

Compte-rendu du séminaire SIST19

Du 5 au 8 novembre 2019
Observatoire Midi-Pyrénées Toulouse



Table des matières

Contexte et objectifs.....	1
Comités.....	2
Comité d'organisation et de programme.....	2
Comité d'organisation locale.....	2
Gestion.....	2
Participant-es.....	2
Contenu du séminaire.....	3
Programme.....	3
Captation vidéo et présentations.....	5
Animation.....	5
Bilan et perspectives.....	6
Projet de groupe de travail « Guide de bonnes pratiques ».....	6
Projets de formations.....	6
Sondage 2019.....	6
Annexe.....	7
Programme détaillé.....	7

Contexte et objectifs

La 4ème édition du séminaire SIST (Séries Interopérables et Systèmes de Traitement) a eu lieu du 5 au 8 novembre 2019 à l'Observatoire Midi-Pyrénées (OMP) à Toulouse.

Ce séminaire a été organisé dans le cadre des actions du réseau technologique SIST des gestionnaires de données des observatoires (OSU et unités de recherche ayant des missions d'observation) soutenu par le CNRS INSU.

Il a fait suite aux trois premières éditions :

- 2015 à l'OSU Pytheas, Marseille, deux jours
- 2016 à l'OSU OREME, Montpellier, deux jours
- 2018 à l'Observatoire de Versailles Saint Quentin (OVSQ), deux jours.

Les objectifs des séminaires SIST sont entre autres de :

- réunir les informaticien·nes gestionnaires de données d'observation (des OSU et des unités de l'INSU, ainsi que d'autres instituts / organismes de recherche) autour de technologies communes ;
- partager la connaissance et les bonnes pratiques sur les aspects de la gestion des données d'observation, et ainsi favoriser le transfert de compétences ;
- assurer une veille sur les technologies émergentes ;
- promouvoir la réalisation de l'interopérabilité dans les formats et échanges des données d'observation.

Le séminaire SIST19 a été financé par le CNRS INSU et par l'Infrastructure de Recherche [Data Terra](#).

Comités

Comité d'organisation et de programme

- Juliette Fabre - OSU OREME
- Maurice Libes - OSU Pytheas
- Olivier Lobry- OSU OREME
- Didier Mallarino - OSU Pytheas
- Joël Sudre - OSU OMP/LEGOS

Comité d'organisation locale

- François André - OSU OMP / SEDOO
- Damien Boulanger - OSU OMP / SEDOO
- Elodie Bourrec - OSU OMP / IRAP
- Étienne Gondet - OSU OMP
- Emilie Lerigoleur - GEODE / OHM Pyrénées Haut Vicdessos et ZA Pygar
- Joël Sudre - OSU OMP / LEGOS

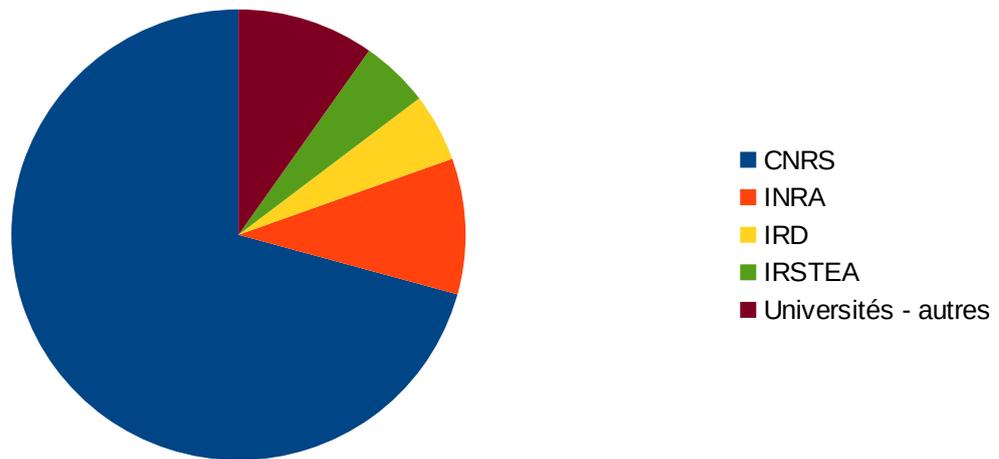
Gestion

- Alexia Vaillé – OSU OREME
- Hajar Zeroili – OSU OREME

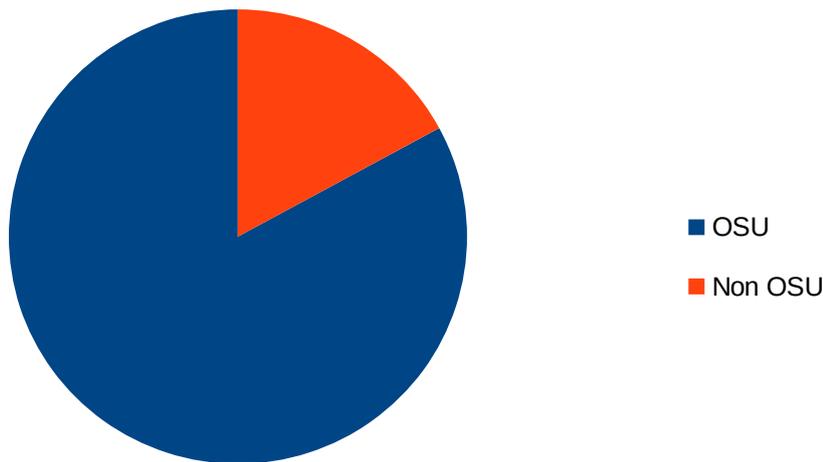
Participant·es

41 personnes se sont rendues à ce séminaire :

- 70 % des participant·es provenaient du CNRS, les autres appartenant à d'autres organismes et EPST comme l'INRA, l'IRD, l'IRSTEA ou différentes universités.



- 83 % participant·es provenaient de laboratoires liés à des OSU, représentant ainsi 11 des 27 [OSU](#) de l'INSU (Ecce Terra, IPGP, IUEM, OSUL, OMP, OREME, OSU Réunion, THETA, OSUNA, OSUR, Pytheas).



Le nombre de participant·es relativement peu élevé comparé aux éditions précédentes (en moyenne autour de 60 - 70 personnes) s'explique notamment par l'organisation simultanée de plusieurs événements sur cette même période (école technique e-Envir, événements PNDB, Theia, Resomar ...). Toutefois, cette taille de groupe a permis d'établir une vraie cohésion de groupe et de mettre en œuvre de nouvelles techniques d'animation (voir plus loin) et a été très appréciée par les participant·es.

Contenu du séminaire

Programme

Le programme complet du séminaire est disponible en annexe ainsi que sur <https://sist19.sciencesconf.org/program>.

Le séminaire SIST19 a proposé comme les années précédentes des retours d'expérience autour de la gestion des données et métadonnées d'observation.

Cette édition a également été l'occasion d'innover et de répondre aux attentes des participant·es formulées lors des dernières éditions :

- ajout d'une journée d'ateliers pratiques autour des différentes problématiques rencontrées par les gestionnaires de données d'observation ;
- allongement de la durée à trois jours (au lieu de deux) ;
- aménagement de temps longs de discussions et d'échanges.

Le séminaire a également consacré une journée à l'Infrastructure de Recherche Data Terra et à ses pôles de données.

Enfin, une dernière matinée hors-programme a permis aux participant-es qui le souhaitaient de bénéficier d'espaces de travail pour organiser ou participer à des groupes de travail libres.

Détails : retours d'expérience

Mots-clés des outils et thématiques abordées :

- outil Apache Airflow
- annotations sémantiques / ontologies
- outil Thredds
- outil istSOS
- outil R geoflow
- outil GeoNetwork
- workflow de données
- l'Atelier des données

Détails : IR Data Terra et pôles de données

La matinée du 6 novembre a été consacrée à des présentations techniques de l'IR Data Terra et des 5 pôles de données : Aeris, Odatis, Theia, Formater et PNDB. Les participant-es réparti-es en cinq groupes de travail ont ensuite préparé et restitué une synthèse comparative du fonctionnement des pôles sur différentes problématiques, donnant lieu à des échanges avec les représentant-es des pôles. Problématiques étudiées :

- chemin de la donnée des producteurs jusqu'aux pôles et au-delà ;
- méthodes d'alimentation (moissonnage, format pivot) ;
- services fournis actuels et prévus (portail de données, catalogues, DOI, standards de diffusion, ..) ;
- référentiels utilisés (vocabulaires contrôlés, thésaurus, ..) ;
- gestion de la qualification de la donnée (niveau de données, confiance, précision).

Détails : Ateliers

Quatre ateliers d'une demi-journée ont été proposés le 7 novembre.

Titre de l'atelier	Animateur·rices	Nb de participant·es
TP Réaliser un fichier NetCDF CF compliant <i>Présentation du format NetCDF, de la convention CF et des API</i>	<i>Maurice Libes, Joël Sudre, Didier Mallarino</i>	12
TP Générer automatiquement des fiches de métadonnées ISO19139 valides Inspire avec R <i>Exporter des jeux de métadonnées et des données dans un catalogue ISO19139 avec la librairie R geoflow</i>	<i>Emilie Lerigoleur, Wilfried Heintz, Juliette Fabre, Julien Barde</i>	14
Atelier Bonnes pratiques pour la gestion des données d'observation <i>Réflexions et élaboration d'un guide de recommandations publié sur le site web SIST</i>	<i>Comité de pilotage SIST</i>	19
"Allô docteur" GeoNetwork <i>Travail en groupe autour des points de blocage sur l'outil GeoNetwork avec l'appui d'un expert</i>	<i>Mathias Rouan</i>	10

Lors de l'atelier **Bonnes pratiques pour la gestion des données d'observation**, quatre groupes ont travaillé successivement sur quatre thématiques. Après avoir restitué son travail, chaque groupe a participé à la rédaction d'une ébauche de synthèse sous la forme d'un [document collaboratif](#) mis à disposition de la communauté et [disponible en téléchargement](#) sur le site web SIST.

Détails : Groupes de travail libres

Des salles ont été mises à disposition lors de la matinée hors-programme du 8 novembre pour les participant·es souhaitant participer ou organiser des groupes de travail libres. Deux groupes se sont créés :

- réunion du comité de pilotage SIST ;
- groupe de travail libre : présentation des outils Opentheso et CollecScience, et entraide autour de la création du SNO Observil.

Captation vidéo et présentations

Les supports de présentations sont disponibles sur le [site de la conférence](#).

L'équipe de l'« Atelier de l'OMP » a enregistré chacune des présentations. Les vidéos seront mises en ligne dès que possible.

Animation

Suite à l'ANF « Animer un réseau métier » (sept. 2019) suivie par l'une des membres du comité

de pilotage SIST, différentes techniques d'animation ont été mises en œuvre :

- création d'îlots de travail (groupes de table) ;
- temps de discussion par îlot après chaque retour d'expérience et avant le temps de question (sessions *Retours d'expérience*) ;
- groupes de travail autour de plusieurs thématiques (session *Pôles de données*) ;
- groupes de travail tournants (« world café ») autour de plusieurs thématiques (atelier *Bonnes pratiques pour la gestion des données d'observation*) ;
- inclusions (brise-glace) et déclusions (questions de fin de journée) chaque jour ;
- production d'un rapport d'étonnement en fin de séminaire.

L'ensemble de ces techniques ont demandé un temps de préparation important en amont (conception des différentes sessions, élaboration d'un déroulé extrêmement détaillé, préparation du matériel ..).

Ce temps a été largement mis à profit puisque ces méthodes ont considérablement facilité l'intégration des nouveaux membres, les interactions entre les participant·es et la mobilisation active de chacun·e, la cohésion du groupe, ainsi que la production de contenu (synthèses comparatives autour des pôles de données et guide de bonnes pratiques). Les retours des participant·es ont été extrêmement positifs.

Bilan et perspectives

Le séminaire a été très apprécié des participant·es, tant sur le contenu que sur la forme. Le comité de pilotage prévoit d'organiser un séminaire en 2020 dans un OSU à définir prochainement.

Projet de groupe de travail « Guide de bonnes pratiques »

Il est envisagé la création d'un groupe de travail pour poursuivre et étoffer la rédaction du guide de bonnes pratiques débutée lors de l'atelier **Bonnes pratiques pour la gestion des données d'observation**. A minima, ce guide pourrait être étoffé et révisé lors de chaque séminaire SIST (ateliers).

Projets de formations

Un projet d'ANF « Gestion des données d'observation : bases et outils de géomatique pour la gestion des données géoréférencées » a été déposé pour 2020.

Enfin, le comité de pilotage envisage une ou plusieurs formations sur les outils liés au standard de diffusion de données SOS (istSOS, 52°North), en 2020 ou 2021.

Sondage 2019

Un sondage sera envoyé d'ici la fin 2019 aux membres du réseau au sujet notamment de

l'organisation de SIST20 et des thèmes à aborder, du groupe de travail « Guide de bonnes pratiques » et des besoins en formation.

Annexe

Programme détaillé

Mardi 5 novembre 2019

Heures	événement
13:30 - 14:00	Accueil - Accueil des participant-es
14:00 - 14:30	Introduction - Joël Sudre - Comité de pilotage SIST
14:30 - 15:00	> Standardisation des séries chronologiques en altimétrie spatiale avec istSOS - <i>STEPHANE DEBARD, Hydrosciences Montpellier</i>
15:15 - 15:45	> Processus d'annotation sémantique pour favoriser l'interopérabilité autour des données de biodiversité et d'études d'écosystèmes au sein de l'infrastructure AnaEE-France. - <i>Rachid Yahiaoui, Unité INFOSOL</i>
16:00 - 16:30	Pause café
16:30 - 17:00	> Eccad, un exemple de mise en oeuvre de Thredds - <i>Sabine Darras, Observatoire Midi-Pyrénées</i>
17:15 - 17:45	> Service de gestion des flux de données basé sur Apache Airflow - <i>Franck Gabarrot, Observatoire des Sciences de l'Université de La Réunion</i>

Mercredi 6 novembre 2019

Heures	événement
09:00 - 10:30	L'infrastructure de Recherche Data Terra, les pôles de données AERIS et THEIA - <i>Richard Moreno, Damien Boulanger et Arnaud Sellé</i>
10:30 - 11:00	Pause café
11:00 - 12:30	Les pôles de données FORM@TER, ODATIS et PDNB - <i>Emilie Deschamps-Ostanciaux, Gilbert Maudire et Yvan Le Bras</i>
12:30 - 14:00	Déjeuner
14:00 - 15:00	Table ronde autour des Pôles de données
15:00 - 15:30	Pause café / Posters
15:30 - 16:00	> Geoflow : orchestrer des workflow en R pour gérer les données spatiales - <i>Julien Barde, UMR MARBEC, Institut de Recherche pour le Développement</i>
16:00 - 16:10	> Mise à disposition de divers Thesauri opérationnels sur GNetwork (fichiers RDF) - <i>Claudia Lavalley, UMR Territoire Environnement Teledetection Information Spatiale</i>
16:10 - 16:20	> Retour d'expérience : comment créer un DOI à partir de l'outil de catalogage Geonetwork - <i>Annick Battais, Géosciences Rennes</i>
16:20 - 16:30	> Génération automatique d'un catalogue standardisé à l'OSU OREME - <i>Juliette Fabre, Olivier Lobry, Observatoire de REcherche Méditerranéen de l'Environnement</i>
16:30 - 16:40	> Acquisition, traitement et diffusion des données de réseaux d'observations

Heures	événement
	océanographiques - SSS et ROSAME - <i>Philippe Téchiné, Laboratoire d'études en Géophysique et océanographie spatiales</i>
16:40 - 16:50	> pratiques de gestion des données dans les réseaux métiers de la MITI - <i>Maurice Libes, Marie-Claude Quidoz</i>
17:00 - 18:00	Discussions - Posters

Jeudi 7 novembre 2019

Heures	événement
09:00 - 12:00	ATELIER 1 : TP Réaliser un fichier NetCDF CF compliant - Joël Sudre, Maurice Libes & Didier Mallarino
09:00 - 12:00	ATELIER 2 : TP Générer automatiquement des fiches de métadonnées ISO19139 avec R - Emilie Lerigoleur, Wilfried Heintz, Juliette Fabre & Julien Barde
12:00 - 13:30	Déjeuner
13:30 - 16:30	ATELIER 3 : table ronde : Bonnes pratiques pour la gestion des données d'observation - Comité de pilotage SIST
13:30 - 16:30	ATELIER 4 : "Allô docteur" GeoNetwork - Mathias Rouan
16:30 - 18:00	Restitution plénière de l'atelier 3, discussions et clôture

Vendredi 8 novembre 2019

Heures	événement
09:00 - 12:30	Groupes de travail libres sur proposition