

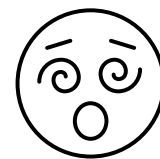
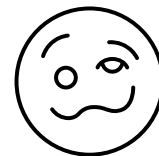
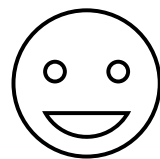
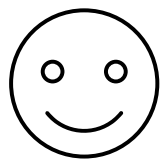
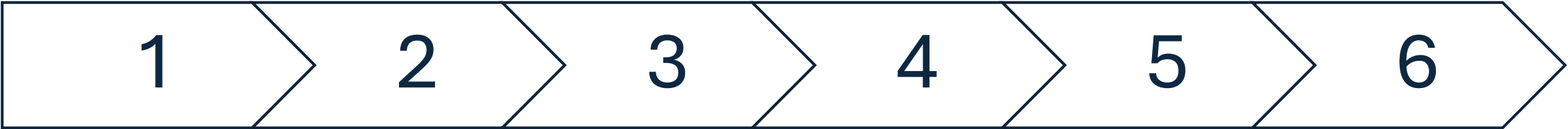
# Retour d'expérience

## Ouvrir les données de la thèse

JTTA 24/06/2025

[clotilde.george@univ-lorraine.fr](mailto:clotilde.george@univ-lorraine.fr)

# Les années de doctorat et la Q de l'ouverture des données





On se lance ! On contacte des futur-es participant-es

1

2

3

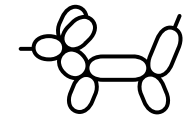
4

5

6



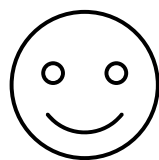
On bricole un contrat avec des éléments inspirés de modèles existants, piochés de part et d'autre



On présente aux participant·es potentiel·les le projet avec les enjeux de publication, et sans les effrayer



(on en effraie quand même)



On filme !

1

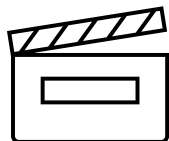
2

3

4

5

6

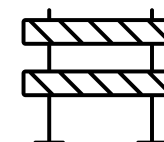


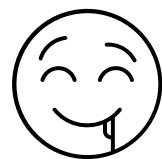
Tout le monde a signé, ça tourne !

Des galères techniques... On travaille dans l'urgence  
(les soirs, les weekends, les vacances)  
mais ça se passe bien avec le « terrain », on est content.e

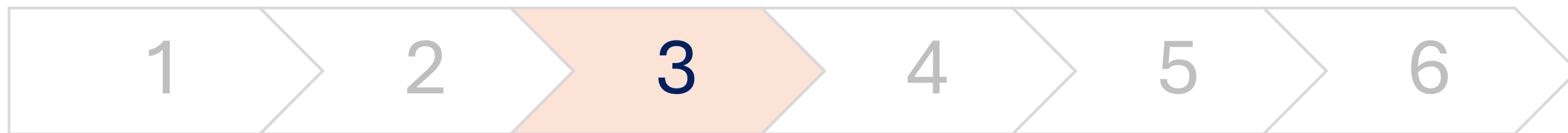


A ceci près : le covid, la fermeture du restaurant, ETC.





Pas grave, on a de quoi traiter !



On en bave à transcrire, annoter... mais ça avance.

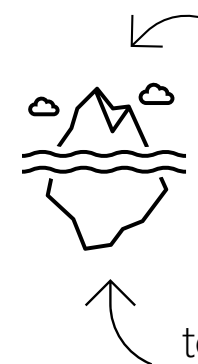
On commence à entendre parler de termes obscurs.



RGP\_ quoi ? PG\_ quoi ? DMP ? DP\_ quoi ?

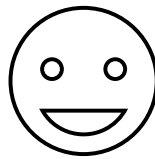
Un comité d'éthique ? Quel comité d'éthique ?

FAIR ? Data papers ?



contrats,  
anonymisation

tout le reste

Et si...  on refilmait ?



Une nouvelle salve de contrats, plus complets

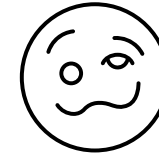
De nouvelles personnes filmées

Une meilleure qualité de captation

Un plan de gestion de données !



Euh:::



1

2

3

4

5

6



Au fait, la publication du corpus, on a le droit ?

Des échanges avec l'institution (DPO)  
en dents de scie

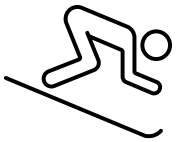


Verdict : les contrats signés en 2019 et 2022 ne  
sont pas valides pour une publication du corpus

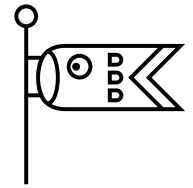
Et il faudrait un contrat pour chaque article publié !!!



On boucle ! 



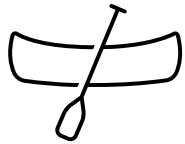
Il faut terminer de rédiger, soutenir la thèse.  
Le reste, on verra après !



Le reste,

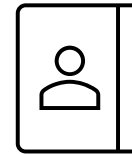
7

c'est maintenant.



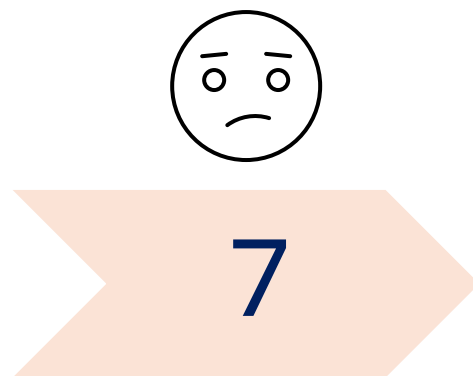
Il faut recontacter tout le monde

Certes, c'est un petit corpus, un petit terrain



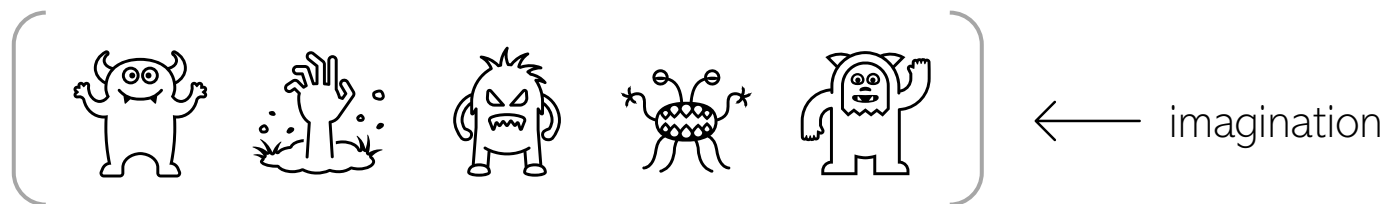
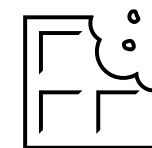
mais aujourd'hui dispersé  
(en France, en Pologne, en Hongrie...)





On a envoyé la thèse aux principaux concernés

C'est un « retour au terrain », on n'est pas serein·e

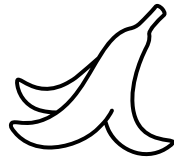




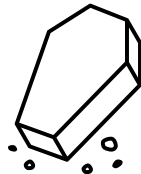
En fait, tout va bien.



...Tout va bien ?

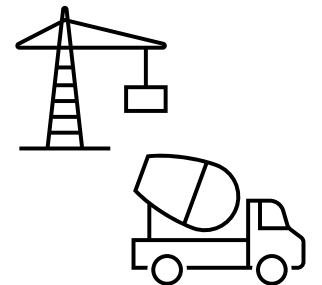


« MAE » ne répond pas



Est-ce la mort de la publication du corpus ?

Non, juste un gros chantier pour tronquer le corpus





7



Heureusement, il y a AVAA

et ses « *remove-sequences-from-corpus* » &  
« *reduce-corpus* » & « *sequences-to-corpus* »



La reproductibilité est atteinte,  
mais les données seront ouvertes

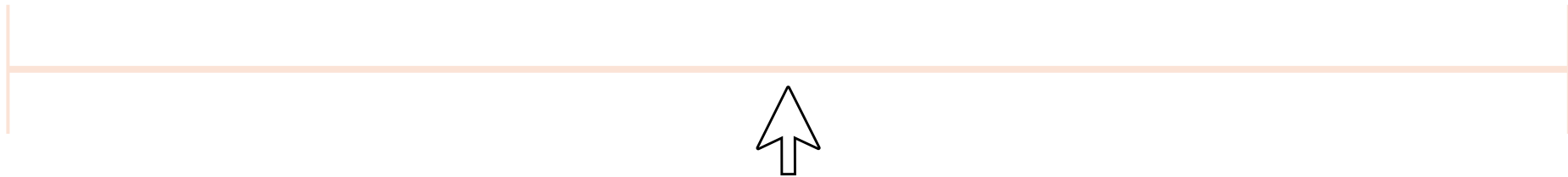


A suivre...

# La science ouverte, sur les épaules de Darwin doctorant·es...

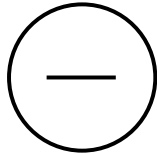
## Une question

« Aussi ouvert que possible, aussi fermé que nécessaire »



Où placer le curseur ?

# La science ouverte, sur les épaules de Darwin doctorant·es...



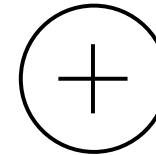
- **Anxiété**

(et si les personnes étaient gêné·es par certaines données ou analyses ? et si elles changeaient d'avis ? etc. )

- **Peu de visibilité** du travail lié au droit des données

(un temps passé à contractualiser les droits et devoirs des différentes parties, qui n'est pas facilement valorisé)

- **Décalage** entre prescription et pratiques réelles



- **Responsabilisation** face aux enjeux et risques de la publication de données

- Une **leçon** : multiplier les moyens de contact avec les participant·es

- **A faire** : un espace-temps d'échange sur ces Q entre doctorant·es ?

Merci !