
DMP du projet "observatoire so-ii"

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "INRAE - Trame Structure" fourni par INRAE - Institut national de recherche pour l'agriculture l'alimentation et l'environnement.

Renseignements sur le plan

Titre du plan	DMP du projet "observatoire so-ii"
Version	Version initiale
Objet/périmètre du plan	première version du plan de gestion des données de so-ii
Domaines de recherche (selon classification de l'OCDE)	
Langue	fra
Date de création	2022-02-15
Date de dernière modification	2022-02-15
Licence	Creative Commons Attribution Non Commercial No Derivatives 4.0 International

Documents

(publications, rapports, brevets, plan expérimental...), sites web associés

- site internet de l'observatoire so-ii : <http://so-ii.org/>
- dataverse so-ii : <https://data.inrae.fr/dataverse/so-ii>
- librairies floodam : <http://www.floodam.org/>

Renseignements sur le projet

Titre du projet observatoire so-ii

Acronyme so-ii

Résumé so-ii est un système pluri-disciplinaire d'observation des impacts des inondations sur un territoire méditerranéen local qui comprend le bassin versant du Lez, le Bassin de l'Or ainsi que leurs bandes littorales (département de l'Hérault, France). L'objectif de cet observatoire de recherche scientifique est de produire des données basées sur l'observation des impacts des inondations de façon fine, continue, la plus exhaustive possible et pluri-disciplinaire en intégrant :

- les conséquences négatives (dommages matériels et sur les écosystèmes, pertes d'activités, impacts humains, pollutions) ainsi que les modifications de trajectoires (adaptations, réparations...) ;
- différents phénomènes d'inondations : par débordement de cours d'eau (fluvial), de ruissellement (pluvial), submersion (intrusions marines) ;
- toutes les amplitudes d'événements : des plus forts jusqu'aux plus faibles.

so-ii doit aussi permettre de caractériser les temporalités des impacts qui se produisent et sont résorbés ainsi que celles des modifications de trajectoires individuelles ou collectives (adaptation / urbanisme).

Il fait aussi partie des sites expérimentaux utilisé par le Groupe de Travail sur l'Analyse Multi-critère (GT-AMC) du Commissariat Général au Développement Durable (CGDD) sur lequel sont appliqués des méthodologies sur les dommages aux inondations. Il est également un service d'observatoire de l'OSU OREME.

Date de début 2020-01-01

Date de fin 2035-12-31

Produits de recherche :

1. Default research output
2. New research output 2

Contributeurs

Nom	Affiliation	Rôles
Balergue Pierre - 1158001	INRAE	<ul style="list-style-type: none">• Coordinateur du projet• Personne contact pour les données (Research Output 2)• Responsable du plan de gestion de données

Droits d'auteur :

Le(s) créateur(s) de ce plan accepte(nt) que tout ou partie de texte de ce plan soit réutilisé et personnalisé si nécessaire pour un autre plan. Vous n'avez pas besoin de citer le(s) créateur(s) en tant que source. L'utilisation de toute partie de texte de ce plan n'implique pas que le(s) créateur(s) soutien(nen)t ou aient une quelconque relation avec votre projet ou votre soumission.

DMP du projet "observatoire so-ii"

Informations sur la structure

Default research output

système d'observation des impacts des inondations (so-ii), c'est un observatoire de recherche scientifique est structuré en plusieurs Tâches d'Observations :

- TO Réseau d'Observateurs des Impacts des inondations (ROI) : afin de créer et suivre à long terme un réseau de personnes (habitants, entrepreneurs, exploitants agricoles) particulièrement exposés aux inondations.
- TO Retour d'Expérience suite à une inondation (REX) : afin de développer une méthodologie de retour d'expérience en réseau avec les différents acteurs du territoire afin de pouvoir vérifier et estimer les impacts ou non d'une inondation d'intensité variable sur la zone d'étude.
- TO Adaptation du bâti : pour faire un suivi de l'état d'adaptation de tout le bâti du territoire en relevant en particulier les protections individuelles aux inondations à travers des enquêtes de terrain non intrusives depuis l'espace public.

- Autre : à préciser dans la zone "Informations supplémentaires"

Observatoire de recherche scientifique. C'est également un service d'observation de l'OSU OREME, voir la page suivante pour plus d'informations : <https://oreme.org/>

pas d'identifiant actuellement

Nom, Prénom	Courriel	Rôle
Balergue Pierre	pierre.balergue@inrae.fr	Chargé de la gestion des données et animateur TO Adaptation du bâti et TO REX-impact
Brémond Pauline	pauline.bremond@inrae.fr	Coordinatrice projets MOOM-Agri et CAFRUA
Grelot Frédéric	frederic.grelot@inrae.fr	Coordinateur de so-ii
Marry Valériane	valeriane.marry@inrae.fr	Animatrice TO roi-habitant
Modjeska Maxime	maxime.modjeska@inrae.fr	Animateur TO roi-agriculture

INRAE / UMR G-EAU

- AQUA : Ecosystèmes aquatiques, ressources en eau et risques

INRAE
CGDD (Commissariat Général au Développement Durable)

New research output 2

Question sans réponse.

Question sans réponse.

Question sans réponse.

Question sans réponse.

Question sans réponse.

Question sans réponse.

Question sans réponse.

Informations sur le plan de gestion

Default research output

pas de DOI actuellement

Date	n° de version	Statuts	Auteur	Affiliation de l'auteur (se reporter à l' annuaire INRAE)	Validé par	Validé le
02/2022	version 1	en cours de validation	Pierre Balzergue	Ingénieur d'Etude INRAE/UMR G-EAU		

New research output 2

Question sans réponse.

Question sans réponse.

Présentation générale des données

Default research output

- Données produites par un tiers
- Données générées par la structure

La production de données primaires dans le cadre de l'observatoire so-ii sont réalisées dans le cadre des Tâches d'Observation décrites auparavant ainsi que de projets ponctuels. En fonction de la nature de la donnée, plusieurs méthodes d'élicitation sont prévues:

- des observations de terrain pour produire des données physiques
- des campagnes d'enquête avec des questionnaires fermés pour des données quantitatives
- des entretiens semi-directifs pour des données qualitatifs
- des modélisations de dommages

L'observatoire utilise également des données secondaires, produites par des structures tierces, pour réaliser des études et analyses préalables à la production de la donnée primaire. Dans le cas où ces données secondaires sont diffusées publiquement, nous pouvons les télécharger directement. Dans le cas contraire, elles sont obtenues et utilisées en établissant des partenariats cadrés par une convention d'échange de données ou bien par demande officielle (comme pour les fichiers fonciers par exemple).

- Simulation, modélisation
- Observation
- Enquête
- Code
- Analyse

Les données sont produites à partir des actions réalisées dans le cadre des Tâches d'Observation ainsi que dans les projets qui s'inscrivent, au moins en partie, dans le cadre de l'observatoire so-ii.

- **enquête**: Des enquêtes de terrain sous forme d'entretiens semi-directifs et qualitatifs sont menées dans la TO ROI, la TO REX-impact et dans plusieurs projets (CAFRUA,...)
- **observation** : des relevés basés sur les observations des dispositifs d'adaptations sur les bâtiments sont réalisés dans la TO Adaptation du bâti
- des **modélisations hydrauliques** de PAPI (Lez et Or) sont utilisées pour des traitements sur le territoire so-ii
- du **code sous forme de script R** de l'équipe est utilisé pour produire des nouveaux traitements, des rendus cartographiques, des rapports,...
- des analyses, études de l'équipe ou externe sont utilisés pour mener des nouveaux travaux

- Text
- Software
- Image
- Event
- Dataset

La majorité des données (primaires et secondaires) sont des jeux de données, des fichiers audiovisuels (photos, enregistrements d'entretiens), des rapports écrits ou des codes informatique (script R)

- données physiques : issues des observations de terrain telles que présence de batardeau et description, laisses de crues, érosion des sols et autres impacts physiques sur les parcelles agricoles et les bâtiments (habitations, exploitations agricoles, entreprises). Ces données sont produites en particulier dans la TO Adaptation du bâti et la TO REX-impact
- données quantitatives : issues des réponses aux questionnaires fermés de la TO REX
- données qualitatives : issues des réponses aux entretiens semi-directifs des entretiens préliminaires et de ceux de suivi de la TO ROI

- comma-separated values (.csv)
- langage R (.R, .rds)
- R Markdown (.rmd)
- shape files (.shp)
- limesurvey survey archive (.lsa)
- Extensible Markup Language (.xml) : pour les formulaires ODK

- Farming Systems and Practices
- Not specified
- Water resources
- Information management
- Economics
- Computer science

New research output 2

Question sans réponse.

Question sans réponse.

Question sans réponse.

Question sans réponse.

Question sans réponse.

Question sans réponse.

Droits de propriété intellectuelle

Default research output

INRAE et l'UMR G-EAU

New research output 2

Question sans réponse.

Confidentialité

Default research output

Type d'information	Nature de données	Type de confidentialité
prénom, nom de personne	primaire	Donnée à caractère personnel
adresse postale	primaire	Donnée à caractère personnel
rapport expertise dommages à l'assurance : déclaration de sinistres, montant de remboursement, délais	primaire	Donnée à caractère personnel
nom police assurance	primaire	Donnée à caractère personnel
raison sociale et SIRENE	primaire	Donnée à caractère personnel
type de contrat d'assurance	primaire	Donnée à caractère personnel
photos des dommages	primaire	Donnée à caractère personnel
point GPS de logements et de parcelles/exploitations agricoles	primaire	Donnée à caractère personnel
photo géolocalisée de la façade	primaire	Donnée à caractère personnel
comportement : installation de protection en cas de vigilance météo	primaire	Donnée à caractère personnel
historique des dommages humains et psychologiques subis : blessures, psychologique/moral, comportement suite à l'inondation,...	primaire	Donnée sensible

Chaque enquêteur signe un engagement de confidentialité (voir le paragraphe "sécurité - intégrité - traçabilité" pour plus d'informations).

Les données à caractère personnel ou sensible sont collectées en utilisant des outils libre et institutionnels comme Limesurvey INRAE et ODK et le serveur OSU OREME.

Ces données sont ensuite stockées sur un dossier confidentiel (dit "data confidential" du nextcloud so-ii) qui est accessible uniquement par trois membres de l'équipe de so-ii.

Une version "anonymisée" des données de collecte est créé et dans laquelle les données à caractère personnel (décrites au paragraphe précédent) sont soit supprimées et l'identité des répondants est rendu sous forme de pseudonyme. Cette version est utilisée pour les traitements.

Concernant les coordonnées géographiques relevés de personnes, parcelles et bâtis d'exploitations agricoles ou entreprises individuelles sont "brouillées" en utilisant la fonction R "jitter" qui créé une approximation des coordonnées tout en permettant de mieux différencier les différents points d'un même secteur sur une représentation cartographique.

Pour les données sensibles : actuellement aucune donnée sensible (impacts psychologique et physique d'une inondation) n'a encore été collectée mais il est prévu de le faire lors de prochaines campagnes. Leur caractérisation et traitement fait l'objet de discussions en cours avec la Déléguée à la Protection des Données d'INRAE, Nathalie Gandon.

Les données produites utilisées pour les analyses sont rendus anonymes et l'identité de la personne est pseudonymisée avant d'être utilisées.

Si des analyses sur les données nécessitaient d'utilisation des données brutes dont celles à caractère personnel ou une analyse de second niveau. Cette analyse serait réalisée par les chercheurs de so-ii en accord avec les finalités de l'observatoire (tel que la personne enquêtées a pu être informée et a donné son consentement). Si la finalité serait différente, les chercheurs de so-ii s'engagent à informer la personne et à lui demander à nouveau son consentement (conformément au RGPD).

New research output 2

Question sans réponse.

Question sans réponse.

Question sans réponse.

Partage des données

Default research output

INRAE a une politique de "science ouverte" qui dans la lignée de la directive européenne INSPIRE recommande de donner accès au maximum aux données et résultats de la recherche scientifique, selon les principes FAIR : Findable, Accessible, Interoperable, Reusable. Pour plus d'informations, consulter la page : <https://datapartage.inrae.fr/Produire-des-donnees-FAIR>

D'un autre côté, les données confidentielles (car non valorisées, tierces dont le producteur ne souhaite pas diffuser leur données) ou à caractère personnel/sensibles ne peuvent être partagées ou alors à un niveau très restreint selon les cas.

-
- évaluation de la vulnérabilité et des dommages liés aux inondations

- bilan sur les dommages d'une inondation
- comportements individuels face au risque inondation
- état des protections individuelles sur le bâti aux inondation

Non

Elles sont publiées sur le dataverse so-ii de data.inrae à : <https://data.inrae.fr/dataverse/so-ii>
Les rapports écrits sont publiés sur le portail d'archive ouverte HAL

-
- Autre
 - Certains jeux de données sont publiées avec tous (open access) quand les données à caractère personnel potentiellement existantes ont été pseudonymisées et qu'elles ont pu être valorisées a minima par l'équipe
 - Certains jeux de données seront publiés uniquement avec les partenaires de l'observatoire ou dans le cadre de projets menés en lien avec l'observatoire so-ii

-
- Licence ouverte <https://www.etalab.gouv.fr/licence-ouverte-open-licence> (compatible CC-BY)

Cette licence est privilégiée actuellement

New research output 2

Question sans réponse.

Question sans réponse.

Question sans réponse.

Question sans réponse.

Question sans réponse.

Question sans réponse.

Organisation et documentation des données

Default research output

L'observatoire so-ii a pour l'objectif de développer des nouvelles méthodologies d'observation. Ce processus se déroule de la façon suivante :

1. collecte des données : en fonction de la méthodologie développée ou utilisée, la collecte est réalisée sur des outils de type relevés d'observation terrain (ODK Collect, QField,...) et d'entretien (Limesurvey).
2. pour les relevés d'observation de terrain, les données sont compilées sur le serveur ODK Aggregate de so-ii géré par l'OSU OREME, des extractions de ce serveur sont compilées dans le dossier "data-common" du serveur nextcloud INRAE. Pour les entretiens, les données qui incluent des données à caractère personnel/sensible des personnes enquêtées sont stockées dans le dossier "data-confidential", une version pseudonymisée (sans les données à caractère personnel) est ensuite sauvegardée dans le dossier "data-common" du serveur nextcloud INRAE.
3. les jeux de données sont publiés sur l'entrepôt de données "Data INRAE" (qui deviendra prochainement "Recherche Data Gouv") dans un premier temps en pré-production puis en production.
4. L'archivage des données est prévue sur le serveur nextcloud INRAE pour la durée d'existence de l'observatoire, estimée à 15 ans, puis de façon pérenne sur des solutions d'archivage INRAE ou de la recherche scientifique française selon l'évolution des entrepôt d'archivage.

	Origine, mode de production des métadonnées (ex : saisie manuelle, annotation automatique...)	Standard, Vocabulaires associés	Conditions ou fréquence de la mise à jour (si applicable) (ex : changement de l'accessibilité)
Citation Metadata	saisie via l'interface Web, API	DDI LiteOuvre une nouvelle fenêtre , DDI 2.5 CodebookOuvre une nouvelle fenêtre , DataCite 3.1Ouvre une nouvelle fenêtre , et Dublin Core's DCMI Metadata TermsOuvre une nouvelle fenêtre . Vocabulaire contrôlé pour la langue (ISO 639-1Ouvre une nouvelle fenêtre).	quand nécessaire
Geospatial Metadata	saisie via l'interface Web, API	DDI Lite, DDI 2.5 Codebook, DataCite, and Dublin Core. Vocabulaire contrôlé pour les pays ISO 3166-1Ouvre une nouvelle fenêtre .	
Social Science & Humanities Metadata	saisie via l'interface Web, API	DDI Lite, DDI 2.5 Codebook, and Dublin Core.	
Life Sciences Metadata	saisie via l'interface Web, API	ISA-Tab SpecificationOuvre une nouvelle fenêtre , . Vocabulaires contrôlés pour différentes métadonnées (sous-ensemble de OBI OntologyOuvre une nouvelle fenêtre et NCBI Taxonomy for OrganismsOuvre une nouvelle fenêtre)	

Les jeux de données (dataset) publiés par so-ii sont accompagnés d'un fichier "readme.txt" qui décrit très concrètement la façon de lire et d'utiliser les données, leur organisation et autres informations utiles.

Les fichiers sont stockés sur le serveur nextcloud INRAE et leur nom sont standardisée en décrivant :

- le mot clé de l'étude et/ou de la TO ou projet
- la date d'acquisition ou d'entretien en format : YYYYMMDD
- le versionnage n'est pas utilisé ou très peu car l'équipe de so-ii travaille avec le logiciel de suivi de versionnage Gitlab INRAE

- TO Adaptation du bâti : pour qualifier la qualité des données physiques issues des observations produites :
 - des campagnes tests sont menées soit en faisant des relevés exhaustifs des bâtiments d'un sous-quartiers ou soit réalisant des relevés de façon aléatoire sur un secteur plus grand : à l'échelle d'un quartier ou d'une commune. Les résultats de ces campagnes tests sont comparés à ceux de la campagne réelle.
 - les données sur l'adaptation du bâti sont adossées à la Base d'Adresse Nationale (BAN) ce qui permet de vérifier que les relevés réalisés n'ont pas omis certaines adresses.
- TO REX : pour qualifier la qualité des données quantitatives issues des questionnaires fermés de la TO REX aux habitants et agriculteurs impactés :
 - les enregistrements audio des entretiens sont écoutés pour vérifier que les réponses écrites sont bien cohérentes avec ce qui a été dit
 - les informations sur les impacts issues des réponses sont recoupées avec les données secondaires sur les impacts (de la Chambre d'Agriculture, Préfecture, SDIS, presse)

New research output 2

Question sans réponse.

Question sans réponse.

Question sans réponse.

Question sans réponse.

Question sans réponse.

Stockage et sécurité des données

Default research output

Les flux de gestion des données (primaires et secondaires) ainsi que leur supports de stockage sont décrits dans le schéma suivant : <http://so-ii.org/wp-content/uploads/2022/03/type-stockage-donnee.png>

Les données primaires occupent actuellement un volume d'environ 28 Go (dossiers "data common" et "data confidential").
Les données secondaires occupent un volume d'environ 400 Go et sont stockées sur un autre serveur distant sécurisé

Il semble qu'INRAE n'ait pas encore publié sa politique de sécurité pour le système d'information hébergeant les données (nextcloud).

Certaines données font l'objet d'échange ou de partage dans le cadre des TO de so-ii (en particulier pour la TO REX-impact) et dans le cadre de projets en lien avec so-ii. Tous ces échanges font l'objet de conventions signées qui défini le cadre de ces échanges et de la diffusion des données.

Un engagement de confidentialité est signé par chacun des chercheurs, stagiaires et étudiants contribuant à l'observatoire so-ii. Il comporte les engagements suivants :

- garder le secret sur toute information relative aux données (document, fichier de données, résultats) dont la personne prendra connaissance, oralement, par support visuel ou par écrit, préalablement, pendant ou à la suite de la mission évoquée ci-dessus, et m'oblige à :
- prendre à l'égard de ces données, toutes les précautions pour empêcher leur divulgation ;
- utiliser uniquement des outils informatiques sécurisés et institutionnels (dont les serveurs sont gérés par le service public français)
- ne pas divulguer ni transmettre à un tiers, directement ou indirectement, et quel que soit le support (papier, fichier électronique, ou autre) tout ou partie de ces données ;
- s'abstenir de toute exploitation directe ou indirecte de ces données en dehors du cadre de la mission ;
- reconnaître que les données demeurent la propriété exclusive de INRAE / UMR G-Eau.
- remettre à la fin de ma mission tous les documents relatifs au projet qui m'ont été transmis en vue de l'exécution de ma mission, y compris les copies qui ont pu en être faites et à ne conserver aucune des données relatives à la mission

New research output 2

Question sans réponse.

Question sans réponse.

Question sans réponse.

Question sans réponse.

Question sans réponse.

--

Archivage et conservation des données

Default research output
Toutes les données élaborées et dérivées sont conservées à long terme ainsi que les données brutes utiles (à une ré-exploitation ultérieure)
Sur le serveur nextcloud de data INRAE en attendant qu'INRAE développe sa solution d'archivage (prévu)
A minima la durée d'existence de l'observatoire so-ii soit environ 15 ans puis il est envisagé d'archiver les données dans un entrepôt de stockage de données pérenne.
Budget propre d'INRAE/ UMR G-Eau

New research output 2
Question sans réponse.
Question sans réponse.
Question sans réponse.
Question sans réponse.