

FICHE PRATIQUE RGPD

Bonnes pratiques pour la rédaction de la thèse

Cette FAQ rassemble quelques conseils pratiques, complémentaires aux obligations légales, pour aider les doctorants à anticiper la soutenance, l'archivage et la diffusion de la thèse en toute sérénité. Il ne s'agit pas des démarches obligatoires de conformité, mais de **bonnes pratiques simples à appliquer** pour éviter les erreurs fréquentes dans le cadre des projets de recherche.

1 - Quand commencer sa conformité RGPD ?

La conformité RGPD doit être intégrée **dès le début de votre projet**, idéalement **dès la conception du protocole** et de la méthodologie de recherche. Cela permet d'anticiper les questions liées aux données personnelles et d'éviter des ajustements de dernière minute une fois la collecte ou l'analyse des données commencée.

La démarche n'est pas particulièrement complexe, mais elle requiert du temps, de la réflexion et un certain recul pour définir la méthodologie la plus adaptée. Elle ne vise pas à ralentir votre travail, mais à l'accompagner : elle s'intègre en parallèle de vos activités et vous aide à structurer vos pratiques dès le départ. Même si certaines questions restent en suspens, commencer tôt facilite vos choix méthodologiques et réduit les risques de non-conformité. Une fois le projet avancé, corriger certains écarts devient beaucoup plus difficile, voire impossible pour certaines données.

Un autre argument pour commencer tôt tient à l'articulation entre le RGPD et l'arrêté encadrant l'obtention du doctorat. Cet arrêté fixe les règles relatives au diplôme national de doctorat, régissant à la fois le dépôt des thèses et leur version diffusée après soutenance, avec des contraintes strictes. En particulier, l'ensemble des thèses déposées à l'archivage, puisse ouverte à la diffusion doivent correspondre exactement (ou presque) à la version validée par le jury.

C'est à ce niveau que surgissent les difficultés. Comme rappelé précédemment, le RGPD doit être intégré dès le début du projet de thèse (« privacy by design »). Or, de nombreuses thèses, déjà soutenues ou en cours de dépôt, n'ont pas pris en compte ces aspects dès la rédaction initiale. Il en résulte souvent des modifications de dernière minute, qui peuvent entrer en tension avec les exigences de l'arrêté, puisque **la version archivée et diffusée doit rester fidèle à celle validée par le jury**, d'où l'intérêt de pouvoir anticiper ces aspects.



2 - Quelles difficultés peuvent survenir si je ne prends pas certaines précautions lors de la rédaction de ma thèse ?

Exemple 1 : Consentement des participants

L'une des principales difficultés liées au consentement apparaît lorsque les pratiques de collecte de données n'ont pas été correctement encadrées dès le départ. Cela concerne particulièrement les projets impliquant des entretiens ou des enquêtes en direct, sans formulaire de consentement complet. Il ne suffit pas de demander aux participants leur accord pour participer à la recherche : il faut également obtenir leur consentement explicite pour la collecte et le traitement de leurs données personnelles, ce dont ils ne sont pas toujours pleinement conscients.

Un consentement insuffisamment détaillé ou imprécis complique fortement la situation, notamment si les données brutes n'ont pas été pseudonymisées au préalable. Le doctorant doit anticiper plusieurs questions essentielles : les participants ont-ils accepté que leurs données soient éventuellement accessibles en ligne ? Autorisent-ils la publication de photos, d'enregistrements ou de vidéos ? Consentent-ils au partage de leurs données avec d'autres chercheurs ou pour de futurs projets scientifiques ? Dans de nombreux cas, il ne s'agit pas d'un refus de la part des participants, mais simplement d'un manque de précision initiale.

Une fois les données collectées sans cadre clair, il devient très difficile, voire impossible, de régulariser la situation : recontacter les participants pour obtenir un consentement rétroactif est souvent chronophage et peu fiable, et certaines personnes peuvent refuser rétroactivement, compromettant la validité des données déjà recueillies.

Exemple 2 : Données personnelles superflues disséminées dans la thèse

Les annexes peuvent devenir un véritable « fourre-tout » où des jeux de données complets sont parfois inclus avec l'identité des participants, sans consentement préalable ni pseudonymisation ou anonymisation. Cela pose un risque important, car ces données peuvent être consultables ou diffusées, et il devient alors difficile de revenir en arrière pour corriger ou retirer les informations une fois la thèse finalisée.

Les remerciements peuvent également poser problème si des personnes y sont citées sans avoir été informées que leur nom apparaîtrait dans la thèse publiée. Ces personnes peuvent exercer leurs droits et demander le retrait de leurs informations, ce qui oblige à modifier une thèse déjà publiée et complique sa diffusion et sa conservation. À l'inverse, la mention des directeurs ou directrices de thèse est généralement considérée comme légitime et moins sensible sur le plan personnel.

3 - Concrètement, quels sont les bons conseils pour la rédaction de ma thèse ?

Bien que le dépôt et la diffusion de votre thèse soient généralement pris en charge par les correspondants thèses et les personnels d'appui à la recherche, il vous revient de **prévoir une version adaptée à ces usages**, idéalement purgée des données personnelles.

La rédaction de votre thèse est donc le moment idéal pour organiser vos contenus et anticiper d'éventuelles difficultés futures. Il est fortement conseillé de distinguer dès la rédaction **deux versions de la thèse** :

La version complète de la thèse, destinée au dépôt officiel et à l'archivage obligatoire, peut contenir des données personnelles si celles-ci sont indispensables. Il est donc recommandé d'adapter cette version afin qu'elle ne contienne que les données personnelles strictement nécessaires à la validation par le jury. Après la soutenance et la validation par le jury, toute modification devient plus complexe, puisque la version soutenue sert de référence pour le dépôt à l'archivage obligatoire.

La seconde, la version publique ou diffusée, est quasi identique à la précédente mais purgée des données personnelles non destinées à être rendues publiques. Elle peut être diffusée en ligne sans risque via Theses.fr ou HAL par les établissements, en accès restreint ou libre, et doit appliquer certaines bonnes pratiques : pseudonymisation ou anonymisation des données en l'absence de consentement des participants, vérification et limitation des informations superflues (notamment dans les annexes et les remerciements), et remplacement ou masquage des éléments sensibles par des mentions neutres telles que « retrait pour confidentialité » par exemple.

Préparer ces deux versions dès la rédaction permet d'anticiper les difficultés futures, de protéger les participants et les personnes citées, et de faciliter à la fois la soutenance, le dépôt officiel et la diffusion de la thèse. Parlez-en avec vos référents en thèse !

Document réalisé par Rebecca Rousseau, adjointe DPO et RSSI Université Paris 1 Panthéon-Sorbonne, sous licence CC BY-NC-SA

