

DÉLIBÉRATION

Conseil d'administration

Séance du 7 juillet 2020

Délibération
n°93-2020
Point 4.4.9.5

Point 4.4.9.5 de l'ordre du jour

Mise en place des enseignements « Approche interdisciplinaire en sciences des données »

EXPOSE DES MOTIFS

Les Sciences des données sont aujourd'hui un enjeu à la fois scientifique, sociétal et économique. Si l'université doit être leader de cette révolution, et en particulier pour la recherche en développant les infrastructures lui permettant d'interroger l'ensemble des données à disposition par une approche fondamentalement pluridisciplinaire, le lien avec la formation est également évident et les enjeux en sont multiples. Aujourd'hui aucune discipline ne peut ignorer les capacités à croissance exponentielle de stockage et de calcul des systèmes d'information. L'accès aux bases de données existantes, la numérisation de collections d'échantillons rares révolutionnent de nombreuses disciplines dans les quatre domaines de l'ESR : ALL, SHS, DEG et STS. Tout étudiant diplômé de l'Unistra doit donc être aujourd'hui familiarisé avec ces problématiques nouvelles.

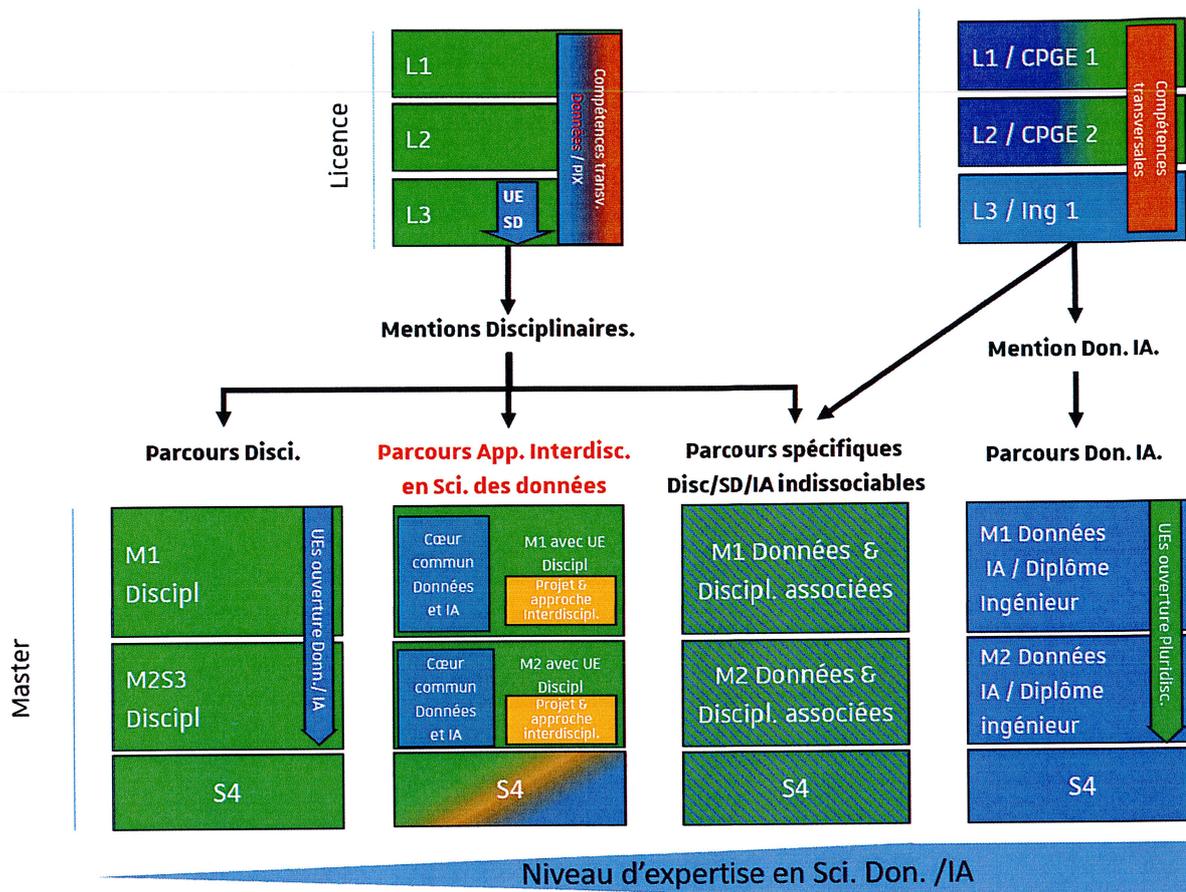
A l'Université de Strasbourg, il existe actuellement 2 types de formation s'intéressant à cette question.

- En premier lieu, l'Unistra propose des **filières dont le cœur de formation est propre aux sciences des données** et à l'intelligence artificielle. (*Parcours Don / IA sur le schéma*)
- Parallèlement, des formations proposent des enseignements en sciences des données dans des **filières dont l'intitulé principal est autre mais où l'acquisition des compétences disciplinaires visées est indissociable d'une formation spécifique et très pointue en traitement des données** et à l'intelligence artificielle (*Parcours spécifiques Disc/SD/IA indissociables*)

Un groupe de travail chargé de faire des propositions pour faire monter en compétences en sciences de la donnée tous les étudiants qui ne sont pas dans des filières décrites ci-dessus a préconisé de développer deux approches complémentaires :

- La première consiste à **proposer à toutes les mentions de master des UE optionnelles ou complémentaires** (*Parcours Disc. Sur le schéma*)
- Pour répondre aux attentes d'étudiants qui souhaiteraient devenir des experts de ces questions dans/pour leur discipline d'origine, le GT formule aussi une 2^{ème} approche, qui sera, en ce sens, la spécificité de l'offre strasbourgeoise. Cette **nouvelle filière existera comme un parcours bien identifié qui pourra être proposé dans toutes les mentions qui le souhaiteront**. Pour toutes les mentions concernées, l'intitulé de ce parcours sera identique : **«Approche interdisciplinaire en Sciences des Données»**.

Le bloc d'UE spécifique Sciences des données aura un volume équivalent à 30 ECTS sur les 120 ECTS que compte le master.



Pour l'année universitaire 2020-2021, il est proposé de créer ces UE et de les intégrer par contrat pédagogiques aux masters existants et qui ont manifesté leur intérêt. Les masters existants bénéficieront des UE ainsi créées, par ajout ou substitution, voire transformation de ces cours à certains des leurs.

Par 26 voix pour, la CFVU a approuvé la mise en place des enseignements «Approche interdisciplinaire en sciences des données».

Délibération

Le Conseil d'administration de l'Université de Strasbourg approuve la mise en place des enseignements «Approche interdisciplinaire en sciences des données».

Résultat du vote :

Nombre de membres en exercice	37
Nombre de votants	27
Nombre de voix pour	27
Nombre de voix contre	0
Nombre d'abstentions	0

Destinataires :

- Madame le Recteur de l'académie, Chancelier des universités
- Direction Générale des Services
- Direction des Finances
- Agence Comptable

Fait à Strasbourg, le 15 juillet 2020

Le Directeur Général des Services



Fredéric DEHAN

DESCRIPTIF DE L'UE

	Formation de rattachement	Intitulé de l'enseignement
	L3 Université de Strasbourg	Initiation à la science des données

RESPONSABLES DE L'ENSEIGNEMENT

Nom/Prénom	Discipline	Composante / Courriel/Tél. (fixe)
MAUMY-BERTRAND MYRIAM	Informatique et mathématique	mmaumy@math.unistra.fr
BERNHARD DELPHINE		dbernhard@unistra.fr
CHAPUIS YVES-ANDRE		ya.chapuis@unistra.fr
JOSSINET FABRICE		f.jossinet@ibmc-cnrs.unistra.fr
LACHICHE NICOLAS		nicolas.lachiche@unistra.fr
MEILLIER CELINE		meillier@unistra.fr

MODALITES D'ORGANISATION DE L'ENSEIGNEMENT

Semestre 6 Licence / 2 heures par semaine UE Libre ou intégrée dans le Semestre 6 de chaque composante.
--

LANGUE ET NIVEAU DE LANGUE DE L'ENSEIGNEMENT :

FRANÇAIS (SUPPORT EN FRANÇAIS)

DESCRIPTION DE L'ENSEIGNEMENT

Dans un contexte scientifique, sociétal et économique où stockage, archivage, traitement et valorisation des données sont devenus des enjeux majeurs, cette U.E. permettra de découvrir les bases de la science des données, c'est-à-dire les méthodes et outils, théoriques et pratiques, pour comprendre les données : analyse, visualisation, apprentissage de modèles... Les cours seront faits en salle informatique, avec un poste de travail par étudiant et des manipulations fréquentes sur des exemples. Toutes les notions du cours seront illustrées à partir d'une base de données sur le cinéma. Aucune expérience en programmation n'est requise, cependant est attendu des étudiants un intérêt et une curiosité pour les outils informatiques et les notions élémentaires de statistiques (moyenne, écart-type...).

Plan du cours :

Initiation au notebook Python

Introduction – les différents types de données

Représentation statistique des données

- A) Représentation numérique des données
- B) Représentation graphique des données
- C) Importance de la représentation graphique

Exploration des données

- A) Recueil des données
- B) Prétraitements
- C) Relation entre les variables

Modélisation, prédiction et apprentissage

- A) Régression linéaire
- B) Classification
- C) Méthodes d'évaluation

Mise en pratique des connaissances à travers un projet encadré permettant d'approfondir les connaissances introduites dans les supports de cours.

Webographie :

Jeux de données et challenges : <https://www.kaggle.com>

Package Python dédié à la science des données : <http://scikit-learn.org> et <https://pandas.pydata.org>

Package Python dédié à la visualisation des données : <https://seaborn.pydata.org>

Bibliographie :

Data science : fondamentaux et études de cas. Machine learning avec Python et R, Eric Biernat et Michel Lutz, Eyrolles.

Big data et machine learning - Le concept et les outils de la data science, 2ème édition, Pirmin Lemberger, Marc Batty, Médéric Morel, Jean-Luc Raffaelli, Dunod.

COMPETENCES VISEES

Objectifs en termes de connaissances

- Contrôler la qualité, la cohérence, la fiabilité, la représentativité des données à traiter.
- Savoir transformer des données pour les rendre utilisables par un algorithme de traitement de données.

- Expliquer et interpréter un jeu de données à partir des résultats fournis par les outils de prédiction, classification, covariance, etc.
- Résumer et expliquer un jeu de données à ses pairs.

Objectifs en termes de compétences

Compétences disciplinaires :

- Lire des données statistiques sous différentes formes, interpréter un tableau croisé, pour produire un document de synthèse de données statistiques.
- Se servir aisément des outils et méthodes de recueil, de traitement et d'analyse des données pour observer et analyser les phénomènes et/ou les comportements du sous domaine.
- Concevoir le traitement informatisé d'informations de différentes natures, telles que des données, des images et des textes.
- Exploiter des logiciels d'acquisition et d'analyse de données avec un esprit critique.
- Mettre en œuvre des techniques de programmation, par l'exemple, notamment pour développer des applications simples d'acquisition et de traitements de données.
- Analyser et interpréter les résultats produits par l'exécution d'un programme.
- Mobiliser les outils mathématiques nécessaires à la modélisation.
- Valider un modèle par comparaison de ses prévisions aux résultats expérimentaux et apprécier les limites de validité.

Compétences transversales :

- Analyser et synthétiser des données en vue de leur exploitation.
- Développer une argumentation avec esprit critique.

ENSEIGNEMENT – ORGANISATION HORAIRE

Titre de l'enseignement	CM	CI	TD	TP
		28h		
Eq. HETD= 35h				

MODALITES D'EVALUATION DES ENSEIGNEMENTS

(session initiale et de rattrapage)

Évaluations écrites multiples (comptant pour 50% de la note finale) réalisées en séances au cours du semestre. Pas de convocation.

Évaluation (qui comptera pour 50% de la note finale) sous la forme d'un rendu de projet (notebook commenté permettant le traitement et l'analyse du jeu de données fourni en projet). Pas de convocation, rendu à la date fixée.

Pas de session de rattrapage.

Mise en place des parcours « Approche interdisciplinaire en sciences des données »

Les Sciences des données sont aujourd'hui un enjeu à la fois scientifique, sociétal et économique. Si l'université doit être leader de cette révolution, et en particulier pour la recherche en développant les infrastructures lui permettant d'interroger l'ensemble des données à disposition par une approche fondamentalement pluridisciplinaire, le lien avec la formation est également évident et les enjeux en sont multiples. Aujourd'hui aucune discipline ne peut ignorer les capacités à croissance exponentielle de stockage et de calcul des systèmes d'information. L'accès aux bases de données existantes, la numérisation de collections d'échantillons rares révolutionnent de nombreuses disciplines dans les quatre domaines de l'ESR : ALL, SHS, DEG et STS. Tout étudiant diplômé de l'Unistra doit donc être aujourd'hui familiarisé avec ces problématiques nouvelles.

A l'université de Strasbourg, il existe actuellement 2 types de formation s'intéressant à cette question.

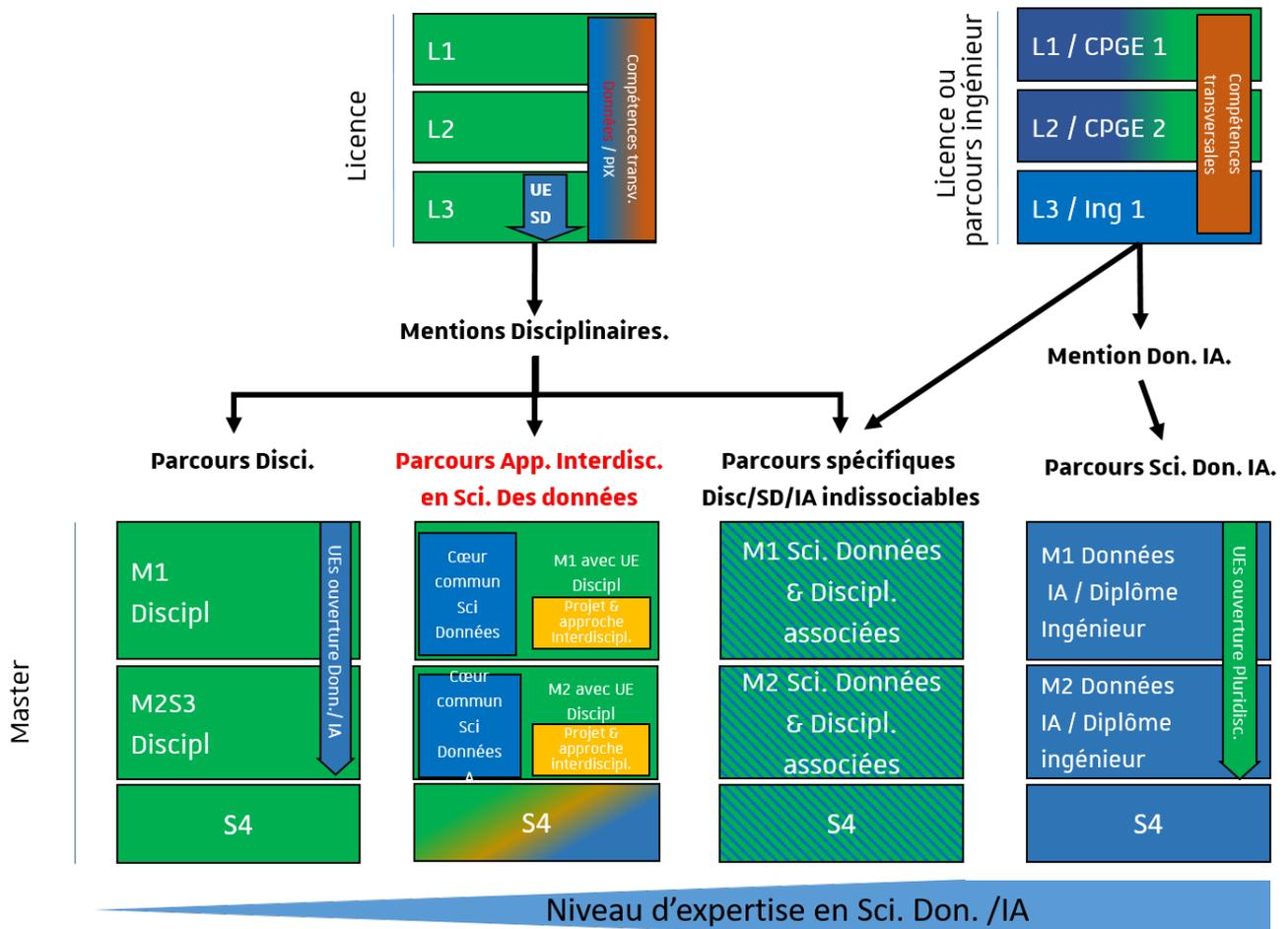
En premier lieu, l'Unistra propose des filières dont le cœur de formation est propre aux sciences des données et à l'Intelligence artificielle qui ont pour objectif de former des spécialistes des mathématiques de l'informatique et en algorithmique au stockage, au traitement, à l'interrogation, et à la fouille des données. Ces diplômés sauront mettre leurs compétences au service de missions très diversifiées dans tous les secteurs socio-économiques et industriels. (*Parcours Don / IA sur le schéma*)

Parallèlement, des formations proposent des enseignements en sciences des données dans des filières dont l'intitulé principale est autre mais où l'acquisition des compétences disciplinaires visées est indissociable d'une formation spécifique et très pointue en traitement des données et à l'intelligence artificielle. Ces enseignements sont donc difficilement mutualisables en dehors du champs disciplinaires visées (astrophysique, démographie, ou génomique par exemple). (*Parcours spécifiques Disc/SD/IA indissociables*)

Le groupe de travail chargé de faire des propositions pour faire monter en compétences en sciences de la donnée tous les étudiants qui ne sont pas dans des filières décrites ci-dessus a préconisé de développer 2 approches complémentaires :

1. La première consiste à proposer à toutes les mentions de master des UE optionnelles ou complémentaires, qui pourront même le cas échéant donner lieu à l'attribution d'un DU, et qui permettront à des futurs diplômés de toutes les formations de master de l'Unistra d'acquérir des compétences en traitement des données et à l'Intelligence artificielle (niveau initiation jusqu'à avancé ?). (*Parcours Disc. Sur le schéma*)
2. Cependant, pour répondre aux attentes d'étudiants qui souhaiteraient devenir des experts de ces questions dans/pour leur discipline d'origine, une 2^{ème} approche, qui sera, en ce sens, la spécificité de l'offre strasbourgeoise, cette nouvelle filière existera comme un parcours bien identifié qui pourra être proposé dans toutes les mentions qui le souhaiteront : Pour toutes les mentions concernées, l'intitulé de ce parcours sera identique : « Approche interdisciplinaire en Sciences des Données ». Tous ces parcours proposeront donc un tronc commun construit autour d'un cœur d'UEs consacrées aux traitements de données & IA – proposé comme un « package complet » aux composantes. Ce cœur étant l'originalité et le point central de la cohérence de la formation, il sera incontournable. Autour de ce cœur, les étudiants recevront un « approfondissement thématique » en lien avec les disciplines des composantes d'origine, qui correspondra dans la plupart des cas à des enseignements spécifiques d'un parcours existant de la mention dans laquelle ils sont inscrits (voir le paragraphe consacré aux modalités d'admission des étudiants). C'est en effet à partir de leurs disciplines d'origine que tous les étudiants qui suivent cette formation entrent dans une étude commune du traitement des données et de l'IA. C'est donc en spécialiste de leur discipline d'origine et avec une expertise commune acquise en

traitement des données et IA qu'ils aborderont la confrontation de leurs points de vue, cette formation sera ainsi un point de nucléation pour de nombreux travaux pluridisciplinaires. ((*Parcours App. Interdisc. en Sci. des Données sur le schéma*)).



Ce nouveau parcours est avant tout mise en place pour répondre aux besoins des formations qui n'intègrent pas pour l'instant d'enseignements en sciences des données. C'est pour cela que l'intitulé se doit de le préciser de façon explicite. L'intitulé unique pour toutes les mentions renforce la volonté de l'établissement de permettre à des diplômés de monter en compétences pour répondre aux attentes de plus en plus forte du monde professionnel en la matière.

Les objectifs pédagogiques des enseignements spécifiques à ce parcours seront fondamentalement axés sur l'acquisition de compétences et sur une approche favorisant l'interdisciplinarité. La pédagogie par projet y aura une place importante.

Le bloc d'UE spécifiques aux sciences des données aura un volume équivalent à 30 ECTS.

A noter que ce parcours n'ouvrira pas à la rentrée 2020 mais à la rentrée 2021. Cependant, dès janvier 2021 (donc pour le prochain semestre pair), des UE pourront être suivies par les étudiants intéressés, et en premier lieu ceux qui ont suivi l'UE proposée cette année en L3.

1) Contenus des UE du bloc spécifique au parcours AISD (présentation non définitive)

Blocs de compétences	Objectifs	UEs
Enseignements Tronc Commun 1 - Concepts généraux (6 ECTS)	Objectif : Donner à l'étudiant une vision globale et transversale de la Science des	C-1 : Concepts, enjeux et transdisciplinarité C-2 : Ethique et droit
Enseignements Tronc Commun 2 - Méthodologies (9 ECTS)	Objectif : Former l'étudiant aux méthodes et outils de base de l'analyse des données et de l'extraction de connaissances à partir de données massives	M-1 : Méthodes d'apprentissage supervisé M-2 : Méthodes d'apprentissage non-supervisé et d'apprentissage profond M-3 : Méthodes symboliques M-4 : Méthodes numériques
Enseignements Tronc Commun 3 - Données : de la collecte aux traitements (6 ECTS)	Objectif : Donner à l'étudiant des capacités à déployer des moyens de collecte, de stockage et d'exploitation des données conformes aux enjeux sociétaux associés	D-1 : Collecte, fiabilité et visualisation des données D-2 : Modélisation et traitement des données
Enseignements SDA : "Sciences des données , mise en oeuvre et applications" (9 ECTS)	Objectif : Confronter l'étudiant à des cas réels d'utilisation et de recherche dans son domaine de prédilection	A-P1 : Projet transdisciplinaire I A-P2 : Projet transdisciplinaire II A-TER : Travail d'Etude et de Recherche A-S4 : Stage (Validation AISD par un jury AISD)

2) Rattachement administratif des UE spécifiques du parcours AISD

Ces UE seront rattachées à l'UFR Mathématique et informatique.

3) Modalités d'admission en parcours AISD (pour la rentrée 2021)

Chaque mention qui le souhaite pourra ouvrir un seul parcours AISD, il n'y a aura pas de parcours AISD spécifique à tel ou tel parcours, mais un parcours par mention.

L'admission en parcours AISD imposera cependant à l'étudiant de préciser le(s) parcours de la mention concernée dans lequel il prendra tout ou partie des 60 ECTS disciplinaires. Le(s) responsable(s) du parcours disciplinaire(s) de rattachement devra(ont) valider la candidature. Si la candidature n'est pas validée par un responsable de parcours disciplinaire de rattachement, elle ne pourra pas être acceptée dans le parcours AISD.

Si l'étudiant n'est pas admis dans le parcours AISD de la mention de son choix, son accès dans un autre parcours ne sera possible que s'il a déposé une autre candidature (hors AISD) pour le(s) parcours envisagé(s).

1^{ère} étape : pré-sélection au sein des mentions : *(Cette partie de la procédure doit être terminée fin mai)*

Les responsables de parcours disciplinaires auront identifié les candidatures qu'ils jugent recevables. Sur cette base, le porteur du parcours AISD (par exemple le responsable de la mention), en collaboration avec les responsables des parcours concernés établira la liste des candidats pré-sélectionnés.

2^{ème} étape : commission inter-mention spécifique aux parcours AISD *(Cette partie de la procédure doit être terminée pour le 15 juin)*

Une commission spécifique, comprenant les porteurs de parcours AISD et des responsables des SD nommés es-qualité (des différentes mentions concernées) se réuniront pour harmoniser les procédures et établir la liste des candidats définitivement retenus pour une inscription dans l'un des parcours AISD.

4) Inscription pédagogique aux 60 ECTS spécifiques aux différentes mentions

Pour tous les parcours qui étudieront des candidatures et qui proposeront donc des UE que les étudiants des parcours AISD suivront, la liste de ces UE devra être arrêtée, par parcours, assez rapidement (avril 2020) pour que la construction administrative des parcours se réalise dans les meilleures conditions.



Programme IdEx Initiative d'excellence

Nom du projet : **Parcours commun de master**

« **Approche Interdisciplinaire en Science des Données** »

Structure porteuse : Vice-Présidence Transformation numérique et innovations pédagogiques / l'UFR de Mathématique et d'Informatique comme faculté porteuse au nom des composantes impliquées

Description du projet :

La Science des données est aujourd'hui un enjeu à la fois scientifique, sociétal et économique. Si l'université doit être leader de cette révolution, et en particulier pour la recherche en développant les infrastructures lui permettant d'interroger l'ensemble des données à disposition par une approche fondamentalement pluridisciplinaire, le lien avec la formation est également évident et les enjeux sont multiples. Aujourd'hui aucune discipline ne peut ignorer les capacités à croissance exponentielle de stockage et de calcul des systèmes d'information. L'accès aux bases de données existantes, la numérisation de collections d'échantillons rares révolutionnent de nombreuses disciplines dans les quatre domaines de l'ESR : ALL, SHS, DEG et STS. Tout étudiant diplômé de l'Unistra doit donc être aujourd'hui familiarisé avec ces problématiques nouvelles.

A l'université de Strasbourg, il existe actuellement 2 types de formation s'intéressant à cette question.

En premier lieu, l'Unistra propose des filières dont le cœur de formation est propre à la science des données et à l'Intelligence artificielle qui ont pour objectif de former des spécialistes des mathématiques de l'informatique et en algorithmique au stockage, au traitement, à l'interrogation, et à la fouille des données. Ces diplômés sauront mettre leurs compétences au service de missions très diversifiées dans tous les secteurs socio-économiques et industriels.

Parallèlement, des formations proposent des enseignements en science des données dans des filières dont l'intitulé principal est autre mais où l'acquisition des compétences disciplinaires visées est indissociable d'une formation spécifique et très pointue en traitement des données et à l'intelligence

artificielle. Ces enseignements sont donc difficilement mutualisables en dehors des champs disciplinaires visés (astrophysique, démographie, ou génomique par exemple).

Cependant, pour répondre aux attentes d'étudiants qui souhaiteraient devenir des experts de ces questions Données dans/pour leur discipline d'origine, l'Université de Strasbourg proposera un 3^e type de formation qui sera la spécificité de l'offre strasbourgeoise.

Cette nouvelle filière existera comme un parcours bien identifié qui pourra être proposé dans toutes les mentions de master qui le souhaiteront. **Pour toutes les mentions concernées, l'intitulé de ce parcours sera identique : « Approche Interdisciplinaire en Science des Données ». Tous ces parcours proposeront donc un tronc commun construit autour d'un cœur d'UEs consacrées aux traitements de données- proposé comme un « package complet » aux composantes.** Ce cœur étant l'originalité et le point central de la cohérence de la formation, il sera incontournable. Autour de ce cœur, les étudiants recevront un « approfondissement thématique » en lien avec les disciplines des composantes d'origine, qui correspondra dans la plupart des cas à des enseignements spécifiques d'un parcours existant de la mention dans laquelle ils sont inscrits. C'est en effet à partir de leurs disciplines d'origine que tous les étudiants qui suivent cette formation entrent dans une étude commune du traitement des données et de l'IA. C'est donc en spécialiste de leur discipline d'origine et avec une expertise commune acquise en traitement des données et IA qu'ils aborderont la confrontation de leurs points de vue, cette formation sera ainsi un point de nucléation pour de nombreux travaux pluridisciplinaires.

Les objectifs pédagogiques des enseignements spécifiques à ce parcours seront fondamentalement axés sur l'acquisition de compétences et sur une approche favorisant l'interdisciplinarité. La pédagogie par projet y aura une place importante.

Première simulation de maquette du bloc d'UE spécifiques aux sciences des données (volume équivalent à 30 ECTS) :

Blocs thématiques	Objectifs	UEs
Concepts généraux (6 ECTS)	Donner à l'étudiant une vision globale et transversale de la Science des Données et de l'Intelligence Artificielle	C-1 : Concepts, enjeux et transdisciplinarité C-2 : Ethique et droit
Méthodologies (12 ECTS)	Former l'étudiant aux méthodes et outils de base de l'analyse des données et de l'extraction de connaissances à partir de données massives	M-1 : Méthodes d'apprentissage supervisé M-2 : Méthodes d'apprentissage non-supervisé et d'apprentissage profond M-3 : Méthodes symboliques M-4 : Méthodes numériques
Données : de la collecte aux traitements (9 ECTS)	Donner à l'étudiant des capacités à déployer des moyens de collecte, de stockage et d'exploitation des données conformes aux enjeux sociétaux associés	D-1 : Collecte, fiabilité et visualisation des données D-2 : Outils statistiques D-3 : Représentation et traitement des données
Applications et mise en oeuvre (3 ECTS)	Confronter l'étudiant à des cas réels d'utilisation et de recherche dans son domaine de prédilection	A-1 : Projet transdisciplinaire A-2 : Stage (Validation AISD par un jury AISD)

Depuis le 2nd semestre de l'année universitaire 2019-2020, une **UE d'initiation est proposée aux étudiants de licence**, notamment en vue de leur orientation vers ce parcours de master à venir.

Objectifs :

L'objectif premier est d'assurer que l'université forme ses étudiants aux compétences Données en mettant en place des formations audacieuses et pluridisciplinaires qui répondent aux besoins socio-économiques tendus actuels et anticipant ceux à venir.

Riche de sa pluridisciplinarité, il s'agit pour l'Unistra de permettre très prochainement à des étudiants spécialistes d'autres disciplines que celles du cœur Données, de devenir dans leur champ de compétence les référents en termes de data management par une formation transversale proposée à des étudiants d'origines disciplinaires différentes.

Ce projet repose sur le travail et l'investissement d'un groupe d'enseignants-chercheurs fort de leur double expertise Données et de leurs différentes disciplines ainsi que d'un portage au plus haut niveau de l'université. Cette approche permet de :

- accompagner et coordonner la mise en place des formations proposées en garantissant une cohérence d'ensemble et une complémentarité dans l'offre de formation
- accompagner les équipes pédagogiques chargées de la mise en place de ces formations pour que leur élaboration soit conduite en mode projet en intégrant une approche programme
- assurer la pertinence et la cohérence des UEs proposées dans les enseignement d'initiative en L3 et dans ceux du parcours commun AISD
- stimuler au sein de ces équipes le développement d'approches pédagogiques innovantes (formation hybride, en mode projet, etc.).

Valeur ajoutée et évolution attendue :

Développer de nouveaux axes stratégiques autour du traitement des données dont l'interdisciplinarité sera la spécificité.

Comme précisé, cette nouvelle formation AISD s'inscrit dans un ensemble plus large de formations ayant trait au traitement massif de données. Il sera important de valoriser **sous une même bannière**, aussi bien ce nouveau parcours différenciant par rapport à d'autres universités que cette offre globale de formation Unistra reflétant la richesse disciplinaire de notre université.

Impact numérique :

Le porteur précisera ici les éventuels développements informatiques et numériques engendrés par le projet nécessitant l'implication de la DUN/DI : achat de logiciels, intégration au réseau, besoin en développement, création/captation d'image et son, création de supports graphiques...

Des achats de postes informatiques, de logiciels sont prévus.

Impact bâtiminaire :

Le porteur précisera ici les travaux ou aménagements engendrés par le projet nécessitant l'implication de la DPI/DALI : commande de mobilier, réfection de salle, travaux de mise aux normes, construction/réfection de bâtiments...

Crucial : l'attribution d'une salle dédiée à équiper

Gouvernance :

Le Parcours commun de master « Approche Interdisciplinaire en Science des Données » est rattaché à l'UFR de Mathématique et d'Informatique et sera sous la responsabilité de Myriam Maumy-Bertrand et Nicolas Lachiche. L'équipe pédagogique sera composée d'enseignants-chercheurs de différentes composantes et renforcée par l'arrivée d'une fellowship IdEx Formation, atout important pour la réussite du projet.

Comme déjà indiqué, le projet est piloté par le Vice-Président Transformation numérique et innovations pédagogiques et porté par un groupe ad-hoc constitué d'une dizaine d'experts issus de différentes disciplines et donc de composantes. Ce corps constitue le comité de pilotage du projet.

Il fera 1 à 2 fois par an un retour sur l'avancée du projet au Comité de Pilotage Sciences et Données Unistra piloté par la Vice-Présidente Stratégie et développement afin de recueillir les recommandations des membres et experts.

Phasage :

Phase 1 – 2020-2021 :

- déploiement à plus grande échelle de l'UE d'initiation en L3
- programme transitoire de master – niveau M1 - avant son ouverture officielle à la rentrée 2021 (notamment pour les étudiants ayant suivi les UEs d'initiation en L3 au 2nd semestre 2019-2020) ; les étudiants seront inscrits dans leurs masters de discipline

Phase 2 – 2021 -2022

- poursuite de l'UE d'initiation en L3
- lancement plein et entier du nouveau parcours commun de master AISD au sein de différentes composantes.

Soutenabilité du projet :

Le porteur précisera ici le devenir du projet au-delà de la période de financement IdEx : arrêt du projet, prolongation sur autre source de financement....

Au vu du retour d'expérience sur l'offre de formation actuelle, le projet pourra s'inscrire pleinement dans la prochaine offre 2023-2028.

Coût global : 250 k€ ; Aide demandée : 150 k€

Parcours AISD (de la licence au master)

Année 2019-2020 :

Heures complémentaires pour les UE de L3 : 4264 €

Années universitaires 2020-2021 et 2021-2022

1/ Fellowship IdEx à hauteur de 116k€

NB : Une enveloppe de 100k€ (pour un an) avait été votée au Copil IdEx de février 2019 ; 16k€ sont sollicités en plus pour ce COPIL IdEx.

2/ Autres coûts

Masses	Rubriques	Année 2020-2021	Année 2021-2022
		Montants en €	Montants en €
Personnel	Heures comp	12 844,00€	38 549,00€
	Responsabilité péda (L3 + 2xM)	3 744,00€	3 744,00€
Equipement	Salles/mobilier	?	
	Postes informatiques	30 000,00€	
	Logiciels	?	
Fonctionnement	Petit budget de fonct.t	1 000,00€	1 500,00€

	Communication	600,00€	1 200,00€
Total par année		48 188,00€	44 993,00€
TOTAL (€)		93 181,00 €	

Bannière Unistra

- Partage d'outils par l'ensemble de la communauté Sciences & Données Unistra à des fins d'enseignement : 30 K€
- Communication globale Unistra « Sciences et Données » en formation (notamment avant l'ouverture de Parcoursup et les candidatures en master et de la remontée d'information sur trouvermonmaster.gouv.fr) : 5 000 €

Compétence	Objectif	Répartition semestrielle et crédits associés				Programme	Forme et projet	
		S1	S2	S3	S4			
Bloc de compétences CTC : "Concepts, outils et défis SDIA" - Tronc commun (18 ECTS)								
CTC 1: Concepts, enjeu et transdisciplinarité	Objectif : Donner à l'étudiant une vision globale et transversale de la Sciences des Données et de l'Intelligence Artificielle	S1 : Introduction - Présentation des domaines d'application et des aspects "Métier" des projets transdisciplinaires (3 ECTS)	S2 : Projet transdisciplinaires I : Analyse du contexte et proposition de solutions - Aspects sociaux et éthiques - Preuve de concepts (3 ECTS)			6	Contenu : - Les principaux paradigmes des données et des traitements - Principaux défis - Approche méthodologique - Vision transdisciplinaire y compris rappel de l'enseignement L3 Contenu : - Outils pour l'analyse et le traitement de données (suite de l'UE statistiques du L3) - Apprentissage automatique et de traitement des données (Classification supervisée et non supervisée, aide à la décision ...) - Méthodes et outils : Clustering, arbres de décision, SVM, apprentissage profond ... Contenu : - Collecte, stockage, partage des données - Protection et sécurité des données - Aspects juridiques - Ethique et déontologie	Forme : - Cours / Séminaire - Projets co-encadrés par les différentes composantes, dans l'objectif d'obtenir une synergie autour de la « Donnée » entre les différents domaines Forme : - Cours / Séminaire - Projets et TP d'utilisation des outils de base tels que Python, R - Mise en oeuvre dans le cadre de TP et mini-projet personnel Forme : - Cours/TD/TP - Liens forts avec les projets transverses
CTC 2 : Approches méthodologiques et outils en Science des Données et IA	Objectif : Former l'étudiant aux méthodes et outils de base de l'analyse des données et de l'extraction de connaissances à partir de données massives	S1 : Outils et méthode de base - Classification non supervisée (3ECTS)	S2 : Apprentissage supervisé (3ECTS)					
CTC 3 : De la collecte des données aux défis sociétaux	Objectif : Donner à l'étudiant des capacités à déployer des moyens de collecte, de stockage et d'exploitation des données conformes aux enjeux sociétaux associés	S1 : Aspects techniques (3ECTS) OU Ethique et droit (3 ECTS) -- A voir ?	S2 : Ethique et droit (3 ECTS) ou Aspects techniques (3ECTS) -- A voir ?					

Facultatif : Stage initiative étudiant (7 à 8 semaines) lié au projet transdisciplinaire

Bloc de compétences CRT : "Représentation et traitement avancés des données" (6 ECTS) dont 3 peuvent venir d'une option propre à la composante								
Option CRT 1: Données symboliques	Objectif : Donner à l'étudiant les connaissances nécessaires à la représentation et au traitement des données symboliques majoritairement rencontrées en Sciences Humaines, Droit, ...		S2 : Aspects théorique et méthodologique (3ECTS)	S3 : Projet transdisciplinaires II : Mise en pratique du projet transdisciplinaire (3 ECTS) -			Contenu : - Collecte, traitement de données symboliques : Texte, opinions, TAL ... - Confiance et qualité des données - Modélisation des données hétérogènes multisources - Représentation des connaissances : ontologies... - Partage des données, web sémantique Contenu : - Collecte, pré-traitement et mise en forme des données - Modélisation des données hétérogènes multisources - Analyse de séries temporelles, d'images ...	Forme : - Cours/TD - TP en liens forts avec le(s) domaine(s) d'application des étudiants - Un projet (type TER : Travail d'Etude et de Recherche) en liens forts avec le(s) domaine(s) d'application des étudiants Forme : - Cours/TD - TP en liens forts avec le(s) domaine(s) d'application des étudiants - Un projet en liens forts avec le(s) domaine(s) d'application des étudiants
Option CRT 2: Données numériques	Objectif : Donner à l'étudiant les connaissances nécessaires à la représentation et au traitement des données numériques majoritairement rencontrées en Ingénierie		S2 : Aspects théorique et méthodologique (3ECTS)	S3 : Projet transdisciplinaires II : Mise en pratique du projet transdisciplinaire (3 ECTS)				

Bloc de compétences SDA : "Sciences des données et applications" (6 ECTS)							
CSA : Mise en oeuvre concrète des SDIA dans un domaine applicatif	Objectif : Confronter l'étudiant à des cas réels d'utilisation et de recherche dans son domaine de prédilection			S3 ou S4 : TER SDIA (6 ECTS)	S4 : Stage (0 ECTS)		Contenu : - Introduction à Recherche (Travail personnel) : Mémoire et présentation sur un aspect SDIA dans le domaine d'application -> 6 ECTS - Stage en entreprise ou laboratoire : Coloration validée par la "Gouvernance" du parcours >SDIA - ... Forme : - Initiation à la recherche, aide à la bibliographie ... - Séminaires dédiés au domaine - Stage "coloré" SDIA

Possibilité de création un CTC 2bis "Approches méthodologiques et outils avancés en Science des Données et IA(S3 ?)" en remplacement de CTC 2 : Accès et calcul sur S1 : 9 ECTS plutôt CM/TD/TP

S2 : 9 ECTS plutôt CM/TD/TP

S3 : 3 ECTS Projet et travail personnel

S4 : 6 ECTS Travail personnel et stage

Répartition des crédits sur le parcours :

Question 2 : Prévoir une formation plus ou moins longue sur Python (ou R) pour pouvoir mettre en place les TP ?

Principe de subsidiarité: Certains enseignements, déjà assurés par un composante, pourront être pris en compte pour valider le parcours (soumis à validation par le jury du parcours D&IA)