



UNIVERSITÉ
TOULOUSE III
PAUL SABATIER



Master 2 SID-Big data

FORMATION INITIALE ET ALTERNANCE

Année 2020-2021

Le M2 SID est la dernière année de la formation SID-Big Data. Depuis la rentrée 2016, les étudiants peuvent s'y inscrire en formation initiale ou en alternance (sous contrat de professionnalisation ou d'apprentissage).

LA FORMATION SID

Le parcours SID (Statistique et Informatique Décisionnelle) du L3 au M2 forme, depuis plus de 15 ans, au métier d'**ingénieur-expert en Sciences des données** au niveau Bac+5. Elle aborde tous les aspects de l'ingénierie des données : collecte, modélisation, administration, exploitation, valorisation et Big Data. Les étudiants y acquièrent une **double compétence** dans les méthodes et outils **statistiques et informatiques** pour l'aide à la prise de décision, avec notamment une spécialisation poussée en machine learning et en bases de données.

Qui recrute ? Tous les secteurs où la modélisation, l'administration et l'exploitation de l'information sont primordiales : Assurances et banques, Aéronautique et transports, Sociétés de service en informatique, Administration, Hôpitaux et secteurs médicaux, industrie agro-alimentaire, Recherche et Développement

Quels emplois ? Data Scientist, Ingénieur d'études, Concepteur et/ou administrateur de bases de données, Statisticien, Ingénieur Business Intelligence, Informaticien spécialiste des systèmes d'information, Ingénieur R&D

De nombreux atouts professionnels

- Nombreuses mises en situation : Stages (10 mois au total minimum sur les 3 années du parcours), projets, bureau d'études, challenge, ...
- Connaissance de l'entreprise, Gestion de projet, Anglais et Communication
- Nombreux intervenants issus du milieu industriel
- Le **label CMI** (Cursus de Master en Ingénierie) du réseau Figure

ADMISSION en M2

- Accès direct pour les titulaires du M1 SID de l'UPS
- Pour d'autres M1, accès sur dossier par validation des pré-requis thématiques.
- Possibilité de Validation des Acquis de l'Expérience (VAE), suite à une expérience professionnelle.

CONTACTS

- Site de la formation SID : <http://www.math.univ-toulouse.fr/sid>
- Cécile Chouquet et Gilles Hubert, co-responsables du M2 SID-Big data
- Mail : sid@math.univ-toulouse.fr

ORGANISATION DES ENSEIGNEMENTS EN M2

L'organisation du M2 permet l'accueil simultané des étudiants en formation initiale et en alternance.

- En alternance : Contrat de professionnalisation ou d'apprentissage d'une année répartie en 17 semaines à l'université et 34 semaines en entreprise (selon le planning fixé par l'équipe pédagogique).
- En formation initiale : A l'université de septembre à mi-mars (avec projets), puis stage de mi-mars à fin août.

Code NSF : 114b, 114g, 326p

Code ROME : H1206, M1805, M1802, M1201

Code NAF : J63.1, H49, H51, K64, K66, M72, M73.2, M74.9

7 COMPETENCES CLEFS POUR LE MASTER SID – Big Data

Au niveau Expertise

- Mettre en oeuvre une étude statistique de la planification jusqu'à l'analyse et la synthèse des résultats (Sondage, Marketing, Etude biomédicale, Maîtrise Statistique des Procédés).
- Concevoir et développer un système d'information (base de données relationnelle ou NoSQL) pour aider au fonctionnement d'une organisation.

Au niveau Maîtrise (en autonomie)

- Extraire de l'information pertinente des sources de données textuelles ou structurées pour les valoriser (aide à la décision, recherche d'information, fouille de données) en entreprise, dans une administration, ou dans un environnement de recherche.
- Analyser des masses de données et construire des tableaux de bord pour le pilotage d'institutions (entreprises, administration, etc.).
- Maintenir et analyser un système décisionnel (entrepôt de données, OLAP, ERP) pour gérer les ressources d'une organisation.
- Travailler en équipe en suivant des méthodologies de gestion de projet.
- Echanger et s'exprimer aisément à l'oral et à l'écrit en français et en anglais, dans un contexte professionnel.

Au niveau Application (avec guidage)

- Construire et valider un modèle mathématique pour traiter des ensembles de données (provenant de sondages, de relevés, de capteurs, etc.), afin de développer des outils d'aide à la décision en entreprise, dans une administration, ou dans un environnement de recherche.

PROGRAMME DES ENSEIGNEMENTS DU M2 (388 heures de formation)

- **Apprentissage pour les données massives**
Méthodes d'agrégation et d'optimisation pour l'apprentissage sur les données massives et flux de données, deep learning, aspects Big Data et Open Data
- **Manipulations de données volumineuses**
BD semi et non-structurées, BD massivement parallèles, Computing cloud, Business Intelligence
- **Analyse et exploitation de documents textes, audio et vidéo**
Objectifs, enjeux et applications pour le traitement d'images, de sons et de textes
- **Modélisation statistique avancée**
Méthodes statistiques spécifiques pour l'analyse de données complexes
- **Statistique pour l'industrie**
Approfondissement en Machine learning, Fiabilité des systèmes, Maîtrise statistique des procédés
- **Evaluation et Certification Logiciel**
Certification logiciel, processus, évaluation, traçabilité des exigences et diagnostic Qualité, défaillance
- **Communication et management en entreprise**
Gestion de projet, propriété intellectuelle, droit des données et techniques de communication
- **NoSQL et expériences industrielles en Science des données**
Panorama et manipulation des bases de données NoSQL - Conférences de partenaires professionnels
- **Recherche, développement et valorisation**
Coordination entre plusieurs rôles d'une équipe sur une mission R&D en Open Data ou Big Data
- **Anglais**

PROJETS EN M2 POUR LES NON-ALTERNANTS

- BE « Statistique & Santé ou Industrie »
- BE « Evaluation et Certification Logiciel »
- TER Challenge « Analyse et exploitation de documents textes, audio et vidéo »