



DELEGUE A LA GESTION DES DONNEES DE LA RECHERCHE - SCIENTIFIC DATA STEWARD



Présentation

Le projet de diplôme transversal « Scientific Data Steward » est fondé sur le constat d'un besoin de formation, exprimé par les communautés scientifiques, besoin résultant des lignes directrices du Plan National Science Ouverte (2018) et de la Commission européenne. Il s'agit de proposer à l'ensemble des communautés scientifiques (enseignants-chercheurs, chercheurs, ingénieurs, doctorants, étudiants).

La politique de science ouverte est également sous-tendue par les enjeux que présente actuellement « l'économie fondée sur les données » (IoT, IA, etc.). Il y a donc lieu aujourd'hui pour les établissements de recherche de produire des données réutilisables par les chercheurs eux-mêmes mais aussi par le secteur privé. Ceci suppose cependant que les données produites puissent être partagées et diffusées et donc « lisibles par machine ». Dans cette optique, l'utilisation de formats a été préconisée par les pouvoirs publics, soutenus par une partie de la communauté des chercheurs (format FAIR : facile à trouver, accessibles, interopérables, réutilisables). Ces contraintes conditionnent aujourd'hui le financement des projets de recherche et sont requises dans l'exécution des LabEx (formation des personnels, élaboration des Plans de Gestion des Données).

La formation, qui a déjà fonctionné quatre années consécutives (2020-2024) et qui a déjà produit une centaine de diplômés, se propose de répondre à une demande forte de formation en « ingénierie des données ». Une formation

permettant de présenter les aspects de gestion des données de la recherche (Scientific Data Steward) sous un angle à la fois technique, juridique et économique s'avère donc aujourd'hui nécessaire.

La formation s'adresse aux communautés scientifiques toutes disciplines confondues et quel que soit leur niveau.

Le DU Scientific Data Steward repose sur 2 piliers :

- Excellence – Des professeurs, ingénieurs de recherche, directeurs de recherche, professionnels de la santé et de la data, sont associés à l'équipe pédagogique du DU de Montpellier.
- Online – Le DU sera numérique et les enseignements en ligne. L'idée est de favoriser la souplesse dans la transmission des savoirs et d'augmenter le nombre d'étudiants.

Objectifs

Il s'agira de présenter l'environnement des données scientifiques sous deux angles :

1/ Un angle doctrinal, lié aux savoirs fondamentaux afin de délivrer un diplôme d'excellence :

La première partie de la formation permet de comprendre l'environnement des données de la science



Seront développés la recherche collaborative, le financement des recherches orientées vers la collecte et l'analyse des données, les stratégies de développement de

projets de data science ainsi que les règles de gouvernance qui s'appliquent aux données.

2/ Un angle pratique de « Scientific Data Steward » recherche

Ces deux angles d'approche seront organisés sous formes de modules qui représentent 186 heures de cours et de contenus pédagogiques en ligne.

Il s'agira de présenter les grands principes de la gouvernance des données de la recherche (I), d'expliquer les outils d'analyse des données de la recherche (II), ainsi que de comprendre comment gérer l'ouverture des données de la recherche (III) (v. maquette jointe).

I. Comprendre la gouvernance des données suppose que soient présentés les grands enjeux de la science ouverte.

II. Maîtriser les outils d'analyse des données scientifiques suppose de comprendre comment les données sont extraites, contextualisées, fouillées et explorées mais aussi de comprendre comment elles peuvent être sécurisées (au plus technique comme au plan juridique) et protégées.

III. Gérer l'ouverture des données scientifiques suppose de savoir comment et quand partager les données, comment les conserver, les stocker et les archiver et comment les valoriser (au plan scientifique comme au plan juridique).

Admission

Conditions d'admission

Pour les étudiants en formation initiale, Master validé.

Pour la formation continue, sur décision du responsable pédagogique.

Droits de scolarité

- **Formation initiale** : 700 €
- **Formation continue** : 3 500 €
- **Tarifs dérogatoires** :

200 € : Boursiers

350 € : Doctorants

2 000 € : Autofinancement

1 500 € : Demandeurs d'emploi, inscrits à Pôle Emploi à la suite d'un devis pour passage en commission de prise en charge

35 € / h € : Inscription à un seul module (Taux horaire augmenté de 28 %)

0 € : Réinscription en 2ème année pour finir le diplôme (chevauchement sur 2 années universitaires)

Pré-requis obligatoires

Pour les étudiants en formation initiale, Master validé.

Pour la formation continue, sur décision du responsable pédagogique.

Et après

Insertion professionnelle

La profession « Délégué à la gestion des données » devrait, à terme, voir le jour dans les laboratoires, les chercheurs et enseignants-chercheurs ne pouvant pas toujours assumer cette tâche. Il s'agit donc d'anticiper cette nécessité. Les tâches relatives à la gestion des données peuvent parfois également être confiées aux doctorants. Ceux-ci doivent donc être en mesure de comprendre quel est le sort des données qu'ils manipulent. Enfin, les chercheurs et enseignants-chercheurs doivent, quant à eux, comprendre



les enjeux et techniques utilisées pour stocker au mieux et archiver les données produites au sein d'un laboratoire par exemple, ou comment ils peuvent ou doivent concevoir les Plans de Gestion de Données. Ils doivent également être mesure de savoir ce qu'ils peuvent ou non faire des jeux de données lorsqu'ils sont en partenariat avec des entreprises du secteur privé. Enfin, les doctorants auxquels la formation est ouverte peuvent avoir intérêts à acquérir une « culture » de la gestion des données scientifiques lorsqu'ils souhaitent poursuivre une carrière dans la recherche ou simplement parce qu'ils manipulent déjà des jeux de données importants au sein de leur laboratoire, ou encore parce qu'ils ont un projet innovant à développer.

Infos pratiques

Contacts

Responsable pédagogique

Agnes Robin

✉ agnes.robin@umontpellier.fr

Lieu(x)

📍 Montpellier - Faculté de Droit et de Science politique



Programme

Organisation

Volume formation : 125 heures

→ Programme :

Module 1 : les grands principes de la gouvernance des données de la recherche

Module 2 : les outils d'analyse des données de la recherche

1. Extraire les données
2. Gérer les données
3. Explorer et analyser les données
4. Sécuriser les données

Module 3 : comprendre comment gérer l'ouverture des données de la recherche

1. Construire un plan de gestion de données
2. Stocker et archiver les données
3. Diffuser et partager les données
4. Valoriser les données