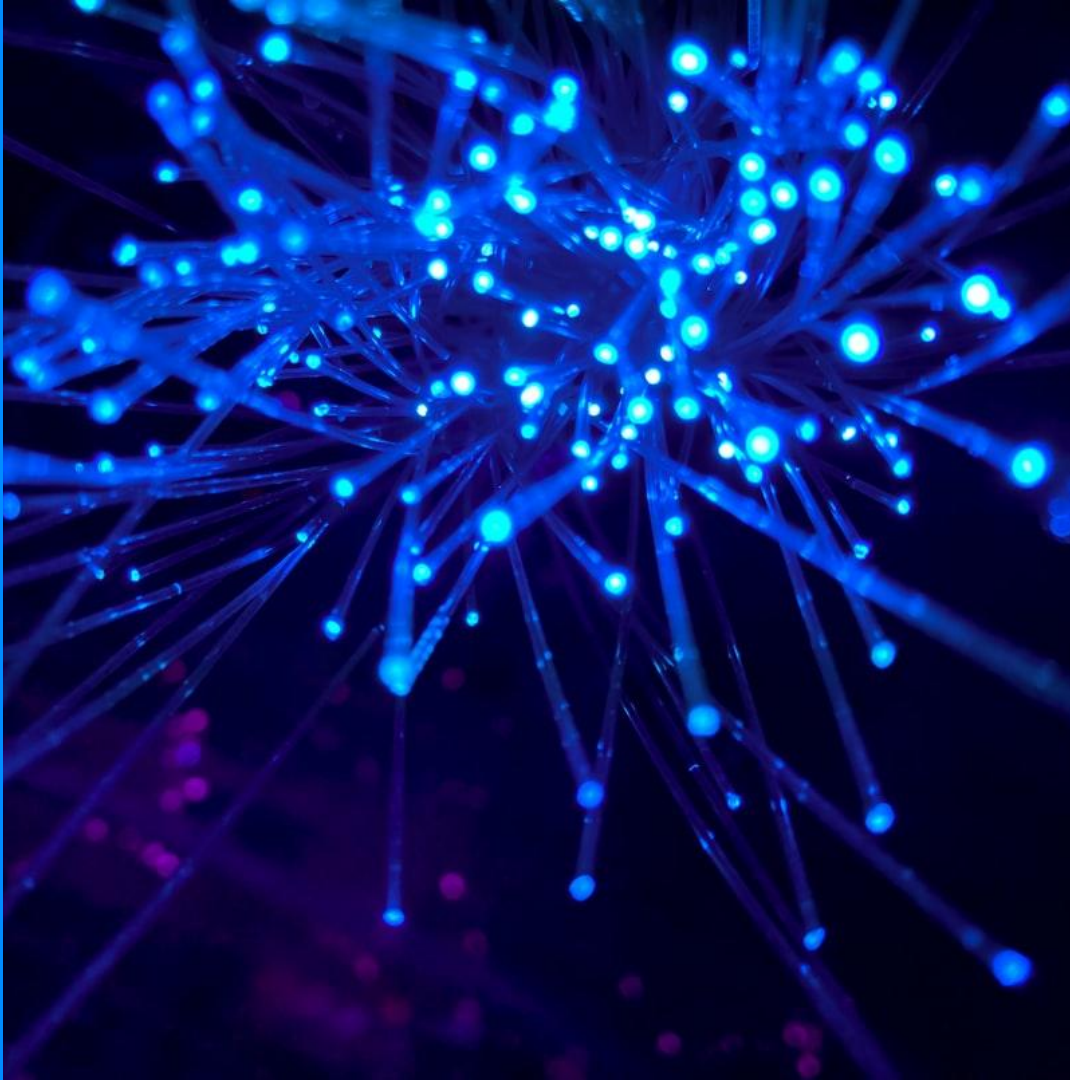


Hackaviz 2022

Les graphes

pour comprendre les aides dispensées
par les organismes de gestion collective
des droits d'auteur

Nicolas Rouyer,
Sales engineer, Neo4j



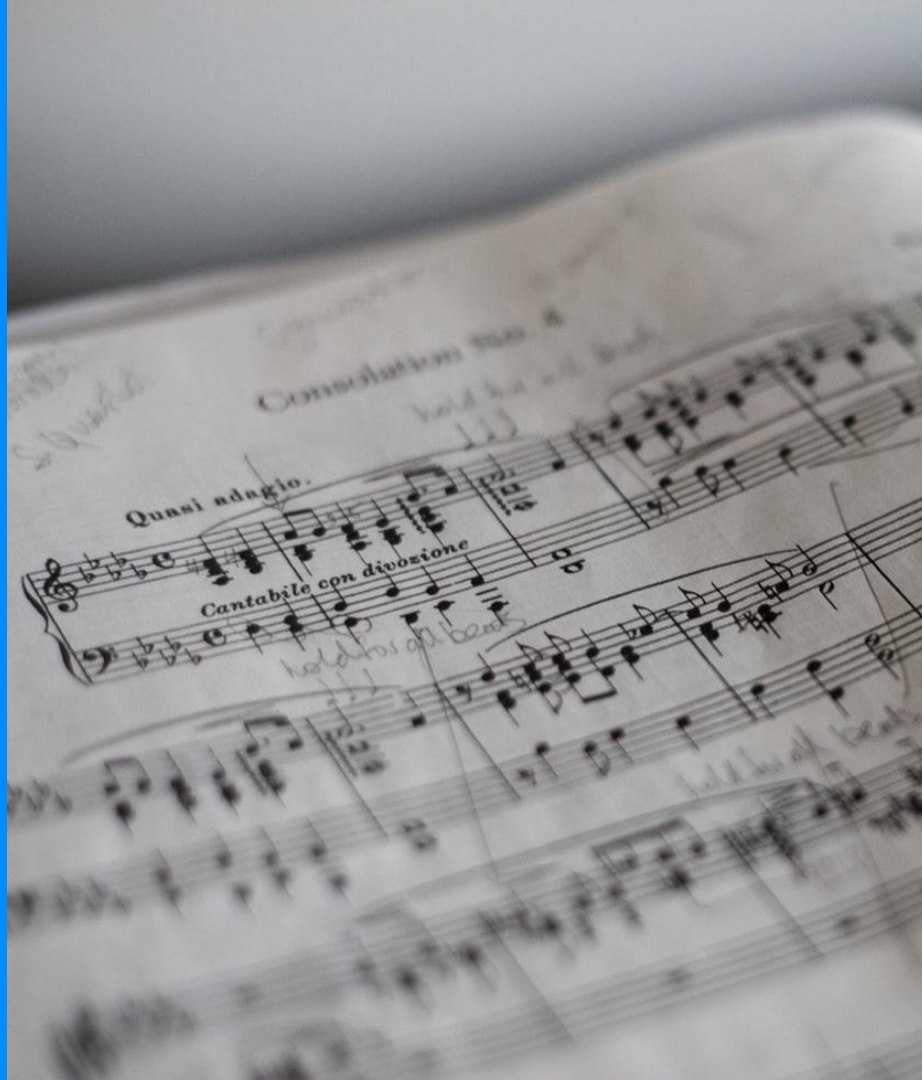
Introduction

Bonjour,

Je vous propose de vous présenter une visualisation sous forme de graphe des données du Hackaviz 2022.

Les graphes ce sont des noeuds, des relations qui les connectent et des propriétés sur les noeuds et les relations.

Les graphes permettent de capturer la complexité du monde et de révéler de la connaissance et de la valeur en exploitant les connections entre les données.



Sommaire

La partition que nous allons jouer est la suivante :

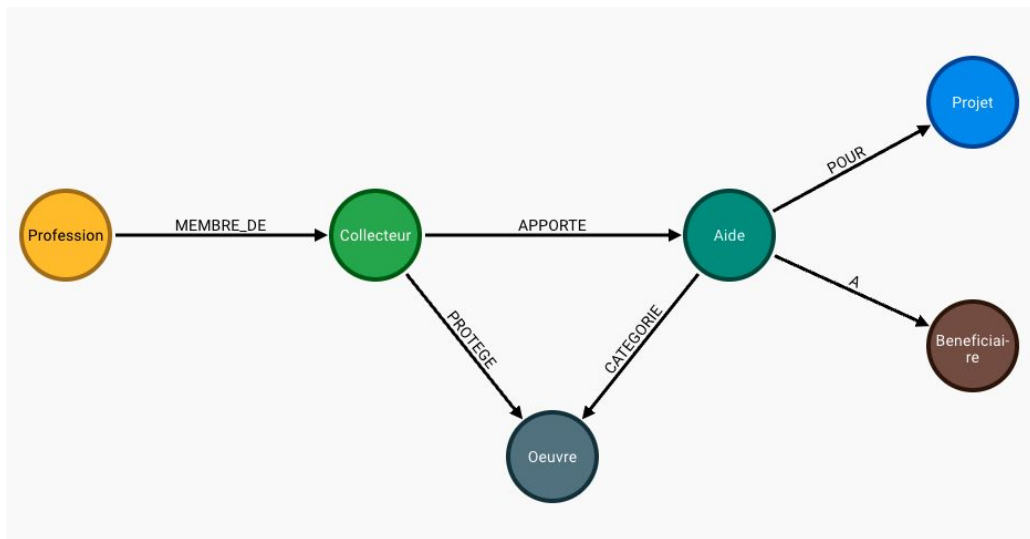
- *Allegro*, le **modèle de données graphe** des collecteurs, des aides
- *Andante*, l'**analyse graphe** une fois chargées les données
- *Presto*, l'utilisation de la graph data science pour déterminer la **similarité des aides**

Modèle de données

Tout est graphe...



Modèle de données



Un **Collecteur** **APPORTE** de l'**Aide** **A** un **Beneficiaire** **POUR** un **Projet**.

Une **Profession** est **MEMBRE_DE** un **Collecteur**.

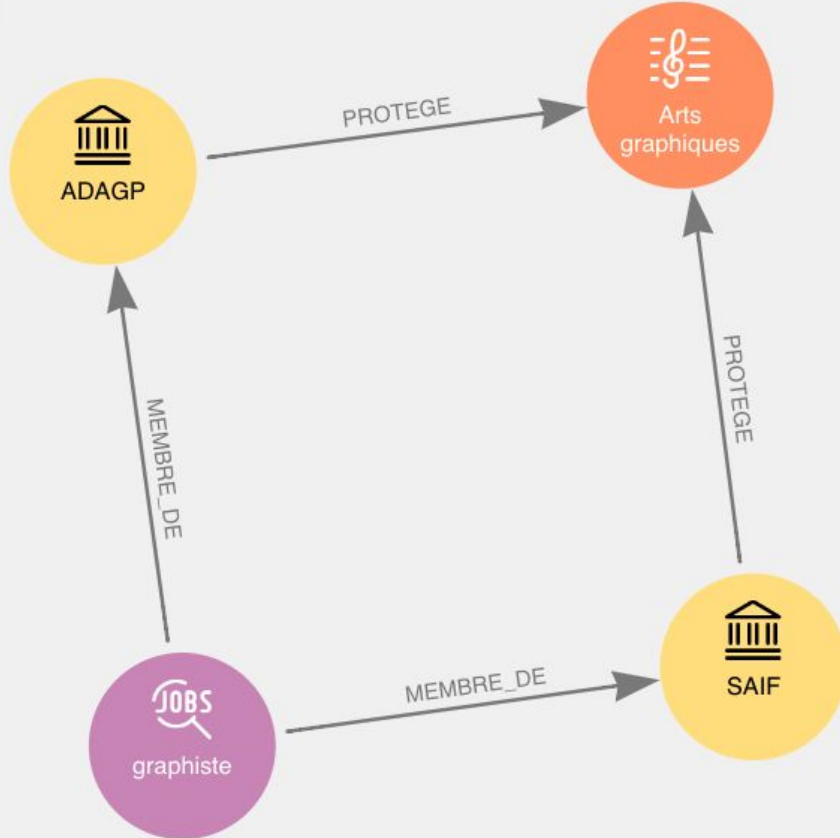
Un **Collecteur** **PROTEGE** une **Oeuvre**. Une **Aide** est classée par **CATEGORIE** d'**Oeuvre**.



graphiste

Collecteur

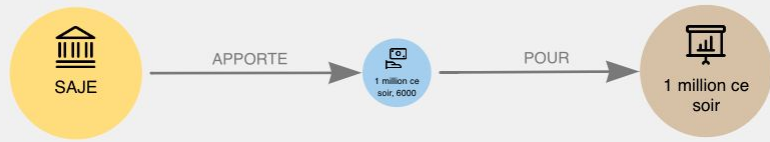
Oeuvre



Exemple

On peut poser une question en langage quasi naturel au graphe Neo4j :

Quelles sont les oeuvres protégées par les organismes collecteurs dont les membres sont des graphistes ?



Exemple #2

Quel est l'organisme collecteur ayant apporté de l'aide au projet de jeu télévisé :
"1 million ce soir ?"

Quelques visualisations statistiques

Avec un outil light de Business Intelligence, NeoDash



Tableau de bord

Page name...

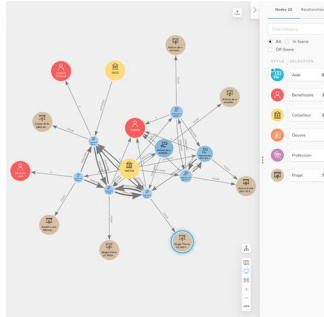
+



Hello 🙌



Tableau de bord Hackaviz 2022

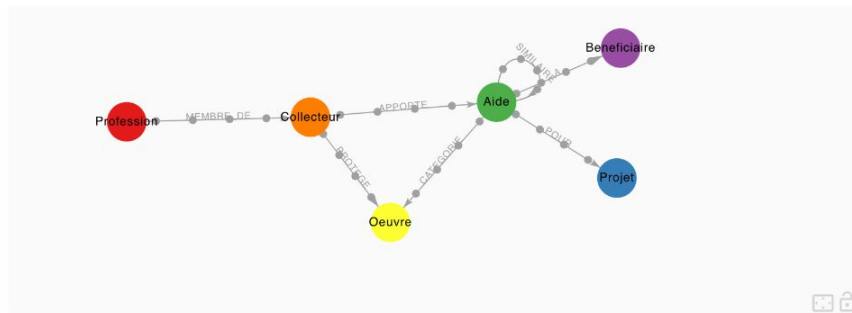


Nombre d'aides



52 740

Modèle graphe



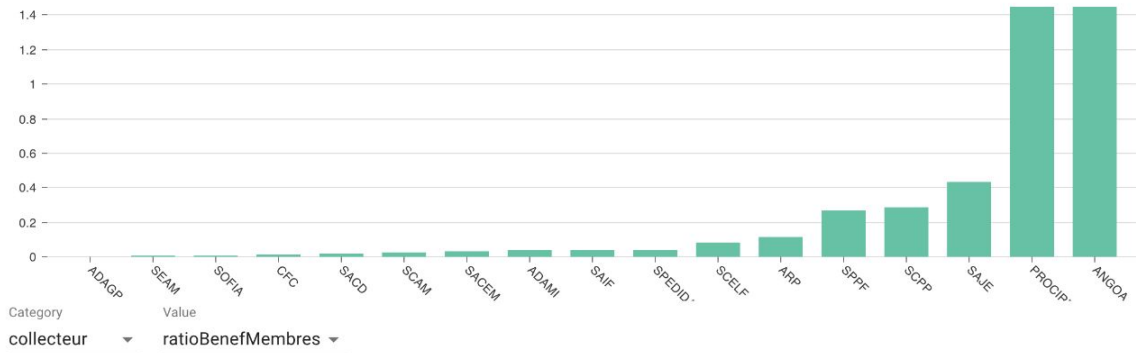
Profession
 Projet
 Aide
 Beneficiaire
 Collecteur
 Oeuvre

TOP 10 des projets les + financés



projet	nombreAides	montantCum...
Aide à la création d'emploi de secrétariat	112	4 510 817
Aide à la défense du droit d'auteur -	66	4 018 447
LE FONDS POUR LA CRÉATION MUSICALI	4	3 065 164
Développement de carrière Musiques actu	328	2 530 290
Formation	113	2 387 935
Valorisation et promotion du répertoire - C	5	2 020 064

Ratio nb bénéficiaires / nbMembres par collecteur



Commentaires sur la visualisation statistique

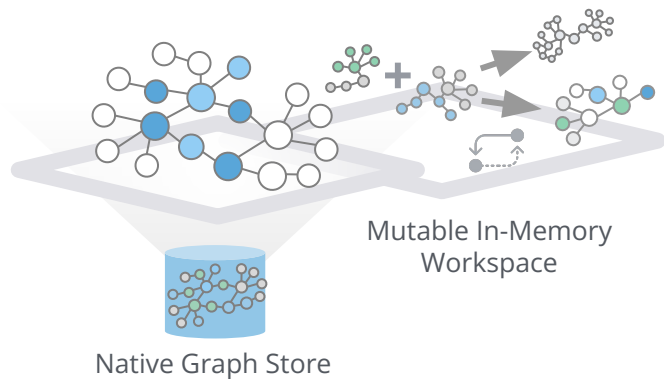
- Le ratio nombre bénéficiaires d'aides / nombre de membres ne dépasse pas 1,4%. Certes, certains bénéficiaires peuvent regrouper plusieurs membres
- On veut aller plus loin en comprenant comment les aides sont décernées, pour cela on va utiliser la graph data science et déterminer une similarité des aides entre elles :
 - Par les montants accordés
 - Par les bénéficiaires touchant ces aides
 - Par les projets financés

Graph Data Science

Des algorithmes de graphe pour comprendre les
connections entre les aides



Computational Graph

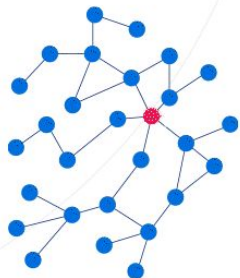


Principes de la Graph Data Science

- On utilise une bibliothèque neo4j pour la Graph Data Science
- Cette bibliothèque permet de projeter en mémoire tout ou partie du graphe, et de passer sur ces projections des algorithmes
- En particulier, nous allons utiliser les **graph embeddings**

neo4j

Graph Embeddings:
AI That Learns
from Your Data to
Solve Problems



Résultats GDS

On peut explorer les communautés d'aides similaires.

La similarité a été calculée grâce au montant des aides, et au “voisinage” des noeuds Aide, ce voisinage est capturé par les “embeddings”.

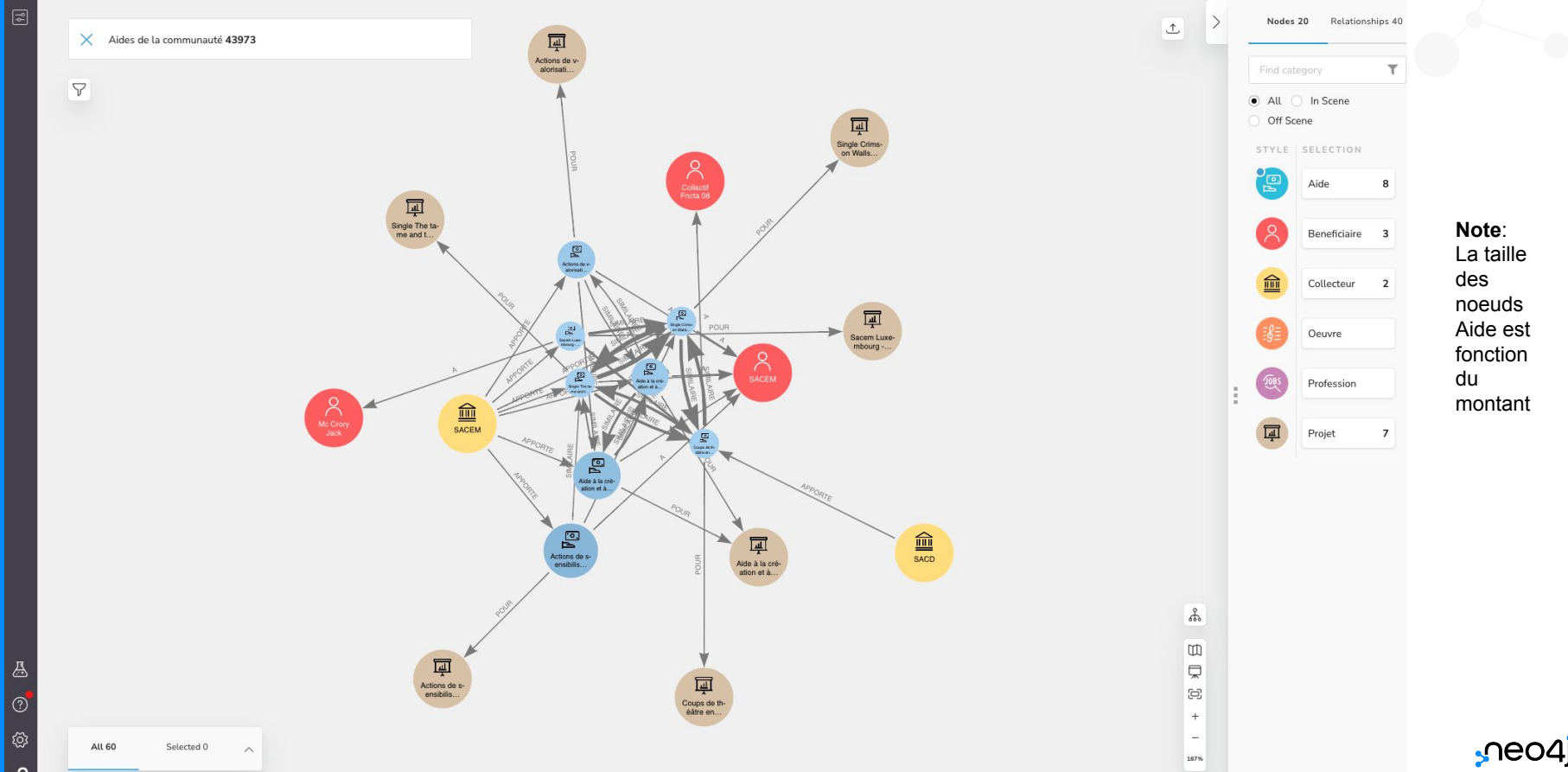
A partir du score de similarité, on lance un algorithme de Louvain pour détecter les communautés, que l'on peut alors explorer en fonction du nombre de membres et d'autres critères.

```
1 MATCH (a:Aide)
2 RETURN a.louvain AS communauté, count(a) AS nbMembres
3 ORDER BY nbMembres DESC
4 LIMIT 500
```

	communauté	nbMembres
1	7338	3648
2	47801	3599
3	11414	3341
4	47594	2575
5	18995	2483
6	46378	2166
7		

Started streaming 327 records in less than 1 ms and completed after 30 ms.

Exploration d'1 communauté d'aides semblables par les montants, les bénéficiaires et les projets



Vous voulez en savoir plus ? RV le 7 avril 2022 Pour 1 démo dynamique



 nicolas.rouyer@neo4j.com