



INFORMATION NUMÉRIQUE

Enjeux et Pratiques

Licence 3 / UE502 - Accès expert à l'information - 2024-2025

OpenRefine 3.9.1 : Présentation et fonctionnalités principales



philippe.houdry@inist.fr
lucile.bourguignon@inist.fr
valerie.bonvallot@inist.fr



Plan de l'atelier OpenRefine

- Présentation d'OpenRefine
- Installation, lancement, désinstallation
- Import des données, gestion des projets
- Espace de travail, exploration et nettoyage des données
- Transformation des données
- Export des données et des traitements
- Introduction à quelques fonctionnalités avancées



Présentation d'OpenRefine

Fonctionnalités principales

- Exploration
- Correction / Harmonisation
- Transformation
- Restructuration
- Enrichissement

Les données sont dans des cellules comme dans un tableur (projet tabulé) mais OpenRefine n'est pas un tableur de type Excel et n'est pas fait pour réaliser des calculs. Il ne réalise aucun graphique.

OpenRefine est utilisé surtout pour traiter des données textuelles. On peut cependant aussi y utiliser des données de type nombre, date et booléen.



Présentation d'OpenRefine

Avantages d'OpenRefine	Inconvénients
Fonctions absentes des tableurs traditionnels	Langage spécifique
Interface graphique	Modèle de données contraint (tabulaire)
Totalement libre et gratuit	Pas adapté au "big data"
Installation PC, Linux, Mac	Ne traite pas de données en flux
Enregistrement des opérations réalisés	Gourmand en mémoire
Maîtrise des données	Pas de fonction collaborative
Large communauté d'utilisateurs	Formats d'imports limités
Extensibilité (plug-ins)	Ergonomie dépassée

Performance : Avec une machine virtuelle de 2 Go, on traite généralement sans problème au moins 100 000 lignes avec de nombreuses colonnes.



Présentation d'OpenRefine

Versions, licence, mode de développement et historique

Dernière version stable : 3.9.1 du 14/03/2025.

OpenRefine est open source et gratuit.

Le développement est communautaire (quelques développeurs principaux et des contributeurs occasionnels) et hébergé sur GitHub. Il est soutenu par l'association [Code for Science & Society](#).

2009 : Logiciel Gridworks par Metaweb (open source)

2010 : Rachat de Metaweb par Google -> Google Refine

2012 : Suite au retrait de Google, développement communautaire
-> OpenRefine

2017 : Financement de la fondation Google News

2019 : Financement de la fondation Chan Zuckerberg Initiative



Présentation d'OpenRefine

Sites de référence et documentations

Site officiel (en Anglais) : <https://openrefine.org/>

Documentation officielle 3.7 : <https://docs.openrefine.org/>

Cours de M. Saby 2020 (pour 3.4) :

<https://msaby.gitlab.io/tutoriel-openrefine/index.html>

Mon cours 2022 (pour 3.5.2) :

<https://www.ahp-numerique.fr/2022/04/26/atelier-openrefine-2022/>

Tes premiers pas avec OpenRefine (14 courtes vidéos pour 3.6.2) :

<https://www.canal-u.tv/chaines/rbdd/tes-premiers-pas-avec-openrefine-0>

Forum OpenRefine (Community, Support,...) : <https://forum.openrefine.org/>

Utile pour savoir sur quoi travaillent les développeurs et trouver de l'aide (en Anglais).

État des traductions (mode collaboratif) :

<https://hosted.weblate.org/projects/openrefine/#languages>

La traduction en français en est à 72 %.



Installation d'OpenRefine

Principes généraux

Monoposte et multiplateforme (Windows, Mac, Linux).
Interface utilisable dans un navigateur internet, Firefox ou Chrome conseillés.

Données, projets et logiciel sur le PC. Tout est en local.

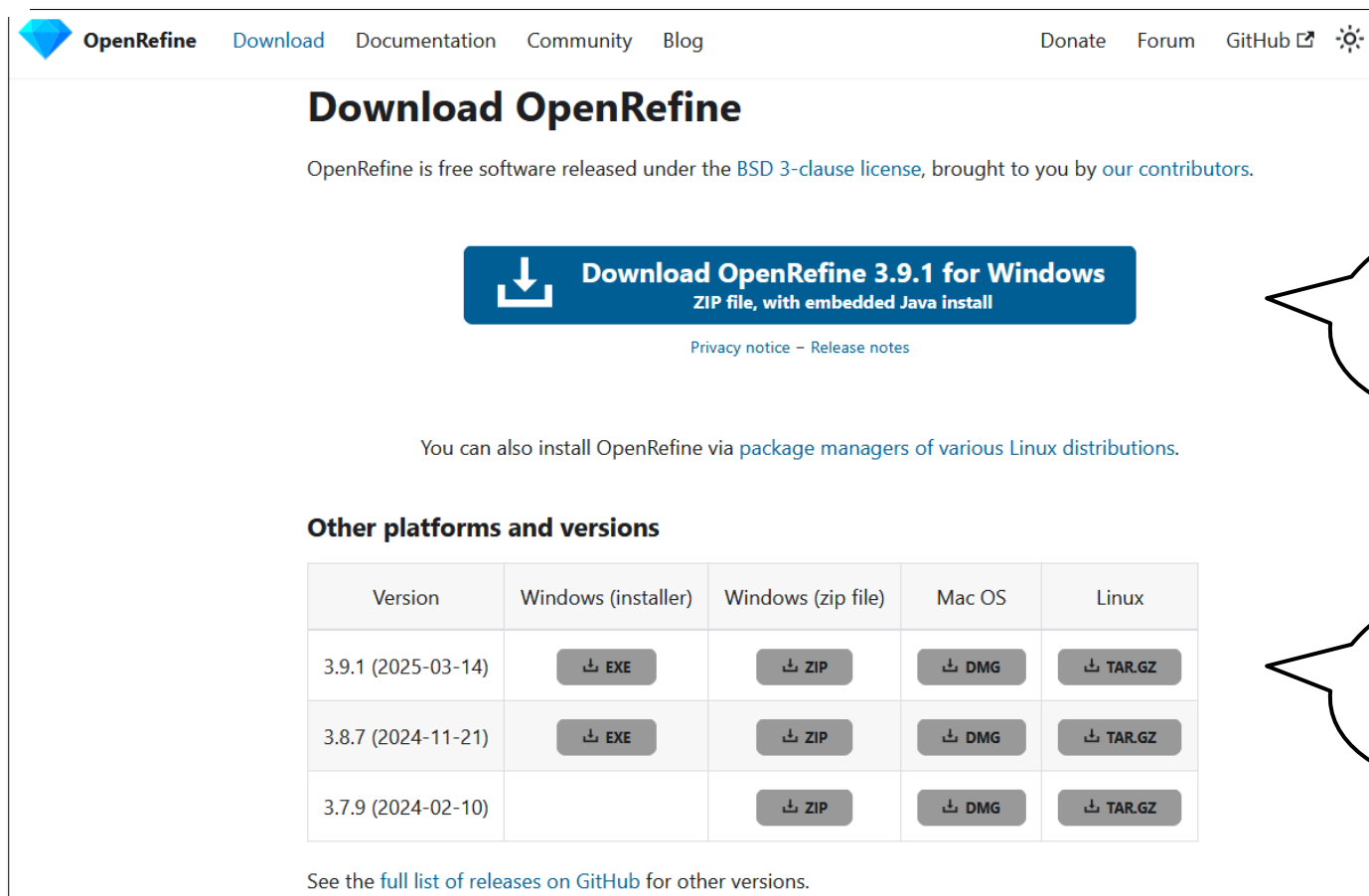
Nombreuses extensions utilisables (attention aux versions compatibles).
Automatisation des traitements possible par API.

OpenRefine est écrit en Java, il nécessite donc un Java installé sur poste.
Il est très conseillé d'avoir un navigateur et Java en 64 bits.

Taille de la machine virtuelle modifiable dans les fichiers de configuration .ini (voir documentation) : ajout de mémoire.



Installation d'OpenRefine



The screenshot shows the OpenRefine website's download page. At the top, there's a navigation bar with links to Download, Documentation, Community, Blog, Donate, Forum, and GitHub. The main heading is "Download OpenRefine". Below it, a text line states: "OpenRefine is free software released under the BSD 3-clause license, brought to you by our contributors." A large blue button with a download icon and the text "Download OpenRefine 3.9.1 for Windows" is prominent, with a subtext "ZIP file, with embedded Java install". Below the button are links for "Privacy notice" and "Release notes". Further down, a text line says: "You can also install OpenRefine via package managers of various Linux distributions." A section titled "Other platforms and versions" contains a table with download links for different operating systems and versions.

Version	Windows (installer)	Windows (zip file)	Mac OS	Linux
3.9.1 (2025-03-14)	EXE	ZIP	DMG	TAR.GZ
3.8.7 (2024-11-21)	EXE	ZIP	DMG	TAR.GZ
3.7.9 (2024-02-10)		ZIP	DMG	TAR.GZ

See the [full list of releases on GitHub](#) for other versions.

Prendre JAVA
embarqué
(embedded)

Raccourci sur le
bureau pour
openrefine.exe

<https://openrefine.org/download.html>



Lancement d'OpenRefine

OpenRefine Un outil puissant pour travailler avec des données désordonnées.

Créer un projet

Ouvrir un projet

Importer un projet

Langues

Créer un projet en important des données. Quelles sortes de données puis-je importer ?

Les documents de type TSV, CSV, *SV, Excel (.xls et .xlsx), JSON, XML, RDF en XML, OpenDocument (.ods) et Google Data sont reconnus nativement. D'autres formats peuvent être ajoutés via des extensions OpenRefine.

Récupérer les données à partir de

Cet ordinateur

Adresses web (URLs)

Presse-papier

Base de données

Google Data

Chercher un ou plusieurs fichiers à charger :

Parcourir...

Aucun fichier sélectionné.

Suivant »

openrefine.exe 3.7.1 - Raccourci

```
09:18:41.854 [ refine_server] Initializing context: '/' from 'C:\Users\houdry\Documents\openrefine-3.7.1\webapp' (38ms)
09:18:43.006 [ refine] Starting OpenRefine 3.7.1 [ecf4ff8]... (1152ms)
09:18:43.007 [ refine] initializing FileProjectManager with dir (1ms)
09:18:43.007 [ refine] C:\Users\houdry\AppData\Roaming\OpenRefine (0ms)
09:18:47.778 [ refine] POST /command/core/load-language (4771ms)
09:18:47.806 [ refine] GET /command/core/get-preference (28ms)
09:18:47.824 [ refine] POST /command/core/load-language (18ms)
09:18:47.831 [ refine] POST /command/core/load-language (7ms)
09:18:47.831 [ command] Language file [pc-axis:translation-fr.json] not found (0ms)
09:18:47.836 [ refine] POST /command/core/load-language (5ms)
09:18:47.907 [ refine] GET /command/core/get-importing-configuration (71ms)
09:18:47.921 [ refine] GET /command/core/get-all-project-tags (14ms)
09:18:47.981 [ refine] GET /command/database/saved-connection (60ms)
09:18:48.006 [ refine] GET /command/core/get-all-project-metadata (25ms)
09:18:48.071 [ refine] GET /command/core/get-csrf-token (65ms)
09:18:48.082 [ refine] GET /command/core/get-languages (11ms)
09:18:48.136 [ refine] GET /command/core/get-version (54ms)
09:20:13.208 [ refine] GET /command/core/get-csrf-token (85072ms)
09:20:13.216 [ refine] POST /command/core/create-importing-job (8ms)
09:20:13.234 [ refine] POST /command/core/importing-controller (18ms)
09:20:14.251 [ refine] POST /command/core/get-importing-job-status (1017ms)
09:20:14.314 [ refine] GET /command/core/get-all-project-tags (63ms)
09:20:14.338 [ refine] GET /command/core/get-csrf-token (24ms)
09:20:14.344 [ refine] POST /command/core/importing-controller (6ms)
09:20:14.401 [ refine] GET /command/core/get-csrf-token (57ms)
09:20:14.406 [ refine] POST /command/core/importing-controller (5ms)
09:20:14.497 [ refine] POST /command/core/get-models (91ms)
09:20:14.523 [ refine] POST /command/core/get-rows (26ms)
```

Version 3.7.1 [ecf4ff8]

Préférences

Aide

À propos

Pour vraiment quitter, fermer cette console.



Mise à jour d'OpenRefine

OpenRefine Un outil puissant pour travailler avec des données désordonnées.

Créer un projet
Ouvrir un projet
Importer un projet
Langues

All

		Dernière modification	Nom	Étiquettes	Créateur	Sujet	Description	Nombre de lignes
X	Détails	2023-03-14 15:45 PM	ReconciliationWikidata csv					4
X	Détails	2023-03-14 15:05 PM	donnees DOAJ					1001

Version 3.7.1 [ecf4ff8]

Préférences
Aide
À propos

Parcourir le dossier de l'espace de travail

« Roaming » OpenRefine

Rechercher dans : OpenRefine

Nom	Modifié le	Type	Taille
2322363880064.project	04/04/2023 09:36	Dossier de fichiers	
2639601575510.project	04/04/2023 09:36	Dossier de fichiers	
dbextension	23/08/2021 13:05	Dossier de fichiers	
ex.json	25/04/2022 15:29	Fichier JSON	2 Ko
workspace.old.json	04/04/2023 09:31	Fichier JSON	4 Ko
workspace.old.json	04/04/2023 09:31	Fichier JSON	4 Ko

Sauvegarder ses projets par sécurité.



Désinstallation d'OpenRefine

Il suffit de supprimer sous Windows le répertoire contenant l'exécutable :

 openrefine-3.3	06/07/2021 17:22	Dossier de fichiers
 openrefine-3.5.2	08/03/2022 13:32	Dossier de fichiers
 openrefine-3.6.2	06/10/2022 09:15	Dossier de fichiers
 openrefine-3.7.1	13/03/2023 09:30	Dossier de fichiers

Cela ne supprime pas les projets. Pour le faire, supprimer les répertoires vus page précédente.



OpenRefine : import des données

Parcourir pour sélectionner le fichier csv sur le Blob (ou glisser-déposer).

Créer un projet

Ouvrir un projet

Importer un projet

Langues

Créer un projet en important des données. Quelles sortes de données puis-je importer ?

Les documents de type TSV, CSV, *SV, Excel (.xls et .xlsx), JSON, XML, RDF en XML, OpenDocument (.ods) et Google Data sont reconnus nativement. D'autres formats peuvent être ajoutés via des extensions OpenRefine.

Récupérer les données à partir de

Cet ordinateur

Adresses web (URLs)

Presse-papier

Base de données

Google Data

Chercher un ou plusieurs fichiers à charger :

Parcourir...

Aucun fichier sélectionné.

Suivant »

export-wos_blob_2012-2022.txt

1	Publication Type	Authors Book Authors	Book Editors	Book Group Authors	Author Full Names	Book Author Full Names	
	Group Authors	Article Title	Source Title	Book Series Title	Book Series Subtitle	Language	Document Type
	Conference Title	Conference Date	Conference Location	Conference Sponsor	Conference Host	Author Keywords	Keywords Plus
	Abstract	Addresses	Affiliations	Reprint Addresses	Email Addresses	Researcher Ids	ORCIDs
	Name Preferred	Funding Text	Cited References	Cited Reference Count	Times Cited, WoS Core	Times Cited, All	Funding Orgs
	Databases	180 Day Usage Count	Since 2013 Usage Count	Publisher	Publisher City	Publisher Address	ISSN
	ISBN	Journal Abbreviation	Journal ISO Abbreviation	Publication Date	Publication Year	Volume	Issue
	Number Supplement	Special Issue	Meeting Abstract	Start Page	End Page	Article Number	DOI
	Early Access Date	Number of Pages	WoS Categories	Web of Science Index	Research Areas	IDS Number	Pubmed Id
	Access Designations	Highly Cited Status	Hot Paper Status	Date of Export	UT (Unique WOS ID)	Web of Science Record	
2	C Schumann, A	"Simos, TE; Tsitouras, C"	Schumann, Andrew	Reversible Logic Gates on Physarum			
	Polycephalum	PROCEEDINGS OF THE INTERNATIONAL CONFERENCE OF NUMERICAL ANALYSIS AND APPLIED MATHEMATICS 2014 (ICNAAM-2014)					
	AIP Conference Proceedings	English Proceedings Paper	International Conference on Numerical Analysis and Applied				
	Mathematics (ICNAAM) SEP 22-28, 2014	Rhodes, GREECE	"Physarum polycephalum; quantum logic gates; reversible logic				
	gates; sequential logic gates"	In this paper, we consider possibilities how to implement asynchronous sequential logic					
	gates and quantum-style reversible logic gates on Physarum polycephalum motions. We show that in asynchronous sequential						
	logic gates we can erase information because of uncertainty in the direction of plasmodium propagation. Therefore						
	quantum-style reversible logic gates are more preferable for designing logic circuits on Physarum polycephalum. Univ						
	Informat Technol & Management, PL-35225 Rzeszow, Poland	University of Information Technology & Management Rzeszow					
	Schumann, A (corresponding author), Univ Informat Technol & Management, Sucharskiego 2, PL-35225 Rzeszow, Poland.						
	"Schumann, Andrew/AAU-7875-2020; Schumann, Andrew/R-9825-2016"	"Schumann, Andrew/0000-0002-9944-8627; "					
13	1	9	24	AMER INST PHYSICS	MELVILLE	2	HUNTINGTON QUADRANGLE, STE 1N01, MELVILLE, NY 11747-4501 USA
	0094-243X	978-0-7354-1287-3	AIP CONF PROC	2015	1648		580011
	10.1063/1.4912819	http://dx.doi.org/10.1063/1.4912819		4	"Mathematics, Applied; Physics, Applied"	Conference	
	Proceedings Citation Index - Science (CPCI-S)	"Mathematics; Physics"	BC7YA				2023-03-16
	WOS:000355339703092	View Full Record in Web of Science					
3	C	"Dimonte, A; Romeo, A; Tarabella, G; D'Angelo, P; Erokhin, V; Iannotta, S"	IEEE	"Dimonte, Alice; Romeo,			

idered a feature of more

Windows (CR LF) ANSI INS

Ecran modifié en 3.9.x.

A noter que le chargement de fichiers JSON ou XML, bien que possible, ralentissent beaucoup le chargement. D'autres manières de faire existent (mais avec GREL).



OpenRefine : prévisualisation après import

CNRS Inist

OpenRefine Un outil puissant pour travailler avec des données désordonnées.

Créer un projet « Recommencer Configurer les options d'analyse syntaxique Nom du projet export vos blob 2012 2022 txt Étiquettes Créer un projet »

Ouvrir un projet
Importer un projet
Langues

	Publication Type	Authors	Book Authors	Book Editors	Book Group Authors	Author Full Names	Book Author Full Names	Group Authors	Article Title
1.	C	Schumann, A		Simos, TE; Tsitouras, C		Schumann, Andrew			Reversible Logic Gates on Physarum Polycephalum

Choisir l'encodage

Encodage courant Tous les encodages

Encodage Alias

ISO-8859-1 819, ISO8859-1, I1, ISO 8859-1:1987, ISO_8859-1, 8859_1, iso-ir-100, latin1, cp819, ISO8859_1, IBM819, ISO_8859_1, IBM-819, csISOLatin1

US-ASCII ANSI_X3.4-1968, cp367, csASCII, iso-ir-6, ASCII, iso_646.inv:1983, ANSI_X3.4-1986, ascii7, default, ISO_646.inv:1991, ISO646-US, IBM367, 646, us

UTF-16 UTF_16, unicode, utf16, UnicodeBig

UTF-16BE X-UTF-16BE, UTF_16BE, ISO-10646-UCS-2, UnicodeBigUnmarked

UTF-16LE UnicodeLittleUnmarked, UTF_16LE, X-UTF-16LE

UTF-8 unicode-1-1-utf-8, UTF8

windows-1252 ibm-1252, cp1252, ibm1252, cp5348

Considérer les données comme

Fichiers CSV / TSV / séparateur

Fichiers texte à base de lignes

Fichiers texte à largeur de champ fixe

PC-Axis text files

Fichiers JSON

Fichiers MARC

Fichiers JSON-LD

Fichiers RDF/N3

Fichiers RDF/N-Triples

Format des caractères WINDOWS-1252

Les colonnes sont séparées par

☐ une virgule (CSV)

☒ une tabulation (TSV)

☐ personnalisé \t

☒ Utiliser le caractère " pour fermer les cellules contenant les séparateurs de colonnes

☐ Supprimer les espaces de début et de fin

Protéger les caractères spéciaux avec \

Ignorer la ou les 0 première(s) ligne(s) du début du fichier

Analyser la ou les 1 ligne(s) suivante(s) comme des entêtes de colonnes

Noms de colonnes (séparés par des virgules)

Ignorer les données

Ignorer les données

Mettre à jour l'aperçu

☐ Désactiver l'aperçu automatique

☐ Analyser le texte des cellules comme nombres

☒ Conserver les lignes vides

☒ Enregistrer les cellules vides comme des valeurs nulles

☐ Indiquer la source du fichier

☐ stocker le fichier d'archive

Changer pour UTF-8.

Extension csv mais bien tsv

En général, à décocher.

Créer le projet quand tout est bon,
le nommer avant de valider.



OpenRefine : espace de travail

OpenRefine export vos blob 2012 2022 Permalien

Ouvrir... Exporter Aide

Facette / Filtre Défaire / Refaire 0 / 0 400 lignes Extensions Wikidata

Voir en: lignes entrées Afficher: 5 10 25 50 100 500 1000 lignes « première < précédente 1 > suivante » dernière »

Utiliser les facettes et les filtres

Utiliser les facettes et les filtres pour sélectionner les sous-ensembles de données à traiter. Choisir les méthodes de facette et de filtre dans les menus situés dans les entêtes de colonne.

Vous ne savez pas par où commencer ?
[Regarder ces tutoriels vidéos](#)

	Toutes	Publication Type	Authors	Book Authors	Book Editors	Book Group Authors	Author Full Names	Book Author Full Names
1.	C		Schumann, A		Simos, TE; Tsitouras, C		Schumann, Andrew	
2.	C		Dimonte, A; Romeo, A; Tarabella, G; D'Angelo, P; Erokhin, V; Iannotta, S		IEEE		Dimonte, Alice; Romeo, Agostino; Tarabella, Giuseppe; D'Angelo, Pasquale; Erokhin, Victor; Iannotta, Salvatore	

Menu colonne contextuel.

Grille de données, facettes/filtres et menus divers.



OpenRefine : Colonne Toutes

OpenRefine export wos blob 2012 2022 Permalien

Ouvrir... Exporter Aide

Facette / Filtre Défaire / Refaire 0 / 0

400 lignes

Voir en: lignes entrées Afficher: 5 10 25 50 100 500 1000 lignes « première « précédente 1 suivante » dernière »

Extensions Wikidata

Utiliser les facettes et les filtres

Utiliser les facettes et les filtres pour sélectionner les sous-ensembles de données à traiter. Choisir les méthodes de facette et de filtre dans les menus situés dans les entêtes de colonne.

Vous ne savez pas par où commencer ?
[Regarder ces tutoriels vidéos](#)

Transformer...
Modifier toutes les colonnes
Facette
Éditer les lignes
Éditer les colonnes
Aperçu

Retrier / supprimer les colonnes...
Recopier les valeurs dans les cellules vides consécutives
Vider les valeurs répétées dans des cellules consécutives

Toutes	Publication Type	Authors	Book Authors	Book Editors	Book Group Authors	Author Full Names	Book Author Full Names	Group Au
Schumann, A				Simos, TE; Tsitouras, C		Schumann, Andrew		
2.	C	Dimonte, A; Romeo, A; Tarabella, G; D'Angelo, P; Erokhin, V; Iannotta, S		IEEE		Dimonte, Alice; Romeo, Agostino; Tarabella, Giuseppe; D'Angelo, Pasquale; Erokhin, Victor; Iannotta, Salvatore		

Voir toutes les colonnes ? Colonne Toutes > Editer les colonnes > Retirer-Supprimer



OpenRefine : Supprimer/Garder des colonnes

CNRS Inist

OpenRefine export vos blob 2012 2022 Permalien

Facette / Filtre Défaire / Refaire 0 / 0 400 lignes

Voir en: lignes entrées Afficher: 5 10 25 50 100 500 1000 lignes

« première » « précédente » 1 « suivante » « dernière »

Toutes Publication Type Authors Book Authors Book Editors Book Group Authors Author Full Names Book Author Full Names Group Authors

Utiliser les facettes et les filtres

Utiliser les facettes et les filtres pour sélectionner les sous-ensembles de données à traiter. Choisir les méthodes de facette et de filtre dans les menus situés dans les entêtes de colonne.

Vous ne savez pas par où commencer ?
[Regarder ces tutoriels vidéos](#)

Transformer...
Modifier toutes les colonnes
Facette
Éditer les lignes
Éditer les colonnes
Aperçu

Retrier / supprimer les colonnes...
Recopier les valeurs dans les cellules vides consécutives
Vider les valeurs répétées dans des cellules consécutives

2. C Dimonte, A; Romeo, A; Tarabella, G; D'Angelo, P; Erokhin, V; Iannotta, S IEEE

javascript:()

Trier / Supprimer des colonnes

Glisser des colonnes pour les trier

Déposer des colonnes ici pour les supprimer

Authors
Author Full Names
Book Author Full Names
Group Authors
Article Title
Source Title
Book Series Title
Book Series Subtitle
Language
Document Type
Conference Title
Conference Date
Conference Location
Conference Sponsor
Conference Host

Publication Type
Book Authors
Book Editors
Book Group Authors


Supprimer toutes Ajouter toutes

OK Annuler

La colonne Toutes permet de supprimer, en une fois, plusieurs colonnes (prudence !...)



OpenRefine : Garder 25 colonnes

 **OpenRefine** export vos blob 2012 2022 [Permalien](#)

Ouvrir...

Exporter ▾

Aide

Facette / Filtre

Défaire / Refaire 1 / 1

400 lignes

Extensions Wikidata ▾

Utiliser les facettes et les filtres

Utiliser les facettes et les filtres pour sélectionner les sous-ensembles de données à traiter. Choisir les méthodes de facette et de filtre dans les menus situés dans les entêtes de colonne.

Vous ne savez pas par où commencer ?

[Regarder ces tutoriels vidéos](#)

Voir en: lignes entrées

Afficher: 5 10 25 50 100 500 1000 lignes

« première

« précédente

1

» suivante

» dernière

Toutes	Authors	Author Full Names	Article Title	Source Title	Language	Document Type	Conference Title	Author Keywords			
		1.	Schumann, A	Schumann, Andrew	Reversible Logic Gates on Physarum Polycephalum	PROCEEDINGS OF THE INTERNATIONAL CONFERENCE OF NUMERICAL ANALYSIS AND APPLIED MATHEMATICS 2014 (ICNAAM-2014)	English	Proceedings Paper	International Conference on Numerical Analysis and Applied Mathematics (ICNAAM)	Physarum polycephalum; quantum logic gates; reversible logic gates; sequential logic gates	In this cons how asyr sequ and i reve Phys motik in as sequ we c infor unce direc prop quar reve are r desig on P poly
		2.	Dimonte, A; Romeo, A; Tarabella, G; D'Angelo, P; Erokhin, V; Iannotta, S	Dimonte, Alice; Romeo, Agostino; Tarabella, Giuseppe; D'Angelo, Pasquale; Erokhin, Victor; Iannotta, Salvatore	Interfacing Physarum polycephalum with organic memristors	2015 INTERNATIONAL CONFERENCE ON MEMRISTIVE SYSTEMS (MEMRISYS)	English edit	Proceedings Paper	International Conference on Memristive Systems (MEMRISYS)	Physarum polycephalum; living organism; memristors; synapses; learning	Sinc mem revo of cr now beer as p for n In pa base the c circ larr Poly is w imple func smal elect

Authors

Author Full Names

Article Title

Source Title

Language

Document Type

Conference Title

javascript:()

Author Keywords
Abstract
Addresses
Affiliations
Publisher
ISSN
eISSN
ISBN
Publication Year
Volume
Issue
Start Page
End Page
DOI
Early Access Date
WoS Categories
Research Areas
UT

Authors
Author Full Names
Article Title
Source Title
Language
Document Type
Conference Title

On garde 25 colonnes seulement, utiles aux travaux qui suivront avec LODEX.



OpenRefine : renommer, réordonner ou supprimer des colonnes

OpenRefine export vos blob 2012 2022 Permalien

Ouvrir... Exporter Aide

Facette / Filtre Défaire / Refaire 1 / 1 400 lignes Extensions Wikidata

Voir en : lignes entrées Afficher: 5 10 25 50 100 500 1000 lignes « première < précédente 1 suivante > dernière »

Utiliser les facettes et les filtres

Utiliser les facettes et les filtres pour sélectionner les sous-ensembles de données à traiter. Choisir les méthodes de facette et de filtre dans les menus situés dans les entêtes de colonne.

Vous ne savez pas par où commencer ?
[Regarder ces tutoriels vidéos](#)

Toutes	Authors	Author Full Names	Article Title	Source Title	Language	Document Type	Conference Title	Author Keywords
1.	Schumann, A	Schumann, Andrew	Facette Filtrer le texte Éditer les cellules Éditer la colonne Transposer Trier... Aperçu Réconcilier	PROCEEDINGS OF THE INTERNATIONAL CONFERENCE OF NUMERICAL ANALYSIS AND	English	Proceedings Paper	International Conference on Numerical Analysis and Applied Mathematics (ICNAAM)	Physarum polycephalum; quantum logic gates; reversible logic gates; sequential logic gates
2.	Dimonte, A; Romeo, A; Tarabella, G; D'Angelo, P; Erokhin, V; Iannotta, S	Dimonte, Alice; Romeo, Agostino; Tarabella, Giuseppe; D'Angelo, Pasquale; Erokhin, Victor; Iannotta, Salvatore	Interfacing Physarum polycephalum with organic memristors			Paper	International Conference on Memristive Systems (MEMRISYS)	Physarum polycephalum; living organism; memristors; synapses; learning



OpenRefine : Utiliser des facettes

OpenRefine export was blob 2012 2022 Permalien

Ouvrir... Exporter Aide

Facette / Filtre Défaire / Refaire 1 / 1

400 lignes

Voir en: lignes entrées Afficher: 5 10 25 50 100 500 1000 lignes

« première » « précédente » 1 « suivante » « dernière »

Extensions Wikidata

Facette / Filtre

Rafraîchir Tout réinitialiser Tout supprimer

Language changer

2 choix Trier par: nom compte Groupe

English 398

Russian 2

Facette par nombre de choix

	Toutes	Authors	Author Full Names	Article Title	Source Title	Language	Document Type	Conference Title	Author Keywords
1.	Schumann, A	Schumann, Andrew	Reversible Logic Gates on Physarum Polycephalum	PROCEEDINGS OF THE INTERNATIONAL CONFERENCE OF NUMERICAL ANALYSIS AND APPLIED MATHEMATICS 2014 (ICNAAM-2014)					Physarum polycephalum; quantum logic gates; reversible logic gates; sequential logic gates
2.	Dimonte, A; Romeo, A; Tarabella, G; D'Angelo, P; Erokhin, V; Iannotta, S	Dimonte, Alice; Romeo, Agostino; Tarabella, Giuseppe; D'Angelo, Pasquale; Erokhin, Victor; Iannotta, Salvatore	Interfacing Physarum polycephalum with organic memristors	2015 INTERNATIONAL CONFERENCE ON MEMRISTIVE SYSTEMS (MEMRISYS)	English	Proceedings Paper	International Conference on Memristive Systems (MEMRISYS)	Physarum polycephalum; living organism; memristors; synapses; learning	

Une facette pour voir les valeurs présentes dans une colonne.

Les opérations portant sur une sélection par facette sont limitées à cette sélection (même après suppression de la facette).



OpenRefine : Remplacer des valeurs

The screenshot shows the OpenRefine web interface. At the top, it says 'OpenRefine export was blob 2012 2022'. Below this, there's a 'Facette / Filtre' section on the left with a 'Language' facet showing 2 choices: English (398) and Russian (2). The main table has 400 lines and shows records with columns: Toutes, Authors, Author Full Names, Article Title, Source Title, Language, Document Type, Conference Title, and Author Keywords. A context menu is open over the 'Language' column of the second record, with options: Facette, Filtrer le texte, Éditer les cellules (selected), Éditer la colonne, Transposer, Trier..., Aperçu, Réconcilier, Transformer..., Transformations courantes, Recopier les valeurs dans les cellules vides consécutives, Vider les valeurs répétées dans des cellules consécutives, Diviser les cellules multivaluées..., Joindre les cellules multivaluées..., Grouper et éditer..., and Remplacer... (highlighted).

Toutes	Authors	Author Full Names	Article Title	Source Title	Language	Document Type	Conference Title	Author Keywords
1.	Schumann, A	Schumann, Andrew	Reversible Logic Gates on Physarum Polycephalum	PROCEEDINGS OF THE INTERNATIONAL CONFERENCE OF NUMERICAL ANALYSIS AND APPLIED MATHEMATICS 2014 (ICNAAM-2014)		ings Paper	International Conference on Numerical Analysis and Applied Mathematics	Physarum polycephalum; quantum logic gates; reversible logic gates; sequential logic gates
2.	Dimonte, A; Romeo, A; Tarabella, G; D'Angelo, P; Erokhin, V; Iannotta, S	Dimonte, Alice; Romeo, Agostino; Tarabella, Giuseppe; D'Angelo, Pasquale; Erokhin, Victor; Iannotta, Salvatore	Interfacing Physarum polycephalum with organic memristors	2015 INTERNATIONAL CONFERENCE ON MEMRISTIVE SYSTEMS (MEMRISYS)	English	Procee		Physarum polycephalum; living organism; memristors; synapses; learning

Puis remplacer les valeurs en Anglais de la colonne Language par leur traduction française.

Menu contextuel : Editer les cellules > Remplacer



OpenRefine : Remplacer des valeurs

OpenRefine export was blob 2012 2022 Permalien

Facette / Filtre Défaire / Refaire 2 / 2 400 lignes

Rafraîchir Tout réinitialiser Tout supprimer

Language changer

2 choix Trier par: nom compte Groupe

Anglais 398

Russian 2

Facette par nombre de choix

Voir en: Remplacer

Chercher Russian

☐ insensible à la casse ☐ mot entier ☐ expression rationnelle

Laisser vide pour ajouter la chaîne de remplacement après chaque caractère.

Cocher « expression régulière » pour chercher des caractères spéciaux (sauts de lignes, tabulations...) ou des motifs complexes.

Remplacer par Russe

☐ utiliser \n pour les sauts de ligne, \t pour les tabulations, \n pour \n, \t pour \t.

Si l'option « expression régulière » est cochée et que le motif recherché contient des groupes délimités par des parenthèses, \$0 renverra la chaîne complète correspondant au motif, et \$1, \$2... les 1er, 2e... groupes.

OK Annuler

Document Type	Conference Title	Author Keywords	Language
Proceedings Paper	International Conference on Numerical Analysis and Applied Mathematics (ICNAAM)	Physarum polycephalum; quantum logic gates; reversible logic gates; sequential logic gates	Anglais
Proceedings Paper	International Conference on Memristive Systems (MEMRISYS)	Physarum polycephalum; living organism; memristors; synapses; learning	Anglais

La facette se met à jour dynamiquement au fur et à mesure des remplacements.



OpenRefine : Remplacer des valeurs

The screenshot shows the OpenRefine interface with a dataset of 400 lines. The 'Document Type' facet is active on the left, showing a list of document types: Article (306), Article; Early Access (1), Article; Proceedings Paper (8), Editorial Material (3), Meeting Abstract (3), Proceedings Paper (67), and Review (12). A context menu is open over the 'Document Type' column, showing options like 'Facette', 'Facette textuelle', 'Facette numérique', 'Facette chronologique', 'Facette de nuage de points...', 'Facette textuelle personnalisée...', 'Facette numérique personnalisée...', and 'Facettes personnalisées'. The main table displays two rows of data. The first row has 'Document Type' as 'Facette' and 'Conference Title' as 'Facette textuelle'. The second row has 'Document Type' as 'Proceedings Paper' and 'Conference Title' as 'International Conference on Memristive Systems (MEMRISYS)'.

	Toutes	Authors	Author Full Names	Article Title	Source Title	Language	Document Type	Conference Title	Author Keywords
1.	Schumann, A	Schumann, Andrew	Reversible Logic Gates on Physarum Polycephalum	PROCEEDINGS OF THE INTERNATIONAL CONFERENCE OF NUMERICAL ANALYSIS AND APPLIED MATHEMATICS 2014 (ICNAAM-2014)	Anglais	Facette	Facette textuelle	um; quantum reversible sequential	
2.	Dimonte, A; Romeo, A; Tarabella, G; D'Angelo, P; Erokhin, V; Iannotta, S	Dimonte, Alice; Romeo, Agostino; Tarabella, Giuseppe; D'Angelo, Pasquale; Erokhin, Victor; Iannotta, Salvatore	Interfacing Physarum polycephalum with organic memristors	2015 INTERNATIONAL CONFERENCE ON MEMRISTIVE SYSTEMS (MEMRISYS)	Anglais	Proceedings Paper	International Conference on Memristive Systems (MEMRISYS)	Physarum polycephalum; living organism; memristors; synapses; learning	

Remplacer la valeur « Article; Early Access » par « Article ».

Faire de même qu'avec les langues pour Document Type (Facette puis Remplacer).



OpenRefine : Transformations standards

OpenRefine export vos blob 2012 2022 Permalien

Facette / Filtre Défaire / Refaire 4 / 4 400 lignes

Voir en: lignes entrées Afficher: 5 10 25 50 100 500 1000 lignes

« première » « précédente » 1 « suivante » « dernière »

Extensions Wikidata

Utiliser les facettes et les filtres

Utiliser les facettes et les filtres pour sélectionner les sous-ensembles de données à traiter. Choisir les méthodes de facette et de filtre dans les menus situés dans les entêtes de colonne.

Vous ne savez pas par où commencer ?
[Regarder ces tutoriels vidéos](#)

Toutes	Authors	Author Full Names	Article Title	Source Title	Language	Document Type	Conference Title	Author Keywords
1	Schumann, A	Schumann, Andrew	Reversible Logic Gates on Physarum Polycephalum	PROCEEDINGS OF THE INTERNATIONAL CONFERENCE OF NUMERICAL ANALYSIS AND	Anglais	Proceedings Paper	International Conference on Numerical Analysis and Applied Mathematics	Facette
2	Dimonte, A; Romeo, A; Tarabella, G; D'Angelo, P; Erokhin, V; Iannotta, S	Dimonte, Alice; Romeo, Agostino; Tarabella, Giuseppe; D'Angelo, Pasquale; Erokhin, Victor; Iannotta, Salvatore	Interfacing Physarum polycephalum with organic memristors	2 IN C M S (I				Physarum polycephalum; living organism; memristors; synapses; learning

Supprimer les espaces de début et de fin
Rassembler les espaces consécutifs
Convertir les entités HTML
Remplacer les guillemets courbés par des guillemets droits
En initiales majuscules (en capitales)
En majuscules
En minuscules
En nombre
En date
En texte
En valeurs nulles
Transformer en chaîne vide

Transformer...
Transformations courantes
Recopier les valeurs dans les cellules vides consécutives
Vider les valeurs répétées dans des cellules consécutives
Diviser les cellules multivaluées...
Joindre les cellules multivaluées...
Grouper et éditer...
Remplacer...

Facette
Filtrer le texte
Éditer les cellules
Éditer la colonne
Transposer
Trier...
Aperçu
Réconcilier

Beaucoup de transformations courantes (les plus utilisées).

Passer en majuscules les Author Keywords avec le menu contextuel.



OpenRefine : Traiter Publication Year

OpenRefine export was blob 2012 2022 Permalien

Facette / Filtre Défaire / Refaire 5 / 6 400 lignes

Rafraîchir Tout réinitialiser Tout supprimer

Voir en: lignes entrées Afficher: 5 10 25 50 100 500 1000 lignes

« première » « précédente » 1 « suivante » « dernière »

Extensions Wikidata

Publication Year changer

12 choix Trier par: nom compte Groupe

2015 64
2016 39
2017 43
2018 35
2019 34
2020 29
2021 24
2022 28
2023 1
(blank) 1

Facette par nombre de choix

es	Affiliations	Publisher	ISSN	eISSN	ISBN	Publication Year	Volume	Issue	Start Page	End Page	DOI
	University of Information Technology & Management Rzeszow	AMER INST PHYSICS	0094-243X		978-0-7354-1287-3						10.1063/1.4912819
e;	Consiglio Nazionale delle Ricerche (CNR); Istituto dei Materiali per l'Elettronica ed il Magnetismo (IMEM-CNR)	IEEE			978-1-4673-9209-9	2015					

Pour un corpus extrait pour une date de publication entre 2014 et 2024, y a-t-il une valeur 2025 ?

On détecte ici une valeur « blank » qui signifie une valeur absente.

Pour remédier à ces deux situations, il faut faire un choix.

Une facette sur une colonne permet de détecter beaucoup d'anomalies.



OpenRefine : Traiter Publication Year

The screenshot shows the OpenRefine interface with the 'Publication Year' facet selected. The facet shows a list of years from 2015 to 2022, with 2023 highlighted. The main table displays one matching line with the following data:

ISBN	Publication Year	Volume	Issue	Start Page	End Page	DOI	Early Access Date	WoS Categories	Research Areas
	2023						NOV 2022	Computer Science, Artificial Intelligence	Computer Science

A modal dialog is open for editing the 'Publication Year' cell, showing the value '2023' and buttons for 'Appliquer', 'Appliquer à toutes les cellules identiques', and 'Annuler'.

Publication Year	Volume	
2022	56	1
Éditer cette cellule		

Pour modifier la valeur d'une seule cellule dans une colonne, cliquer sur le bouton bleu « edit » qui apparaît au survol de la souris.

On utilisera l'année du champ « Early Access Date » pour traiter les 2 anomalies recensées.

Pour sélectionner une ligne (ou plusieurs) selon la valeur dans une facette, cliquer sur la valeur choisie.



OpenRefine : Regrouper des valeurs proches

CNRS Inist

OpenRefine export was blob 2012 2022 Permalien

Facette / Filtre Défaire / Refaire 5 / 5 400 lignes Extensions Wikidata

Rafraîchir Tout réinitialiser Tout supprimer Voir en: lignes entrées Afficher: 5 10 25 50 100 500 1000 lignes « première » « précédente » 1 « suivante » dernière »

Publisher changer

88 choix Trier par: nom compte Groupe

IOP PUBLISHING LTD 30
IEEE 29
TAYLOR & FRANCIS LTD 22
OLD CITY PUBLISHING INC 21
SPRINGER 21
ELSEVIER 17
ROYAL SOC 17
ELSEVIER SCI LTD 16
IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC 12
WILEY 12

for Keywords	Abstract	Addresses	Affiliations	Publisher	ISSN	eISSN	ISBN	Publication Year	Volume	Issue
UM HALUM; M LOGIC REVERSIBLE ATES; IAL LOGIC	In this paper, we consider possibilities how to implement asynchronous sequential logic gates and quantum-style reversible logic gates on Physarum polycephalum motions. We show that in asynchronous sequential logic gates we can erase information because of uncertainty in the direction of plasmodium propagation. Therefore quantum-style reversible logic gates are more preferable for designing logic circuits on Physarum polycephalum.	Univ Informat Technol & Management, PL-35225 Rzeszow, Poland	University of Information Technology & Management Rzeszow	Facette	Facette textuelle		4-1287-3	2015	1648	
UM HALUM; RGANISM; ORS; ES; LEARNING	Since their prediction, memristive devices revolutionized the world of computing and nowadays they have been widely considered as promising candidate for mimicking synapses. In particular, organic-based memristors allow the construction of circuits capable of learning. Physarum Polycephalum slime mold is well suited for the implementation of the functional properties of smart living systems into electronic devices	[Dimonte, Alice; Romeo, Agostino; Tarabella, Giuseppe; D'Angelo, Pasquale; Erokhin, Victor; Iannotta, Salvatore] CNR, IMEM, Parco Area Sci 37-A, Parma, Italy	Consiglio Nazionale delle Ricerche (CNR); Istituto dei Materiali per l'Elettronica ed il Magnetismo (IMEM-CNR)	Facette	Facette textuelle					

Facette
Filtrer le texte
Éditer les cellules
Éditer la colonne
Transposer
Trier...
Aperçu
Réconcilier

Facette textuelle
Facette numérique
Facette chronologique
Facette de nuage de points...
Facette textuelle personnalisée...
Facette numérique personnalisée...
Facettes personnalisées

Après une facette textuelle sur la colonne Publisher, cliquer sur le bouton Groupe.

Certaines données textuelles peuvent avoir de nombreuses formes variantes.



OpenRefine : Regrouper des valeurs proches

CNRS Inist

Regrouper & Éditer une colonne "Publisher"

Cet outil vous aide à identifier des groupes de cellules ayant des valeurs différentes mais qui peuvent correspondre à des représentations alternatives de la même valeur. Par exemple, les deux chaînes "New York" et "new york" n'ont qu'une différence de casse et font très certainement référence à la même ville. "Gödel" et "Godel" se réfèrent probablement à la même personne. [En savoir plus...](#)

Méthode collision de clés Fonction de codage Daitch-Mokotoff 9 grappes trouvées

Taille du groupe	Nombre de lignes	Valeurs dans le groupe	Fusionner ?	Nouvelle valeur
8	51	<ul style="list-style-type: none">• SPRINGER (21 lignes)• SPRINGER-VERLAG BERLIN (8 lignes)• SPRINGER INTERNATIONAL PUBLISHING AG (7 lignes)• SPRINGER HEIDELBERG (6 lignes)• SPRINGER WIEN (4 lignes)• SPRINGER BASEL AG (2 lignes)• SPRINGER NATURE (2 lignes)• SPRINGER INT PUBLISHING AG	<input type="checkbox"/>	SPRINGER
3	13	<ul style="list-style-type: none">• ELSEVIER SCIENCE BV (8 lignes)• ELSEVIER SCIENCE INC (3 lignes)• ELSEVIER SCIENCE SA (2 lignes)	<input type="checkbox"/>	ELSEVIER SC
2	5	<ul style="list-style-type: none">• WORLD SCIENTIFIC PUBL CO PTE LTD (4 lignes)• PARLAR SCIENTIFIC PUBLICATIONS (P S P)	<input type="checkbox"/>	WORLD SCIEI
2	6	<ul style="list-style-type: none">• OXFORD UNIV PRESS (5 lignes)• OXFORD UNIV PRESS INC	<input type="checkbox"/>	OXFORD UNIV

Choix dans le groupe
2 — 8

Lignes dans le groupe
2 — 51

Longueur moyenne des choix
18 — 38

Variabilité de la longueur des choix
0 — 8.16

Tout sélectionner Enlever toutes les sélections Exporter les groupes Fusionner la sélection & regrouper Fusionner la sélection & fermer Fermer

Méthode collision de clés

Taille du groupe collision de clés
Voisin le plus proche

Fonction de codage Daitch-Mokotoff

- Empreinte
- empreinte N-grammes
- metaphone3
- phonétique de Cologne
- Daitch-Mokotoff
- Beider-Morse



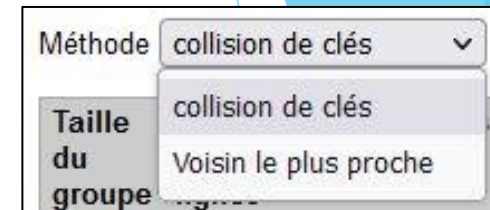
OpenRefine : Regrouper des valeurs proches

CNRS Inist

<https://openrefine.org/docs/manual/cellediting#clustering-methods>

Ces algorithmes peuvent faire des rapprochements non pertinents (par exemple deux numéros de téléphones se « ressemblent »). Mais aussi des formes proches peuvent ne pas être rapprochées.

La collision de clés est rapide mais basique. La méthode du plus proche voisin est plus élaborée mais plus lente.



La fonction de codage empreinte fait des regroupements sur de simples différences de casse, d'accentuation, de ponctuation, de nombre d'espaces, etc...

La fonction empreinte N-grammes regroupe des expressions à N termes, par exemple en bigramme (paramètre 2) elle regroupe « paris-plage » et « paris plage ».

Les suivants peuvent faire des regroupements phonétiques (anglais pour metaphone3) ou autres (voir documentation).



OpenRefine : Annuler ou rejouer un Traitement - Défaire / Refaire

OpenRefine export was blob 2012 2022 Permalien

Ouvrir... Exporter Aide

Facette / Filtre Défaire / Refaire 8 / 8 400 lignes

Extensions Wikidata

Voir en: lignes entrées Afficher: 5 10 25 50 100 500 1000 lignes

« première » « précédente » 1 « suivante » « dernière »

	Toutes	Authors	Author Full Names	Article Title	Source Title	Language	Document Type	Conference Title	Author Keywords
1.	Schumann, A	Schumann, Andrew	Reversible Logic Gates on Physarum Polycephalum	PROCEEDINGS OF THE INTERNATIONAL CONFERENCE OF NUMERICAL ANALYSIS AND APPLIED MATHEMATICS 2014 (ICNAAM-2014)	Anglais	Proceedings Paper	International Conference on Numerical Analysis and Applied Mathematics (ICNAAM)	PHYSARUM POLYCEPHALUM; QUANTUM LOGIC GATES; REVERSIBLE LOGIC GATES; SEQUENTIAL LOGIC GATES	
2.	Dimonte, A; Romeo, A; Tarabella, G; D'Angelo, P; Erokhin, V; Iannotta, S	Dimonte, Alice; Romeo, Agostino; Tarabella, Giuseppe; D'Angelo, Pasquale; Erokhin, Victor; Iannotta, Salvatore	Interfacing Physarum polycephalum with organic memristors	2015 INTERNATIONAL CONFERENCE ON MEMRISTIVE SYSTEMS (MEMRISYS)	Anglais	Proceedings Paper			

0. Create project

1. Reorder columns

2. Text transform on 398 cells in column Language: value.replace("English","Anglais")

3. Text transform on 2 cells in column Language: value.replace("Russian","Russe")

4. Edit single cell on row 339, column Document Type

5. Text transform on 322 cells in column Author Keywords: value.toUpperCase()

6. Edit single cell on row 15, column Publication Year

7. Fill down 0 cells in column Publication Year

8. Edit single cell on row 339, column Publication Year

Cet historique n'enregistre pas tout (par ex. ni facette ni tri).

ATTENTION : perte des opérations après un retour en arrière.



OpenRefine : Données multivaluées

OpenRefine export vos blob 2012 2022 Permalien

Facette / Filtre Défaire / Refaire 8 / 8 400 lignes Extensions Wikidata

Voir en: lignes entrées Afficher: 5 10 25 50 100 500 1000 lignes « première » « précédente » 1 « suivante » « dernière »

Utiliser les facettes et les filtres

Utiliser les facettes et les filtres pour sélectionner les sous-ensembles de données à traiter. Choisir les méthodes de facette et de filtre dans les menus situés dans les entêtes de colonne.

Vous ne savez pas par où commencer ?
[Regarder ces tutoriels vidéos](#)

Facette

Filtrer le texte

Éditer les cellules

Éditer la colonne

Transposer

Trier...

Aperçu

Réconcilier

Transformer...

Transformations courantes

Recopier les valeurs dans les cellules vides consécutives

Vider les valeurs répétées dans des cellules consécutives

Diviser les cellules multivaluées...

Joindre les cellules multivaluées...

Grouper et éditer...

Remplacer...

Diviser les cellules multivaluées

Comment diviser des cellules à plusieurs valeurs

☒ par séparateur

Séparateur ; ☐ expression rationnelle

☐ par les longueurs de champs

Liste les longueurs en les séparant par des virgules, par exemple 5, 7, 15

☐ par le passage des minuscules aux majuscules [11Abc, Def22] ☐ Inverser l'ordre de division [11A, bcD, ef22]

☐ par le passage des chiffres aux lettres [11, AbcDef22] ☐ Inverser l'ordre de division [11AbcDef, 22]

OK Annuler

Découpage en axe vertical, souvent sur un délimiteur particulier (ici le ;).



OpenRefine export was blob 2012 2022
 Permalien

Facette / Filtre

Défaire / Refaire 9 / 9

Rafraîchir

Tout réinitialiser

Tout supprimer

Author Full Names

changer

912 choix

Trier par: nom compte

Groupe

Abdel-Mottaleb, Noha 1

Adamatzky, Andrew 2

Adamatzky, A. 2

Adamatzky, A. I. 2

Adamatzky, Andrew 36

Adamatzky, Andrew I. 4

Adamatzky, Andrew, I 1

Adamatzky, Andy 1

Aihara, Kazuyuki 1

Akinshin, Stepan 1

Akita, Dai 1

1451 lignes

Extensions Wikidata

Voir en: lignes entrées

Afficher: 5 10 25 50 100 500 1000 lignes

« première

« précédente

1

suivante »

dernière »

Toutes	Authors	Author Full Names	Article Title	Source Title	Language	Document Type	Conference Title	Author Keywords	Abstract		
☆	🔍	3.	Romeo, Agostino								
☆	🔍	4.	Tarabella, Giuseppe								
☆	🔍	5.	D'Angelo, Pasquale								
☆	🔍	6.	Erokhin, Victor								
☆	🔍	7.	Iannotta, Salvatore								
☆	🔍	8.	Cifarelli, A; Berzina, T; Erokhin, V	Cifarelli, Angelica	Bio-organic memristive device: polyaniline-Physarum polycephalum interface	PHYSICA STATUS SOLIDI C: CURRENT TOPICS IN SOLID STATE PHYSICS, VOL 12, NO 1-2	Anglais	Proceedings Paper	European-Materials-Research-Society Spring Meeting Symposium S on Memristor Materials, Mechanisms and Devices for Unconventional Computing	MEMRESISTIVE DEVICE; POLYMERS; PHYSARUM POLYCEPHALUM; HYDROGELS	Organic memristive device is an artificial system mimicking biological synapses. This work is the realization of a single device/Physarum polycephalum and studying its electrical properties was found to be a promising mold re...

Après une facette textuelle sur la colonne découpée verticalement, on peut voir les possibles formes variantes des données isolées.

Est-ce une bonne
idée d'uniformiser
des noms d'auteurs
qui semblent être
identiques ou
voisins ?

Le découpage vertical est très utile pour explorer ces données en facette.



OpenRefine : Données multivaluées

Extensions: Wikidata

mière < précédente 1 de 401 pages suivante > dernière »

	ISSNs	Publisher	Citation	Licence
icejQ	1099-4300	MDPI AG	Facette	CC BY
cellular S	2077-0472	MDPI AG	Facette	CC BY

Diviser en plusieurs colonnes...

Joindre des colonnes...

Ajouter une colonne en fonction de cette colonne...

Ajouter une colonne en moissonnant des URL...

Ajouter des colonnes à partir de valeurs réconciliées...

Renommer cette colonne

Supprimer cette colonne

Déplacer la colonne en premier

Déplacer la colonne en dernier

Déplacer la colonne à gauche

Déplacer la colonne à droite

Diviser la colonne Citation en plusieurs colonnes

Méthode de division de la colonne

☒ par séparateur
Séparateur ☐ expression rationnelle
Diviser en colonnes au plus (laisser vide pour ne pas limiter)

☐ par les longueurs de champs

Liste les longueurs en les séparant par des virgules, par exemple 5, 7, 15

Après la division

☐ Deviner le type de cellule

☐ Supprimer cette colonne

OK Annuler

Extensions: Wikidata

première < précédente 1 de 401 pages suivante > dernière »

Citation	Citation 1	Citation 2	Citation 3	Citation 4	Licence
Entropy, Vol 17, Iss 12, Pp 7848-7858 (2015)	Entropy	Vol 17	Iss 12	Pp 7848-7858 (2015)	CC BY
Agriculture (Basel), Vol 5, Iss 4, Pp 1172-1182 (2015)	Agriculture (Basel)	Vol 5	Iss 4	Pp 1172-1182 (2015)	CC BY
	Vol 16	Iss 12	Pp 28285-28295 (2015)		CC BY

En général, à décocher.

Exemple non extrait du fichier Blob. Pourquoi ?

Le découpage horizontal est très utile pour séparer des données de nature différente.



OpenRefine : Exporter des données

OpenRefine export vos blob 2012 2022 [Permalien](#)

Facette / Filtre Défaire / Refaire 8 / 9 400 lignes

Voir en: lignes entrées Afficher: 5 10 25 50 100 500 1000 lignes

Utiliser les facettes et les filtres

Utiliser les facettes et les filtres pour sélectionner les sous-ensembles de données à traiter. Choisir les méthodes de facette et de filtre dans les menus situés dans les entêtes de colonne.

Vous ne savez pas par où commencer ? [Regarder ces tutoriels vidéos](#)

	Toutes	Authors	Author Full Names	Article Title	Source Title	Language	Document Type
1.	Schumann, A	Schumann, Andrew	Reversible Logic Gates on Physarum Polycephalum	PROCEEDINGS OF THE INTERNATIONAL CONFERENCE OF NUMERICAL ANALYSIS AND APPLIED MATHEMATICS 2014 (ICNAAM-2014)	Anglais	Proceedings Paper	
2.	Dimonte, A; Romeo, A; Tarabella, G; D'Angelo, P; Erokhin, V; Iannotta, S	Dimonte, Alice; Romeo, Agostino; Tarabella, Giuseppe; D'Angelo, Pasquale; Erokhin, Victor; Iannotta, Salvatore	Interfacing Physarum polycephalum with organic memristors	2015 INTERNATIONAL CONFERENCE ON MEMRISTIVE SYSTEMS (MEMRISYS)	Anglais	Proceedings Paper	

Archive de projet OpenRefine

Valeurs séparées par des tabulations

Valeurs séparées par des virgules

Table HTML

Excel (.xls)

Excel 2007+ (.xlsx)

Classeur OpenDocument ODF (.ods)

Export tabulaire personnalisé...

Export SQL...

Patrons...

Projet OpenRefine archivé dans Google Drive...

Google Sheets...

Contributions Wikibase...

Fichier QuickStatements

Schéma Wikibase

Conference on Memristive Systems (MEMRISYS)

POLYCEPHALUM; LIVING ORGANISM; MEMRISTORS; SYNAPSES; LEARNING

En sélectionnant certaines lignes du projet ou datagrid, on peut faire un export partiel.



OpenRefine : Exporter les traitements

OpenRefine export was blob 2012 2022 Permalien

Ouvrir... Exporter Aide

Facette / Filtre Défaire / Refaire 8 / 9 400 lignes Extensions Wikidata

Extraire... Appliquer...

Voir en: lignes entrées Afficher: 5 10 25 50 100 500 1000 lignes « première < précédente 1 suivante > dernière »

Toutes Authors Author Full Names Article Title Source Title Language Document Type Conference Title Author Keywords

1. Schumann, A Schumann, Andrew Reversible Logic Gates PROCEEDINGS OF THE Anglais Proceedings Paper International Conference on POLYCEPHALUM In thi

2. Dimonte, A; Romeo, A; Tarabella, G; D'Angelo, P; Erokhin, V; Iannotta, S Dimonte, Alice; Romeo, Agostino; Tarabella, Giuseppe; D'Angelo, Pasquale; Erokhin, Victor; Iannotta, Salvatore

0. Create project

1. Reorder columns

2. Text transform on 398 cells in column Language: value.replace("English","Anglais")

3. Text transform on 2 cells in column Language: value.replace("Russian","Russe")

4. Edit single cell on row 339, column Document Type

5. Text transform on 322 cells in column Author Keywords: value.toUppercase()

6. Edit single cell on row 15, column Publication Year

7. Fill down 0 cells in column Publication Year

8. Edit single cell on row 339, column Publication Year

9. Split multi-valued cells in column Author Full Names

Extraire l'historique des opérations

Extraire et enregistrer des sous-parties de l'historique des opérations au format JSON afin de les réappliquer dans ce projet ou de les réutiliser ultérieurement dans d'autres projets.

- ☒ Reorder columns
- ☒ Text transform on cells in column Language using expression value.replace("English","Anglais")
- ☒ Text transform on cells in column Language using expression value.replace("Russian","Russe")
- ☒ Edit single cell on row 339, column Document Type
- ☒ Text transform on cells in column Author Keywords using expression value.toUppercase()
- ☒ Edit single cell on row 15, column Publication Year
- ☒ Fill down cells in column Publication Year
- ☒ Edit single cell on row 339, column Publication Year

Tout sélectionner Enlever toutes les sélections Exporter

```
{
  "expression": "value.toUppercase()",
  "onError": "keep-original",
  "repeat": false,
  "repeatCount": 10,
  "description": "Text transform on cells in column Author Keywords using",
  "op": "core/fill-down",
  "engineConfig": {
    "facets": [
      {
        "type": "list",
        "name": "Publication Year",
        "expression": "value",
        "columnName": "Publication Year",
        "invert": false,
        "omitBlank": false,
        "omitError": false,
        "selection": [],
        "selectBlank": true,
        "selectError": false
      }
    ]
  },
  "mode": "row-based",
  "columnName": "Publication Year",
  "description": "Fill down cells in column Publication Year"
}
```

Fermer

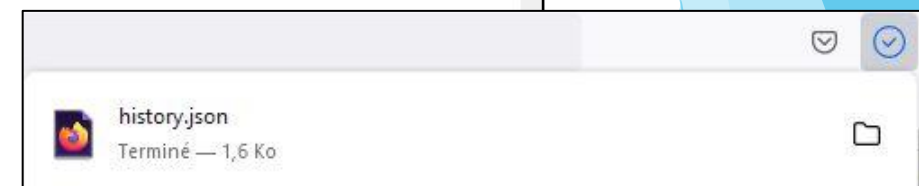
Copier la sélection (en JSON) : Ctrl-C, dans Notepad++ par ex., ou Exporter.



OpenRefine : Exporter les traitements

```
1  [
2  {
3    "op": "core/column-reorder",
4    "columnName": [
5      "Authors",
6      "Author Full Names",
7      "Article Title",
8      "Source Title",
9      "Language",
10     "Document Type",
11     "Conference Title",
12     "Author Keywords",
13     "Abstract",
14     "Addresses",
15     "Affiliations",
16     "Publisher",
17     "ISSN",
18     "eISSN",
19     "ISBN",
20     "Publication Year",
21     "Volume",
22     "Issue",
23     "Start Page",
24     "End Page",
25     "DOI",
26     "Early Access Date",
27     "WoS Categories",
28     "Research Areas",
29     "UT (Unique WOS ID)"
30   ],
31   "description": "Reorder columns"
32 },
33 {
34   "op": "core/text-transform",
35   "engineConfig": {
```

Ou exporter ces données dans le fichier history.json



Copier la sélection (en JSON) dans Notepad++ par exemple : Ctrl-V (sauvegarde / échange)



OpenRefine : (Re)Appliquer les traitements

CNRS Inist

OpenRefine export was blob 2012 2022 Permalien

Ouvrir... Exporter Aide

Facette / Filtre Défaire / Refaire 0 / 9 400 lignes Extensions Wikidata

Extraire... Appliquer...

Voir en: lignes entrées Afficher: 5 10 25 50 100 500 1000 lignes « première < précédente 1 suivante > dernière »

Filter

0. Create project

1. Reorder columns

2. Text transform on 398 cells in column Language: value.replace("English","Anglais")

3. Text transform on 2 cells in column Language: value.replace("Russian","Russe")

4. Edit single cell on row 339, column Document Type

5. Text transform on 322 cells in column Author Keywords: value.toUppercase()

6. Edit single cell on row 15, column Publication Year

7. Fill down 0 cells in column Publication Year

8. Edit single cell on row 339, column Publication Year

9. Split multi-valued cells in column Author Full Names

Toutes	Publication Type	Authors	Book Authors	Book Editors	Book Group Authors	Author Full Names	Book Author Full Names	Group Au
1.	C	Schumann, A		Simos, TE; Tsitouras, C		Schumann, Andrew		
2.	C	Dimonte, A; Romeo, A; Tarabella, G; D'Angelo, P; Erokhin, V; Iannotta, S		IEEE		Dimonte, Alice; Romeo, Agostino; Tarabella, Giuseppe; D'Angelo, Pasquale; Erokhin, Victor; Iannotta, Salvatore		

A noter qu'il n'est pas encore proposé de charger le fichier history.json.

Copier le fichier JSON à partir de Notepad++ par exemple : Ctrl-V (sauvegarde / échange).



OpenRefine : Utiliser GREL

GREL est un langage de script utilisable dans OpenRefine, voulu proche de Javascript et de Python (mais en plus simple).

Il permet des transformations personnalisées, parfois très puissantes comme le passage de données en arbre JSON ou XML.

Historiquement, GREL voulait dire « Google Refine Expression Language » et maintenant le G est là pour « General ».

Ajouter une colonne en fonction d'une colonne Volume

Nouveau nom de colonne

En cas d'erreur ☒ vider la cellule ☐ conserver l'erreur ☐ copier la valeur depuis la colonne originale

Expression Langue General Refine Expression Language (GREL) ▾

Pas d'erreur de syntaxe.

Aperçu Historique Étoilée Aide

row	value	"vol." + value
1.	1648	vol.1648
2.		vol.
3.	12	vol.12
4.		vol.
5.	36	vol.36
6.	550	vol.550

OK Annuler

Informations détaillées : <https://docs.openrefine.org/manual/grel>



OpenRefine : Autres services / extensions

cnrs Inist

Il existe d'autres services de réconciliation mais, comme ils ont été développés pour d'anciennes versions d'OpenRefine, il faut bien les tester avant toute chose.

Par ex. : <http://refine.codefork.com/>

refine.codefork.com

This server provides [OpenRefine](#) reconciliation services for the following data sources:

- [VIAF](#) - Virtual International Authority File
- [ORCID](#) - digital identifiers for researchers
- [Open Library](#) - an open, editable library catalog (a project of the Internet Archive)
- more to come...



OpenRefine VIB-BITS plugin

Copyright © 2014 VIB

Cette ancienne extension fonctionne correctement avec les versions d'OpenRefine jusqu'à la 3.4.1

Il existe aussi d'assez nombreuses extensions pour OpenRefine mais, là aussi, bien faire attention aux versions concernées. De vieilles extensions peuvent parfois être migrées vers la version courante d'OpenRefine (voir documentation)... ou pas.

<https://openrefine.org/extensions>