

Bilan scientifique Datacc'+

Août 2020-Août 2022

I- Objectifs

Datacc'+ propose un dispositif complet d'accompagnement des chercheurs en matière de gestion des données, depuis la production et la collecte jusqu'à la publication et la préservation de celles-ci. Le projet repose sur une approche originale, volontairement disciplinaire et conçue en adéquation avec les principes de la science ouverte. L'une des particularités du dispositif consiste à améliorer la traçabilité des travaux scientifiques, leur reproductibilité et l'efficacité de la recherche à travers la publication de résultats négatifs, en parallèle du déploiement de solutions de cahiers de laboratoire électroniques. Cette deuxième phase fait suite à un premier projet amorcé grâce au concours du GIS-CollEx-Persée en 2019, en vue de consolider le service mis en place tout en le diversifiant.



II- Réalisations

1- Sensibilisation à la diffusion des résultats négatifs

Production de ressources didactiques

Les premiers mois ont été consacrés à la conception et la rédaction de ressources sur l'enjeu de la diffusion des résultats négatifs, qui fait désormais l'objet d'une [rubrique dédiée](#) sur le site de Datacc.org. Cette première étape a permis la publication, fin 2020, d'un [article de vulgarisation](#) sur cet enjeu, constamment mis à jour depuis et mettant en avant les revues disciplinaires accueillant ce type de résultats. Un travail complété par la production d'une [infographie ludique](#), grâce au double appui de Doranum et de Thierry Billard, chercheur en chimie à Lyon 1 et coordinateur scientifique de Datacc. Cette thématique a également fait l'objet d'une veille régulière dans le cadre de [six bulletins](#) abordant la question des résultats non-concluants, entre autres sujets.

Publication d'un article de revue

A partir du premier semestre 2021, l'équipe-projet s'est pleinement mobilisée afin de lancer une enquête en direction des chercheurs issus de quatre communautés (chimie, physique, sciences de l'ingénieur et sciences de l'environnement). L'objectif principal consistait à mieux cerner les pratiques des chercheurs tout en appréciant l'opportunité de créer une plateforme favorisant le dépôt de ce type de résultats dans une discipline donnée. L'enquête visait ainsi à répondre à la série de questionnements suivants :

- Quelle valeur et quelle utilité les chercheurs accordent-ils aux résultats négatifs ?
- Quelles sont leurs pratiques de communication de ce type de données ?
- Quelle proportion de chercheurs soutient l'idée de création d'une plateforme spécialisée dans le partage de résultats négatifs ?

Composée d'une trentaine de questions, l'enquête a attiré près de 800 réponses, dont 310 étaient pleinement exploitables. Les conclusions sont riches d'enseignements. Parmi les chiffres les plus éloquentes, nous pouvons citer que plus de 8 chercheurs sur 10 ont déjà produit des résultats négatifs qu'ils jugent utiles pour leur communauté. En revanche, seuls 12% ont eu l'occasion de publier ce type de résultats dans des revues à comité de lecture. Faute de diffusion adéquate, près de 30% des répondants rapportent avoir découvert qu'une autre équipe avait effectué les mêmes expériences aboutissant aux mêmes résultats infructueux.

La richesse des résultats obtenus et le caractère inédit de ce type d'enquête auprès des communautés précitées a trouvé un écho auprès de la revue *Learned Publishing*, qui a accepté le manuscrit. La version finale (en anglais) est [librement accessible sur le site de l'éditeur](#) depuis janvier 2022. Le pré-print, en français, a quant à lui été mis en ligne sur [HAL](#) et [Zenodo](#) dès l'automne 2021. La publication de cet article a permis de diffuser amplement les fruits de l'enquête auprès des différentes communautés intéressées. Elle a également déclenché des sollicitations d'interventions dans le cadre de conférences sur la science ouverte.

Participation à des séminaires et mise en place de partenariats

Ce travail de recherche a donné lieu à une communication dans le cadre d'un séminaire organisé par la Faculté des Sciences et Ingénierie de Sorbonne Université le 7 décembre dans le cadre du cycle de rencontres baptisé « le 13h de la science ouverte ». Il a également nourri l'intérêt de l'émission scientifique « Grand Labo » diffusée sur Youtube, lors de son édition spéciale du 25 janvier consacrée aux résultats négatifs, à laquelle la représentante du projet Datacc a été conviée.

Enfin, les travaux menés ont permis un rapprochement avec l'équipe de chercheurs à l'origine du projet Gaffex (Gathering of Fundamentally Failed EXperiments), qui vise à lancer une plateforme de signalement des expériences infructueuses en chimie. La coopération menée avec les porteurs de ce projet nous a amenés à soutenir le développement de cette dernière plutôt qu'à créer une nouvelle infrastructure ad hoc.

2- Développement de l'usage des cahiers de laboratoire électroniques

La deuxième phase du projet a par ailleurs été l'occasion de poursuivre l'action menée sur les cahiers de laboratoire électroniques, en franchissant le cap de la souscription à plusieurs solutions : Mbook (10 licences annuelles réparties entre trois équipes de chimie de synthèse), Sciformation (1 licence perpétuelle) et elabFTW (qui comptabilise près de 150 comptes utilisateurs issus d'établissements variés, dont les principaux sont Lyon 1, l'UGA, et l'Université de Savoie). La pérennisation des financements au-delà des moyens mis à disposition dans le cadre de l'appel à projet CollEx-Persée a été anticipée, avec le projet de solliciter des crédits à partir de 2023 dans le cadre du plan de développement concerté de l'UCBL (dotation sur projet).

Le laboratoire SIMAP, à Grenoble, a expérimenté ElabFTW dans le cadre du projet CAOLILA (CAhier de laboratoire Ouvert et Libre à destination des LABoratoires). Une instance d'ElabFTW a donc été déployée à l'UGA grâce à une collaboration entre l'UAR Gricad, la cellule data Grenoble Alpes et les laboratoires en science des matériaux. Ce projet, toujours en phase de test, fait l'objet d'un [retour d'expérience](#) très positif lors du séminaire "Pour une Recherche Reproductible" du 5 mai 2022 : "*Mise en œuvre du cahier de laboratoire eLabFTW à l'UGA/GINP : le projet CAOLILA*". Après la phase de test, il est prévu à l'UGA de développer l'usage des carnets de laboratoires électroniques auprès des équipes

de recherche volontaires en proposant un accompagnement le plus personnalisé possible. Il s'agit d'une des priorités affichées par la cellule data Grenoble Alpes.

L'intérêt pour les cahiers de laboratoire électronique s'est confirmé au gré des événements organisés par Datacc, dont le webinaire du 3 février 2021, en présence de Sylvain Monnier, chercheur en biophysique à Lyon 1 (laboratoire ILM), qui a rassemblé 120 participants issus de différents établissements de l'ESR. Entre 2021 et 2022, l'équipe-projet a organisé ou participé à 10 séminaires ou ateliers (en ligne et en présentiel) en liaison avec la question des cahiers de laboratoire électroniques, rassemblant au total un public de plus de 400 personnes sensibilisées à cet enjeu.

Type d'événement	Date
Webinaire sur elabFTW	03/02/2021
Séminaire en ligne sur elabFTW auprès du LMI (labo Lyon 1)	28/05/2021
Atelier de formation en présentiel sur elabFTW à la BU Sciences Lyon 1	28/09/2021
Atelier de formation en présentiel sur elabFTW à la BU Sciences Lyon 1	12/10/2021
Séminaire en présentiel sur Mbook à l'ICBMS	18/11/2021
Séminaire en présentiel à l'ISCR de Rennes I	26/11/2021
Atelier de formation en présentiel auprès d'une équipe du LEM (labo Lyon 1)	06/12/2021
Séminaire en ligne pour les ambassadeurs des données de l'Université de Lorraine	16/12/2021
Séminaire en ligne « Pour une recherche reproductible »	05/05/2022
Webinaire sur elabFTW	28/06/2022

La poursuite du projet Datacc a non seulement conforté la pertinence de l'axe sur les cahiers de laboratoire électroniques, mais elle a également permis la tenue d'ateliers de formation d'un genre nouveau, puisque la BU Sciences de Lyon 1 n'avait jusqu'ici jamais eu l'occasion de former les chercheurs à l'utilisation de ce type de logiciels. Cette action inédite a obtenu un bon taux de fréquentation, avec 24 chercheurs présents sur les 2 sessions organisées en septembre et octobre 2021. L'action menée a par ailleurs ouvert la voie à des projets d'applications pédagogiques de l'utilisation d'elabFTW dans le cadre de TP en biochimie, pour l'année 2022-2023.

Enfin, le travail mené dans le cadre du projet Datacc a été mis à profit dans le cadre de la réflexion nationale portée par le Comité pour la Science Ouverte. La représentante de Datacc et son adjoint ont ainsi participé à la rédaction du [rapport national](#) sur les cahiers de laboratoire électroniques, paru en janvier 2022.

3- Publication de lettres d'information et partage de la veille

Entre septembre 2020 et mai 2022, 18 lettres d'information et autant de bulletins de veille ont été conçus et adressés à la communauté. L'ensemble de ces productions peut être consulté librement sur le site de Datacc (archives [newsletters](#) et [bulletins de veille](#)). Ces ressources recensent utilement et à un rythme mensuel l'ensemble des informations relatives à la gestion des données et à la science ouverte, en prenant soin d'apporter un éclairage disciplinaire sur ces enjeux. La priorité est ainsi donnée aux avancées et défis de la gestion des données de recherche en chimie et en physique, même si les développements constatés dans d'autres disciplines peuvent ponctuellement émailler et enrichir les bulletins de veille. En juillet 2022, la newsletter comptabilise 317 inscrits, issus d'établissements très variés et répartis entre les personnels de recherche et d'IST.

La veille réalisée par l'équipe projet permet une mise à jour régulière de la plateforme : la liste - et la description - des entrepôts disciplinaires permettant de trouver (actuellement une soixantaine d'entrepôts signalés) ou de déposer (une quarantaine d'entrepôts signalés) des données fait l'objet d'un enrichissement permanent. Ce travail de mise à jour a été l'occasion de solliciter deux réseaux de documentalistes de l'IN2P3 (réseau Democrite) et de l'INSU et de nouer des contacts intéressants.

4- Accompagnement à la rédaction de plans de gestion de données

En cohérence avec l'action engagée lors de la première phase du projet, l'équipe-projet a poursuivi et amplifié le service rendu en direction des équipes de recherche amenées à devoir élaborer un plan de gestion de données. L'équipe-projet à Lyon 1 a animé deux ateliers consacrés à cet enjeu fin 2020 et comptabilise 12 accompagnements à la rédaction de PGD entre 2021 et 2022.

A Grenoble, l'équipe-projet a animé 1 atelier sur les PGD et 1 atelier sur l'outil DMP OPIDor en 2022. 15 accompagnements à la rédaction de PGD ont été menés entre 2021 et 2022. Ces accompagnements ont été l'occasion de présenter rapidement la plateforme Datacc', qui a rencontré l'intérêt des porteurs de projet.

Conclusion

On peut tout d'abord souligner que l'équipe projet travaille en permanence à améliorer la visibilité de Datacc' dans les communautés concernées et dans le réseau des personnels d'accompagnement à la gestion des données, au travers notamment du bulletin de veille.

L'une des vertus principales du projet repose sur le travail étroit mené entre les BU et les équipes de recherche. Ce partenariat requiert une adaptabilité forte de la part du bibliothécaire qui doit faire cas des pratiques de chaque spécialité disciplinaire ou sous-disciplinaire afin d'ajuster le service en fonction. L'offre de cahiers de laboratoire électroniques introduite par la BU de Lyon 1 a ainsi pris en compte les besoins hétérogènes des chercheurs en ne proposant pas une solution unique.

Le lancement d'une réflexion approfondie sur la diffusion des résultats négatifs et l'étude qui en a découlé ont permis d'enrichir l'état des connaissances en la matière, tout en sensibilisant les communautés.

De manière plus globale, les thématiques d'intervention sont en cohérence avec l'agenda national qui inclut désormais la question des cahiers de laboratoire électroniques tout comme celle des résultats négatifs. En témoignent la publication, au printemps 2022, de recommandations du CNRS sur les solutions électroniques de cahiers de laboratoire ou encore la parution du décret du 3 décembre 2021 sur l'intégrité scientifique, qui encourage la communication des résultats négatifs.