

L'enquête GDR...

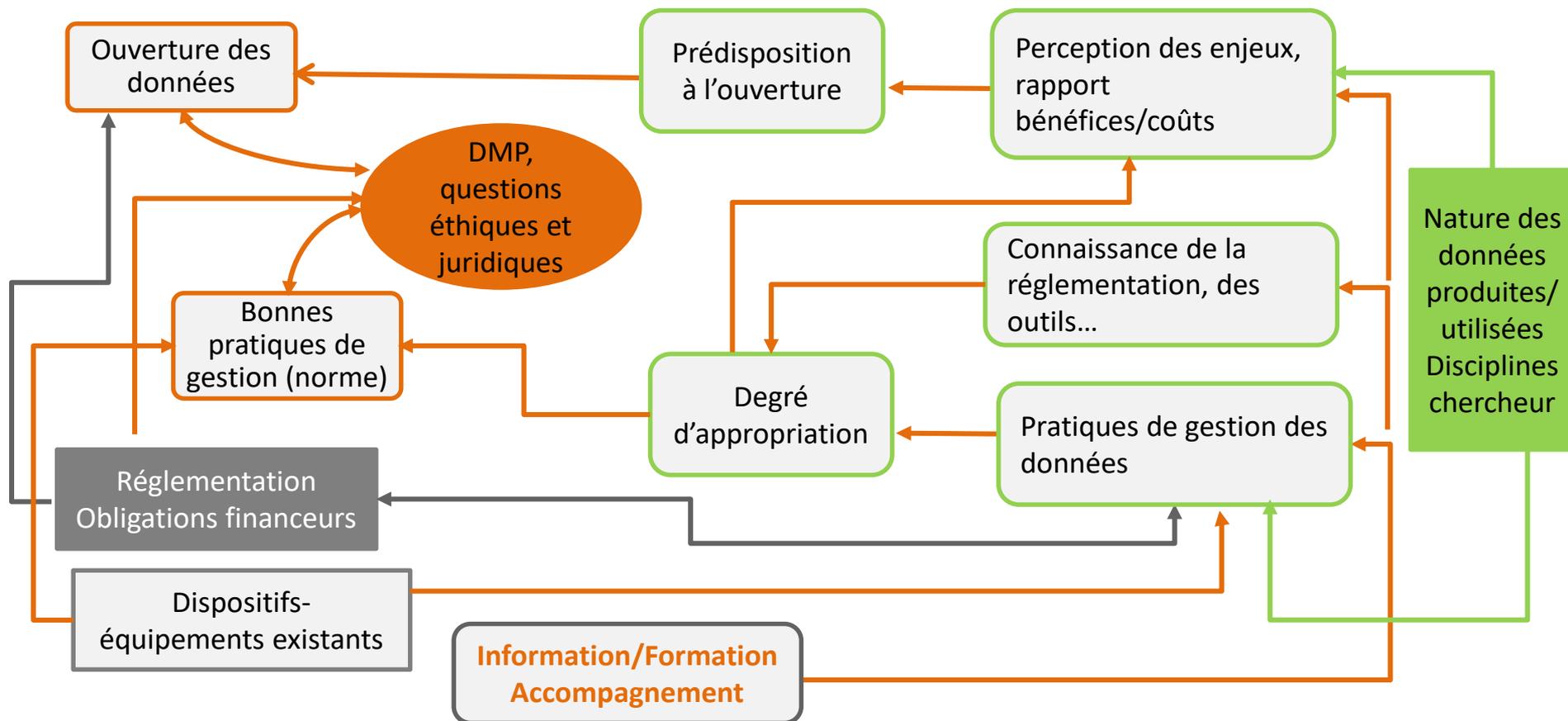
Vers les premiers contours d'une stratégie de laboratoire en matière d'ouverture des données de recherche

Journées annuelles 2021 MATE-SHS « Science Ouverte »

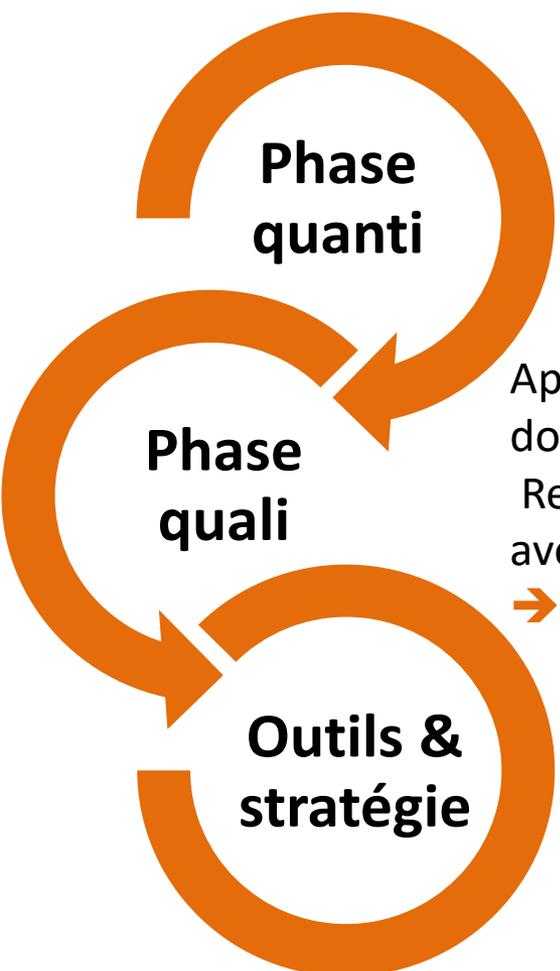
Isabelle André-Poyaud, Sandrine Astor et Guillaume Collain,
Univ. Grenoble Alpes, CNRS, Sciences Po Grenoble, Pacte

Le chercheur au centre du modèle

L'ouverture des données : modèle simplifié (Pacte 2021)



La méthode



Phase quanti

Dresser un état des lieux de l'existant : mesurer les connaissances, comprendre les pratiques actuelles et évaluer les besoins en matière de gestion des données à Pacte

→ Enquête par questionnaire auprès des producteurs/utilisateurs de données à Pacte

Phase quali

Approfondir les atouts et freins que la gestion et l'ouverture des données génèrent

Recueillir des témoignages, retours d'expérience (REX) à partager avec les membres du laboratoire

→ Entretiens semi-directifs sur un thème précis du questionnaire

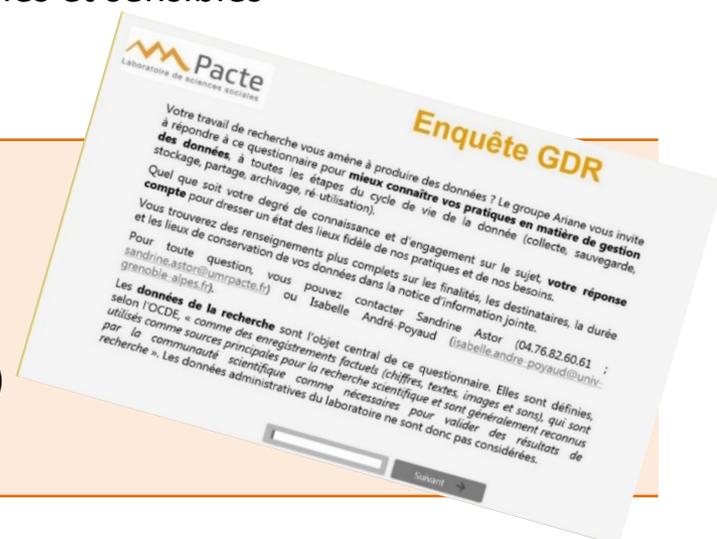
Outils & stratégie

Réaliser des fiches-outils sur différentes phases du cycle de vie de la donnée à l'aide des REX

Construire une plateforme collaborative alimentée par tous ces matériaux

La science ouverte à Pacte : état des lieux...

Passation	En ligne, via SphinxOnline de fin juillet 2020 à mi-janvier 2021 (4 relances)
Population	Chercheurs, enseignants-chercheurs, personnels d'appui à la recherche permanents et doctorants à partir de la 3 ^{ème} année, soit 184 personnes
Questionnaire	Révisé par un GT « Experts », 150 questions environ sur les connaissances, usages GDR et besoins en formation et accompagnement : <ul style="list-style-type: none">- Les données de vos recherches (primaires, secondaires, dérivées)- Vos pratiques de gestion des données exploitation, stockage, sauvegarde, archivage, partage et diffusion, Plan de Gestion de Données, données personnelles et sensibles- Les besoins- Profil (+coordonnées pour entretien)
Taux de réponse	86 questionnaires complets, soit 47% dont : <ul style="list-style-type: none">- 14 chercheurs sur 19 (74%)- 51 enseignants-chercheurs sur 85 (60%)- 16 doctorants de 3^{ème} année sur 72 (22%)- 5 IT-Recherche sur 8 (63%)



La diversité des données primaires

	Effectifs	%
enregistrements audio d'entretiens, de focus groups	64	78
notes d'observation de terrain ou d'entretien	61	74
données d'enquêtes par questionnaire	50	61
textes produits par d'autres personnes que vous (compte rendus de réunions, rapports d'enquêtes publiques, articles de lois, articles de presse, articles scientifiques...)	48	59
photographies, images	41	50
vidéos	25	31
matériaux extraits de fonds d'archives	25	31
données issues du web (comportements d'internautes ou informations sur des personnes ou des choses, collectées manuellement ou par une machine)	21	26
dessins, croquis, cartes dessinées à la main	20	24
données spatialisées	17	21
autres enregistrements sonores	10	12
comptages ou relevés effectués sur le terrain avec ou sans l'aide d'appareils ou d'outils (températures, pluviométrie, levés GPS, levés laser, captation par drone, ...)	10	12
données d'expérience (personnes mises en situation de test ou de jeu, utilisation d'équipements de laboratoire)	10	12
mesures de capteurs portés par la personne (GPS, accéléromètre, objets connectés, ...)	6	7
données générées par un ordinateur (modélisation, simulation multi-agents, ...)	6	7
autre type de données	5	6
Total	82	0

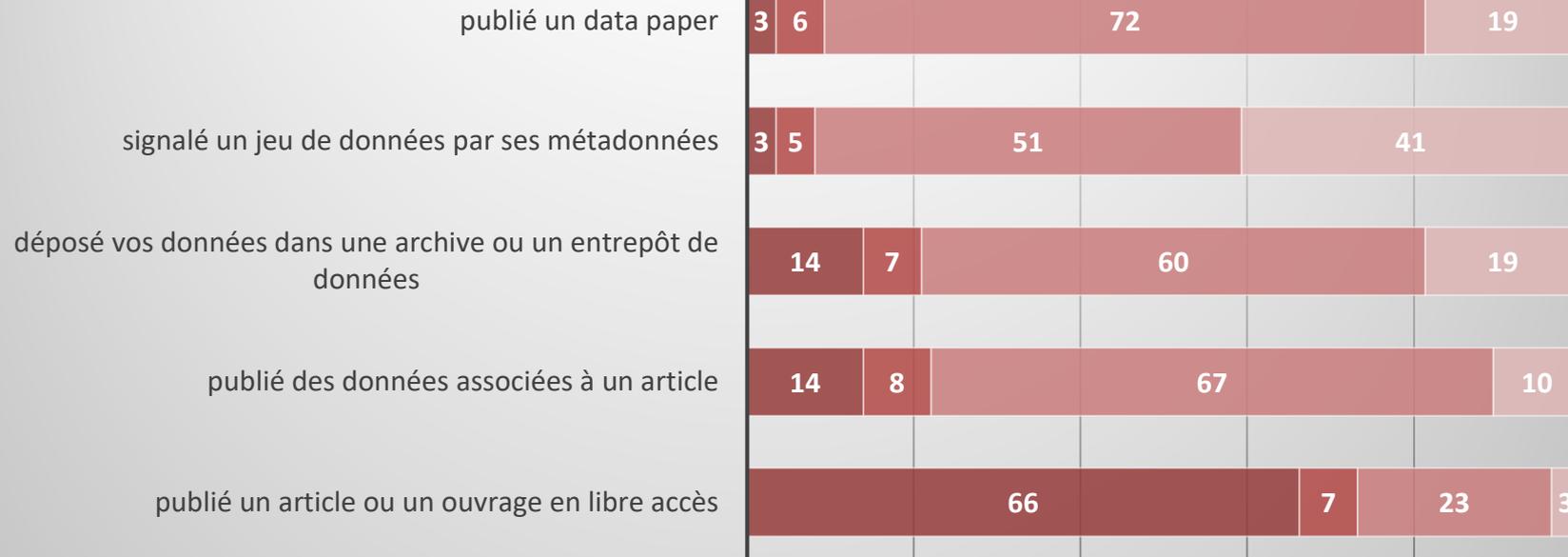
84 % collectent ou ont déjà collecté des données personnelles et 51% des données sensibles

Enquête GDR 2020

Chiffres clés : la diffusion des données à Pacte

Pratiques de diffusion de données et de publications en libre accès

A déjà...

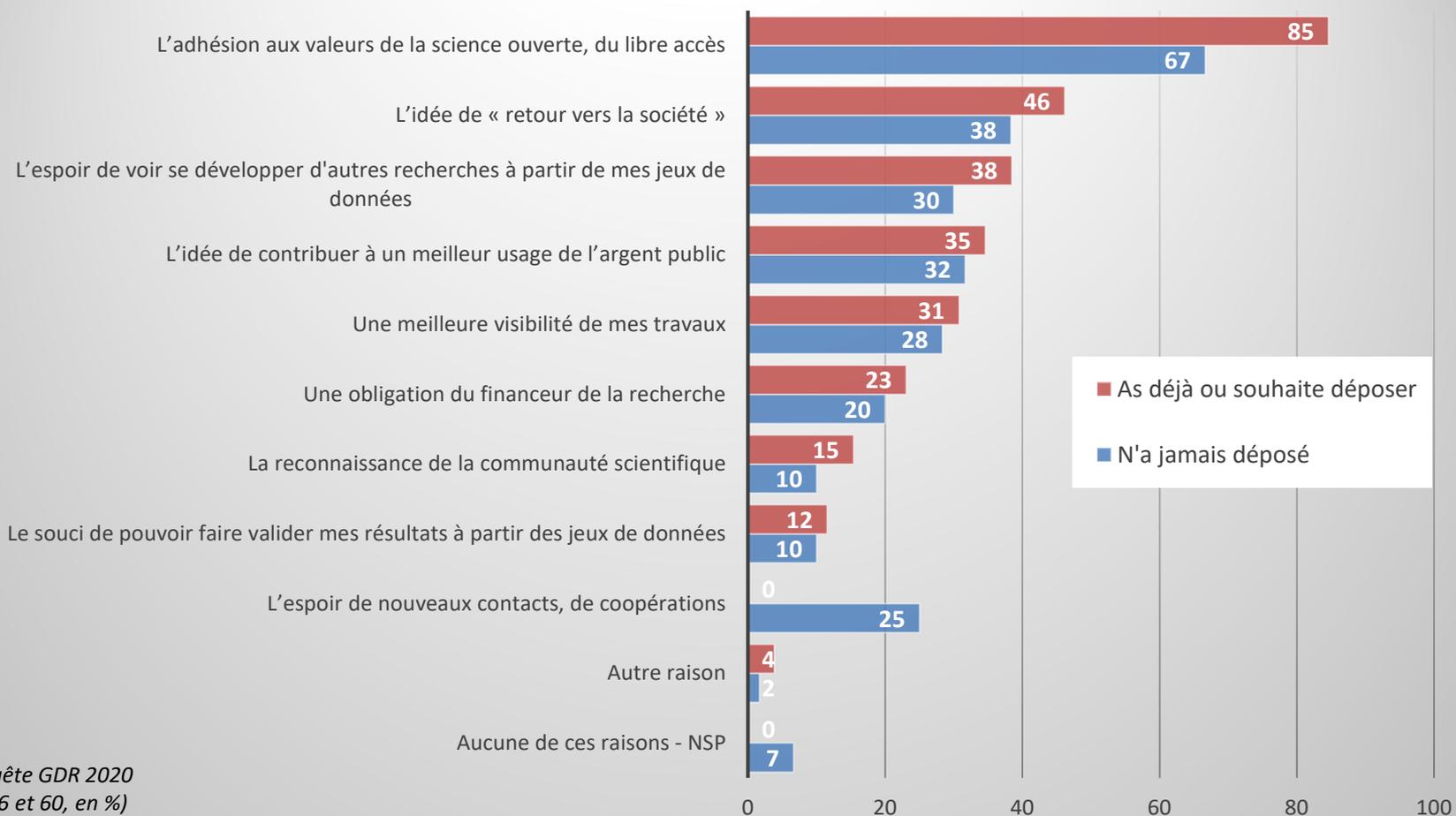


Enquête GDR 2020
(n=86, en %)

■ Oui ■ Non, mais j'ai l'intention de le faire ■ Non ■ Je ne connaissais pas cette possibilité

Chiffres clés : la diffusion des données à Pacte

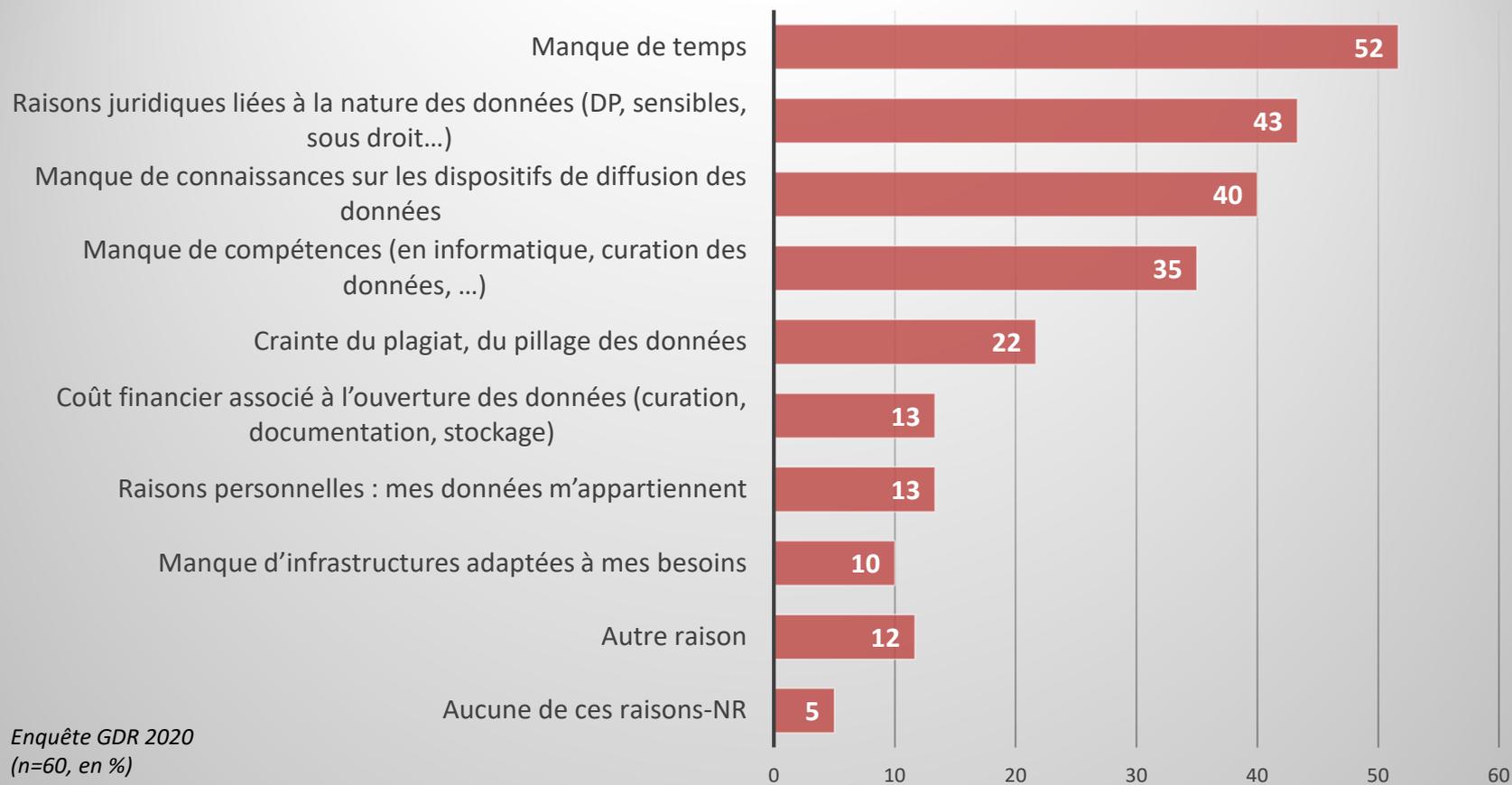
Les raisons pour diffuser ses données en libre accès



Enquête GDR 2020
(n=26 et 60, en %)

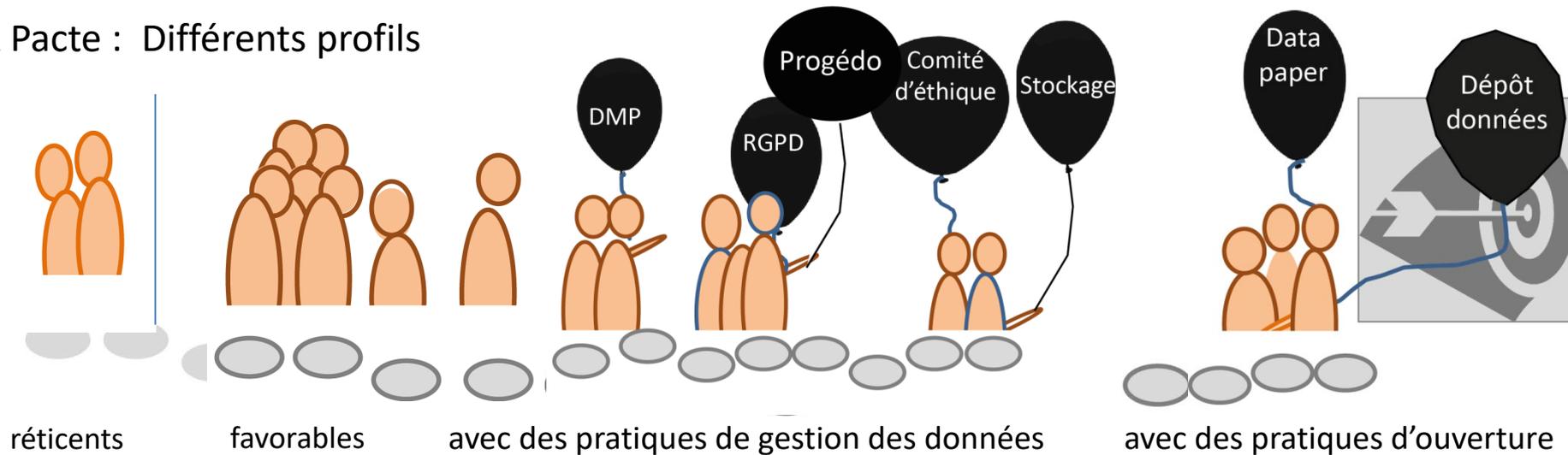
Chiffres clés : la diffusion des données à Pacte

Les freins à la diffusion des données en libre accès



La phase qualitative

A Pacte : Différents profils



Accompagner la gestion des données sur les étapes du cycle de vie... et l'ouverture des données par des retours d'expérience (près de 10 entretiens semi-directifs courts)

pour :

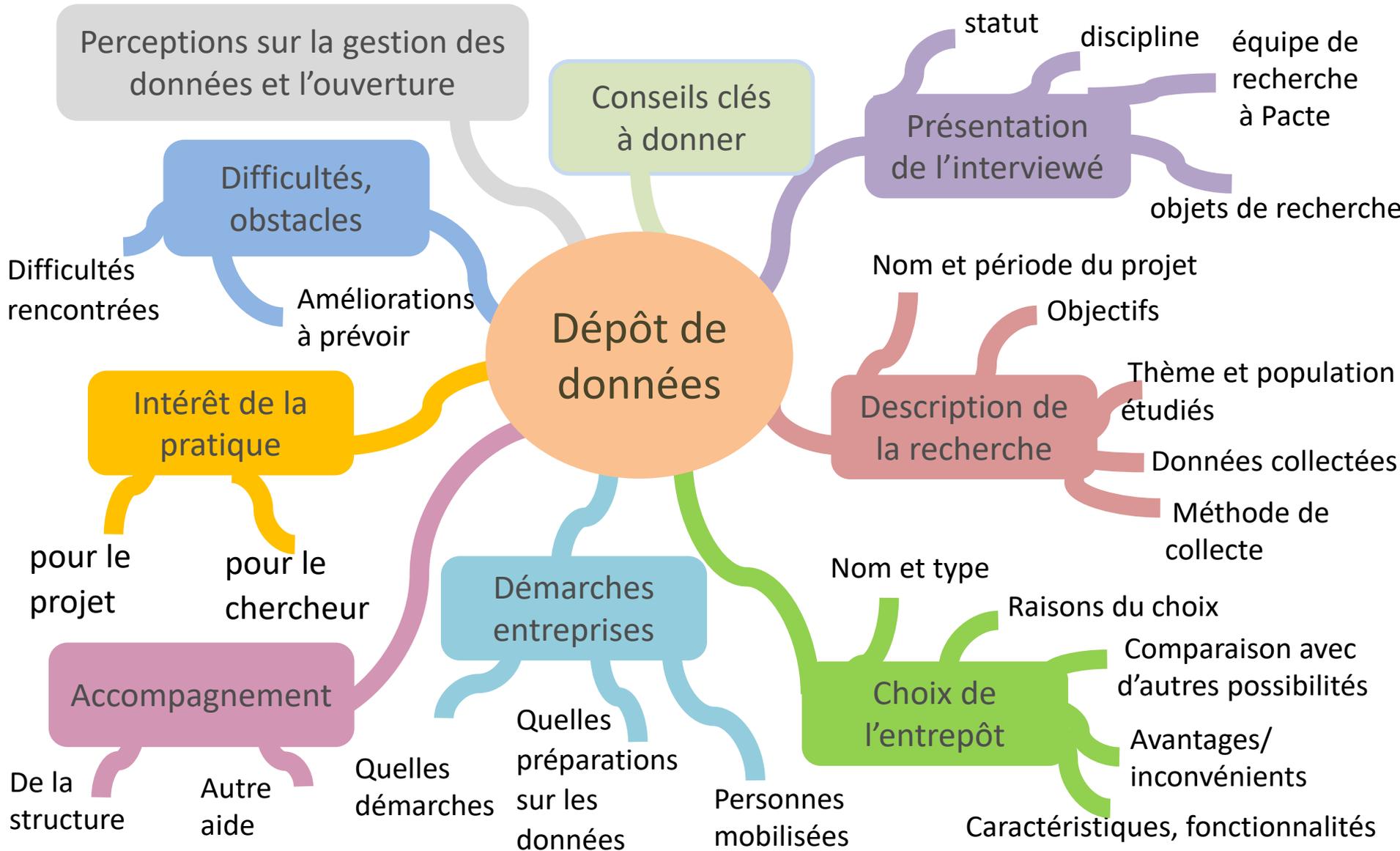
Approfondir des pratiques particulières :

Utilisation de données secondaires ; Partage et diffusion de données (dépôt de données, data paper) ; Rédaction de DMP ; Déclaration de traitement de données personnelles ; Soumission à un comité d'éthique ; Pratiques de stockage

Saisir les perceptions autour de ces nouveaux enjeux

Comprendre les réticences, freins

Un exemple de grille d'entretien



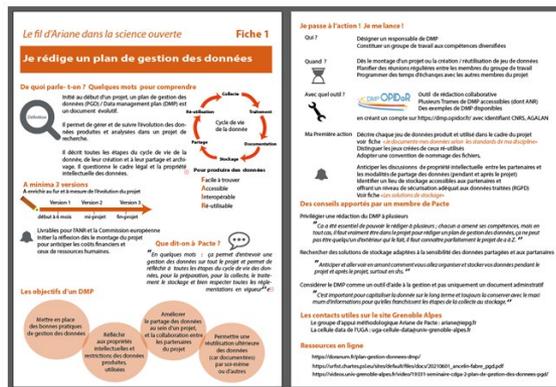
Vers la production de fiches-outils

Rédiger des fiches-outils en utilisant les témoignages recueillis pour

- Sensibiliser les membres du laboratoire à ces pratiques
- Donner des clés pour se lancer

Structure d'une fiche

- Définition et principales caractéristiques de la pratique
- Les finalités
- Les informations clés pour se lancer
- Des conseils issus des retours d'expérience
- Les contacts utiles sur le site Grenoble Alpes
- Des ressources en ligne



Liste non exhaustive

Je collecte des données à caractère personnel

Je demande un avis au comité éthique

Je rédige un Plan de Gestion de données

Je documente mes données selon les standards de ma discipline

Je choisis un entrepôt de données adapté

Je rédige un data paper

Je demande l'attribution d'un DOI

Je dépose mes données dans un entrepôt de données

Un exemple de fiche

La science ouverte à Pacte

Fiche GDR

Je rédige un plan de gestion des données

De quoi parle-t-on ? Quelques mots pour comprendre



Initié au début d'un projet, un plan de gestion des données (PGD) / Data management plan (DMP) est un document évolutif.

Il permet de gérer et de suivre l'évolution des données produites et analysées dans un projet de recherche.

Il décrit toutes les étapes du cycle de vie de la donnée, de leur création et à leur partage et archivage. Il questionne le cadre légal et la propriété intellectuelle des données.

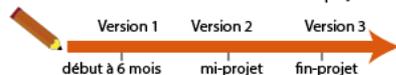


☐ Pour produire des données

- Facile à trouver
- Accessible
- Interopérable
- Ré-utilisable

A minima 3 versions

A enrichir au fur et à mesure de l'évolution du projet



Livrables pour l'ANR et la Commission européenne
Initier la réflexion dès le montage du projet pour anticiper les coûts financiers et ceux de ressources humaines.

Que dit-on à Pacte ?

"En quelques mots : ça permet d'entrevoir une gestion des données sur tout le projet et permet de réfléchir à toutes les étapes du cycle de vie des données, pour la préparation, pour la collecte, le traitement, le stockage et bien respecter toutes les réglementations en vigueur et notamment tout ce qui tourne autour du RGPD."

Les objectifs d'un DMP

Mettre en place des bonnes pratiques de gestion des données

Réfléchir aux propriétés intellectuelles et restrictions des données produites, utilisées

Améliorer le partage des données au sein d'un projet, et la collaboration entre les partenaires du projet

Permettre une réutilisation ultérieure des données (car documentées) par soi-même ou d'autres

Je passe à l'action ! Je me lance !

Qui ?

Désigner un responsable de DMP
Constituer un groupe de travail aux compétences diversifiées

Quand ?



Dès le montage d'un projet ou la création / réutilisation de jeu de données
Planifier des réunions régulières entre les membres du groupe de travail
Programmer des temps d'échanges avec les autres membres du projet

Avec quel outil ?



Outil de rédaction collaborative
Plusieurs trames de DMP accessibles (dont ANR)
Des exemples de DMP disponibles

en créant un compte sur <https://dmp.opidor.fr/> avec identifiant CNRS, AGALAN

Ma première action

Décrire chaque jeu de données produit et utilisé dans le cadre du projet
fiche «Je documente mes données selon les standards de ma discipline»
Distinguer les jeux créés de ceux ré-utilisés
Adopter une convention de nommage des fichiers fiche «Je nomme mes fichiers»



Anticiper les discussions de propriété des données entre les partenaires et leurs modalités de partage (pendant et après le projet)
Identifier un lieu de stockage accessibles aux partenaires et offrant un niveau de sécurisation adéquat aux données traitées (RGPD)
fiche «Je choisis une solution de stockage»

Des conseils apportés par un membre de Pacte

Privilégier une rédaction du DMP à plusieurs

"Ca a été essentiel de pouvoir le rédiger à plusieurs ; chacun a amené ses compétences, mais en tout cas, il faut vraiment être dans le projet pour rédiger un plan de gestion des données, ça ne peut pas être quelqu'un d'extérieur qui le fait, il faut connaître parfaitement le projet de A à Z."

Rechercher des solutions de stockage adaptées à la sensibilité des données partagées et aux partenaires

"Anticiper et aller voir en amont comment vous allez organiser et stocker vos données pendant le projet et après le projet, surtout en shs."

Considérer le DMP comme un outil d'aide à la gestion et pas uniquement un document administratif

"C'est important pour capitaliser la donnée sur le long terme et toujours la conserver avec le maximum d'informations pour qu'elles franchissent les étapes de la collecte au stockage."

Les contacts utiles sur le site Grenoble Alpes

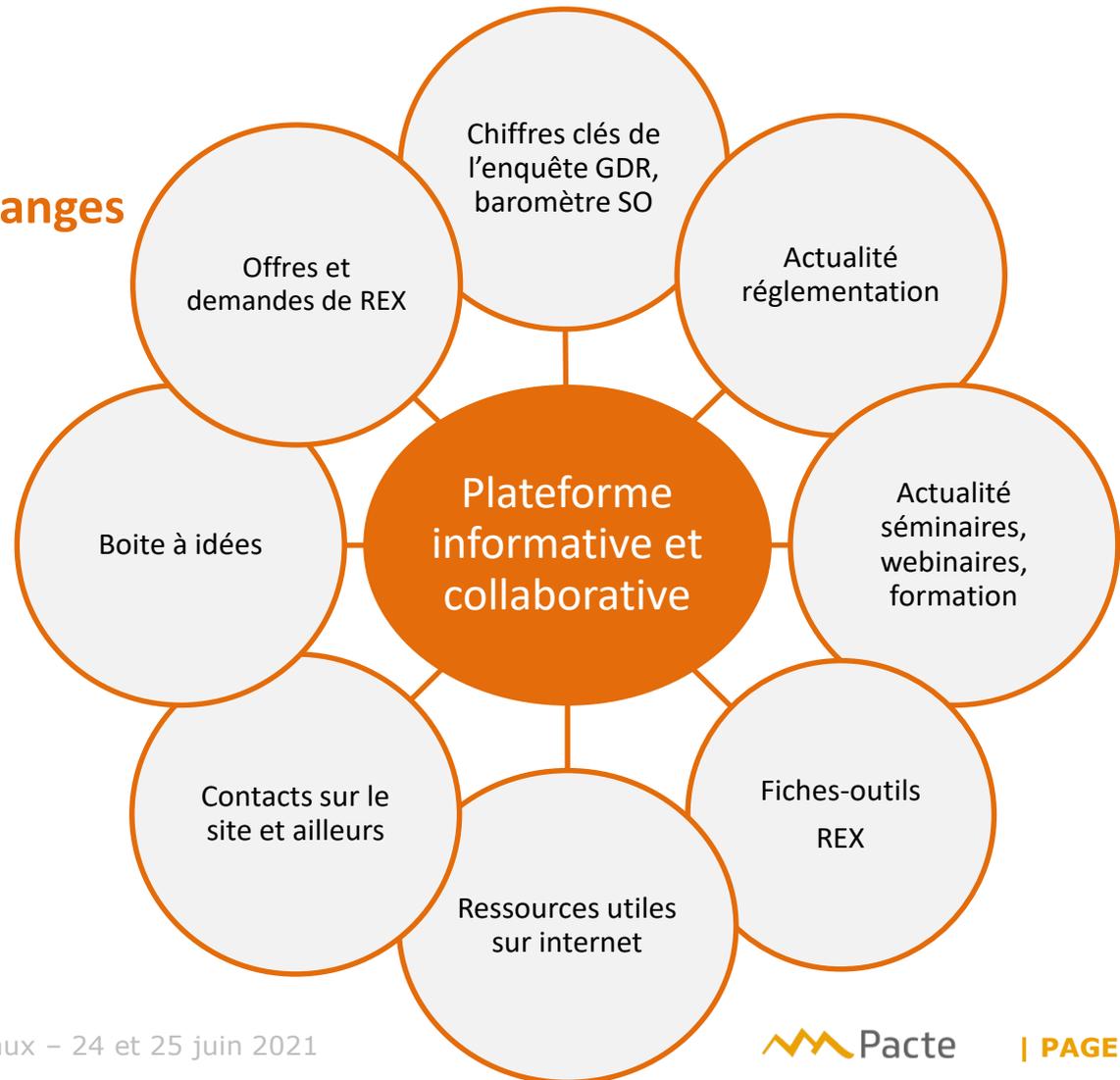
Le groupe d'appui méthodologique Ariane de Pacte : ariane@iepg.fr
La cellule data de l'UGA : uga-cellule-data@univ-grenoble-alpes.fr

Ressources en ligne

<https://doranum.fr/plan-gestion-donnees-dmp/>
https://urfst.chartes.psl.eu/sites/default/files/docs/20210601_ancelin-fabre_pgdp.pdf
<https://videos.univ-grenoble-alpes.fr/video/19331-seminaire-cdga-2-plan-de-gestion-des-donnees-pgdp/>

Vers une plateforme collaborative pour la science ouverte à Pacte

- Pour informer
- Pour former
- Pour générer des échanges de savoir-faire



Merci de votre attention...



Des questions ?