

## 1. Bases de données

Agriculture\_alimentation\_monde:

<http://faostat3.fao.org/download/E/EF/F>

Import-export mondial:

<http://comtrade.un.org/data/>

**Trade Nosis, COMEX** : information gratuite import-export pour 243 pays, basé sur les infos officielles de chaque pays

<http://trade.nosis.com/es/Comex/Importacion-Exportacion/Suiza/oro--oro-incluido-el-oro-platinado-en-bruto-semilabrado-o-en-polvo/CH/7108>

<http://www.nosis.com/es/exi/exiexplorer/queesexiexplorer>: version payante, permet de voir qui sont les exportateurs des USA et du mercosur

Camp de réfugiés, bases militaires, stades:

<http://virtualglobetrotting.com/search/?cx=partner-pub-9765959110898159%3Aaliduc9jj2sn&cof=FORID%3A9&ie=utf8&t=combined&q=camp+refugees>

Archive.org : capture certains sites a différentes dates et permet de retrouver des pages qui ont disparu

Evolution d'une ville:

Google (historical) earth

Registres du commerce et documents de justice dans 141 sources (y compris les Panama papers ou interpol):

<https://data.occrp.org/>

Interpol:

<http://www.interpol.int/notice/search/wanted/2015-4581>

Liens d'intérêt des parlementaires suisses : <https://www.parlament.ch/centers/documents/de/rar-nr-interessen.pdf>

<https://lobbywatch.ch/fr>

Liste des personnes accréditées au parlement suisse :

<https://www.parlament.ch/centers/documents/de/zutrittsberechtigigte-nr.pdf>

Base de données des objets parlementaires : <https://www.parlament.ch/fr/ratsbetrieb/suche-curia-vista#k=>

Votation des parlementaires :

<https://www.parlament.ch/fr/ratsbetrieb/abstimmungen/abstimmungs-datenbank-nr>

Données douanières suisse:

<https://www.swiss-impex.admin.ch/>

Entreprises suisses cotées en bourse, exportable en csv : [http://www.six-swiss-exchange.com/shares/companies/issuer\\_list\\_fr.html](http://www.six-swiss-exchange.com/shares/companies/issuer_list_fr.html)

Actionnaires par entreprises cotées en bourse : <https://www.six-exchange-regulation.com/fr/home/publications/significant-shareholders.html>

Marchés publics suisses

<https://www.simap.ch/shabforms/COMMON/search/searchresult.jsf>

version scrappable: <https://beschaffungen.fdn.iwi.unibe.ch/zuschlaege.php>

Registre du commerce

<https://zefix.ch>

outil ringier opensource, ressort des CSV:

<http://tools.letemps.ch/zefix/>

## **2. Outils**

### **a. Créer une base de données:**

**Googleform** : permet de faire un questionnaire pour collecter des données

### **b. Scraper une base de données:**

Outwithub (scraper qui contient aide au codage)

<https://scraperwiki.com/help/corporate/>

<http://lessourceurs.com/les-articles/> (trucs et astuces)

<http://www.downdthemall.net/> (add de firefox)

scraper (add to chrome):

<https://chrome.google.com/webstore/detail/scraper/mbigbapnjcgaffohmbkdlecaccepngjd>

### **c. Convertir des fichiers:**

**Nitro** (payant) : permet de convertir des formats de fichiers, même des pdf image en word

**Import.io** (gratuit) : convertit des fichiers

**Omnipage** : transforme des PDF en text (OCR)

**Nitro Pro 10** : OCR, 90 dollars, fonctionne bien. Convertit même des images semble-t-il.

**Zamzara** : convertit des PDF en word ou en Excel

**Tabula** : convertisseur de PDF contenant des tableaux en tableau excel, gratuit et open source, développé par La Nacion, N-Y Times et Pro publica. Ne fonctionne que s'il s'agit de tableau inséré dans un pdf.

**Vozdata** : développé par ProPublica, permet aux lecteurs et citoyens de participer à l'ouverture de fichier. Le programme est opensource mais il faut l'aide d'un programmeur pour le connecter au serveur

<https://opennews.org/> <http://vozdata.lanacion.com.ar/> <https://github.com/crowdata/crowdata>

#### **d. nettoyer et analyser une base de données**

**excel**: pour la plupart des petites bases de données

**openrefine ou googlerefine**: pour les grandes bases de données et un nettoyage très rapide. Aussi pour scraper à l'aide d'API

#### **c. Dessiner des réseaux**

**Xmind**: dessiner des relations, adjoindre des références web, no de tel, etc. Simple à utiliser et gratuit. Permet d'extraire un csv.

**NODEXL** : a partir d'un fichier excel dessine un réseau complexe (twitter, facebook). Version gratuite et version payante. Fonctionne avec windows

**Gephi**: Fait la même chose que NODEXL

#### **d. Archiver des données**

**Archivarius 3000** : archive et convertit plusieurs formats de documents pour pouvoir rechercher par mots-clés. (windows)

#### **e. analyse de discours**

**Worldnet** : permet de copier un discours et sort un fichier excel avec le nombre de mots qui sortent dans ces discours et combien de fois ces mots sont utilisés

**Tagxedo** : compte combien de fois deux mots sortent en même temps dans un discours

#### **f. Faire des graphs**

Datawrapper

Highchart

Tableau Public

Cartodb: pour les cartes

### **3.ouvrages**

**Guide du datajournalisme, Nicolas Kayser-Bril** / Traduction du Datajournalism handbook reprenant de la théorie et des cas pratiques dans les plus grandes rédactions du monde  
Gratuit

<http://jplusplus.github.io/guide-du-datajournalisme/>

**Manuel de periodismo de datos iberoamericano:** exemples pratiques dans presque tous les pays d'Amérique latine, en espagnol

Gratuit

<http://manual.periodismodedatos.org/>

Plus d'infos sur ce que j'ai fait:

<http://tedxtalks.ted.com/video/Investigative-Data-Journalism-M>

<http://pages.rts.ch/la-1ere/programmes/medialogues/7704556-medialogues-du-21-05-2016.html>