



Analyse de l'enquête sur les cahiers de laboratoire électroniques au CNRS

Nathalie Léon – Domenico Libri

Avec le soutien des membres du groupe de travail dédié

Le cahier de laboratoire doit permettre à tous ceux qui réalisent des travaux de recherche (chercheurs, ingénieurs, techniciens, thésards, stagiaires...) de consigner au jour le jour le détail de leurs travaux, de rendre compte du cheminement et de l'expérimentation scientifique, de l'idée à la conclusion. Outil de transmission du savoir, il s'avère également très utile pour la rédaction d'un brevet ou pour prouver une antériorité.

Elaboré par le MESR (Ministère de l'Enseignement Supérieur et de la Recherche), en collaboration avec l'INPI (Institut National de la Propriété Intellectuelle) et en concertation avec les organismes de recherche publics, le cahier de laboratoire dit « national » permet à tous ceux qui réalisent des travaux de recherche (chercheurs, ingénieurs, techniciens, thésards, stagiaires...) de consigner au jour le jour le détail de leurs travaux. Il constitue en ce sens un véritable outil scientifique répondant également aux obligations légales et contractuelles en apportant la preuve de l'invention et de ses inventeurs.

Toutefois le cahier de laboratoire « national » est un outil papier qui apparaît de moins en moins adapté aux pratiques de la recherche compte tenu de l'évolution de la nature et du volume des données produites. Les apports du numérique sont en effet multiples en améliorant la traçabilité des recherches, la lutte contre la fraude et la gestion des données à travers par exemple, le partage de l'information avec un rattachement des données brutes, une recherche d'informations facilitée et une datation des expériences par l'horodatage.

Aussi, le CNRS a lancé une réflexion sur la mise en place de cahiers de laboratoires électroniques suite aux besoins remontés par les agents en laboratoire en alternative au cahier de laboratoire national (format papier).

Comme le cahier de laboratoire national, les solutions électroniques doivent également permettre l'identification des contributions intellectuelles et techniques des agents et équipes de recherche à l'obtention des résultats issus des travaux internes et collaboratifs des unités.

L'objectif de cette réflexion est d'identifier des solutions électroniques adaptées aux pratiques de la recherche et prenant en compte la diversité de nos activités tout en assurant la mémoire, la traçabilité, la sécurité, la confidentialité et la pérennité des résultats de recherche et en respectant les exigences liées à la Science Ouverte et à la protection du patrimoine scientifique et technique de l'établissement.

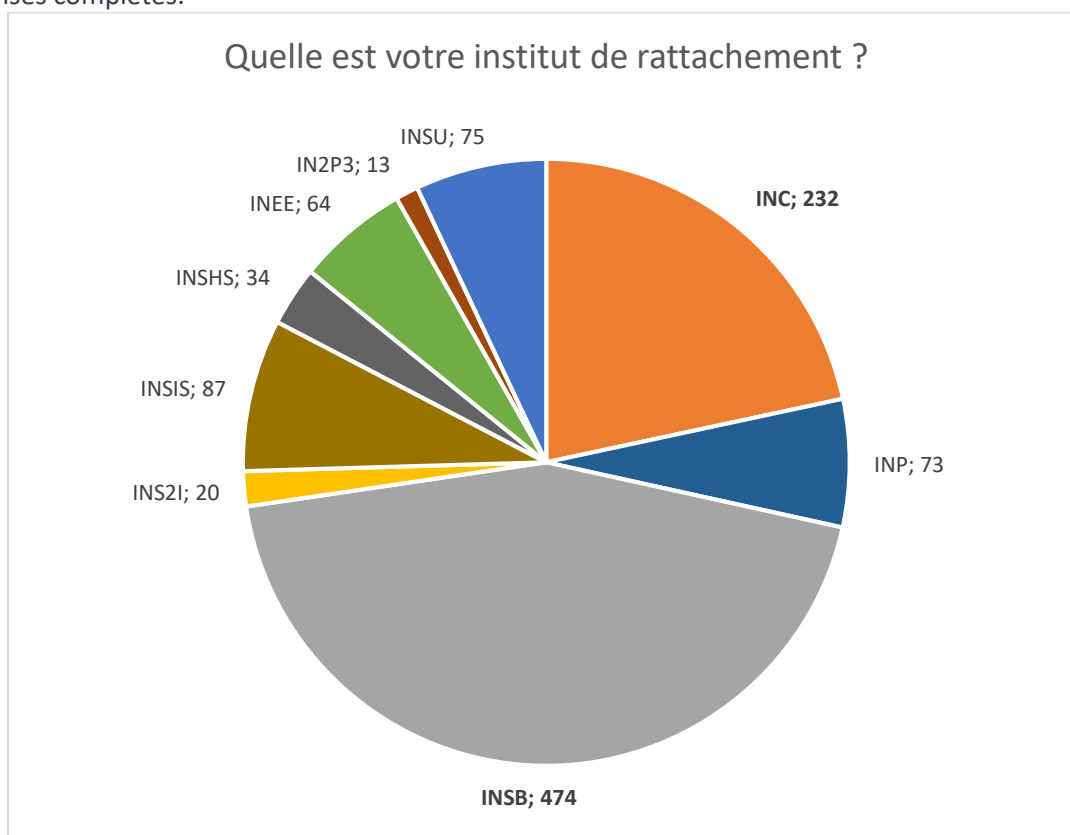
A la demande conjointe du DGD-S et du DGD-R, un groupe de travail réunissant les expertises ci-dessous a été mandaté pour se pencher sur cette réflexion :

- Scientifique (Instituts)
- Juridique et archivage (DAJ)
- De la qualité (MPR & réseau QeR)
- Des systèmes d'informations (DSI)
- De la direction des données ouvertes de la recherche (DDOR) avec les correspondants données des instituts et les correspondants Information Scientifique et Technique (IST) des instituts (CORIST)
- De la sécurité (DIRSU/FSD & RSSI)
- De la protection des données (DPD)

La première action de ce groupe a été de déployer une enquête ayant les objectifs suivants :

1. Informer la communauté de la réflexion en cours.
2. Inventorier :
 - le nombre de personnes potentiellement intéressées par les versions électroniques des cahiers de laboratoire,
 - les solutions déjà utilisées et leur nombre d'utilisateurs,,
 - les fonctionnalités souhaitées, la variété et la volumétrie de données et des résultats de recherche à stocker ,
 - les besoins d'encadrement, de formation et de support,
 - les point(s) éventuel(s) d'inquiétude(s) ou de vigilance quant à l'utilisation d'un tel outil.

Cette enquête de 19 questions a été transmise par le DGD-S aux directeurs d'Instituts le 25 juin 2020 pour diffusion aux unités. L'enquête a été clôturée le 4 septembre 2020. Elle comptabilise 1072 réponses complètes.



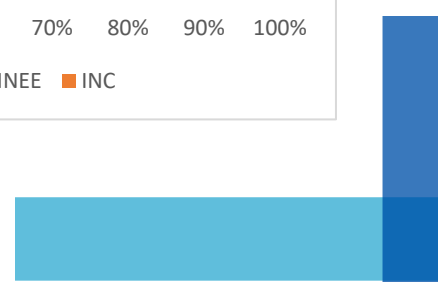
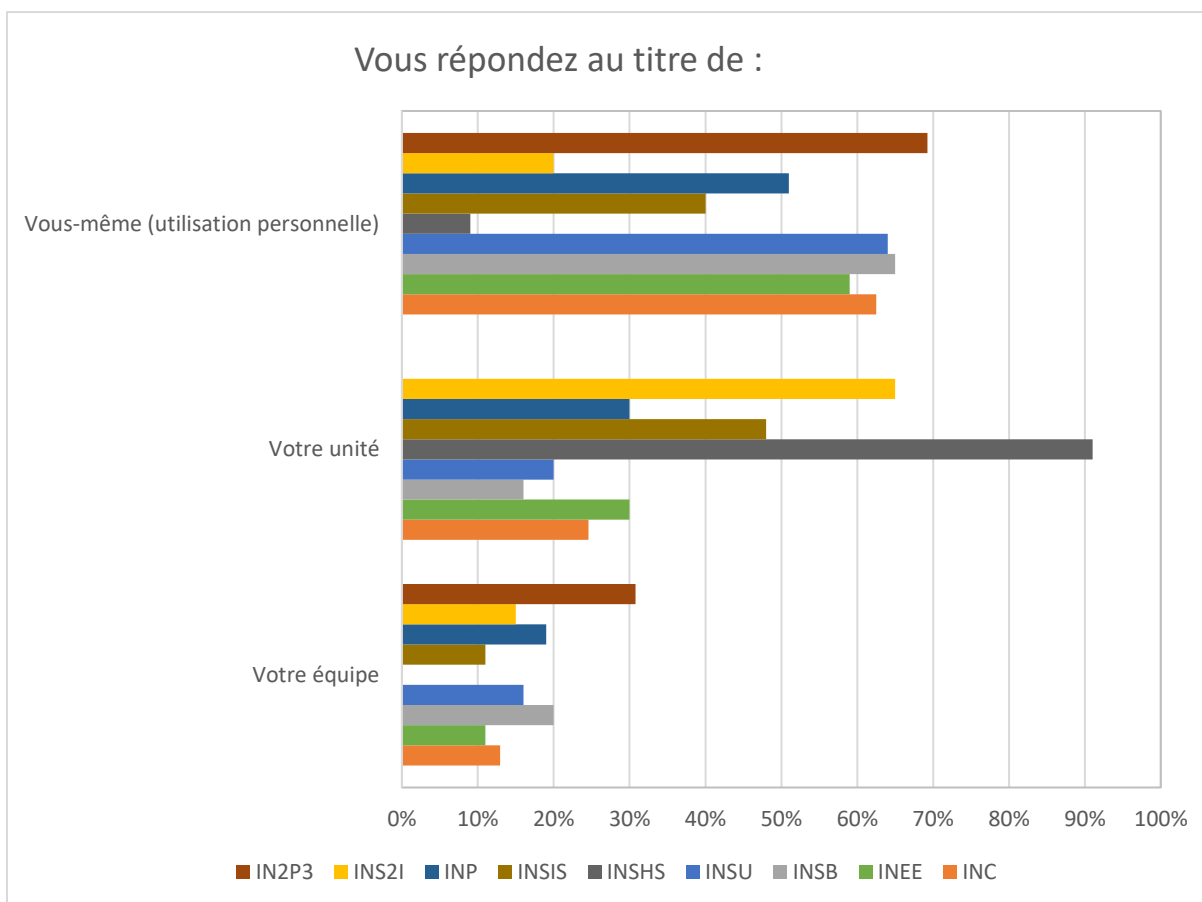
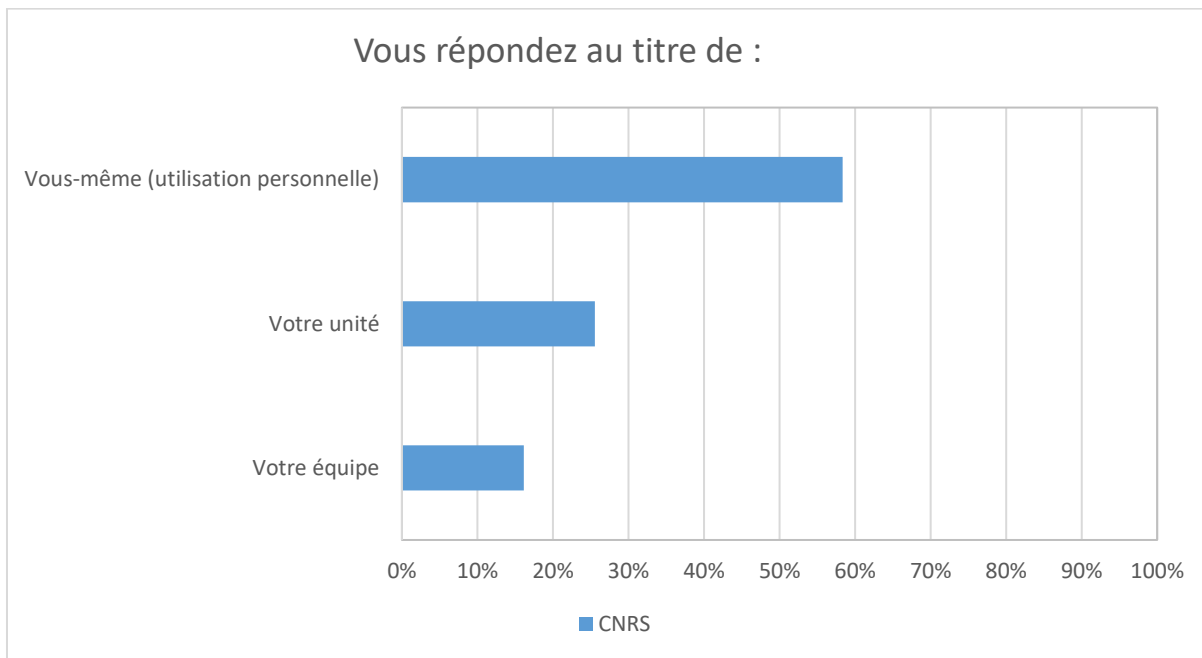
Les deux instituts les plus représentés sont l'INSB et l'INC. Treize réponses d'une même entité de l'IN2P3 ont été enregistrées. Elles sont à interpréter au titre de l'INSB du fait de leurs domaines de recherche.

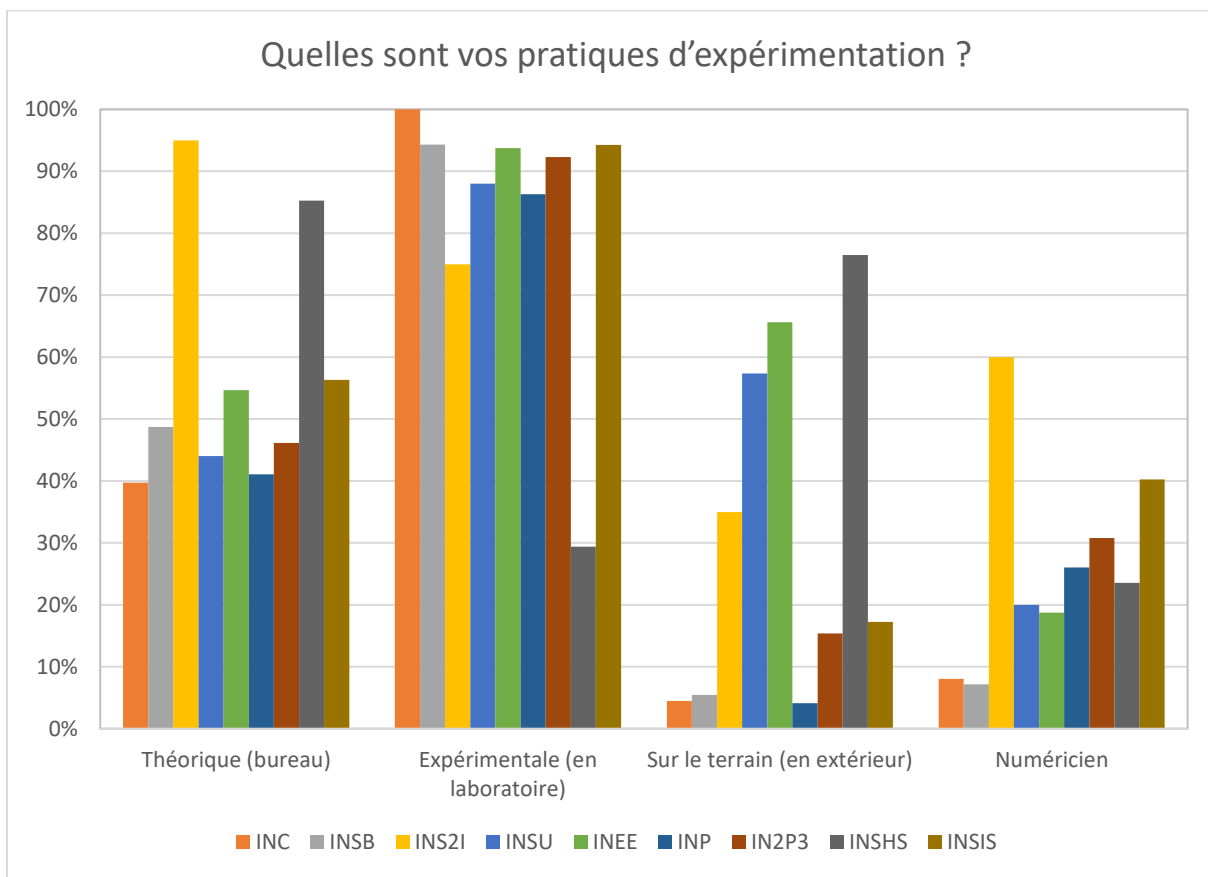
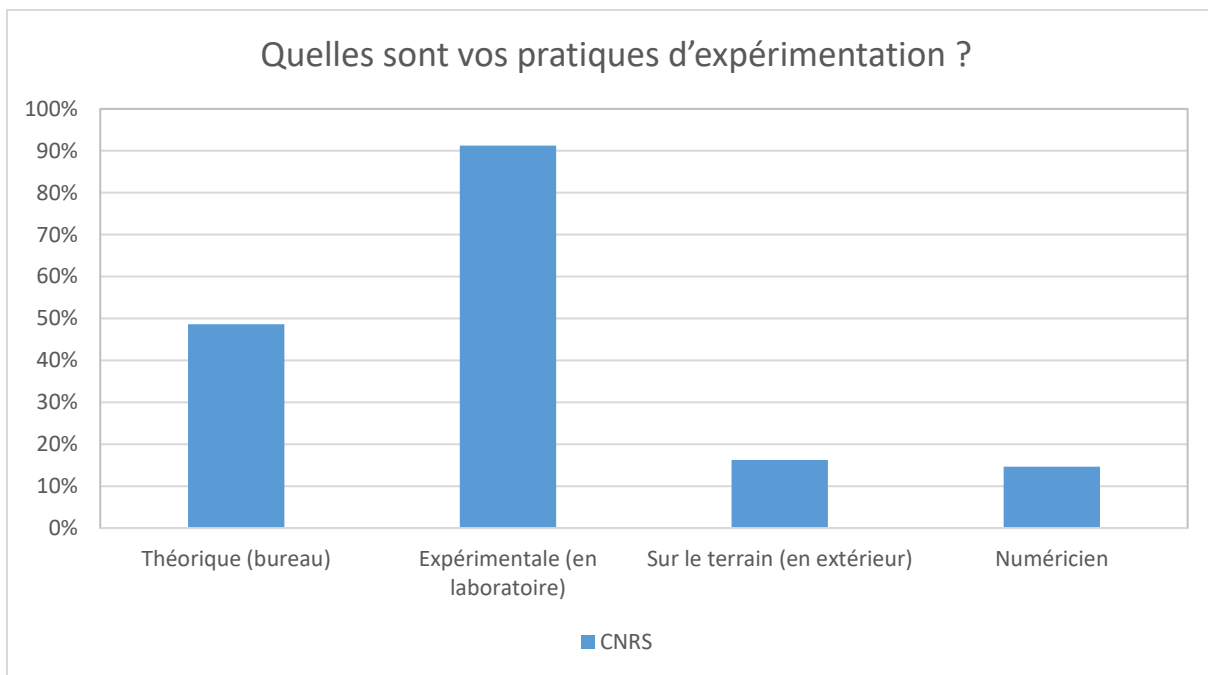
Les réponses issues de 432 unités de recherche ont été enregistrées soit 38% des 1143 unités de recherche et unités de service. 258 réponses l'ont été au nom du DU, soit 23% des 1143 unités de recherche et unités de service.

La plupart des répondants ont indiqué avoir d'autres tutelles que le CNRS, telles que :

- Les universités et écoles,
- INSERM,
- INRAe,
- CEA,
- IRD.

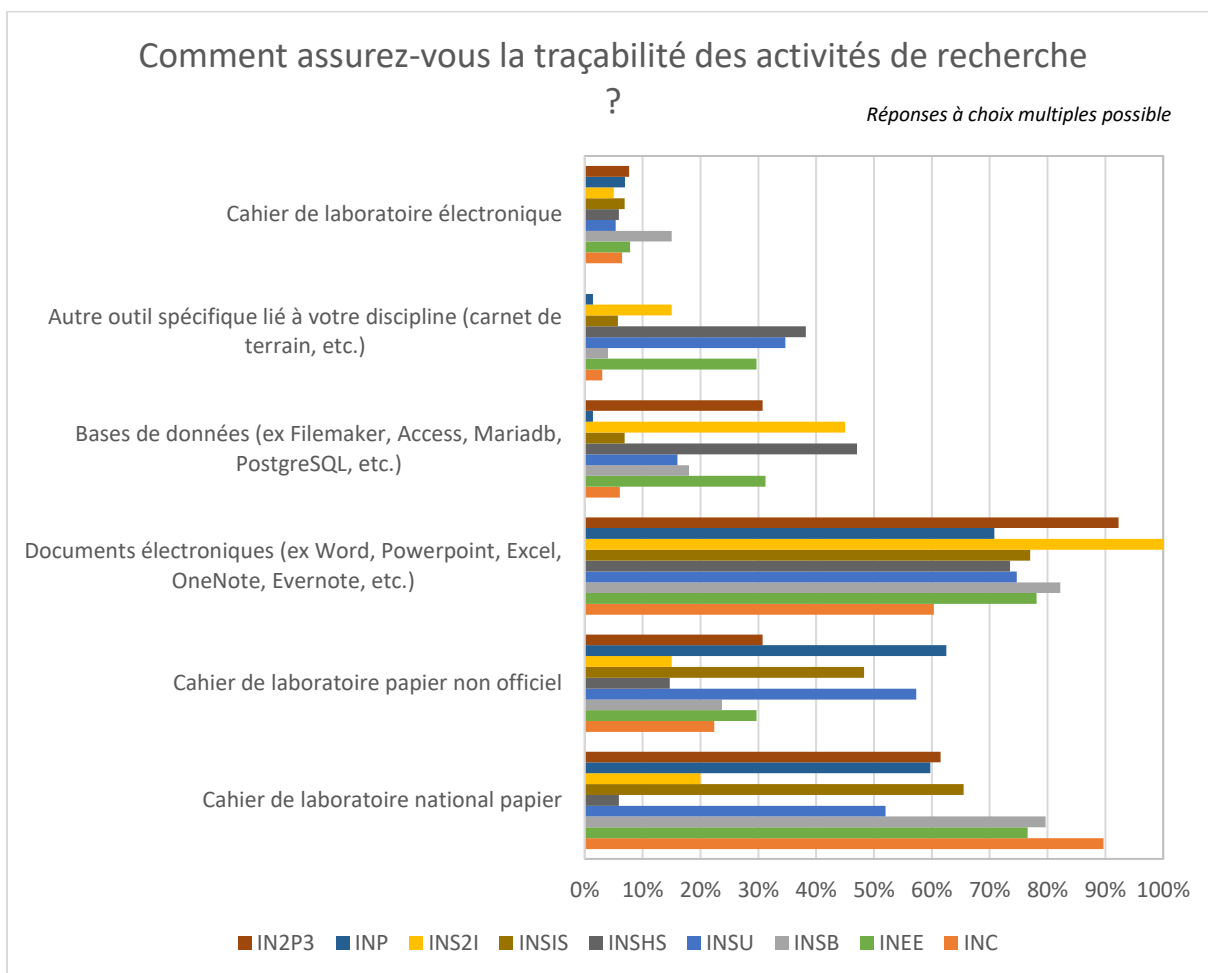
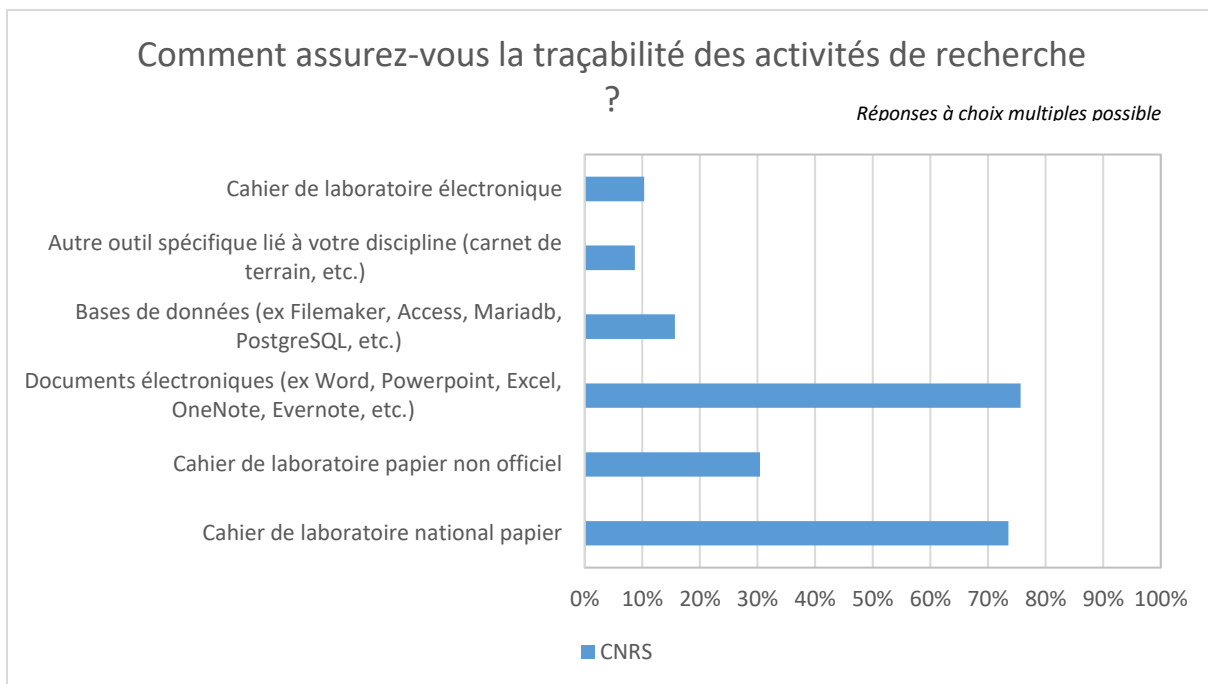
En réponse aux questions posées, il est estimé qu'environ 9000 équipes de recherche sont identifiées dans l'ensemble des structures (UMR, UPR, etc.) et concernent entre 90 000 et 100 000 personnes (chercheurs, enseignants-chercheurs, IT, doctorants, titulaires et contractuels) susceptibles d'utiliser un cahier de laboratoire électronique. La plupart des répondants ont répondu à titre individuel ou de leur unité.





Les pratiques de recherche des répondants sont principalement expérimentales (en laboratoire) ou théoriques (au bureau) avec les personnes dépendant de l'INSHS, l'INEE et l'INSU exerçant en grande partie sur le terrain.



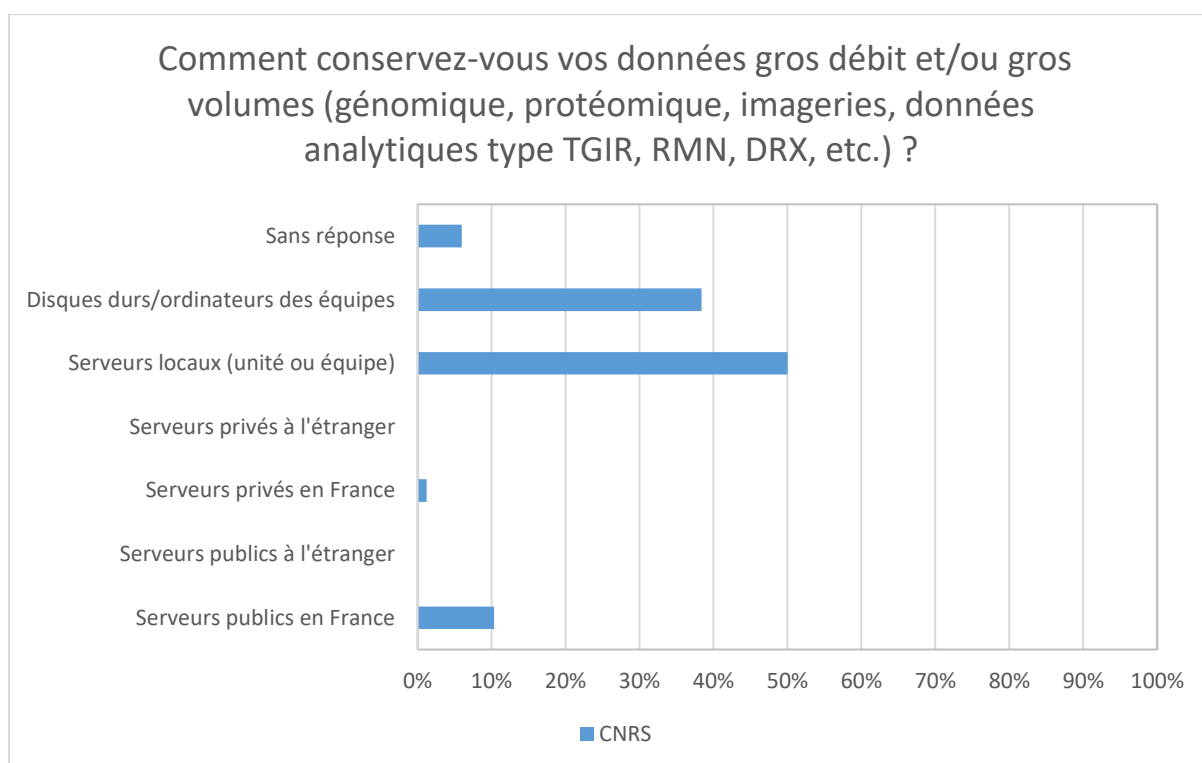


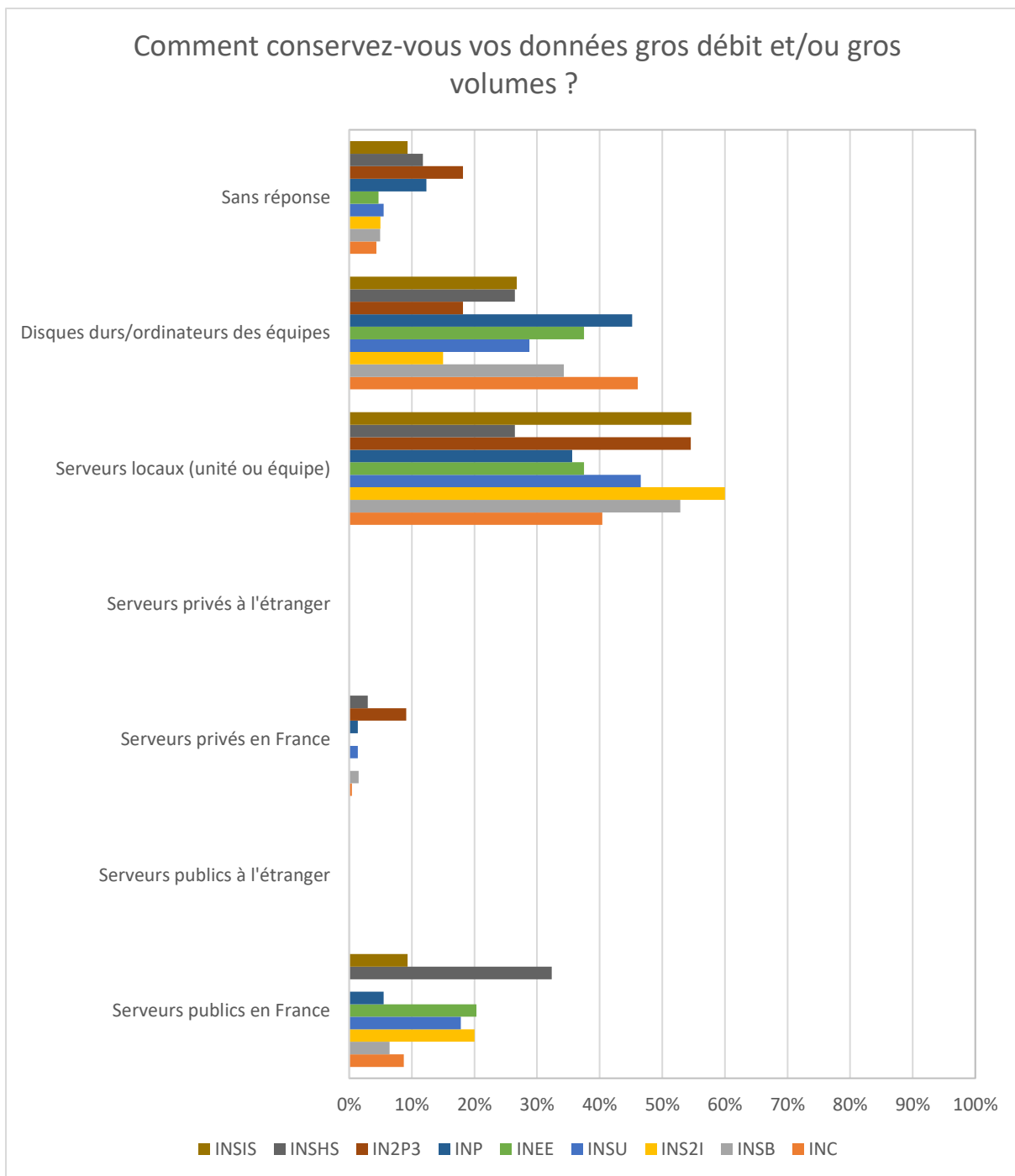
Le cahier de laboratoire papier « national » est fortement utilisé (74%) ainsi que les documents bureautiques (76%) tels que Word, Excel, PowerPoint, OneNote etc.

10% des répondants utilisent des cahiers de laboratoires électroniques comme :

- eNovalys (outil commercial utilisé par des unités à la DR10),
- eLabFTW (outil Open Source),
- Scinote,
- Chemotion (outil Open Source),
- Labguru (outil déployé par l'INSERM),
- LabCollector (outil partiellement utilisé par le CEA),
- Benchling,
- Findings,
- BIOVIA,
- Labstep,
- ISIA (consortium AnaEE France).

De plus, des carnets de terrains et fiches d'observations papiers, des logiciels de gestion de versions décentralisés (du type git) ainsi que des logiciels interactifs (Notebook, etc.) ou de gestion de projet sont utilisés.



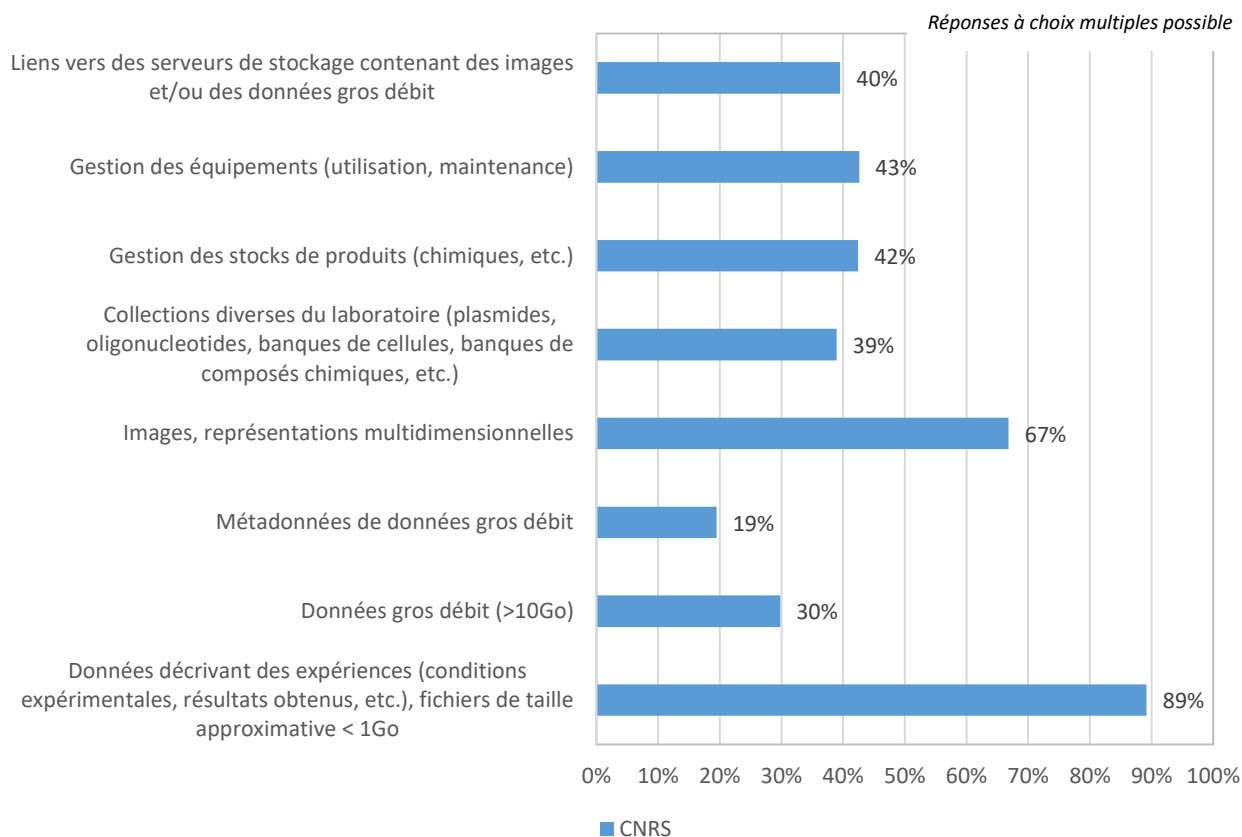


Les données présentant des volumes importants sont majoritairement conservées sur des serveurs locaux d'unité ou d'équipe (50%).

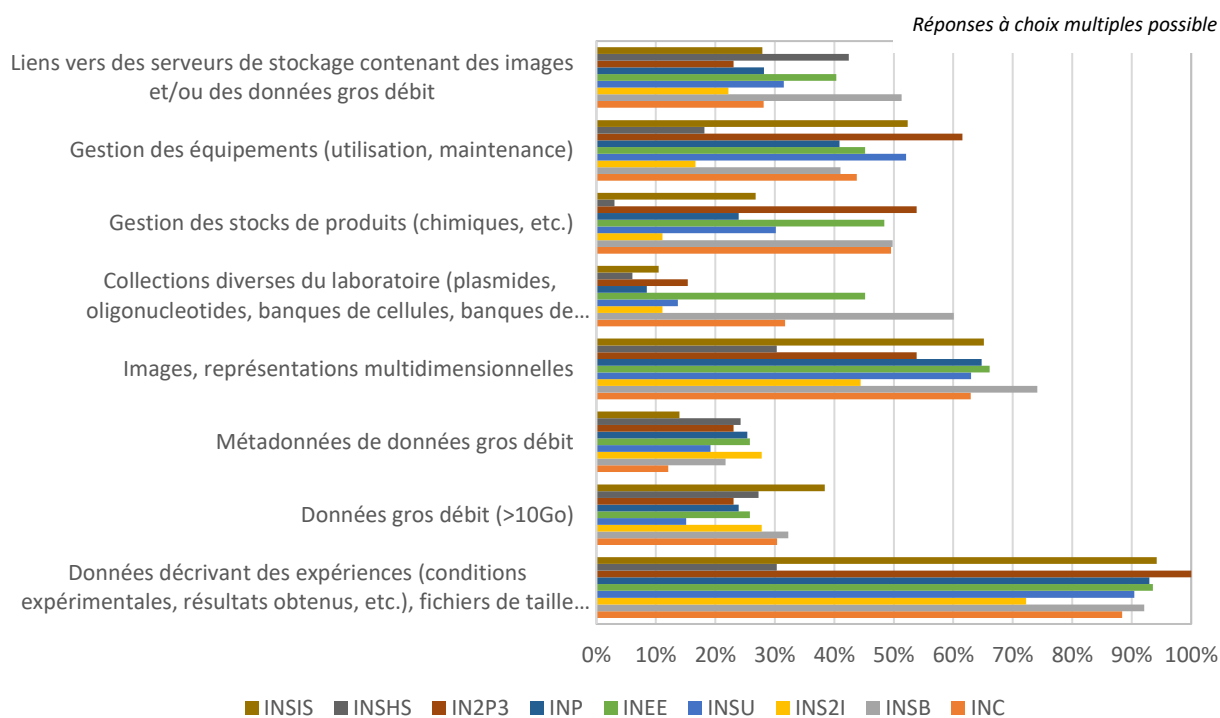
Malgré les consignes du CNRS en la matière, on note que 38% des répondants stockent leurs données sur des disques durs et ordinateurs d'équipe, ne correspondant pas de ce fait à une bonne gestion de leurs données.



Quel(s) type(s) de données répertoriez-vous (ou souhaiteriez-vous répertorier) dans un cahier de laboratoire électronique ?

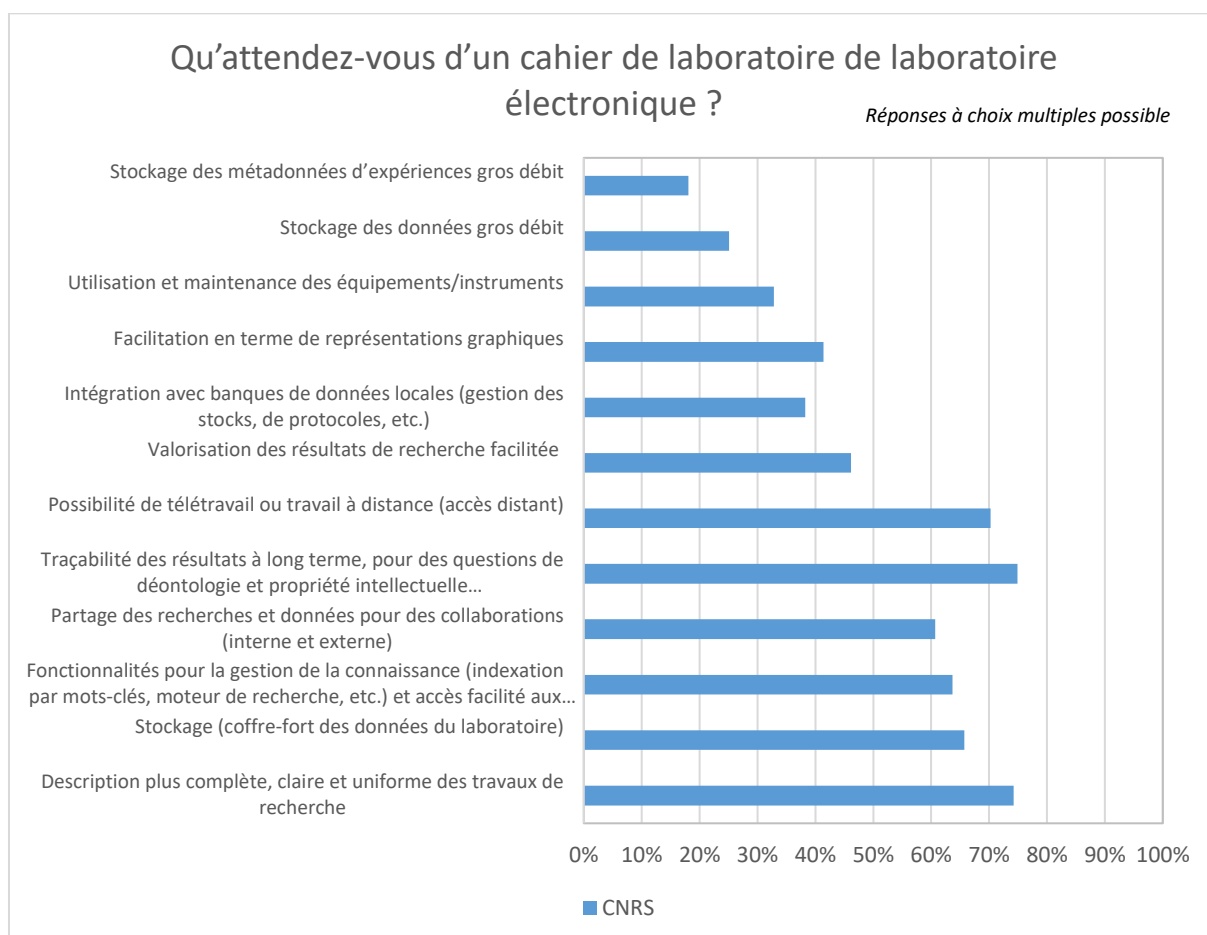


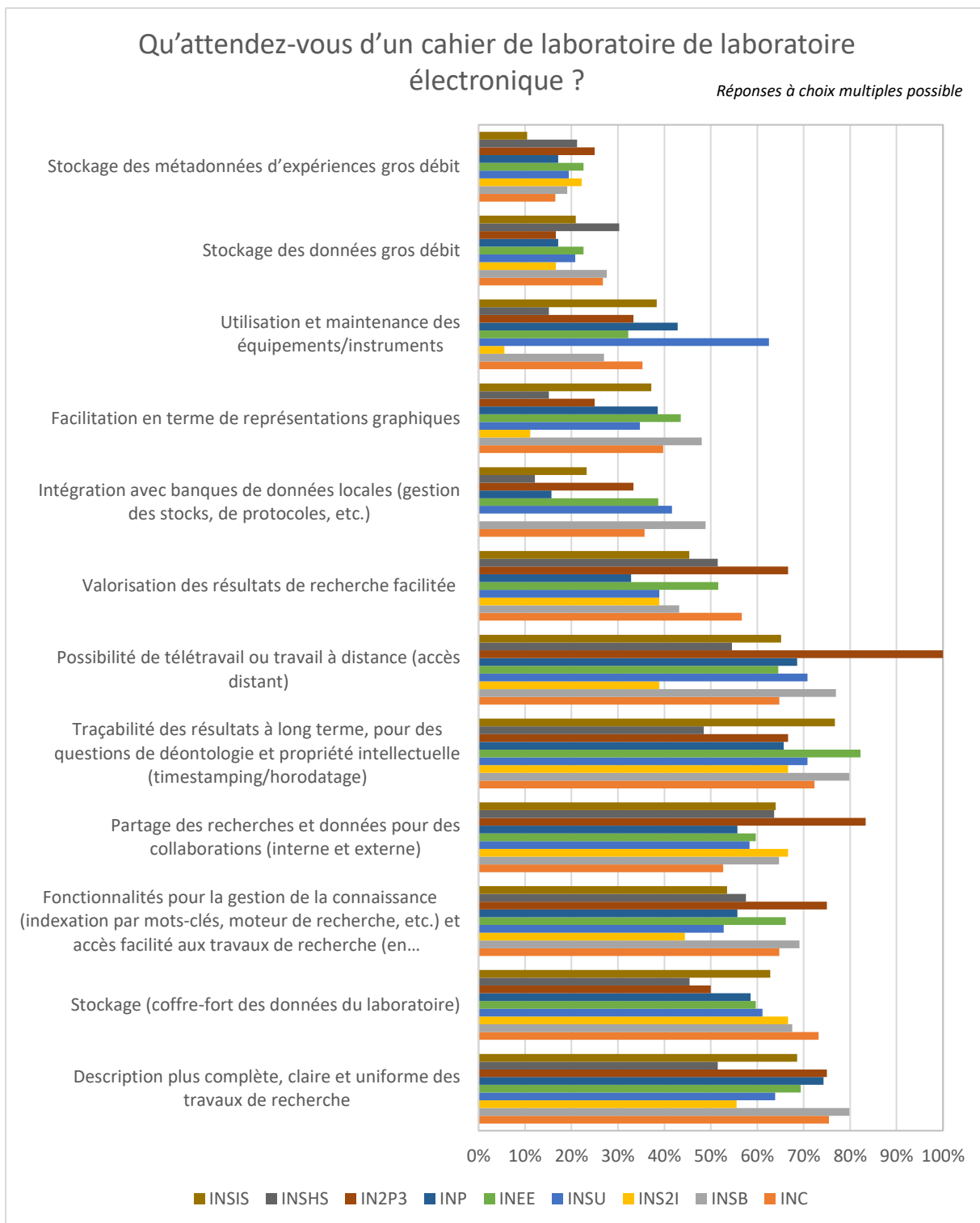
Quel(s) type(s) de données répertoriez-vous (ou souhaiteriez-vous répertorier) dans un cahier de laboratoire électronique ?



Les données répertoriées ou que les répondants souhaiteraient répertorier sont préférentiellement liées aux données décrivant des expériences (89%) ainsi qu’aux images et représentations tridimensionnelles (67%).

Autres types de données mentionnées (une ou plusieurs fois, par institut)									
Institut	INSIS	INEE	INSHS	INS2I	IN2P3	INC	INSU	INP	INSB
Protocoles (expérimentaux, procédures, etc.)		X	X				X	X	X
Notes et activités quotidiennes (réunions, annotations, schémas ou mentions manuscrites)	X	X	X				X		X
Données (tout autre format numérique : PDF, EXCEL, PPT, CAO, spectres, chromatogrammes, petites et moyennes données, etc.)	X	X	X	X		X	X	X	X
Code (calculs, formules, scripts, équations mathématiques, paramétrage, etc.)	X			X		X			X
Collection de lignées, base de données, donnée codes de simulation, métadonnées, stocks de réactifs, réservation des équipements commun, toxicité des composés / FDS	X	X	X	X		X	X	X	
Autres (vidéos, audio, etc.)									X



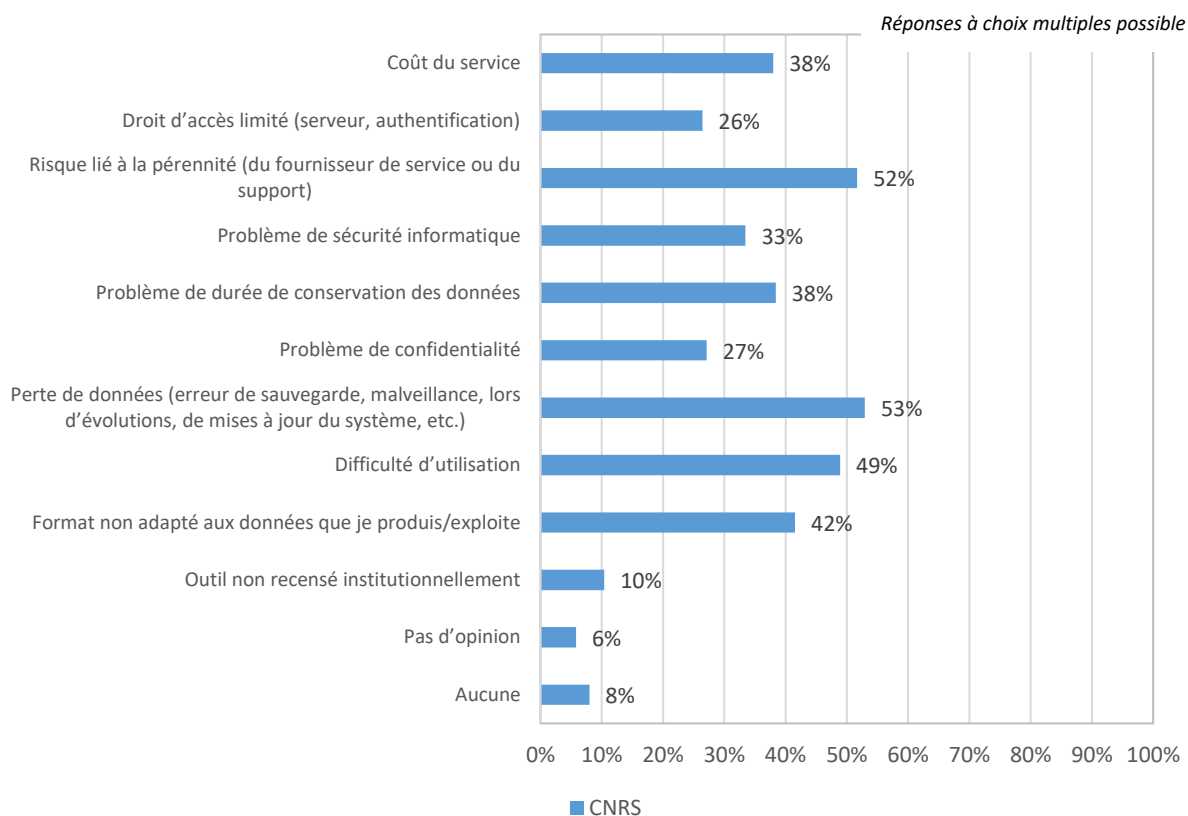


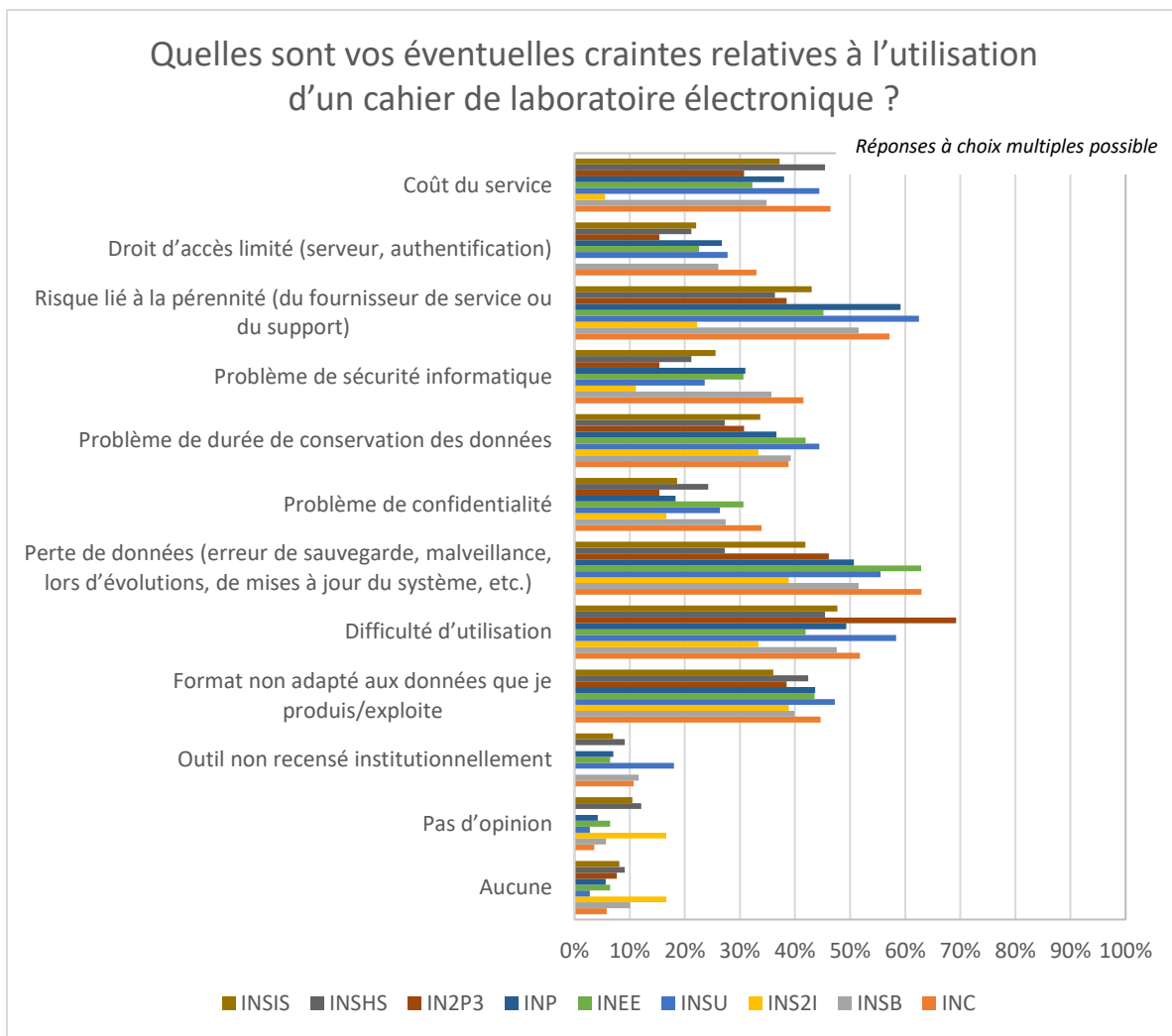
Les répondants attendent que les cahiers de laboratoires électroniques permettent une traçabilité des résultats à long terme, pour des questions de déontologie et de propriété intellectuelle (timestamping/horodatage) (75%) et qu'ils permettent le stockage (coffre-fort des données du laboratoire) (66%). Ces outils devront permettre une description plus complète, claire et uniforme des travaux de recherche (74%) et favoriser les possibilités de télétravail ou de travail à distance (70%).

De manière générale, le cahier de laboratoire électronique doit permettre un gain de temps dans la formalisation des résultats de recherche.

Autres attentes mentionnées (une ou plusieurs fois, par institut)									
Institut	INSIS	INEE	INP	INSHS	INS2I	IN2P3	INC	INSU	INSB
Ré-exploitation données	x	x							x
"page blanche", pouvoir prendre des notes à la volée et disposer d'un outil toujours disponible et le plus simple possible (type Noteshelf)		x	x					x	x
Support de rédaction de publication multimédia				x		x		x	
Interopérabilité avec standard métadonnées ou données bibliographiques (type support Labtex.bib Tex)	x						x	x	
Meilleure organisation des données : fonctionnalités de recherche (par structure chimique par exemple), classement, centralisation des informations							x		x
Autres : certification électronique, preuve d'existence, partage des protocoles, sauvegarde et extraction possible, calcul, mesure de l'empreinte environnementale, autre langue, utilisable sur tablette									

Quelles sont vos éventuelles craintes relatives à l'utilisation d'un cahier de laboratoire électronique ?



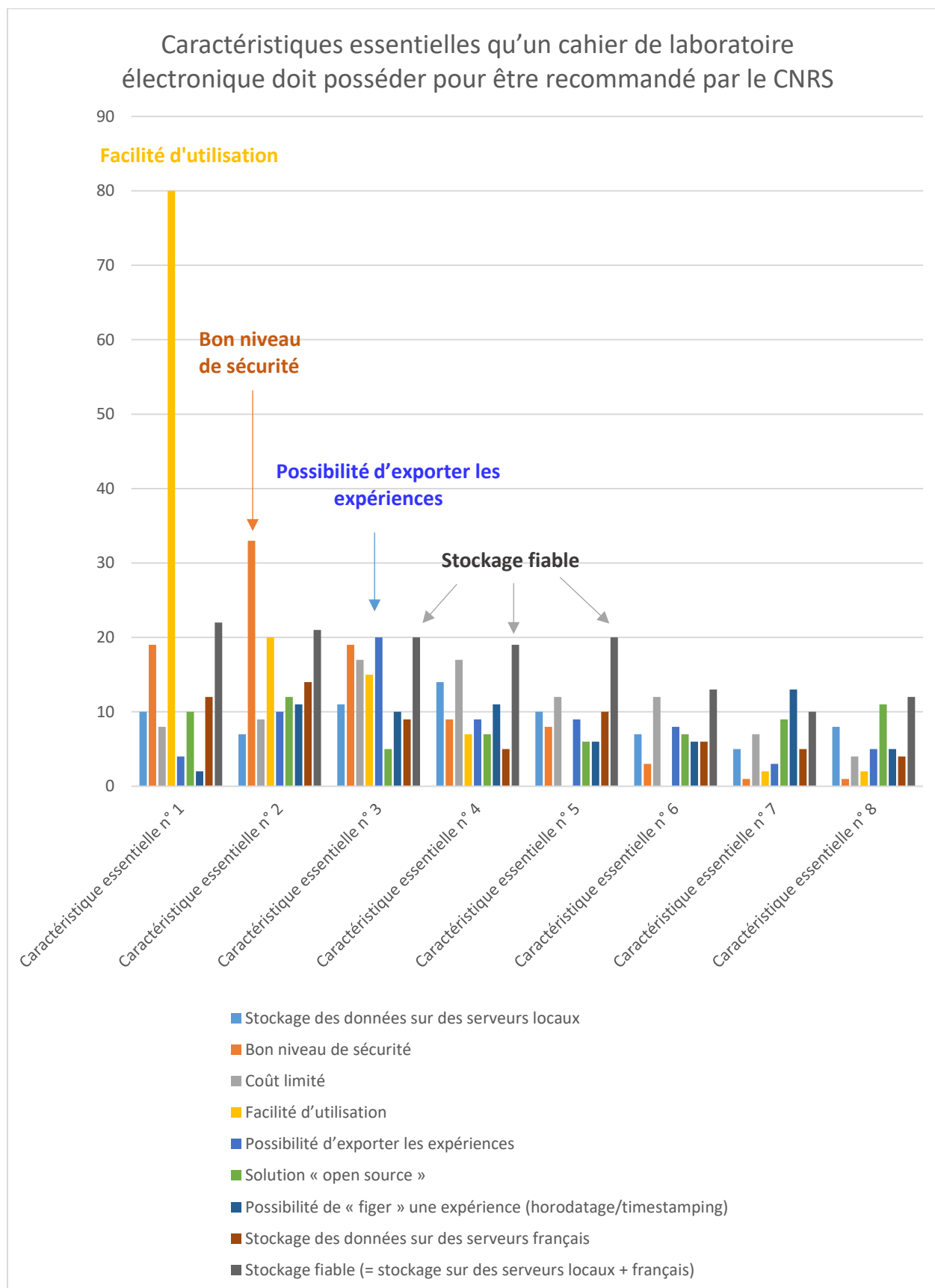


Les répondants craignent la perte de données (53%) et le risque lié à la pérennité (du fournisseur de service ou du support) (52%) ainsi que des difficultés d'utilisation (49%).

Autres craintes mentionnées (une ou plusieurs fois, par institut)									
Institut	INSIS	INEE	INP	INSHS	INS2I	IN2P3	INC	INSU	INSB
Surplus administratif, lourdeur, procédures, "usine à gaz "	x	x	x	x			x	x	x
Outil sans valeur juridique (en cas de problème de PI)									x
Compatibilité avec documents PDF, ou autres sources	x								
Non adapté aux pratiques expérimentales (paillasse, terrain, etc.), dépendance support électronique (physique, batteries, électricité, serveur, réseau)	x	x	x				x	x	x

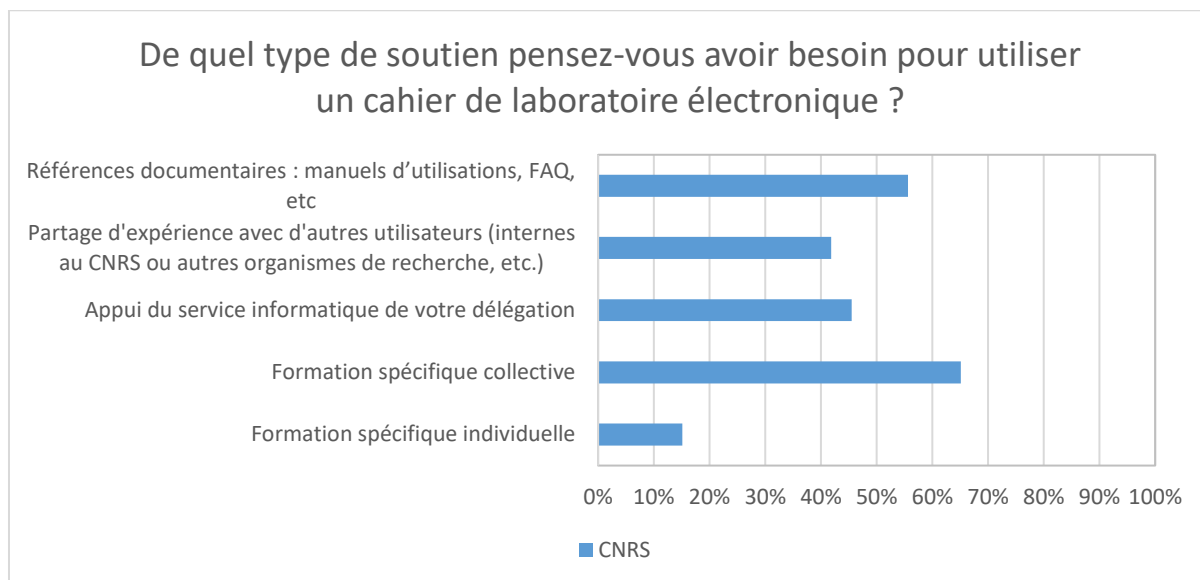
Traçabilité = pas de modification invisible a posteriori		x							
Pas de rapidité pour faire des schémas			x					x	
Manque d'interopérabilité			x						
Outil non reconnu par les partenaires (autres tutelles, universités, etc.) et/ou non homogène				x			x		x
Problème de droit pour les images par exemple				x					
Non accessibilité pour partenaire étranger ou non-académiques et quid de la gestion des accès quand départ d'une unité ou fermeture d'unité (accès aux résultats ?)				x			x		x
Besoin supplémentaires de serveurs locaux							x		
Dévoisement vers une surveillance de la "productivité" du personnel, Formatage de la pensée des chercheurs par la définition a priori des champs à renseigner dans le CLE : perte d'inventivité, Perte de communication entre les personnes		x					x		
Erreurs liées au copier-coller par oubli de modifier certaines valeurs							x		
Impact environnemental							x	x	
Outil peu convivial, non ergonomique, non adapté aux tablettes et applications							x		x
Logiciel propriétaire (code fermé, outil non évolutif et non modulaire ou personnalisable)							x		x
Pour non francophone : gestion de plusieurs langues									x
Pas d'authentification moderne (CAS, Janus, LDAP, AD, two factor)							x		
Inscrire que les expériences réussies et non pas décrire les problèmes ou manipulations "ratées"									x

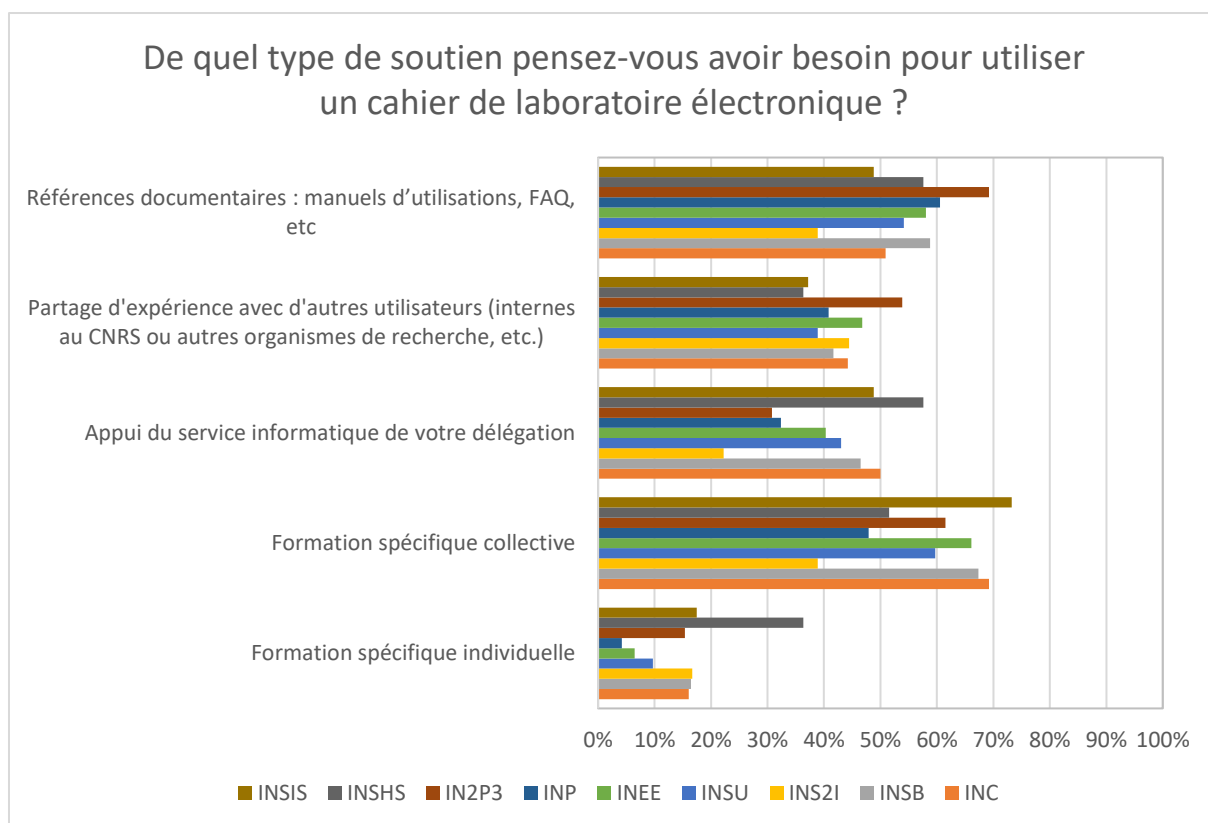
Compte tenu d'un mauvais fonctionnement informatique de la fonctionnalité de tri de l'outil de sondage, les statistiques issues des réponses de cette question sont basées sur les **176 classements** ayant été enregistrés. En voici les résultats :



Autres caractéristiques essentielles citées (une ou plusieurs fois, par institut)									
Institut	INP	INEE	INSHS	INS2I	IN2P3	INC	INSU	INSIS	INSB
Compatibilité / interopérabilité	x	x	x			x			
Fonctionnalité interface avec analyse des données ou autres outils	x		x			x			
Pérennité des résultats	x		x		x	x			
Logiciel français	x		x						x
Partage / accès distants / gestion accès	x	x	x			x	x		
Insertion prises de notes manuscrites (schémas, etc.)	x		x			x	x		
Citabilité des contenus			x	x					
Réplication des travaux empiriques			x						
Même système pour tous les EPST						x			
Accès hors ligne						x			
Suivi d'équipement		x				x			
Modularité/personnalisation/flexibilité		x				x	x	x	x

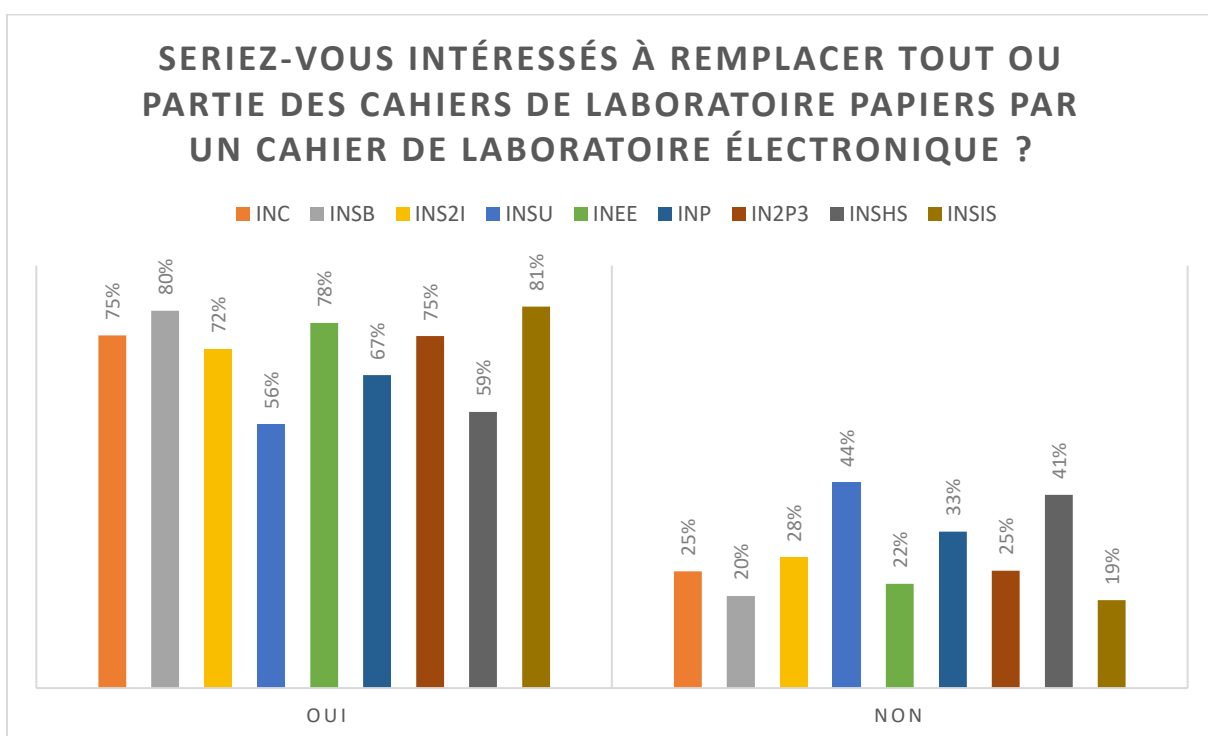
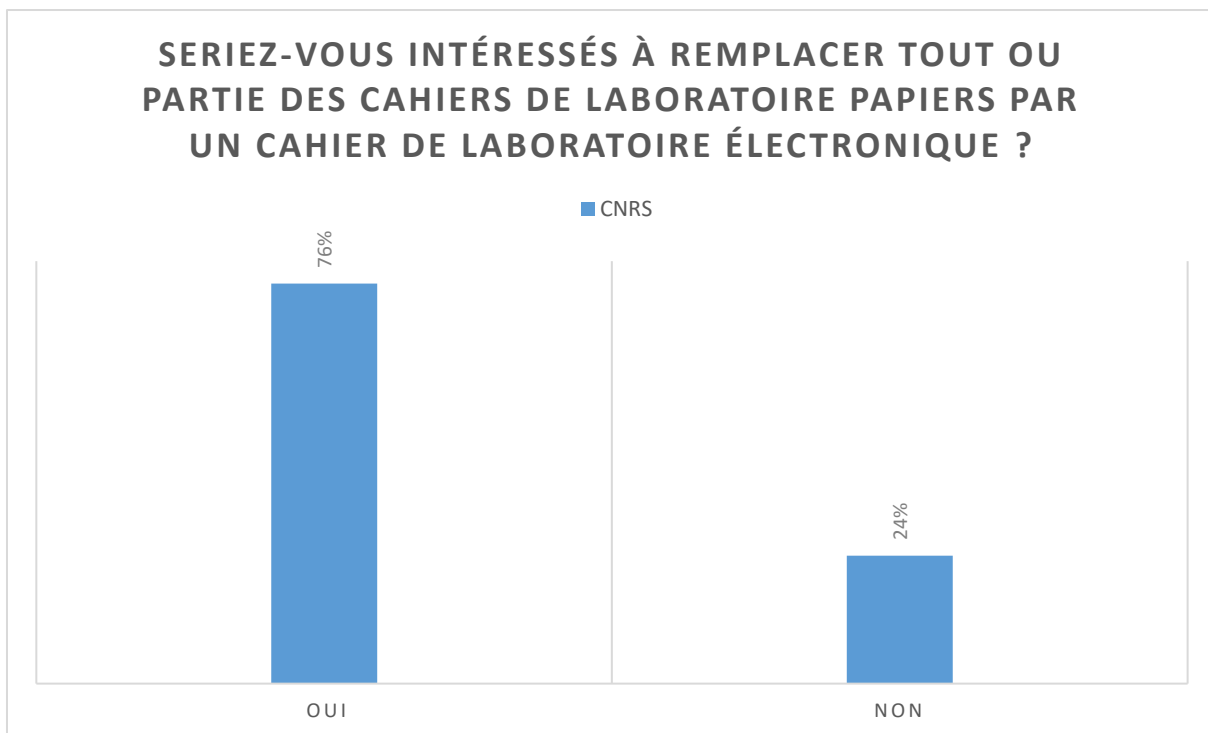
Globalement, les répondants souhaitent que le ou les outils traitent les données selon les principes FAIR (Findability, Accessible, Interoperable, Reusable).





Les formations collectives (65%) et la mise à disposition de documentation (56%) sont à privilégier selon les répondants. On note toutefois dans les réponses le souhait d'un outil facile d'utilisation et ergonomique permettant de limiter le besoin de soutien.

Autres soutiens mentionnés (une ou plusieurs fois, par institut)								
Institut	INSIS	INEE	INP	INSHS	INS2I	IN2P3	INC	INSU
Dépannage	x	x				/		x
Support	x	x				/		x
Ergonomie	x	x		x		/		
Compatibilité (.tex. python, ..)	x					/		x
Outil intuitif sans formation							x	
Tutoriels vidéos / MOOC pour validation du savoir-faire des étudiants(es).							x	



En complément, il a été demandé aux répondants d'indiquer les raisons pour lesquelles ils serait intéressant de remplacer tout ou partie du cahier de laboratoire papier par un cahier de laboratoire électronique, et a contrario, pour quelles raisons ceci aurait peu ou difficilement d'intérêt.

De façon générale, selon l'environnement de travail : salle blanche, salle de manipulation (présence de poussières, etc.), laboratoire de chimie (environnement acide, etc.), sur le terrain, il apparaît que l'utilisation de l'un ou l'autre des outils, qu'il soit papier ou informatique, peut ne pas être possible.

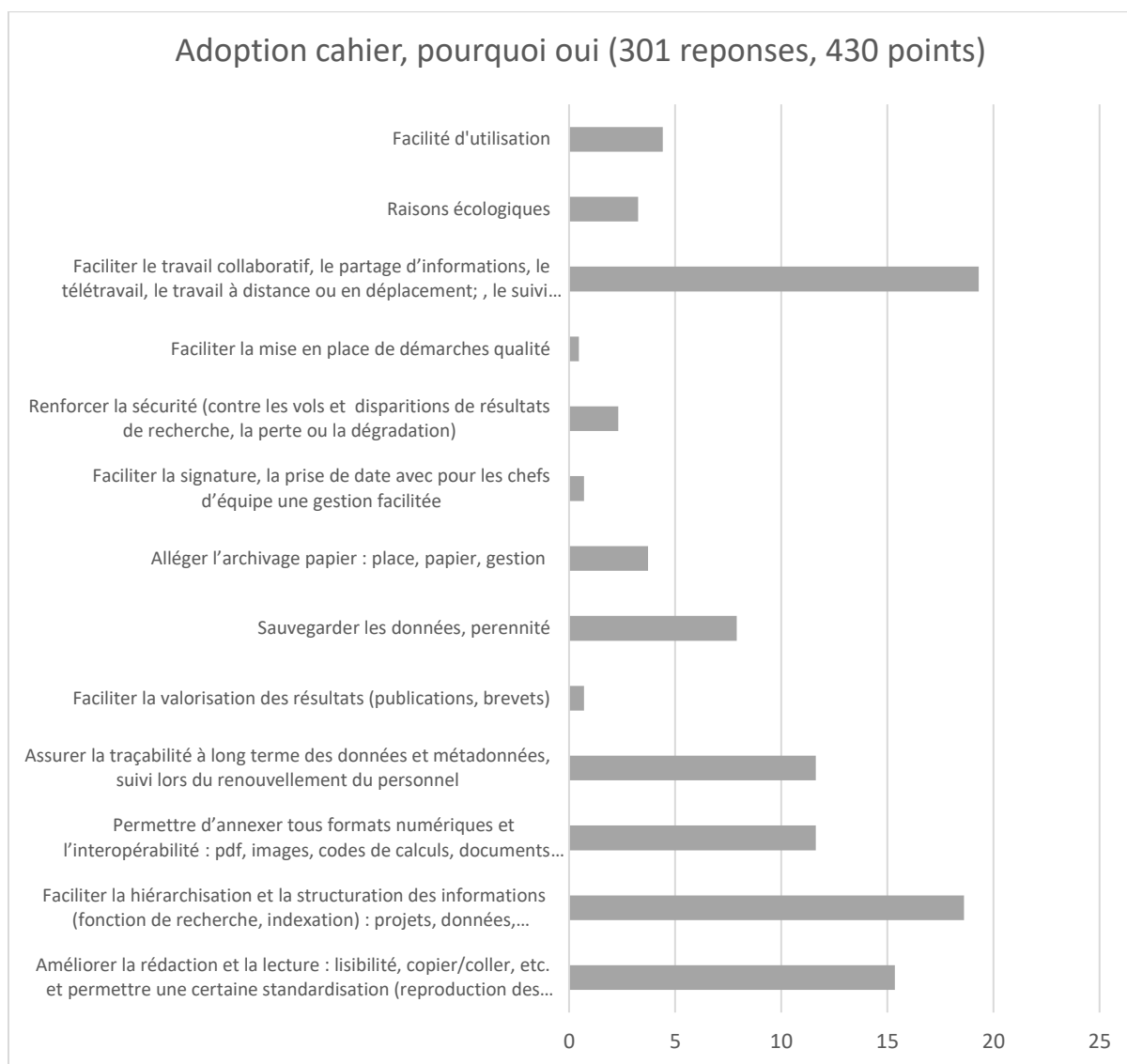
Intérêt à remplacer tout ou partie des cahiers de laboratoires papiers par un cahier de laboratoire électronique car ceux-ci peuvent :	Non intérêt à remplacer tout ou partie des cahiers de laboratoires papiers par un cahier de laboratoire électronique car ceux-ci peuvent :
Eléments communs à tous les instituts	
Améliorer la rédaction et la lecture : lisibilité, copier/coller, etc. et permettre une certaine standardisation (reproduction des expériences)	Ne pas garantir l'agilité de la partie manuscrite (près de la paillasse ou autre) : notes, dessins, etc.
Faciliter la hiérarchisation et la structuration des informations (fonction de recherche, indexation) : projets, données, métadonnées	Ne pas être ergonomique et lourd à utiliser
Permettre d'annexer tous formats numériques (pdf, images, codes de calculs) et permettre d'interfacer logiciellement le CLE (interopérabilité) à d'autres outils (logiciel de gestion expérimentale, outil de développement logiciel, stockage/base/analyse de données, ...)	Augmenter la charge administrative par la nécessité de gestion des utilisateurs et des accès
Assurer la traçabilité à long terme des données et métadonnées ; ainsi que le suivi lors de renouvellement de personnel	Limiter l'utilisation en cas de non accès physique au PC, à l'électricité ou au réseau
Faciliter la valorisation des résultats (publications, brevets)	Risquer la perte d'informations et la charge de travail lors des éventuelles retranscriptions
Sauvegarder les données et assurer leur pérennité	Nécessiter des ressources trop importantes en termes de licences et formations
Alléger l'archivage papier : place, papier, gestion	Ne pas répondre aux attentes des domaines ou des types de recherches diverses
Faciliter la signature, la prise de date avec pour les chefs d'équipe une gestion facilitée	Présenter un risque en termes de piratage, perte de la maîtrise de l'information, de stockage numérique (à court, moyen et long terme) selon la pérennité du et des supports informatiques
Renforcer la sécurité (contre les vols et disparitions de résultats de recherche, la perte ou la dégradation)	Avoir un impact environnemental important, d'où le besoin d'en réaliser la mesure
Faciliter le travail collaboratif, le partage d'informations, le télétravail, le travail à distance ou en déplacement (notamment le suivi des expériences par les responsables) et la portabilité	
Faciliter la mise en place de démarches qualité	
Spécificités selon les instituts	
INEE	Nécessité d'un outil adapté aux missions de terrain (accès hors ligne, accès au travers d'interface type tablette, possibilité d'annexer facilement des photos, etc.)

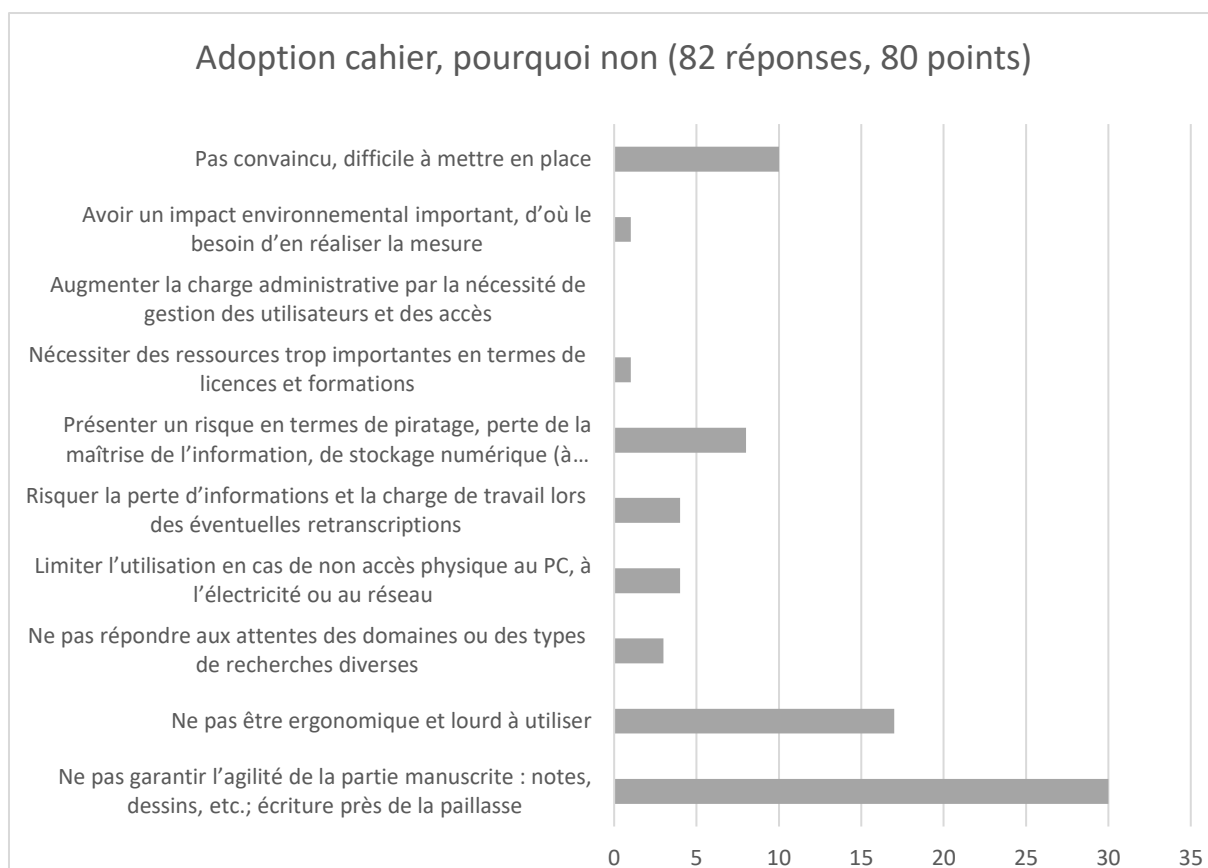
INSHS	<p>Pas ou peu d'utilisation des cahiers de laboratoire papier du MESRI. Le suivi journalier peut s'effectuer dans certaines disciplines sous forme de « journal de bord » ou de « carnet de terrain. Beaucoup de répondants ne se sentent pas concernés par l'outil. Certains expriment cependant un a priori favorable. L'utilisation est parfois itinérante, sans possibilité de connexion au réseau électrique ou internet.</p> <p>En SHS des pratiques se développent autour de la plateforme de carnets de recherche hypotheses (https://hypotheses.org/) et de la grille de services de la TGIR Huma-Num (https://www.huma-num.fr/services-et-outils) qui peuvent se rapprocher fonctionnellement des cahiers de labo dans les autres disciplines, sans être complètement équivalentes. Si l'on considère que l'enjeu organisationnel est de faire émerger au sein des laboratoires de SHS la question des données numériques comme un sujet collectif, on constate que les réponses à l'enquête ne rendent pas compte de cet aspect de la problématique. Mais le champ SHS est tout sauf homogène et la production de données expérimentales (autour de plateformes et d'équipements mutualisés) existe tout autant que dans d'autres domaines disciplinaires des sciences exactes et expérimentales. Celles-ci pourraient constituer une cible prioritaire pour la mise en place de cahiers de laboratoires. Nécessité d'un outil adapté aux missions de terrain (accès hors ligne, accès au travers d'interface type tablette, possibilité d'annexer facilement des photos, etc.).</p>
IN2P3	<p>Peu d'utilisation des cahiers de laboratoire papiers. Existence de plusieurs systèmes de conservation des informations mais utilité de centraliser toutes les informations pouvant être contenues dans un cahiers de laboratoire en un endroit unique.</p> <p>Le moins : cela engendre l'utilisation d'un outil informatique complémentaire</p> <p>Quid des collaborations internationales et règles différentes ?</p>
INC	<p>Besoin éventuel de lier à un système de gestion des stocks et des équipements et de dessin de molécules *</p>
INP	<p>S'inspirer des outils et applications existants et déjà fortement utilisés selon les domaines scientifiques</p>

INSU	Besoin éventuel de lier à un système de gestion des stocks et des équipements * Nécessité d'un outil adapté aux missions de terrain (accès hors ligne, accès au travers d'interface type tablette, possibilité d'annexer facilement des photos, etc.)
INSB	Nécessité d'un outil adapté aux missions de terrain (accès hors ligne, accès au travers d'interface type tablette, possibilité d'annexer facilement des photos, etc.)

* cette fonctionnalité optionnelle peut être prise en charge par l'outil du moment que l'accès aux inventaires de stocks (de produits chimiques, biologiques, de matières nucléaires, etc.) et d'équipements reste limité aux personnes autorisées afin de respecter la confidentialité de ces informations pouvant être sensibles et réglementées

Analyse plus détaillée des résultats de l'INSB :





De façon globale, les principales attentes et craintes sont communes pour l'ensemble des répondants, avec toutefois quelques spécificités propres à certains Instituts.

Malgré une adhésion majoritaire au passage au cahier de laboratoire électronique (76% des répondants sont intéressés par remplacer tout ou partie des cahiers de laboratoire papier par des cahiers de laboratoire électroniques), un accompagnement au changement sera à considérer (au travers de présentations, de formations, de communautés d'utilisateurs, de documentation...). Les craintes principales concernent les aspects informatiques (cf. réponses à la question « quelles sont vos éventuelles craintes relatives à l'utilisation d'un cahier de laboratoire électronique ? ») : ergonomie des outils, difficultés de mise en œuvre notamment lors de missions de terrain, les coûts associés à la fois financier (licences, équipements) et humain (compétences techniques, accompagnement, support) ainsi qu'un manque de confiance dans la fiabilité des outils informatiques et la pérennité des résultats de la recherche.

Un point de vigilance à signaler concerne la diversité des outils déjà existants selon la politique des différentes tutelles. Il est indiqué que l'intention initiale est que le demandeur en laboratoire doit être le plus facilement possible orienté vers un outil répondant à son besoin.

Par ailleurs, il semblerait pertinent de (re) sensibiliser les utilisateurs aux bonnes pratiques de rédaction d'un cahier de laboratoire, qu'il soit papier ou électronique. Un frein tout particulier concerne les équipes qui ne sont actuellement pas utilisatrices de cahiers de laboratoire papiers car elles considèrent que c'est un outil totalement inadapté à leur activité et n'appréhendent pas l'intérêt que peut avoir un cahier de laboratoire électronique pour la protection et la valorisation de leurs résultats.

Si le cahier de laboratoire électronique doit permettre une traçabilité des résultats de recherche, il devra également être une passerelle vers d'autres outils comme les entrepôts de données en permettant de gérer de gros volumes de données (infrastructures de recherche...).

Pour finir, il conviendra de prendre en compte l'impact environnemental, notamment compte tenu du grand nombre d'utilisateurs possibles, localement ou internationalement et de la diversité et l'importance des formats numériques pouvant être partagés.

Il est important de rappeler qu'un livrable, quel que soit sa forme, doit avant tout être utile et utilisable, le risque étant que les utilisateurs ne s'approprient pas les propositions de l'établissement et se tournent vers des outils qui ne répondraient pas aux recommandations du CNRS.

Dans notre cas, l'utilité de ce type d'outil est inhérente au contexte de la demande (cf. introduction du présent document) ainsi qu'au taux des répondants (75%) exprimant être prêts à remplacer tout ou partie des cahiers de laboratoire papiers par des cahiers de laboratoires électroniques.

Pour que le cahier de laboratoire électronique puisse être utilisable, il devra être adapté aux différents domaines de recherche et disposer de facilités de mise en œuvre tout en respectant les caractéristiques techniques auxquelles il devra répondre (en termes de sécurité, stockage, documentation, moyens techniques et financiers, etc.). Il se devra d'être facile d'installation et d'utilisation : une réflexion pourra être menée sur la possibilité de fournir un package aux utilisateurs, comprenant l'outil informatique, la documentation, un espace de stockage ainsi que des moyens de garantie de propriété intellectuelle et d'authentification.