

## L'IA qui valorise les données de la recherche animale



Les études précliniques produisent une quantité massive de données. Pourtant, une partie de leur valeur se perd : informations incomplètes, protocoles difficilement traçables, résultats non réutilisables. Certaines expériences doivent être répétées faute de pouvoir exploiter celles déjà réalisées. À l'heure où l'IA promet d'accélérer l'analyse scientifique, un préalable s'impose : disposer de données bien décrites, contextualisées et interopérables.

C'est précisément l'objectif de [Metadatapp](#), une solution logicielle pensée pour structurer les métadonnées et rendre les résultats d'expériences plus exploitables dans le temps, au service de la qualité scientifique et des 3R. Pour en parler, nous avons interrogé Damien Huzard, docteur en neurosciences comportementales et fondateur du projet.

△ *Cet article ne fait l'objet d'aucune collaboration commerciale.*

### À retenir

- Un volume considérable de données précliniques est produit, mais une partie de sa valeur se perd car les informations de contexte sont incomplètes, dispersées ou introuvables ;
- En reliant et en standardisant les données et leurs métadonnées, on les rend traçables, comparables et exploitables dans le temps, ce qui permet notamment d'éviter des expériences redondantes ;
- Metadatapp connecte les outils du laboratoire, automatise la collecte des métadonnées et génère un fichier unique, interopérable, prêt à être stocké, partagé et analysé (y compris par l'IA), au service de la qualité scientifique et des 3R.



### Qui êtes-vous ? Quel est votre projet ?

**Damien Huzard** : Pendant plusieurs années, j'ai travaillé dans des laboratoires publics, en Suisse et en France, où j'étudiais le comportement des animaux pour mieux comprendre certaines maladies humaines et plus précisément les troubles du neurodéveloppement.

Dans ce contexte, j'ai été confronté à une réalité peu visible du grand public : **les expérimentations animales génèrent une immense quantité de données**, mais une partie significative de leur potentiel reste trop souvent inexploité. Non pas parce que les expériences sont mal faites, mais parce que les informations qui décrivent précisément comment ces données ont été produites sont incomplètes, dispersées, mal organisées ou simplement impossibles à retrouver.

C'est de là qu'est née Metadatapp : une solution logicielle conçue pour connecter, structurer et valoriser les résultats d'études de recherche préclinique.

Mon objectif initial était très simple : **tirer le maximum de connaissances de chaque animal utilisé, afin d'en utiliser moins, et mieux**. C'est à la fois une exigence scientifique et une responsabilité éthique, pleinement alignée avec **le principe des 3R** (Remplacer, Réduire, Raffiner). Metadatapp est une solution pragmatique et immédiate pour concrétiser ce principe éthique au quotidien dans les laboratoires.



### Qu'est-ce que Metadatapp ?

**DH** : Les **métadonnées**, ce sont les données sur les données. Par exemple pour un livre, ce sont le nom de l'auteur, le titre, le genre, etc. Pour une expérience de recherche préclinique, ce sont l'âge et le sexe de l'animal testé, son environnement de vie, les détails du protocole expérimental suivi, le type de matériel utilisé, ou encore les conditions d'hébergement (température, cycle lumineux, enrichissement, densité, etc.).

Sans ces informations, une donnée scientifique perd une grande partie de son sens. De la même manière qu'il est impossible de gérer une bibliothèque sans métadonnées sur les livres, il devient extrêmement difficile d'organiser, d'archiver et surtout de réutiliser des résultats scientifiques sans métadonnées structurées.

Quelques chiffres parlants pour décrire l'ampleur du problème :

- [Environ 70 % des chercheurs déclarent ne pas réussir à reproduire les expériences de leurs pairs](#), et plus de la moitié ne parviennent pas à reproduire leurs propres résultats (ces chiffres ne concernent pas spécifiquement la recherche animale mais la recherche en général).
- [Plus de 25% des échecs au stade clinique du développement de médicaments sont attribués à des résultats issus des phases précliniques ingérables](#).
- Chaque année, [la non-reproductibilité des études précliniques est estimée à un coût de 28 milliards de dollars aux États-Unis et environ 10 milliards d'euros en Europe](#).

Avec Metadatapp, nous construisons donc une plateforme qui aide toutes les personnes impliquées dans des projets de recherche à décrire clairement et de manière standardisée leurs expériences, afin que les résultats puissent être compris, comparés et réutilisés, même des années plus tard.

Pendant ma carrière de chercheur, j'ai effectué beaucoup d'expériences scientifiques, certaines ont été publiées et sont accessibles (ou en accès limité par les éditeurs), mais **la grande majorité des résultats, jugés "non-concluants", sont seulement stockés de façon éclatée et très difficilement accessible**. Je sais pertinemment que, dans cet état, personne ne pourra les réutiliser correctement. **C'est un énorme gâchis de temps, d'argent, et surtout un gaspillage de vies animales.**

J'ai donc conceptualisé cet outil pour simplifier le travail du chercheur : en se connectant aux outils existants (logiciels d'analyse, cahiers de laboratoire électroniques, systèmes de gestion animale), Metadatapp automatise la collecte des détails expérimentaux. L'information circule ainsi de façon horizontale, générant un fichier unique, structuré et traçable. Ce fichier peut ensuite être stocké, partagé, et compris par des systèmes automatisés, ouvrant la voie à des analyses ultérieures plus puissantes et plus responsables.

Sans ce travail préalable, même les meilleures intelligences artificielles produisent des résultats fragiles ou trompeurs.

## Quel rôle joue l'intelligence artificielle ?

**DH** : Les progrès récents en intelligence artificielle (IA) et notamment l'émergence des LLMs (Large Languages Models, tels que ChatGPT, Gemini ou Le Chat de Mistral), ont été un levier important pour Metadatapp. Ces outils sont capables d'analyser rapidement de grandes quantités d'informations, de détecter des incohérences ou d'aider à organiser des données complexes. Ils ont aussi démocratisé l'accès à l'analyse : aujourd'hui, n'importe quel chercheur peut interroger des données via un simple prompt.

Mais il y a des limites et des précautions à prendre. Un point fondamental à comprendre est qu'**un LLM fonctionnera d'autant moins bien que les données qu'on lui fournit sont mal décrites**. Une IA ne "comprend" pas une donnée floue. Les algorithmes sont beaucoup plus efficaces lorsque les informations sont bien structurées, définies de manière non ambiguë et reliées logiquement entre elles.

C'est là qu'interviennent deux autres concepts-clés : la sémantique du web et l'interopérabilité.

La sémantique du web peut se résumer très simplement : c'est une façon de décrire les données pour que les ordinateurs comprennent ce qu'elles veulent dire, et pas seulement à quoi elles ressemblent. Par exemple, si un fichier indique « 20 », un humain peut comprendre qu'il s'agit de 20 °C dans une animalerie, mesuré à un moment précis. Un ordinateur, lui, a besoin qu'on lui dise explicitement : 'Température', '20°C', 'Mesurée dans l'animalerie X', 'le 08/03/2024'. La sémantique du web permet justement d'ajouter ce contexte de manière

standardisée. On peut la voir comme une forme d'intelligence artificielle "de base", qui ne fait pas de prédictions, mais qui permet aux machines de raisonner correctement sur les données, en comprenant leurs relations et leur signification.

L'interopérabilité, c'est la conséquence directe de cela : c'est la capacité de différents logiciels, laboratoires ou pays à utiliser et comprendre les mêmes données sans ambiguïté. Sans interopérabilité, chaque système parle son propre langage, et les données restent enfermées dans des silos difficiles voire impossibles à comparer ou à combiner.

Metadatapp crée ce langage commun en s'appuyant sur des standards partagés et des vocabulaires communs. **Concrètement, cela permet à des données produites dans des contextes différents — par exemple dans deux laboratoires ou avec deux logiciels différents — d'être interprétées de la même manière et analysées ensemble.**

Bien organiser les données aujourd'hui, c'est donc les rendre réellement exploitables par l'intelligence artificielle. Dans ce cadre, l'IA agit comme un amplificateur de connaissances : elle peut analyser plus vite et plus largement, mais le résultat ne sera pertinent que si les données sont bien décrites, vérifiées et traçables.

Metadatapp intervient en amont de ces approches, en garantissant que les données analysées sont fiables, comparables et correctement contextualisées. Sans ce travail préalable, même les meilleures modèles de LLMs produisent des résultats fragiles ou trompeurs.

Cela montre que certaines méthodes alternatives permettant de réduire le nombre d'animaux utilisés, ne reposent pas nécessairement sur un nouvel appareil ou un nouveau test spectaculaire, mais parfois sur des infrastructures numériques invisibles, enrichies par l'IA, essentielles à la recherche moderne — et malheureusement encore trop difficiles à financer.

**La version prototype de Metadatapp est connectée à plusieurs applications couramment utilisées dans les animaleries de recherche :**

- [SoftMouse.NET](#)
- [ElabFTW](#)
- [Fair3R](#)
- [OSF](#)
- [preclinicaltrials.eu](#)

**Quels sont les bénéfices concrets, aujourd'hui et demain ?**

**DH** : À court terme, **Metadatapp permet aux utilisateurs d'éviter de refaire des expériences simplement parce que les données existantes sont inutilisables ou introuvables.** Des données bien organisées restent exploitables dans le temps. La plateforme améliore également l'interprétation des résultats grâce à une meilleure traçabilité, c'est-à-dire la possibilité de retracer précisément comment une donnée a été produite, ce qui renforce la qualité scientifique des études impliquant des animaux. Un scientifique peut ainsi exporter un seul fichier contenant toutes les informations relatives aux expériences présentées dans un article de recherche et le joindre à la publication.

À long terme, notre ambition est plus globale : **permettre la mutualisation des données entre laboratoires, des analyses plus robustes, et le développement de modèles prédictifs soutenant la réduction de l'engagement animal en recherche, comme les groupes contrôles virtuels ou les Nouvelles Approches Méthodologiques (NAM).**

Un exemple concret est celui des groupes contrôles virtuels. Lorsqu'un grand volume de données bien contrôlées et structurées est collecté de façon routinière dans plusieurs centres, il devient possible de comparer une nouvelle expérience à des résultats préexistants très proches en termes de contexte expérimental. Dans certains cas, cela peut éviter la création de nouveaux groupes animaux.

Cela agit directement sur les 3R :

- **Remplacement** : données suffisamment riches pour tester certaines hypothèses sans relancer systématiquement de nouvelles expérimentations animales.

- **Réduction** : moins d'animaux nécessaires grâce à la réutilisation de données existantes.
- **Raffinement** : la plateforme améliore également l'interprétation des résultats grâce à une meilleure traçabilité, permettant de retracer précisément comment une donnée a été produite et donc d'optimiser les protocoles.

Mettre en valeur les bonnes pratiques de gestion des données aurait des retombées très positives sur la reconnaissance du travail réalisé dans les laboratoires.

### Quelles difficultés avez-vous rencontrées ?

**DH** : L'une des principales difficultés est culturelle. **La gestion des (méta)données est encore souvent perçue comme une contrainte administrative supplémentaire**, alors qu'elle constitue en réalité **un levier majeur de progrès scientifique et éthique**.

Il existe aussi un problème lié à la façon dont les projets de recherche sont évalués et récompensés. **Le système actuel valorise avant tout les résultats "spectaculaires" publiés dans des revues prestigieuses, bien plus que la qualité du travail méthodologique, de documentation et de gestion des données sous-jacentes.**

Il est temps de changer cette tendance. C'est pourquoi le soutien d'acteurs comme le Gircor est essentiel pour faire reconnaître la bonne gestion des données comme une exigence scientifique de premier plan, au même titre qu'une publication prestigieuse. **Mettre en valeur les bonnes pratiques de gestion des données aurait des retombées très positives sur la reconnaissance du travail réalisé dans les laboratoires et sur la robustesse globale de la recherche.**

Le défi de Metadatapp est donc de proposer des outils simples, utiles et acceptables pour les chercheurs, sans alourdir leur charge de travail quotidienne.

### De quoi avez-vous besoin aujourd'hui pour continuer ?

**DH** : Pour poursuivre le développement de Metadatapp, trois types de soutien sont essentiels.

Tout d'abord, des financements, afin de consolider le développement de l'outil, de le tester à grande échelle et d'en garantir la pérennité.

Ensuite, **nous recherchons activement des partenariats** pour des projets pilotes avec des plateformes animales et des institutions engagées dans les 3R, afin de prouver l'impact concret de notre solution sur la réduction du nombre d'animaux utilisés. Aussi nous planifions de collaborer avec des acteurs du numérique scientifique et de la sémantique du web, afin d'ancrer l'outil dans la réalité du terrain et d'en renforcer la robustesse technologique.

Enfin, une reconnaissance institutionnelle, gouvernementale et réglementaire est primordiale pour faire reconnaître que mieux gérer et mieux exploiter les données constitue une méthode alternative à part entière, contribuant directement à l'innovation médicale, au même titre que de nouveaux équipements ou de nouveaux modèles expérimentaux.

**Metadatapp travaille à un partenariat avec Tecniplast, un constructeur de cages, afin d'utiliser les données récoltées par ces dernières.**

**D'autres partenariats se créeront aussi avec des logiciels de gestion d'animaleries de recherche et avec l'IGF de Montpellier.**

---

Cet article a été téléchargé depuis [gircor.fr](https://www.gircor.fr)