

Autoformation Callisto

Plan du cours « Corpus et données : comment utiliser les données ouvertes de Wikimedia »(V1, 2024)
– Pierre-Yves Beaudouin, chargé de mission RéSoNA



URFIST
bordeaux

université
de BORDEAUX

Plan du cours

Corpus et données : comment utiliser les données ouvertes de Wikimedia

Chapitre 0. Synopsis

Chapitre 1. Dataviz

1. Types de graphiques

- 1.1. Déviation
- 1.2. Corrélation
- 1.3. Classement
- 1.4. Distribution
- 1.5. Évolution dans le temps
- 1.6. Ordre de grandeur
- 1.7. Spatial
- 1.8. Partie d'un tout
- 1.9. Flux

2. Outils Wikimedia

2.1. Pageviews Analysis

2.2. Wikistats

2.3. XTools

2.4. Wikiscan

2.5. Humaniki – Wikimedia Diversity Dashboard Tool

2.6. Who Wrote That?

2.7. Wikimedia Grafana

3. Auto-évaluation des connaissances

4. Ressources supplémentaires

Plan du cours

Chapitre 2. Données

1. Introduction
2. Sources de données
3. Données de consultation
 - 3.1. Statistiques de fréquentation
 - 3.2. Appareils uniques
 - 3.3. Flux de clics
 - 3.4. Données technologiques
4. Corpus
 - 4.1. Wikipedia
 - 4.2. Wikidata
5. Métadonnées

6. Divers jeux de données

- 6.1. Références bibliographiques
 - 6.2. Personnes notables
 - 6.3. Liens Wikipédia sur Twitter
 - 6.4. Enquête sur les usages de Wikipedia
 - 6.5. Rencontres entre wikimédiens
7. Bonnes pratiques de réutilisation des données
 8. Auto-évaluation des connaissances

Plan du cours

Chapitre 3. Outils d'extraction

1. Introduction
2. Techniques basiques
 - 2.1. Copier-coller
 - 2.2. Export XML
 - 2.3. Moteur de recherche
 - 2.4. PetScan
3. Outils d'interrogation de base de données
 - 3.1. Quarry (SQL)
 - 3.2. Query (SPARQL)
4. API

5. Outils de programmation

- 5.1. Python
- 5.2. R
- 5.3. PAWS (Jupyter Notebook)
6. Flux des modifications récentes
7. Auto-évaluation des connaissances

Plan du cours

Chapitre 4. Outils d'enrichissement

1. Ajout de données structurées
 - 1.1. QuickStatements
 - 1.2. EditGroups
 - 1.3. Mix'n'match
 - 1.4. OpenRefine
2. Import de fichiers
 - 2.1. Wikipédia
 - 2.2. Wikimedia Commons (Upload Wizard)
 - 2.3. Pattypan
3. Auto-évaluation des connaissances

Chapitre 5. Annuaire des outils

Chapitre 6. Où trouver de l'aide ?

Chapitre 7. Bibliographie et crédits

Chapitre 8. Sur le même thème

Crédits

IBM 96-hole punched card from 1969, Marcin Wichary, CC-BY-SA-2.0