

Questions

https://huit.re/questions_2018-03-05

Documenter : commentaire de fichiers readme

Les 3 fichiers readme considérés proviennent des sources suivantes.

1. Exemple 1

Wilkinson, T. C. (2017, 24 novembre). *Project Panormos Archaeological Survey: Method Design (survey-design)* [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.1063932>

2. Exemple 2

De Valença, J. C., Kurniawan, A., Wagterveld, R. M., Wood, J. A., & Lammertink, R. G. H. (2017). *Influence of Rayleigh-Bénard convection on electrokinetic instability in overlimiting current conditions* [Dataset].

<https://doi.org/10.4121/uuid:b55184a0-930d-4b5c-940b-fd3038e05d83>

3. Exemple 3

4TU Centre for Research Data. (2017, décembre). Guidelines for creating a README file. Repéré à http://researchdata.4tu.nl/fileadmin/editor_upload/pdf/README/Guidelines_for_creating_a_README_file.pdf

- Quels points communs repérez-vous entre ces exemples?
- Quelles différences observez-vous?

Diagnostic : étape 1

Complétez la rubrique **Projet** : description globale allégée = principales caractéristiques de votre projet (ou d'un de vos projets ou de votre activité), sous forme formelle et informelle.

- Formelle = rubrique
 - titre
 - durée
 - contexte : membres du projet
 - types de données
 - types de fichiers
- Informelle = 2 ou 3 phrases descriptives

Diagnostic : étape 2 : structure

- Complétez la rubrique **Structure actuelle** en indiquant la façon dont vous structurez vos dossiers actuellement.

- Reproduisez le schéma général de votre arborescence si possible,
- sinon précisez selon quelle logique vous procédez.
- Précisez le cas échéant les différents systèmes que vous utilisez s'ils obéissent à des logiques différentes.
- Quel est le principal avantage de votre système actuel?
- Quel est son principal inconvénient?

Automatiser l'inventaire

Pour vous aider dans cette tâche, [Q-Dir](#) permet d'exporter sous la forme d'un fichier .csv le contenu d'un dossier.

- Un clic-droit sur les colonnes permet de sélectionner les colonnes à afficher.
- Le menu **Fichier > Exporter** permet de lancer l'export.

Diagnostic : étape 2 : structure : nouvelle

Complétez la rubrique **Dossiers**.

- Décrivez l'organisation générale de votre structure cible et sa logique.
- Précisez si cette structure correspond aux conventions de votre discipline, de votre structure, etc.
- Précisez comment ces dossiers se distinguent des dossiers que vous utilisez pour d'autres projets.

Complétez la rubrique **Structure cible**.

Rédigez votre arborescence sous la forme suivante.

```
dossier1\sous-dossier1.1
dossier1\sous-dossier1.2
dossier2
dossier3
```

Générer automatiquement une arborescence avec Text 2 Folders

[Text 2 Folders](#) permet de créer très rapidement des dossiers à partir d'un fichier .txt.

- Créez un fichier .txt en copiant l'arborescence que vous avez défini à l'étape 2.
- Précisez un dossier racine.
- Créez les dossiers.

Nommage des fichiers

A partir de ce que nous avons déjà vu, vous ne devriez pas avoir de difficultés à identifier le meilleur nom de fichier parmi les quelques exemples suivants.

Diagnostic : étape 3 : nommage

- Complétez la rubrique **Schéma actuel**.
 - Indiquez le schéma de nommage si applicable,
 - sinon précisez selon quelle logique vous procédez.
 - Précisez le cas échéant les différents systèmes que vous utilisez s'ils obéissent à des logiques différentes
- Quel est le principal avantage de votre système actuel?
- Quel est son principal inconvénient?

Diagnostic : étape 3 : nommage : nouveau

Complétez la rubrique **Fichiers > Nommage**.

- Décrivez la logique générale de votre schéma de nommage cible.
- Précisez si ce schéma correspond aux conventions de votre discipline, de votre structure, etc.
- Précisez comment les fichiers se distinguent des fichiers d'autres projets.

Complétez la rubrique **Schéma cible**.

Indiquez votre (ou vos) schéma(s) de nommage, avec un exemple pour chaque schéma.

Automatiser : nommage par lot

[Ant Renamer](#) permet de renommer par lot vos fichiers et dispose de fonctionnalités étendues.

On utilise aujourd'hui une fonctionnalité très simple, AntRenamer permet de faire des traitements bien plus sophistiqués et complexes.

Renommez le lot d'images :

- en ajoutant une énumération sur 3 caractères à la fin du nom de fichier, sous la forme `_001.jpg` , `_002.jpg` , etc. ,
- en remplaçant la chaîne de caractères précédant cette énumération par la chaîne de caractères : `windows_echantillons_2018-02` ;
- un nouveau nom de fichier ressemblera à : `windows_echantillons_2018-02_001.jpg`

Diagnostic : étape 4 : versionnage : nouveau

Complétez la rubrique **Fichiers > Versionnage**.

- Le contrôle de versions est-il géré manuellement ou automatiquement?
- Comment documentez-vous le versionnage (indication du numéro de version + statut + identification des modifications)?
- Quelle est le destin des différentes versions? Quelles versions intermédiaires sont-elles supprimées par exemple?
- Quels critères font-ils passer d'une version majeure à l'autre? Quels critères constituent-ils un changement

majeur (et non mineur)?

Bonnes pratiques pour les données tabulaires : les noms de variable

Pouvez-vous améliorer les en-têtes de colonne ci-dessous?

- max tempt °C
- airport/faa code
- M/F
- 12. quel est votre profil?

Bonnes pratiques pour les données tabulaires : faire le ménage

Le fichier `survey_data_spreadsheet_messy.xls` n'applique pas toutes les bonnes pratiques.

Pouvez-vous créer dans ce classeur une feuille `tidy` présentant les données selon les bonnes pratiques?

Diagnostic : évaluer les risques

Considérez chacun des items de la fiche *Data loss scenari*

- Avez-vous déjà expérimenté l'une de ces situations?
- Dans quels scénarii ne perdez-vous aucune donnée?
- Au contraire, quel scénario occasionnerait la plus grande perte de données pour vous?
- Une autre situation vous a-t-elle conduit à perdre des données?