



HAL
open science

FAIR-Aware, un outil d'auto-évaluation de ses pratiques de gestion de données

Justine Fabre

► **To cite this version:**

Justine Fabre. FAIR-Aware, un outil d'auto-évaluation de ses pratiques de gestion de données. 2021. hal-04211423

HAL Id: hal-04211423

<https://hal.science/hal-04211423>

Submitted on 19 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

FAIR-Aware, un outil d'auto-évaluation de ses pratiques de gestion de données

Le [deuxième Plan national pour la science ouverte](#), présenté le 6 juillet dernier par la ministre de l'Enseignement supérieur, de la Recherche et de l'Innovation Frédérique Vidal, se fixe comme objectif dans son deuxième axe de "Structurer, partager et ouvrir les données de la recherche". Plus encore, ce Plan insiste sur la nécessité de "faire en sorte que les données produites par la recherche publique française soient progressivement structurées en conformité avec les principes FAIR (faciles à trouver, accessibles, interopérables, réutilisables)".

Cette préoccupation se retrouve également au cœur des exigences des financeurs de la recherche, et notamment de la Commission européenne. Celle-ci invite en effet depuis plusieurs années les chercheurs dont elle subventionne les projets à s'attacher à FAIRiser leurs données, et a choisi de structurer explicitement [ses modèles de plan de gestion de données](#) autour de ces principes, dans une rubrique "FAIR data".

Genèse de l'outil

Pour aider les chercheurs à respecter ces exigences, et faciliter la mise à disposition de données de qualité, le projet [FAIRsFAIR](#) ("Fostering FAIR Data Practices in Europe") s'efforce de leur fournir divers outils et solutions pratiques, élabore des normes mondiales pour la certification FAIR des référentiels et des données qu'ils contiennent, et contribue aux politiques et pratiques qui feront du futur European Open Science Cloud (EOSC) une infrastructure opérationnelle¹.

Dans cette optique, trois des partenaires du projet (le [Digital Curation Centre](#) au Royaume-Uni, l'[Université de Brême](#) en Allemagne et le [Data Archiving and Networked Services](#) aux Pays-Bas) ont publié en juin 2020 l'outil [FAIR-Aware](#), un petit questionnaire d'auto-évaluation en ligne visant à sensibiliser ses utilisateurs à l'importance des principes FAIR dans la gestion et la valorisation de leurs données, en lien avec leur diffusion dans des entrepôts.

FAIR-Aware se présente sous la forme d'un formulaire de 10 questions, consacrées à chacun des principes FAIR, et propose des explications complémentaires et suggestions d'approfondissement sur l'ensemble des thématiques abordées (identifiants pérennes, archivage, etc.). Il aide ainsi les chercheurs à mieux comprendre les enjeux d'une bonne gestion de données, et s'efforce de proposer des pistes de perfectionnement concrètes.

Il peut aussi être utilisé lors d'une formation ou d'un atelier de sensibilisation à la bonne gestion des données de recherche, en servant de trame au propos de l'animateur : cette façon de faire a été testée avec succès à l'URFIST de Paris...

Naissance d'une version française

Devant l'intérêt flagrant de l'outil, il a vite semblé dommage qu'il ne soit disponible qu'en anglais, au risque de décourager certains utilisateurs potentiels. Aussi a-t-il vite semblé intéressant à l'URFIST d'en proposer une traduction en français, et c'est en ce sens que des contacts ont été pris dès l'été 2020 avec l'équipe de DANS, qui suit de près le développement et l'adoption de FAIR-Aware. Le [service IST de l'Institut de Recherche pour le Développement](#) ayant eu la même idée, nous avons travaillé conjointement à l'adaptation de l'outil, entre décembre 2020 et mars 2021. Contexte sanitaire oblige, tous les échanges URFIST-IRD ainsi que les points réguliers avec DANS ont eu lieu à distance, mais sans que cela n'ait d'impact particulier sur le projet, qui se prêtait assez bien à un tel mode opératoire ! Même la langue n'a pas été une barrière lors des discussions en visioconférence,

même si les notions de français de certains membres de l'équipe de DANS nous ont été précieuses pour éclaircir certains points...

C'est que l'adaptation de FAIR-Aware n'a pas posé uniquement des questions lexicales. Ce travail mené en commun a également été l'occasion pour nos collègues de DANS d'avoir des retours sur les formulations choisies dans le texte original, et de réfléchir à des modifications utiles pour leurs propres usages. Les passages de FAIR-Aware relatifs aux vocabulaires sémantiques, notamment, ont posé question : les chercheurs ont-ils tous une bonne idée de ce que peut être une ontologie, un vocabulaire contrôlé ? Jusqu'où peut-on se permettre de simplifier certains concepts, tout en préservant leur richesse ?

D'autres questions se sont également posées sur le référentiel à adopter pour le choix du domaine de recherche. Même si certains référentiels "français" comme celui du Hcéres ou de HAL auraient pu paraître plus adaptés aux pratiques des utilisateurs, il a finalement été décidé de rester fidèles au référentiel original, basé sur la classification Re3data², de façon à pouvoir comparer les réponses au questionnaire sans avoir à fournir de table de mapping entre les disciplines. Certaines adaptations ont en revanche été nécessaires pour les propositions de réponse aux questions relatives au poste occupé par le répondant, ou à son employeur³.

Ce genre d'interrogation, qui s'était déjà posé lors de la conception de la version originale de l'outil, est la raison pour laquelle FAIR-Aware comprend également un sous-ensemble de questions visant à recueillir l'avis de ses utilisateurs sur l'outil, notamment sur la formulation des différents points abordés et leur pertinence. Les informations ainsi récupérées serviront à améliorer l'outil, mais aussi à mieux connaître les pratiques de ses utilisateurs en matière de gestion des données de recherche.

De la traduction à la mise en ligne

En parallèle de la traduction proprement dite s'est posée la question du vecteur de diffusion de l'outil. La plateforme [DoRANum](#), qui hébergeait ou signalait déjà un certain nombre d'outils utiles à la gestion des données de recherche, s'est très vite imposée comme une évidence, et un accord de principe avec l'équipe responsable de la plateforme (Inist-CNRS et réseau des URFIST) a pu être conclu dès janvier 2021. Des membres du service Formation-DoRANum ont ainsi rejoint le projet, avec le soutien du service web de l'Inist-CNRS.

Le code-source de FAIR-Aware, prévu dès le départ pour être très simple d'utilisation et d'intégration au sein d'un site internet, est [librement téléchargeable sur Github](#), ce qui a beaucoup facilité le travail de l'équipe DoRANum. Après quelques nouvelles réunions en visio associant les équipes techniques de DANS, [la version française de FAIR-Aware](#) a ainsi pu être mise en ligne début juillet 2021, dans la rubrique "[Enjeux et bénéfices](#)" de DoRANum aux côtés des explications sur les [principes FAIR](#) !

Et maintenant ?

L'outil FAIR-Aware est appelé à évoluer, en fonction des retours de ses utilisateurs. Les partenaires du projet FAIRsFAIR vont continuer à l'améliorer, et s'efforceront d'en proposer de nouveaux cas d'usage : il se murmure qu'à l'automne prochain, il devrait être possible pour un formateur de lancer une session personnalisée de FAIR-Aware, et d'en recueillir les réponses pour mieux connaître les usages de son groupe de stagiaires...

D'une manière générale, les réponses au questionnaire en français seront collectées par DANS, en même temps que les réponses au questionnaire en anglais. Du fait de l'alignement des questions entre les deux versions, cette mutualisation permettra de générer des statistiques globales. Les données produites lors des utilisations de la version française de l'outil seront également transmises deux fois par an à l'Inist-CNRS, pour mieux connaître les pratiques et attentes propres à la

communauté francophone, ce qui permettra de faire évoluer la version française de l'outil en parallèle de la version anglaise.

N'hésitez pas à vous emparer de l'outil, et à nous transmettre vos retours d'expérience à son sujet !

De l'autre côté du projet : le retour d'expérience de Linas Cepinkas ([DANS](#))

Avez-vous reçu des demandes de traduction de FAIR-Aware dans d'autres langues ?

Depuis le lancement de FAIR-Aware, nous avons en effet reçu plusieurs demandes concernant la traduction de l'outil. Nous espérons que cette collaboration réussie avec nos collègues français inspirera d'autres enthousiastes de la Science Ouverte à rendre FAIR-Aware disponible dans différentes langues, et à atteindre un public plus large en Europe et au-delà.

Qu'est-ce que ce genre de projet de traduction peut apporter à la version originale de FAIR-Aware ?

La promotion multilingue des pratiques FAIR de gestion des données nous permet d'augmenter considérablement notre audience, ce qui ne serait pas possible avec la seule version en anglais. Comme le montre l'exemple de la version française de FAIR-Aware, la traduction permet également de faire émerger des besoins spéciaux et de s'adapter aux contextes particuliers des différents utilisateurs. C'est d'autant plus important que le niveau de connaissance et de compréhension des principes FAIR est assez hétérogène, et les infrastructures de soutien à ces principes très variées. Chaque traduction peut être adaptée pour des communautés particulières d'utilisateurs, ce qui garantit que l'outil reste pertinent et à jour.

Bien que l'équipe du projet n'ait pas prévu initialement la traduction de FAIR-Aware dans d'autres langues, cette collaboration réussie avec [l'Inist-CNRS](#), [l'URFIST de Paris](#) et le [service IST de l'IRD](#) démontre l'intérêt croissant et la nécessité de mettre de tels outils à la disposition des chercheurs dans différents pays.

Quels sont les autres outils proposés par FAIRsFAIR qu'il serait intéressant de découvrir en parallèle de FAIR-Aware ?

Parallèlement à FAIR-Aware, FAIRsFAIR propose [F-UJI](#), un service en ligne permettant d'évaluer automatiquement le caractère FAIR de jeux de données spécifiques, sur la base d'indicateurs développés par le projet [FAIRsFAIR](#).