Projet LIASD

Présenté par:

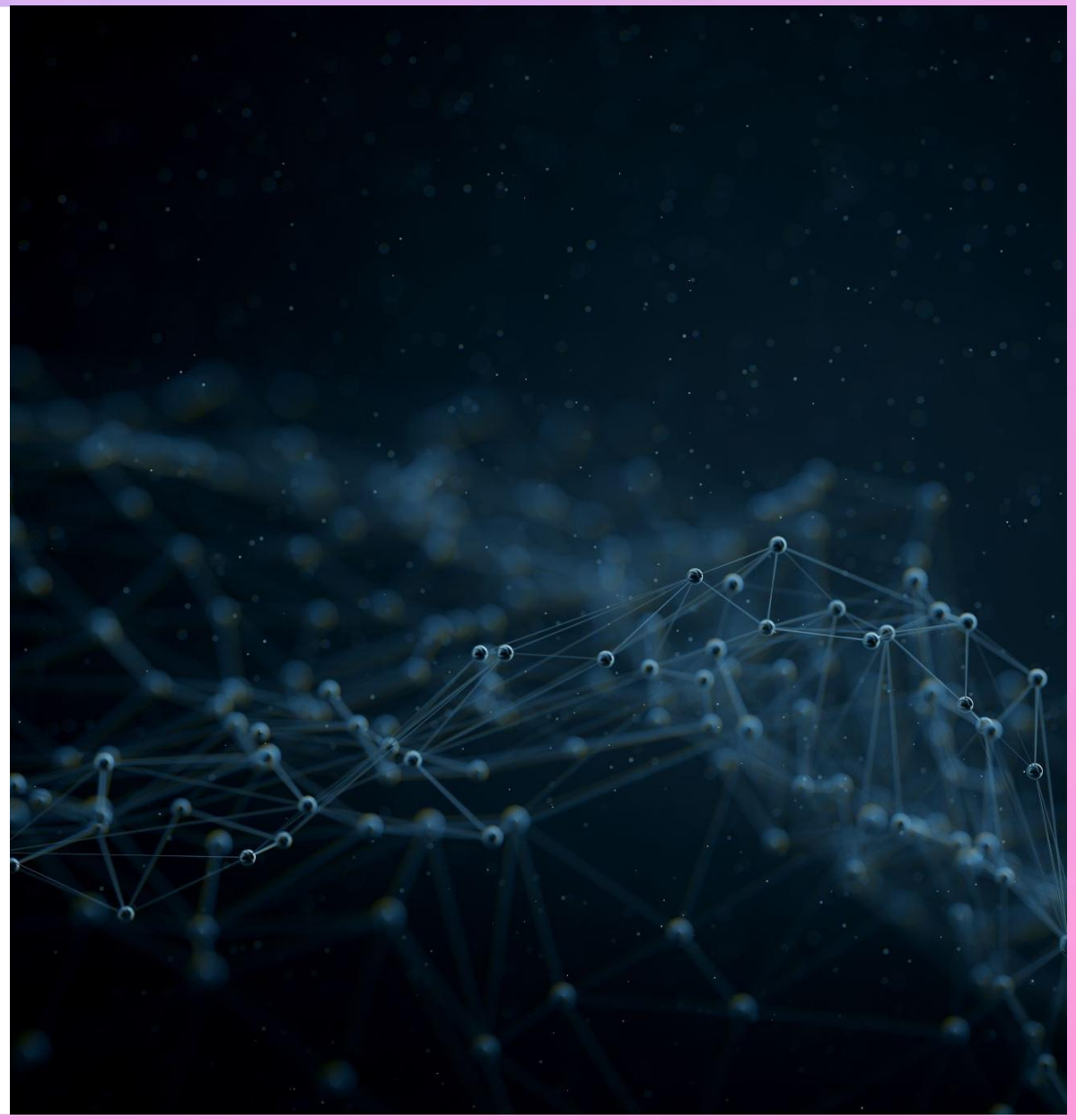
Lynda Haddar et
Rouguiyatou Ndoye

Plan

- Présentation de LIASD
- Les sites d’entrée utilisés
- Le déroulement du Crawl
- Le processus de nettoyage
- La consultation du graphe et de ses différents paramétrages
- Ce que vous avez pu faire sur Gephi.
- Les éventuelles questions et difficultés

Présentation de LIASD

Le LIASD (laboratoire d’intelligence artificielle et sémantique des données)ressemble des enseignants chercheurs et doctorants en informatique. Le laboratoire est composé de trois axes



Les axes de LIASD

EID(Espaces intelligents des Données est spécialisé en ingénierie des données massives.

IUSD(Informatique Ubiquitaire et science de données) est membre de LIASD. Ses travaux visent le développement d’algorithmes et des méthodologies pour les applications centrées sur l’humain

PASTIS(Programming, artificial, intelligence, security, texts, images, simulation)

Les sites d’entrée utilisés

L'IEEE(l’institut des ingénieurs électriciens et électroniciens) est la plus grande organisation professionnelle technique au monde dédiée à l'avancement de la technologie au profit de l'humanité

ACM, la plus grande société informatique éducative et scientifique au monde, fournit des ressources qui font progresser l'informatique en tant que science et profession

Hal est une plateforme en ligne développée par la communication scientifique du CNRS. Nous avons ajouté à nos sources la page Hal du laboratoire LIASD

L'Agence nationale de la recherche (ANR) est un établissement public à caractère administratif, placé sous la tutelle du ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation.

La commission européenne finance des projet de recherches

Le site de PASTIS qui est l’un des membres du laboratoire LIASD

Déroulement du Crawl

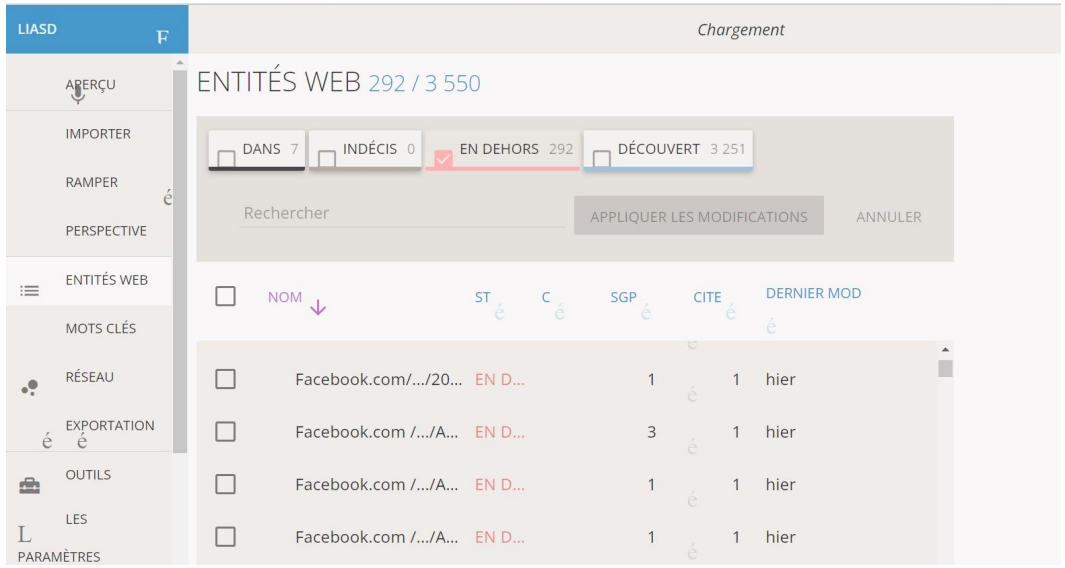
Nous avons créé notre projet sur Hyphe sous le nom de LIASD, nous avons mis les sites entrée

Le résultat du crawl

Après cette étape, nous avons paramétré le crawleur à 2 niveaux avant de lancer le crawl. Hyphe à crawler 3421 liens.

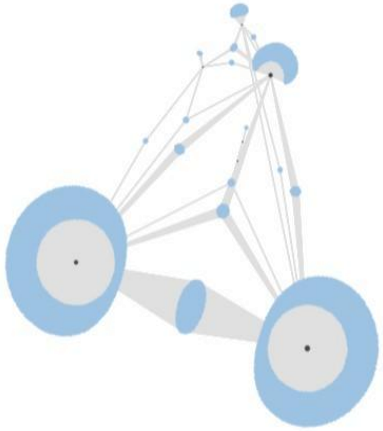
Le processus de nettoyage

- Supprimer tous les liens liés aux GAFAM, et liens qui nous mènent vers certaines pages de réseaux sociaux comme Instagram, Facebook, twitter, YouTube, télégramme.
- Déplacer vers la ” Case out” créant un total de 292



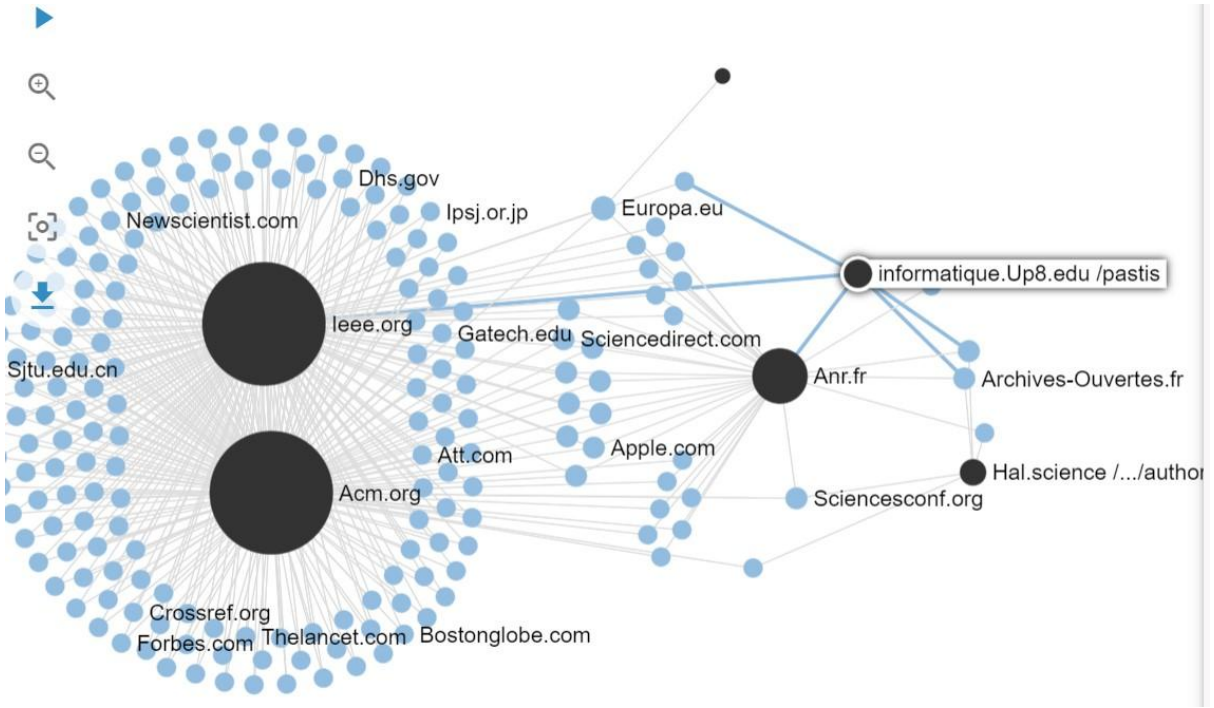
Consultation du graphique et ses différents paramétrage

le graphique numéro 1 qui est sans filtre. Il contient l’ensemble des liens qui sont dans IN, Undecided et Discovered



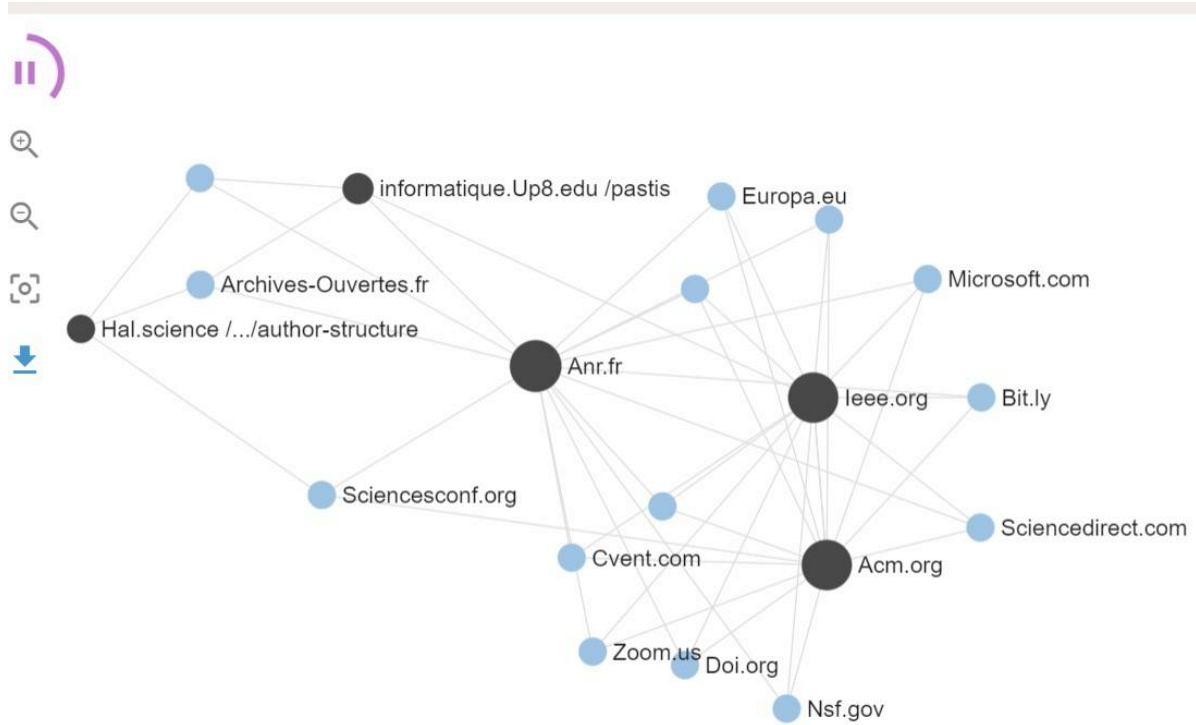
graphique et ses différents paramétrage

Sur ce deuxième graphique, nous avons appliqué des filtres : DANS, INDECIS, DECOUVERT, Afficher uniquement les entités web avec 2+liens et Indegree

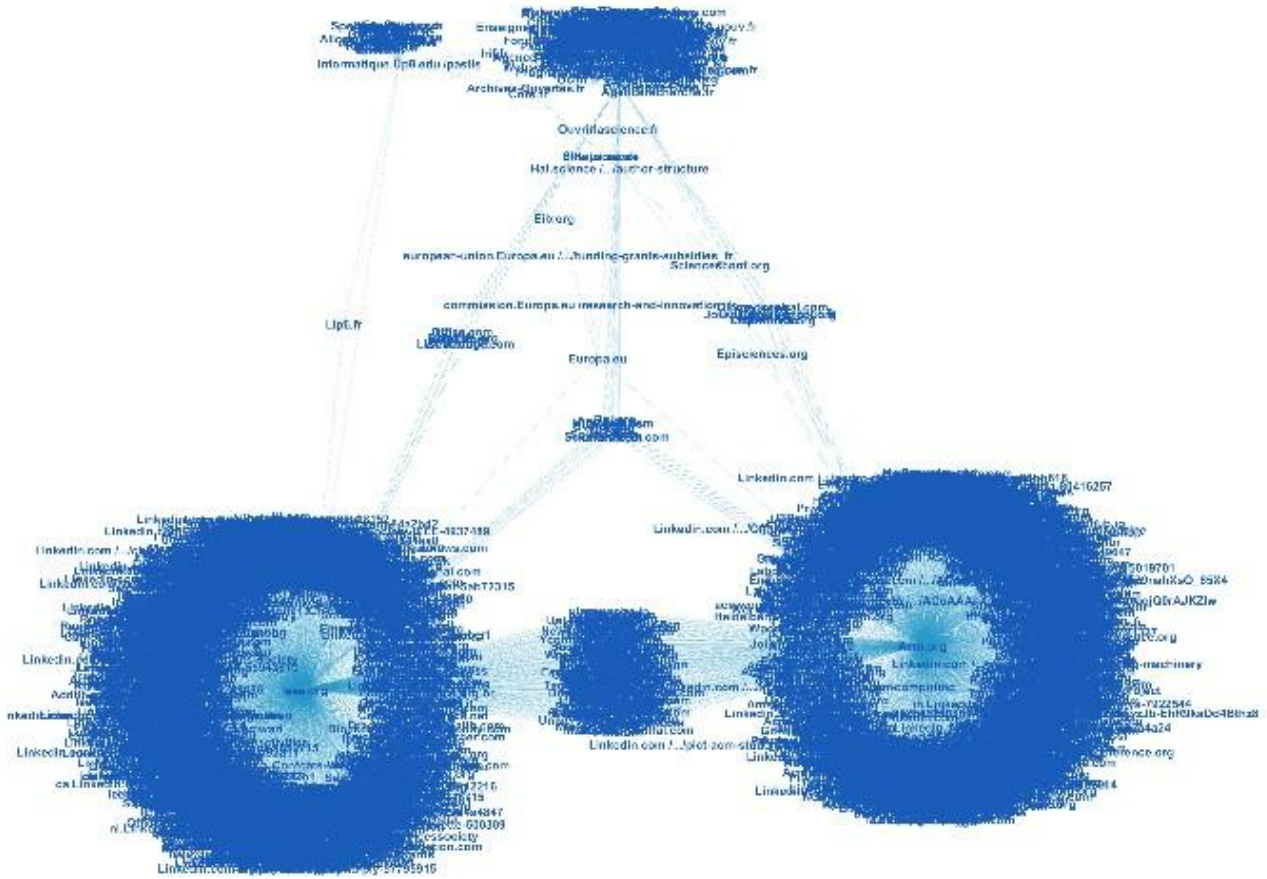


graphique et ses différents paramétrage

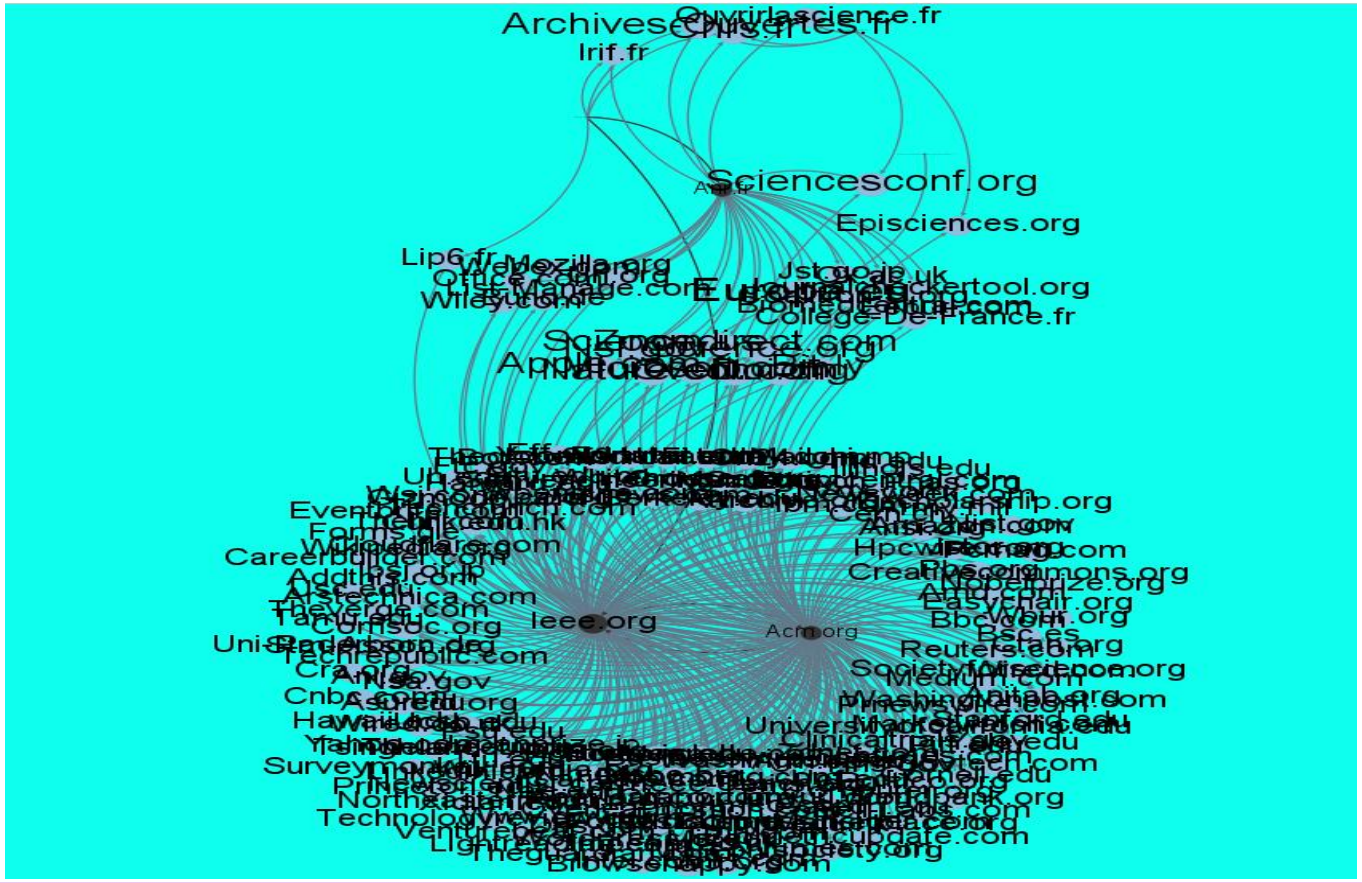
Nous avons également affiné notre graphique en filtrant par: DANS, INDECIS, DECOUVERT, Afficher uniquement les entités web avec 3+liens et Indegree
Voici le résultat obtenu



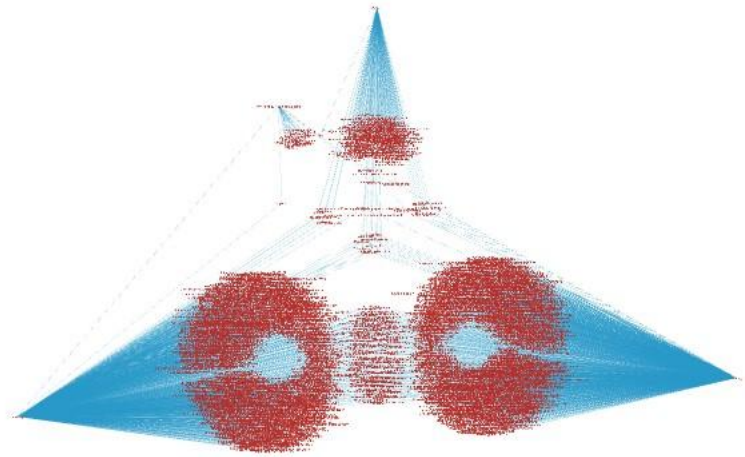
Cartographie sur Gephie



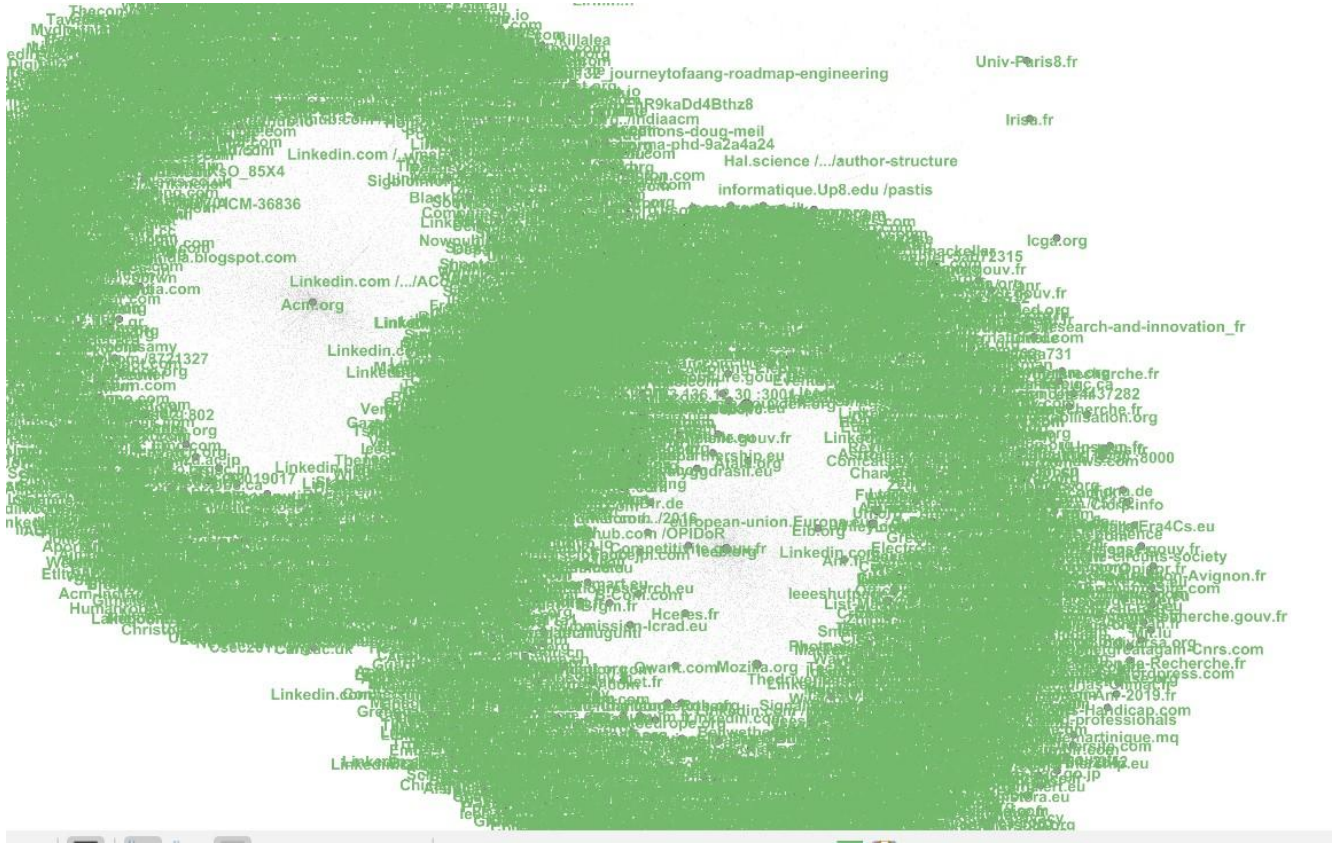
Cartographie sur Gephie



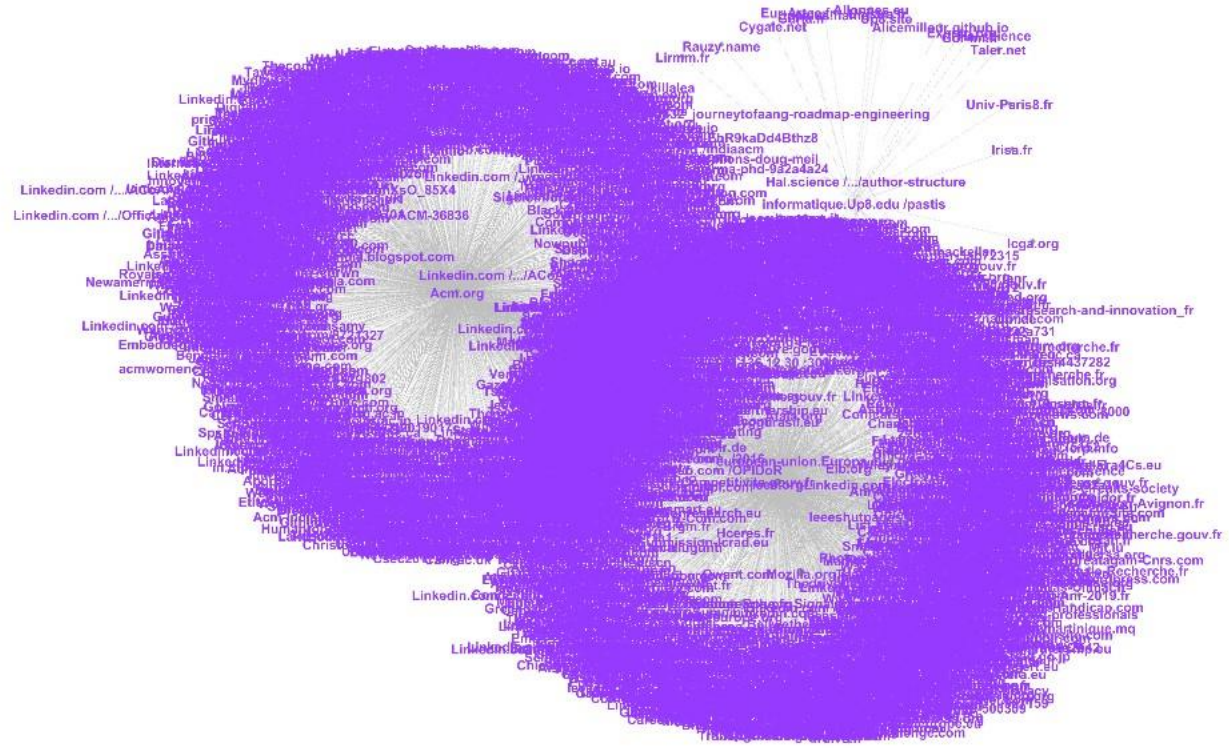
Cartographie sur Gephie



Cartographie sur Gephie



Cartographie sur Gephie



Les difficultés rencontrés

- URL qui bloquent au moment de crawler
- Problème de manipulation de Gephie

Lien vers notre site

https://samszo.github.io/M2GSI_22-23/pages/LIASD.html

Merci de votre attention