

Projet Analyse des réseaux

HAMMOUCH Siham

1 Introduction

Dans ce projet, je vais étudier l'évolution d'un réseau de collaboration scientifique au sein d'un cluster de chercheurs, observé à deux moments : avant et après un choc.

Chaque chercheur est décrit par plusieurs attributs : son genre , son statut de membre "core" , ainsi que son éventuel rôle de coordinateur. Le genre est représenté par la variable *gender* qui vaut 1 si la personne est une femme. Son statut de membre "core" est lui représenté par la variable *core*, qui vaut 1 si l'individu a été identifié avant traitement comme un membre central du cluster dans l'espace d'expertise. L'espace d'expertise désigne le domaine scientifique spécifique sur lequel porte l'activité du cluster (par exemple une spécialité de recherche). Être central dans cet espace signifie que l'individu joue un rôle clé par son savoir ou ses travaux dans ce domaine, indépendamment de sa position dans le réseau de collaborations. Pour finir, son éventuel rôle de coordinateur est illustré par la variable *coord*, qui vaut 1 si la personne est désignée comme coordinateur du cluster. Être coordinateur n'implique pas nécessairement que la personne soit présente dans les données, car il arrive que le coordinateur ne figure pas dans la base fournie.

Les liens du réseau représentent des collaborations scientifiques actives entre chercheurs.

Ce projet examine comment un choc collectif influence l'organisation d'un réseau de collaborations scientifiques. Il s'agit d'analyser l'évolution des positions individuelles, en particulier pour les femmes et les membres du "core", et d'observer si le traitement tend à renforcer ou réduire les inégalités de centralité et la segmentation du réseau.

Dans une première partie, je vais analyser l'évolution des positions centrales individuelles. J'étudierais ensuite le rôle du statut *core* dans la dynamique du réseau. Enfin, nous explorerons la structuration en sous-communautés et l'homophilie de genre.

2 Visualisation générale des réseaux

Avant d'analyser en détail les mesures de centralité, je présente une visualisation globale du réseau de collaborations scientifiques, avant et après le choc.

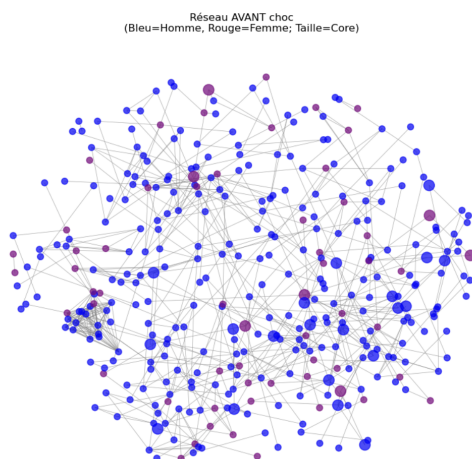


Figure 1: Graphe des collaborations avant choc (genre & coordination)

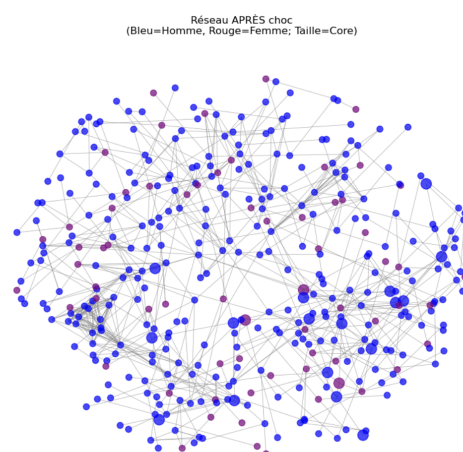


Figure 2: Graphe des collaborations après choc (genre & coordination)

La visualisation du réseau avant choc montre que les femmes (nœuds violets) sont réparties dans tout le graphe, mais apparaissent souvent en position périphérique. Peu de femmes occupent les zones centrales les plus denses.

Après choc, la structure globale du réseau reste assez similaire : le graphe reste peu connecté, mais j'observe quelques zones où les liens se densifient légèrement. Les femmes et les hommes sont toujours répartis de manière relativement homogène, avec des coordinatrices et coordinateurs visibles par une taille de nœuds plus conséquente. Certains individus, femmes ou hommes, semblent avoir gagné en visibilité, ce que j'approfondirais dans l'analyse des centralités.

Nombre de liens avant choc : 607
Nombre de liens après choc : 639

Figure 3: Évolution du nombre de liens

Le nombre total de collaborations actives passe de 607 avant choc à 639 après choc. Cette légère augmentation traduit un élargissement modéré du réseau.

3 Évolution de la centralité des individus avant et après choc

Dans cette première partie, j'analyse comment les positions centrales des chercheurs évoluent à travers le réseau avant et après le choc. Je mobilise trois indicateurs classiques de centralité : le degré (nombre de connexions directes), la proximité (closeness) et l'intermédierité (betweenness).

3.1 Centralité de degré

Le degré mesure pour chaque chercheur le nombre de connexions directes qu'il entretient dans le réseau. Après calcul, je constate que les scores moyens de degré sont relativement faibles,

traduisant un réseau peu dense. J'analyse également quels chercheurs occupent les positions les plus connectées avant et après le choc.

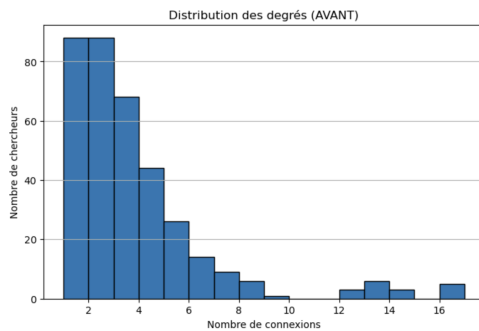


Figure 4: Distribution des degrés avant choc

Degré moyen AVANT : 3.36
Femmes : 3.25 – Hommes : 3.39

Figure 5: Degré moyen avant choc par genre

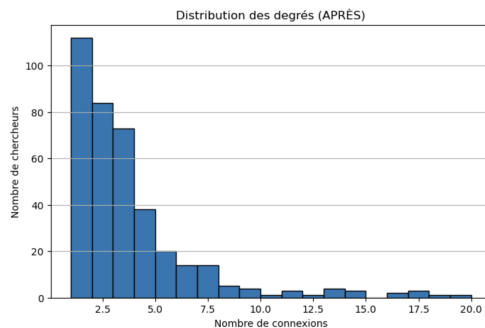


Figure 6: Distribution des degrés après choc

Degré moyen APRÈS : 3.34
Femmes : 2.67 – Hommes : 3.48

Figure 7: Degré moyen après choc par genre

Avant choc, la distribution des degrés est fortement concentrée autour de 2 à 4 connexions, avec très peu de chercheurs hautement connectés. Après choc, la forme générale de la distribution est conservée, mais le pic autour des faibles degrés devient plus marqué, et quelques individus atteignent des degrés plus élevés (jusqu'à 20). Cela traduit une accentuation de l'asymétrie du réseau : une majorité de chercheurs restent faiblement connectés, tandis qu'une minorité consolide des positions très centrales. Le choc semble ainsi creuser les écarts de connectivité entre chercheurs, en concentrant davantage les liens autour d'un petit nombre d'individus très connectés.

Du côté du genre, les femmes avaient déjà un degré moyen un peu plus faible que les hommes avant le choc (3.25 contre 3.39). Après traitement, cet écart devient plus marqué : le degré moyen des femmes baisse à 2.67, alors qu'il augmente chez les hommes (3.48). Cela montre que les hommes ont élargi leur réseau de collaborations, tandis que les femmes ont, en moyenne, perdu des connexions. Le choc semble donc avoir renforcé les différences entre les deux groupes.

3.2 Centralité de proximité

Après avoir étudié la connectivité directe des chercheurs, j'analyse leur accessibilité globale dans le réseau à travers la centralité de proximité (closeness). Cette mesure évalue à quelle distance moyenne un chercheur se trouve des autres membres du réseau : plus un individu est proche de tous les autres, plus sa valeur de closeness est élevée.

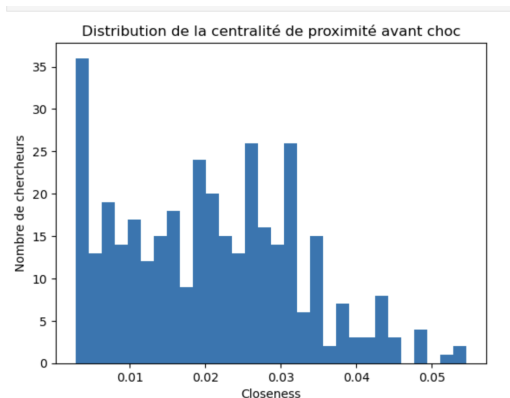


Figure 8: Distribution de la centralité de proximité avant choc

Moyenne (avant) femmes : 0.0226 – hommes : 0.0202
 Top 5 :
 ID 11194 – Closeness : 0.0546
 ID 47723 – Closeness : 0.0543
 ID 11193 – Closeness : 0.0524
 ID 47635 – Closeness : 0.0494
 ID 47611 – Closeness : 0.0491

Figure 9: Centralité de proximité moyenne avant choc par genre

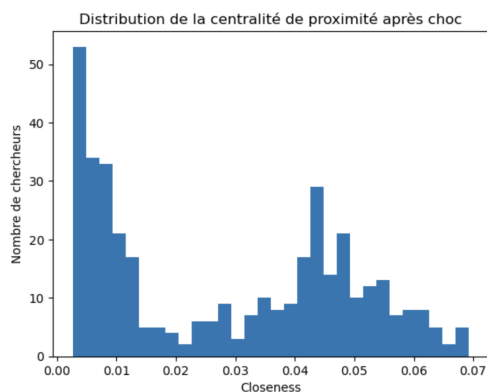


Figure 10: Distribution de la centralité de proximité après choc

Top 5 – Centralité de proximité (après) :
 ID 47624 – Closeness : 0.0693
 ID 47685 – Closeness : 0.0693
 ID 77973 – Closeness : 0.0690
 ID 47724 – Closeness : 0.0687
 ID 47720 – Closeness : 0.0673
 Closeness moyenne Femmes : 0.0246
 Closeness moyenne hommes : 0.0291

Figure 11: Centralité de proximité moyenne après choc par genre

Avant le choc, la centralité de proximité est relativement équilibrée entre chercheurs, avec une majorité d'individus présentant une accessibilité modérée. Après le choc, on observe une légère augmentation de la proximité pour certains chercheurs, traduisant une amélioration de l'accessibilité globale pour une minorité, tandis que d'autres restent en périphérie du réseau.

Concernant les différences de genre, avant choc, les femmes présentaient une centralité de proximité moyenne légèrement supérieure à celle des hommes. Après choc, cet avantage s'inverse : les hommes enregistrent une proximité moyenne plus élevée que les femmes, traduisant un repositionnement plus périphérique des femmes dans la structure du réseau.

Enfin, l'analyse du Top 5 des chercheurs les plus centraux montre un renouvellement complet après le choc : les individus ayant les plus fortes accessibilités ne sont plus les mêmes qu'avant choc. Cela suggère que le choc a profondément redistribué les rôles centraux dans le réseau, en permettant à de nouveaux chercheurs de gagner en visibilité.

3.3 Centralité d'intermédiation

Après avoir étudié la connectivité directe (degré) et l'accessibilité globale (proximité), j'analyse la centralité d'intermédiation (betweenness). Cette mesure indique dans quelle position se trouve un chercheur par rapport aux chemins les plus courts entre les autres membres du réseau. Plus elle est élevée, plus l'individu joue un rôle de passeur stratégique dans la circulation de l'information.

Moyenne femmes : 0.0013 – hommes : 0.0009
 Top 5 :
 ID 11194 – Betweenness : 0.0237
 ID 47723 – Betweenness : 0.0209
 ID 11193 – Betweenness : 0.0203
 ID 22031 – Betweenness : 0.0187
 ID 11191 – Betweenness : 0.0121

Figure 12: Distribution de la centralité d'intermédiation avant choc

Top 5 – Centralité d'intermédiation (après) :
 ID 47685 – Betweenness : 0.1171
 ID 47724 – Betweenness : 0.1148
 ID 47624 – Betweenness : 0.1137
 ID 77973 – Betweenness : 0.1137
 ID 47717 – Betweenness : 0.1080
 Betweenness moyenne Femmes : 0.0028
 Betweenness moyenne hommes : 0.0062

Figure 13: Top 5 et moyennes par genre de la centralité d'intermédiation après choc

Avant choc, la majorité des chercheurs présente une centralité d'intermédiation très faible, avec des valeurs maximales autour de 0.02, traduisant un réseau relativement distribué sans forte dépendance à quelques passeurs. Après choc, quelques chercheurs acquièrent un rôle d'intermédiaire beaucoup plus marqué, avec des scores de betweenness dépassant 0.11, tandis que la majorité reste faiblement positionnée.

Concernant les différences de genre, avant choc, les femmes présentaient une centralité d'intermédiation moyenne légèrement supérieure à celle des hommes (0.0013 contre 0.0009). Après choc, cette tendance s'inverse : les hommes enregistrent une betweenness moyenne deux fois supérieure à celle des femmes (0.0062 contre 0.0028), traduisant un retrait relatif des femmes des positions stratégiques du réseau.

Enfin, l'analyse du Top 5 des chercheurs les plus intermédiaires révèle un renouvellement important après choc : les individus occupant les rôles d'intermédiaires stratégiques ne sont plus les mêmes qu'avant traitement. Cela suggère que le choc a profondément redistribué les rôles d'intermédiation, avec l'émergence de nouveaux acteurs clés au sein du réseau.

3.4 Synthèse sur l'évolution des centralités

L'analyse des trois mesures de centralité montre que le choc a renforcé certaines inégalités dans le réseau de collaboration scientifique.

Concernant la centralité de degré, la majorité des chercheurs reste faiblement connectée, mais quelques individus concentrent davantage de liens après choc. La centralité de proximité révèle que certains chercheurs ont amélioré leur accessibilité dans le réseau, tandis que

d'autres se sont marginalisés. La centralité d'intermédiation indique que de nouveaux individus acquièrent des rôles très stratégiques pour la circulation de l'information, traduisant une redistribution partielle des rôles au sein du réseau.

En termes de genre, avant choc, les femmes présentaient des niveaux de centralité comparables, voire légèrement supérieurs à ceux des hommes en termes de proximité et d'intermédiation. Après choc, elles voient leur position centrale reculer, en particulier sur les dimensions stratégiques du réseau (intermédiation), où les hommes deviennent plus dominants. Cela suggère que le choc a contribué à creuser les écarts de visibilité et d'accessibilité, en particulier au détriment des chercheuses.

Mon analyse met aussi en évidence un phénomène classique connu sous le nom de friendship paradox : en moyenne, les chercheurs les plus centraux du réseau ont davantage de connexions que leurs voisins. Ce paradoxe s'explique ici par la structure même du réseau : la majorité des chercheurs ont peu de connexions, tandis qu'une minorité concentre un grand nombre de liens. Après le choc, cette asymétrie s'accroît, car les liens se reforment surtout autour de ces individus très connectés. Ce phénomène souligne que l'inégalité d'accès aux collaborations est structurelle, et tend à se renforcer, concentrant les interactions autour d'un nombre limité de chercheurs.

4 Analyse du rôle du statut Core / Non-Core

Après avoir analysé l'évolution des positions centrales individuelles, je m'intéresse maintenant à l'impact du choc sur les collectifs structurants du réseau. Je distingue pour cela les chercheurs appartenant au core du réseau de ceux dits non-core.

4.1 Définition du statut Core

Le core désigne l'ensemble des chercheurs les plus centraux et les plus actifs du réseau avant le choc.

Analyser séparément les trajectoires des membres du core et des non-core permet de mieux comprendre comment un choc collectif impacte la hiérarchie sociale préexistante dans un réseau scientifique.

4.2 Comparaison Core vs Non-Core avant choc

Je compare ici les positions structurelles des chercheurs appartenant au core et des non-core avant le choc. J'analyse plusieurs indicateurs : le degré moyen, la centralité de degré, la centralité de proximité (closeness) et la centralité d'intermédiation (betweenness).

Degré moyen Core	: 3.11//	Non-core	: 3.38
Centralité degré Core	: 0.0086//	Non-core	: 0.0094
Closeness Core	: 0.0221//	Non-core	: 0.0205
Betweenness Core	: 0.0032//	Non-core	: 0.0008

Figure 14: Comparaison des indicateurs de centralité Core vs Non-Core avant choc

Avant le choc, les membres du core présentent une accessibilité globale (closeness) et une centralité d'intermédiation (betweenness) supérieures à celles des non-core, traduisant des

positions plus stratégiques pour la circulation de l'information. En revanche, leur degré moyen est légèrement inférieur à celui des non-core, indiquant que les membres core entretiennent moins de connexions directes mais occupent des rôles plus centraux dans la structure du réseau. Cependant, l'écart global reste modéré : la différence entre core et non-core existe mais n'est pas extrêmement marquée, traduisant un réseau relativement ouvert avant choc, où la hiérarchie sociale n'est pas encore complètement rigide.

4.3 Comparaison Core vs Non-Core après choc

Après le choc, j'observe une évolution plus marquée dans les positions des chercheurs du core et du non-core. J'utilise les mêmes indicateurs qu'avant choc pour évaluer ces différences.

Degré moyen Core : 2.65// Non-core : 3.37
 Centralité degré Core : 0.0069// Non-core : 0.0088
 Closeness Core : 0.0303// Non-core : 0.0282
 Betweenness Core : 0.0034// Non-core : 0.0057

Figure 15: Comparaison des indicateurs de centralité Core vs Non-Core après choc

Après le choc, les membres du core voient leur accessibilité globale (closeness) légèrement augmenter, traduisant une meilleure position en termes de proximité dans le réseau. Cependant, leur degré moyen diminue, et leur rôle d'intermédiaire (betweenness) devient inférieur à celui des non-core, indiquant une perte relative d'influence stratégique dans la circulation de l'information.

Cette évolution montre que si les core restent mieux accessibles, ils ne concentrent plus les connexions ni les positions d'intermédiaires aussi fortement qu'avant traitement, contrairement aux non-core qui gagnent du terrain.

4.4 Synthèse sur l'impact du choc sur les statuts Core et Non-Core

Le choc a eu un impact différencié selon le statut des chercheurs :

- Avant choc, les chercheurs non-core avaient des degrés légèrement plus élevés que les core, mais une accessibilité et un rôle d'intermédiation plus faibles.
- Après choc, les chercheurs core perdent en nombre de connexions, améliorent leur accessibilité (closeness), mais se retrouvent devancés par les non-core en termes de centralité d'intermédiation (betweenness).

Cela suggère que le choc a partiellement redistribué l'importance stratégique au sein du réseau : les core conservent une meilleure accessibilité globale, mais les non-core renforcent leur rôle dans la circulation de l'information.

5 Structure du réseau : communautés et homophilie

Après avoir étudié les positions individuelles, j'analyse maintenant la structuration collective du réseau à travers deux aspects :

- la détection de sous-communautés (groupes fortement connectés),
- l'évaluation de l'homophilie de genre (tendance à collaborer avec des personnes du même genre).

5.1 Détection des sous-communautés

Pour identifier les sous-structures du réseau, j'utilise l'algorithme de Louvain, une méthode de clustering optimisant la modularité. La modularité mesure à quel point le réseau est divisé en communautés denses.

Modularité du clustering (AVANT) : 0.9124

Modularité du clustering (APRÈS) : 0.8954

Figure 16: Modularité du clustering avant choc

Figure 17: Modularité du clustering après choc

Avant le choc, la modularité du découpage atteint 0.9124, ce qui est très élevé : cela indique des communautés très fortement structurées. Après le choc, la modularité reste élevée (0.8954), mais diminue légèrement, traduisant une ouverture relative entre les communautés. Ainsi, malgré l'élargissement du réseau, celui-ci conserve une organisation interne forte, centrée sur des sous-groupes denses.

Le niveau très élevé de modularité observé (proche de 0.9) est en nette rupture avec ce que l'on attendrait d'un graphe aléatoire (Erdős-Rényi), qui présente en général une modularité très faible. Cela confirme que notre réseau est fortement structuré en communautés, et qu'il ne peut être assimilé à un réseau aléatoire classique.

Pour compléter l'analyse de la structure locale du réseau, nous avons calculé le coefficient de clustering moyen.

Avant choc, le clustering moyen s'établit à 0,472, traduisant une forte propension des chercheurs connectés à partager aussi des liens mutuels. Cela signifie qu'en moyenne, lorsqu'un chercheur est connecté à deux collègues, il existe dans près d'un cas sur deux une collaboration directe entre ces deux collègues.

Après choc, ce coefficient diminue légèrement à 0,418, indiquant une ouverture relative du réseau : les collaborations tendent à devenir un peu moins systématiquement triangulaires. Malgré cette évolution, le réseau conserve une structure locale relativement dense.

5.2 Homophilie de genre

J'étudie maintenant dans quelle mesure les chercheurs ont tendance à se connecter à d'autres chercheurs du même genre.

Je commence par observer la composition des communautés détectées :

Seuil	Période	Communautés très féminisées	Communautés très core
0.75	Avant choc	1	0
0.70	Avant choc	1	2
0.75	Après choc	0	0
0.70	Après choc	1	0

Table 1: Détection de communautés homogènes selon le genre ou le statut core, avant et après choc

Je remarque que la communauté très féminisée détectée avant choc (seuil 0,75) est de très petite taille (2 chercheurs), ce qui relativise son poids structurel dans le réseau. À l'inverse, les deux communautés très core détectées au seuil 0,7 avant choc regroupent respectivement 7 et 4 chercheurs, ce qui leur confère une importance plus notable dans la structuration locale du réseau.

L'observation de la composition des communautés montre qu'avant comme après choc, très peu de sous-communautés sont homogènes en genre ou en statut core. Même en abaissant le seuil d'homogénéité à 70 %, seules quelques micro-communautés apparaissent très féminisées ou très "core". Cela suggère qu'avant traitement, le réseau présentait déjà une organisation largement mixte, sans regroupement fort par genre ou statut, et que le choc n'a pas significativement renforcé l'homophilie locale.

Cependant, la forte modularité observée après choc (capacité du réseau à se diviser en sous-groupes denses) ne s'explique pas principalement par des regroupements de genre : même au sein des communautés, les liens de collaboration restent relativement mixtes.

5.3 Homophilie globale

Ensuite, pour évaluer l'homophilie à l'échelle de l'ensemble du réseau, je calcule le coefficient d'assortativité par genre.

```
nx.attribute_assortativity
```

```
-0.014745916515426677
```

Homophilie globale (genre)

Figure 18: Coefficient d'assortativité par genre

Le coefficient d'assortativité obtenu est de $-0,015$ (environ $-0,0147$), une valeur très proche de zéro. Cela indique qu'il n'existe ni tendance significative à collaborer avec des chercheurs du même genre (homophilie), ni tendance inverse (hétérophilie). Les collaborations scientifiques apparaissent ainsi largement indépendantes du genre, sans préférence particulière.

5.4 Synthèse sur la structuration du réseau

L'analyse des sous-communautés et de l'homophilie confirme que le réseau reste structurellement très mixte, malgré le choc. La forte modularité observée traduit une organisation en groupes bien définis, mais ces groupes ne sont pas exclusivement genrés ni marqués par le statut core.

Au niveau local (au sein des communautés) comme global (ensemble du réseau), le genre n'apparaît pas comme un facteur déterminant dans la formation ou l'évolution des liens de collaboration. Cela confirme que les collaborations restent ouvertes, sans préférence ni évitement lié au genre.

6 Conclusion

Ce projet a permis d'analyser l'évolution d'un réseau de collaborations scientifiques avant et après un choc collectif, en mobilisant des outils classiques d'analyse des réseaux.

Le nombre total de collaborations augmente légèrement après choc, traduisant un élargissement modéré du réseau. L'étude des centralités (degré, proximité, intermédiarité) montre que le choc a renforcé certaines inégalités : la majorité des chercheurs reste faiblement connectée, tandis qu'une minorité consolide des positions très stratégiques. On observe également une redistribution des rôles : de nouveaux chercheurs émergent au centre du réseau, notamment sur les mesures d'accessibilité et d'intermédiarité.

En termes de genre, les femmes, qui présentaient initialement des niveaux de centralité comparables ou légèrement supérieurs, voient leur position centrale se dégrader après choc, particulièrement en termes d'intermédiarité. Le phénomène de *friendship paradox* souligne aussi que les chercheurs les plus centraux concentrent davantage de collaborations, accentuant les inégalités d'accès.

S'agissant des statuts core et non-core, l'impact du choc est contrasté : les core conservent une meilleure accessibilité globale, mais perdent en influence stratégique, laissant aux non-core un rôle croissant dans la circulation de l'information.

Au niveau structurel, malgré quelques évolutions locales, le réseau conserve une forte modularité après choc. Cependant, cette structuration n'est pas portée par des regroupements de genre ou de statut core : l'organisation reste largement mixte. Le coefficient d'assortativité proche de zéro confirme que les collaborations ne dépendent pas du genre.

J'ai mis en évidence que le réseau présente une structuration particulièrement forte (modularité très élevée et coefficient de clustering moyen important). Cela invite à s'interroger sur les facteurs sous-jacents à cette organisation : comme l'ancienneté des chercheurs, leur âge, ou leur domaine scientifique, qui ne sont pas documentés dans les données disponibles. Ces dimensions pourraient constituer des pistes d'approfondissement pour mieux comprendre les trajectoires individuelles et collectives après le traitement.

Annexe : Code Python

Le code suivant présente les principales étapes de l'analyse de réseau réalisée :

1. Chargement et nettoyage des données

```
import pandas as pd

nodes = pd.read_csv("nodes_cluster_A.csv", sep=";")
edges = pd.read_csv("edges_cluster_A.csv", sep=";")

# Vérif rapide des types de variables
print(nodes.dtypes)
print(edges.dtypes)

# Suppression des doublons (parceque non dirigé)
edges['pair'] = edges.apply(lambda row: tuple(sorted([row['source'], row['target']])), axis=1)
edges = edges.drop_duplicates(subset='pair').drop(columns='pair')

# On garde que ceux qui ont collaboré (si ils n'ont pas collaboré avant ou parès alors on supprime)
edges = edges[(edges['pre_link'] == 1) | (edges['post_link'] == 1)]

print("Nœuds (agents) :")
print(nodes.head(), "\n")

print("Liens (dyades) :")
print(edges.head())
```

Visualisation générale des réseaux

```
import networkx as nx

# On trie en avant et après
edges_pre = edges[edges['pre_link'] == 1][['source', 'target']]
edges_post = edges[edges['post_link'] == 1][['source', 'target']]

# On créé deux graphes (non orientés)
G_pre = nx.Graph()
G_pre.add_edges_from(edges_pre.values)

G_post = nx.Graph()
G_post.add_edges_from(edges_post.values)

# Attributs (coord, core, gender)
attr_dict = nodes.set_index('id').to_dict('index')
nx.set_node_attributes(G_pre, attr_dict)
nx.set_node_attributes(G_post, attr_dict)

# Nombre de nœuds et d'arêtes
print(f"Pre : {G_pre.number_of_nodes()} nœuds, {G_pre.number_of_edges()} arêtes")
print(f"Post : {G_post.number_of_nodes()} nœuds, {G_post.number_of_edges()} arêtes")

# Visualisation par genre & coordination (après)
import matplotlib.pyplot as plt

# Nombre liens avant et après choc
nb_liens_pre = G_pre.number_of_edges()
nb_liens_post = G_post.number_of_edges()

print(f"Nombre de liens avant choc : {nb_liens_pre}")
print(f"Nombre de liens après choc : {nb_liens_post}")
import networkx as nx
```

```

import matplotlib.pyplot as plt

# Création d'un graphe qui contient tous les nœuds
G_union = nx.compose(G_pre, G_post)

# Position commune pour tous les nœuds (avant et après)
pos = nx.spring_layout(G_union, seed=42, k=0.3)

plt.figure(figsize=(18, 8))

# Avant
plt.subplot(1, 2, 1)
colors_pre = ['purple' if G_pre.nodes[n]['gender'] == 1 else 'blue' for n in G_pre.nodes()]
sizes_pre = [50 + G_pre.nodes[n]['core'] * 100 for n in G_pre.nodes()]
nx.draw(G_pre, pos,
        node_color=colors_pre,
        node_size=sizes_pre,
        edge_color='gray',
        width=0.5,
        alpha=0.7,
        with_labels=False)

plt.title("Réseau AVANT choc\n(Bleu=Homme, Rouge=Femme; Taille=Core)", fontsize=12)

# Après
plt.subplot(1, 2, 2)
colors_post = ['purple' if G_post.nodes[n]['gender'] == 1 else 'blue' for n in G_post.nodes()]
sizes_post = [50 + G_post.nodes[n]['core'] * 100 for n in G_post.nodes()]
nx.draw(G_post, pos,
        node_color=colors_post,
        node_size=sizes_post,
        edge_color='gray',
        width=0.5,
        alpha=0.7,
        with_labels=False)

plt.title("Réseau APRÈS choc\n(Bleu=Homme, Rouge=Femme; Taille=Core)", fontsize=12)

plt.tight_layout()
plt.show()

```

Évolution de la centralité des individus avant et après choc

Centralite de degrés

```

degrees = [deg for node, deg in G_pre.degree()]

plt.figure(figsize=(8,5))
plt.hist(degrees, bins=range(1, max(degrees)+2), edgecolor='black')
plt.title("Distribution des degrés (AVANT)")
plt.xlabel("Nombre de connexions")
plt.ylabel("Nombre de chercheurs")
plt.grid(axis='y')
plt.show()

degrees = [deg for node, deg in G_post.degree()]

plt.figure(figsize=(8,5))
plt.hist(degrees, bins=range(1, max(degrees)+2), edgecolor='black')
plt.title("Distribution des degrés (APRÈS)")
plt.xlabel("Nombre de connexions")
plt.ylabel("Nombre de chercheurs")
plt.grid(axis='y')
plt.show()

```

```

# Pre
gender_pre = nx.get_node_attributes(G_pre, "gender")

degrees_pre = dict(G_pre.degree())

# Moy générale
mean_degree_pre = sum(degrees_pre.values()) / len(degrees_pre)

# Moy genre
femmes_pre = [n for n in G_pre.nodes() if gender_pre.get(n) == 1]
hommes_pre = [n for n in G_pre.nodes() if gender_pre.get(n) == 0]

mean_deg_femmes_pre = sum(degrees_pre[n] for n in femmes_pre) / len(femmes_pre)
mean_deg_hommes_pre = sum(degrees_pre[n] for n in hommes_pre) / len(hommes_pre)

print(f"Degré moyen AVANT : {mean_degree_pre:.2f}")
print(f"Femmes : {mean_deg_femmes_pre:.2f} - Hommes : {mean_deg_hommes_pre:.2f}")

# Post
gender_post = nx.get_node_attributes(G_post, "gender")

degrees_post = dict(G_post.degree())

# Moy générale
mean_degree_post = sum(degrees_post.values()) / len(degrees_post)

# Moy genre
femmes_post = [n for n in G_post.nodes() if gender_post.get(n) == 1]
hommes_post = [n for n in G_post.nodes() if gender_post.get(n) == 0]

mean_deg_femmes_post = sum(degrees_post[n] for n in femmes_post) / len(femmes_post)
mean_deg_hommes_post = sum(degrees_post[n] for n in hommes_post) / len(hommes_post)

print(f"Degré moyen APRÈS : {mean_degree_post:.2f}")
print(f"Femmes : {mean_deg_femmes_post:.2f} - Hommes : {mean_deg_hommes_post:.2f}")

```

Centralite de proximit 

```

# Centralit  de proximit 
closeness_pre = nx.closeness centrality(G_pre)
closeness_post = nx.closeness centrality(G_post)

# Hist avant
plt.hist(list(closeness_pre.values()), bins=30)
plt.title('Distribution de la centralit  de proximit  avant choc')
plt.xlabel('Closeness')
plt.ylabel('Nombre de chercheurs')
plt.show()

# Hist apr s
plt.hist(list(closeness_post.values()), bins=30)
plt.title('Distribution de la centralit  de proximit  apr s choc')
plt.xlabel('Closeness')
plt.ylabel('Nombre de chercheurs')
plt.show()

# Moy par genre
gender_pre = nx.get_node_attributes(G_pre, "gender")
gender_post = nx.get_node_attributes(G_post, "gender")

femmes_pre = [n for n in G_pre.nodes() if gender_pre.get(n) == 1]
hommes_pre = [n for n in G_pre.nodes() if gender_pre.get(n) == 0]

```

```

mean_close_femmes_pre = sum(closeness_pre[n] for n in femmes_pre) / len(femmes_pre)

# Avant
close_centrality_pre = nx.closeness centrality(G_pre)
top_close_pre = sorted(close_centrality_pre.items(), key=lambda x: x[1], reverse=True)[:5]
mean_close_femmes_pre = sum(close_centrality_pre[n] for n in femmes_pre) / len(femmes_pre)
mean_close_hommes_pre = sum(close_centrality_pre[n] for n in hommes_pre) / len(hommes_pre)

print(f"Moyenne (avant) femmes : {mean_close_femmes_pre:.4f} - hommes : {mean_close_hommes_pre:.4f}")
print("Top 5 :")
for node, val in top_close_pre:
    print(f"ID {node} { Closeness : {val:.4f}")

gender_post = nx.get_node_attributes(G_post, "gender")
close_centrality = nx.closeness centrality(G_post)

# Top 5
top_close = sorted(close_centrality.items(), key=lambda x: x[1], reverse=True)[:5]
print("Top 5 { Centralité de proximité (après) :")
for node, val in top_close:
    print(f"ID {node} { Closeness : {val:.4f}")

# Moy genre
gender_post = nx.get_node_attributes(G_post, "gender")
femmes = [n for n in G_post.nodes() if gender_post.get(n) == 1]
hommes = [n for n in G_post.nodes() if gender_post.get(n) == 0]
mean_close_femmes = sum(close_centrality[n] for n in femmes) / len(femmes)
mean_close_hommes = sum(close_centrality[n] for n in hommes) / len(hommes)

print(f"Closeness moyenne Femmes : {mean_close_femmes:.4f}")
print(f"Closeness moyenne hommes : {mean_close_hommes:.4f}")

```

Centralité d'intermédierité

```

# Apres
between centrality = nx.betweenness centrality(G_post)

# Top 5
top_bet = sorted(between centrality.items(), key=lambda x: x[1], reverse=True)[:5]
print("Top 5 { Centralité d'intermédierité (après) :")
for node, val in top_bet:
    print(f"ID {node} { Betweenness : {val:.4f}")

# Moy genre
mean_bet_femmes = sum(between centrality[n] for n in femmes) / len(femmes)
mean_bet_hommes = sum(between centrality[n] for n in hommes) / len(hommes)

print(f"Betweenness moyenne Femmes : {mean_bet_femmes:.4f}")
print(f"Betweenness moyenne hommes : {mean_bet_hommes:.4f}")

# Avant
between centrality_pre = nx.betweenness centrality(G_pre)
top_bet_pre = sorted(between centrality_pre.items(), key=lambda x: x[1], reverse=True)[:5]
mean_bet_femmes_pre = sum(between centrality_pre[n] for n in femmes_pre) / len(femmes_pre)
mean_bet_hommes_pre = sum(between centrality_pre[n] for n in hommes_pre) / len(hommes_pre)
print(f"Moyenne femmes : {mean_bet_femmes_pre:.4f} - hommes : {mean_bet_hommes_pre:.4f}")
print("Top 5 :")
for node, val in top_bet_pre:
    print(f"ID {node} { Betweenness : {val:.4f}")

```

Core/non-Core

```

core_pre = nx.get_node_attributes(G_pre, "core")

```

```

core_nodes_pre = [n for n in G_pre.nodes() if core_pre.get(n) == 1]
noncore_nodes_pre = [n for n in G_pre.nodes() if core_pre.get(n) == 0]

degrees_pre = dict(G_pre.degree())
mean_deg_core_pre = sum(degrees_pre[n] for n in core_nodes_pre) / len(core_nodes_pre)
mean_deg_noncore_pre = sum(degrees_pre[n] for n in noncore_nodes_pre) / len(noncore_nodes_pre)

# Centralité de degré
deg_centrality_pre = nx.degree_centrality(G_pre)
mean_deg_cent_core_pre = sum(deg_centrality_pre[n] for n in core_nodes_pre) / len(core_nodes_pre)
mean_deg_cent_noncore_pre = sum(deg_centrality_pre[n] for n in noncore_nodes_pre) / len(noncore_nodes_pre)

# Closeness
closeness_pre = nx.closeness_centrality(G_pre)
mean_closeness_core_pre = sum(closeness_pre[n] for n in core_nodes_pre) / len(core_nodes_pre)
mean_closeness_noncore_pre = sum(closeness_pre[n] for n in noncore_nodes_pre) / len(noncore_nodes_pre)

# Betweenness
betweenness_pre = nx.betweenness_centrality(G_pre)
mean_bet_core_pre = sum(betweenness_pre[n] for n in core_nodes_pre) / len(core_nodes_pre)
mean_bet_noncore_pre = sum(betweenness_pre[n] for n in noncore_nodes_pre) / len(noncore_nodes_pre)

print(f"Degré moyen Core : {mean_deg_core_pre:.2f} // Non-core : {mean_deg_noncore_pre:.2f}")
print(f"Centralité degré Core : {mean_deg_cent_core_pre:.4f} // Non-core : {mean_deg_cent_noncore_pre:.4f}")
print(f"Closeness Core : {mean_closeness_core_pre:.4f} // Non-core : {mean_closeness_noncore_pre:.4f}")
print(f"Betweenness Core : {mean_bet_core_pre:.4f} // Non-core : {mean_bet_noncore_pre:.4f}")

core_post = nx.get_node_attributes(G_post, "core")
core_nodes = [n for n in G_post.nodes() if core_post.get(n) == 1]
noncore_nodes = [n for n in G_post.nodes() if core_post.get(n) == 0]

degrees_post = dict(G_post.degree())
mean_deg_core = sum(degrees_post[n] for n in core_nodes) / len(core_nodes)
mean_deg_noncore = sum(degrees_post[n] for n in noncore_nodes) / len(noncore_nodes)

# Centralité de degré
deg_centrality_post = nx.degree_centrality(G_post)
mean_deg_cent_core = sum(deg_centrality_post[n] for n in core_nodes) / len(core_nodes)
mean_deg_cent_noncore = sum(deg_centrality_post[n] for n in noncore_nodes) / len(noncore_nodes)

# Closeness
closeness_post = nx.closeness_centrality(G_post)
mean_closeness_core = sum(closeness_post[n] for n in core_nodes) / len(core_nodes)
mean_closeness_noncore = sum(closeness_post[n] for n in noncore_nodes) / len(noncore_nodes)

# Betweenness
betweenness_post = nx.betweenness_centrality(G_post)
mean_bet_core = sum(betweenness_post[n] for n in core_nodes) / len(core_nodes)
mean_bet_noncore = sum(betweenness_post[n] for n in noncore_nodes) / len(noncore_nodes)

print(f"Degré moyen Core : {mean_deg_core:.2f} // Non-core : {mean_deg_noncore:.2f}")
print(f"Centralité degré Core : {mean_deg_cent_core:.4f} // Non-core : {mean_deg_cent_noncore:.4f}")
print(f"Closeness Core : {mean_closeness_core:.4f} // Non-core : {mean_closeness_noncore:.4f}")
print(f"Betweenness Core : {mean_bet_core:.4f} // Non-core : {mean_bet_noncore:.4f}")

```

Structure du réseau

Sous-communautés

```

import community as community_louvain
# Détection des communautés
partition_post = community_louvain.best_partition(G_post)
partition_pre = community_louvain.best_partition(G_pre)

```

```
modularity_pre = community_louvain.modularity(partition_pre, G_pre)
print(f"Modularité du clustering (AVANT) : {modularity_pre:.4f}")

modularite_post = community_louvain.modularity(partition_post, G_post)
print(f"Modularité du clustering (APRÈS) : {modularite_post:.4f}")
```

Homophilie

```
# DataFrame pour G_pre
df_nodes_pre = pd.DataFrame({
    "id": list(partition_pre.keys()),
    "community": list(partition_pre.values()),
    "gender": [G_pre.nodes[n]["gender"] for n in partition_pre.keys()],
    "core": [G_pre.nodes[n]["core"] for n in partition_pre.keys()]
})

# Moy par communauté
stats_com_pre = df_nodes_pre.groupby("community").agg(
    taille=("id", "count"),
    part_femmes=("gender", "mean"),
    part_core=("core", "mean")
).sort_values(by="taille", ascending=False)

# Communautés homogènes
print("Communautés très féminisées (> 0.75) AVANT :") # nous fixons le seuil, un autre aurait pu faire l'affaire mais on garde
print(stats_com_pre[stats_com_pre["part_femmes"] > 0.75])

print("\nCommunautés très core (> 0.75) AVANT :")
print(stats_com_pre[stats_com_pre["part_core"] > 0.75])

# Communautés homogènes
print("Communautés très féminisées (> 0.7) AVANT :") # nous fixons le seuil, un autre aurait pu faire l'affaire mais on garde
print(stats_com_pre[stats_com_pre["part_femmes"] > 0.7])

print("\nCommunautés très core (> 0.7) AVANT :")
print(stats_com_pre[stats_com_pre["part_core"] > 0.7])

df_nodes_post = pd.DataFrame({
    "id": list(partition_post.keys()),
    "community": list(partition_post.values()),
    "gender": [G_post.nodes[n]["gender"] for n in partition_post.keys()],
    "core": [G_post.nodes[n]["core"] for n in partition_post.keys()]
})
stats_com_post = df_nodes_post.groupby("community").agg(
    taille=("id", "count"),
    part_femmes=("gender", "mean"),
    part_core=("core", "mean")
).sort_values(by="taille", ascending=False)

print("Communautés très féminisées (> 0.75) APRÈS :")
print(stats_com_post[stats_com_post["part_femmes"] > 0.75])

print("\nCommunautés très core (> 0.75) APRÈS :")
print(stats_com_post[stats_com_post["part_core"] > 0.75])

print("Communautés très féminisées (> 0.7) APRÈS :")
print(stats_com_post[stats_com_post["part_femmes"] > 0.7])

print("\nCommunautés très core (> 0.7) APRÈS :")
print(stats_com_post[stats_com_post["part_core"] > 0.7])

nx.attribute_assortativity_coefficient(G_post, "gender") #hmzphilie globale
```