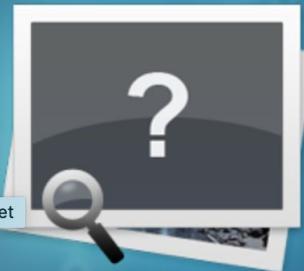
Les formats bureautique Microsoft (XLSX, DOCX...) sont standardisés (ISO/IEC 29500). Mais les logiciels semblent parfois s'écartier du standard

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#) [OTHER](#)**WELCOME TO DOTWHAT? ... THE LEADING FILE EXTENSION RESOURCE**

Thanks to years of research and help from our loyal visitors, we now have one of the world's largest and most detailed databases of file extension information, covering multiple operating systems from Microsoft's Windows, Apple's OS X and all variations of Unix to those used on the latest mobile devices and phones.

EVERYTHING YOU NEED TO KNOW! IF NOT, JUST ASK!

We try to provide as much information on each file extension as possible and we encourage visitors to contact us if they have any additional information on an extension or if they think a new file extension should be added to the database. Alternatively, each entry can be edited and visitors have the option of adding a comment, question or tip!

**Sections**

Software Developers



Software Products



Common File Extensions

Categories

3D/CAD Files



Audio Files



Backup Files



Compressed Files



Configuration Files



Data Files

Organisation des données



Organisation des données

- Définissez une politique d'organisation de vos données pour chaque projet
- Documentez et diffusez votre politique au sein de l'équipe
- La cohérence prime sur la préférence personnelle

Name	Modified
Ariane.jpg	2021-03-15 08:14 AM
Audrey.jpg	2021-03-15 08:51 AM
celia.JPG	2021-03-15 08:53 AM
christophe2.jpg	2021-03-15 09:23 AM
Claire.jpg	2021-03-14 10:11 PM
Dominique.jpg	2021-03-15 06:48 AM
Fred.JPG	2021-03-09 03:40 PM
Hélène.jpeg	2021-03-14 11:01 PM
jef.jpg	2021-03-15 08:50 AM
julien.JPG	2021-03-15 09:16 AM
loraine.jpg	2021-03-14 09:08 PM
Magali.JPG	2021-03-15 08:25 AM
Maxime.jpg	2021-03-15 09:31 AM
morganeT.JPG	2021-03-15 07:07 AM

Organisation des données

Organisation des dossiers

- Limitez le nombre de dossiers par niveau (5 ou 6 max)
- Allez du général au spécifique
- Choisissez des noms de dossiers explicites

Pour chaque dossier, ajoutez un fichier README:

- Choisissez un format simple et ouvert (par exemple Markdown ou TXT)
- Indiquez un minimum de méta-données concernant le dossier et son contenu :
 - Titre
 - Date de création / réception des données
 - Origine/Source des données
 - Version
 - Propriétaire/responsable des données
 - Organisation des données
 - Méthode de réception/téléchargement des données

Organisation des données

Exemple :

Un dossier par projet

 Un sous-dossier par type de manip (microscopie, séquençage, phénotypage)

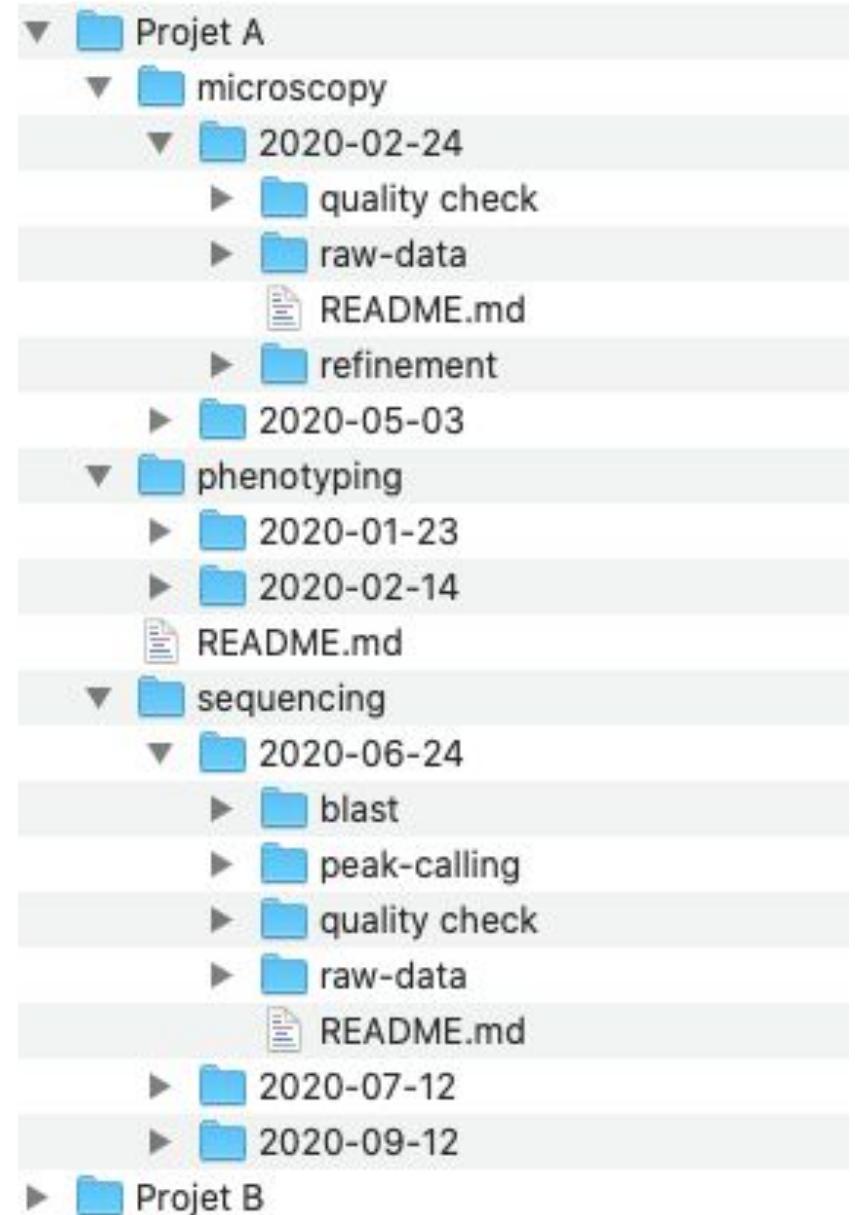
 Un sous-dossier par date (2020-02-24, 2020-05-03)

 Un sous-dossier pour les données brutes

 Un sous-dossier par analyse (contrôle qualité, nettoyage statistique, raffinement)

 Un sous-dossier par publication

 Un lien symbolique vers chaque dossier données ou analyse associé à la publication



Un cas d'usage

Situation : une partie de vos données est considérée comme données sensitives.

Q : quelles actions sont-elles à déployer pour garantir votre contrôle sur les accès à ces données ?

Protéger ses données



Intégrité des données

Identifier et contrôler la corruption des données

- Corruption : introduction de modifications non intentionnelles des données

Les données peuvent être corrompues par :

- des modifications non souhaités (ransomware, collègue...)
- un transfert de données défectueux
- un plantage d'un disque dur
- ...

Intégrité des données

Identifier et contrôler la corruption des données

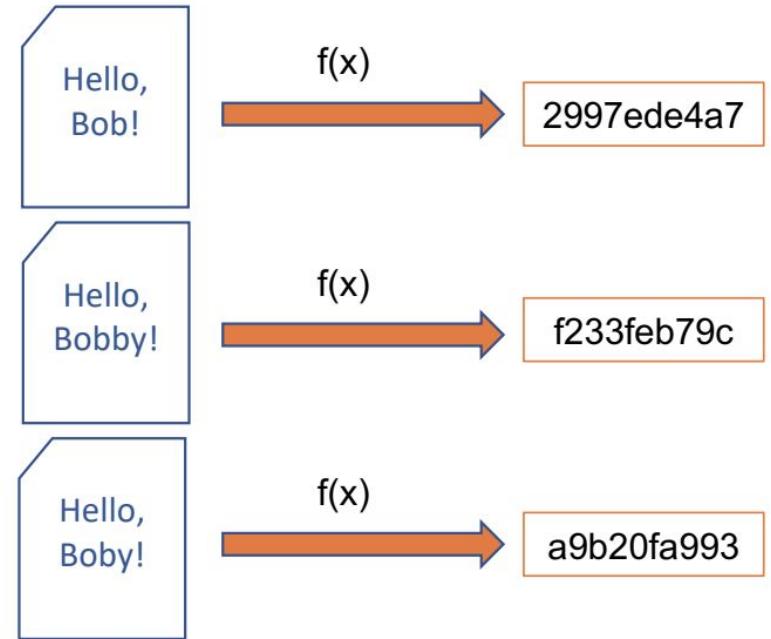
Solution 1 : générer des sommes de contrôles

Comment ?

- Linux / macOS : md5sum, sha256sum
- Windows : certutil

Quand ?

- Avant un transfert de données
 - Lorsqu'on réceptionne un nouveau jeu de données d'un collaborateur
 - Lorsqu'on transfert des données sur un stockage distant
- Stockage à long terme
 - La version principale de chaque dataset
 - Les extraits de données utilisés dans les publications



Intégrité des données

Identifier et contrôler la corruption des données

Solution 2 : utilisez le contrôle d'accès

N'accordez que les permissions d'accès nécessaire :

- Limitez le nombre d'utilisateurs ayant accès à vos données
- Limitez la visibilité des données (réseau interne vs internet)
- N'utilisez jamais de partage public sans chiffrement des données !

Mettez les données brutes en lecture seule

L'accès aux données sensibles doit être documenté

Gérer l'accès à son projet

Project mytest

Management console

New project member

User id

Request member

Remove project member

▼

Request member removal

Copie des données

Limitez les copies au maximum !

- Copie principale (master)
 - Également appelé donnée “source” ou “brute”
 - Stratégie 3-2-1
- Copie de travail
 - A éviter au maximum
 - Utilisez des liens symboliques vers la copie principale
- Copie de sauvegarde
 - Ne travaillez jamais sur votre copie de sauvegarde

Un cas d'usage

Situation : vous êtes régulièrement obligés (par votre institution etc...) de procéder au nettoyage des données que vous sauvegardez sur le serveur. Cette obligation s'accompagne de la nécessité de justifier la raison pourquoi les dossiers listés doivent être conservés.

A cet effet, les dossiers sont taggés/catégorisés : par exemple “pour publication”, etc...

Quelles catégories seraient pertinentes pour justifier la conservation des données ?

La suppression des données



Suppression des données

Est-ce que ces données peuvent être supprimés ?

Le stockage des données a un coût financier et écologique.

- Distinguez clairement la copie principale (master) de ses dérivés
- Organisez régulièrement une revue des données
- Récupérer rapidement les données sur supports externes (disque ou clé USB)



Un petit exercice :

Quels jeux de données puis-je supprimer ?
(on va pas être d'accord)

<https://www.wooclap.com/JNHKXR>

Conservation des données

“Je veux garder mes données pour l'éternité”

Ne manquez pas le module 4...

- Quels sont vos obligations en terme de rétention de données
- Dans quelles conditions allez-vous les archiver ?
- Avez-vous documenter clairement vos données ?
- Que se passera-t-il si vous partez (pour l'éternité) ?

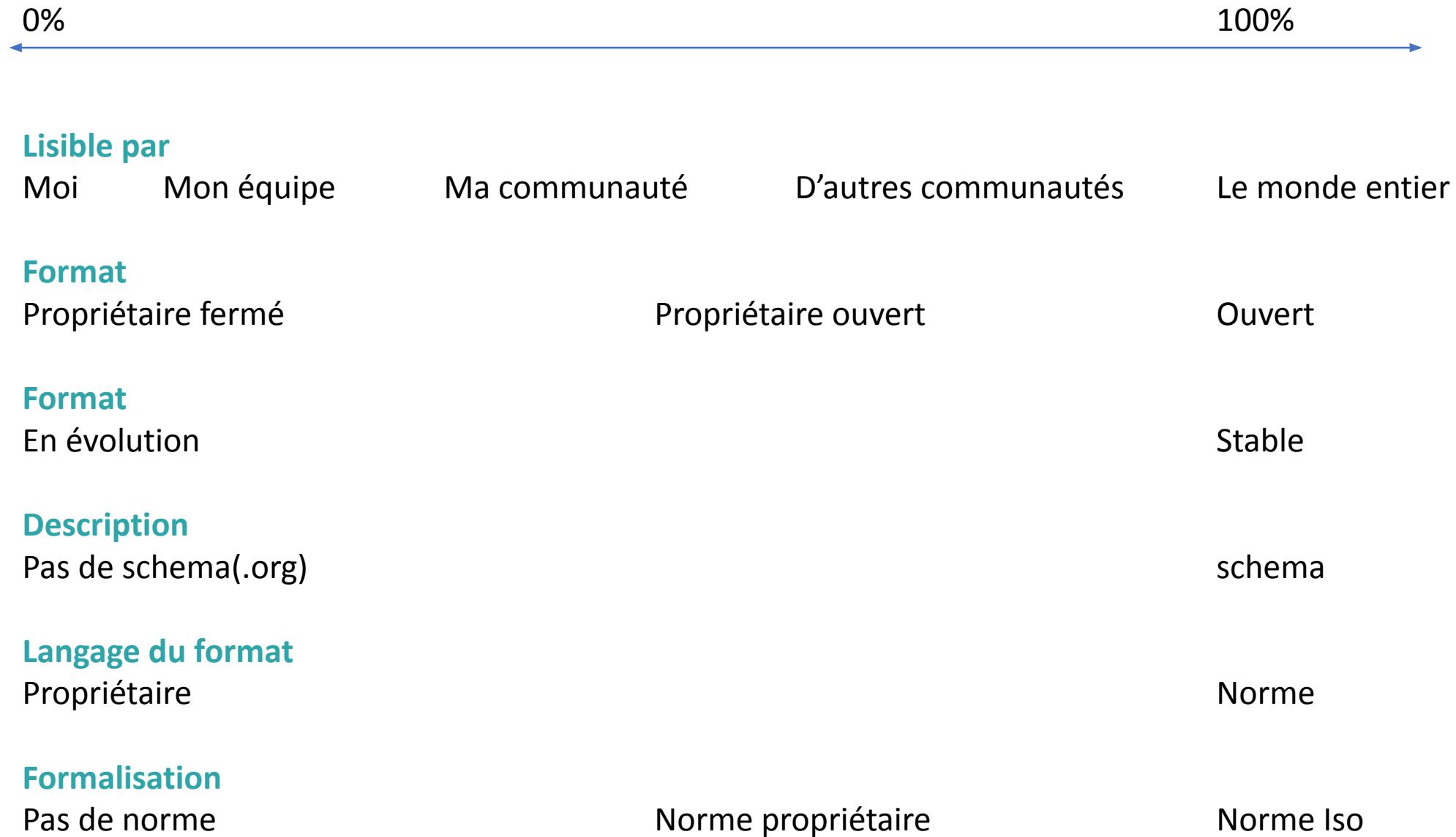
Les infrastructures de stockage sont vos amies

- Politique de sauvegarde professionnel et cohérente
- Nombre de copies minimum (stratégie 3-2-1)
- Gestion claires des droits d'accès
- Haute disponibilité et accessibilités
- Sécurité

Essayons de nous améliorer

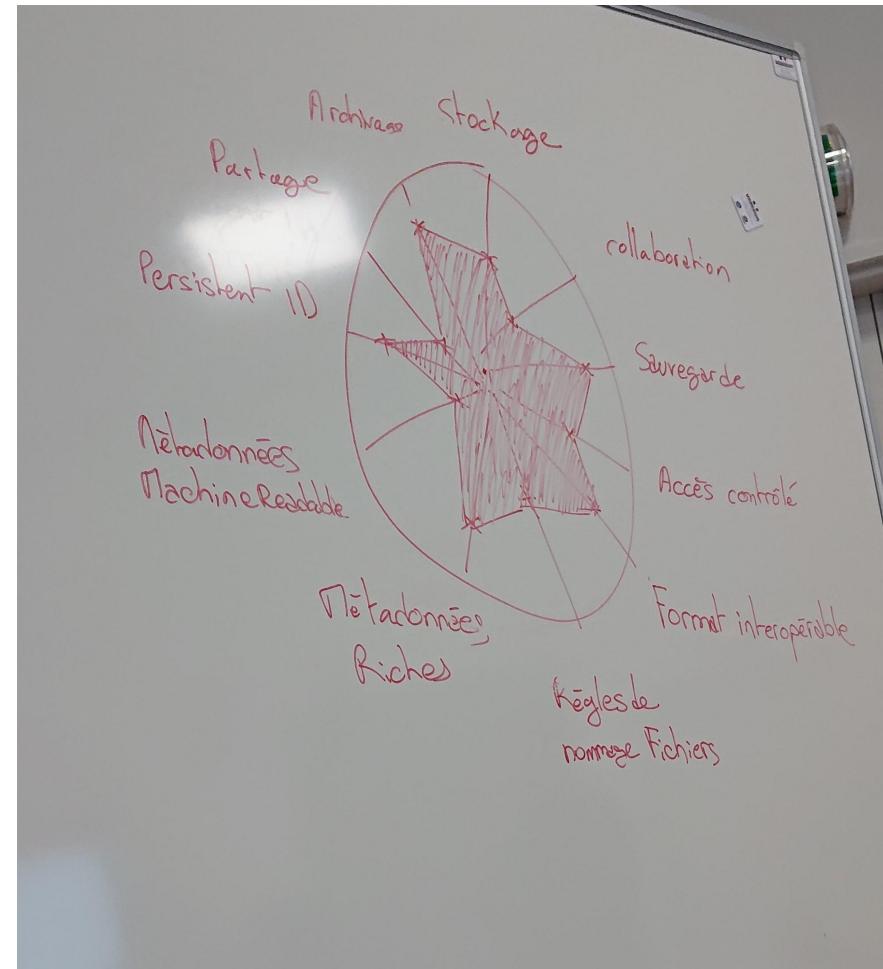


Où se situe mon fichier ?



Exercice

- Télécharger la matrice Excel **modèle radar.xlsx** sur osf.io
- Donnez une note de 0 à 5 pour chaque critère pour votre fichier



Gestion Electronique de Documents



Espace de :
Préservation
et de
Partage
du
Savoir
du
Groupe





Tout le monde à sa GED

Gratuite ou payante
Bonne ou mauvaise
Choisie ou imposée

Espace personnel cirad

Afficher la barre latérale

Vie pratique

Documents > 6Bioinfo : MolInfo13_CreaCompteUtilisCC2_V2.pdf

MolInfo13_CreaCompteUtilisCC2_V2.pdf Modifié par Patricia Turquay le 16/06/2013 à 16:00

CIRAD Frédéric

Filex : Échangez vos fichiers

Cette application permet d'échanger des fichiers avec les autres applications.

Pour en savoir plus

Activités de l'application

16:00, 4 juin, 2013 16:00 Catherine Turquay

15:00
15:41 Catherine Turquay
15:41 Catherine Turquay
15:33 Catherine Turquay
15:32 Catherine Turquay
15:20 Patricia Turquay

11:00
11:52 Catherine Turquay

17:00, 3 juin, 2013 17:00 Najate Jaffar

16:00
16:12 Catherine Turquay
16:10 Catherine Turquay
16:08 Catherine Turquay
16:07 Catherine Turquay
16:07 Catherine Turquay

15:00
15:58 Catherine Turquay
15:57 Catherine Turquay
15:29 Dominik

16:00, 30 mai, 2013 16:32 Najate Jaffar
16:13 Najate Jaffar Commentaires

14:00, 29 mai, 2013 Ajouter un commentaire
14:10 Catherine Turquay Pas de commentaire
14:04 Catherine Turquay
14:04 Catherine Turquay
14:04 Catherine Turquay
14:03 Catherine Turquay

13:00
13:57 Catherine Turquay
13:57 Catherine Turquay
13:55 Catherine Turquay
13:25 Stéphanie

11:00
11:42 François

Historique Référence Version Version 00 Version 01 Version 02 Confidentialité Sommaire 1. Object 2. Contac 3. Mots clés 4. Cratique 5. Gestio 6. Supr 7. Historial 8. Docum Fiche de vi...
1. Objets L'objet compte
2. Contacts DS1 : Philip
Sectré
Sylvie
Agnès
3. Mots clés Se réf...

Workflow

Ce document ne fait partie d'aucun workflow.

Historique des versions

Dernière version

2.0 MolInfo13_CreaCompteUtilisCC2_V2.pdf

Patricia Turquay il y a un an environ (Pas de commentaire)

Versions antérieures

1.4 MolInfo13_CreaCompteUtilisCC2_V2.pdf

Patricia Turquay il y a un an environ (Pas de commentaire)

1.3 MolInfo13_CreaCompteUtilisCC2_V2.pdf

Patricia Turquay il y a un an environ (Pas de commentaire)

1.2 MolInfo13_CreaCompteUtilisCC2_V1.pdf

Patricia Turquay il y a plus de 4 ans (Pas de commentaire)

1.1 MolInfo13_CreaCompteUtilisCC2_V1.pdf

Patricia Turquay il y a plus de 4 ans (Pas de commentaire)

1.0 MolInfo13_CreaCompteUtilisCC2_V0.pdf

Patricia Turquay il y a plus de 4 ans (Pas de commentaire)

Rechercher des personnes, sites et documents

Télécharger

is?nodeRef=workspace://SpacesStore/d38-751a-4101-9e22-ac217e0cc742/MacesStore/9549bd38-751a-4101-9e22-

83

Gérez vos données de la recherche - Formation

Contributors: Jean-François Martin, alexandre dehne garcia, Frédéric de Lamotte, Victor REYS, Jonaz Vasquez-Villegas, Jeff Rouri, Urcel Kalenga, Dalbard Swanr, Alexandre Benzari, GILLES Andre, Mariam BAHRO, Charlotte, Johanna Girodolle, Germain Valentin Faity, Liyan OUYANG, chayma ben maamer

Date created: 2020-01-06 05:03 PM | Last Updated: 2020-06-26 03:29 PM

Category: Project

Description: Add a brief description to your project

Wiki

Page d'accueil

Par ordre alphabétique écrivez vos initiales suivies de votre prénom et nom Gérez vos données de la recherche - Formation

- ADG : Alexandre Dehne Garcia
- GVF : Germain Valentin Faity
- JFM : Jean-François Martin
- F2L : Frédéric de Lamotte
- LKM : Charlotte Kinowski-Moysan

[Read More](#)

Files

Click on a storage provider or drag and drop to upload

Name	Modified
📁 Gérez vos données de la recherche - Formation	
- OSF Storage (Germany - Frankfurt)	
└ 00_AtelierFilRouge.pptx	2020-01-22 04:49 PM
└ 01-En route vers l'open science.pptx	2020-01-22 02:08 PM
└ 02- les connées de la recherche et l'open data.pptx	2020-01-22 02:06 PM
└ 03- La Vie Des Données.pptx	2020-01-22 03:30 PM
└ 04- PlanGestionDonnees.pptx	2020-01-22 04:49 PM
└ 05- gestion des données pendant le projet(2).pptx	2020-01-23 01:58 PM
└ 07_Nommage_format.pptx	2020-01-23 01:59 PM
└ 08_Metadata.pptx	2020-01-23 01:59 PM
└ 09- Diffusez et partagez les données de recherche.pptx	2020-01-23 01:59 PM
└ 10- Droit des données - Cas pratique.pptx	2020-01-24 09:13 AM
└ 4.1_dmpLifeCycleMatrix.xlsx	2020-01-22 04:33 PM
└ entrepot_doc_parametres_synop_168.pdf	2020-01-23 11:29 AM

Mendeley

Enter citation style (e.g. "APA")

Citation	Actions
□ Marx, V. (2013). Biology: The big challenges of big data. <i>Nature</i> , 498(7453), 255–...	□ ⌂

Citation

Components

Add components to organize your project.

Tags

academic training x data management x open data x open science x Add a tag

Recent Activity

Frédéric de Lamotte linked Mendeley folder DataViz to Gérez vos données de la recherche - Formation 2020-06-26 03:29 PM

Frédéric de Lamotte authorized the Mendeley addon for Gérez vos données de la recherche - Formation 2020-06-26 03:28 PM

Frédéric de Lamotte added add-on Mendeley to Gérez vos données de la recherche - Formation 2020-06-25 01:54 PM

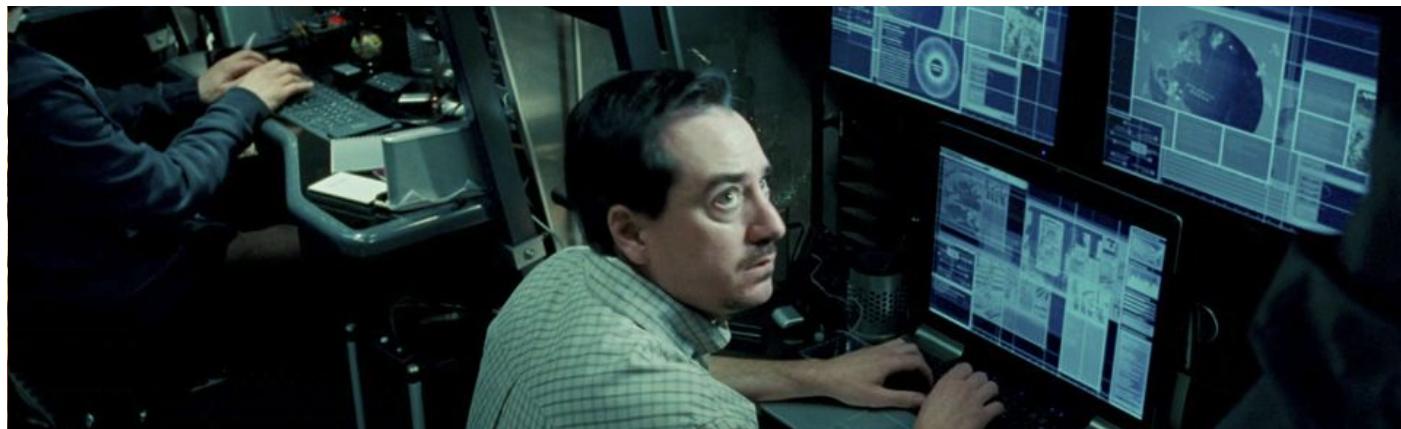
Germain Valentin Faity updated wiki page Home to version 6 of Gérez vos données de la recherche - Formation 2020-02-06 04:45 PM

Frédéric de Lamotte updated wiki page j'ai des problèmes ! to version 11 of Gérez vos données de la recherche - Formation 2020-02-04 01:55 PM

Frédéric de Lamotte updated wiki page j'ai des problèmes ! to version 10 of Gérez vos données de la recherche - Formation 2020-02-03 05:11 PM

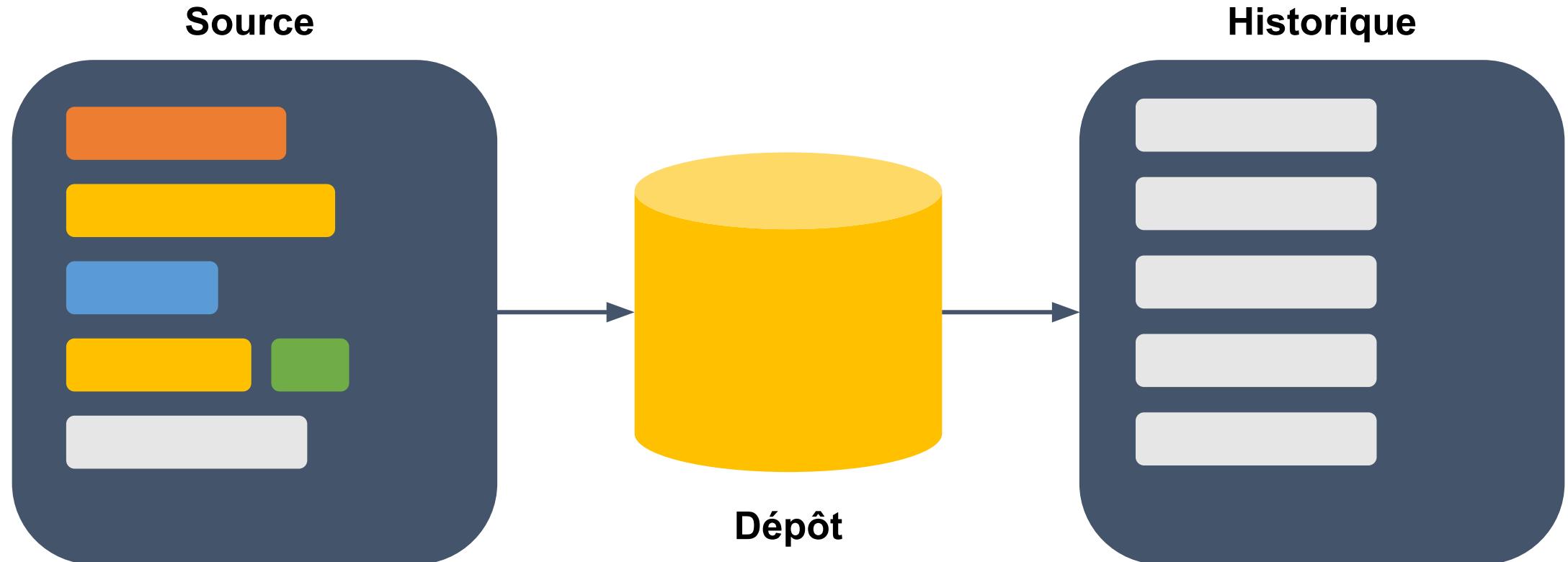
< 1 2 3 4 ... 14 >

Gestion des codes sources



C'est quoi Git ?

Git est un système de gestion de versions (version control system)



C'est quoi Git ?

v1.0.0

v1.0.1

v1.0.2

C'est quoi Git ?

Alice

Bob

v1.0.0

v1.0.0

v1.0.1

v1.0.1

v1.0.2

v1.0.2

C'est quoi Git ?

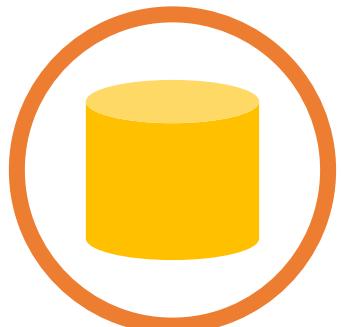
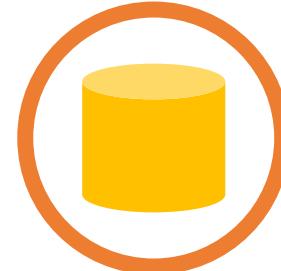
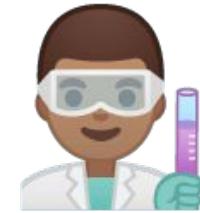
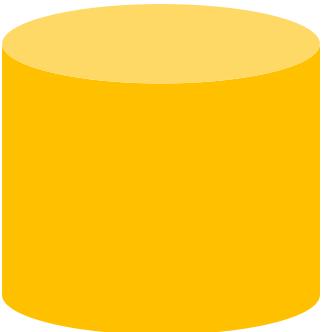
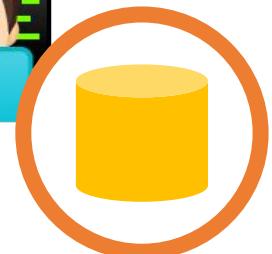
Système de gestion des versions

Suivre les changements

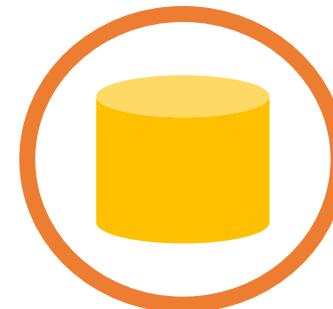
Travailler ensemble

C'est quoi Git ?

Git est un système de gestion de versions DISTRIBUÉ

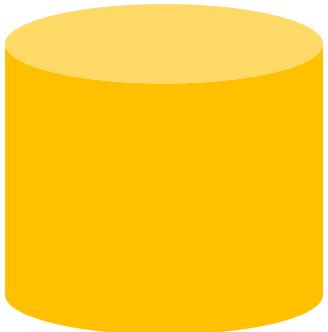
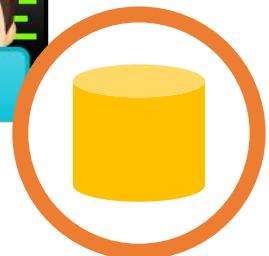


Serveur Git

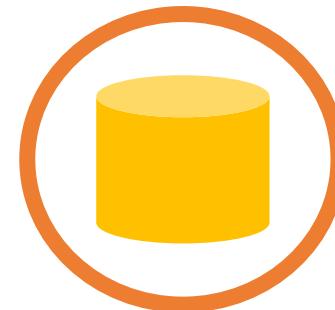
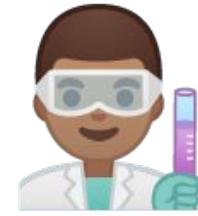
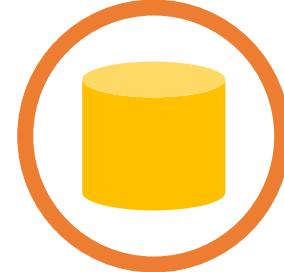


C'est quoi Git ?

Git est un système de gestion de versions DISTRIBUÉ

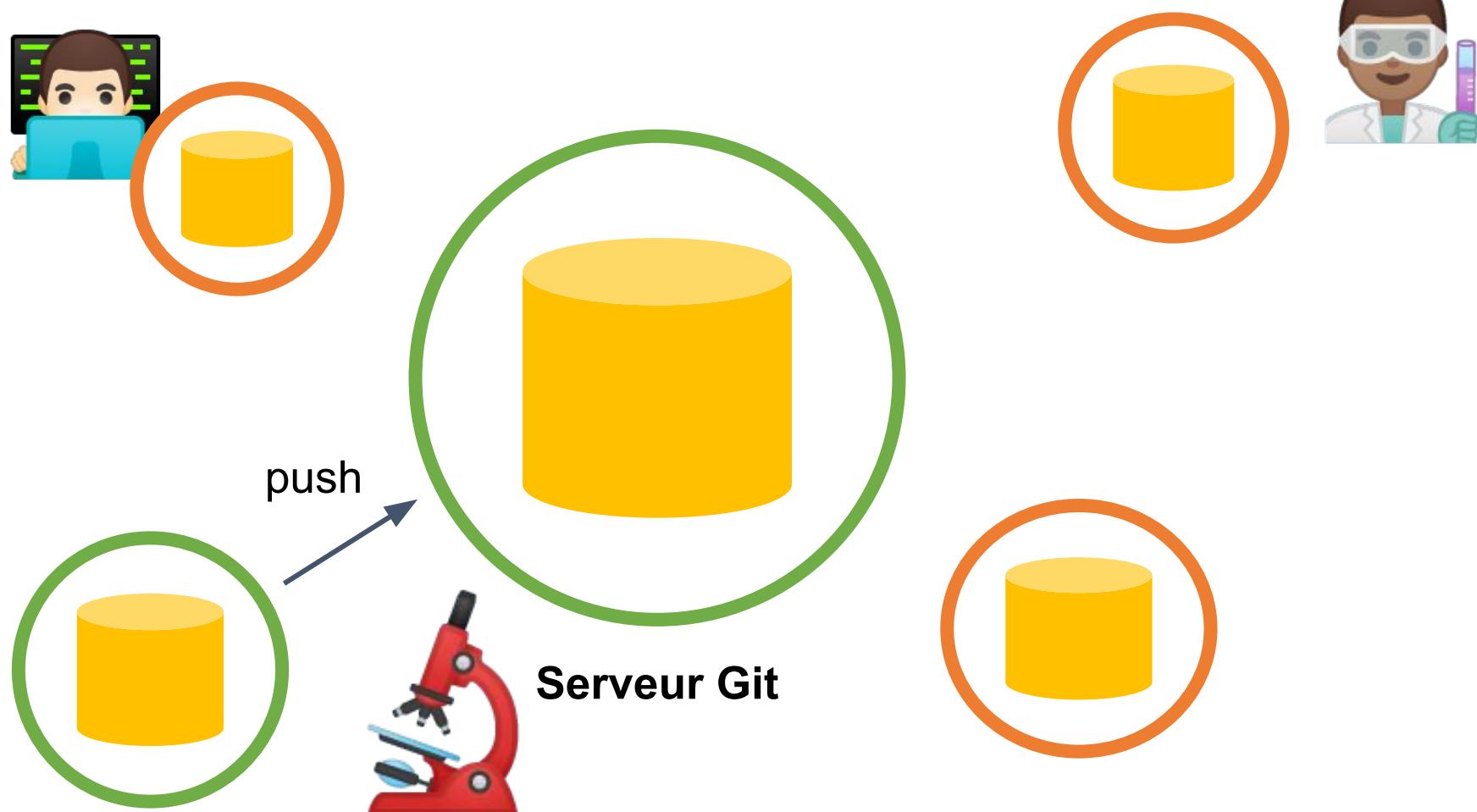


Serveur Git



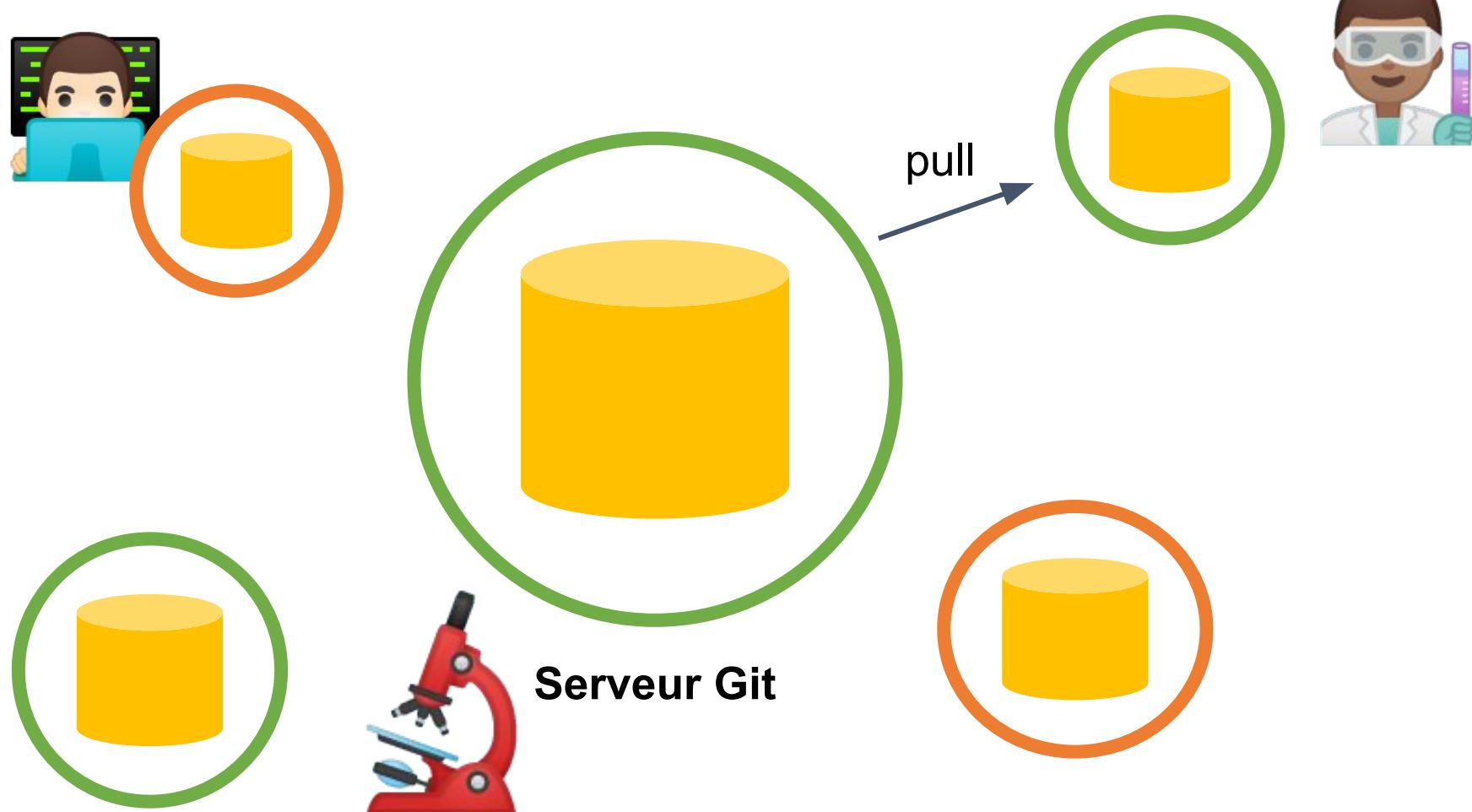
C'est quoi Git ?

Git est un système de gestion de versions DISTRIBUÉ



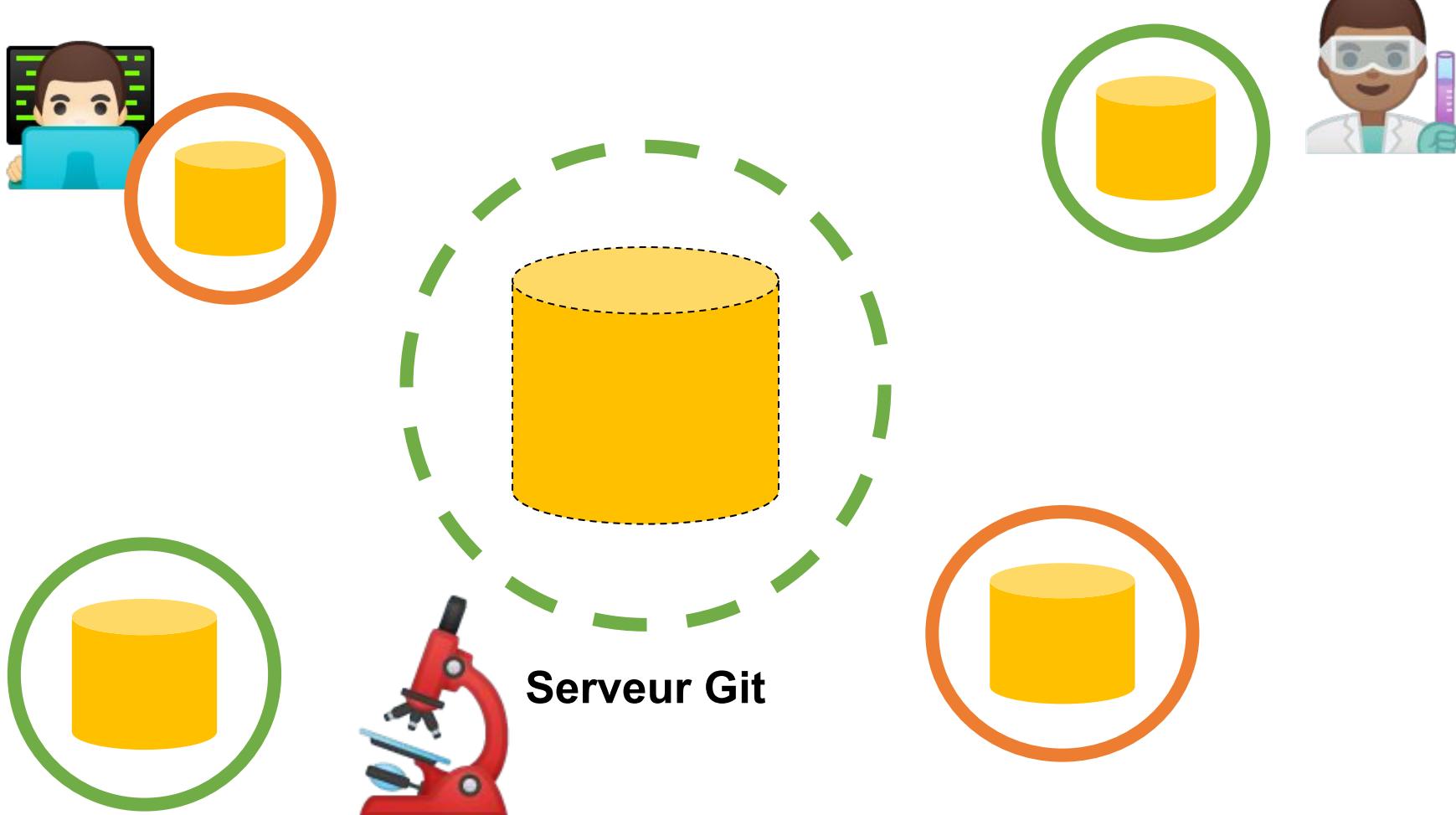
C'est quoi Git ?

Git est un système de gestion de versions DISTRIBUÉ



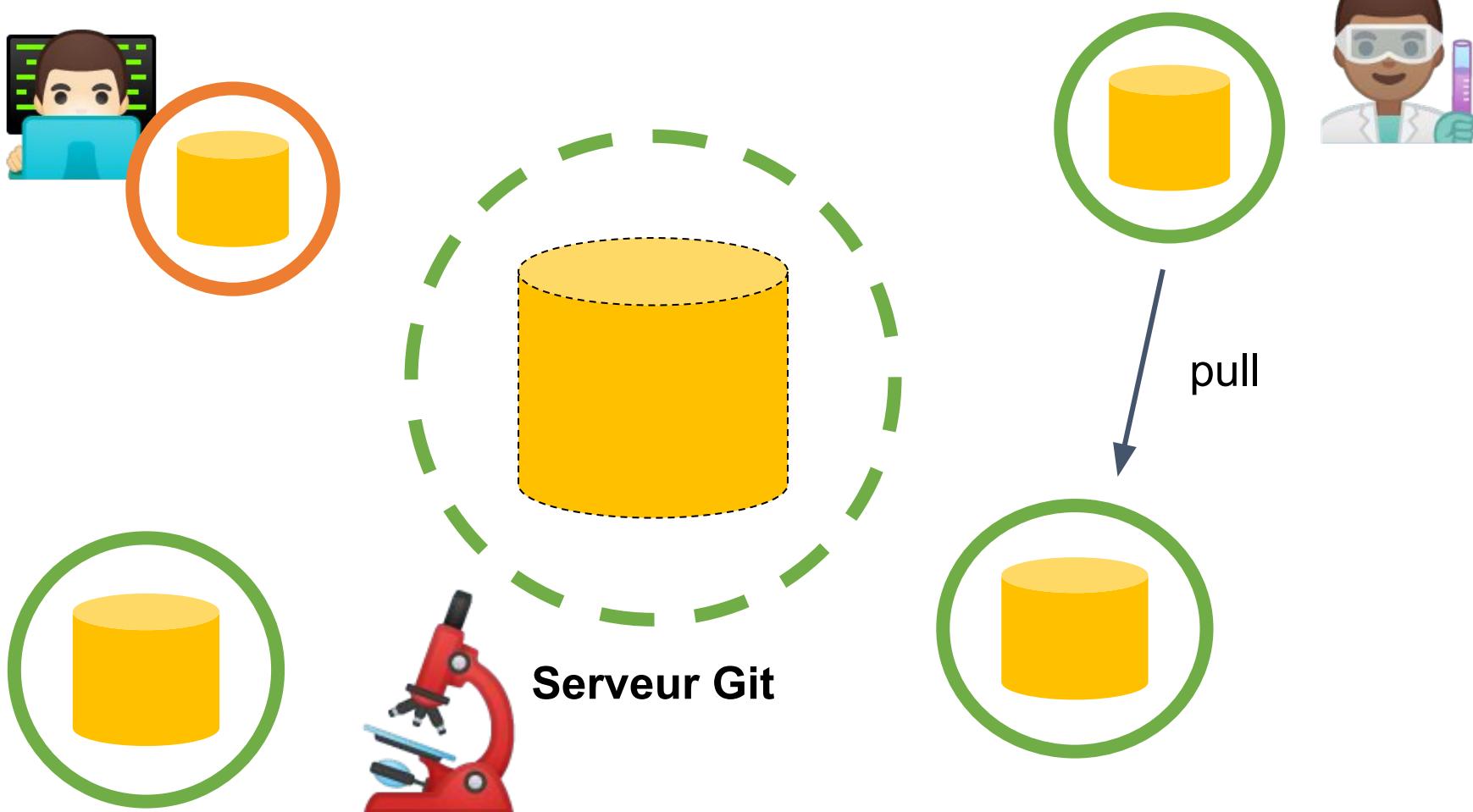
C'est quoi Git ?

Git est un système de gestion de versions DISTRIBUÉ

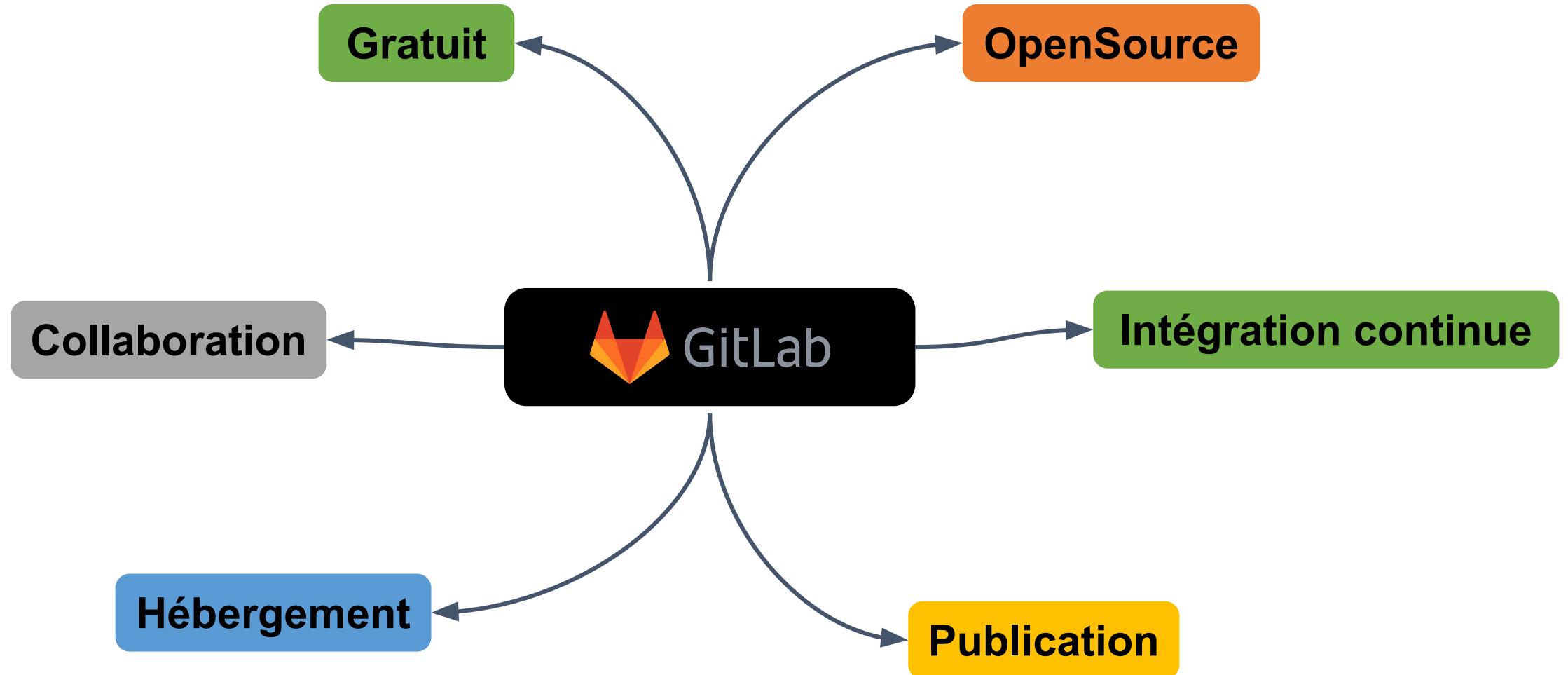


C'est quoi Git ?

Git est un système de gestion de versions DISTRIBUÉ



C'est quoi Git ?



**Merci !
mais au fait...
vous les avez tous ?**



La vie aquatique



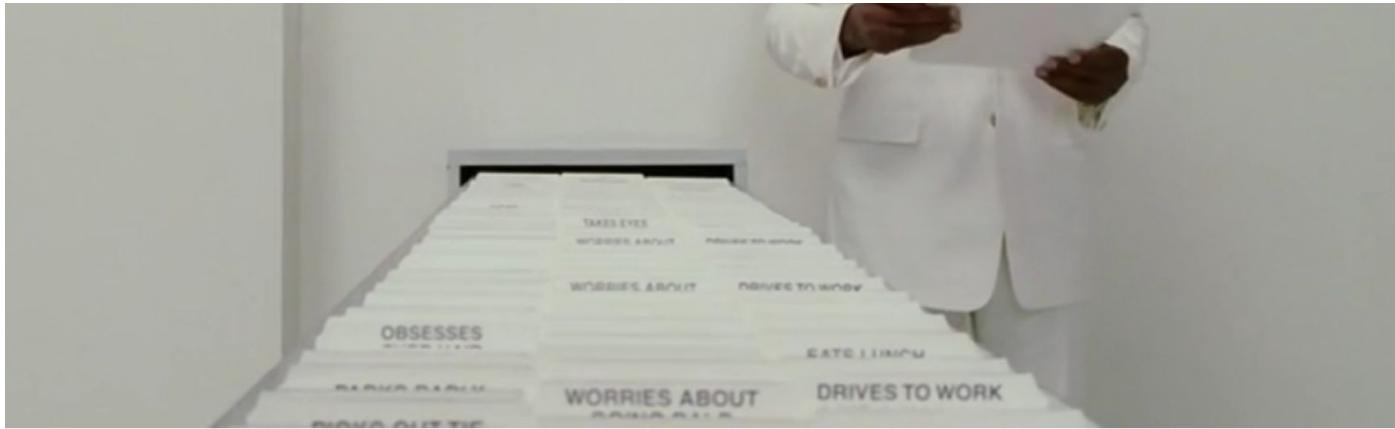
Contagion



Alien, le retour



Indiana Jones



Bruce Tout Puissant



Pulp Fiction



La Momie



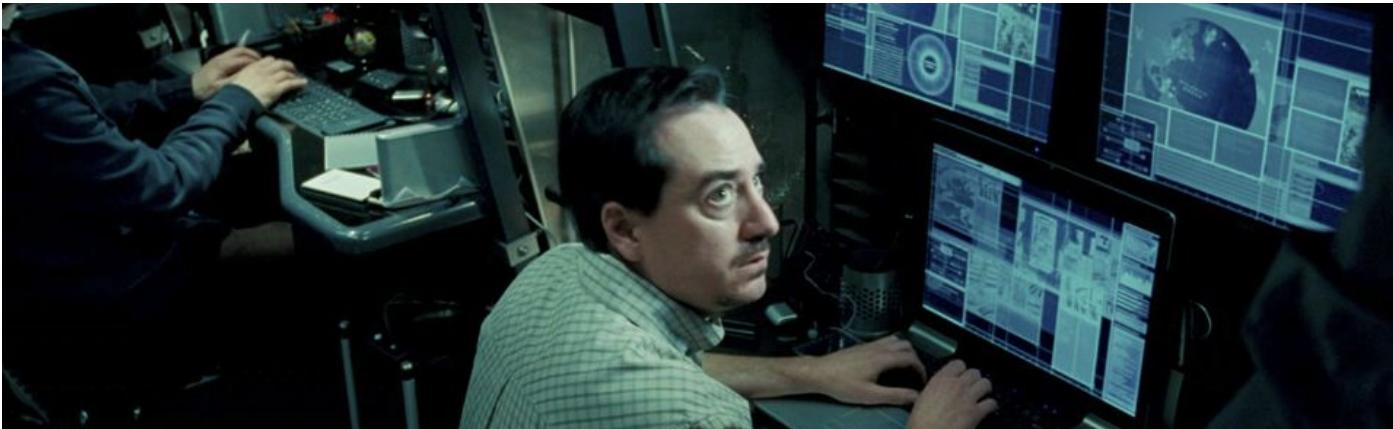
Matrix, reloaded



Wall-E



Lucy



Die Hard, retour en enfer