

Fouiller son terrain en explorant le web avec Hyphe

Benjamin Ooghe Tabanou, Viviane Le Hay

Atelier, jeudi 15 octobre 9h30-12h45

Présentation générale

Le web nous oppose des défis à la fois méthodologiques et technologiques. Le médialab de Sciences Po a développé, et publié sous la forme d'un logiciel libre, Hyphe, un robot amasseur de données web aussi appelé «crawler», ajusté aux besoins de la recherche en sciences sociales. Il s'adresse aux sociologues qui veulent investiguer le web comme terrain d'enquête qualitative et en tirer des indicateurs quantitatifs. S'appuyant sur le modèle du web «en couches», il guide son utilisateur pour offrir aux chercheurs et étudiants un outil de création et nettoyage, itération après itération, un corpus de ressources et/ou d'acteurs.

La séance proposera tout d'abord une présentation générale de la méthode et de l'outil, ainsi qu'une rapide démonstration d'usage de Hyphe avant une mise en pratique par les participants.

Environnement informatique

Un ordinateur personnel avec un navigateur web connecté à Internet.

Type de données traitées - droit d'accès

Métadonnées sur des sites web et réseaux de liens entre les pages de ces sites. Libre accès.

Niveau requis

Tout public, simple connaissance des bases de fonctionnement du web préférable.

Objectifs

Découverte et utilisation exploratoire de l'outil Hyphe avec réalisation d'un premier corpus personnel sur la thématique de son choix.

Formule pédagogique

Travaux pratiques en petits groupes sur le logiciel après un exemple d'usage.

Références bibliographiques

- OOGHE-TABANOU, Benjamin, JACOMY, Mathieu, GIRARD, Paul & PLIQUE, Guillaume, "Hyperlink is not dead!", In Proceedings of the 2nd International Conference on Web Studies (WS.2 2018), Everardo Reyes, Mark Bernstein, Giancarlo Ruffo, and Imad Saleh (Eds.). ACM, New York, NY, USA, 12-18. DOI: <https://doi.org/10.1145/3240431.3240434>
- JACOMY, Mathieu, GIRARD, Paul, OOGHE-TABANOU, Benjamin, et al, "Hyphe, a curation-oriented approach to web crawling for the social sciences.", in International AAAI Conference on Web and Social Media. Association for the Advancement of Artificial Intelligence, 2016.

Liens

<http://hyphe.medialab.sciences-po.fr/demo/>

<https://github.com/medialab/hyphe-browser/releases/tag/v1.1>