



HAL
open science

Guide pour la documentation et la diffusion de données d'enquêtes avec la suite logicielle Colectica

Lucie Marie

► **To cite this version:**

Lucie Marie. Guide pour la documentation et la diffusion de données d'enquêtes avec la suite logicielle Colectica. Centre de données socio-politiques (CDSP). 2022. hal-03845847

HAL Id: hal-03845847

<https://sciencespo.hal.science/hal-03845847v1>

Submitted on 9 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License

Guide pour la documentation et la diffusion de données d'enquêtes avec la suite logicielle Colectica

Auteur : Lucie MARIE

Validation : Alina Danciu
Date : 25 octobre 2022, Version 1.0

Ce guide a été produit à partir d'une expérimentation menée au Centre de Données Socio-Politiques dans le cadre du WP2 du projet *UpMet (Upscaling Metadata for increasing reuse in social science)* financé par l'Agence National de la Recherche. Cette procédure s'appuie sur la version 5.5.7410 de Colectica et permet de produire une documentation respectant le standard DDI 3.2.

Glossaire	4
Procédure de documentation	5
1.1. Principes DDI-Lifecycle	5
1.2. Inventaire et curation	5
1.3. Référentiel de métadonnées	5
1.4. Connexion des éléments	5
1.5. Publication	6
La suite logicielle Colectica	6
2.1. Les outils	6
2.2. L'interface d'édition	7
Référentiel de métadonnées	8
3.1. Eléments	8
3.2. Organisation	10
3.3. Règles de nommage	12
3.4. Ajout de métadonnées	12
Documentation d'un jeu de données	14
4.1. Série	14
4.2. Etude	15
4.2. Questionnaire	15
4.3. Jeu de données	16
4.4. Variable	17
Stockage et diffusion	20
6.1. Association des outils	20
6.2. Publication sur le portail en ligne	20
6.3. Export de la documentation	21

Glossaire

Colectica Designer : Editeur de métadonnées qui, connecté au *repository*, permet de créer et modifier la documentation présente dans l'entrepôt.

Colectica Portal : Interface web qui permet la visualisation et l'exploration des données présentes dans l'entrepôt (*repository*).

Colectica Repository : Espace de stockage de la suite Colectica. Outre la conservation des données, il permet le versioning, la réutilisation, ainsi que le travail collaboratif.

Concordance : Être semblable, correspondre aux mêmes idées, tendre au même résultat.

DDI (Data Documentation Initiative) : Norme internationale permettant de décrire les données issues d'enquêtes et d'autres méthodes d'observation en sciences sociales, comportementales, économiques et de la santé. DDI peut documenter et gérer différentes étapes du cycle de vie des données de recherche, telles que la conception, la collecte, le traitement, la diffusion, la découverte et l'archivage.

Interopérabilité : Capacité d'un système informatique à fonctionner avec d'autres systèmes existants ou futurs. C'est la possibilité qu'ont des systèmes à fonctionner ensemble, à « communiquer » entre eux.

Métadonnées : Eléments descriptifs d'une ressource - ici les données d'enquête. Les métadonnées facilitent la recherche en donnant des informations utiles pour la compréhension et la réutilisation de ces ressources.

Niveau étude : Informations relatives à l'enquête (ex. titre, producteur(s), mode de collecte).

Niveau variable : Informations relatives au contenu de la base de données (ex. nom et label d'une variable, liste de code, valeurs manquantes)

XML (Extensible Markup Language) : Langage de structuration de données, utilisé notamment pour la gestion et l'échange d'informations sur Internet. XML (ou le XML) est un langage de description des données, plus puissant que HTML.

Légende



Point clé



Astuce



Avertissement

Procédure de documentation

1.1. Principes DDI-Lifecycle

Data Documentation Initiative est un standard de documentation international. Il permet de décrire les données à partir de nombreuses spécifications. Pour en savoir plus <https://ddialliance.org/>

La version *Lifecycle* de DDI a pour but de documenter et de gérer les données tout au long de leur cycle : de la conception de l'enquête jusqu'à la publication des données, leur analyse et au-delà. Il contient toutes les spécifications présentes dans le standard DDI-Codebook. Basé sur des schémas XML, DDI-Lifecycle est modulable et extensible, permettant notamment une réutilisation des métadonnées.

Dans une documentation DDI-L, chaque élément constitue une balise dotée d'un identifiant unique, permettant l'adoption du principe "Saisir une fois, réutiliser plusieurs fois" (DRY : *don't repeat yourself*). Ainsi, cette version convient parfaitement pour les enquêtes de type longitudinal, barométrique ou répété, car elle permet une documentation harmonisée entre les vagues de créer des concordances entre les différentes vagues. Elle répond au principe de métadonnées actives.

1.2. Inventaire et curation

L'inventaire des métadonnées consiste à identifier et documenter les éléments réutilisables, les liens, mais aussi les disparités entre les différentes enquêtes que l'on souhaite intégrer dans la base de métadonnées. Ce travail de recensement et d'analyse du potentiel de documentation en DDI-L, permet d'évaluer en amont la charge de travail à accomplir. Par exemple, si une enquête a des questions et des variables très similaires, le travail de documentation sera moindre.

La curation consiste à rassembler les métadonnées selon leur nature (catégories, questions, etc.) dans une base pour ne conserver qu'une seule et unique version de chacune d'entre elles pour éviter les doublons, et ainsi pouvoir importer les métadonnées uniques et réutilisables dans le logiciel.



Pour créer des métadonnées uniques et réutilisables, il faut leur attribuer un nom en respectant une règle de nommage définie et harmonisée.

1.3. Référentiel de métadonnées

Après avoir rassemblé, nettoyé et nommé les métadonnées pour les rendre réutilisables, il est recommandé de les importer au sein d'un référentiel en respectant une classification pour assurer une meilleure gestion à posteriori.

1.4. Connexion des éléments

Les éléments présents dans le référentiel peuvent ensuite être réutilisés et associés pour construire la documentation d'une ou de plusieurs enquêtes, ainsi que pour créer des concordances (liens) entre les éléments de la base de métadonnées.

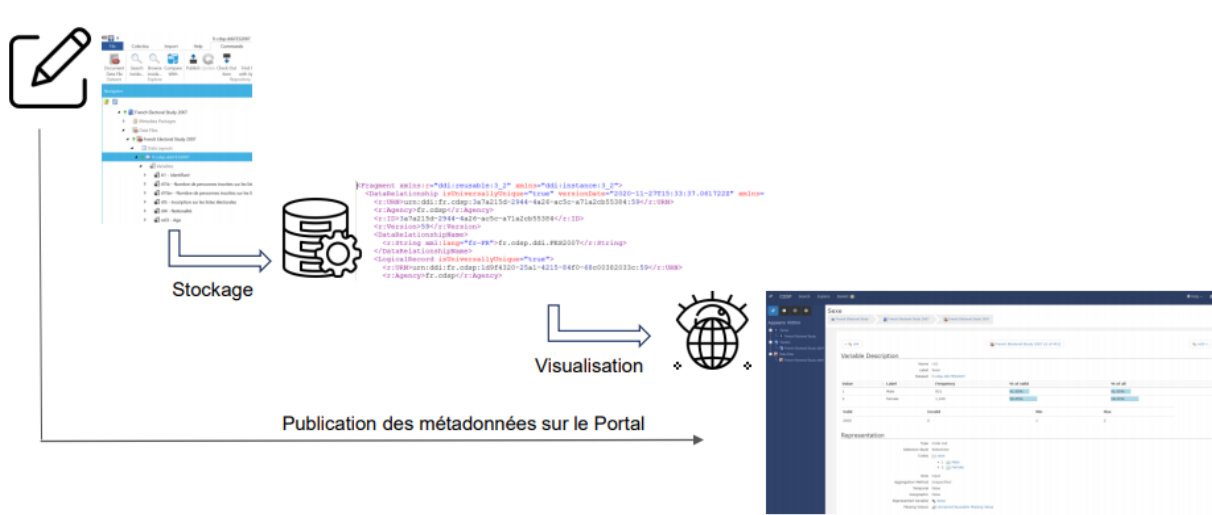
1.5. Publication

La publication permet la sauvegarde des métadonnées dans l'entrepôt, leur réutilisation à partir de plusieurs postes permettant ainsi le travail collaboratif, et leur visualisation sur le portail.

La suite logicielle Colectica

2.1. Les outils

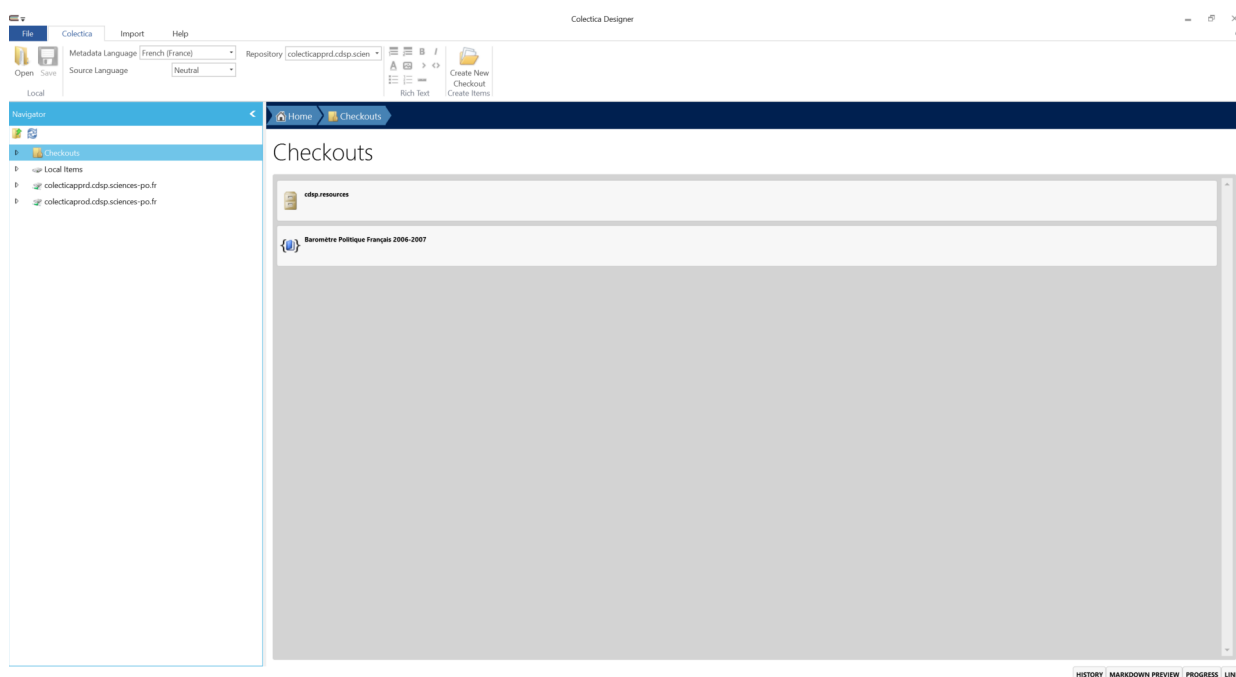
Designer, *Repository* et *Portal* sont les trois composants de la suite logicielle Colectica. L'éditeur de métadonnées (*Designer*) permet l'import et l'export de fichiers de documentation, d'instruments d'enquêtes et de bases de données dans plusieurs formats. Le portail (*Portal*) permet l'exploration du contenu de l'entrepôt de stockage (*Repository*) pouvant être édité, modifié et enrichi via l'application d'édition (*Designer*). Pour en savoir plus : <https://www.colectica.com/>



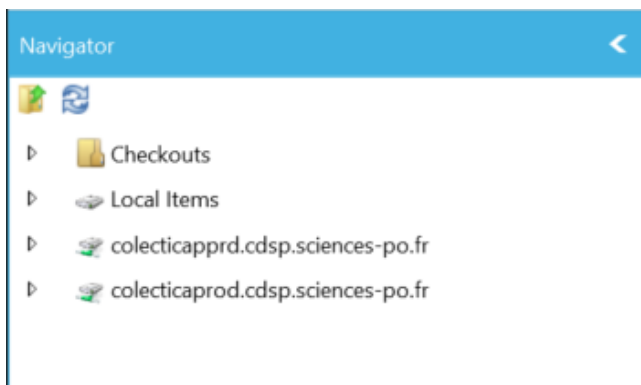
Source : Marie, Lucie. (2020, December 1). Using Colectica to document socio-political surveys in DDI-Lifecycle: feedback and benefits. 12th Annual European DDI User Conference (EDDI20). Zenodo. <https://doi.org/10.5281/zenodo.4300216>

2.2. L'interface d'édition

2.2.1. Les environnements de travail



Voici un aperçu de la page d'accueil de l'outil d'édition.



Dans le ruban de navigation latéral gauche, 4 espaces :

- **checkouts** : espace d'édition dans lequel les métadonnées peuvent être créées, importées, modifiées et publiées.
- **local items** : espace de stockage interne.
- **colecticapprd.cdsp.sciences-po.fr** : entrepôt de pré-production (version test du stockage), son contenu est disponible via <https://colectica-portal-pprd.cdsp.sciences-po.fr/>
- **colecticaprod.cdsp.sciences-po.fr** : entrepôt de production (version finale du stockage), son contenu est disponible publiquement via <https://colectica-portal.cdsp.sciences-po.fr/>

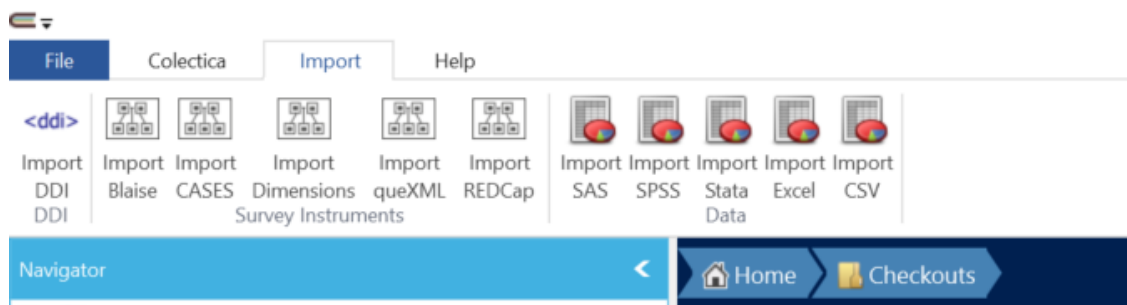


Pour accéder au portail de pré-production, il faut être connecté au VPN de Sciences Po.

2.2.2. Fonctionnalités principales

Import

Colectica Designer permet d'importer des fichiers de différentes natures (documentation, instrument d'enquête, base de données) et de formats divers (cf. capture).



Réutilisation

L'intérêt de cette application est qu'elle permet de respecter le principe de réutilisation de DDI-L grâce à des fonctionnalités de recherche dans l'entrepôt disponibles à différentes étapes de la documentation. Le document "[Metadatavariety](#)" est un répertoire des champs disponibles pour la documentation.



Lorsque l'on ré-utilise un élément, il faut penser à le mettre à jour dans son emplacement initial avec la fonction « Update » pour inclure ses nouvelles interactions dans l'entrepôt.

Référentiel de métadonnées

Cet outil assure la bonne gestion de l'entrepôt puisqu'il permet d'organiser et de stocker des objets. Par bonne gestion, nous pensons à l'harmonisation des objets, la réduction du stockage ou encore du temps de documentation. Un paquet de métadonnées (*metadatapackage*) est nécessaire pour construire cet environnement. Il permet d'organiser et de stocker des métadonnées uniques et réutilisables pour la documentation des enquêtes et des variables.

3.1. Eléments

Voici les principaux éléments utilisés pour la documentation d'une variable issue d'une enquête par questionnaire accompagnés de leur définition (source : DDI Data Documentation <https://ddialliance.github.io/ddimodel-web/f5c0935ed38da008d1db27ed54750b3a46696f9d/#ddi-data-documentation-3-3>).

Catégories : *“A description of a particular category or response. OECD Glossary of Statistical Terms: Generic term for items at any level within a classification, typically tabulation categories, sections, subsections, divisions, subdivisions, groups, subgroups, classes and subclasses. In addition to the standard name, label, and description, a category may contain a reference to a defining concept, provide information on how the category was generated (membership defined), and indicate if it is the description of a type of missing value.”*

<https://ddialliance.github.io/ddimodel-web/f5c0935ed38da008d1db27ed54750b3a46696f9d/item-types/Category/?highlight=categorie>

Questions

<https://ddialliance.github.io/ddimodel-web/f5c0935ed38da008d1db27ed54750b3a46696f9d/item-types/Question/>

- **Question simple (Question Item)** : *“Structure a single Question which may contain one or more response domains (i.e., a list of valid category responses where if “Other” is indicated a text response can be used to specify the intent of “Other”). The structure provides detail on the intent of the question, the text of the question, the valid response options and the number of allowed responses, references to external aids and instructions, and an estimation of the time needed to respond to the question. Note that the QuestionItem is a reusable format for use in any number of applied uses. External aids, instructions, response sequencing etc. should contain information consistent with the general use of the QuestionItem. Additional materials and information can be added within the QuestionConstruct which is the applied use of a question.”*

<https://ddialliance.github.io/ddimodel-web/f5c0935ed38da008d1db27ed54750b3a46696f9d/item-types/QuestionItem/>

- **Question choix multiple (Question Grid)** : *“Structures the QuestionGrid as an NCube-like structure providing dimension information, labeling options, and response domains attached to one or more cells within the grid. Provides the intent of the QuestionGrid, input and output parameters for the grid, the question text for the grid, details on the dimensions, allowed responses, and additional cell contents of the grid, references to external aids and instructions, and an estimate of the time required to complete the grid. Note that the QuestionGrid is a reusable format for use in any number of applied uses. External aids, instructions, response sequencing etc. should contain information consistent with the general*

use of the QuestionGrid. Additional materials and information can be added within the QuestionConstruct which is the applied use of a question.”

<https://ddialliance.github.io/ddimodel-web/f5c0935ed38da008d1db27ed54750b3a46696f9d/item-types/QuestionGrid/>

- **Question Block** : “A QuestionBlock is a specific structure used in educational and other types of testing where an object (Stimulus Material) is provided and a set of questions are asked regarding the object. The QuestionBlock generally has related QuestionBlocks that measure similar skills or aptitudes and is used randomly within a set of questionnaires to create multiple versions of a single questionnaire that can be used with large groups for testing purposes. Assembly of the QuestionBlocks into a questionnaire may be the result of selection based on an experimental design model. It contains information on what the QuestionBlock is intended to measure, input and output parameters for the QuestionBlock, a description of the stimulus material and the questions related to it, instructions on sequencing and number of allowed responses, references to external aids and instructions, and an estimate of the time needed to complete the question. Note that the QuestionBlock is a reusable format for use in any number of applied uses. External aids, instructions, response sequencing etc. should contain information consistent with the general use of the QuestionBlock (QuestionItems and QuestionGrids will contain information specific to the individual question). Additional materials and information can be added within the QuestionConstruct which is the applied use of a question.”

<https://ddialliance.github.io/ddimodel-web/f5c0935ed38da008d1db27ed54750b3a46696f9d/item-types/QuestionBlock/>

Liste de codes (codelist) : “A structure used to associate a list of code values to specified categories. May be flat or hierarchical. This is a maintainable object. In addition to the standard name, label, and description the CodeList specifies a recommended data type, the hierarchical nature of the CodeList, level descriptions, individual code values and associated category, and whether the CodeList contents are used to represent system specific missing values. May include another CodeList by reference. If including other CodeLists make sure there are no code conflicts in the overall content. A code value must be unique within the CodeList. May also include a reference to a default CategoryScheme.”

<https://ddialliance.github.io/ddimodel-web/f5c0935ed38da008d1db27ed54750b3a46696f9d/item-types/CodeList/?highlight=codelist>

Variable représentative (represented variable) : “Describes a RepresentedVariable contained in the RepresentedVariableScheme. In addition to the standard name, label, and description a RepresentedVariable contains a reference to the Concept and Universe (or ConceptualVariable) as well as the representation of the RepresentedVariable. Representation may be provided in-line or by reference to a managed representation. RepresentedVariables are the core reusable parts of a Variable. RepresentedVariable maps to the GSIM Represented Variable. In addition to the standard name, label, and description.”

<https://ddialliance.github.io/ddimodel-web/f5c0935ed38da008d1db27ed54750b3a46696f9d/item-types/RepresentedVariable/>



Colectica se base sur les catégories de réponse pour associer les champs, donc les valeurs des codes peuvent être différentes.

Variable conceptuelle (conceptual variable) : “Describes a *ConceptualVariable* which provides the link between a concept to a specific universe (object) that defines this as a *ConceptualVariable*. In addition to the standard name, label, and description, it provides the two primary components of a *ConceptualVariable* by referencing a concept and its related universe. Note that the concept referenced may itself contain sub-concepts and/or references to similar concepts. This maps to the *GSIM ConceptualVariable* and has a basis in the *ISO/IEC 11179 RepresentedVariableConcept*.”

<https://ddialliance.github.io/ddimodel-web/f5c0935ed38da008d1db27ed54750b3a46696f9d/item-types/ConceptualVariable/>

3.2. Organisation

Afin de faciliter la réutilisation des objets du référentiel, les catégories (*categories*), variables conceptuelles (*conceptual variables*), variables représentées (*represented variables*) et les listes de codes (*codelists*) ont été organisés à partir d’une classification dans des ensembles (*sets*) et des sous-ensembles (*groups*) de métadonnées. Cette classification a été créée pour répondre au besoin des enquêtes et répliquée pour les ensembles et sous-ensembles de métadonnées, mais il est également possible de réutiliser des classifications existantes.

3.2.1. Classification des domaines

- ▲ Represented Variable Sets
 - representedvariables:activity - activity
 - representedvariables:binary - binary
 - representedvariables:date - date
 - representedvariables:frequency - frequency
 - representedvariables:media - media
 - representedvariables:missingValue - missingValue
 - representedvariables:numeric - numeric
 - representedvariables:opinion - opinion
 - representedvariables:other - other
 - representedvariables:political - political
 - representedvariables:resources - resources
 - representedvariables:respondent - respondent
 - ↑ representedvariables:sociodemo - sociodemo
 - representedvariables:territory - territory

3.2.2. Classification des sous-domaines

- ▲ ↑ ↻ representedvariables:political - political
 - ▲ ↻ Represented Variable Groups
 - ▷ ↻ representedvariables:political.issue - issue
 - ▷ ↻ representedvariables:political.opinion - opinion
 - ▷ ↻ representedvariables:political.organization - organization
 - ▷ ↑ ↻ representedvariables:political.participation - participation
 - ▷ ↻ representedvariables:political.party - party
 - ▷ ↻ representedvariables:political.personality - personality
 - ▷ ↻ representedvariables:political.position - position

Les questions sont regroupées par séries et organisées par type : *simpleQuestions*, *multipleQuestions* et *questionBlock*. Le choix de les distinguer par séries est fait pour conserver le nom (*name*) des questions qui varie peu dans les questionnaires d'une vague à l'autre. Cependant, le potentiel de réutilisation des questions est très faible en dehors d'une série.



Les questions sont regroupées par séries et organisées par type : *simpleQuestions*, *multipleQuestions* et *questionBlock*. Le choix de les distinguer par séries est fait pour conserver le nom (*name*) des questions qui varie peu dans les questionnaires d'une vague à l'autre. Cependant, le potentiel de réutilisation des questions est très faible en dehors d'une série.

3.2.3. Réutilisation d'une ontologie existante

Des organismes ou des projets ont mis au point des classifications qu'il est possible de réutiliser. L'avantage de s'appuyer sur un cadre commun utilisé par d'autres projets est de garantir l'harmonie des métadonnées et peut permettre une éventuelle fusion à posteriori entre différents jeux de données documentées de façon similaire. Cette option est conseillée notamment dans des projets de post-harmonisation à grande échelle. Par exemple le CESSDA (*Consortium of European Social Science Data Archives*) propose une liste thématique de vocabulaires contrôlés (<https://vocabularies.CESSDA.eu/vocabulary/TopicClassification>) ou encore le projet *True European Voter* qui a mis au point une arborescence conceptuelle pour la post-harmonisation des enquêtes électorales.

3.3. Règles de nommage



Deux types d'éléments peuvent se distinguer dans un référentiel. Les éléments dit "uniques", c'est-à-dire qui sont pas amenés à avoir plusieurs versions, par exemple une catégorie constituée d'un nom propre. D'autres peuvent connaître un nombre plus ou moins important de versions. Ce type d'éléments "versionnables" sont par exemple des variables représentatives lorsqu'une même variable conceptuelle a été mesurée à l'aide de listes de codes différentes.

Pour tous les éléments (uniques ou versionnables), le libellé de l'élément est préfixé par le domaine, et le sous-domaine et séparés par un point. Puis d'élément est introduit à la suite d'un double-point ":" et enseigné en minuscule sans caractère spécial. Pour les éléments avec plusieurs versions, il est recommandé d'ajouter un numérotation unique basé sur une séquence chronologique continue et

sans rupture, soit une suite ininterrompue afin d'éviter la possibilité de doublons et de suivre la chronologie des ajouts.

Modèle : "domaine.sousdomaine:libelle-de-la-categorie-sans-majuscule-ni-caractere-special"

Exemple : "political.personality:jose-bove"



Il est essentiel de mettre en place et de respecter une règle de nommage pour maintenir une cohérence au sein du référentiel. Un élément avec un nom unique facilite son identification et sa réutilisation.

3.4. Ajout de métadonnées

Trois commandes permettent d'ajouter des objets au référentiel de métadonnées :

- créer un élément nouveau via la commande, option 1 via l'icône +
- réutiliser un élément existant d'un autre entrepôt, option 2 via l'icône loupe
- importer des bases de métadonnées au format tableur, option 3 via la commande d'import Wizard



Vérifier qu'un élément n'est pas déjà présent dans le référentiel avant de le recréer. Cela risquerait de générer des doublons.

3.4.1. Option 1 : icône +

Home > Checkouts > cdsp.resources > categories:political - political > categories:political.personality...

categories:political.personality - personality

Members Description

Description

Group Type

Categories Filter

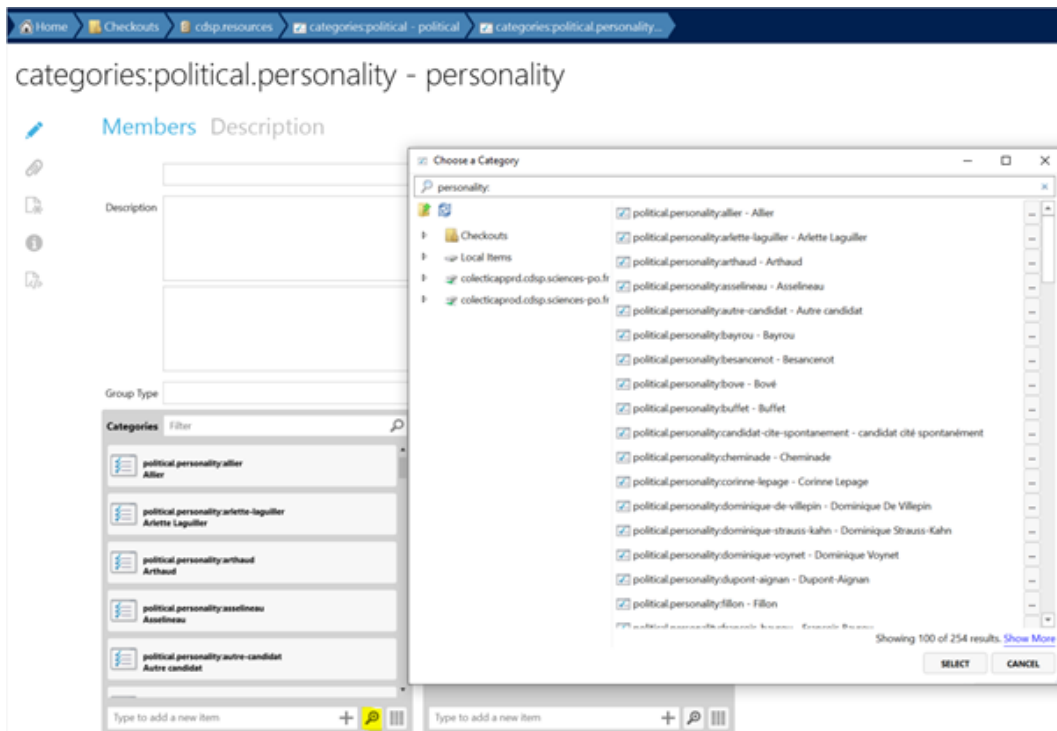
- political.personality:alier
Alier
- political.personality:arlette-laguiller
Arlette Laguiller
- political.personality:arthaud
Arthaud
- political.personality:asselineau
Asselineau
- political.personality:autre-candidat
Autre candidat

Type to add a new item + 🔍 ☰

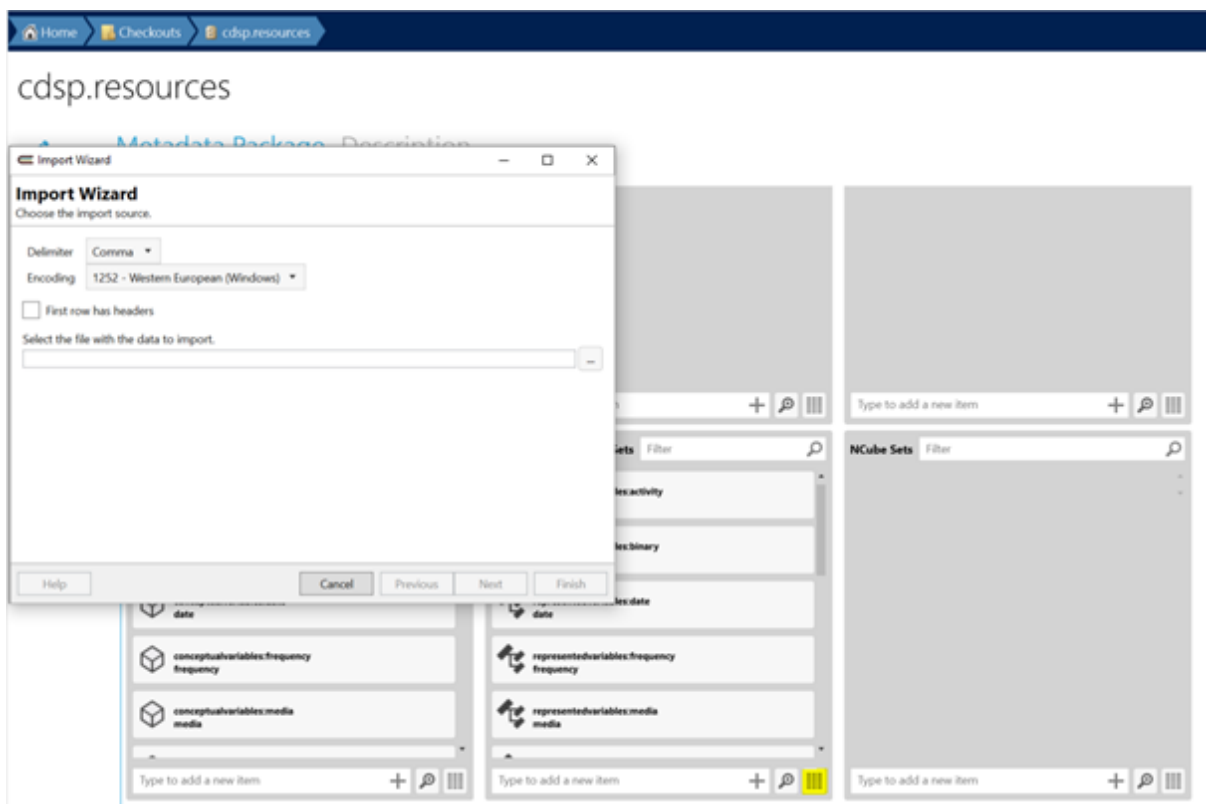
Category Groups Filter

Type to add a new item + 🔍 ☰

3.4.2. Option 2 : icône loupe



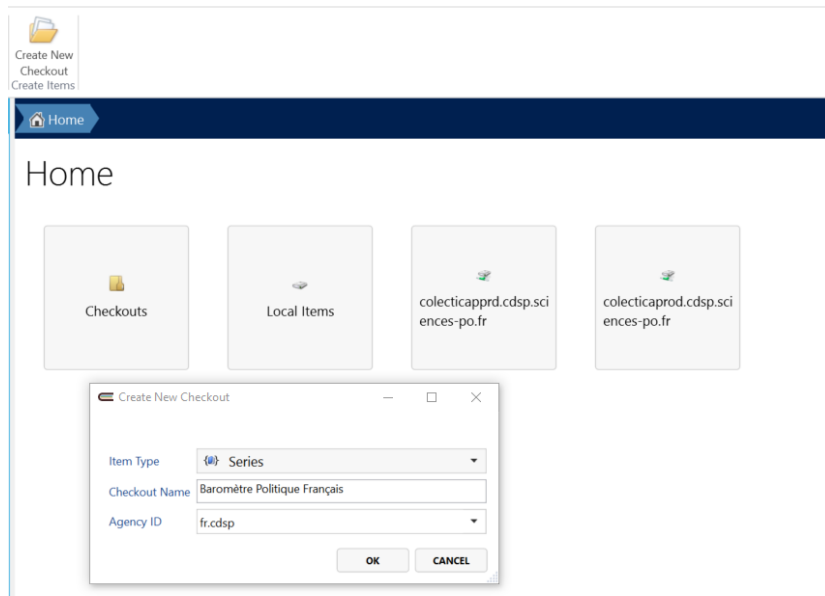
3.4.3. Option 3 : import wizard



Documentation d'un jeu de données

4.1. Série

Créer une série dans l'espace de travail "checkouts".



4.2. Etude

Ajouter une ou plusieurs études dans la série.



4.2. Questionnaire

questionnaire.bpf.r1 - Questionnaire - Baromètre Politique Français - Vague 1

Instrument Summary Questions Flowchart Details

A TOUS

Description

Q0 **Quelle est votre nationalité ?**
 1 sociodemo.nationality:française - Française
 2 territory.international:etrangere - Étrangère

HORS CIBLE
 hors cible si Q0==2
 0 items

Q0b **Etes-vous inscrit(e) sur les listes électorales pour pouvoir voter ?**
 1 respondent.declaration:oui-dans-cette-commune - Oui, dans cette commune
 2 respondent.declaration:oui-dans-une-autre-commune - Oui, dans une autre commune
 3 binary.text:non - Non
 4 opinion.scale:ne-sait-pas - Ne sait pas

HORS CIBLE
 hors cible si Q0b==3 or 4
 0 items

RS1 **Sexe**
 1 sociodemo.sex:homme - Homme
 2 sociodemo.sex:femme - Femme

RS2 **En quelle année êtes-vous né(e) ?**
 Numeric

+ Add an item

Il est possible de réutiliser des questions simples, multiples, des blocs de questions, des séquences ou des actions. Pour cela, cliquez sur la fonction « Add an item », puis « Reuse existing items ».

The screenshot shows the Collectica Designer interface. A 'Choose a Question' dialog box is open, displaying a list of questions with their IDs and descriptions. The questions listed include:

- Q17 - Pendant la campagne électorale pour les dernières législatives, diriez-vous qu'il y a eu entre les partis politique...
- E5A - En pensant à [PROG : AFFICHER REPONSE DONNEE EN F1] quel était selon vous le candidat à l'élection prés...
- Q21a - Si vous aviez voté, parmi les candidats suivants, pour qui auriez-vous voté ?
- Q21b - En plus de votre revenu régulier, est-ce que vous avez ...
- Q06b - Etes-vous très satisfait(e), assez satisfait(e), peu satisfait(e) ou pas du tout satisfait(e) du nouveau gouvernem...
- E6A - Et en pensant à [PROG : AFFICHER REPONSE DONNEE EN E2] quel était selon vous le candidat à l'élection pr...
- Q21 - Beaucoup d'électeurs n'ont pas voté au premier tour des élections législatives. Vous-même, pouvez-vous me...
- Q14 - Parmi les questions évoquées à l'occasion des dernières élections législatives, laquelle est la plus importa...
- Q21c - Y a-t-il un ou plusieurs autres candidats pour lequel ou lesquels vous avez sérieusement envisagé de voter ?
- Q23 - Beaucoup d'électeurs n'ont pas voté au premier tour de l'élection présidentielle. Vous-même, pouvez-vous...
- Q28a - Selon vous, dans quelle mesure la question du chômage est-elle une question importante pour la France ?...
- Q30 - Et en pensant à [PROG : METTRE LA REPONSE DONNEE EN Q28 = SECOND PROBLEME LE PLUS IMPORTANT...]
- Q22a - Si vous étiez allé voter, parmi les candidats suivants, pour qui auriez-vous voté ?
- Q23b - Parmi les candidats suivants, pour qui avez-vous voté ?
- Q25e - L'Etat aide financièrement les partis politiques.
- Q20 - D'une manière générale, y a-t-il un parti ou un mouvement politique dont vous sentez proche ?
- Q21d - Lequel ? Qui d'autre ?

The background shows the instrument editor with a tree view on the left and a main workspace displaying the 'A TOUS' instrument structure.

4.3. Jeu de données

4.3.1. Import du dataset

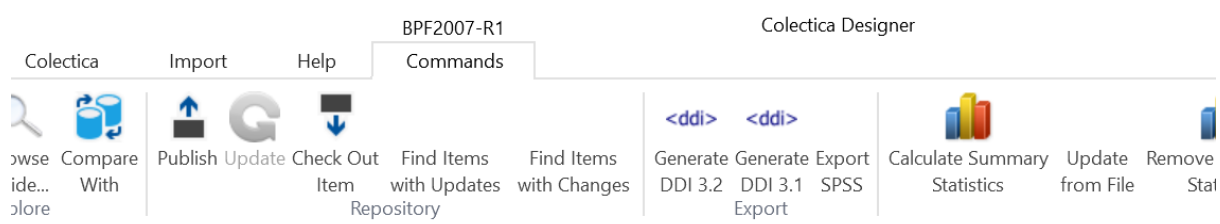
Métadonnées récupérées : les objets variables du jeu de données, ainsi que les libellés



Si votre jeu de données contient déjà des métadonnées, il faudra effectuer un nettoyage afin de supprimer tout risque de doublons.

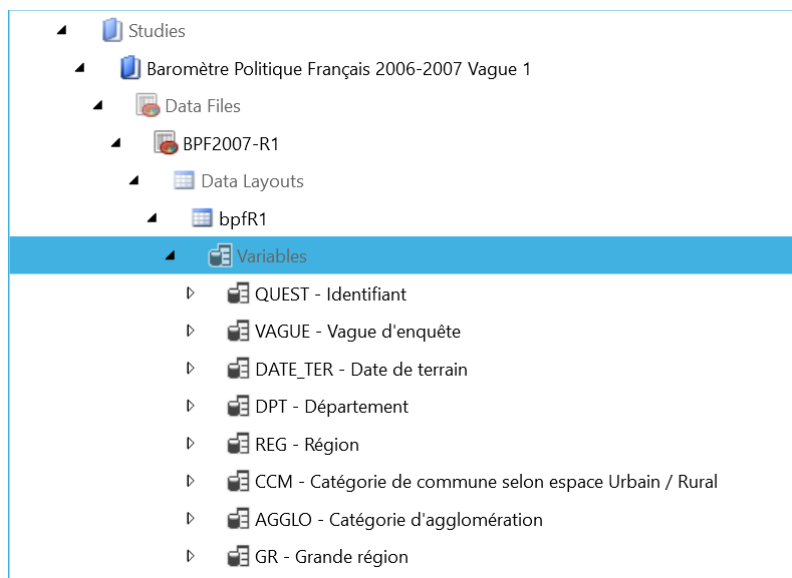
La description du jeu de données est disponible dans la partie “*Data Layouts*”

4.3.2. Calcul des variables



4.3.3. Localisation

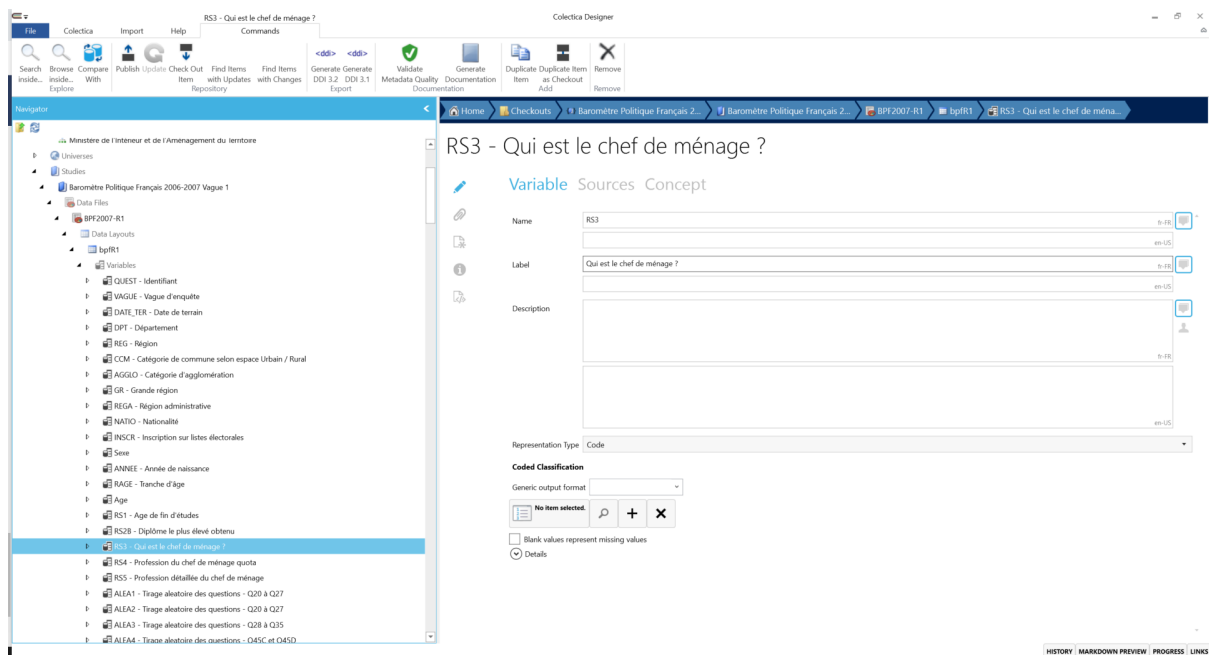
Il va s’agir ici de réutiliser et d’associer les éléments présents dans le référentiel afin de reconstruire la documentation des variables de la base de données (Study > dataset > datalayout > variables).




4.4. Variable

4.4.1. Etape 1 : Identification


Chaque variable doit obligatoirement être documenter d’un nom (*Name*), d’un libellé (*Label*), ainsi que le type de représentation (*Code*, *Date/Time*, *Numeric*, *Text*).



 Le nom de la variable doit correspondre à celui de la base de données.

4.4.2. Etape 2 : Liste de codes

Pour les variables de type catégorielle, il est possible de créer une liste de code ou d'associer une liste de codes.

 Si la variable est issue d'une question, réutiliser la même liste de codes que pour la question.

RS3 - Qui est le chef de ménage ?

Variable Sources Concept

Name: RS3

Label: Qui est le chef de ménage ?

Description:

Representation Type: Code

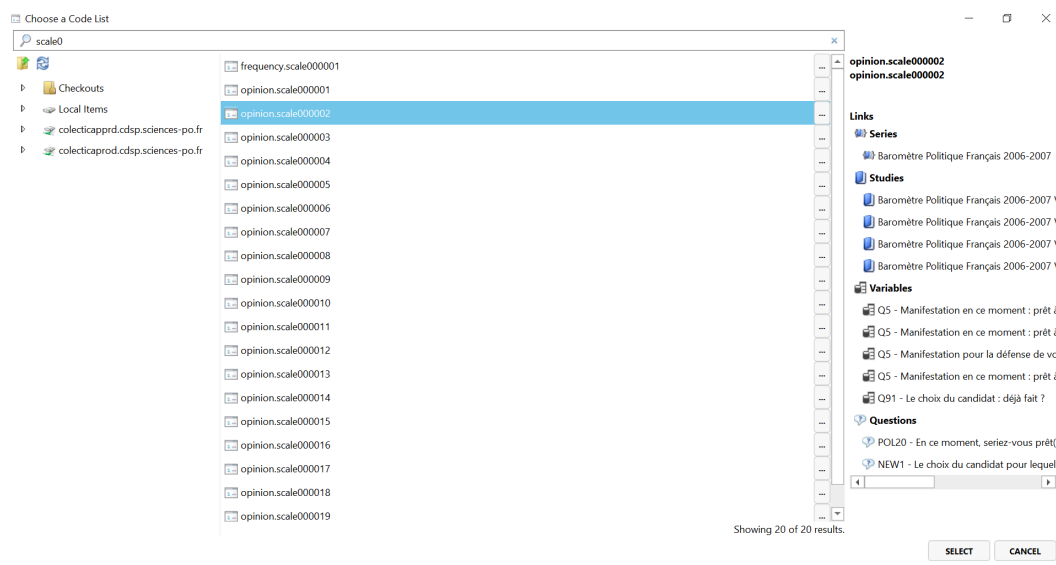
Coded Classification

Generic output format: [dropdown]

No item selected. [search] [add] [remove]

Blank values represent missing values

[Details](#)



A droite de la fenêtre se trouvent les éléments associés à cette liste de code.

4.4.3. Etape 3 : Valeurs manquantes

Une fois la liste de codes attribuée, le champ des valeurs manquantes (*Missing Values*) apparaît. Cette fonction de gestion des valeurs manquantes permet d'affecter des listes de codes contenant exclusivement ce type spécifique de catégorie (non concerné, n'a pas répondu, question non posée, etc.)



De nouvelles catégories peuvent être reconstruites suite à la collecte pour identifier le type de valeurs manquantes (ex. non concerné, n'a pas répondu, etc.)

Coded Classification

Generic output format

sociodemo.householdComposition000001

Blank values represent missing values

Missing Values No item selected.

Role

Analysis Unit

4.4.4. Etape 4 : Sources

La variable peut être construite à partir d'un questionnaire ou d'autres variables (ex. recodage). L'onglet "Sources" permet d'indiquer les éléments qui ont permis la construction de cette variable. Si par exemple, la variable est produite à partir d'une question, la fonction "loupe" permet de retrouver l'élément dans le repository, ce qui crée les connexions dans la base de métadonnées.

RS3 - Qui est le chef de ménage ?

Variable Sources Concept

Source Questions Filter

RS0
Qui est le chef de ménage ?

Type to add a new item + 🔍 📄

Source Question Grids Filter

Type to add a new item + 🔍 📄

Source Question Blocks Filter

Type to add a new item + 🔍 📄

Source Variables Filter

Type to add a new item + 🔍 📄

4.4.5. Etape 5 : Concordance

Cette étape permet d'harmoniser et de connecter des variables entre les différents jeux de données. L'harmonisation s'effectue via les variables représentées, c'est-à-dire qui ont le même type de représentation. La connexion s'effectue via les variables conceptuelles, c'est-à-dire qui mesure le même concept mais pas avec les mêmes représentations. Les variables conceptuelles peuvent être rattachées à des concepts.

Pour en savoir plus

<https://docs.colectica.com/designer/manage-content/data/harmonize-variables/?highlight=represented#overview>

Q13 - Jugement sur action du gouvernement

Variable Sources Concept

Represented Variable JGMTGOUVDDV000001 - Jugement sur action du g... 🔍 + ✕

Conceptual Variable JGMTGOUVDDV - Jugement sur action du gouverne... 🔍 + ✕

Concept No item selected. 🔍 + ✕

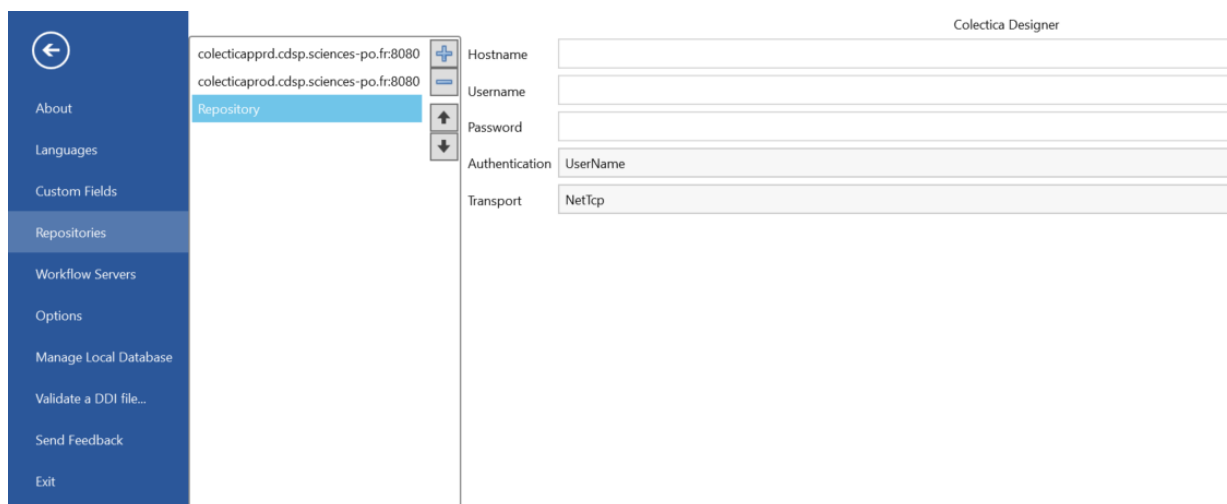
Universes Filter

Type to add a new item + 🔍 📄

Stockage et diffusion

6.1. Association des outils


Afin d'associer (ou supprimer) un entrepôt à l'outil d'édition, rendez-vous dans le menu "File" ; "Repositories", puis indiquez les informations nécessaires à la connexion.



6.2. Publication sur le portail en ligne

Pour publier les métadonnées sur le portail, rendez-vous dans le menu "Commands" et cliquez sur "Publish".



 Les flèches vertes dans le bandeau de navigation signalent que des modifications ont eu lieu depuis la dernière publication.

6.3. Export de la documentation

6.3.1. Fichiers DDI

Pour exporter une documentation DDI au format XML, rendez-vous dans le menu “*Commands*” de *Colectica Designer* et cliquez sur “*Generate DDI*”.



6.3.2. Livre de codes

Pour exporter un livre de codes au format PDF ou XML, rendez-vous dans le menu “*Commands*” et cliquez sur “*Generate Documentation*”.

