



Écoles thématiques 2021 - Description du projet (en français)

Pièce 1

Titre court (maximum 15 caractères) : e-ENVIR 2021

Titre complet de l'école (en majuscules) : Les données ouvertes en sciences environnementales : exploration et cas d'études et applications

Responsable scientifique porteur du projet (un seul) :

Nom Prénom	Fonction	Code et intitulé du laboratoire	Adresse	Téléphone	Adresse électronique
CHARPENTIER Isabelle	DR2	Icube UMR 7357 & ZAEU	2 rue, Boussaingault 67000, Strasbourg	0652843857	icharpentier@unistra. fr

Délégation régionale (DR) du CNRS organisatrice : DR10-Alsace

Conseiller(e) formation de la DR en charge de l'école : Rebecca Omnes

Lieu : Strasbourg, Centre Saint Thomas (2 rue de la Carpe Haute 67000 Strasbourg ; <https://www.centre-st-thomas.fr/>)

Date de début : 02/11/2021

Date de fin : 05/11/2021

Durée : 4 (jours)

S'agit-il d'une école récurrente ?

oui

non

Si oui, **Code de l'école de la session précédente :**

Titre de la session précédente : e-ENVIR, Les données ouvertes en sciences environnementales : concepts et méthodes de qualification et de partage des données à l'ère du big data

NOMBRE DE PARTICIPANTS (stagiaires uniquement – hors intervenants et organisateurs)

Nombre total de participants prévus	40
Nombre d' agents CNRS prévus	20

Attention : "Les agents CNRS ne sont pas seulement des agents appartenant à une unité CNRS, ils doivent être RÉMUNÉRÉS par le CNRS "

BUDGET PREVISIONNEL

Budget global prévu	18 300 €
Budget demandé au CNRS	9 985 €
Nom du gestionnaire financier pressenti dans le laboratoire du responsable scientifique porteur du projet	Magali Darrieumerlou

INSTITUT(S) DU CNRS CONCERNE(S) :

- Principal (1 seul) : INEE
- Secondaires (le cas échéant) : INSU

PROPOSITION DES SECTIONS DU COMITE NATIONAL LES PLUS PERTINENTES POUR EXAMEN DU PROJET PAR ORDRE DE PRIORITE (maximum 6 sections) :

29 (Biodiversité, évolution et adaptations biologiques : des macromolécules aux communautés)	30 (Surface continentale et interfaces)	19 (Système Terre : enveloppes superficielles)
--	---	--

1/ PILOTAGE DE L'ACTION

Comité scientifique (cf. annexe pour des détails sur THEIA/OZCAR et BED/RZA) :

Nom Prénom	Fonction	Code et intitulé du laboratoire
CHARPENTIER Isabelle	Co-animatrice du réseau BED/RZA	ICUBE UMR 7357
PIGNOL Cécile	Co-animatrice du réseau BED/RZA	EDYTEM UMR 5204
HEINTZ Wilfried	Membre BED/RZA	DYNAFOR INRAE
LERIGOLEUR Wilfried	Membre BED/RZA	GEODE UMR 5602
DAMY Sylvie	Membre BED/RZA	CHRONO-ENV UMR 6249
SARRAMIA David	Membre BED/RZA	LPC UMR UMR 6533
LE MOAL Françoise	Membre BED/RZA	ECOBIO UMR 6553
GLATRON Sandrine	Membre BED/RZA	DYNAME UMR 7367
ROUAN Mathias	Membre BED/RZA	LETG UMR 6544
PLUMEJAUD Christine	Membre BED/RZA	LIENS UMR 7266
BRAUD Isabelle	Co-animatrice du réseau OZCAR et co-resp Theia/OZCAR	RiverLy INRAE
GALLE Sylvie	Co-resp. Theia/OZCAR	IGE UMR 5001
CHAFFARD Véronique	Responsable technique Theia/OZCAR	IGE UMR 5001
LE BRAS Yvan	Resp. PNDB	PNDB UMS

Comité d'organisation :

Nom Prénom	Fonction	Code et intitulé du laboratoire
GIRARD Virginie	Animatrice/Coordinatrice réseau eLTER FR	LECA UMR 5553
BRAUD Isabelle	Co-animatrice du réseau OZCAR	RiverLy INRAE
CHARPENTIER Isabelle	Co-animatrice du réseau BED/RZA	ICUBE UMR 7357

Calendrier des réunions préparatoires :

Cinq réunions préparatoires seront organisées avec le comité d'organisation et le comité scientifique par visioconférence sur un rythme régulier (oct. 2020 – jan. 2021 – avr. 2021 – aou. 2021 – oct. 2021) afin de finaliser le programme de l'école thématique et son déroulé avant diffusion, et de préciser les supports pédagogiques.

- **oct. 2020** : Validation de la synthèse e-ENVIR 2019 (version 1) + Confirmation de la disponibilité des intervenants + Finalisation programme + Finalisation des documents de communication + Lancement de l'invitation (save the date)
- **jan. 2021** : Réunion avec les intervenants (pool 1) + identification des études de cas et du détail des sessions + Lancement des inscriptions.

- **avr. 2021** : Réunion avec les intervenants (pool 2) + identification des études de cas et du détail des sessions + Clôture des inscriptions et choix des participants.
- **aoû. 2021** : Mise en place et premier test des TD pour passer en revue les compléments à ajouter sur les supports des TD et/ou des cours magistraux.
- **oct. 2021** : Finalisation des supports pédagogiques + Scénarios prises vidéos pour matériels pédagogiques et de communication post-ET

NB. L'équipe organisatrice et le comité scientifique respecte l'équilibre des genres.

2/ SITUATION SCIENTIFIQUE ET OBJECTIFS

2.1 Les enjeux (Justification de l'intérêt du thème et des objectifs de l'école sur le plan de la recherche)

Le projet d'école thématique E-ENVIR donne suite à celle ayant eu lieu en octobre 2019. Force a été de constater que les différentes communautés partageaient des difficultés très proches dans l'élaboration et la mise en place de solutions de gestion de la donnée (ex. thésaurus, ontologie) pour lesquelles une mutualisation des moyens à travers l'appui des pôles de données et les relais institutionnels est cruciale.

D'autre part, il a été souligné une absence notable des scientifiques producteurs de données et notamment ceux issus des sciences humaines et sociales. Les IR OZCAR et RZA, miroir français de l'Infrastructure de Recherche européenne eLTER-RI ont un besoin impératif d'homogénéiser et de co-construire leurs (nouvelles) pratiques de gestion des données à des fins de collaborations à l'échelle nationale et européenne. En particulier, eLTER vise à mettre en place des services d'accès aux sites d'observation (Transnational Access) et aux bases de données (Virtual Access) pour lesquels il sera indispensable d'offrir des outils partagés, reposant notamment sur une nomenclature commune aux niveaux national et européen. Le projet d'école thématique souhaite en premier lieu rassembler les deux communautés des IR OZCAR et IR RZA, et s'ouvrir aux autres infrastructures du domaine de l'environnement du CNRS (ACTRIS, AnaEE France, IAGOS, ICOS, Data Terra, PNDB).

L'école thématique entend enfin favoriser la recherche interdisciplinaire par la présentation et la structuration des données des chercheurs. Notamment, l'IR OZCAR et l'IR RZA balaient un large ensemble de disciplines (hydro-météo-géologie-géophysique, géochimie, socio-écologie, écologie, archéologie, sociologie...) et présentent la particularité de disposer de données sur le long terme, où la gestion et le partage des données sont des éléments clefs.

2.2 Objectifs scientifiques de l'école.

Les objectifs scientifiques de l'école thématique consistent à **l'acquisition des principes du FAIR (Findable, Accessible, Interoperable and Reusable) dans la gestion** des grands jeux de données des sciences de l'environnement, notamment :

- Identifier et partager les enjeux que représente le partage des données pour consolider les résultats scientifiques et traiter les problématiques complexes des sciences de l'environnement ;
- Développer la culture scientifique et technique sur la production et l'utilisation de données FAIR et sur les opportunités offertes par le partage des données ;
- Informer sur les stratégies et outils de valorisation de données FAIR afin de faire émerger des questions de recherche et projets basés sur ces ressources ;
- Connaître les référentiels notamment sémantiques de caractérisation des 'données' afin de permettre leur bonne utilisation dans les projets de recherche mais aussi favoriser la contribution des scientifiques français à ces référentiels ;
- Faire mieux connaître le cadre organisationnel de la communauté européenne du réseau eLTER.

2.3 Objectifs de formation

Le choix du format d'une école thématique est justifié par la portée interdisciplinaire de la formation rassemblant une grande communauté scientifique et technique du domaine en complément des actions de formation des réseaux déjà constituées. Les acteurs concernés partagent tous le même intérêt méthodologique mais maîtrisent de manière très hétérogène les connaissances liées à la gestion des données. Le projet pédagogique cohérent proposé par cette école thématique permettra d'asseoir et d'enrichir le socle de connaissances. Le format d'école permettra aussi aux participants d'avoir du temps pour se connaître et échanger sur leurs pratiques respectives, afin de permettre l'émergence d'un réseau de compétences pluridisciplinaire. Elle complètera les actions de communication et de formation des réseaux nationaux comme RBDD ou SIST en visant notamment le public eLTER et en abordant les outils d'automatisation dans la gestion de la donnée (geoflow...). Ainsi, elle contribuera aussi à créer une communauté autour de cas concrets, nécessaire pour eLTER France (ex. porter à connaissance de données existantes pour mieux les partager ; réflexion sur comment documenter des données de nature différente associées à un objet commun).

2.4 Public concerné

- Prioritairement : des chercheurs et des ingénieurs impliqués dans la gestion de données au sein des différentes infrastructures de recherche du domaine de l'environnement, notamment les personnels impliqués dans les projets nationaux ou Européens associés aux observatoires RZA et OZCAR.
- Secondairement : des chercheurs et des ingénieurs responsables de programmes de recherche produisant des données à forte valeur ajoutée ou en charge de systèmes d'information dans des laboratoires de recherche publique.
- A partir de quels constats, analyses, diagnostics ce public est-il ciblé?
Il est important d'associer chercheurs et ingénieurs à la construction des outils pour qu'ils répondent d'une part aux besoins des scientifiques, et, d'autre part, qu'ils respectent les standards et utilisent les technologies les plus en pointe. Le public visé par cette école thématique rassemblera donc les deux publics.
- Prérequis
Les participants doivent être sensibilisés à la question de la gestion des données. Il sera notamment demandé aux participants de s'appropriier les documents de synthèse fournis au-préalable et les supports de l'ET e-ENVIR 2019. Une connaissance du logiciel libre R sera souhaitable pour faciliter la participation aux TD.

3/ CONSÉQUENCES ATTENDUES

Le programme de l'École thématique est conçu de manière à favoriser la prise en main des solutions et favoriser les échanges, impliquant des « travaux appliqués » sur la base des jeux de données représentatifs du réseau eLTER France, identifiés au préalable pour partie, et pour certains à l'interface avec les infrastructures du domaine de l'environnement (ACTRIS, AnaEE France, IAGOS, ICOS, Data Terra, PNDB). Ce projet d'école thématique entend donc aller plus loin que la précédente en identifiant des cas d'études existants et potentiels, afin de caractériser les difficultés rencontrées, et en réponse suggérer les ressources qu'il faudrait mobiliser. Il vise également à rassembler des scientifiques et des gestionnaires de données, expérimentés et néophytes, pour faciliter la compréhension des contraintes de

chacun (contraintes organisationnelles, matérielles...). Notamment, cette école souhaite contribuer au développement d'une communauté pluridisciplinaire (eLTER FR), et favoriser le transfert de méthodes et techniques (Geoflow, Metashark pour l'automatisation des pratiques).

Ainsi, une conséquence immédiate sera une meilleure valorisation des données produites dans nos communautés et de manière plus concrètes, les participants devront savoir-faire à la fin de la formation :

- identifier les vocabulaires contrôlés/ontologies à utiliser
- savoir renseigner des métadonnées sur un jeu de données
- savoir identifier les entrepôts/portails de données où ils peuvent déposer leurs données
- savoir produire un plan de gestion de données
- savoir établir un contrat d'enquête (ou notice d'information et de consentement) associé au RGPD.

4/ GRANDS AXES DU PROGRAMME

La première demi-journée sera consacrée à une vue d'ensemble des étapes de gestion de la donnée, et à une rapide présentation du projet eLTER et du réseau LTER France pour en comprendre les tenants et aboutissants au prisme des données. Ensuite, chaque demi-journée du programme sera consacrée à une étape de la gestion des données / un thème et composée de courtes interventions introductives (dédiés en particulier aux personnes n'ayant pas participé à e-ENVIR 2019) suivi de un ou deux cas d'études, d'un TD, et d'une restitution collective pour identifier les spécificités et les lacunes... Six ateliers seront proposés (ordre non chronologique) :

- Un atelier de **production et actualisation de fiches de métadonnées** à partir d'un fichier Excel décrivant des jeux de données de manière standard et d'une routine R (Geoflow). Enjeux : automatisation, réactualisation de fiches de méta-données, échelle de production (découverte, granularité, objets d'intérêt).
- Un atelier sur le **vocabulaire contrôlé**. Identification des référentiels clefs des disciplines, exploration du thesaurus THEIA/OZCAR complété de la partie biodiversité sur la base de l'existant, définitions...
- Un atelier sur la **DOI-isation et les entrepôts de données** avec un jeu de données fourni pour ceux/celles qui n'en n'auraient pas. Les pôles de données nationaux et locaux, et des soutiens à plus ou moins long terme seront évoqués notamment pour répondre aux questions : Comment gérer les données des projets pluridisciplinaires ? faut-il les disperser entre plusieurs pôles ?
- Rédaction et partage d'un **plan de gestion de données** en particulier pour les observatoires sur le long terme (plan gestion de la donnée) et les projets pluriannuels ; les étapes pour y parvenir, et les ressources nécessaires.
- Un atelier sur les **données SHS, d'enquêtes** : enjeu de stockage, celles issues des portails nationaux tel que l'INSEE, etc.
- Un échange et mise en application de l'**interopérabilité des jeux de données à l'échelle européenne** avec DEIMS

5/ MODALITÉS PÉDAGOGIQUES

5-1 - Modalités pédagogiques

Pour les interventions « introductives » et « cas d'études », un format homogène sera fourni, notamment avec des synthèses concises des points clefs à retenir. Une synthèse de e-ENVIR 2019 sera envoyée également au préalable aux participants. Les « cas d'études » serviront de bases pour le TD, pour lequel des données et des accès aux outils en ligne seront fournis. A l'issue de chaque TD, une restitution collective permettra d'avoir les éléments de discours pour la production de tutoriels vidéo. La chaîne de

traitement Geoflow (librairie R en libre accès) sera privilégiée dans plusieurs TD car elle permet une automatisation de la production de métadonnées, de flux pour les géoserveur, de création de plan de gestion de la donnée et de DOI, etc. Un espace numérique préconfiguré sera proposé via OpenStack, afin d'accéder rapidement aux logiciels des TD. Enfin, chaque session sera pilotée par un binôme chercheur-gestionnaires de données pour veiller à la bonne formulation des éléments fournis en terme de pédagogie.

5-2 -Rythme des séances de travail

(cf. annexe pour le détail des intervenants)

Module	Intervenant pressentis	Forme
J-30 Synthèse ET E-ENVIR 2019 envoyée aux participants	<i>La communauté BED/RZA et OZCAR/THEIA en appui avec l'équipe organisatrice et les intervenants de l'ET E-ENVIR 2019</i>	<i>Rapport avec en Annexes les présentations de l'ET E-ENVIR 2019</i>
Jour 1		
Après-Midi		
**Introduction : e-ENVIR 2019, enjeux ; programme, présentation du projet eLTER et du réseau eLTER France	Virginie Girard (eLTER FR), Isabelle Charpentier (RZA/BED), Isabelle Braud (OZCAR/THEIA)	Plénière
**Présentation et mise en place de la chaîne de traitement automatisée avec l'outil Geoflow et MétaShARK	Emmanuel Blondel (expert indépendant) & Yvan Le Bras (PDNB)	Cours magistral + TD
Jour 3		
Matin		
**DOI-isation, data-paper et entrepôts (Geoflow)	Yvan Le Bras (PDNB)	Cours magistral + TD
Après-Midi		
**Vocabulaires contrôlés	Chloé MARTIN (BBEES)	Cours magistral + TD
Fin de journée		
**Echanges-débats informels / Ice Breaker		
Jour 4		
Matin		
**Plan Gestion de données (Geoflow)	Chloé MARTIN (BBEES)	Cours magistral + TD
**Production de MD : quelles données ? quelles finalités ?	Chloé MARTIN (BBEES)	Cours magistral + TD
Après-Midi		
**Données SHS : RGPD, et outils de référencement et de stockage (SYGADE, etc.)	Stéphane Pouyllau (HumaNum)	Cours magistral + TD
Jour 4		
Matin		
**Gestion des données à l'échelle du réseau eLTER (DEIMS-SDR)	Christopher Wohner (DEIMS-SDR ; LTER) & Virginie Girard (eLTER FR)	Cours magistral
**Conclusion générale	Virginie Girard (eLTER FR),	Table ronde avec les

	Isabelle Charpentier (RZA/BED), Isabelle Braud (OZCAR/THEIA)	intervenants
--	---	--------------

5-3 -Supports pédagogiques

- Utilisés pour l'enseignement : présentations orales, logiciels et tutoriels associés, démonstrations sur jeux de données à l'aide de tutoriels
- Distribués aux participants : documents de cours (supports de présentations), documents de synthèse (rapports nationaux ou internationaux), ouvrages collectifs et articles marquants du domaine

5-4- Publication de documents (forme, titre et année)

Des **tutoriels vidéos** (tuto-vidéos) seront mis à disposition dans les réseaux pour faciliter la dissémination des connaissances sur l'exemple des jeux de données utilisés en TD. Notamment, ces tuto-vidéos tiendront compte des questions récurrentes abordées. Chaque tuto-vidéo sera réalisé dans un format court (5 à 10 min) ; le titre sera défini par session et type de questions. Date de restitution et mise en ligne : Juillet 2022.

6/ PROCÉDURE D'ÉVALUATION

Le comité d'organisation assurera le suivi et l'évaluation de l'Ecole thématique et produira une **enquête de satisfaction** qui sera diffusée auprès des participants à la fin de l'école thématique. Les résultats de l'enquête seront attendus pour début 2022.

Les **tutos-vidéos** seront également évalués selon les **outils statistiques** mise à disposition par le média utilisé (ex. youtube : nombre de like ; site internet : nombre de clics...)

7/ MODALITÉS PRATIQUES

- Justification du choix du lieu : Le Centre Culturel Saint Thomas permet de centraliser les déplacements des participants sur Strasbourg tout en bénéficiant de conditions de travail en groupe isolés à des prix très raisonnables.
- Communication prévue (*annonces, plaquettes, affiches, etc.*) :
 - Mail d'annonce « Save the date » + « inscription »
 - Sites web et liste de diffusion des IR OZCAR et RZA, moyens de communication de l'INSU ou de l'INEE
 - Plaquette de communication

ANNEXES

****THEIA/OZCAR**

THEIA/OZCAR est un système d'information qui collecte les données in-situ de tous les observatoires OZCAR. Il a été construit dans le cadre du pôle Data Terra, lequel facilite l'accès aux données du système terre. Theia est l'un des quatre pôles de Data Terra dédié aux données de surface terrestre. L'objectif de THEIA/OZCAR est de rendre plus visible les données de la communauté OZCAR.

Publication à consulter:

Braud I., V. Chaffard, C. Coussot, S. Galle, P. Juen, and the Theia/OZCAR team (25 co-authors) (2020) : [Building the information system of the French Critical Zone Observatories network: Theia/OZCAR-IS](#). Hydrological Sciences Journal, Special Issue "Hydrological data: opportunities and barriers". DOI: 10.1080/02626667.2020.1764568

****BED/RZA**

La gestion des données au sein du RZA s'inscrit dans une feuille de route nationale et européenne (eLTER) et relève du projet et collectif « **Bancariser les Données Ensemble** » (**BED**). BED est un collectif de personnels multi-institution (CNRS, INRAe, université) existant depuis 2017, dont les compétences et l'expérience en gestion de données sont reconnues au sein du RZA.

Site internet : <http://www.za-inee.org/fr/data-md-intro>

****CV des intervenants**

Emmanuel BLONDEL, expert indépendant, spécialisé dans la gestion des données, il est à l'origine du développement du package Geoflow sur R. Il travaille notamment depuis 1 an aux côtés de BED/RZA (Page RDA : <https://www.rd-alliance.org/users/emmanuel-blondel>)

Yvan LE BRAS, Chef de projet de l'"Infrastructure de Recherche Pôle national Données Biodiversité". Il contribue au développement de solutions pour la gestion automatisée des données dans le domaine de la biodiversité et plus largement, à travers des interactions avec les autres Infrastructures de Recherches sur les données. (Page du site internet : <http://www.patrinat.fr/fr/pole-national-de-donnees-de-biodiversite-pndb-6256>)

Chloé MARTIN, Ingénieure de Recherche à UMS BBEES, en charge entre autres de la Coordination du réseau Bases de données (rBDD) du CNRS depuis 2012. Elle a notamment co-organisé l'ET e-ENVIR 2019. (Page professionnelle : <https://bbees.mnhn.fr/node/23>)

Stéphane POUYLLAU, Ingénieur de Recherche au CNRS, il est directeur adjoint du pôle de données en SHS HumaNum. (Page professionnelle : <https://www.huma-num.fr/personne/stephane-pouyllau>)

Christopher WOHNER, Chargé de projet au sein de l'Environment Agency Austria. Il s'occupe notamment de la gestion du catalogue des sites et des bases de données 'DEIMS-SDR' pour les sites LTER (Page du site internet : <https://deims.org/about>).



Pièce 2

Ecoles thématiques 2021 Fiche budgétaire prévisionnelle

Titre court (maximum 15 caractères) :		E-ENVIR	
Titre complet de l'école :		LES DONNÉES OUVERTES EN SCIENCES ENVIRONNEMENTALES : EXPLORATION DE CAS D'ÉTUDES ET APPLICATIONS	
Nom et prénom du responsable scientifique porteur du projet		Isabelle Charpentier	
DETAIL DES COUTS			MONTANT
Durée de l'école en jours	4		
Montant journalier (2 repas + nuitée)		65	
Frais de déplacement (organisation)			
	nbre		
- Intervenants	5		1 000,00
- Organismes	3		600,00
Frais d'hébergement (organisation)			
	nbre		
- Intervenants	5		1 300,00
- Organismes	3		780,00
Frais d'hébergement des participants (hors intervenants et organisateurs)			
	nbre		
- Participants salariés du CNRS	20		5 200,00
- Participants non salariés CNRS	20		5 200,00
Frais de location de salles			2 020,00
Frais de fonctionnement (détail)			
1- : matériels pédagogiques, fournitures diverses			200,00
2- : prise vidéo et montage pour produire des docs pédagogiques in fine			2 000,00
3- :			
4- :			
TOTAL DES COUTS			18 300,00
DETAIL DES RECETTES			MONTANT
Droits d'inscription			
	nbre	Tarif	
Participants d'autres établissements publics	17	260	4 420,00
Participants d'établissements privés	2	410	820,00
Doctorants	1	75	75,00
Autres			0,00
Financements			
Organismes publics (préciser) : IR OZCAR et IR RZA			3 000,00
Organismes privés (préciser) :			
Soutien sollicité auprès du CNRS sur le budget formation permanente			9 985,00
TOTAL DES RECETTES			18 300,00

Ne mentionner que les frais de mission des intervenants et organisateurs; ceux des participants CNRS sont pris en charge par la délégation d'origine ou le laboratoire.

Attention ! Que l'école soit résidentielle ou non, les frais d'hébergement des salariés du CNRS, quelque soit leur statut (titulaire, contractuel, BDI ...), seront à prendre en charge sur le budget de l'école.

Attention ! Tous les frais devront être clairement détaillés. Sinon le montant ne sera pas pris en compte et la subvention CNRS ajustée en conséquence

Dans la mesure où le CNRS accorde une subvention, les agents CNRS sont dispensés des droits d'inscription.

- G11: **Ne mentionner que les frais de mission des intervenants et organisateurs; ceux des participants CNRS sont pris en charge par la délégation d'origine ou le laboratoire.**
- G19: **Attention ! Que l'école soit résidentielle ou non, les frais d'hébergement des salariés du CNRS, quelque soit leur statut (titulaire, contractuel, BDI ...), seront à prendre en charge sur le budget de l'école.**
- G24: **Attention ! Tous les frais devront être clairement détaillés. Sinon le montant ne sera pas pris en compte et la subvention CNRS ajustée en conséquence**
- G33: **Dans la mesure où le CNRS accorde une subvention, les agents CNRS sont dispensés des droits d'inscription.**



Écoles thématiques 2019 – BILAN

INFORMATIONS GENERALES :

Code ET :	Titre court : e-ENVIR		
Titre Complet : Les données ouvertes en sciences environnementales: concepts et méthodes de qualification et de partage des données à l'ère du big data			
Responsable scientifique porteur du projet :	Nom Prénom	Fonction	Adresse électronique
	Le Galliard Jean-François Martin Chloé	CR CNRS IR CNRS	cereep@biologie.ens.fr chloe.martin@mnhn.fr

Conseiller(e) formation en charge de l'école :	Nom, Prénom	DR
	Najat ESSAIFI	IFSEM

Institut(s) Scientifique(s) du CNRS concerné(s) :	Principal	Secondaires			
	INEE	INSU			

Lieu	Date de début	Date de fin	Durée (jours)
Campus CNRS de Gif sur Yvette	28/10/2019	31/10/2019	4

PARTICIPANTS : (Ne pas inclure les intervenants et les organisateurs)

	Nombre total de participants	Nombre de participants salariés du CNRS
Prévisionnel (cf fiche projet)	40	25
Réalisé	34 (2 annulations)	20 (1 annulation)

INTERVENANTS-ORGANISATEURS (ayant assisté à la totalité de l'ET : 8 (1 annulation + 4 intervenants ayant assisté seulement en partie))

BUDGET :

Budget global initial (Cf. fiche projet)	Budget global final réalisé	Dont Financement CNRS
20 460 €	18 444	12 787 €

1- Rappel des objectifs de l'école (10 lignes max)

Les objectifs scientifiques de l'école thématique consistent à l'acquisition des connaissances et des compétences du FAIR dans la gestion des grands jeux de données des sciences de l'environnement, notamment identifier et partager les enjeux du partage des données ; développer la culture scientifique et technique sur la production de données FAIR et sur les opportunités offertes par le partage des données ; informer sur les stratégies et outils de valorisation de données FAIR afin de faire émerger des questions de recherche et projets basés sur ces ressources ; connaître les référentiels notamment sémantiques de caractérisation des 'données' afin de permettre leur bonne utilisation dans les projets de recherche mais aussi favoriser la contribution des scientifiques français à ces référentiels, et faire mieux connaître le cadre législatif national et international.

2- Bilan pédagogique :

Le programme de l'école thématique a permis d'organiser (i) 8 cours magistraux suivis de temps de question et d'échanges dont deux tables rondes dédiées à des discussions collectives avec l'ensemble de participants, (ii) plusieurs retours d'expériences organisés sous forme de présentations courtes d'une demi-heure en général suivies de temps de questions ainsi que (iii) deux ateliers dont un organisé sous la forme d'un jeu coopératif. Cette première école thématique avait des objectifs très généraux dans le but de favoriser la mise à connaissance de savoirs génériques transversaux au domaine notamment le cadre juridique, les normes et les standards et les technologies. Les retours d'expérience ont été spécifiquement choisis pour couvrir différentes situations pertinentes à la communauté scientifique et technique de l'école thématique depuis la production jusqu'à la valorisation des données. Le nombre d'ateliers était volontairement limité afin de favoriser les temps d'échange pendant les interventions et les pauses. Cependant, deux ateliers pratiques supplémentaires ont été proposés en option en fin de deux journées pour compléter le programme. Environ la moitié des participants ont suivi des ateliers optionnels. Les intervenants ont disposé de temps libre importants au moment des pauses et des repas pour discuter librement avec les participants.

3- Synthèse de l'évaluation à chaud :

Une enquête a été effectuée qui a permis de quantifier la satisfaction des participants. Les conclusions quantitatives de cette enquête sont résumées dans le texte ci-dessous qui propose aussi une évaluation plus qualitative.

a) Aspect scientifiques

Tous les sujets proposés dans le programme prévisionnel de l'école thématique ont été abordés avec une prépondérance accordée à la problématique scientifique de l'annotation, de la sémantique et des vocabulaires contrôlés. Les cours ont généralement alterné des présentations très générales et des présentations très spécialisées autour de trois thématiques (plan de gestion FAIR des données, annotation et sémantique, mise à disposition et entrepôts de données). Les retours d'expérience ont permis de présenter des cas d'étude pertinents et les temps d'échange ont favorisé pour la première fois la constitution d'une large communauté scientifique et technique associant toutes les infrastructures de recherche du CNRS et de ses partenaires dans le domaine de l'environnement. Ces retours d'expérience ont été très largement appréciés, notamment des intervenants qui ont suivi l'intégralité de la formation (6 personnes au total).



Écoles thématiques 2019 – BILAN

L'évaluation quantitative du programme scientifique de l'école thématique a été effectuée sur une échelle de 1 à 5 et donne les résultats suivants :

- Cours sur les plans de gestion des données :
- Cours sur la sémantique et l'annotation :
- Cours sur les entrepôts de données :
- Table ronde :
- Table ronde :
- Ateliers :

Points positifs :

Les points positifs identifiés par les participants sont :

Points à améliorer :

Les points à améliorer proposés par les participants sont :

b) Aspects organisationnels

L'École thématique a été organisée sur le campus CNRS de Gif-sur-Yvette offrant un complexe de salles de réunions dans le « château », une offre d'hébergement en chambres simples ou doubles et une offre de restauration. L'organisation de l'école thématique a été facilitée par ce choix qui permettait de plus une réduction significative des coûts d'hébergement et de restauration. Ce lieu a aussi permis de travailler dans des conditions adaptées et conviviales

L'évaluation quantitative des conditions pratiques de l'école thématique a été effectuée sur une échelle de 1 à 5 et donne les résultats suivants :

Points positifs :

Les points positifs identifiés par les participants et les organisateurs sont :

- Accueil particulièrement professionnel et disponibilité de l'équipe de la DR4 ;
- Cadre de travail agréable et au calme avec tout sur place ;
- Un prix très concurrentiel ;

Points à améliorer :

Les points à améliorer proposés par les participants et les organisateurs sont :

- Une salle un peu limitée pour la taille de notre groupe (impossible de mettre des tables) et difficilement adaptable à des travaux pratiques en informatique ;
- Une offre de restauration, imposée par le prestataire, de type « restaurant professionnel » alors qu'on aurait souhaité une solution bio plus en phase avec nos exigences et celles de nos participants ;
- Une aide très limitée des services de la délégation organisatrice dans la logistique et la communication ;

4- Analyse et commentaires de la délégation régionale : *(A remplir par le Conseiller formation de la DR)*



Écoles thématiques 2019 – BILAN

5- Suite envisagée

La discussion avec les participants a permis d'identifier le besoin d'une suite à l'Ecole thématique orientée sur les démarches de « qualité » de productions de données comprenant des cours théoriques et de nombreux cours pratiques d'annotation, de nettoyage et de validation des bases de données en environnement dans la perspective de leur réutilisation. Ce projet pourrait être porté par le même collectif des infrastructures de recherche de l'INEE et de l'INSU.



Écoles thématiques 2019 – BILAN

Liste des participants :

Nom	Prénom	Employeur	Email
Abbassi	Sarra	CNRS	Sarra.Abbassi@univ-savoie.fr
Arnaud	Fanny	CNRS	fanny.arnaud@ens-lyon.fr
Boër	Michel	CNRS	michel.boer@anaee.eu
Braud	Isabelle	Irstea	isabelle.braud@irstea.fr
Damy	Sylvie	Université de Bourgogne Franche-Comté	sylvie.damy@univ-fcomte.fr
Girard	Virginie	CNRS	virginie.girard@univ-grenoble-alpes.fr
Hajj-Hassan	Hicham	CNRS	hishamhh@cnrs.edu.lb
Heintz	Wilfried	INRA	wilfried.heintz@inra.fr
Henon	Amandine	Muséum national d'Histoire naturelle	amandine.henon@mnhn.fr
Hoebeke	Mark	CNRS	mark.hoebeke@sb-roscoff.fr
Jacqueton	Céline	INRA	celine.jacqueton@inra.fr
Jean-Charles	Arnaud	CNRS	jeancharles@eccorev.fr
Joly	Dominique	CNRS	dominique.joly@cnrs-dir.fr
Kouadio	Sekedoua	IFSTTAR	jules.kouadio@ifsttar.fr
Lafont	Sébastien	INRA	sebastien.lafont@inra.fr
Le Moal	Francoise	CNRS	francoise.le-moal@univ-rennes1.fr
Lerigoleur	Emilie	CNRS	emilie.lerigoleur@univ-tlse2.fr
Liger	Lucie	CNRS	lucie.liger@irstea.fr
Martel	5 Julien	Institut de Radioprotection et de Sûreté Nucléaire (IRSN)	julien.martel@irsn.fr
Massol	Florent	CNRS	florent.massol@ens.fr
Moulard	Nathalie	CNRS	nathalie.moulard@cnrs-orleans.fr
Ogereau	Francis	Université Clermont-Ferrand Auvergne	francis.ogereau@uca.fr
Ohanessian	Gilles	CNRS	gilles.ohanessian@polytechnique.edu
Pons	Marie-Noëlle	CNRS	marie-noelle.pons@univ-lorraine.fr
Ramet	Anaïs	CNRS	anais.ramet@cefe.cnrs.fr
Ravel	Olivier	CNRS	olivier.ravel@ecotron.cnrs.fr
Rouan	Mathias	CNRS	mathias.rouan@univ-brest.fr
Sarramia	David	Université Clermont Auvergne	david.sarramia@clermont.in2p3.fr
Souton	Corentin	CNRS	souton@biologie.ens.fr
Tarniewicz	Jérôme	UVSQ	jerome.tarniewicz@lsce.ipsl.fr
Vannièrè	Boris	CNRS	boris.vanniere@univ-fcomte.fr
Boulangèr	Damien	CNRS	damien.boulangèr@obs-mip.fr
Chaffard	Véronique	IRD	veronique.chaffard@ird.fr
Vachèz	Dominique	CNRS	dominique.vachèz@inist.fr

Statut : Directeur de recherche, chargé de recherche, Professeur, maître de conférences, Ingénieur, assistant-ingénieur, technicien, post-doctorant, doctorant, autres



Écoles thématiques 2019 – BILAN

Liste des intervenants et organisateurs (ayant assisté à la totalité de l'école) :

Nom	Prénom	Statut	Employeur	Email
Barde	Julien	IR	IRD	julien.barde@ird.fr
Clastre	Philippe	IR	INRA	philippe.clastre@inra.fr
Garnier	Eric	DR	CNRS	eric.garnier@cefe.cnrs.fr
Jonquet	Clément	Mcf	Université de Montpellier	jonquet@lirmm.fr
Quidoz	Marie-Claude	IR	CNRS	Marie-Claude.Quidoz@cefe.cnrs.fr
Ziebelin	Danielle	PR	Université Grenoble Alpes	danielle.ziebelin@imag.fr
Le Galliard	Jean-François	CR	CNRS	cereep@biologie.ens.fr
Martin	Chloé	IR	CNRS	chloe.martin@mnhn.fr

Statut : Directeur de recherche, chargé de recherche, Professeur, maître de conférences, Ingénieur, assistant-ingénieur, technicien, post-doctorant, doctorant, autres



Écoles thématiques 2019 – BILAN

Bilan financier :

Ecoles thématiques 2019- Fiche budgétaire finale	
Titre court (maximum 15 caractères) :	e-ENVIR
Titre complet de l'école : Les données ouvertes en sciences environnementales: concepts et méthodes de qualification et de partage des données à l'ère du big data	
Nom du responsable scientifique porteur du projet	Le Galliard Jean-François
DETAIL DES DEPENSES	MONTANT
Frais de déplacement (intervenants et organisateurs)	2 900
Frais d'hébergement (intervenants et organisateurs)	3 360
Frais d'hébergement des participants (hors intervenants et organisateurs)	11 424
Frais de fonctionnement (détailler)	
Frais de communication	100
Frais de location (salle et matériels pédagogiques)	460
Frais de fonctionnement	200
7	
TOTAL DES COUTS	18 444,00
DETAIL DES RECETTES	MONTANT
Droits d'inscription pour les non CNRS	350
Participants d'autres établissements publics	4 900
Participants d'établissements privés	0
Doctorants	0
Autres	
Financements	
Organismes publics (préciser) : AnaEE France Infrastructure	757
Organismes privés (préciser) :	0
Subvention CNRS	12787
TOTAL DES RECETTES	18 444,00



Écoles thématiques 2019 – BILAN



Écoles thématiques 2019 – BILAN

Annexe : Bilan scientifique (format libre, 2 à 3 pages maximum)