



# ARDoISE

2<sup>e</sup> JOURNÉE D'ÉTUDE

17 décembre 2024

9h > 17h

Un enjeu de la reproductibilité de la science :  
l'ouverture des codes sources et logiciels

ARDoISE est un projet financé avec le soutien du Fonds National pour la Science Ouverte

## PROGRAMME DU MATIN

8h45 • Café d'accueil



9h00 • OUVERTURE DE LA JOURNÉE

- Patrick Gros, directeur du centre Inria de l'Université de Rennes (sous réserve)
- Mickaël Kepenekian, Université de Rennes, vice-président chargé de la science ouverte

9h45 • RETOUR SUR L'ÉTUDE MENÉE SUR LA PRODUCTION ET LA VALORISATION DES LOGICIELS ISSUS DE LA RECHERCHE PUBLIQUE FRANÇAISE

- Isabelle Blanc, administratrice ministérielle des données, algorithmes et codes sources au Ministère de l'Enseignement Supérieur et de la Recherche - DGRI

10H30 • LE CODE AU-DELÀ DU FAIR. MISSIONS ET ACTIVITÉS DU COLLÈGE «CODE SOURCES ET LOGICIELS» AU SEIN DU COMITÉ POUR LA SCIENCE OUVERTE

- Nicolas Rougier, centre Inria de l'Université de Bordeaux  
« Les principes FAIR sont un ensemble de lignes directrices visant à faciliter l'ouverture des données scientifiques afin d'encourager leur réutilisation et leur reproductibilité. Cependant, les logiciels de recherche diffèrent considérablement des données par leur nature vivante et leur relation avec les logiciels libres et open source. C'est la raison pour laquelle nous proposons une feuille de route à plusieurs niveaux pour améliorer la situation des logiciels de recherche, incluant toutes les parties prenantes de leur écosystème : le personnel scientifique, mais aussi les institutions, les bailleurs de fonds, les bibliothèques et les éditeurs. »

11h15 • Pause

11H45 • EXISTE T-IL UNE CRISE DE LA REPRODUCTIBILITÉ EN SCIENCE ?

- Florian Naudet, Inserm  
« La présentation traitera des facteurs ayant conduit la communauté scientifique à évoquer une crise de la reproductibilité en science. Après avoir défini quelques concepts-clés, nous examinerons pourquoi la reproductibilité constitue un critère fondamental de fiabilité dans la démarche scientifique, en abordant ses différentes dimensions. L'objectif sera également d'explorer les stratégies et approches proposées pour améliorer la reproductibilité de nos travaux de recherche. »

12h30 • Pause déjeuner (buffet)





# ARDoISE

2<sup>e</sup> JOURNÉE D'ÉTUDE

17 décembre 2024

9h > 17h

Un enjeu de la reproductibilité de la science :  
l'ouverture des codes sources et logiciels

ARDoISE est un projet financé avec le soutien du Fonds National pour la Science Ouverte

## PROGRAMME DE L'APRÈS-MIDI

### 14h00 • ARVEST : DÉVELOPPEMENT D'UN OUTIL NUMÉRIQUE POUR LA GESTION DE COLLECTIONS MULTIMODALES

• Clarisse Bardiot et Jacob Hart, Université Rennes 2, CNRS, UR Arts Pratiques et Poétiques  
« Depuis 2022, notre équipe travaille à la conception d'un outil destiné à la création, la gestion et la navigation de corpus multimodaux pour les chercheurs en sciences humaines. Après avoir exposé les enjeux qui sous-tendent ce projet et son intégration dans le cadre du projet ERC From Stage to Data, nous présenterons les principales fonctionnalités de l'outil ainsi que les étapes clés de son développement. Nous expliquerons notamment comment nous avons abordé les questions de science ouverte, de reproductibilité et de pérennité. »

### 14H45 • ENVIRONNEMENT LOGICIEL REPRODUCTIBLE, À QUELLE ÉCHÉANCE ? LES SERVICES OFFERTS PAR SOFTWARE HERITAGE, ARCHIVE OUVERTE DES CODES SOURCES ET LOGICIELS

• Simon Tournier, Université Paris Cité, Institut de Recherche Saint Louis, Inserm US53, CNRS  
« Aujourd'hui traiter des données numériques est une activité de tous les champs de la recherche. Ce traitement implique nécessairement un environnement logiciel. Nous discuterons les difficultés que cela pose en terme de méthode scientifique et aborderons quelques contre-mesures (bonnes pratiques et/ou outils) contribuant à une recherche plus robuste, donc mieux reproductible. »

### 15h30 • Pause

### 16H00 • LA LINGUISTIQUE À L'HEURE DU CODE OUVERT

• Thomas Gaillat et Cyriel Mallart, Université Rennes 2, UR Linguistique, Ingénierie et Didactique des Langues (LIDILE)  
« Le traitement linguistique de corpus textuels impose la mise en œuvre de méthodes TAL (Traitement Automatique des Langues) reposant sur du code pour produire des jeux de données. Les analyses linguistiques et statistiques qui en découlent dépendent donc du code, de sa disponibilité, de sa clarté et de sa fiabilité. Ces trois éléments sont les conditions de la reproductibilité des résultats d'expérience. Ils permettent aussi la transmission de la méthodologie et ainsi, l'évolutivité des questions de recherche sous-tendues par les expériences. Le code partagé facilite le dialogue scientifique au sein d'une équipe multidimensionnelle rassemblant des domaines de recherche variés. »

### 16h45 • CONCLUSION DE LA JOURNÉE

• Grand témoin : Thierry Fournier, Université de Rennes, directeur adjoint du Service Commun de la Documentation

À BIENTÔT