

Tâche du défi

Juan-Manuel TORRES

juan-manuel.torres@univ-avignon.fr

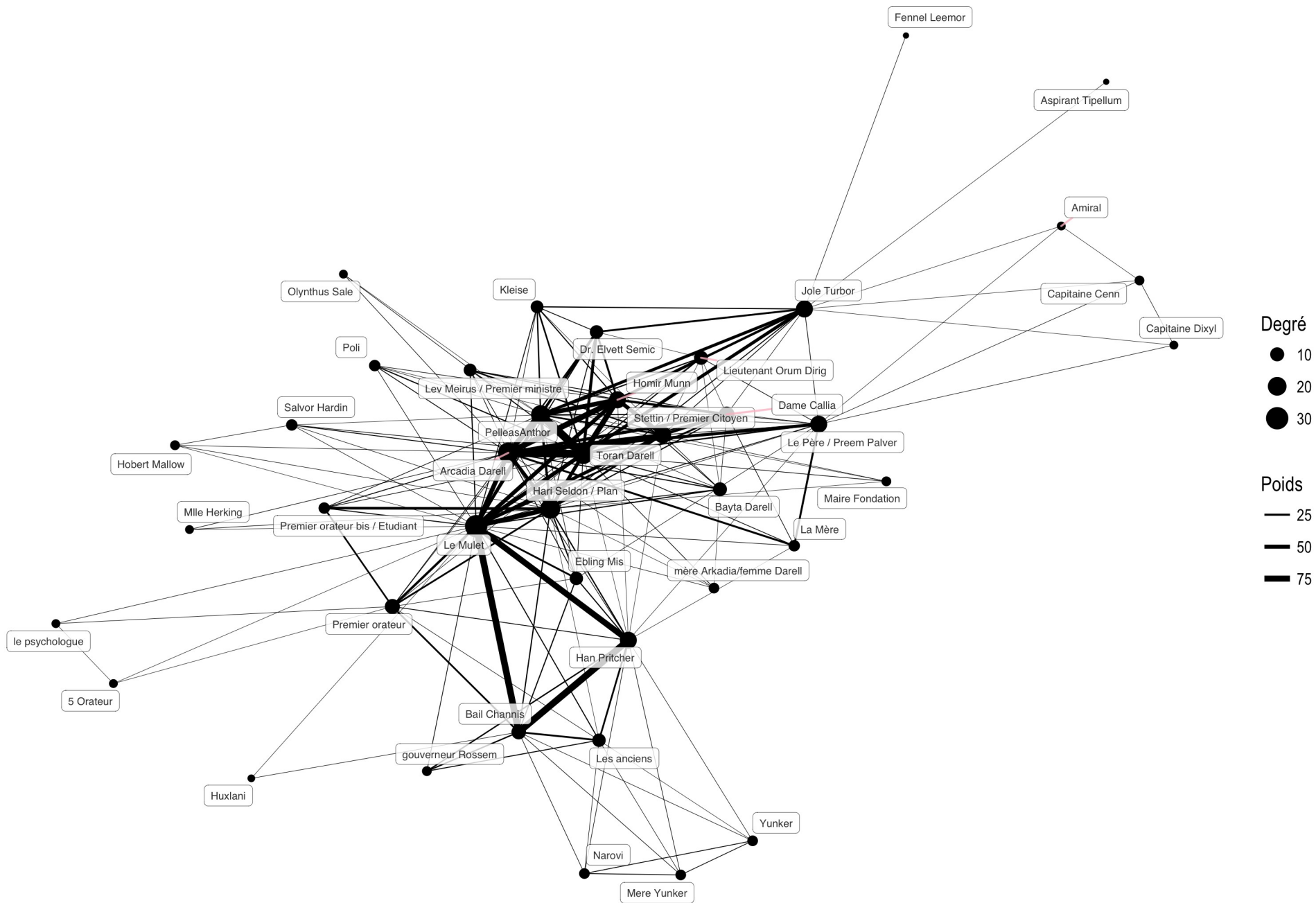


Tâche : Génération de réseaux de personnages

- Extraire des entités nommées des personnages
- Générer un réseau des personnages (avec des liens de proximité)
- Comment ... ?

Tâche : Génération de réseaux de personnages

- Comment ?
 - Utiliser des méthodes hybrides :
 - Symboliques
 - Probabilistes
 - Intelligence artificielle
 - Utiliser des ressources linguistiques : Freeling ...



Mais... comment faire ?





Isaac ASIMOV (Исаák Азі́мов)



Información personal

Nombre de nacimiento	Исаák Азі́мов
Nacimiento	2 de enero de 1920 Petróvichi (Rusia Soviética)
Fallecimiento	6 de abril de 1992 (72 años) Manhattan (Nueva York, Estados Unidos)
Causa de muerte	Enfermedad relacionada con el SIDA, insuficiencia renal e insuficiencia cardíaca
Residencia	Brooklyn , Petróvichi , Boston , Oeste de Filadelfia y Manhattan
Nacionalidad	Estadounidense
Religión	Ninguna (ateo)
Lengua materna	Yidis

Información profesional

Ocupación	Bioquímico, novelista , prosista , autobiógrafo , escritor de ciencia ficción , escritor de ciencia , escritor de no ficción , profesor universitario , periodista , guionista , escritor y escritor de ficción de detectives
Área	Bioquímica
Años activo	1939-1992
Empleador	Universidad de Boston
Seudónimo	Paul French y H. B. Ogden
Género	Ciencia ficción
Obras notables	Serie de la Fundación Serie de los robots Anochecer Yo, Robot El hombre bicentenario Los propios dioses

https://es.wikipedia.org/wiki/Isaac_Asimov

La Saga Fondation d'Isaac Asimov



L'ordre chronologique...

L'ordre chronologique des événements des livres de la saga Fondation d'Isaac Asimov

1. **Prélude à Fondation**
2. **L'Aube de Fondation**
3. **Fondation**
4. **Fondation et Empire**
5. **Seconde Fondation**
6. **Fondation Foudroyée**
7. **Terre et Fondation**

<https://chronolivres.com/quel-ordre-lire-fondation-isaac-asimov/>

Audiolivres (Plusieurs heures... courage!)

Fondation

<https://www.youtube.com/watch?v=HrKUFYPmbGU>

Fondation et Empire

<https://www.youtube.com/watch?v=DRz7LXphEwU>

Seconde Fondation

<https://www.youtube.com/watch?v=m6rXUrWwUc0>

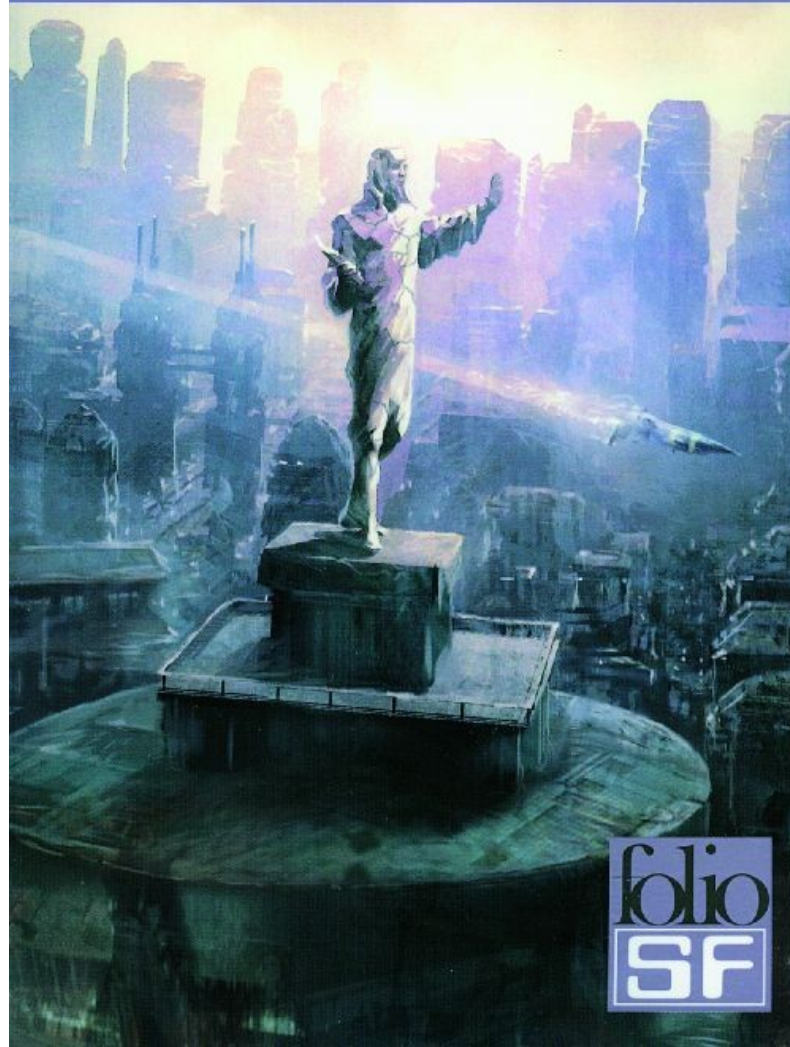
Terre et Fondation

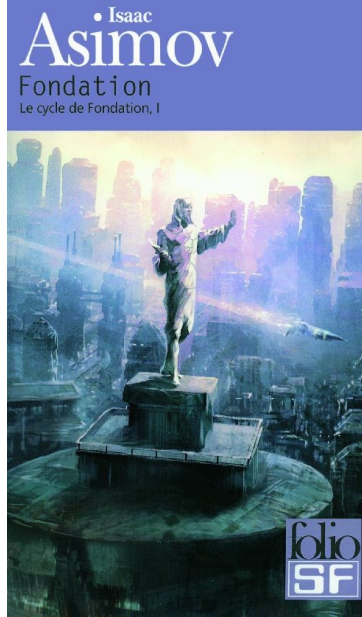
<https://www.youtube.com/watch?v=X7ZRL552ej0>
<https://www.youtube.com/watch?v=XqTTWT79Zuk>
<https://www.youtube.com/watch?v=JVQNdLSUb4Q>

• Isaac
Asimov

Fondation

Le cycle de Fondation, I

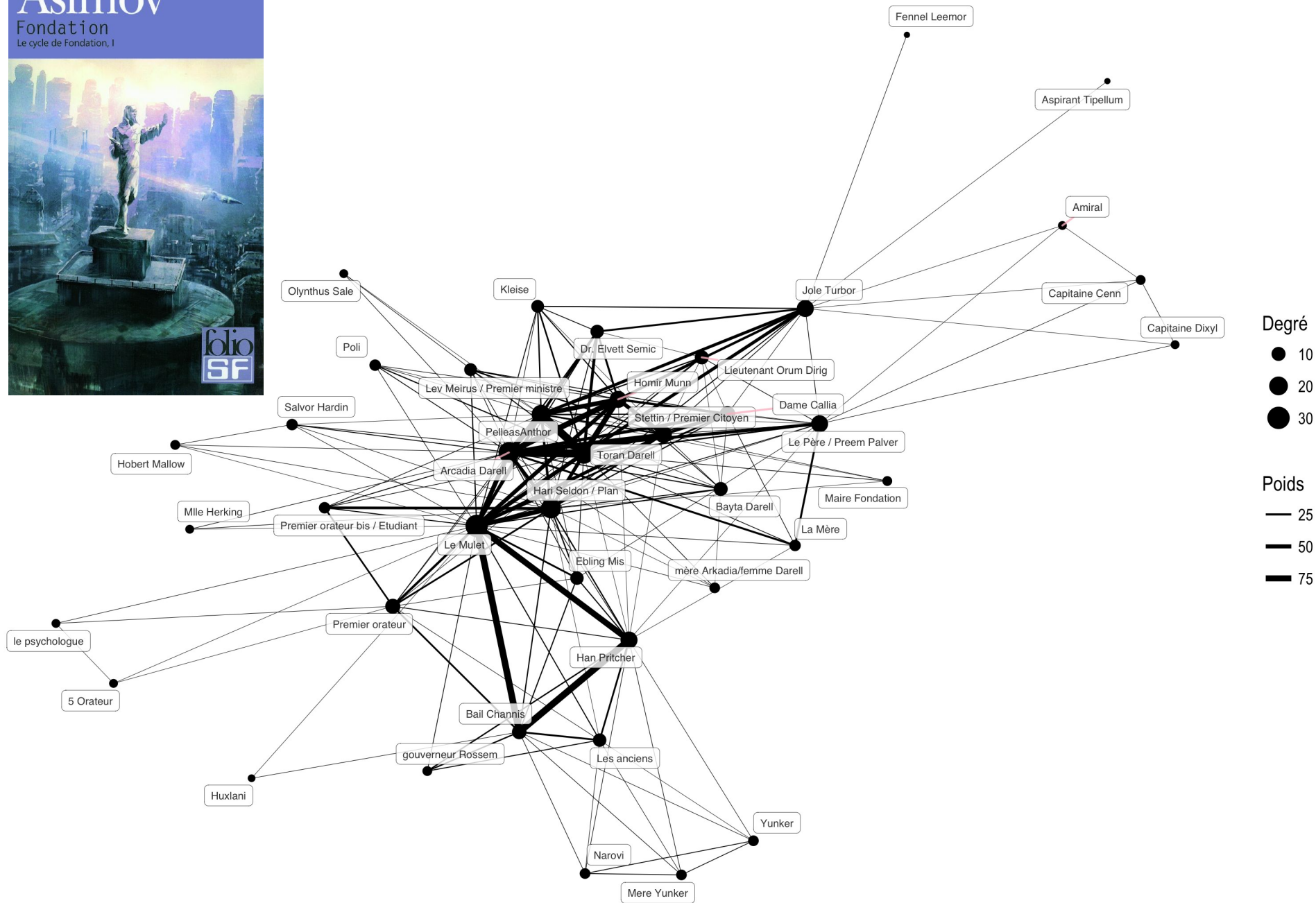




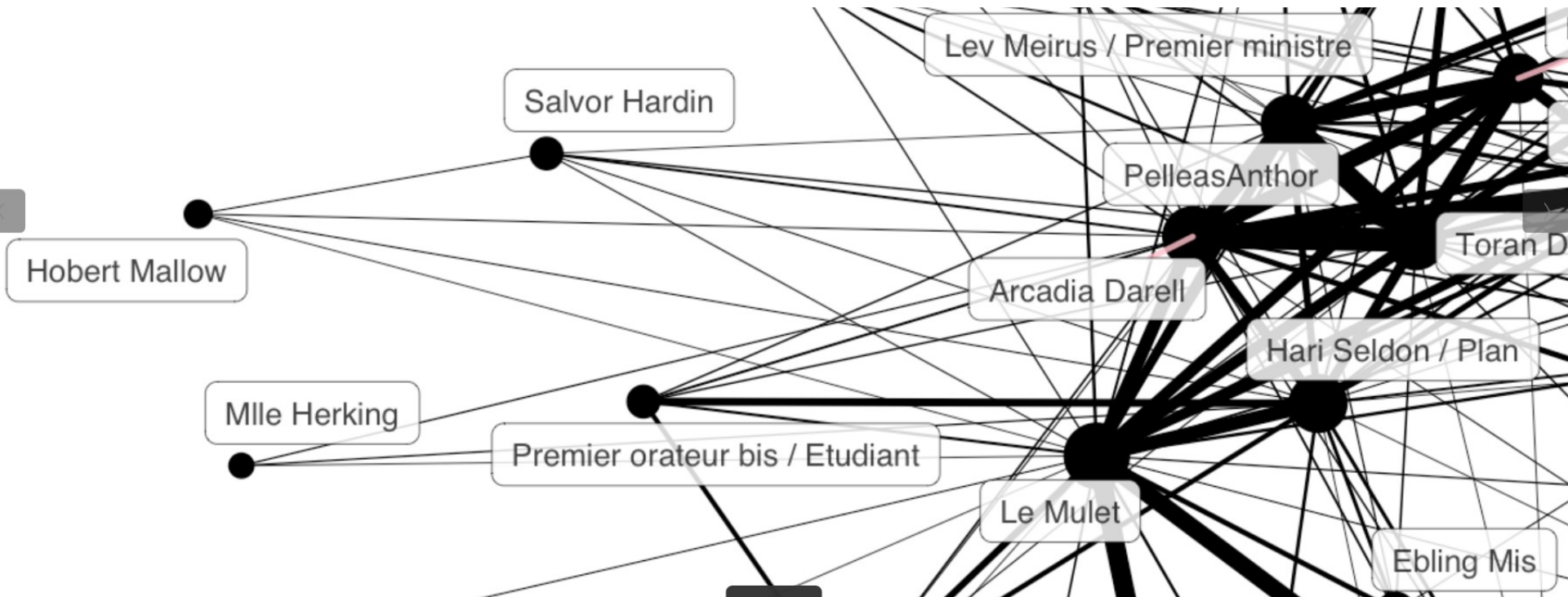
Nous sommes au treizième millénaire de l'Ère Galactique sur la planète Trantor, la capitale de l'Empire. Un mathématicien du nom de Hari Seldon invente la psychohistoire une science fondée sur les mathématiques qui peut prédire l'avenir.

D'après Seldon et sa psychohistoire, L'Empire est sur le point de connaître des jours noirs qui mèneront à trente mille ans de chaos et de barbarie. Afin d'éviter ce funeste futur et réduire au mieux ces trente mille ans de chaos à seulement mille ans, Seldon entreprend d'amasser tout le savoir de l'humanité dans une grande Encyclopédie.

Pour ce faire, il rassemble tous les savants de l'Empire, du scientifique au marchand. Il décide de ramener tout ce beau monde sur Terminus, une petite planète située quelque part dans la galaxie. Il appellera ce rassemblement de savants, **La Fondation**.



Zoom...



Comment construire un réseau de personnages ?

Identifier les entités nommées (EN)...

Identifier les EN candidates à être des noms de personnes...

Calculer les liens entre les EN personnages

Calculer le graphe

Visualiser le graphe

Entités nommées



WIKIPÉDIA
L'encyclopédie libre

Reconnaissance d'entités nommées

🌐 16 langues

Article [Discussion](#)

Lire


[Modifier](#)

[Modifier le code](#)

[Voir l'historique](#)



[Outils](#)

 *Ne doit pas être confondu avec [Annotation sémantique](#).*

La **reconnaissance d'entités nommées** est une sous-tâche de l'activité d'extraction d'information dans des [corpus](#) documentaires. Elle consiste à rechercher des objets textuels (c'est-à-dire un mot, ou un groupe de mots) catégorisables dans des classes telles que noms de personnes, noms d'organisations ou d'entreprises, noms de lieux, quantités, distances, valeurs, dates, etc.

https://fr.wikipedia.org/wiki/Reconnaissance_d'entités_nommées

Exemple

Principe [\[modifier | modifier le code \]](#)

À titre d'exemple, on pourrait donner le texte qui suit, étiqueté par un système de reconnaissance d'entités nommées utilisé lors de la campagne d'évaluation **MUC**:

Henri a acheté 300 actions de la société AMD en 2006

<ENAMEX TYPE="PERSON">Henri</ENAMEX> a acheté <NUMEX TYPE="QUANTITY">300</NUMEX> actions de la société <ENAMEX TYPE="ORGANIZATION">AMD</ENAMEX> en <TIMEX TYPE="DATE">2006</TIMEX>.

Comment faire ?

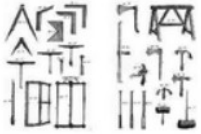
Outils

- Embeddings
- N-grammes
- Modele de langue
- Corpus POS
- ...
- Projections
- Graphes
- Composants principales
- Visualisations
- ...

<https://medium.com/@dhifallah.othmen/extraction-et-entrainement-des-entites-nommees-avec-spacy-90f7e1a219d7>

Outils classiques...

Open Source Natural Language Processing Tools



LIA_PHON

A French Text-To-Phoneme converter, compatible with MBROLA

[DOWNLOAD lia_phon v1.2](#)

LIA_TAGG

A simple POS tagger (French, English) (specific features for French: guesser for unknown words + reaccentueur)

[DOWNLOAD lia_tag v1.1](#)

LIA_SCT

an implementation of Semantic Classification Tree, with some extra features

[DOWNLOAD lia_sct v1.1](#)

LIA_NE

a Named-Entity tagger for French on text, ctm and stm files.

[DOWNLOAD lia_ne v2.2](#)