

LABORATOIRE
PARAGRAPHE

Cartographies des formations en Humanités numériques en France

Orélie Desfriches Doria

Laboratoire Paragraphe

Elie Allouche

Direction du numérique pour l'Education (MENESR)

Atelier DAHLIA

Conférence EGC

28 janvier 2020 – ICHEC- Bruxelles, Belgique

UNIVERSITÉ
PARIS 8
VINCENNES-SAINT-DENIS

Tâtonnements : le positionnement des formations en H.N. par rapport aux formations en général

- Interrogations sur l'architecture des formations :
 - Quels intitulés pour communiquer ?
 - UFR de rattachements
 - Disciplines de rattachement
 - Architecture Mention/parcours
- Publics visés et débouchés :
 - Conditions d'accès, pré-requis, public visé
 - Débouchés affichés
- Modalités pédagogiques
 - Adossement à la recherche/ laboratoires
 - Dimension internationale dans les cursus
 - Partenariats institutionnels
 - Pédagogie par projet
 - Stage

Que nous disent les formations en Humanités Numériques à propos des H.N ?

- Méthodologie :
 - Extraction des données
 - Dariah
 - Trouver mon Master (<https://www.trouvermonmaster.gouv.fr/>)
 - Segmentation des données
 - Reformatage
 - Insertion dans l'outil de cartographie
 - Travail sur la visualisation et algorithmes de spatialisation
 - Analyse des points saillants

Données hétérogènes

	Dariah Course Registry	Trouver mon Master
Nombre de résultats bruts en France	28	26
Nombre d'établissements différents	18	23
Nombre de Masters	13	23
Nombre de Parcours	/	57

Cartographie globale des formations par type de formation

Majorité des formations en master (en bleu) dans cette cartographie
D'après les données de cette base, c'est donc le type de formation majoritairement proposé en H.N. en France.



<https://graphcommons.com/graphs/9372006f-1640-47d5-b23d-1456d39d330f>

Cartographie des disciplines

Cartographie par discipline

Les disciplines mentionnées sont nombreuses : cette nébuleuse renvoie à la question des fondations méthodologiques communes et des pratiques convergentes à travers cette diversité disciplinaire au sein des H.N. (Desfriches Doria et al., 2018).

- 17 Institution
Size by metric_weighted_in_degree
- 14 Discipline
Size by metric_weighted_in_degree
- 11 Master Programme
Size by Degree Centrality
- 9 Course
Size by Degree Centrality
- 4 Bachelor Programme
Size by Degree Centrality
- 2 Research Master
Size by Degree Centrality



Quel positionnement disciplinaire des données sur les disciplines ?

- Les mots clés décrivent des disciplines dont les formations en H.N. se réclament? Ces termes visent –ils à décrire des disciplines connexes ?
- **Quelle est cette nomenclature** de disciplines ? Par qui est-elle proposée ?;
- Les données issues de “Trouver Mon Master” concernant les licences d’origine des étudiants ne peuvent pas être équivalentes : difficultés de rapprochement
- A l’état “brut”, chaque ligne des fichiers correspond à une formation, chaque discipline est classée dans une colonne ordonnée.
 - Cela signifie-t-il que **les disciplines sont ordonnées par ordre d’importance lors de leur saisie ?**
 - pose les questions de l’aspect contributif de la base de données Dariah,
 - de l’encadrement des pratiques d’indexation et de description des formations ;

De fortes disparités dans les données sur les disciplines

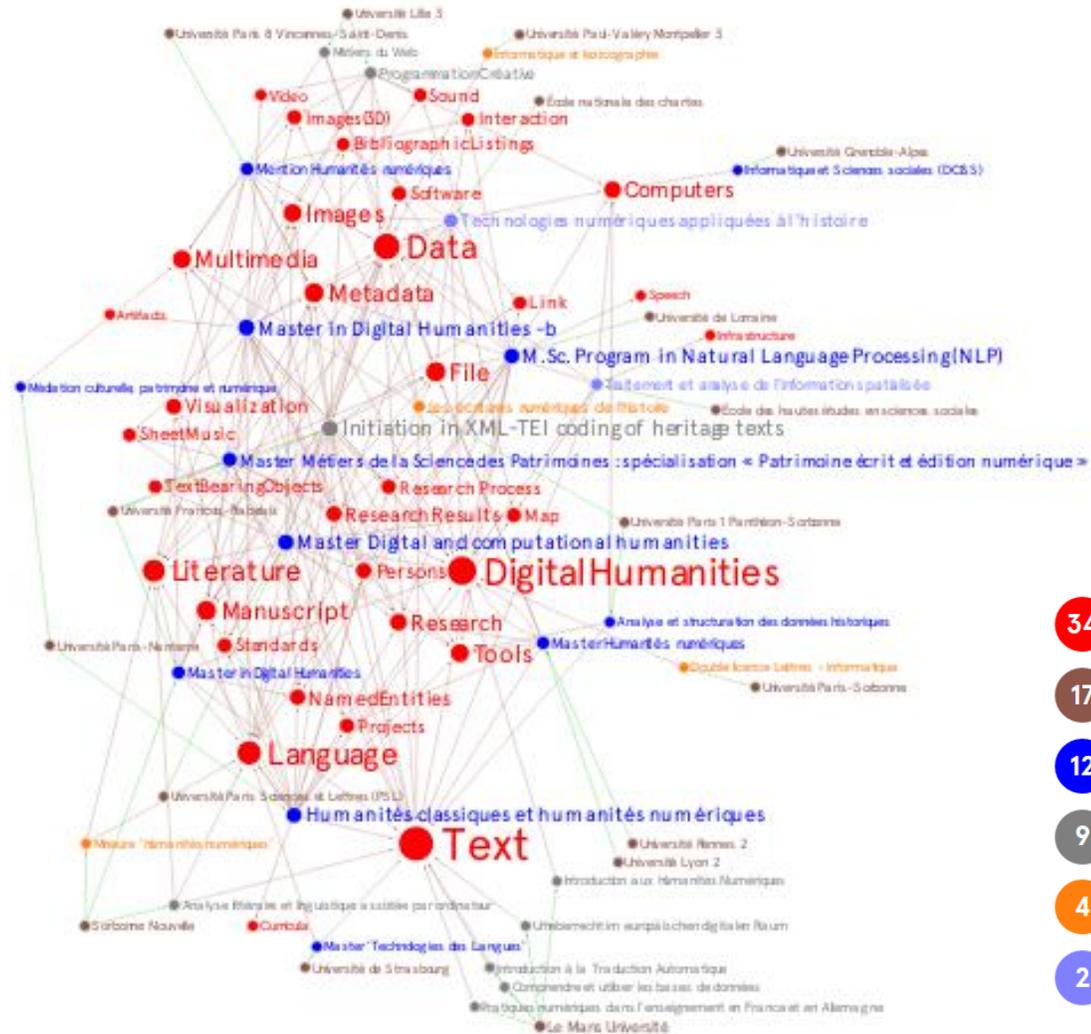
- **Variabilité des données et représentativité des données :**
 - de 1 mot-clé à 8 mots-clés pour décrire les disciplines.
 - 7 formations ne présentent qu'un mot clé, quand 3 formations en présentent 8.
 - Les formations qui ne sont décrites qu'à l'aide d'un seul mot clé, sont exclusivement de type "course" ou "bachelor programme" : pose la question de l'hétérogénéité des données en fonction du type de formation
 - de la représentativité de la quantité des citations de mots clés
 - Les Masters font aussi l'objet de disparités dans la description, leur nombre de mots-clés varie de 2 à 8
- **Des problèmes de lexique :**
 - Au niveau de la généricité des mots-clés employés :
 - l'utilisation d'un mot-clé "Social Science" alors que d'autres mots-clés plus précis existent dans les données comme "Histoire", "Archéologie", "Philosophie" par exemple.
 - Cette nomenclature est proposée pour décrire les disciplines mais il existe des **subtilités dans la formulation des disciplines au niveau international** : exemple les SIC

Cartographie des objets

Cartographie par objet

Une des dominantes observées porte sur le travail sur les données, métadonnées, les fichiers, le texte et plus globalement tout ce qui touche au traitement de corpus, sans que ce dernier terme (corpus) soit inclus dans les objets cités.

Les humanités numériques apparaissent elles-mêmes comme objet d'étude, confirmant ainsi le côté réflexif des pratiques développées et la dynamique interdisciplinaire



- 34 Objet
Size by metric_weighted_in_degree
- 17 Institution
Size by Degree Centrality
- 12 Master Programme
Size by Degree Centrality
- 9 Course
Size by Degree Centrality
- 4 Bachelor Programme
Size by Degree Centrality
- 2 Research Master
Size by Degree Centrality

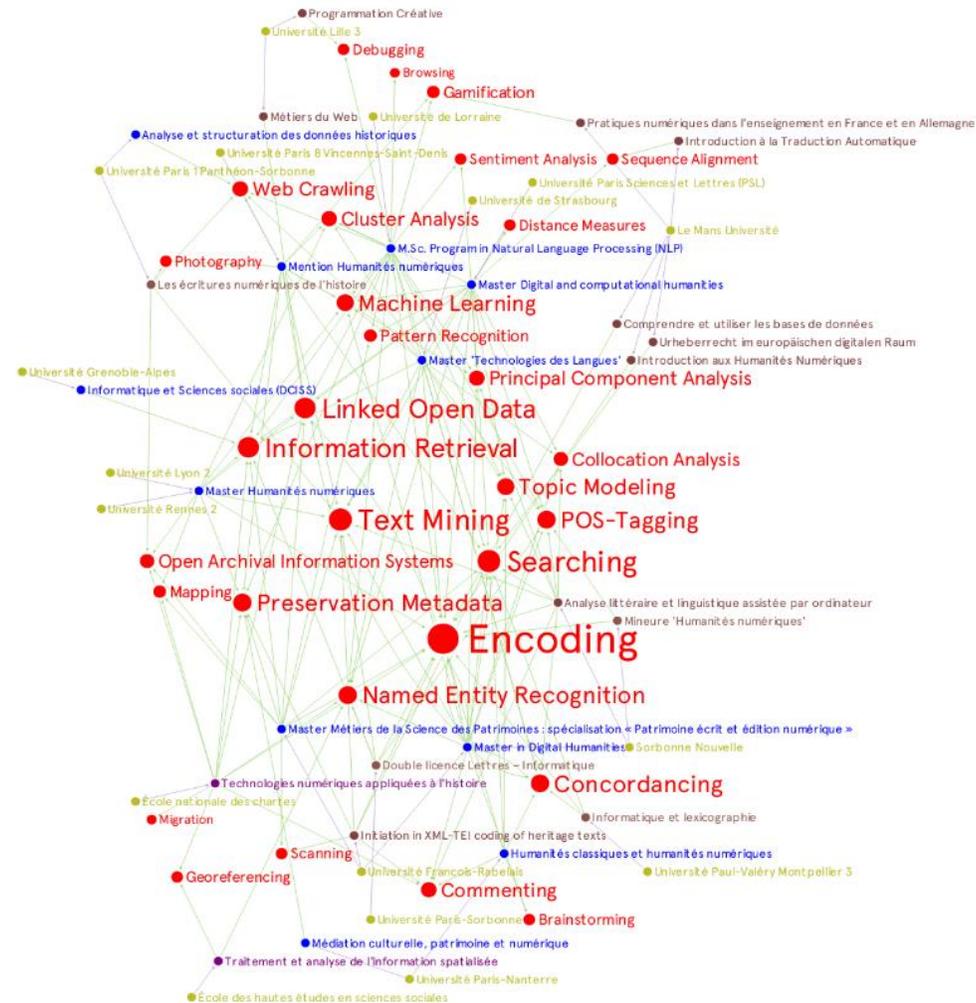
Questionnements sur la nature des termes

- La nomenclature des objets pose question : **certains objets désignent aussi des technologies.**
- La **généricité** de certains termes pose question :
 - l'utilisation des termes "data" ou "files" ou "tools" n'est pas spécifique aux H.N
- **Quels usages des « langages documentaires »**
 - la description des formations à l'aide des termes utilisés est-elle vue comme un système de descripteurs où chaque terme pris séparément fait sens précisément pour chaque item décrit, ou, est-ce l'adjonction conjointement exhaustive des termes qui permet la description complète des items ?
- Renvoie à **la pertinence du traitement** mené ici :
 - le découpage en fonction de la fréquence des citations des mots-clés pour réaliser la cartographie.
 - Le sens des termes utilisés, s'il se définit en référence à l'univers sémantique par proximité avec les autres termes utilisés pour un item, est perdu à l'occasion de ce découpage, par la suppression, via ce procédé, de l'environnement sémantique créé par les autres mots-clés utilisés pour décrire un item.

Cartographie des techniques

Cartographie des techniques

- 30** **Technique**
Size by metric_weighted_in_degree
- 17** **Institution**
Size by Neutral
- 11** **Master Programme**
Size by Neutral
- 9** **Course**
Size by Neutral
- 4** **Bachelor Programme**
Size by Neutral
- 2** **Research Master**
Size by Neutral



Une terminologie discutable

- Même type de disparités dans les données que précédemment présentées : le nombre de termes pour décrire les techniques allant d'un seul terme à 18 termes
- Plus globalement, **une forte hétérogénéité des sortes de techniques** incluses dans les termes pour décrire les formations concernant cette dimension « Techniques ».
 - « Linked open data » ou « Brainstorming » relèvent-ils bien d'une technique ?
 - Le terme lui-même de « Technique » est suffisamment ambigu pour inclure tout autant une méthode de réflexion (brainstorming), qu'une politique de publication, d'éditorialisation et de partage de descriptions et de liaisons de données (Linked open data).
- La **variabilité dans la spécificité des termes** pose elle aussi question :
 - est-il possible de mettre au même niveau des techniques de « machine learning » et d'« information retrieval » ?
 - Que signifie l'emploi d'un terme « Photography » dans le champ « Techniques » ? Ce terme ne serait-il pas plus approprié pour décrire les « Objets » des formations ?
- Revient à la question de la **qualité des données** utilisées pour réaliser ces visualisations et de la **variabilité dans la précision des termes** due aux **aspects contributifs de l'indexation** dans la base de données Dariah.

Cartographie des licences d'origine

Quelques remarques sur le rapprochement disciplines / filières

- Les **disparités dans les données rapprochées nécessitent une requalification** des disciplines mentionnées pour les filières d'origine des étudiants dans la nomenclature des disciplines de la base de données Dariah (ou l'inverse)
- Pourquoi les SIC sont-elles plus visibles dans la carte des filières ?
 - Est-ce l'appétence pour le numérique des étudiants de cette filière qui en fait un vivier favorable pour l'entrée dans des Master HN ?
 - ou bien la nature interdisciplinaire de cette section académique ?
 - ou encore l'intérêt pour l'écrit, le texte, la philologie, liée à la branche documentation des SIC?
- On observe également que **les Arts se situent en 5ème rang dans la cartographie des filières tandis qu'elle est en deuxième rang dans la cartographie des disciplines.**
- La **Philosophie passe du dernier rang** dans la cartographie des disciplines au **deuxième rang dans la cartographie des filières d'origine** des étudiants :
 - un effet d'aubaine pour l'affichage de débouchés porteurs pour les étudiants de filières dont on sait qu'elles sont difficiles en termes d'insertion professionnelle en France ?

Questionnements subsistants

- Comment recenser des formations dans un domaine hétérogène tel que celui des H.N.?
- Comment proposer des langages documentaires de description de ces formations, qui fassent consensus dans ce domaine émergent, et en recomposition constante dans les contextes locaux de mise en œuvre et de projets ? Et dans le contexte international ?
- Quelles sont les entrées pertinentes pour décrire ces formations ?
 - est-il souhaitable d'opter pour une approche par compétences ?

Perspectives

- Pour compléter ce travail de cartographie du champ des H.N. à travers les formations un programme semble intéressant :
 - le recensement des contenus de formations (analyse des maquettes)
 - leur décryptage à l'aide d'entretiens avec les responsables de ces formations,
 - Afin de recenser la diversité des approches, des contenus de formations, et des visions des H.N. qui leur sont corrélées.
 - Cette poursuite du travail sous forme d'enquête qualitative pourrait être complétée d'un travail maïeutique sur les trajectoires professionnelles des enseignants responsables de ces formations et sur leur profil académique.

Conclusion et critiques réflexives

- Les données sont partielles et incomplètes : elles correspondent à un état à un instant T des données présentes dans les bases prises comme sources (Web)
- Certaines formations changent d'intitulé et de programme au fur et à mesure des évolutions des disciplines, mais également des technologies numériques, et ce, au fil du déploiement international des conceptions des H.N
- Travail comparatif des données à l'international permettrait d'observer l'existence hypothétique de caractéristiques nationales dans l'appréhension du champ émergent des H.N.
- L'appartenance disciplinaire est le lieu d'enjeux de prises de position, compte tenu de l'opacité du fonctionnement des financements nationaux de la recherche en France.
- La description de ces formations dans les bases étudiées relève de processus de communication à l'intention des futurs étudiants
- La question du niveau de complexité et de détail à adopter selon les publics et les stratégies de positionnement disciplinaire nous semble cruciale

Merci pour votre attention !