

PRÉSENTATION ÉTUDE EN SYSTÈME D'INFORMATION ET PROGRAMMATION WEB

Par Mariette Yaméogo et Hélénya Jacquet

LES SITES D'ENTRÉES UTILISÉS ET LES ÉTAPES DU CRAWL

- Sujet de l'étude : Le laboratoire Paragraphe de l'Université Paris 8
- 3 étapes de crawl :
 - **Première étape** : 3 pages en lien avec notre sujet : la page du laboratoire paragraphe sur le site de Paris 8, la page d'accueil du site du laboratoire Paragraphe et la page du laboratoire sur les archives ouvertes HAL + 3 pages communes avec les autres groupes : la page d'accueil de l'agence nationale de la recherche, la page de la commission européenne consacrée à la recherche et à l'innovation et la page de financement, d'aides et de subventions du site de l'union européenne.
 - **Résultats** : Environ 3400 pages trouvées mais pas de résultat pour le site du laboratoire Paragraphe
 - **Deuxième étape** : Crawl des axes de recherche du laboratoire, des pages de conférences *web studies*, *existences collectives* et *H2PTM* et la page d'accueil de la chaire de l'*UNESCO*.
 - **Résultats** : 129 liens trouvés pour le site de la chaire de l'*UNESCO*, mais pas de résultats pour les pages présentes sur le site du laboratoire :



laboratoire-paragraphe.Univ-Paris8.fr

Programmé hier

Fini en 7 minutes

🔍 Profondeur 3

INFRUCTUEUX

0 page explorée

| NOM DE L'ENTITÉ WEB | Statut | PI | PC ✓ | PC X | PROGRAMMÉ | DURÉE |
|---|---------------|------|------|------|-------------------|-----------|
| Francearchives.gouv.fr /.../34638 | 3 ATTEINT | 1 | 1 | 0 | 81 86 hier | 7 minutes |
| laboratoire-paragraphe.Univ-Paris8.fr /.../Dispositifs... | 3 INFRUCTUEUX | 0 | 0 | 0 | 0 0 hier | 2 heures |
| laboratoire-paragraphe.Univ-Paris8.fr /.../colloque-i... | 3 INFRUCTUEUX | 0 | 0 | 0 | 0 0 hier | 2 heures |
| Mshparisnord.fr | 3 ATTEINT | 1360 | 1312 | 48 | 13549 131655 hier | 2 heures |
| laboratoire-paragraphe.Univ-Paris8.fr /.../Mediation... | 3 INFRUCTUEUX | 0 | 0 | 0 | 0 0 hier | 13 heures |

Collaborations académiques avec des institutions culturelles

Le laboratoire Paragraphe entretient des relations avec de nombreuses institutions au niveau national. Ces relations se traduisent par l'organisation de manifestations et des réponses à des projets de recherche (voir Liste Évènements) : Centre Pompidou (IRCAM, IRI), BNF, INHA, MSH-Plaine Saint- Denis, INA, Réunion des Musées Nationaux (Grand Palais), direction des Musées de France, Ministère de la Culture, Centre des Monuments Historiques, Archives nationales etc...

Source: <https://laboratoire-paragraphe.univ-paris8.fr/bienvenue/ecosysteme-recherche/article/collaborations-academiques-avec-des-institutions-culturelles>

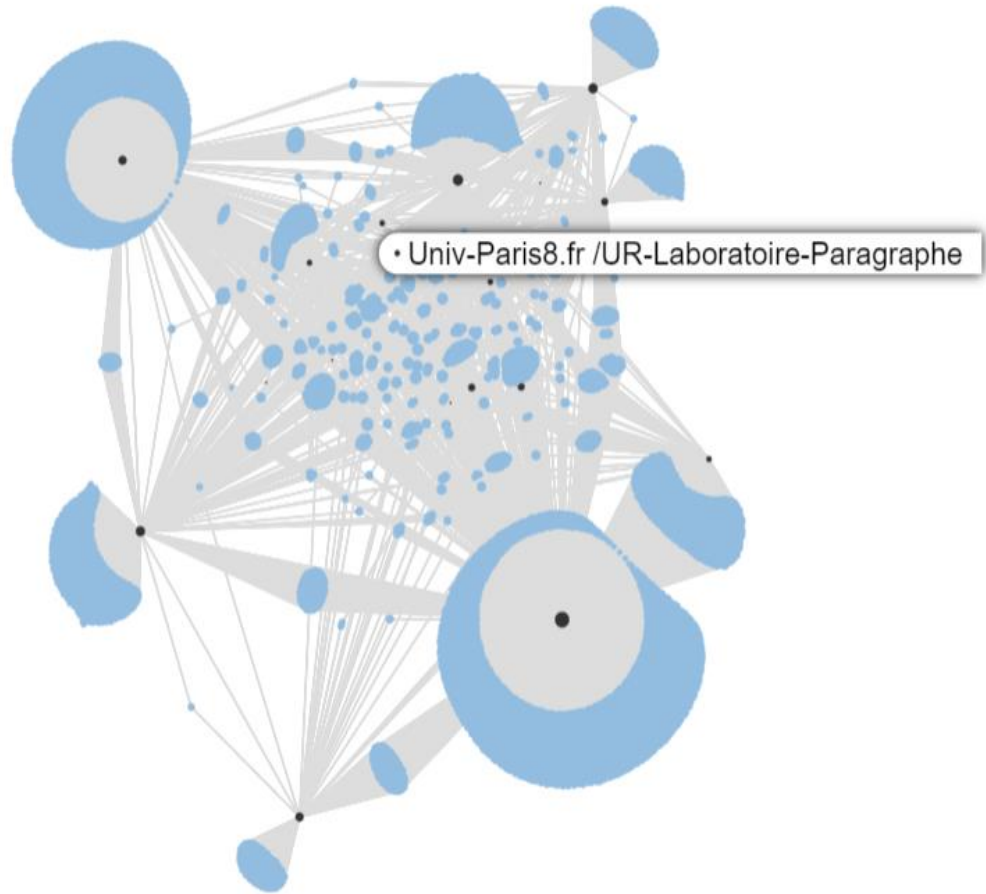
- **Troisième étape** : Crawl des liens de la page d'accueil de chaque site partenaire du laboratoire (indiqués sur la page des collaborations). Les partenaires sont mentionnés au nombre de 11.
- **Résultats** : 32666 pages trouvées
- **Bilan** : Peu de liens de sites extérieurs sont présents sur le site du laboratoire et notamment aucuns liens vers les partenariats mentionnés dans l'écosystème de recherche : nuit à la visibilité des collaborations.

LE PROCESSUS DE NETTOYAGE

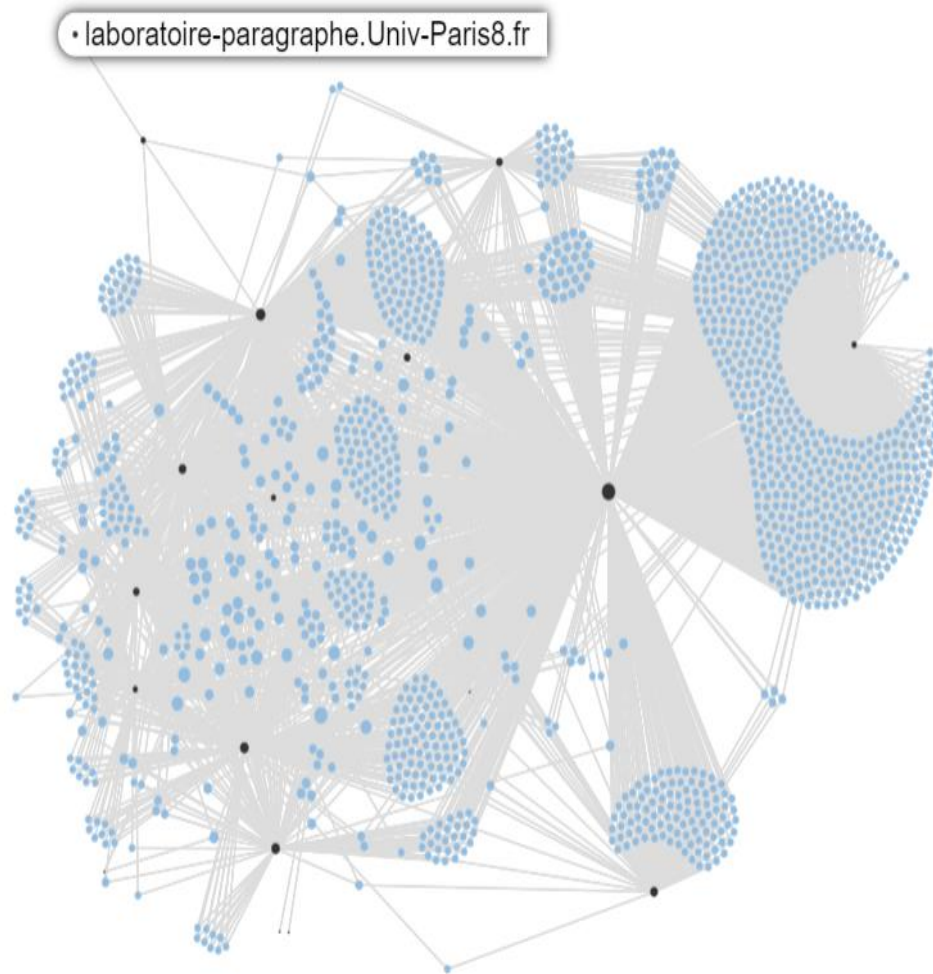
| <input type="checkbox"/> | DANS 27 | INDÉCIS 0 | <input checked="" type="checkbox"/> EN DEHORS 196 | <input type="checkbox"/> DÉCOUVERT 7 952 | | |
|--------------------------|----------------------|-----------|---|--|-----------------------------|---------------|
| Rechercher | | | | | APPLIQUER LES MODIFICATIONS | ANNULER |
| <input type="checkbox"/> | NOM ↓ | ST é | C é | SGP é | CITE é | DERNIER MOD é |
| <input type="checkbox"/> | 01Net.com | EN D... | | 1 | 1 | hier |
| <input type="checkbox"/> | Ajaccio-Tourisme.com | EN D... | | 1 | 1 | hier |
| <input type="checkbox"/> | Ain-Tourisme.com | EN D... | | 2 | 2 | hier |
| <input type="checkbox"/> | Aladin.co.kr | EN D... | | 2 | 1 | hier |
| <input type="checkbox"/> | Adobe.com | EN D... | | 19 | 3 | hier |

- **Liens retirés** : réseaux sociaux, grandes plateformes (Adobe, Wordpress, YouTube, Vimeo, etc.), sites de streaming musical, sites de voyage et touristiques

ANALYSE DES GRAPHES SUR HYPHE

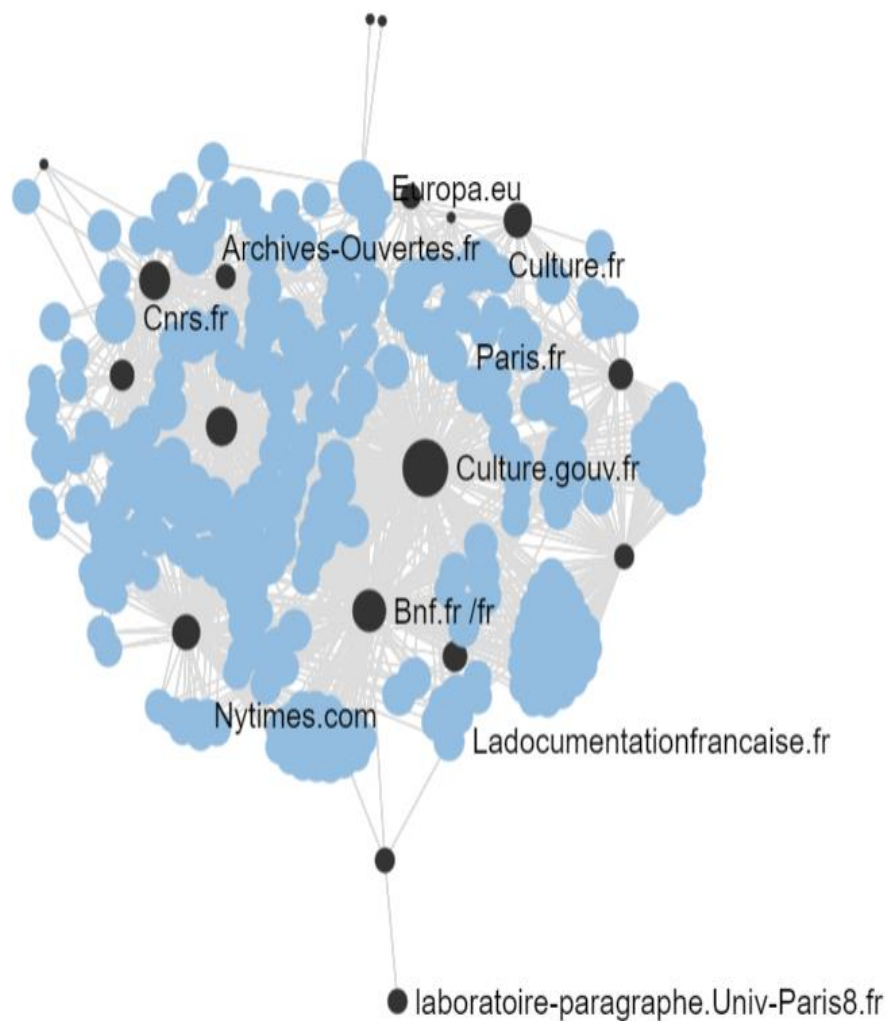


Notre hypothèse se confirme également à la lecture du graphe.
le site du laboratoire paragraphe ne fait pas partie des nœuds principaux, il n'est pas relié a beaucoup de liens comparativement aux sites comme culture. Gouv, l'institut national de l'audiovisuel, la BNF, l'ANR, le musée des sciences de l'homme, les monuments nationaux et l'Ircam.



De plus, lorsque l'on filtre le graphe avec les sites qui sont reliés à au moins 2 liens, on remarque que le site du laboratoire est seulement relié à la page consacrée à ce dernier sur le site de l'université Paris 8. C'est ensuite au deuxième niveau de profondeur qu'on voit qu'il est indirectement relié à la BnF, au site de la documentation française et au site de l'université Paris 8. On peut donc noter le manque d'exhaustivité au niveau du nombre de liens auquel il est rattaché

On remarque également sur ce graphe que de nouveaux nœuds ressortent avec de nouveaux sites mis en avant tels que le site de la réunion des musées nationaux et du grand palais des champs Elysées, le site de l'institut de recherche et d'innovation du centre Georges Pompidou, le site de la chaire de l'UNESCO, la page des musées nationaux présente sur le site du gouvernement et plus précisément au niveau de la page du ministère de la culture et également sur le site de l'institut national de l'histoire de l'art.



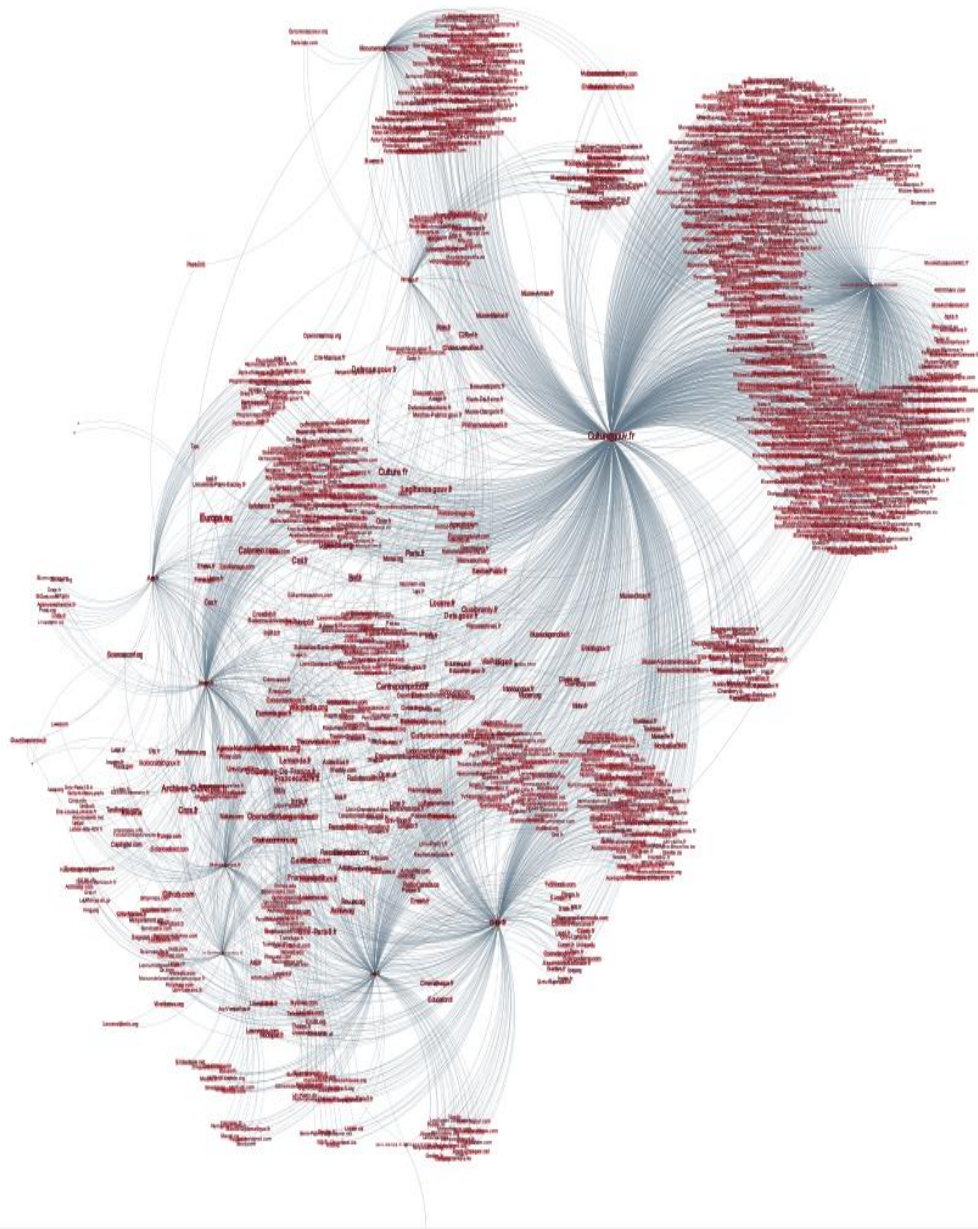
Lorsqu'on filtre selon les sites qui sont reliés à au moins 3 liens, on peut voir tout d'abord que le site du laboratoire paragraphe est à part des autres sites puisqu'il est relié comme dit précédemment qu'à la page qui lui est dédiée sur le site de Paris 8.

On remarque également que de nouveaux sites ressortent comme le site de presse New York Times, le site de la ville de Paris, celui du CNRS, le site des archives ouvertes HAL, le site du ministère de la culture et le site de l'union européenne.

ANALYSE DES GRAPHS SUR GEPHY



Ce graphe présente différentes teintes de “vert” pour mettre en avant les pages qui sont reliées à un plus grand nombre de liens. Plus une page est reliée à des liens plus la couleur verte du point qui représente la page est foncée.



Ce graphe nous permet de visualiser le nom des liens présents dessus et de faire ressortir à l'aide de la couleur rouge les amas de liens situés autour des différents nœuds ce qui permet ainsi de les mettre en valeur.

ETAPES D'UTILISATION DU LOGICIEL GITHUB

- Exportation du graphe réalisé sur *Hyphe* aux formats « .md » « .gexf » et « .json » dans l'espace « data » du logiciel en respectant les normes de nommage informatique.
- Création de pages de code avec les langages de programmation HTML + CSS + Javascript
- Visualisation des graphes GEPHI à partir du code préalablement créé
- Retrouvez ci-joint :
 - Notre page GitHub mise en forme : https://samszo.github.io/M2GSI_22-23/pages/Mariette_et_Helenya.html

MERCI DE VOTRE
ÉCOUTE!



UNIVERSITÉ
PARIS8
VINCENNES-SAINT-DENIS