MASTER EN SCIENCES INFORMATIQUES À FINALITÉ SPÉCIALISÉE EN SOFTWARE ENGINEERING

Mise en oeuvre du RGPD dans un système d'information à travers l'analyse des flux de données

Montulet, Mathilde

Award date:
2019

Awarding institution:
Universite de Namur

UNIVERSITÉ DE NAMUR

FACULTÉ D'INFORMATIQUE

Mise en oeuvre du RGPD dans un système d'information à travers l'analyse des flux de données

Mathilde Montulet

RUE GRANDGAGNAGE, 21 ● B-5000 NAMUR(BELGIUM)

Université de Namur
Faculté d'informatique
Année académique 2018–2019

Mise en oeuvre du RGPD dans un système d'information à travers l'analyse des flux de données

Mathilde Montulet

Maître de stage : Jean-Noël Colin
Promoteur : _____ (Signature pour approbation du dépôt - REE art. 40)
_Jean-Noël Colin

Mémoire présenté en vue de l'obtention du grade de
Master en Sciences Informatiques.

Résumé/Summary

Résumé

Ce mémoire s'intéresse à l'application d'un sous-ensemble du RGPD grâce à l'analyse des flux de données dans un contexte national. Le RGPD, le Règlement Général sur la Protection des Données, est un texte de loi qui s'applique en Europe. La réflexion se base sur un sous-ensemble de ce règlement qui est détaillée dans le travail et s'applique particulièrement au cas des PME et des ASBL.
L'objectif de ce mémoire a été de trouver un outil méthodologique qui puisse aider un informaticien à déterminer des Zones de risque de non-conformité dans le système qu'il analyse.
Pour cela il a été décidé d'étendre la méthodologie de *threat modelling* LINDDUN. LINDDUN repose sur l'utilisation de diagramme de flux de données. Chacune des six étapes de cette dernière méthodologie a été examinée et étendue au RGPD. Les différents types de menaces pour la vie privée existant dans LINDDUN ont été adaptés au RGPD et deux types de menaces spécifiques ont été ajoutés. La méthodologie a ensuite été testée sur trois scénarios pratiques regroupés sous forme d'un scénario anonyme.

Summary

This master's thesis looks at how the analysis of data flow can be used to apply a subpart of the GDPR in a national context. The GDPR, General Data Protection Regulation, is a law text which applies in Europe. The reflexion is based on a subpart of this regulation which is detailled in this master's thesis and which particularly applies to SME and non-profit organizations.
This master's thesis aimed to find a methodological tool which could help a computer scientist to identify risk zones of non-compliance in the system undergoing analysis.
To get to this result, we have chosen to extend the threat modelling method LINDDUN. LINDDUN is based on the use of dataflow diagram. All of the six steps of this methodology were adapted to the GDPR and two new threat types were added. The methodology was then tested on three concrete scenarios taking the form of an anonymous scenario.

Remerciements

Je remercie le Professeur Colin pour ses conseils et ses relectures lors de la rédaction de ce mémoire.

Je remercie l'équipe de recherche InfoSec avec laquelle j'ai eu l'occasion de faire mon stage.

Je remercie Mr Ponsard et Mr Grandclaudon du CETIC pour leurs conseils au démarrage de ce travail.

Je remercie la DPO de l'UNamur K. Rosier pour la discussion enrichissante que j'ai eue avec elle au sujet de ce mémoire.

Je remercie également les 3 personnes que j'ai interviewées dans le cadre de ce travail.

Je remercie mes parents et ma soeur qui m'ont soutenue pendant mes études.

Table des matières

Table des figures

Chapitre 1

Introduction

1.1 Pourquoi s'intéresser à la vie privée ? Mise en contexte

La vie privée et son respect intéressent une partie importante de la population. Depuis l'entrée en vigueur du RGPD le 25 mai 2018, plus de 95 000 personnes ont porté plainte en Europe dans le cadre de la protection de leurs données personnelles. 3 amendes en ont déjà résulté, l'une d'entre elles, adressée à Google, s'élevant à 50 000 000 d'euros. Les représentants de la Commission Européenne ont déclaré à l'occasion de la "Journée européenne de la protection des données", au sujet du RGPD, que "*Ce qui est en jeu, c'est non seulement la protection de notre vie privée, mais aussi la protection de nos démocraties et la durabilité de nos économies fondées sur les données.*" [3] En France, 6 mois après l'arrivée du RGPD, ce n'est pas moins de 66% de français qui se considèrent plus intéressés par les questions de vie privée qu'auparavant.[4]

La collecte de données privées par des entreprises, lorsque des personnes naviguent sur internet, est connue depuis longtemps. Plus encore que la question de la récolte des données, c'est l'usage qui en est fait qui peut être l'objet de questions. La collecte, l'analyse et l'exploitation des données, si elles peuvent être justifiées et nécessaires à l'accomplissement de la tâche de service proposée, peuvent aussi avoir des conséquences négatives. Les médias ont déjà parlé de cas comme celui de Catherine Taylor qui a été catégorisée comme "ayant tenté de créer des méthamphétamines" à tort et a dû faire face à des conséquences négatives dans sa vie quotidienne. Le problème soulevé est donc un traitement qui peut parfois être superficiel ou donner des résultats erronnés. Le cas le plus connu de ces dernières années reste l'affaire Cambridge Analytica où les données personnelles de millions de personnes ont été récoltées sans leur accord. Ainsi, certains algorithmes peuvent considérer des corrélations comme des preuves irréfutables, par exemple, qu'une personne va bientôt se trouver en difficulté financière. Dans ce cas, la personne peut alors se voir refuser un emprunt sur base des conclusions de l'algorithme défectueux et non de sa situation effective. L'un des droits donné par le RGPD est le droit de l'utilisateur ou l'utilisatrice de demander à connaitre les traitements effectués sur ses données ce qui peut éviter certaines de ces situations.[5]

Shoshana Zuboff a écrit un livre parlant du "*capitalisme de surveillance*" qui repose sur l'utilisation de données personnelles qui font partie de la vie privée des individus dans une optique de création d'argent et de revenus, une "*surveillance commerciale*" comme elle la nomme. Ainsi la pratique de la récolte de données a évolué avec le temps et les acteurs qui s'y sont intéressés. Google a fait breveter en 2003 une méthode permettant de présenter aux utilisateurs et utilisatrices de ses services des publicités en lien direct avec leurs goûts. Le but est d'inciter et de modifier le comportement des utilisateurs et des utilisatrices à cliquer plus fréquemment sur les publicités et donc de générer plus de revenu pour l'entreprise. Cela n'étonnera personne qu'une quantité massive de données doit être récoltée et traitée pour atteindre ce résultat. Cette récolte est souvent faite sous couvert d'une personnalisation de l'expérience utilisateur. L'autrice présente les choses sous un autre jour et s'interroge sur le fait que le trait de génie économique de Google pourrait avoir été de se rendre compte que les prédictions du comportement d'un individu rapportent plus que l'individu ou que la recherche de la complétion de ses besoins... L'autrice présente ainsi deux impératifs qui caractérisent cette économie particulière (celui de l'"*extraction*" de ces données et celui de leur "*approfondissement*"). Ils permettent de découvrir des

pans entiers de la vie personnelle d'une personne sous prétexte de personnaliser son expérience. La situation qui pose question à l'autrice est aussi le fait que ce nouveau système économique, qui repose entièrement sur la donnée personnelle, peut aussi servir à influencer l'utilisateur. Par exemple, cela peut permettre de faire apparaître à l'écran des informations particulières qui ont été déduites comme pouvant aider à infléchir son comportement, ou, dans des cas plus extrêmes, prendre potentiellement des décisions à sa place comme celle d'éteindre le téléviseur pour forcer le "propriétaire" de l'objet à aller dormir plus tôt. [6]

Le droit à la vie privée avait déjà été évoqué dans la convention européenne des droits de l'homme en 1950. Le terme "*donnée personnelle*" apparait en 1981 dans le droit européen à l'occasion de la "*Convention pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel*". La réflexion sur le sujet de la vie privée n'est donc pas nouvelle, ainsi en mai 2014 la Cour De Justice Européenne adoptait le "*droit à l'oubli*" et en avril 2014 elle publiait un arrêt qui rendait "*invalide*" la directive européenne de 2006 qui concernait la "*conservation des données personnelles*". Contrairement à ce que certaines personnes peuvent penser, un agrégat de données personnelles peut permettre d'en savoir plus sur un individu que ce que ces simples données personnelles affichent. Il est en effet possible de déterminer par exemple, les habitudes et les cercles sociaux de cette personne. L'auteur, Jérôme Thorel, avance même que la possibilité de suivre les individus de la sorte peut aller jusqu'à menacer "*la liberté de réunion et d'association, à la liberté de conscience et d'opinion ainsi qu'à la liberté de mouvement.*" [7]

1.2 Le règlement général sur la protection des données (RGPD)

Le 25 mai 2018, le RGPD entrait en vigueur en Europe. Ce règlement traite de la protection des données et a pour but de réglementer, suite à son adoption par le Parlement Européen en 2016, la façon dont les données sont récoltées et utilisées de manière plus uniforme pour l'Europe. Ce droit n'existe pas dans d'autres pays, y compris les pays leaders en matière technologique.

Le RGPD porte sur les données personnelles des citoyens de l'Union Européenne et concerne toutes les entreprises qui traitent des données personnelles : de la PME d'un ardoisier couvreur de village qui fait une mailing-list, à la Région Wallonne, en passant par une entreprise présente exclusivement sur internet comme par exemple Zooplus ou une pharmacie en ligne, ... Ses objectifs sont, entre autres, de responsabiliser les entreprises d'un point de vue digital et de redonner confiance aux consommateurs. Globalement, le RGPD porte sur les droits d'accès, d'information, de rectification, d'opposition, de limitation du traitement des données et à l'oubli. Il porte également sur la portabilité des données et la demande de notification des violations de données à caractère personnel. Avant l'application du RGPD, près de 80% des entreprises selon une estimation n'avaient pas conscience de son existence ou des changements profonds qu'il amènerait dans la récolte et le traitement quotidien des données. L'Union Européenne donnait seulement 2 ans pour qu'elles se mettent à niveau et prévoit des sanctions financières en cas de non-conformité.[8] Il était aussi estimé que 28.000 Data Protection Officers devraient être nécessaires rien qu'en Europe et aux Etats-Unis.[9]

La nature et le nombre de(s) données collectées ainsi que les traitements parfois très différents effectués sur ces données rendent l'uniformisation de l'application informatique du RGPD encore plus difficile.

1.3 Le RGPD et l'informatique

Le droit européen a accordé une attention particulière à la notion de vie privée dans la conception du RGPD qui vise à définir le traitement légal des données personnelles des utilisateurs et des utilisatrices. Dans ce texte, la vie privée est définie par l'intermédiaire de la définition des données personnelles qui

sont "*des informations se rapportant à une personne vivante identifiée ou identifiable. Différentes informations, dont le regroupement permet d'identifier une personne en particulier, constituent également des données à caractère personnel.*"[10] ainsi que les données "pseudonymisées" qui peuvent être reliées à une personne particulière. Le RGPD ne distingue pas les données personnelles provenant de la vie professionnelle et celles provenant de la vie personnelle.[8] Le RGPD ne s'applique par contre pas aux données des personnes qui ne sont plus en vie ou à celles des personnes morales.[11] Nous explorerons dans ce mémoire la signification de ces termes.

Du point de vue informatique, cette définition va au-delà des moyens techniques qui sont utilisés pour traiter les données. Cela demande notamment aux entreprises de recenser toute donnée qu'elles veulent traiter et de définir précisément quels sont leurs traitements et quels sont les buts de ces traitements. L'exigence porte aussi sur l'information du public à qui appartient ces données et son consentement pour ces traitements. Les entreprises devront de plus s'assurer de sécuriser les données collectées, de ne les garder que le temps légalement permis, de garder les documents qui sont en rapport avec les traitements effectués et de vérifier que leurs sous-traitants se conforment aussi aux exigences du RGPD. D'autres exigences plus précises peuvent éventuellement concerner une entreprise en fonction de divers facteurs, tel le nombre d'employés ou le type d'entreprise, par exemple.[12]

1.4 Le mémoire

Ce mémoire va s'intéresser à l'application du RGPD, cela signifie que la réflexion a lieu avant l'utilisation du système informatique de la PME ou de l'ASBL, au moment de la création ou de la modification du système. Il n'est pas possible de couvrir tout le RGPD et tous les cas possibles dans un mémoire. Cela explique que ce mémoire se concentre principalement sur l'utilisateur et la manière dont le système devait être capable de répondre à ses droits. Le contexte est ici national, sans transferts internationaux ou appel à des sous-traitants. Le mémoire s'intéresse principalement à des PME ou des ASBL qui gérent un nombre restreint d'utilisateurs mais tend à proposer le début d'une démarche qui pourrait, si la recherche était poussée plus loin, être appliquée à un plus grand nombre d'organisations.

Le mémoire commencera dans le chapitre 2 par présenter comment différentes techniques informatiques liées de près ou de loin à l'analyse des flux de données ont pu être utilisées pour s'intéresser au RGPD et à son application. Le détail des articles qui vont être effectivement traités et une brève discussion de leur pertinence face à la délimitation du sujet seront présentés dans le chapitre 3. Le chapitre 4 présentera la réflexion théorique basée sur l'une des techniques présentées dans le chapitre 2 et en utilisant le sous-ensemble du RGPD présenté dans le chapitre 3. Le chapitre 5 présentera ensuite comment la réflexion théorique pourrait être appliquée dans un contexte pratique et tirera des conclusions quant à certaines améliorations qui pourraient être apportées à la méthodologie développée dans le chapitre 4. Le chapitre 6 concluera le mémoire en discutant ce qui a été accompli et en mentionnant comment la réflexion pourrait être continuée dans d'autres travaux.

1.5 Disclaimer

L'autrice de ce mémoire n'est pas une spécialiste en droit mais en informatique. Ce mémoire n'a pas pour objectif de prouver ou non le respect du RGPD par les PME. L'autrice ne peut pas être tenue responsable pour l'utilisation faite a posteriori de ce mémoire.

Chapitre 2

État de l'Art

2.1 Introduction

Les recherches concernant l'application du RGPD de manière informatique et principalement vis-à-vis de l'analyse des flux de données d'une entreprise sont assez récentes étant donné que le RGPD est entré en application en mai 2018, moins d'un an avant la rédaction de cet état de l'art.

Aborder un sujet qui concerne le RGPD n'est pas chose aisée, il s'agit en effet d'un texte de loi dont l'interprétation et la mise en oeuvre ont commencé relativement récemment. De ce fait, l'Autorité de Protection des Données (APD) belge explique ne pas encore être capable de répondre à toutes les questions qui peuvent se présenter alors qu'elle est notamment composée d'experts juridiques.[13]

Les audits automatiques dans le cadre du RGPD ne sont pas encore très répandus. Les ordinateurs ne peuvent pas déterminer si les données sont nécessaires ou non à un traitement particulier ou comprendre quelle est la nature d'une donnée particulière (publicité ou propagande ?). Il s'agit d'une question de sens et non de technique. La question du but de la récolte de l'information est aussi une question qui existe au-delà de l'ordinateur et de la programmation pure. Une solution pourrait donc être des audits seulement soutenus par ordinateur et non totalement automatisés.[14]

Le travail effectué sur le RGPD est aussi compliqué par le fait qu'il s'agit d'un texte de loi qui doit être analysé et modélisé pour être intégré à une approche informatique. La plupart des analyses restent manuelles mais il existe quelques méthodes d'analyses (partiellement) automatisées. L'une d'entre elles, la méthode "ConRelMiner" a pour but d'être capable de dégager la structure d'un document et de mettre en évidence des phrases similaires pour aider le lecteur à appréhender le texte. Cette méthode cherche aussi à repérer d'éventuels changements entre différentes versions d'un texte de loi et à regrouper des phrases similaires sous un sujet commun. [15]

La suite de cet état de l'art présentera d'autres méthodes déjà existantes qui se sont penchées sur l'application du RGPD et sa traduction en informatique.

Il existe différents types de méthodes. Les méthodes explicitées ici sont toutes liées de près ou de loin à l'analyse du flux de données dans le cadre du RGPD. L'état de l'art se concentre d'abord sur l'analyse de flux de données par l'intermédiaire des *next generation firewalls* et des *network intrusion detection systems*. Puis, le cas de l'analyse statique est abordé. Ensuite, l'utilisation de la modélisation business est présentée. Par après, le cas de la modélisation STS est rapidement exposé. Enfin, LINDDUN une méthode de *threat modelling* est développée. Il s'agit de la méthode utilisée dans ce mémoire.

L'état de l'art s'intéresse aux recherches portant sur les méthodes liées à l'application du RGPD. L'idée est ici d'adopter une vue panoramique de ce qui a déjà été fait pour aider à appliquer le RGPD. Cet état de l'art va donc adopter différents points de vue, certains d'entre eux de haut niveau et d'autres de plus bas niveau. Cela permettra une vue d'ensemble des recherches qui ont déjà eu lieu pour aider à appliquer le RGPD à des systèmes d'information. Les méthodes présentées dans la suite mettent en

avant les données et la manière dont elles peuvent être observées dans le but d'appliquer le RGPD. Ces méthodes utilisent des approches différentes pour atteindre leur objectifs qui sont tous liés de près ou de loin au RGPD et à son application.

Le tableau suivant synthétise les idées et méthodes qui seront développées dans cet état de l'art.

Approche adoptée \ Nom de la méthode	ConRelMiner	Firewalls et NIDS	Analyse statique	Archimate	ODRL	Labnaf	Modélisation STS (étendue au RGPD)	Treat modelling (LINDDUN)
Modéliser le texte de loi	X			X				
Suivre le flux de données d'une entreprise		X (en s'assurant que les données la quittant sont des données autorisées à la quitter au sens du RGPD)						X (Le mémoire va étendre la réflexion pour englober le RGPD)
Evaluer la conformité			X (en repérant les fuites de données éventuelles)	X				
Gérer les changements dans le consentement de l'utilisateur					X			
Modéliser l'entreprise elle-même				X (pour localiser les données personnelles et les accès à ces données)		X	X (L'attention est portée sur le caractère social de l'entreprise et les intentions des acteurs)	

Figure 2.1: Tableau de présentation de l'état de l'art

2.2 Analyse de flux de données

2.2.1 Firewalls et logiciels NIDS

Un moyen de suivre les flux de données d'une entreprise consiste à passer par l'utilisation d'un firewall. Ces logiciels capables de monitorer le trafic sur un réseau et de prendre des mesures si nécessaires peuvent aussi, dans le cas de firewalls déjà avancés, permettre d'inspecter le réseau avec précision. Parmi les firewalls existants, certains firewalls professionnels tels que ceux de Cisco [16], Juniper [17] ou encore Palo Alto [18] répondent à ce dernier critère. Il s'agit de firewalls faisant partie des "next generation" firewalls (NGFW). Ces firewalls se distinguent notamment de firewalls classiques parce qu'il permettent une analyse beaucoup plus détaillée des transferts d'information. Ils permettent ainsi de mener une inspection au niveau applicatif, d'aider à éviter des intrusions et d'aller chercher et de se procurer de l'information qui provient de l'extérieur. [19]

Snort est un logiciel open-source de "Network Intrusion Detection System" (NIDS) [20]

2.2.2 Firewalls, NIDS et le RGPD

Peu de recherches semblent avoir été menées à ce jour au niveau de l'utilisation de firewalls ou de NIDS de cet acabit dans le cadre de l'application du RGPD. Elles observent, entre autres, que les firewalls peuvent notamment aider l'entreprise à s'assurer que les données qui quittent l'entreprise soient des données autorisées par le RGPD. Ainsi, AlgoSec se propose de vérifier que les mesures prises pour assurer la compliance face au RGPD sont bien suffisantes et s'intéresse pour cela aux règles établies pour les firewalls, mais cela est mis directement en rapport avec la sécurité réseau de l'entreprise et non avec le suivi des données traitées. [21]

Ce genre de technologie permet de plus de mieux sécuriser les données ou de détecter de possibles brèches dans la sécurité de l'entreprise, ce qui correspond à l'une des exigences du RGPD. Elles sont donc utiles à la mise en conformité au RGPD.[22]

Dans le cadre du RGPD, on peut faire l'hypothèse qu'il serait possible d'utiliser les firewalls plus avancés comme un moyen de suivre le déplacement des données à caractère personnel sur le réseau d'une entreprise, même si pour le moment peu de recherches semblent avoir exploré cette alternative. Les entreprises proposant les firewalls plus avancés, comme ceux qui ont été cités plus haut, ne semblent pas non plus prendre en compte cette utilisation particulière de leurs produits pour le moment. [22]

2.3 Analyse statique

2.3.1 Analyse statique et RGPD

Une recherche de l'"Università di Verona"[23] s'est déjà penchée sur la possibilité d'utiliser une analyse statique dans l'évaluation de la conformité au RGPD. L'analyse proposée génère un rapport en combinant de la *taint analysis* [1] avec un algorithme de *slicing* [2] qui est exécuté depuis les données repérées par la taint analysis pour retrouver les endroits du programme sujets à un risque de fuite de données. L'analyse vie privée cherche ici principalement à repérer des fuites de données au moment de la compilation. Ces fuites de données représenteraient un non-respect du RGPD et risqueraient d'amener à des brèches de sécurité une fois le système déployé. Le but de l'étude est de trouver comment l'analyse statique peut servir à une analyse concernant la vie privée. Dans une approche de *privacy by design*,

[1] La taint analysis est une méthode d'analyse qui exécute une analyse sur le flux d'informations provenant d'applications "untrusted" pour déterminer si les informations reçues peuvent être utilisées dans des opérations sensibles sans représenter une menace pour la sécurité du système.[24]

[2] Les algorithmes de slicing permettent de sélectionner un point p et une variable x à partir desquels on sélectionne les instructions du programme qui peuvent avoir une influence sur la valeur de la variable x au point p du programme.[25].

elle permet de recenser les différentes analyses statiques appliquées dans une optique de respect de la vie privée. Cette étude s'intéresse aussi de très près à la nature des informations suivies dans la démarche.[23]

Le suivi de ces informations permet le signalement des fuites de données aux 4 rôles distingués dans cette étude. Ces rôles sont ceux des "*Data Protection Officer*" (DPO), "*Chief information security officers*", "*project managers*" et "*développers*". Chacun de ces rôles ayant besoin d'une vision plus précise de la fuite de données potentielle que le rôle précédent pour pouvoir y remédier efficacement. Le DPO a ainsi besoin d'une vision très haut niveau alors que le développeur a besoin d'une vision précise du flux de données amenant à un fuite de données.[23]

Les analyses de vie privée existantes ont la capacité de signaler que des informations pourraient fuiter à un certain point du programme. Il serait intéressant de connaître précisément la nature des informations concernées et le détail du workflow qui pourraient mener à une fuite. Dans le cadre du RGPD, la constitution des informations à destination du développeur consiste à retrouver, à l'aide d'un algorithme de slicing, le flux précis de données considérées comme sensibles au départ de la taint analysis qui a eu lieu. La recherche de l'"Università di Verona" s'est ici concentrée sur la récupération des données précises nécessaires au développeur et à l'abstraction par étape de ces données pour arriver à la vision de très haut niveau nécessaire au DPO. [23]

Il existe aussi des solutions commerciales qui peuvent aider à prévenir les fuites de données (le "*Data Leak Prevention*""). Ces solutions peuvent notamment permettre de s'intéresser aux données lorsqu'elles sont stockées, lorsqu'elles sont transférées sur le réseau et lorsqu'elles sont utilisées.[26]

2.4 Modélisations business

2.4.1 ArchiMate

Le langage de modélisation d'architecture d'entreprise ArchiMate permet de faire des modélisations business. La version d'Archimate présentée ici est la version 3.

Il repose sur 3 niveaux qui sont le niveau "Business" (qui s'intéresse aux relations business qui sont proposées aux clients), le niveau "Application" (qui concerne ce qui soutient le travail et les applications qui le rendent possibles) et le niveau "Technology" (qui s'intéresse aux technologies existantes comme le stockage, le processing, et la communication entre composants). Ces 3 niveaux sont une spécialisation des relations qui sont représentées par ce langage de modélisation et qui lient les éléments génériques avec comme relation la plus proéminente, le "service".[27]

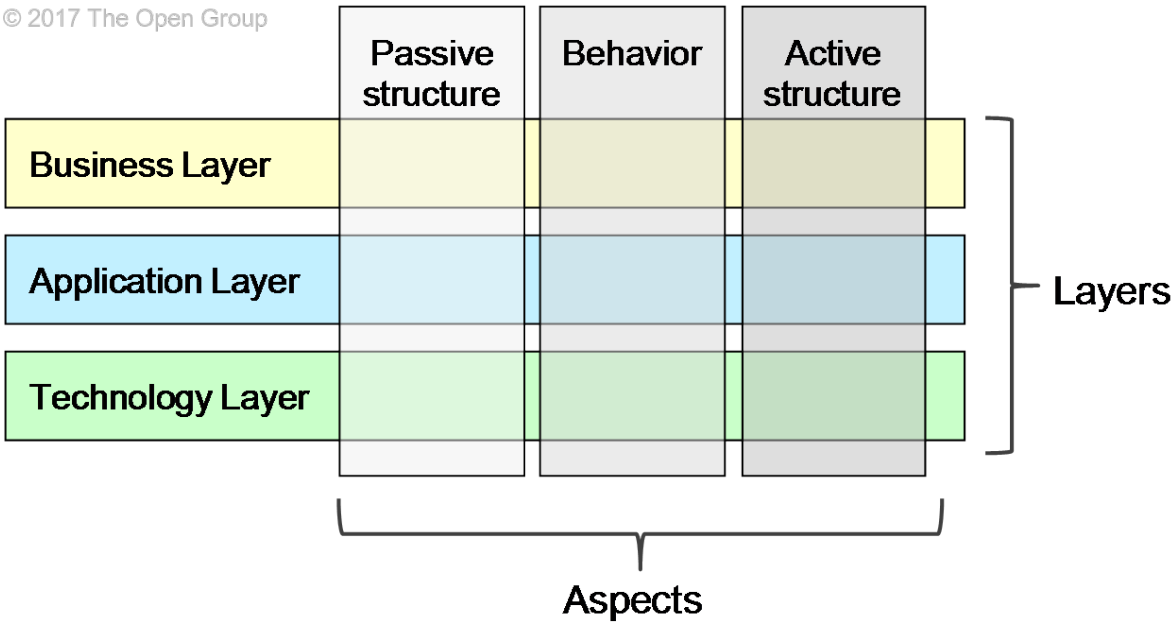


Figure 2.2: Le core framework d'Archimate

Ces trois niveaux sont traversés par trois types d'aspects identifiés et reconnus par ce langage de modélisation : la structure active, la structure passive et le comportement. La structure active contient les éléments de structure qui sont les sujets, la source, d'un ou de plusieurs comportement(s). La structure passive, quant à elle, recouvre tous les objets de l'action des éléments de structure regroupés dans la structure active. Le comportement concerne la manière dont agissent les éléments du système qui sont présentés dans la structure active sur les éléments qui se trouvent dans la structure passive. ArchiMate ne définit pas d'espace spécifique qui regrouperait les éléments de type "information", cela est dû au fait qu'ils peuvent déjà trouver une place dans la structure passive.[27]

La frontière entre les aspects et les niveaux n'est pas une frontière nette à cause de la nature de ce qui doit être modélisé. Il est effectivement possible que des éléments se trouvent à l'intersection de 2 aspects et/ou de 2 niveaux. On appelle, d'ailleurs, élément composite, un élément qui se trouve entre 2 aspects du framework présenté ici. Il s'agit du "core framework", ArchiMate est aussi présenté sous forme d'un framework plus complet qui ajoute des éléments dont un aspect de "motivation" et des niveaux "stratégie", "physique" et "implémentation & migration".[27]

Le logiciel open-source Archi permet une modélisation efficace respectant ce langage de modélisation. Il a la particularité d'être un logiciel cross-platform, open-source et permettant une modélisation basée sur Archimate 3.[28] Il s'agit aussi d'un outil utilisé dans des conditions réelles pour des modélisations d'architecture d'entreprise.[29]

2.4.2 Archimate et le RGPD

La modélisation business appliquée au RGPD et les modélisations concernant la conformité au RGPD ont déjà commencé a être explorées. Ainsi ODRL[3] a été utilisé dans le but de modéliser les changements dans le consentement des utilisateurs au sens du RGPD.[31]

ArchiMate peut aussi se révéler intéressant dans le cadre de l'analyse des flux de données dans une entreprise et être appliqué dans une étude de l'application du RGPD.

En effet, certaines recherches ont déjà exploré la manière dont Archimate pouvait être utilisé pour

[3]L'Open Digital Rights Language (ODRL) permet de modéliser des politiques d'usage d'un contenu ou d'un service.[30]

modéliser le RGPD.[4]
Par exemple, il existe une recherche sur la manière dont ArchiMate pouvait être utilisé pour modéliser les exigences du RGPD dans le contexte de la procédure d'acréditation des hôpitaux. Cet article scientifique commence par faire une brève présentation du RGPD et de la manière dont il va demander aux entreprises de revoir leur business model, ainsi qu'une rapide introduction au langage d'Archimate et au standard TOGAF pour poser les bases de la manière dont il sera utilisé dans le reste de l'article. [5] Il est néanmoins intéressant de mentionner que le choix de se concentrer sur ce type de processus provient du fait que dans les textes régissant les accréditations et les aspects légaux des hôpitaux, il est possible de trouver une estimation assez précise du nombre de processus mis en place dans l'entreprise. Il aborde le fait que les principes qui doivent régner sur le traitement des données sont : "*lawfulness, fairness, transparency, purpose limitation, data minimization, accuracy, storage limitation, integrity and confidentiality, and finally accountability.*"[32]

L'idée est de réussir à modéliser les processus repérés pour permettre de pouvoir repenser les processus et trouver où sont stockées les données à caractère personnel, où elles sont traitées, par qui elles le sont et qui a l'accréditation nécessaire pour les consulter. Cette recherche particulière met en exergue que les tâches à réaliser (tels que le respect de la vie privée, le management du risque et le management de la sécurité, pour reprendre les exemples de l'article) peuvent l'être en s'appuyant sur la liste de processus qui concernent l'hôpital.[32]

Un mémoire de fin de Master, écrit par Charlie Palmér pour la KTH en Suède se concentre sur la modélisation du RGPD grâce à l'usage de l'architecture d'entreprise. Il a la particularité de s'intéresser à la manière dont cette modélisation peut servir pour prendre la mesure du respect des règles du RGPD. Ce mémoire utilise la version 2.1 du langage de modélisation d'Archimate. [33]

Le mémoire de Charlie Palmér commence par définir ses 2 principaux buts qui, s'ils ont des points communs avec ceux de ce mémoire, proposent un angle d'approche différent. En effet, son mémoire passe aussi par une phase d'analyse du RGPD qui se concentre sur des articles qui ont un lien évident avec l'aspect business des entreprises. Le nombre d'articles slectionnés a encore été réduit afin de créer un modèle de conformité plus étendu. Les articles ont tout d'abord été présentés et leur choix motivé avant d'être modélisés. Le modèle qui en est ressorti a ensuite été appliqué à un cas concret pour qu'il puisse être testé.[33]

Certains outils payants permettent une modélisation en lien avec le RGPD notamment en utilisant ArchiMate tel que ce qui est proposé par IServer, qui offre un outil pour aider à la mise en conformité avec le RGPD (appelé "*GDPR Accelerator*").[34]

2.4.3 Labnaf

Labnaf est un langage de modélisation mis en place par A. De Preter, qui a pour objectif de fournir une solution unique qui rassemble de nombreux langages de modélisation ainsi que plusieurs standards. Parmi ceux-ci on compte notamment ArchiMate, UML et TOGAF.[35] Le but est d'uniformiser la démarche et de proposer un unique outil qui puisse prendre de multiples paramètres en compte et aider à modéliser, par exemple, une évolution de l'entreprise et créer des architectures du point de vue informatique ou business. Ce langage est implémenté comme une extension de Sparx, un logiciel qui se concentre sur l'architecture d'entreprise. [36] Labnaf présente l'information et la modélisation en utilisant un système de hiérarchie et en présentant une multitude de points de vue ainsi que les influences qui les relient de manière lisible et compréhensible pour diverses parties prenantes. Le métamodèle ainsi créé peut ensuite servir plusieurs objectifs comme faire de la validation automatique de modèles.[37]

[4] Il existe aussi beaucoup de recherches s'intéressant à la modélisation de textes légaux, cependant, comme ce mémoire s'intéresse en priorité au RGPD, les recherches et travaux suivants se sont principalement penchés sur le cas de ce texte particulier.

[5] Cet article présente aussi la modélisation des processus d'accréditation ayant lieu dans les hôpitaux, ce qui ne se rapporte pas à ce mémoire et ne sera donc pas résumé ici.

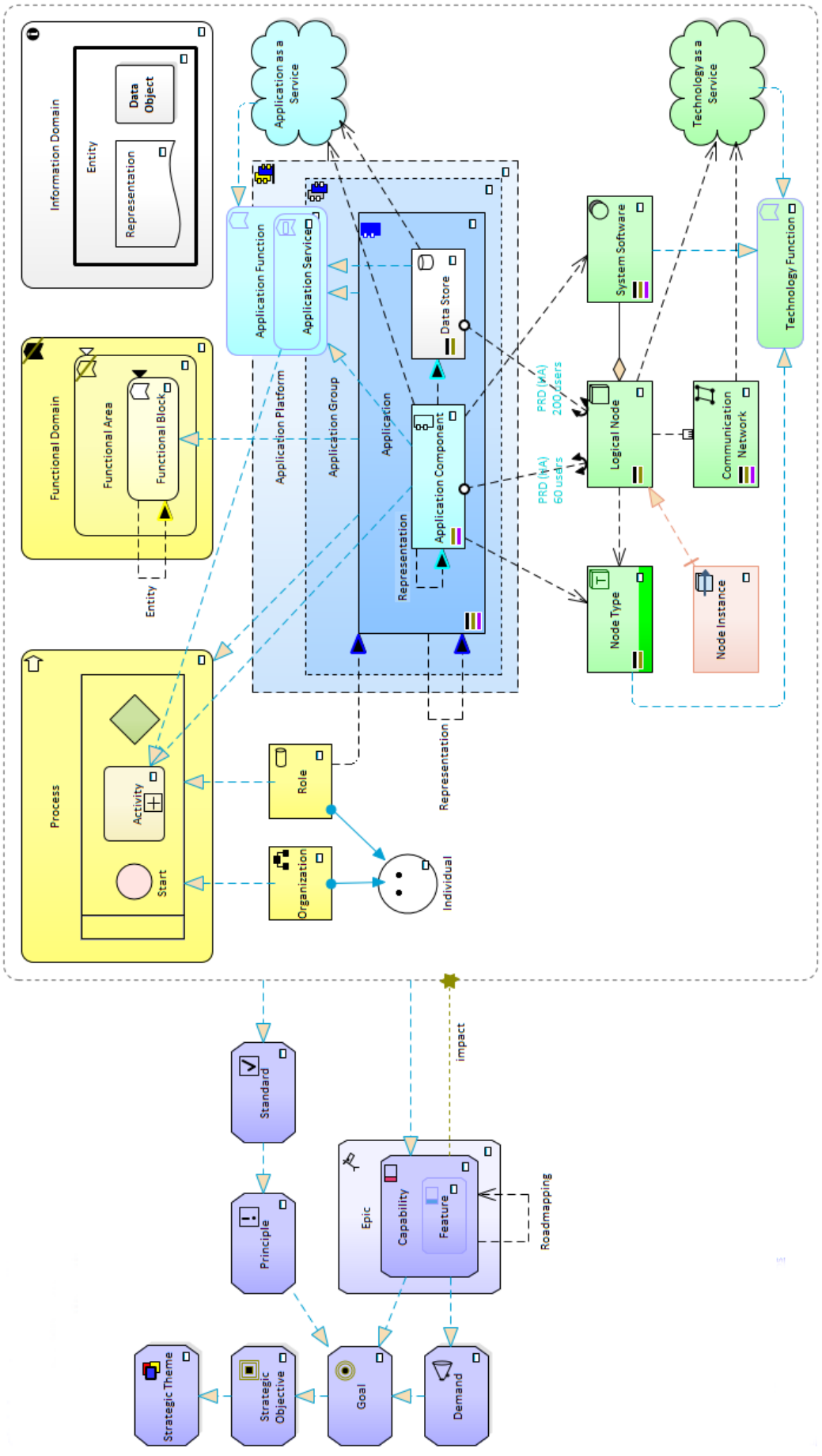


Figure 2.3: Vue d'ensemble du langage de modélisation Labnaf[1]

2.4.4 Labnaf et le RGPD

Aucune recherche ne semble encore s'être consacrée à l'utilisation de Labnaf dans l'optique du RGPD à l'heure où est écrit ce mémoire. Cependant, le site officiel de labnaf fait mention du RGPD dans une des images de présentation qui défilent sur la page d'accueil.[35] Il est donc possible que des recherches tentent d'utiliser ce langage pour proposer une solution pouvant s'appliquer à plusieurs entreprises existent.

2.5 Modélisation STS

2.5.1 Modélisation Socio-Technical Security (STS) adaptée au RGPD

Un grand nombre d'entreprises socio-techniques, c'est-à-dire qui mettent en relation des personnes physiques et des systèmes techniques par le biais d'échanges informationnels sont concernées par l'application du RGPD. La modélisation STS part de cette observation.[38]

La STS met à disposition un langage appelé STS-ml qui permet la modélisation de relations entre acteurs avec une attention particulière portée sur le caractère intentionnel des acteurs. Ces acteurs se trouvent dans la "*social view*" du modèle. Chacun d'entre eux possède des buts définis ("*goals*") ainsi que des ressources mises à leur disposition ("*documents*") et ils reçoivent tous des "rôles" et peuvent être considérés comme des "agents" (c'est-à-dire des acteurs déjà présents lors du design du système). Une autre vision présentée dans ce type de modèle est l'"*information view*" qui regroupe les données concernant ces informations (telles que le nom de leur propriétaire, par exemple). La dernière vision est l'"*authorization view*" qui présente les actions et les autorisations relatives à ces actions.[38]

La recherche a donc décidé de mettre en avant un framework qui vise à aider à mettre en place de nouveaux systèmes d'information qui respectent le RGPD. Leur méthode se base sur la STS (qui s'intéresse principalement, et même exclusivement, aux questions de confidentialité) en lui ajoutant les notions de "*personal data*", "*user consent*", "*legal basis*", ainsi que les notions de "*data controller*" et de "*data processor*" au sens du droit, de *"data owner*" et d'"*employment relationship*".[38]

La recherche explique aussi comment une formalisation de ce langage peut permettre de modéliser certaines exigences du RGPD de manière à ce qu'elles puissent être vérifiées en faisant appel à des processus automatiques. Elle présente cette vérification comme pouvant aider mais n'étant pas capable de garantir un respect total et complet du RGPD. Cette formalisation existe sous la forme d'un langage logique. La recherche définit les prédicats et les notions nécessaires à la formalisation des exigences du RGPD tel que, par exemple, le conflit d'autorisations. La recherche illustre ensuite son propos en formalisant les notions d'"*employment*", de "*Data controller*" et de "*legitimation*".[38]

2.6 Threat modelling

Le *threat modelling* est constitué d'un ensemble de techniques qui permettent de mieux comprendre les notions qui entourent les questions de sécurité d'un système. Le but est de développer une compréhension plus complète des différentes manières dont le système peut être attaqué pour pouvoir y remédier. Ces techniques peuvent adopter le point de vue d'un attaquant du système ou d'une personne le défendant. Un certain nombre de techniques existent, comme les *Attack Trees* ou STRIDE[6]. [40]

[6]STRIDE est une méthode de threat modelling mise au point par Microsoft qui se concentre sur la modélisation de menaces liée à la sécurité du système analysé. L'acronyme STRIDE référence les 6 types de menaces mis en avant : *"Spoofing", "Tampering", "Repudiation", "Information disclosure", "Denial of service"* et *"Elevation of privilege"*. [39]

2.6.1 LINDDUN

LINDDUN[41] fait partie de ces techniques de *threat modelling* et se base sur STRIDE. Cette méthode a été développée par K. Wuyts, R. Scandariato, W. Joosen du groupe de recherche Distrinet de la KU Leuven en collaboration avec M. Deng et B. Preneel.[41] LINDDUN a pour but de permettre une modélisation de menaces qui soit orientée entièrement vers la vie privée. Son nom est l'acronyme des types de menaces qu'elle veut prévenir : **Linkability** (qui peut donner l'occasion à quelqu'un de mal-intentionné de découvrir si 2 ou plusieurs composants sont liés dans le système), **Identifiability** (qui peut permettre de relier un objet du système à une personne), **Non-repudiation** (qui est une menace pour la vie privée principalement et qui permet d'établir avec une certitude relative qu'un sujet est l'auteur d'une action ou d'un message), **Detectability** (qui signale qu'un objet peut être reconnu comme existant par un attaquant potentiel), **information Disclosure** (qui concerne la divulgation d'informations caractérisées comme personnelles), **content Unawareness** (qui englobe à la fois le fait qu'un utilisateur peut donner des informations erronées volontairement ou qu'il peut donner un nombre trop conséquent d'informations et risquer dans ce cas de fournir à un attaquant assez de matière pour découvrir son identité.), **policy and content Non-compliance** (qui peut indiquer un non-respect de la politique d'utilisation présentée par le système à ses utilisateurs ou des lois existantes).[41]

Elle se décompose en 6 étapes principales dont les 3 premières sont considérées comme faisant partie de la définition du problème. Les 3 étapes suivantes sont considérées comme faisant partie de la définition de la solution.[41]

• La première est la "*création du data flow diagram*" qui voit le jour grâce à la description du système étudié.[2]

• La deuxième concerne le *"lien entre les éléments du dit-diagramme et des menaces concernant la vie privée"*, ces menaces sont celles qui sont présentées plus haut et qui sont celles sur lesquelles LINDDUN se concentre. Ce lien prend la forme d'un tableau particulier à LINDDUN dont le template a été défini dans la technique elle-même (et est présenté dans la figure 2.4).[2]

	L	I	N	D	D	U	N
Entity	X	X				X	
Data store	X	X	X	X	X		X
Data flow	X	X	X	X	X		X
Process	X	X	X	X	X		X

Figure 2.4: Exemple de tableau faisant le lien entre le data flow diagram et les types de menaces lors de l'utilisation de la méthode LINDDUN

• La troisième est l'*identification des scénarios de menaces"*. Pour ce faire, plusieurs sous-étapes sont nécessaires. Ainsi la technique commence par énumérer les menaces possibles en les présentant sous forme de "*privacy threat tree*" de manière à exprimer la suite d'étapes la plus probable suivie dans un cas d'attaque. Pour aider à ne pas oublier certaines menaces potentielles, LINDDUN propose une liste préexistante d'arbres de menaces. Chaque feuille de l'arbre peut potentiellement être appliquée sur l'un des éléments du data flow diagram. Il est alors nécessaire de repérer ces feuilles et de les inclure dans la liste des menaces potentielles. Pour les décrire, LINDDUN préfère l'utilisation de *misuse cases* à d'autres méthodes existantes.[2]

• La quatrième étape est ainsi la *"priorisation des menaces identifiées"*.[2]

• La cinquième est l'*"identification des stratégies d'atténuation"* qui considère chaque menace potentielle dans l'ordre de priorité préalablement établi.[2]

• Quant à la sixième, il s'agit de la *"sélection des PETs"* (*Privacy-Enhancing Technology*) correspondantes ce qui permet de proposer une solution à chaque menace relevée précédemment. [2]

3 types de solutions sont abordés : le premier type implique de mettre au courant les utilisateurs. Il ne concerne donc que les menaces potentielles dont le risque reste assez peu important, le deuxième propose d'éteindre ou de supprimer la fonctionnalité incriminée pour atteindre un risque nul. Le troisième envisage l'utilisation de *Privacy Enhancing Technologies*s qui agiraient en prévention ou de manière réactive en cas de problème.[41] Ces *Privacy Enhancing Technologies* peuvent par exemple prendre la forme de *"Privacy enhancing managment system"* ou d'un chiffrement à clé publique.[42]

2.6.2 LINDDUN et le RGPD

L'utilisation de LINDDUN dans le cadre du RGPD n'a pas encore fait l'objet d'un grand nombre de recherches. Cela malgré le fait que les 2 sont liés par leur intérêt pour la question de vie privée dans le cadre technologique.

Une recherche de Martín et del Álamo[43] a mis en place un framework potentiellement capable de méta-modéliser toutes les méthodes de privacy engineering. Ce framework est une extension de SEMDM, le métamodèle décrit dans ISO24744 qui a pour sujet la modélisation de méthodologies de développement. Leur recherche cite à la fois LINDDUN et le RGPD comme exemples de sujets à métamodéliser et peut donc potentiellement être utilisé pour faire le lien entre les deux.

Si les termes "GDPR" et "LINDDUN" peuvent se retrouver relativement fréquemment associés dans un article, rares sont ceux qui explicitent l'utilisation de LINDDUN dans une démarche d'application d'un sous-ensemble du RGPD à partir de l'observation du flux de données d'une organisation. Les méthodes de threat modelling, dont LINDDUN fait partie, sont aussi relativement présentes dans les recherches s'intéressant au respect du RGPD d'un point de vue informatique.

Chapitre 3

Le RGPD

3.1 Introduction

Le RGPD est un texte conséquent. Réussir à le traiter dans son intégralité demanderait des recherches plus poussées qu'un mémoire seul ne peut pas couvrir. Il a donc été nécessaire de choisir quelles parties traiter.[1] Par ailleurs, si le mémoire dit s'intéresser à un traitement belge, l'accent est porté sur le fait que le traitement reste dans un cadre purement national mais le texte référencé dans ce mémoire reste le RGPD, pas le détail des modifications que la loi belge lui a apporté.

L'intérêt s'est porté sur la relation système-utilisateur du point de vue informatique[2]. Il s'est aussi axé sur les PME et les ASBL. En effet, le RGPD demande de prendre en compte des éléments vis-à-vis de cet utilisateur, de ce qu'il doit savoir et de ce que le système d'information doit être capable de faire pour respecter ses droits. Le choix de se concentrer sur l'utilisateur permet aussi de répondre à l'un des objectifs vers lesquels ce mémoire veut tendre : la sensibilisation desdits utilisateurs.

Ce choix reste aussi cohérent avec le fait que le mémoire s'intéresse aux flux de données car, comme cela sera illustré plus en détail dans la suite, pour un informaticien cherchant à respecter le RGPD, il est nécessaire de savoir comment les données se déplacent dans le système.

Chaque point sera illustré par des citations issues du RGPD et par des exemples pour plus de lisibilité.

3.2 La mindmap du RGPD

Pour permettre de clarifier quelle est la partie du RGPD qui nous intéresse, commençons par présenter une mindmap des thèmes du RGPD les plus intéressants dans le cadre des PME. Il a été réalisé grâce au vade-mecum à l'intention des PME que l'Autorité de Protection des Données belge (APD dans la suite) [44] a publié. Ce choix comme première référence est dû au fait qu'il correspond aux limites choisies pour ce travail : un contexte belge qui concerne notamment les PME. S'il ne reprend pas toutes les idées développées dans le RGPD, il permet néammoins un large balayage des articles s'appliquant dans le contexte choisi et de créer la mindmap du RGPD qui suit.

[1] Cela vaut aussi pour les lois pouvant s'appliquer en plus du RGPD. Les PME sont en effet soumises à d'autres lois que le RGPD, qui s'intéressent à la légalité de leurs activités dans son ensemble (loi sur le travail, ...). Le RGPD laisse également aux pays membres une certaine latitude en ce qui concerne certains points (possibilité de considérer qu'un enfant de plus de 13 ans peut donner son consentement sans accord parental alors que le RGPD considère que seuls les enfants de plus de 16 ans peuvent faire cela).

[2] Par exemple, comment modifier le système d'information pour répondre à la demande d'un utilisateur faisant appel à son droit à la portabilité des données ?

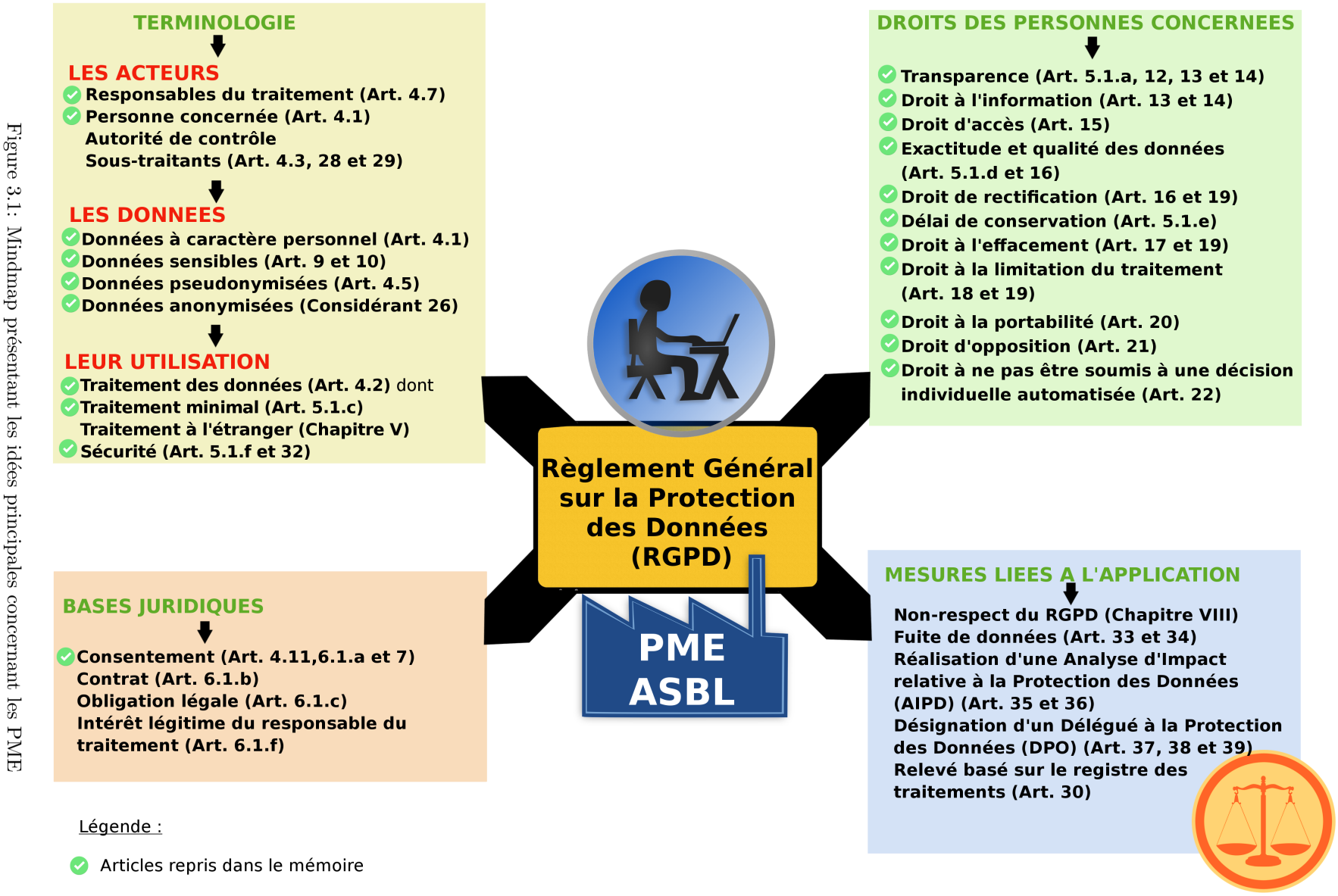


Figure 3.1: Mindmap présentant les idées principales concernant les PME

Chaque point sélectionné, représenté par un "v" vert sur le mindmap, va être présenté plus en détails.

3.3 Terminologie récurrente

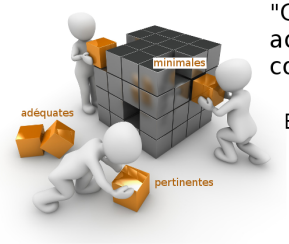
3.3.1 Personne concernée

Personne concernée
Homme, femme
Enfant de plus de 16 ans
Avec accord du titulaire de la responsabilité parentale si moins de 16 ans

Selon le RGPD, la "personne concernée" est une "personne physique[3] identifiable, c'est-à-dire *"une personne physique qui peut être identifiée, directement ou indirectement, notamment par référence à un identifiant, tel qu'un nom, un numéro d'identification, des données de localisation, un identifiant en ligne, ou à un ou plusieurs éléments spécifiques propres à son identité physique, physiologique, génétique, psychique, économique, culturelle ou sociale"* comme l'a défini l'article 4.1.[46]

Il s'agit de la personne (homme, femme, enfant de plus de 16 ans et enfant de moins de 16 ans avec l'accord du titulaire de la responsabilité parentale) qui donne accès à ses données personnelles. Il ne s'agit pas des PME mais de ceux qui utilisent leurs services.

3.3.2 Traitement et traitement minimal



Données : traitement et traitement minimal
"Collecte, enregistrement, organisation, structuration, conservation, adaptation, modification, extraction, consultation, utilisation, communication, diffusion, effacement, destruction. (Art.4.2)

Ex : inscription en ligne pour effectuer des achats, fiches papier d'une PME reprenant l'identité de ses clients (adresse, informations bancaires, téléphone...), gestion des données numériques dans un centre sportif via les cartes magnétiques, inscription obligatoire pour pouvoir faire une activité en ligne etc...

[3] *"Les deux catégories de personnes juridiques*
Les personnes sont des sujets de droit, et en droit, le mot "personne" a un sens plus large que dans le langage courant. On distingue ainsi deux catégories de personnes : les personnes physiques et les personnes morales.
a. Les personnes physiques
Une personne physique est un être humain vivant, sans distinction de sexe, de race, et de religion, conformément au préambule de la Constitution.
b. Les personnes morales
Une personne morale est un groupement d'individus réunis dans un intérêt commun. Par exemple, trois amis se sont associés pour créer une société de services informatiques. Cette société est une personne morale.
On distingue deux sortes de personnes morales :
- *les personnes morales de droit public ;*
- *les personnes morales de droit privé.*

Les personnes morales de droit public regroupent les collectivités publiques (l'Etat, les régions, les départements, les communes), les établissements publics (universités, hôpitaux...).
Les personnes morales de droit privé sont créées par la volonté de certains individus. Cela peut être une société, une association, un syndicat..." [45]

Il est défini à l'article 4.2 tel que "*toute opération ou tout ensemble d'opérations effectuées ou non à l'aide de procédés automatisés et appliquées à des données ou des ensembles de données à caractère personnel, telles que la collecte, l'enregistrement, l'organisation, la structuration, la conservation, l'adaptation ou la modification, l'extraction, la consultation, l'utilisation, la communication par transmission, la diffusion ou toute autre forme de mise à disposition, le rapprochement ou l'interconnexion, la limitation, l'effacement ou la destruction*".[46]

Le traitement sur les données dans le but d'atteindre les finalités présentées doit rester aussi minimal que possible, c'est exprimé dans l'article 5.1.c) qui précise que les données doivent être *"adéquates, pertinentes et limitées à ce qui est nécessaire au regard des finalités pour lesquelles elles sont traitées (minimisation des données);"*[46]

Cela peut donc potentiellement dépasser le champ de l'informatique seul étant donné que tout traitement portant sur des données à caractère personnel est soumis au RGPD que ce traitement soit manuel ou automatique. Cela peut même concerner le traitement sur support papier.

En informatique, principalement dans le cas des traitements automatisés, une intervention humaine va être nécessaire dans une certaine mesure pour déterminer quelles sont les données nécessaires à l'atteinte de la finalité visée. Les ordinateurs font ce pourquoi ils sont programmés et n'ont pas l'intelligence de pouvoir discuter si une donnée est vraiment nécessaire ou non. Si on leur demande de traiter une photo d'identité lorsqu'un utilisateur veut acheter des chaussures en ligne, il serait très étonnant qu'ils commence à protester quant à la nécessité d'une telle action...

De plus, il est possible que certains traitements de données personnelles puissent être réalisés grâce à l'utilisation d'un nombre réduit de données ou avec des données moins "sensibles". Ainsi, l'utilisation d'un numéro de registre national pourrait par exemple être remplacé par un identifiant donné par le système d'information qui n'aurait dès lors plus besoin de garder en mémoire le numéro de registre national pour effectuer ses traitements de manière satisfaisante. Cela va, bien sûr, dépendre du contexte de chaque système, mais la question vaut la peine d'être posée pour trouver comment minimiser les données.

3.3.3 Responsable du traitement

PME en tant que responsable de traitement
La PME (ou l'asbl) qui "détermine les finalités et les moyens de traitement". (Art. 4.7)

Ex : la PME est responsable du traitement des données concernant son personnel et sa clientèle, sa patientèle ou ses membres.....

Le responsable du traitement est aussi un terme qui va être utile dans la suite puisqu'il désigne *"la personne physique ou morale, l'autorité publique, le service ou un autre organisme qui, seul ou conjointement avec d'autres, détermine les finalités et les moyens du traitement; lorsque les finalités et les moyens de ce traitement sont déterminés par le droit de l'Union ou le droit d'un État membre, le responsable du traitement peut être désigné ou les critères spécifiques applicables à sa désignation peuvent être prévus par le droit de l'Union ou par le droit d'un État membre;"* (Article 4.7)[46]

L'informaticien est responsable du traitement informatique des données mais il n'est *a priori* pas responsable de l'utilisation qui est faite de ce traitement informatique sur les données. Les décisions ne sont donc pas prises par l'informaticien seul car il doit agir en accord avec ce qui a été décidé pour l'organisation dans son ensemble et sur conseil du délégué à la protection des données (dont la version anglaise est raccourcie en DPO, terme aussi utilisé par la suite). Cependant, dans le cadre des PME

et des ASBL, il est possible que l'informaticien joue un rôle dans les décisions étant donné la taille de ces entreprises. Ceci pose une question : quel est le rôle éthique et juridique de l'informaticien dans l'application du RGPD ?

3.4 Les différents types de données dans le RGPD

3.4.1 Données à caractère personnel

Donnée à caractère personnel des clients de la PME

"Toute information se rapportant à une personne physique identifiée et identifiable". (Art. 4.1)

Ex : nom, prénom, adresse, téléphone, mail, listes et historiques d'achats, données de localisation, photographies et vidéos, date de naissance, etc ....

La donnée à caractère personnel est définie dans l'article 4.1 comme "*[...] toute information se rapportant à une personne physique identifiée ou identifiable (ci-après dénommée "personne concernée");* ".[46]

En informatique, c'est la catégorie de données qui nous intéresse le plus car la plupart des données qui seront évoquées ici vont rentrer dans cette catégorie particulière.

3.4.2 Données sensibles

Données sensibles

En principe, le traitement en est interdit. Données qui peuvent potentiellement amener des "risques significatifs". (Art. 9 et 10)

Ne sont pas reprises dans ce mémoire

Ex : origine raciale ou ethnique, opinion politique, convictions religieuses ou philosophiques, données génétiques, biométriques, santé, vie et orientation sexuelle..... via l'extrait de casier judiciaire, une carte d'appartenance à un groupe politique, un certificat médical....

La donnée sensible fait partie d'un sous-ensemble de données à caractère personnel qu'il est théoriquement interdit de traiter à moins de se trouver dans l'un des cas d'exception défini dans le RGPD (principalement par les articles 9 et 10). Il s'agit des données qui peuvent potentiellement amener des "*risques significatifs*" si elles font l'objet d'un traitement, ainsi que les données à caractère personnel donnant la possibilité de retrouver les données sensibles.[44]

L'article 9 définit le *"Traitement portant sur des catégories particulières de données à caractère personnel"*[46]. Dans le premier sous-point de cet article, on détaille les données concernées : *"Le traitement des données à caractère personnel qui révèle l'origine raciale ou ethnique, les opinions politiques, les convictions religieuses ou philosophiques ou l'appartenance syndicale, ainsi que le traitement des données génétiques, des données biométriques aux fins d'identifier une personne physique de manière unique, des données concernant la santé ou des données concernant la vie sexuelle ou l'orientation sexuelle d'une personne physique sont interdits."*[46].

L'article 10, quant à lui, définit le *"Traitement des données à caractère personnel relatives aux condamnations pénales et aux infractions"*[46] en précisant qu'il *"ne peut être effectué que sous le contrôle de l'autorité publique, ou si le traitement est autorisé par le droit de l'Union ou par le droit d'un État membre qui prévoit des garanties appropriées pour les droits et libertés des personnes concernées. Tout registre complet des condamnations pénales ne peut être tenu que sous le contrôle de l'autorité publique."*[46]

S'il existe des cas dans lesquels ces données peuvent être traitées (qui peuvent potentiellement différer d'un état membre à l'autre comme précisé par l'article 9.4), elles restent particulières et doivent faire l'objet de mesures particulière.

Ces catégories de données particulières sortent du champ d'investigation de ce mémoire. Les rajouter à la réflexion pourraient être l'une des améliorations futures de la méthodologie mais pour le moment, les données sensibles ne seront pas prises en compte. Cela est aussi dû au fait que si la plupart des entreprises traitent des données rentrant dans le champ des données personnelles, un nombre beaucoup plus restreint d'entre elles s'aventure à traiter les données sensibles. Il a donc été décidé de se concentrer sur l'ensemble des données personnelles en excluant les données sensibles.

3.4.3 Données pseudonymisées

Données pseudonymisées
Ce sont les données qui ne peuvent pas "être attribuées à une personne sans avoir recours à des informations supplémentaires" (Art. 4.5) (comme une clé de chiffrement par exemple)
Ex : les statistiques sur les ventes d'un magasin et les projections sur le nombre de visites la semaine suivante, un numéro d'identification détaché d'une base de données reprenant les données nécessaires à l'identification de la personne

Un autre sous-type de données à caractère personnel sont les données pseudonymisées qui sont abordées dans l'article 4.5) du RGPD : *"'pseudonymisation', le traitement de données à caractère personnel de telle façon que celles-ci ne puissent plus être attribuées à une personne concernée précise sans avoir recours à des informations supplémentaires, pour autant que ces informations supplémentaires soient conservées séparément et soumises à des mesures techniques et organisationnelles afin de garantir que les données à caractère personnel ne sont pas attribuées à une personne physique identifiée ou identifiable;"*[46]

Dans une optique informatique, ce qu'on peut relever principalement de cet article est la nécessité de séparer les données pseudonymisées des données permettant de les retrouver. Les données pseudonymisées demandent donc une attention particulière par rapport aux données anonymisées qui sont abordées ci-dessous car elles ne doivent pas être reliées à une personne physique. Cela peut aussi poser question si de futures données collectées peuvent aider à identifier la personne en question...

3.4.4 Données anonymisées

Données anonymisées

Ce sont les données qu'on ne peut pas relier à une personne particulière. (Considérant 26)

Ex : les données agrégées en statistiques, visites sur les sites qui ne comptent que la visite sans connaître la provenance, questionnaires anonymes etc...

Le dernier type de données personnelles sont les données anonymisées qui sont des données qu'on ne peut pas relier à une personne particulière. Ce type de données est abordé dans le considérant (26) du RGPD qui explique qu'*"il n'y a dès lors pas lieu d'appliquer les principes relatifs à la protection des données aux informations anonymes, à savoir les informations ne concernant pas une personne physique identifiée ou identifiable, ni aux données à caractère personnel rendues anonymes de telle manière que la personne concernée ne soit pas ou plus identifiable. Le présent règlement ne s'applique, par conséquent, pas au traitement de telles informations anonymes, y compris à des fins statistiques ou de recherche."*[46]

Ces données rentrent dans le cadre du mémoire mais restent secondaires étant donné qu'elles ne demandent pas d'appliquer les règles définies par le RGPD.

3.5 Le consentement comme base juridique

Consentement, contrat, respect d'une obligation légale, intérêt légitime.

Chaque PME doit travailler avec la base qui se rapproche le plus de son activité de traitement.

Le mémoire se concentre sur le consentement. Il s'agit de "toute manifestation de volonté, libre, spécifique, éclairée et univoque par laquelle la personne concernée accepte, par une déclaration ou par un acte positif clair, que des données à caractère personnel la concernant fassent l'objet d'un traitement".(Art. 4.11) La personne peut retirer ce consentement à tout moment (Art. 7.3)

Ex : déclaration écrite claire et facile à comprendre par l'utilisateur, le fait de cocher volontairement une case spécifique pour recevoir la newsletter d'un site etc...

Le vade-mecum de l'APD précise que les PME sont principalement concernées par 4 des 6 bases juridiques mises en avant par le RGPD dans l'article 6 : le "*consentement*", le "*contrat*", le "*respect d'une obligation légale*" et l'"*intérêt légitime*". La brochure explique que ces 4 bases sont "*équivalentes*" et que chaque PME doit travailler avec celle qui se rapproche le plus de ses activités de traitement. Chaque base juridique a ses propres spécificités qui peuvent être nombreuses.[44] Le consentement est la seule base juridique qui sera prise en compte ici car il est couramment utilisé pour de nombreux traitements et demande une action directe de l'utilisateur.

Le consentement est *"toute manifestation de volonté, libre, spécifique, éclairée et univoque par laquelle la personne concernée accepte, par une déclaration ou par un acte positif clair, que des données à caractère personnel la concernant fassent l'objet d'un traitement;"*[46] selon l'article 4.11 du RGPD.

L'article 7 détaille les *"Conditions applicables au consentement"*[46] : *"le responsable du traitement est en mesure de démontrer que la personne concernée a donné son consentement au traitement de*

données à caractère personnel la concernant." (article 7.1), *"Si le consentement de la personne concernée est donné dans le cadre d'une déclaration écrite qui concerne également d'autres questions, la demande de consentement est présentée sous une forme qui la distingue clairement de ces autres questions, sous une forme compréhensible et aisément accessible, et formulée en des termes clairs et simples. Aucune partie de cette déclaration qui constitue une violation du présent règlement n'est contraignante."* (article 7.2), *"La personne concernée a le droit de retirer son consentement à tout moment. Le retrait du consentement ne compromet pas la licéité du traitement fondé sur le consentement effectué avant ce retrait. La personne concernée en est informée avant de donner son consentement. Il est aussi simple de retirer que de donner son consentement."* (article 7.3) et *"Au moment de déterminer si le consentement est donné librement, il y a lieu de tenir le plus grand compte de la question de savoir, entre autres, si l'exécution d'un contrat, y compris la fourniture d'un service, est subordonnée au consentement au traitement de données à caractère personnel qui n'est pas nécessaire à l'exécution dudit contrat."* (article 7.4)[46].[4]

Au niveau de l'informatique, le système d'information doit donc mettre en place des mesures particulières lors de la demande de consentement, par exemple en la mettant en exergue par rapport à d'autres considérations pour aider l'utilisateur à la voir. Cette mise en exergue doit continuer dans la période qui suit cette demande de consentement (en permettant à l'utilisateur de le retirer à tout moment ou au responsable du traitement de prouver que le consentement a bien été reçu).

Il est cependant à noter que la question du consentement dépasse le simple cadre informatique et pourrait faire l'objet d'un travail plus poussé concentré sur cette seule notion. Par la suite, seul le contenu des articles présentés dans cette partie sera pris en compte.

3.6 Les finalités des traitements et leur compatibilité

Finalités des traitements et leur compatibilité

Les données doivent être "collectées pour des finalités déterminées, explicites et légitimes et ne pas être traitées ultérieurement d'une manière incompatible avec ces finalités". (Art. 5.1.b)

Ex : si l'utilisateur donne son adresse mail pour recevoir la newsletter, il ne faut pas revendre cette adresse à des organismes de publicité

Les finalités sont présentées par l'article 5.1.b) qui précise que les données doivent être *"collectées pour des finalités déterminées, explicites et légitimes, et ne pas être traitées ultérieurement d'une manière incompatible avec ces finalités; le traitement ultérieur à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques n'est pas considéré, conformément à l'article 89, paragraphe 1, comme incompatible avec les finalités initiales (limitation des finalités);"*[46].

La compatibilité entre deux finalités différentes peut aussi être évaluée par le ou la responsable du traitement. La marche à suivre est présentée dans l'article 6.4.[5]

[4] Notons l'existence de l'article 8 qui précise les règles spécifiques au consentement des enfants : *"en ce qui concerne l'offre directe de services de la société de l'information aux enfants, le traitement des données à caractère personnel relatives à un enfant est licite lorsque l'enfant est âgé d'au moins 16 ans. Lorsque l'enfant est âgé de moins de 16 ans, ce traitement n'est licite que si, et dans la mesure où, le consentement est donné ou autorisé par le titulaire de la responsabilité parentale à l'égard de l'enfant. Les États membres peuvent prévoir par la loi un âge inférieur pour ces finalités pour autant que cet âge inférieur ne soit pas en-dessous de 13 ans."* (article 8.1)[46], *"Le responsable du traitement s'efforce raisonnablement de vérifier, en pareil cas, que le consentement est donné ou autorisé par le titulaire de la responsabilité parentale à l'égard de l'enfant, compte tenu des moyens technologiques disponibles."* (article 8.2)[46]. Le sous-point 3 de l'article 8 précise aussi que le RGPD ne va pas à l'encontre des lois de ses états membres en ce qui concerne des situations comme les contrats passés avec des enfants.

[5] *"Lorsque le traitement à une fin autre que celle pour laquelle les données ont été collectées n'est pas fondé sur le consentement de la personne concernée ou sur le droit de l'Union ou le droit d'un État membre qui constitue une mesure nécessaire et proportionnée dans une société démocratique pour garantir les objectifs visés à l'article 23, paragraphe 1, le responsable du traitement, afin de déterminer si le traitement à une autre fin est compatible avec la finalité pour laquelle*

Cela apprend à l'informaticien que chacun des traitements conçu doit être relié à une finalité préalabement définie. L'informaticien n'est pas compétent pour être le garant des exigences du RGPD et de l'utilisation des données qui en est faite. Il doit être attentif à garder cela en mémoire et à faire appel au DPO lorsque cela est nécessaire.

3.7 Sécurité des données

Sécurité des données

La PME doit traiter les données de façon à garantir une sécurité appropriée des données à caractère personnel, tant du point de vue organisationnel que technique (Art. 5.1.f et Art. 32)

Ex : contrôle d'accès aux données, formation sécurité pour le personnel qui a accès à ces données, chiffrement des données, accès aux données avec identifiant individuel etc...

Le RGPD met en avant la nécessité d'assurer la sécurité des données récoltées. Ici, la mention faite dans l'article 5.1.f et dans l'article 32 sera évoquée.

Ces 2 articles précisent qu'il faut que les données soient *"traitées de façon à garantir une sécurité appropriée des données à caractère personnel, y compris la protection contre le traitement non autorisé ou illicite et contre la perte, la destruction ou les dégâts d'origine accidentelle, à l'aide de mesures techniques ou organisationnelles appropriées (intégrité et confidentialité);"* (article 5.1.f)[46] et que *"1.Compte tenu de l'état des connaissances, des coûts de mise en œuvre et de la nature, de la portée, du contexte et des finalités du traitement ainsi que des risques, dont le degré de probabilité et de gravité varie, pour les droits et libertés des personnes physiques, le responsable du traitement et le sous-traitant mettent en œuvre les mesures techniques et organisationnelles appropriées afin de garantir un niveau de sécurité adapté au risque, y compris entre autres, selon les besoins:*
a) la pseudonymisation et le chiffrement des données à caractère personnel;
b) des moyens permettant de garantir la confidentialité, l'intégrité, la disponibilité et la résilience constantes des systèmes et des services de traitement;
c) des moyens permettant de rétablir la disponibilité des données à caractère personnel et l'accès à celles-ci dans des délais appropriés en cas d'incident physique ou technique;
d) une procédure visant à tester, à analyser et à évaluer régulièrement l'efficacité des mesures techniques et organisationnelles pour assurer la sécurité du traitement.
2. Lors de l'évaluation du niveau de sécurité approprié, il est tenu compte en particulier des risques que présente le traitement, résultant notamment de la destruction, de la perte, de l'altération, de la divulgation non autorisée de données à caractère personnel transmises, conservées ou traitées d'une autre manière, ou de l'accès non autorisé à de telles données, de manière accidentelle ou illicite. [...]
4. Le responsable du traitement et le sous-traitant prennent des mesures afin de garantir que toute personne physique agissant sous l'autorité du responsable du traitement ou sous celle du sous-traitant, qui a accès à des données à caractère personnel, ne les traite pas, excepté sur instruction du responsable du traitement, à moins d'y être obligée par le droit de l'Union ou le droit d'un État membre."[46]

Le paragraphe 3 de l'article 32 n'a pas été repris car il mentionne le fait qu'un "code de conduite approuvé" ou qu'un "mécanisme de certification approuvé" peut aider à prouver le respect du RGPD. Cela

les données à caractère personnel ont été initialement collectées, tient compte, entre autres:
a) de l'existence éventuelle d'un lien entre les finalités pour lesquelles les données à caractère personnel ont été collectées et les finalités du traitement ultérieur envisagé;
b) du contexte dans lequel les données à caractère personnel ont été collectées, en particulier en ce qui concerne la relation entre les personnes concernées et le responsable du traitement;
c) de la nature des données à caractère personnel, en particulier si le traitement porte sur des catégories particulières de données à caractère personnel, en vertu de l'article 9, ou si des données à caractère personnel relatives à des condamnations pénales et à des infractions sont traitées, en vertu de l'article 10;
d) des conséquences possibles du traitement ultérieur envisagé pour les personnes concernées;
e) de l'existence de garanties appropriées, qui peuvent comprendre le chiffrement ou la pseudonymisation."[46]

sort du cadre de ce mémoire étant donné que celui-ci s'intéresse à la phase de création ou d'adaptation du système d'information.

La necessité de sécuriser les traitements et les données conservées n'est pas nouvelle **en informatique**. Le RGPD précise clairement les contours de ce qu'il attend. Une certaine liberté est laissée à l'informaticien dans le choix des techniques utilisées.

Le RGPD insiste également dans cet article sur le fait que le "*niveau de sécurité*" doit être "*adapté au risque*", cela peut faire appel à des méthodes d'analyse des risques, dont le threat modelling fait partie.

3.8 Les droits des personnes concernées

Les articles suivants vont contenir des informations pertinentes quant aux **mesures informatiques** à mettre en place pour permettre aux utilisateurs de faire appel à leurs droits.

L'article 12[46] parle des modalités qui doivent être respectées lors de l'application des droits. Si certaines parties de cet article peuvent avoir un impact au point de vue informatique, il concerne principalement un point plus en aval de celui auquel s'intéresse ce mémoire.

L'une des choses qui impacte de manière importante l'informatique dans l'application de ces droits va notamment être de prévoir tous les cas possibles pour permettre à chaque utilisateur de choisir ou non de faire appel à l'un de ses droits. Tout système va dès lors devoir garder en mémoire les préférences de traitements de chacun de ses utilisateur pour pouvoir personnaliser leurs expériences.

3.8.1 Transparence, droit à l'information et droit d'accès

Transparence, droit à l'information pour la personne concernée, droit d'accès à ses données par la personne concernée
La PME doit traiter les données de façon à garantir à l'utilisateur ces 3 droits.
(Art. 5.1.a, Art. 13, Art. 14, Art. 15)

Ex :
- **transparence** : les informations données à l'utilisateur par la PME doivent être claires et faciles à comprendre. Pas de phrases juridiques ou à l'inverse trop évasives du genre "vous nous autorisez à utiliser vos données".
- **droit à l'information et droit d'accès (voir tableaux suivants)**

• L'article 5.1.a)[46] précise que la transparence est l'une des conditions que doit respecter tout traitement.

• Le droit à l'information est lié à cette notion de transparence et est repris principalement dans les articles 13 et 14[46].
L'article 13[46] fait état des informations qui doivent être transmises à la personne concernée si la collecte a lieu auprès d'elle.
L'article 14[46], qui est très proche dans sa structure et dans son contenu de l'article 13, liste les informations qu'il faut donner à la personne concernée dans le cas d'une récolte de données à caractère personnel qui n'a pas eu lieu auprès d'elle. Le détail des modalités à respecter lors de la transmission d'informations à la personne concernée se trouve dans l'article 14.3)[46] et les conditions de non-

application dans l'article 14.5)[46].

L'informatique ici a un rôle de transmission de l'information. L'informaticien doit être conscient de l'importance de son rôle dans cette transmission.

• Le droit d'accès est abordé dans l'article 15 : *"1. La personne concernée a le droit d'obtenir du responsable du traitement la confirmation que des données à caractère personnel la concernant sont ou ne sont pas traitées et, lorsqu'elles le sont, l'accès auxdites données à caractère personnel ainsi que les informations suivantes: [...]"*[46].

L'article 15.1[46] fait état des informations qui doivent être fournies à la personne concernée s'il s'avère que ses données sont effectivement traitées. Le point 2 de cet article n'a pas été repris parce qu'il concerne les sous-traitants qui ne font pas partie du cadre du mémoire. La suite de cet article est également intéressante pour ce travail :
"3. Le responsable du traitement fournit une copie des données à caractère personnel faisant l'objet d'un traitement. Le responsable du traitement peut exiger le paiement de frais raisonnables basés sur les coûts administratifs pour toute copie supplémentaire demandée par la personne concernée. Lorsque la personne concernée présente sa demande par voie électronique, les informations sont fournies sous une forme électronique d'usage courant, à moins que la personne concernée ne demande qu'il en soit autrement.
4. Le droit d'obtenir une copie visé au paragraphe 3 ne porte pas atteinte aux droits et libertés d'autrui."[46]

Pour l'informatique, cela demande de pouvoir retrouver toutes les données concernant un seul individu à n'importe quel moment et de pouvoir les rassembler en un format électronique (à moins que l'utilisateur exprime sa volonté qu'il en soit autrement).

• Pour plus de lisibilité, le détail des informations à transmettre dans le cadre de ces 3 articles se trouve dans les tableaux de synthèse ci-après. Les informations concernant les transferts de données personnelles vers l'étranger n'ont pas été reprises dans ce tableau comme le mémoire reste dans un cadre national.

	Droit d'information		Droit d'accès
	Article 13 Récolte auprès de la personne concernée	Article 14 Récolte qui n'a pas eu lieu auprès de la personne concernée	Article 15 Accès de la personne à ses données
Identité et coordonnées du responsable du traitement (ou de son représentant)	x	x	
Coordonnées du DPO (s'il a été désigné)	x	x	
Finalité du traitement	x	x	x
Base juridique (ici, consentement)	x	x	
Destinataires ou catégories de destinataires (s'ils existent)	x	x	x (que ces destinataires aient reçu les données dans le passé ou les recevront dans le futur)
Durée de conservation des données la concernant (ou critères pour la déterminer)	x	x	x
Existence des différents droits de la personne concernée	x (droits cités : accès, rectification, effacement, limitation du traitement, opposition, portabilité des données)	x (droits cités : accès, rectification, effacement, limitation du traitement, opposition, portabilité des données)	x (droits cités : rectification, effacement, limitation du traitement, opposition)
Fait que la personne concernée peut retirer son consentement à tout moment (et que tout traitement fait avant ce retrait reste légal)	x	x	
Droit de la personne concernée à "*introduire une réclamation auprès d'une autorité de contrôle*"	x	x	x

Figure 3.2: Tableau de synthèse des informations à donner à l'utilisateur dans les différents cas possibles (article 13, 14 et 15) - première partie

	Droit d'information		Droit d'accès
	Article 13 Récolte auprès de la personne concernée	Article 14 Récolte qui n'a pas eu lieu auprès de la personne concernée	Article 15 Accès de la personne à ses données
Les "*conséquences éventuelles*" si les données ne sont pas données alors que cela est nécessaire pour la conclusion du contrat	x		
Si oui ou non les données pourraient être utilisées dans le futur pour une autre finalité. Si oui, préciser les données nécessaires dans ce cas.	x	x	
Existence d'une décision automatisée (comme le profilage) et les "*informations utiles concernant la logique sous-jacente, ainsi que l'importance et les conséquences prévues de ce traitement pour la personne concernée*"	x	x	x
Catégories de données personnelles récoltées		x	x
Détails sur l'origine des données		x (en précisant si elles sont accessibles publiquement)	x (si les données ne proviennent pas de la personne concernée : "*toute information disponible quant à leur source*")

Figure 3.3: Tableau de synthèse des informations à donner à l'utilisateur dans les différents cas possibles (article 13, 14 et 15) - deuxième partie

3.8.2 Exactitude des données et droit de rectification

Exactitude des données et droit de rectification

La PME doit traiter les données de façon à les tenir à jour pour qu'elles soient exactes. Elle doit donner l'occasion à l'utilisateur de les rectifier s'il le veut. (Art. 5.1.d et Art. 19)

Ex : Quand l'utilisateur communique un changement de statut, d'adresse etc..., la PME doit mettre son dossier à jour

Le RGPD demande que les données soient *"exactes et, si nécessaire, tenues à jour; toutes les mesures raisonnables doivent être prises pour que les données à caractère personnel qui sont inexactes, eu égard aux finalités pour lesquelles elles sont traitées, soient effacées ou rectifiées sans tarder (exactitude);"* dans l'article 5.1.d)[46].

Il confère à l'utilisateur un droit de rectification qui s'exprime dans l'article 16 : *"La personne concernée a le droit d'obtenir du responsable du traitement, dans les meilleurs délais, la rectification des données à caractère personnel la concernant qui sont inexactes. Compte tenu des finalités du traitement, la personne concernée a le droit d'obtenir que les données à caractère personnel incomplètes soient complétées, y compris en fournissant une déclaration complémentaire."*[46].

Notons aussi l'existence de l'article 19 qui dit que *"Le responsable du traitement notifie à chaque destinataire auquel les données à caractère personnel ont été communiquées toute rectification ou tout effacement de données à caractère personnel ou toute limitation du traitement effectué conformément à l'article 16, à l'article 17, paragraphe 1, et à l'article 18, à moins qu'une telle communication se révèle impossible ou exige des efforts disproportionnés. Le responsable du traitement fournit à la personne concernée des informations sur ces destinataires si celle-ci en fait la demande."*[46]. Cet article qui demande de répercuter des changements faits sur les données pour que toutes les copies potentielles restent à jour interviendra aussi dans certains autres droits de l'utilisateur.

Le système informatique doit donc être capable de présenter à l'utilisateur un moyen de rectifier ses données dans les meilleurs délais.

3.8.3 Délai de conservation et droit à l'effacement des données

Délai de conservation et droit à l'effacement

La PME peut conserver les données "pendant une durée n'excédant pas celle nécessaire au regard des finalités pour lesquelles elles sont traitées" (Art. 5.1.e) et l'utilisateur a le droit d'obtenir "l'effacement de ses données dans les meilleurs délais". (Art. 17)

Ex : - une PME ne doit effacer les données à caractère personnel reprises dans la comptabilité qu'après 7 ans (Art. III.88 du Code du droit économique). Le même raisonnement s'applique aux documents tels que les factures.
- la PME doit détruire les informations lorsqu'un candidat n'est pas engagé ou en cas de licenciement

(exemples tirés du Vade-mecum à l'intention des PME de l'APD)

Les données doivent être conservées d'une manière particulière présentée dans l'article 5.1.e) pour respecter le RGPD c'est-à-dire : *"sous une forme permettant l'identification des personnes concernées pendant une durée n'excédant pas celle nécessaire au regard des finalités pour lesquelles elles sont traitées; les données à caractère personnel peuvent être conservées pour des durées plus longues dans la mesure où*

elles seront traitées exclusivement à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques conformément à l'article 89, paragraphe 1, pour autant que soient mises en œuvre les mesures techniques et organisationnelles appropriées requises par le présent règlement afin de garantir les droits et libertés de la personne concernée (limitation de la conservation);"[46]

En plus de ces règles, l'utilisateur a aussi le droit de faire appel à son droit à l'effacement des données qui est explicité dans l'article 17 :
"1. La personne concernée a le droit d'obtenir du responsable du traitement l'effacement, dans les meilleurs délais, de données à caractère personnel la concernant et le responsable du traitement a l'obligation d'effacer ces données à caractère personnel dans les meilleurs délais"[46][6]

L'article 19[46] s'applique aussi en demandant au responsable du traitement de contacter tout destinataire des données pour répercuter une suppression de données lorsqu'elle a lieu.

En informatique, laisser la possibilité à l'utilisateur de pouvoir effacer toutes ses données n'est pas toujours la première chose à être implémentée. Cela peut demander des dispositions particulières pour s'assurer que le système reste cohérent lorsque les données ne sont plus présentes.

[6] *"lorsque l'un des motifs suivants s'applique:*
a) les données à caractère personnel ne sont plus nécessaires au regard des finalités pour lesquelles elles ont été collectées ou traitées d'une autre manière;
b) la personne concernée retire le consentement sur lequel est fondé le traitement, conformément à l'article 6, paragraphe 1, point a), ou à l'article 9, paragraphe 2, point a), et il n'existe pas d'autre fondement juridique au traitement;
c) la personne concernée s'oppose au traitement en vertu de l'article 21, paragraphe 1, et il n'existe pas de motif légitime impérieux pour le traitement, ou la personne concernée s'oppose au traitement en vertu de l'article 21, paragraphe 2;
d) les données à caractère personnel ont fait l'objet d'un traitement illicite;
e) les données à caractère personnel doivent être effacées pour respecter une obligation légale qui est prévue par le droit de l'Union ou par le droit de l'État membre auquel le responsable du traitement est soumis;
f)les données à caractère personnel ont été collectées dans le cadre de l'offre de services de la société de l'information visée à l'article 8, paragraphe 1.
2.Lorsqu'il a rendu publiques les données à caractère personnel et qu'il est tenu de les effacer en vertu du paragraphe 1, le responsable du traitement, compte tenu des technologies disponibles et des coûts de mise en œuvre, prend des mesures raisonnables, y compris d'ordre technique, pour informer les responsables du traitement qui traitent ces données à caractère personnel que la personne concernée a demandé l'effacement par ces responsables du traitement de tout lien vers ces données à caractère personnel, ou de toute copie ou reproduction de celles-ci.
3. Les paragraphes 1 et 2 ne s'appliquent pas dans la mesure où ce traitement est nécessaire:
a) à l'exercice du droit à la liberté d'expression et d'information;
b) pour respecter une obligation légale qui requiert le traitement prévue par le droit de l'Union ou par le droit de l'État membre auquel le responsable du traitement est soumis, ou pour exécuter une mission d'intérêt public ou relevant de l'exercice de l'autorité publique dont est investi le responsable du traitement;
c) pour des motifs d'intérêt public dans le domaine de la santé publique, conformément à l'article 9, paragraphe 2,points h) et i), ainsi qu'à l'article 9, paragraphe 3;
d) à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques conformément à l'article 89, paragraphe 1, dans la mesure où le droit visé au paragraphe 1 est susceptible de rendre impossible ou de compromettre gravement la réalisation des objectifs dudit traitement; ou
e) à la constatation, à l'exercice ou à la défense de droits en justice."[46]

3.8.4 Droit à la limitation du traitement des données

Droit à la limitation du traitement des données

"Le marquage des données à caractère personnel conservées, en vue de limiter leur traitement futur". (Art. 4.3)
C'est une forme de traitement alternatif des données, spécifique par rapport à la suppression, qui s'exercera par exemple pour des raisons liées à un procès.

N'est pas repris dans ce mémoire

Ex : - une personne demande à ce que ses données soient conservées et vérifiées car elle n'est pas d'accord avec elles (on dit qu'elle est cohabitante alors qu'elle vit seule ; la banque informe son conjoint qu'elle est décédée alors que ce n'est pas le cas...)

L'article 4.3 définit la limitation du traitement comme *"le marquage de données à caractère personnel conservées, en vue de limiter leur traitement futur;"*[46]. Cette notion est centrale dans l'explication du droit à la limitation du traitement qui est présentée dans l'article 18.[7]

Il s'agit de la limitation dans le temps et sur le contenu du traitement des données. L'objectif est d'éviter le détournement des données à caractère personnel, ce qui *a priori* ne concerne pas directement l'architecture logicielle. **L'informaticien** n'a pas un rôle de gestionnaire de litige. Cet article ne sera donc pas repris dans la méthodologie.

3.8.5 Droit à la portabilité des données

Droit à la portabilité des données

Les données personnelles doivent être mises à disposition de l'utilisateur "dans un format structuré, couramment utilisé et lisible par machine". Par ailleurs, l'utilisateur a "le droit de transmettre ces données à un autre responsable de traitement sans que le responsable du traitement auquel les données à caractère personnel ont été communiquées y fasse obstacle" (Art 20)

Ex : - un utilisateur veut télécharger les données à son domicile pour en vérifier l'exactitude
- pouvoir récupérer ses données d'un premier organisme pour pouvoir les donner à un autre organisme

[7] *"1. La personne concernée a le droit d'obtenir du responsable du traitement la limitation du traitement lorsque l'un des éléments suivants s'applique:*
a) l'exactitude des données à caractère personnel est contestée par la personne concernée, pendant une durée permettant au responsable du traitement de vérifier l'exactitude des données à caractère personnel;
b) le traitement est illicite et la personne concernée s'oppose à leur effacement et exige à la place la limitation de leur utilisation;
c) le responsable du traitement n'a plus besoin des données à caractère personnel aux fins du traitement mais celles-ci sont encore nécessaires à la personne concernée pour la constatation, l'exercice ou la défense de droits en justice;
d) la personne concernée s'est opposée au traitement en vertu de l'article 21, paragraphe 1, pendant la vérification portant sur le point de savoir si les motifs légitimes poursuivis par le responsable du traitement prévalent sur ceux de la personne concernée.
2. Lorsque le traitement a été limité en vertu du paragraphe 1, ces données à caractère personnel ne peuvent, à l'exception de la conservation, être traitées qu'avec le consentement de la personne concernée, ou pour la constatation, l'exercice ou la défense de droits en justice, ou pour la protection des droits d'une autre personne physique ou morale, ou encore pour des motifs importants d'intérêt public de l'Union ou d'un État membre.
3. Une personne concernée qui a obtenu la limitation du traitement en vertu du paragraphe 1 est informée par le responsable du traitement avant que la limitation du traitement ne soit levée."[46]

Ce droit s'exprime dans l'article 20 :
"1. Les personnes concernées ont le droit de recevoir les données à caractère personnel les concernant qu'elles ont fournies à un responsable du traitement, dans un format structuré, couramment utilisé et lisible par machine, et ont le droit de transmettre ces données à un autre responsable du traitement sans que le responsable du traitement auquel les données à caractère personnel ont été communiquées y fasse obstacle, lorsque:
a) le traitement est fondé sur le consentement en application de l'article 6, paragraphe 1, point a), ou de l'article 9, paragraphe 2, point a), ou sur un contrat en application de l'article 6, paragraphe 1, point b); et
b) le traitement est effectué à l'aide de procédés automatisés.
2. Lorsque la personne concernée exerce son droit à la portabilité des données en application du paragraphe 1, elle a le droit d'obtenir que les données à caractère personnel soient transmises directement d'un responsable du traitement à un autre, lorsque cela est techniquement possible.
3. L'exercice du droit, visé au paragraphe 1 du présent article s'entend sans préjudice de l'article 17. Ce droit ne s'applique pas au traitement nécessaire à l'exécution d'une mission d'intérêt public ou relevant de l'exercice de l'autorité publique dont est investi le responsable du traitement.
4. Le droit visé au paragraphe 1 ne porte pas atteinte aux droits et libertés de tiers."[46]

"Le G29 avance des pistes de mise en œuvre concrète du droit à la portabilité.
Notamment, il appelle à la création de services 'd'entrepôts de données personnelles', qui permettrait aux personnes concernées de stocker leurs données personnelles, puis de donner accès au cas par cas aux données pertinentes. Cet entrepôt pourrait être basé sur un serveur SFTP, une interface de programmation applicative (API) en ligne ou un portail internet sécurisé."[47]

Pour l'informatique, ce droit peut être la source de remises en question d'architectures déjà existantes. Retrouver les données d'une personne peut s'avérer assez complexe, par exemple si l'architecture dans laquelle les données se déplacent est ne serait-ce qu'un peu décentralisée.

3.8.6 Droit d'opposition et droit à ne pas être soumis à une décision individuelle automatisée

Droit d'opposition et droit à ne pas être soumis à une décision individuelle automatisée

Le droit d'opposition permet à l'utilisateur d'échapper aux traitements liés à la publicité et au marketing direct, et plus généralement de s'opposer au fait que ses données soient utilisées dans des traitements avec lesquels il n'est pas d'accord.
Le droit de ne pas être soumis à une décision individuelle automatisée demande de fournir la possibilité à tout utilisateur de pouvoir se soustraire par exemple à un profilage.

Ex : un utilisateur achète des aliments pour animaux en ligne, puis reçoit des mails de différentes entreprises du même secteur. Il écrit à l'entreprise qui fournit les aliments et dit qu'il ne veut plus de pubs de leurs partenaires. La PME doit s'exécuter.

Le droit d'opposition est présenté dans l'article 21 : *"1. La personne concernée a le droit de s'opposer à tout moment, pour des raisons tenant à sa situation particulière, à un traitement des données à caractère personnel la concernant fondé sur l'article 6, paragraphe 1, point e) ou f), y compris un profilage fondé sur ces dispositions."*[46][8]

[8] *"Le responsable du traitement ne traite plus les données à caractère personnel, à moins qu'il ne démontre qu'il existe des motifs légitimes et impérieux pour le traitement qui prévalent sur les intérêts et les droits et libertés de la personne concernée, ou pour la constatation, l'exercice ou la défense de droits en justice.*
2. Lorsque les données à caractère personnel sont traitées à des fins de prospection, la personne concernée a le droit de s'opposer à tout moment au traitement des données à caractère personnel la concernant à de telles fins de prospection, y compris au profilage dans la mesure où il est lié à une telle prospection.

Le droit d'opposition permet à l'utilisateur d'échapper aux traitement liés à la publicité et plus généralement de s'opposer au fait que ses données soient utilisées dans des traitements avec lesquels il ou elle n'est pas d'accord.

Le droit à ne pas être soumis à une décision individuelle automatisée est détaillé dans l'article 22: *"1. La personne concernée a le droit de ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé, y compris le profilage, produisant des effets juridiques la concernant ou l'affectant de manière significative de façon similaire."*[46][9]

Ce droit à ne pas être soumis à une décision individuelle automatisée demande de fournir la possibilité à tout utilisateur et à toute utilisatrice de pouvoir se soustraire, notamment, à un profilage. Cela peut dépasser le système d'information lui-même car cela peut demander une intervention humaine dans le processus.

Pour l'infomaticien, ceci a comme conséquence que le système doit être capable d'omettre des utilisateurs dans le cadre de certains traitements.

3.9 Conclusion de la présentation du RGPD

Le RGPD concerne des données privées d'individus qui les fournissent en fonction des situations (livraisons, mailing lists, prises de rendez-vous, inscriptions à des newsletters, ...). Le RGPD a pour objectif de défendre la vie privée de ces individus. Du point de vue de la mise en oeuvre informatique de ces articles, le rôle de l'informaticien est de respecter l'esprit de la loi tout en s'adaptant au contexte dans lequel il travaille.

Pour l'informatique, le RGPD amène de multiples défis techniques.
La gestion des données, par exemple, amène de nombreux défis d'automatisation étant donné que les programmes ne sont pas en mesure de sélectionner eux-mêmes les données absolument nécessaires au traitement et le tag automatique des données comme "sensibles" ou "pseudonymisées" peut apporter ses propres challenges.

---

3. Lorsque la personne concernée s'oppose au traitement à des fins de prospection, les données à caractère personnel ne sont plus traitées à ces fins.
4. Au plus tard au moment de la première communication avec la personne concernée, le droit visé aux paragraphes 1 et 2 est explicitement porté à l'attention de la personne concernée et est présenté clairement et séparément de toute autre information.
5. Dans le cadre de l'utilisation de services de la société de l'information, et nonobstant la directive 2002/58/CE, la personne concernée peut exercer son droit d'opposition à l'aide de procédés automatisés utilisant des spécifications techniques.
6. Lorsque des données à caractère personnel sont traitées à des fins de recherche scientifique ou historique ou à des fins statistiques en application de l'article 89, paragraphe 1, la personne concernée a le droit de s'opposer, pour des raisons tenant à sa situation particulière, au traitement de données à caractère personnel la concernant, à moins que le traitement ne soit nécessaire à l'exécution d'une mission d'intérêt public."[46]

[9] *"2. Le paragraphe 1 ne s'applique pas lorsque la décision:*
a) est nécessaire à la conclusion ou à l'exécution d'un contrat entre la personne concernée et un responsable du traitement;
b) est autorisée par le droit de l'Union ou le droit de l'État membre auquel le responsable du traitement est soumis et qui prévoit également des mesures appropriées pour la sauvegarde des droits et libertés et des intérêts légitimes de la personne concernée; ou
c) est fondée sur le consentement explicite de la personne concernée.
3. Dans les cas visés au paragraphe 2, points a) et c), le responsable du traitement met en œuvre des mesures appropriées pour la sauvegarde des droits et libertés et des intérêts légitimes de la personne concernée, au moins du droit de la personne concernée d'obtenir une intervention humaine de la part du responsable du traitement, d'exprimer son point de vue et de contester la décision.
4. Les décisions visées au paragraphe 2 ne peuvent être fondées sur les catégories particulières de données à caractère personnel visées à l'article 9, paragraphe 1, à moins que l'article 9, paragraphe 2, point a) ou g), ne s'applique et que des mesures appropriées pour la sauvegarde des droits et libertés et des intérêts légitimes de la personne concernée ne soient en place."[46]

L'informaticien doit aussi agir en accord avec le contexte dans lequel il se trouve, ses décisions quant aux mesures techniques à adopter doivent respecter les décisions prises par l'organisation dans laquelle il travaille.
Certains articles du RGPD peuvent aussi recouvrir des démarches informatiques demandant des ressources conséquentes. La gestion de la sécurité, qui apparaît principalement dans l'article 32, est une démarche informatique en elle-même. Tout comme la gestion du droit à la portabilité des données de l'article 20 et comme de nombreux autres articles.
Le cadre du RGPD qui tend à s'appliquer quelles que soient les technologies utilisées demande aussi une réflexion informatique plus large car s'il est possible de créer des modèles et des méthodes de haut niveau, chaque technologie est susceptible d'avoir des spécificités auxquels ces modèles ne peuvent pas répondre parfaitement.

Chapitre 4

Élaboration de la méthodologie

4.1 Introduction

La méthodologie qui sera élaborée a pour objectif de permettre à des informaticiens de pouvoir repérer des Zones de risque potentiels, c'est-à-dire où il y a un risque de ne pas respecter le RGPD. Le but n'est pas de pouvoir certifier que le traitement respecte ou non le RGPD. La démarche élaborée ne peut que donner des outils de réflexion.

Le cadre de la réflexion établi dans l'introduction et dans le chapitre 3 prend tout son sens ici. Il s'agit sur ces bases de modéliser le flux d'informations dans un système informatique et non le RGPD lui-même.

Le challenge repose donc sur le fait qu'il existe une multitude de PME et d'ASBL qui traitent un large éventail de types de données pour des raisons et des buts différents. Leur taille varie d'une manière qui peut être importante d'une PME ou ASBL à l'autre. Réussir à créer une solution unique "one-size" qui conviendrait à toutes les PME ou ASBL est une tâche vraisemblablement impossible.

C'est la méthode de modélisation de menaces LINDDUN[2] qui a été choisie pour guider l'élaboration de la méthodologie de ce mémoire. LINDDUN a été présentée brièvement dans l'état de l'art. LINDDUN base son analyse du degré de vie privée sur un diagramme de flux de données. C'est particulièrement intéressant dans le cadre de ce travail.

Les autres méthodes présentées dans l'état de l'art auraient aussi pu être choisies car elles apportent toutes quelque chose. Cependant, LINDDUN est la méthode qui était la plus en accord avec l'objectif du mémoire, c'est-à dire la mise en oeuvre d'un sous-ensemble du RGPD grâce à l'analyse des flux de données. Le fait que LINDDUN est une méthodologie dont le principe pouvait s'appliquer au RGPD permettait d'avoir une base solide sur laquelle étayer la réflexion qui va suivre.

L'état de l'art a abordé le fait qu'il était aussi possible d'utiliser les Firewalls et les Network Intrusion Detection Systems dans le but de suivre le flux des données. Cependant, lorsque ces technologies ont été mises en avant dans l'état de l'art, elles s'intéressaient principalement à l'entrée et à la sortie des données du réseau de l'entreprise. Si cela peut être intéressant du point de vue du RGPD, se baser sur une réflexion comme LINDDUN permet d'inclure le détail des traitements effectués sur les données lors de leur passage dans l'entreprise et donc d'apporter une vue d'ensemble plus large des parties du système soumises au RGPD.

Les recherches concernant LINDDUN et le RGPD existent et prouvent de cette manière qu'il est théoriquement possible d'utiliser cette méthode particulière dans le cadre du RGPD. Cependant, aucun des articles découverts lors de la rédaction de l'état de l'art ne semble pointer vers une utilisation s'intéressant aux mêmes points particuliers d'application que ce que ce mémoire cherche à faire. Il est dès lors intéressant de voir comment LINDDUN peut être utilisé dans le contexte du RGPD. La démarche présentée dans les points qui suivent consiste à reprendre la base proposée par LINDDUN pour y incorporer les notions spécifiques au RGPD sélectionnées précédemment.

La méthodologie de ce mémoire étend LINDDUN. LINDDUN offre une méthode d'analyse de menaces liées à la vie privée très complète. Cependant, le RGPD rajoute des notions qui sont propres à ce règlement. Par exemple, les différents droits auxquels l'utilisateur doit pouvoir faire appel. Si ces droits pouvaient déjà avoir été implémentés par certaines organisations, ce n'était pas toujours le cas. Par exemple, un site internet pouvait mettre à disposition de l'utilisateur un moyen de supprimer ses données personnelles en quelques clics alors que d'autres sites ne laissaient pas la possibilité de supprimer facilement son propre compte. Le RGPD rend ces droits obligatoires. La méthodologie suivante va se baser sur LINDDUN pour observer quels sont les points communs entre cette méthodologie et le sous-ensemble du RGPD traité avant d'étendre LINDDUN de manière à englober les notions propres à la partie du RGPD examinée.

Ce chapitre présente les différentes étapes qui ont été suivies dans le cadre de l'élaboration théorique d'une méthodologie.

4.2 Structure de la méthodologie

Les 3 premières étapes de LINDDUN ont pour but une analyse de la situation et une mise en évidence de menaces potentielles.[2] La méthodologie définie ci-dessous va suivre le processus défini par LINDDUN tout en l'étendant pour englober le RGPD. Les articles utilisés sont référencés à chaque étape.
Le premier point présente une analyse de la situation qui doit être étendue au détail des traitements internes importants pour le RGPD.
Le deuxième point montre une première sélection des catégories de zones de risque de non-conformité.
Le troisième point explicite le détail des zones de risque qui ont été sélectionnées[2].

Les 3 dernières étapes de LINDDUN ont pour but de trouver un plan d'action pour répondre à ces zones de risque.[2]
Le quatrième point cherche à donner à chaque zone de risque une priorité.
Le cinquième point veut trouver les mesures qui peuvent être mises en place.
La sixième et dernière étape va sélectionner les PETs[1] qui peuvent s'appliquer pour aider à implémenter les mesures choisies.[2]

4.3 Etape 1 : Construction du dataflow diagram

4.3.1 Critères nécessaires à l'élaboration du diagramme de flux de données dans le cadre du RGPD

Pour la modélisation sous forme de diagramme de flux de données, les notations reprises dans la documentation de LINDDUN sont celles qui seront utilisées. Elles sont présentées dans le graphique suivant.

[1] Privacy enhancing tools

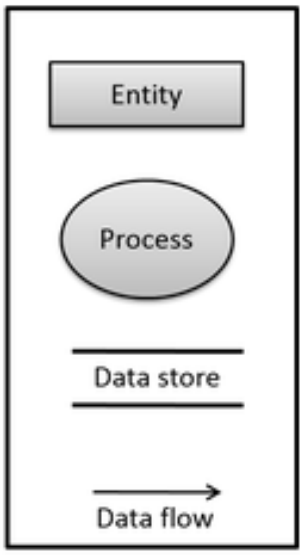


Figure 4.1: Notation (convention graphique) pour le diagramme de flux de données de LINDDUN[2]

Le type des données sera ajouté sur le graphique. Par exemple, on entend par type le fait que "Jane" est un prénom et "Doe" un nom de famille.

La différence entre l'approche LINDDUN et l'approche qui sera présentée ici se trouve au niveau de ce qui est modélisé. LINDDUN s'intéresse principalement aux données introduites dans le système, à la manière dont elles sont stockées et aux informations partagées avec des "entités externes". Les traitements internes au système ne sont pas détaillés.[48] Ici, les traitements effectués de manière interne au système se révèleront très intéressants car faisant partie des informations relevant du RGPD. Il sera important que le dataflow diagram précise les traitements effectués dans ce cas.

2 autres informations seront intéressantes dans le cadre de l'analyse du RGPD :
• La finalité de chaque traitement ou une mention que ce traitement a une finalité déterminée. L'informaticien devrait garder à l'esprit que les finalités doivent être déterminées par le responsable du traitement de la PME ou ASBL étant donné qu'elles ne concernent pas seulement les moyens informatiques mis en oeuvre pour réaliser le traitement.
• Le type des données au sens du RGPD. Les 4 types de données du RGPD (données à caractère personnel, données pseudonymisées, données anonymisées et données sensibles ont été détaillées dans le point 3.4).

4.4 Etape 2 : Relier les menaces concernant la vie privée aux éléments du diagramme de flux de données

4.4.1 Cerner le contexte des menaces pour la vie privée déterminées par LINDDUN

L'étape suivante permet de commencer à cerner les menaces possibles pour la vie privée provenant de LINDDUN et pouvant être adaptées au RGPD.

Les 7 types de menaces relevées par LINDDUN et présentées dans le catalogue de menaces disponibles sur le site officiel et dans le point 2.6.1 de l'état de l'art peuvent déjà être reliées au RGPD mais peuvent être spécifiées grâce au texte de loi. Pour rappel, le but n'est pas ici de rappeler l'intégralité de la méthodologie développée par LINDDUN dans son catalogue de menaces[49] mais de se concentrer sur

comment cette méthodologie particulière peut s'articuler avec le RGPD pour être utile à un informaticien lors de la création d'une nouvelle application.

Le point de vue adopté lors d'une analyse utilisant LINDDUN est celui d'une attaque potentielle du système. Le RGPD, quant à lui, adopte un point de vue différent étant donné qu'il s'agit d'une loi qui prescrit ce qu'il faut faire pour protéger la vie privée, dans le cas particulier qui nous intéresse ici, celui de l'informaticien concevant son système. Ces 2 points de vues doivent être conciliés pour permettre l'élaboration d'une méthodologie cohérente.

Le RGPD répond par ses prescriptions à certaines menaces relevées par LINDDUN. Les illustrations de ceci seront explicitées plus bas dans le point de l'étape 3.

Tous les articles du RGPD considérés ici ne correspondent pas nécessairement à une menace particulière de LINDDUN. Certains d'entre eux peuvent être reliés à une ou plusieurs menaces de LINDDUN mais d'autres concernent de nouveaux points à analyser. En effet, dans l'étape 1 il a été remarqué que LINDDUN se concentrait principalement sur l'entrée et la sortie des informations personnelles et sur leur stockage. Le RGPD, quant à lui, s'intéresse aussi aux traitements effectués sur ces données et présente les droits qu'a une personne dans le cadre de la protection de ses données personnelles.

Il est aussi à noter que l'article 32 du RGPD[46] qui s'intéresse à la nécessité de mettre en place des mesures de sécurité, englobe d'une certaine manière à la fois les analyses faites par LINDDUN dans le cadre de la vie privée et même celles de STRIDE dans un cadre plus général de sécurité informatique. L'article 32 ne se rattache donc pas à une menace particulière de LINDDUN mais l'englobe.

Le point de vue choisi dans la suite est donc plus proche de celui du responsable du traitement que de celui d'une attaque potentielle. Cela est dû au fait que l'intérêt de la méthodologie développée ici est d'aider à identifier des Zones de risque de non-respect du RGPD pour un informaticien qui chercherait à créer ou à maintenir un système concerné par le RGPD. Cependant, étant donné que la méthodologie suivante est basée sur LINDDUN, les parties spécifiques à cette méthodologie particulière garderont bien évidemment le point de vue spécifié par ceux qui l'ont mise au point.

Si la suite du raisonnement s'appuie sur les articles du RGPD, tous les articles ne sont pas repris et seul un juriste peut être à même de répondre aux questions plus précises que pourrait se poser l'informaticien et attester d'une conformité ou de son absence.

4.4.2 Menaces applicables au RGPD

• Identifiability
L'"**identifiability**" se concentre sur le risque de pouvoir retrouver l'identité d'une personne, donc de la lier avec des objets du système. L'"anonymité" représente l'absence de ce lien. La "pseudonymité", au sens présenté par LINDDUN, s'incarne dans le fait qu'il est possible de rassembler des informations autour d'un pseudo pour créer une réputation, ou désigne simplement l'utilisation d'un pseudo.[48] L'identifiabilité représente aussi le risque que l'article 4.5) du RGPD veut éviter en s'assurant que les données pseudonymisées (au sens du RGPD) "*soient conservées séparément et soumises à des mesures techniques et organisationnelles afin de garantir que les données à caractère personnel ne sont pas attribuées à une personne physique identifiée ou identifiable*"[46]. L'article 11.1 du RGPD pousse la recherche de l'évitement de l'identifiability lorsque les données personnelles ne sont plus nécessaires au traitement en l'écrivant ainsi "*Si les finalités pour lesquelles des données à caractère personnel sont traitées n'imposent pas ou n'imposent plus au responsable du traitement d'identifier une personne concernée, celui-ci n'est pas tenu de conserver, d'obtenir ou de traiter des informations supplémentaires pour identifier la personne concernée à la seule fin de respecter le présent règlement.*"[46]

• Unawareness
La "**content unawareness**" est un manque de connaissance des conséquences possibles pouvant survenir après le partage de certaines informations personnelles. Le système devrait chercher à utiliser aussi

peu de données personnelles que possible dans l'accomplissement de sa tâche et à aider l'utilisateur à prendre conscience des données qu'il ou elle partage et des conséquences qui peuvent en découler. LINDDUN veut permettre de donner l'occasion de changer la perception que l'utilisateur est le seul responsable de sa vie privée et donner l'occasion au système d'aider l'utilisateur à apréhender les décisions relevant de la vie privée en faisant notamment preuve de transparence. [48] Enormément d'articles du RGPD font référence à la nécessité de transmettre des informations à l'utilisateur dans le but de l'informer sur les sujets qui concernent ses données personnelles et les droits qu'il possède sur ses données. Parmi les articles qui peuvent être reliés à cette menace particulière, il est possible de citer les articles 5.1.a), 7.2, 7.3, 13, 14, 15.1, 15.3, 19, 21.4.[46]. Le détail des différentes informations à fournir dans les différents cas sera détaillé dans l'étape 3 et dans les arbres spécifiques au RGPD.

- **Non-compliance**

La conformité au RGPD se retrouve dans la "**policy and content non-compliance**", si on considère à un haut niveau cette conformité, étant donné que ce type de menace recouvre le fait de respecter notamment les législations en place.[50] La vision présentée par ce type de menace se confronte au fait qu'il est nécessaire de connaître les différentes lois en place et la nécessité de faire appel à un expert légal pour s'assurer d'une conformité parfaite. Cela est renforcé par le fait que pour un produit particulier, plusieurs lois peuvent être susceptibles de s'appliquer. Etant donné que l'informaticien n'est pas juriste, et que cette menace traite de la conformité à la loi, il est nécessaire de faire appel à des experts pour évaluer la conformité au RGPD.
Ceci est à moduler concernant le consentement. En effet, le RGPD contient quand même des articles qui peuvent être associés à cette volonté de se conformer aux lois en vigueur. Il s'agit de l'article 7 du RGPD[46] qui concerne les points sur lesquels le responsable du traitement doit pouvoir démontrer des conditions particulières. Encore une fois, le détail sera proposé dans l'étape 3.

4.4.3 Menaces non-applicables au sous-ensemble du RGPD examiné

- **Detectability**

L'une des menaces de LINDDUN qui peut être reliée à des idées générales mais difficilement être reliée à un article particulier du RGPD, du moins dans le sous-ensemble d'articles considéré ici, est la **"Detectability"**. La **"Detectability"** permet de repérer l'existence d'un élément particulier, même si son contenu reste inconnu d'un attaquant potentiel. Le problème qui peut menacer la vie privée avec la détectabilité est le fait que connaître l'existence d'un objet d'un système peut permettre d'inférer une ou plusieurs informations (LINDDUN prend l'exemple de l'existence du dossier d'une célébrité dans un centre de désintoxication qui peut permettre de conclure que cette célébrité a (ou a eu) une addiction).[48]

- **Linkability**

Une autre menace est la "**linkability**", c'est-à-dire la possibilité de lier un objet du système à un autre. LINDDUN précise que cela devient un problème pour la vie privée lorsque cela peut aboutir à l'identification de personnes utilisant le système.[48]. Comme aucun des articles du sous-ensemble du RGPD examiné n'a été relié à cette menace, elle n'a pas été retenue ici.

- **Information Disclosure**

L'"**information disclosure**"a à voir avec les droits d'accès, car elle représente le cas où quelqu'un n'ayant pas les autorisations nécessaires se retrouve en présence d'informations qu'il n'est pas censé voir.[48]. Cette menace provient de STRIDE. LINDDUN ne porte plus l'attention dessus autant qu'auparavant parce qu'il insiste sur la nécessité d'effectuer une analyse de sécurité en complément de LINDDUN. La même recommandation est toujours valable ici.

- **Non-Repudiation**

La menace identifiée par LINDDUN est la "**non-repudiation**" qui empêche tout doute quant à l'identité de l'auteur d'une action particulière. LINDDUN rappelle que cette menace est souvent un but recherché dans un contexte de sécurité. Cette menace est donc soit un attribut recherché (LINDDUN prend

l'exemple d'un vendeur pouvant prouver qu'il a bien remis un produit à un acheteur grâce à une preuve de paiement) soit une menace à la vie privée (LINDDUN prend le cas d'un vote exécuté de manière numérique où pouvoir relier une personne à un choix politique est une menace à la vie privée).[48] **L'article 30 du RGPD**[46] qui concerne la tenue du registre des traitements peut sembler répondre à la *non-repudiation* comme objectif de sécurité. Il ne s'agit donc pas exactement de la manière dont LINDDUN envisage la non-repudiation étant donné que LINDDUN est orienté vers l'analyse de menaces vis-à-vis de la vie privée.[48] Cette menace n'a donc pas été reprise dans la suite.

4.4.4 Menaces spécifiques au RGPD non identifiées par LINDDUN

2 types de menaces ont été ajoutées aux menaces déterminées par LINDDUN car elles sont spécifiques au RGPD. Elles sont présentées de manière générale ici et seront détaillées dans l'étape 3.

- **Menace "Processing & Purposes"**

La première concerne les **Traitements et Finalités**. Comme détaillé plus haut, LINDDUN ne s'intéresse théoriquement pas en détail aux traitements effectués dans un système analysé. Le RGPD, lui, s'y intéresse. La notion de traitement est en effet une des notions importantes du RGPD qui en détaille notamment les *"principes relatifs au traitement des données à caractère personnel"* dans les articles 5, 6, 7 et 8.[46] La notion de traitement ne pouvant pas être dissociée de la notion de finalité, cette dernière est aussi contenue dans cette menace potentielle pour la vie privée.

- **Menace "Rights of Data Subject"**

La deuxième concerne les **Droits des personnes concernées**. En effet, le RGPD met en avant une série de droits appartenant à la personne dont les données personnelles sont traitées par le système. Ces droits peuvent avoir des conséquences dans la manière dont doit être pensé le système informatique qui doit permettre aux personnes concernées d'utiliser ces droits comme indiqués dans le RGPD. Ils sont repris dans les articles 15, 16, 17, 19, 20, 21 et 22 du RGPD.[46]

4.4.5 Tableau des menaces pouvant concerner le système

Une fois les types de menaces[2] déterminés, il est possible de remplir un tableau reposant sur le framework proposé originellement par LINDDUN[48] en incluant les ajouts déterminés plus haut.

Il s'agit donc d'étendre le pattern de tableau de menaces proposé par LINDDUN avec les 2 nouvelles classes de menaces définies ci-dessus. De cette manière, un analyste pourra réfléchir à la problématique de l'application du RGPD avec les notions communes aux 2 démarches et avec celles qui sont spécifiques à la compliance au RGPD.

[2]Pour rappel : I = Identifiability, U = Unawareness, N = policy Non-compliance, P&P = Traitements de données et Finalités et R = Droits des personnes concernées

	I	U	N	P&P	R
Entity					
Data store					
Data flow					
Process					

Figure 4.2: Template du tableau de menaces reprises de LINDDUN complété par les "Processing & Purposes" et les "Rights of Data Subject"

4.5 Etape 3 : Identifier les scénarios de menace

4.5.1 Préambule

• **Identification des arbres de menaces**
Cette section va permettre d'identifier les scénarios de menaces plus précis qui pourraient potentiellement présenter un risque pour la vie privée (threat modelling). LINDDUN prévoit parfois plusieurs arbres de menaces pour une seule menace générale. Dans ce cas, le mémoire en fera rapidement état avant de se concentrer sur celui qui sera complété par les notions du RGPD. La suite ne présente que les graphiques des arbres de menaces qui peuvent être complétés et ceux qui ont été rajoutés dans leur intégralité pour permettre l'adaptation au contexte du RGPD.

Il arrive que 2 arbres de menaces possèdent la même feuille, cela est dû au fait que lors d'une analyse, seul un sous-ensemble d'arbres va être passé en revue. Dupliquer une feuille permet donc de diminuer la probabilité qu'elle ne soit pas prise en compte par celui qui conduit l'analyse alors qu'elle aurait pu être pertinente dans son analyse.

• **Convention d'écriture**
Dans les diagrammes suivants, une convention d'écriture va être utilisée pour réduire l'espace nécessaire à la création des diagrammes et privilégier la clarté des informations présentées. Ainsi :
- **PC** sera la personne concernée, l'utilisateur qui donne accès à ses données personnelles.
- **RT** sera le responsable du traitement au sens du RGPD.
Les définitions de ces 2 termes ont été présentées dans les point 3.3.1 et 3.3.3 respectivement.

Notons aussi que les diagrammes complétés ou créés dans la suite sont des arbres de menaces, c'est-à-dire des arbres présentant les cas qui ne sont pas souhaités et pourraient avoir des conséquences négatives sur la vie privée (dans le cas de LINDDUN) et du respect du RGPD. Pour cela, les menaces rajoutées formulent les exigences du RGPD en utilisant la négation. Par exemple : "PC ne fait pas cela (contrairement à l'article x)". Cela permet de respecter le formalisme étali par LINDDUN et donc maintient la cohérence interne des diagrammes complétés. Le sens des articles a été au maximum préservé, cependant, l'autrice du mémoire n'est pas juriste et une erreur peut s'être glissée quelque part. Les articles ont aussi été mentionnés, sous leur forme positive originale dans les explications qui suivent chacun des arbres et sont, bien sûr, présents dans le RGPD.

4.5.2 "Identifiability"

L'Identifiability est présentée en 4 arbres par LINDDUN, ceux de l'"Identifiability of process", de l'"Identifiability of dataflow", de l'"identifiability of datastore" et de l'"Identifiability of entity"[3].

L'arbre de menaces portant sur l'"Identifiability of process" n'a pas été repris ici car, tout comme LINDDUN l'explique, cela ne peut dériver que d'une **"Information Disclosure"** et LINDDUN renvoie à STRIDE pour le détail lié à cette mesure particulière.[52]

L'"Identifiability of dataflow"[53] et l'"Identifiability of data store"[54] ne sont pas celles qui ont été mises au centre des préoccupations du RGPD.

Le sous-ensemble du RGPD traité dans ce mémoire va aborder l'"Identifiability of entity"[51]. Le segment du RGPD sur lequel se base ce travail s'intéresse principalement à la relation entre le responsable du traitement et la personne concernée. Le responsable du traitement doit parfois pouvoir identifier la personne concernée pour lui rendre le service. Cependant, dans le cas du traitement des données pseudonymisées et anonymisées, le responsable du traitement doit prendre des mesures pour s'assurer que les données traitées ne puissent pas identifier la personne à qui elles appartiennent.

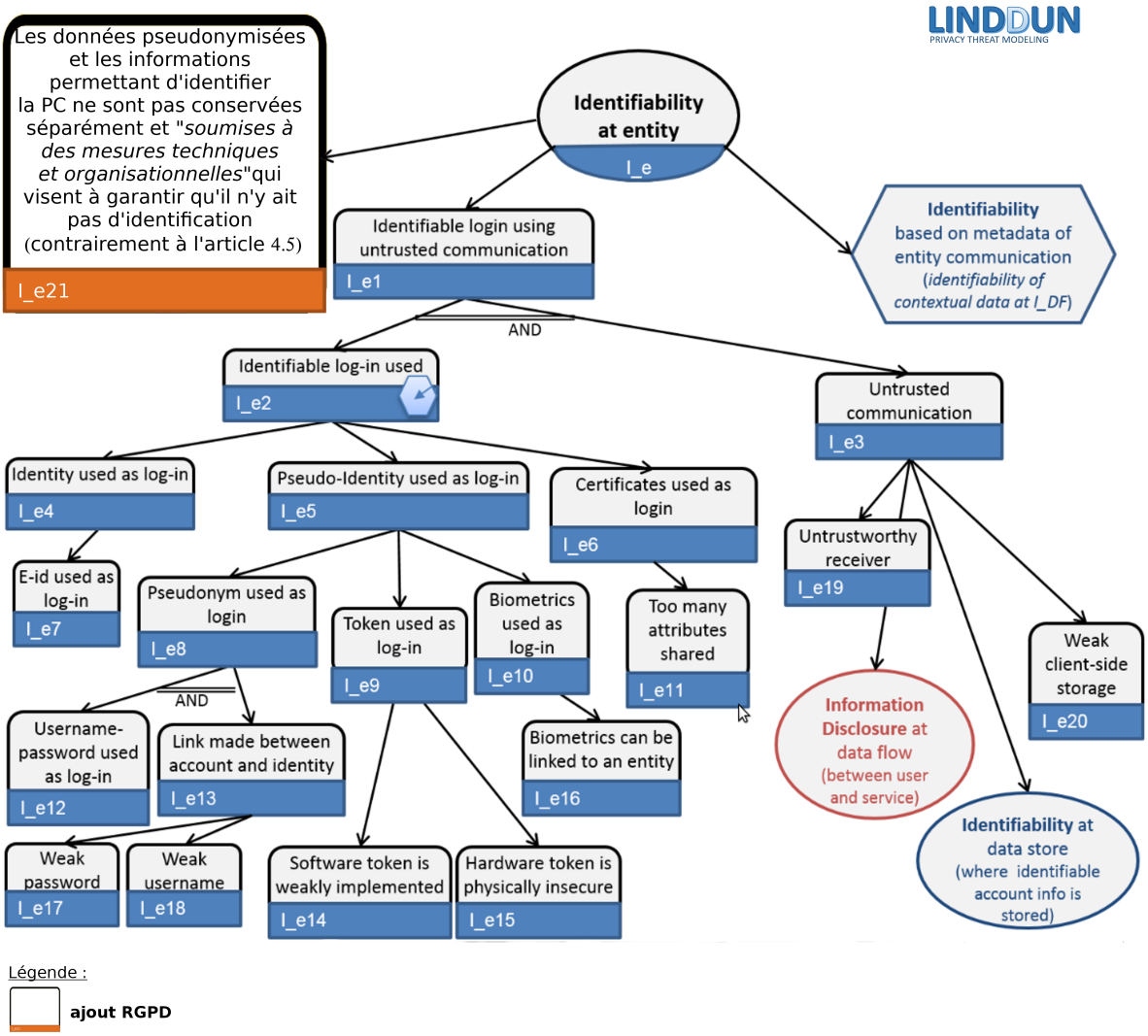


Figure 4.3: Diagramme d'"identifiability" de LINDDUN complété

[3]LINDDUN utilise à la fois les termes "Identifiability **of** entity" et "Identifiability **at** entity" dans son catalogue de menaces.[51]

L'arbre présenté ci-dessus comporte l'arbre de menaces de l'**"identifiability of entity"** défini par LINDDUN [55] ainsi que la menace pour la vie privée tirée de la partie du RGPD analysée qui est présentée en orange pour être facilement distinguée de l'arbre original.

Le détail du graphique propre au RGPD va être présenté ci-après. Pour des raisons de concision et de clarté, la nouvelle feuille de cet arbre sera désignée par son identifiant (sous forme I_eX pour garder le formalisme préalablement établi par LINDDUN).

I_e21 : le RGPD insiste sur le fait que les données pseudonymisées et les données permettant de relier les données pseudonimisées à un individu doivent être conservées séparément pour que les données *"ne soient pas attribuées à une personne physique identifiée ou identifiable"* (article 4.5).[46]

4.5.3 "Content Unawareness"

La **"Content Unawareness"** est une menace à laquelle le RGPD tend à répondre, c'est d'ailleurs la menace de LINDDUN qui est liée au plus grand nombre d'articles parmi le sous-ensemble d'articles du RGPD sur lesquels se penche cette méthodologie.

L'arbre présenté ci-dessous ne comporte que ce qui peut potentiellement représenter une zone de risque de non-conformité. La formalisation utilisée est la même que celle proposée par LINDDUN et l'arbre de menaces que ce point vient compléter se trouve sur la page correspondante dans le catalogue de LINDDUN[55]

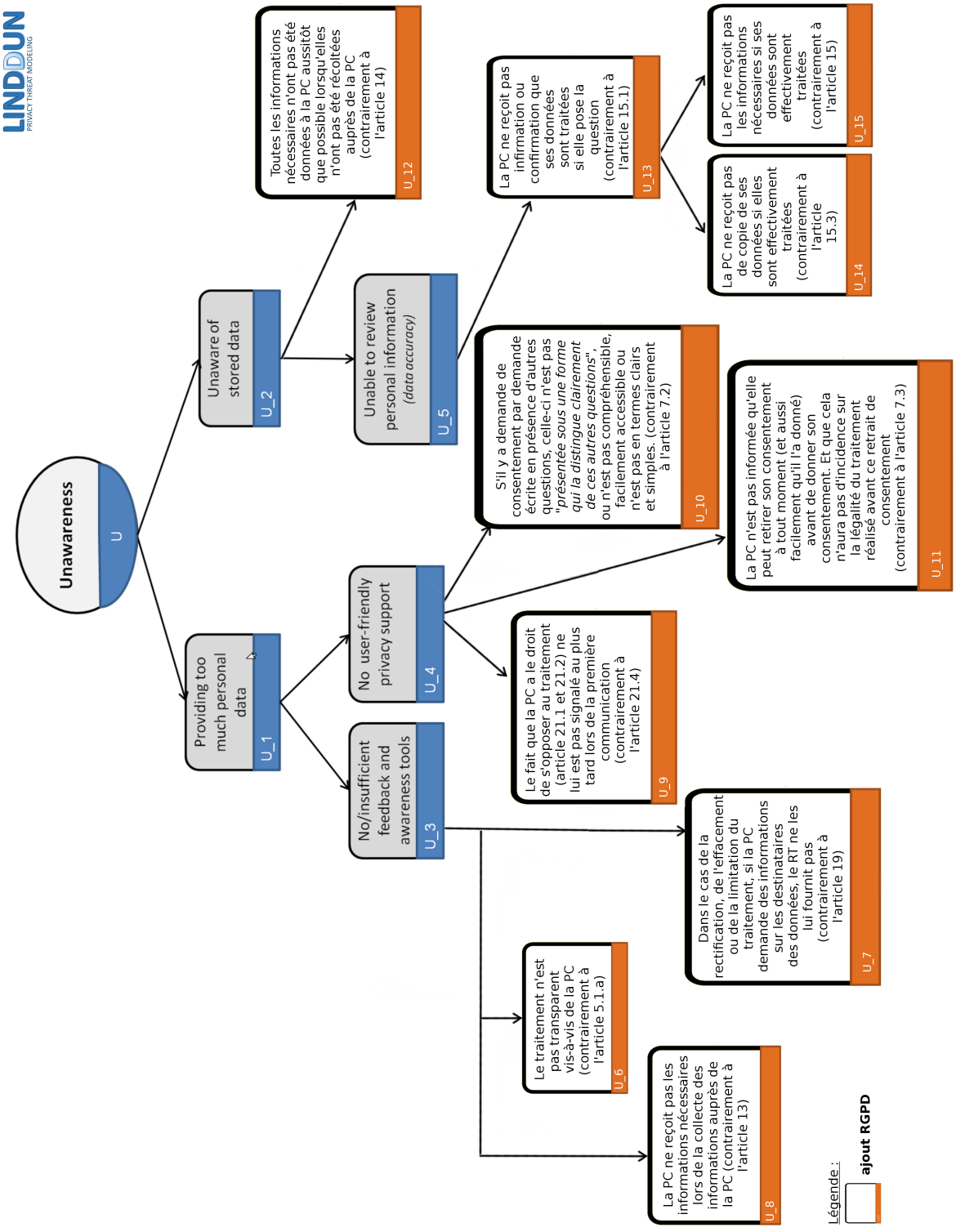


Figure 4.4: Diagramme d'"Unawareness" de LINDDUN complété

Le détail des éléments ajoutés à l'arbre de menaces de départ va à présent être détaillé.

U_6 : l'article 5.1.a)[46] s'intéresse à l'une des conditions que doit respecter le traitement. Ces conditions sont développées plus en détails dans l'élément *TF_1* de "Traitements et finalités" ci-dessous.

U_7 : l'article 19[46] expose le fait que toute rectification, effacement ou limitation de traitement doit être signalé aux destinataires des données pour que cela soit répercuté.

U_8 : cet élément parle de la nécessité de donner accès à la personne concernée aux informations précisées dans l'article 13[46]. Le détail de ces informations a déjà été synthétisé dans les figures 3.2 et 3.3 et ne sera donc pas repris ici.

U_9 : l'article 21.4)[46] rappelle qu'il faut signaler à la personne concernée son droit d'opposition au traitement (qui est détaillé dans les articles 21.1) et 21.2)[46]).

U_10 : l'article 7.2[46] met en avant le fait que le consentement doit être présenté à part de toute autre question lorsqu'il y a demande du consentement.

U_11 : l'article 7.3[46] expose le fait que la personne concernée doit pouvoir retirer son consentement à tout moment, aussi facilement qu'il l'a donné et qu'il doit apprendre cela avant que le consentement soit donné.

U_12 : cet élément parle de la nécessité de donner accès à la personne concernée aux informations précisées dans l'article 14[46]. Le détail de ces informations a déjà été synthétisé dans les figures 3.2 et 3.3 et ne sera donc pas repris ici.

U_13 : l'article 15.1[46] signale le droit de la personne concernée de demander si oui ou non ses informations personnelles sont traitées et d'avoir une réponse à ce sujet.

U_14 : l'article 15.3[46] expose le droit de la personne concernée à recevoir une copie de ses données personnelles qui sont traitées.

U_15 : cet élément parle de la nécessité de donner accès à la personne concernée aux informations précisées dans l'article 15.1[46]. Le détail de ces informations a déjà été synthétisé dans les figures 3.2 et 3.3 et ne sera donc pas repris ici.

4.5.4 "Policy and Content Non-Compliance"

La **"Policy and Content Non-Compliance"** concerne principalement le respect du RGPD dans son ensemble. Les articles qui ont quand même été repris ici (et qui ne sont qu'une partie de la sélection d'articles abordés dans ce mémoire) énoncent donc le besoin de respecter ce règlement ou une partie plus précise de ce règlement. LINDDUN considère que chacun des éléments de cet arbre de menaces s'applique au système dans son ensemble[50], cela reste aussi vrai pour les 2 feuilles rajoutées qui concernent le consentement et le respect du RGPD.

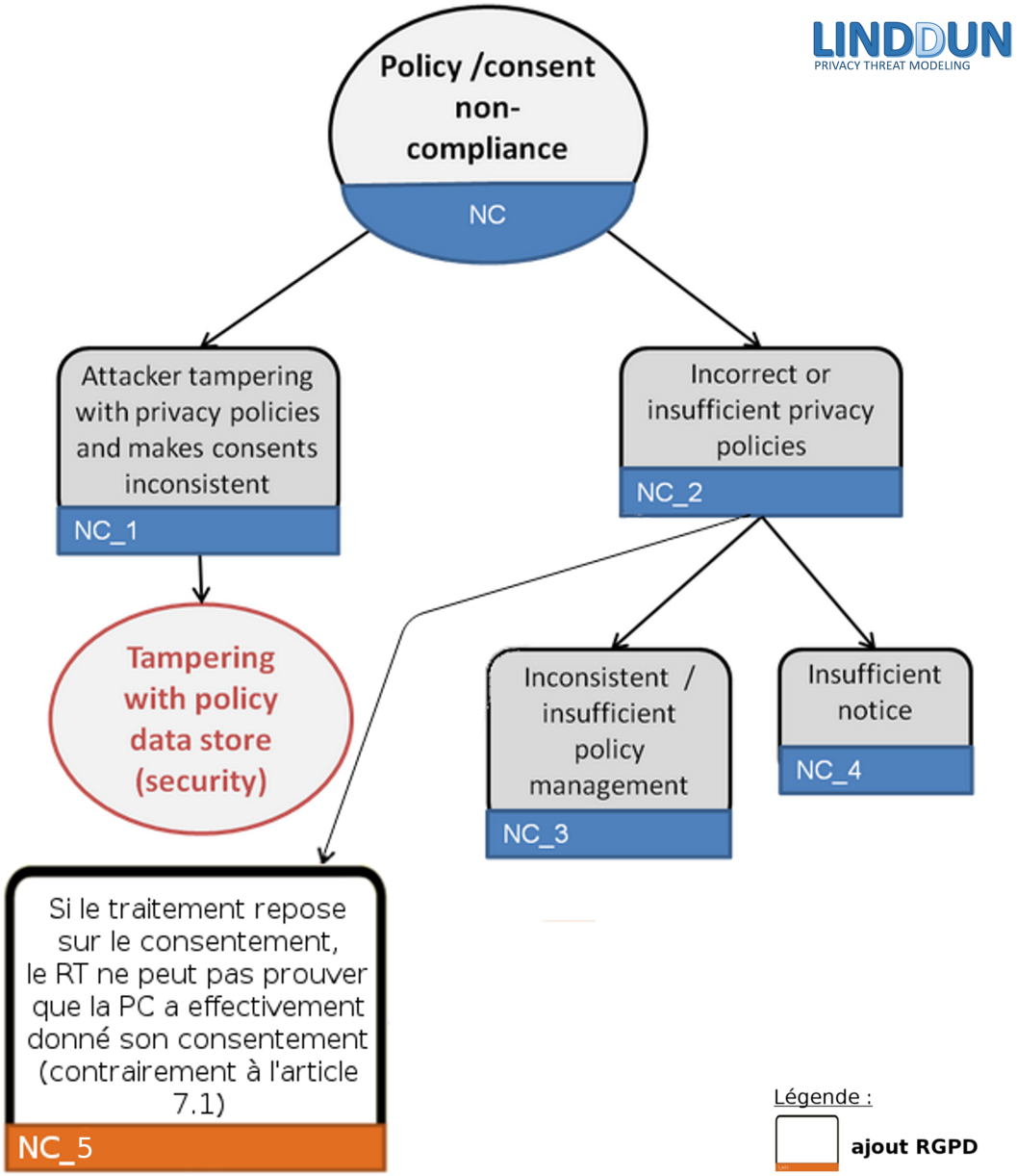


Figure 4.5: Diagramme de "Non-Compliance" de LINDDUN complété

NC_5 : l'article 7[46], qui est titré *"Conditions applicables au consentement"*[46], détaille les règles à suivre par rapport au consentement. En plus du fait d'*"être en mesure de démontrer que la personne concernée a donné son consentement au traitement des données à caractère personnel la concernant"*[46], l'article présente le fait que la demande de consentement doit respecter certaines conditions (détaillées notamment dans l'article 7.1[46])) et pouvoir être révoqué par la personne concernée à tout moment (article 7.3[46])). Il faut aussi noter la nécessité si on veut *"déterminer si le consentement est donné librement"* d'être attentif au fait que le service puisse n'être rendu que si la personne concernée accepte le traitement de données personnelles lui appartenant mais dont le traitement n'est pas obligatoire pour que le service puisse être rendu (article 7.4[46])).[4]

[4]Notons aussi la présence de l'article 8 qui traite exclusivement des *"Conditions applicables au consentement des enfants en ce qui concerne les services de la société de l'information"*.[46]

4.5.5 Processing & Purposes

Les **Traitements et les Finalités** reprennent les articles qui explicitent les bases sur lesquelles doivent reposer les traitements des données et les finalités de ces traitements. Il reprend également les règles spécifiques à leur mise en oeuvre. Ces règles spécifiques ne se retrouvent pas toujours sous une forme similaire dans LINDDUN car elles relèvent de la réglementation et peuvent donc dépasser la définition de la vie privée mise en place par cette méthode de modélisation de menaces.[46]

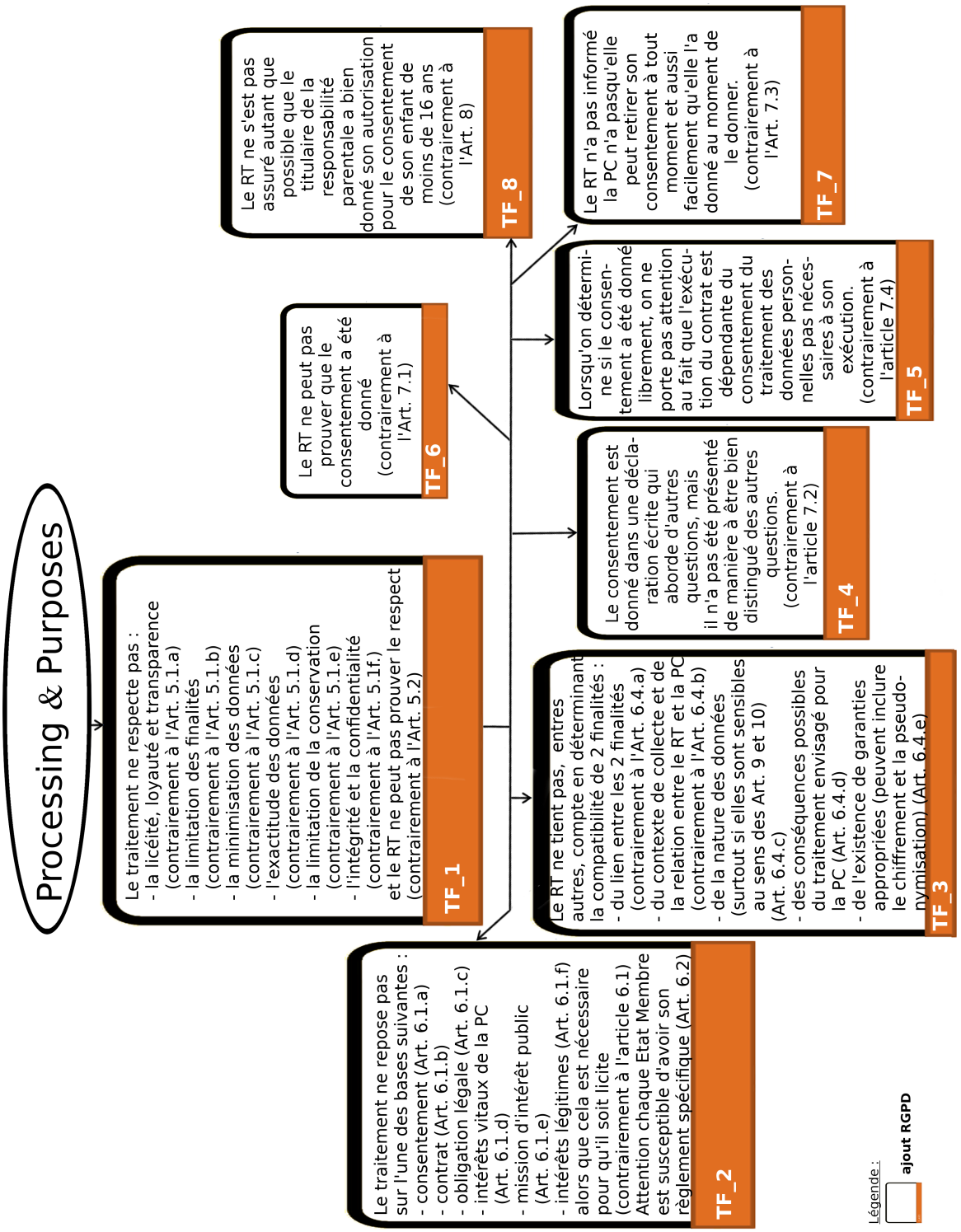


Figure 4.6: Diagramme des "Traitements et Finalités"

TF_1 : ce premier élément de l'arbre s'intéresse aux conditions qu'un traitement doit respecter pour être en conformité avec le RGPD. Elles sont détaillées dans l'article 5[46]. Certains de ces éléments sont approfondis dans d'autres articles. Ainsi, certaines autres feuilles présentes sur ce diagramme exposent les approfondissements en question. Tous les éléments que doit respecter le traitement ont été repris ici dans un souci d'exhaustivité et pour rappeler le contexte qui doit être celui de tout traitement. Même

si ce mémoire n'a pas pu traiter en profondeur tous ces éléments, ils sont tous nécessaires à chaque traitement et doivent être présents dans l'application du RGPD. Les indiquer ici permet donc de les garder à l'esprit même s'ils ne sont pas détaillés dans ce travail particulier.

TF_2 : l'article 6[46] définit les bases juridiques reconnues par le RGPD sur lesquelles peut reposer le traitement.

TF_3 : reprend le sous-point de l'article 6 qui s'intéresse aux paramètres que le RGPD cite comme devant être pris en compte avec les autres paramètres pertinents lors de la détermination de la compatibilité possible entre 2 finalités. Cela peut avoir lieu, par exemple, pour déterminer si les données récoltées pour une première finalité pourraient potentiellement être utilisées dans le cadre d'une nouvelle finalité.[46]

TF_4 : Le détail de cet article a été repris dans l'élément *U_12* se trouvant dans le point traitant du "Content Unawareness" ci-dessus.

TF_5 : l'article 7.4[46] détaille ce qui doit être pris en compte lors d'une détermination de la liberté de donner le consentement.

TF_6 : L'article 7.1[46] demande de garder le moyen de démontrer que le consentement a bien été reçu de la personne concernée pour le traitement en question.

TF_7 : Le détail de cet article a été repris dans l'élément *U_13* se trouvant dans le point traitant du "Content Unawareness" ci-dessus.

TF_8 : Cet élément, tiré de l'article 8[46], s'intéresse au consentement des enfants (c'est-à-dire ceux et celles qui ont moins de 16 ans selon le texte du RGPD, cela peut différer en fonction des pays comme cela est précisé dans le même article) et au fait que le responsable du traitement doit mettre en place les mesures nécessaires pour s'assurer que *"le titulaire de la responsabilité parentale à l'égard de l'enfant"* a donné son accord ou son consentement au traitement.

4.5.6 Rights of Data Subject

Les **Droits des personnes concernées** sont des droits qui sont donnés par le RGPD et peuvent donc posséder des exigences propres à ce règlement que la définition de la vie privée sur laquelle LINDDUN se base n'englobe pas entièrement.

Figure 4.7: Diagramme des Rights of Data Subject - première partie

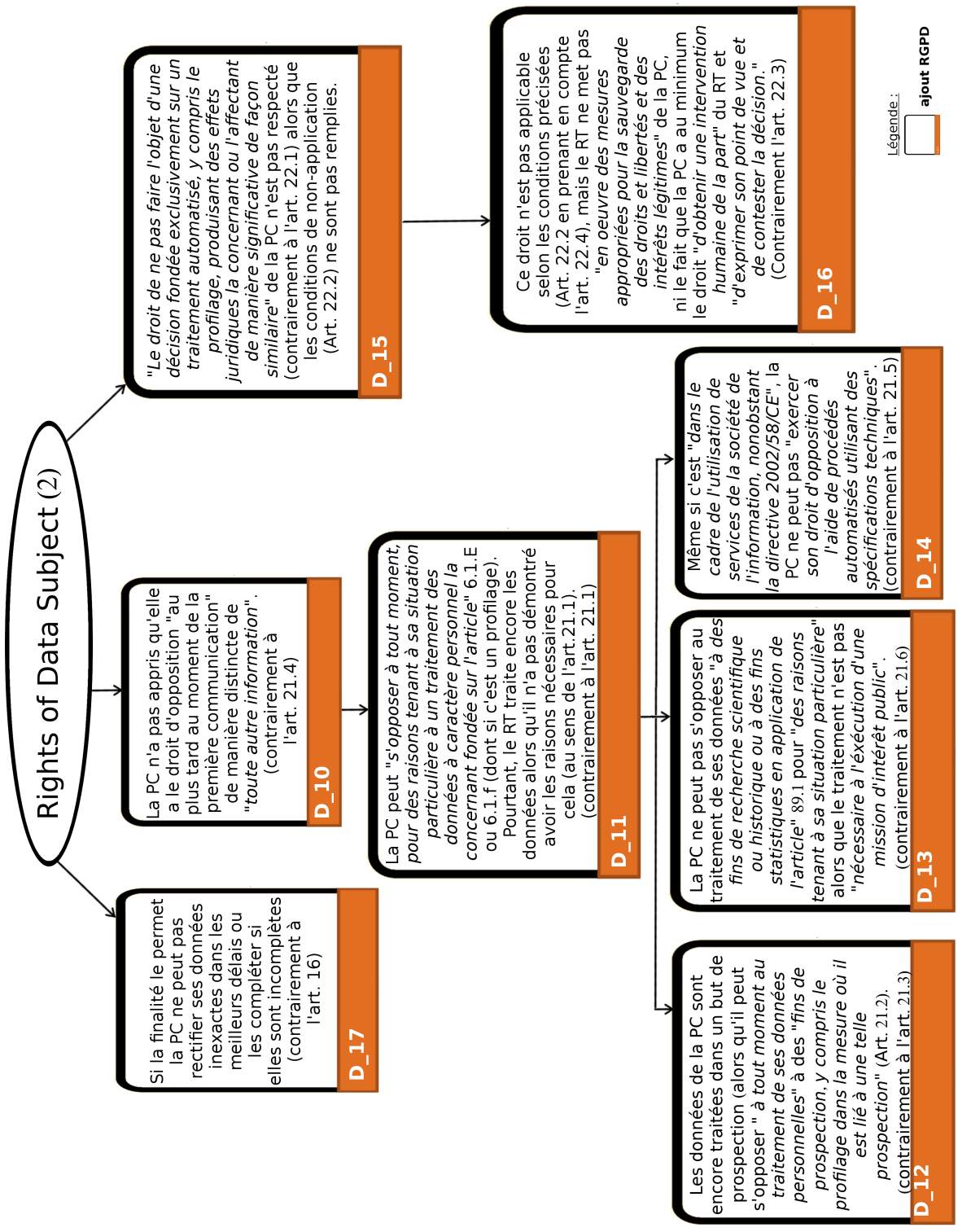


Figure 4.8: Diagramme des Rights of Data Subject - deuxième partie

- **"Droit d'accès de la personne concernée"**

D_1 : l'article 15.1[46] a déjà été détaillé dans l'"Unawareness" dans les éléments *U_15* et *U_17* qui a reprécisé les informations qui doivent être transmises à la personne concernée quant au fait que ses données soient traitées (ou non).

D_2 : Cet élément a déjà été détaillé dans l'élément *U_16* de l'"Unawareness.

D_3 : L'article 15.3)[46] fait notamment état du fait que les données doivent être transmises à la personne concernée de la manière dont elle en fait la demande ou par *"voie électronique"*[46] si elle a elle-même fait sa demande par *"voie électronique"*[46]

• **"Droit à la rectification"**
D_17 : l'article 16[46] stipule que la personne concernée a le droit de rectifier ses données inexactes *"dans les meilleurs délais"*[46] et de compléter des données si elles sont incomplètes et si cela est cohérent avec les finalités.

• **"Droit à l'effacement"**
D_4 : l'article 17.1[46] stipule que le ou la responsable du traitement doit effacer les données de la personne concernée si celle-ci se trouve dans l'un des cas où ce droit s'applique (comme expliqué dans les sous-points de l'article 17.1[46])

D_5 : l'article 17.2 concerne le cas où les données qui doivent être effacées ont été rendues accessibles au public. Dans ce cas, il faut prendre des *"mesures raisonnables"*[46] en mettant en place des moyens selon les *"technologies disponibles"*[46] et en prenant en compte le coût que cela suppose. Le cas des sous-traitants est laissé de côté comme cela ne rentre pas dans le cadre du mémoire.

• **"Obligation de notification en ce qui concerne la rectification, ou l'effacement de données à caractère personnel ou la limitation du traitement"**
D_6 : l'article 19[46] précise que le responsable du traitement doit notifier les rectifications, l'effacement ou la limitation du traitement si cela est possible.

D_7 : Le responsable du traitement doit fournir à la personne concernée des informations sur le ou les destinataires si elle le demande.[46]

• **"Droit à la portabilité des données"**
D_8 : Si la personne concernée demande au responsable du traitement de transférer ses données personnelles à un second responsable du traitement, le responsable du traitement originel doit transférer les données comme cela est exposé dans l'article 20.2)[46]. Le détail concernant les conditions d'application de ce droit se trouvent dans les articles 20.3) et 20.4)[46].

D_9 : l'article 20.1[46] expose le droit de la personne concernée à recevoir ses données dans un format *"structuré, couramment utilisé et lisible par machine"* et à pouvoir les transmettre à un second responsable du traitement si la personne concernée le veut sans que le premier responsable du traitement ne s'y oppose à condition que le traitement ait lieu sur base du consentement (comme cela est détaillé dans l'article 21.1.a)[46]) ou soit effectué à base de procédés automatisés (comme détaillé dans l'article 21.1.b)[46]).

• **"Droit d'opposition"**
D_10 : la personne concernée doit apprendre l'existence de son droit d'opposition au traitement *"au plus tard au moment de la première communication"* et cela *"de manière distincte à toute autre information"* comme cela est établi par l'article 21.4[46].

D_11 : la personne concernée peut s'opposer à tout moment à un traitement si celui-ci est basé sur la nécessité *"[...] d'exécuter une mission d'intérêt public ou relevant de l'exercice de l'autorité publique dont est investi le responsable du traitement"*[5] ou *"nécessaire aux fins d'intérêts légitimes poursuivis par le responsable du traitement ou par un tiers [...]"*[6]. Le RGPD précise aussi que cela concerne le profilage. Le responsable du traitement doit dès lors arrêter de traiter les données à moins qu'il ne prouve qu'il a les raisons nécessaires pour cela. Ces éléments sont détaillés dans l'article 21.1)[46].

[5] article 6, paragraphe 1, point e)[46]
[6] article 6, paragraphe 1, point f)[46]

D_12 : la personne concernée peut s'opposer *"à tout moment au traitement des données à caractère personnel la concernant à de telles fins de prospection"*. Le RGPD précise de nouveau que cela inclut le profilage dans ce but. Le détail se trouve à l'article 21.2[46]. Si la personne concernée décide de faire cela, le responsable du traitement doit arrêter de traiter les données dans ce but comme cela est précisé dans l'artricle 21.3)[46].

D_13 : l'article 21.6[46] précise que *"Lorsque des données à caractère personnel sont traitées à des fins de recherche scientifique ou historique ou à des fins statistiques en application de l'article 89, paragraphe 1, la personne concernée a le droit de s'opposer, pour des raisons tenant à sa situation particulière , au traitement de données à caractère personnel la concernant, à moins que le traitement ne soit nécessaire à l'exécution d'une mission d'intérêt public."*[46]

D_14 : l'article 21.5[46] expose que *"dans le cadre de l'utilisation de services de la société de l'information et nonobstant la directive 2002/58CE, la personne concernée peut exercer son droit d'opposition à l'aide de procédés automatisés utilisant des spécifications techniques"*[46]. Il faut donc prévoir ce mécanisme de manière informatique.

• **"Décision individuelle automatisée, y compris le profilage"**
D_15 : la personne concernée a *"le droit de ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé, y compris le profilage, produisant des effets juridiques la concernant ou l'affectant de manière significative de façon similaire"*[46] comme détaillé dans l'article 22.1[46]. Il existe cependant dans le RGPD, des conditions sous lesquelles ce droit n'est pas applicable qui sont exposées dans l'article 22.2[46] en prenant en compte l'article 22.4)[46].

D_16 : Dans le cadre des conditions de non-application de ce droit de ne pas *"faire l'objet d'une décision fondée exclusivement sur un traitement automatisé, y compris le profilage"*[46] définies aux article 22.2.a) et 22.2.c) et 22.4)[46], le reponsable du traitement doit prendre "des mesures appropriées pour la sauvegarde des droits et libertés et des intérêts légitimes de la personne concernée".

4.5.7 Déroulement de l'étape 3 selon LINDDUN

La suite de ce point va être organisé en 3 sous-points qui sont les 3 sous-étapes déterminées par LINDDUN.[48]

"Refine threat via threat pattern"
Chaque croix faite dans le tableau de l'étape 2 peut être précisée grâce à ces arbres de menaces qui sont présents sur le site officiel de LINDDUN. [48] Dans le cadre de la méthodologie, les arbres utilisés sont ceux présentés ci-dessus.

"Document assumptions"
A partir des arbres déterminés comme potentiellement intéressants juste au dessus, il est à présent temps de déterminer quelles sont les branches ou les feuilles de l'arbre qui sont réellement pertinentes dans le cas du système analysé. Certains choix vont reposer sur des hypothèses qui doivent être documentées (LINDDUN parle d'un format texte et/ou d'un *misuse case* pour des raisons liées à la traçabilité). L'idée est aussi de pouvoir suivre des changements éventuels dans les hypothèses posées.[48]

Il semble important de décrire le raisonnement suivi avec la méthodologie étendue dans le cadre du RGPD. Dans ce cadre, il serait avisé de faire appel au délégué à la protection des données qui a une meilleure connaissance du texte de loi. Son rôle peut être crucial dans la réflexion puisqu'il est censé être capable d'établir avec certitude qu'un élément respecte le RGPD ou doit être amélioré.

"Document threat using threat template"
LINDDUN propose l'utilisation de *misuse cases* pour décrire les menaces et propose une manière de formater la présentation des différentes menaces. Le détail de cette manière de formater n'a pas été repris ici.[48]

Les *misuse cases* pourraient se révéler moins cruciaux dans le cadre d'une analyse portée sur le RGPD. En effet, les différences de point de vue global qui existent entre LINDDUN et l'analyse seulement concentrée sur la tentative de déterminer des Zones de risque de non-respect du RGPD rendent les *misuse cases* potentiellement différents. Pour rappel, LINDDUN prend le point de vue d'un attaquant visant le système alors que notre démarche vise à identifier les Zones de risque propres au RGPD internes au système lui-même. Nous n'avons pas cherché à développer les *misuse cases* interférant avec le RGPD dans ce travail.

4.5.8 Conclusion de l'étape 3

A présent que la situation a été analysée pour ce qu'elle est et que de potentielles Zones de risque ont été repérées, les étapes qui suivent vont tenter de guider l'adoption de mesures concrètes permettant d'amoindrir ces Zones de risque.

4.6 Etape 4 : Prioriser les menaces

L'étape 4 va guider la priorisation des menaces en fonction du RGPD. LINDDUN ne propose pas une méthode de priorisation des menaces et se limite à rappeller une des manières préexistantes de calculer le risque (Risk = likehood x impact).[48]

4.6.1 Utiliser la définition du risque de l'AIPD ?

Du point de vue du RGPD, il existe une manière de classer des risques d'atteinte à la vie privée selon certains critères définis. Pour cela, la notion d'AIPD est intéressante.

L'AIPD est définie dans l'article 35 du RGPD[46] et est un processus particulier qui n'est nécessaire que lorsque le traitement particulier envisagé *"est susceptible d'engendrer un risque élevé pour les droits et libertés des personnes physiques"*.

S'il n'est donc pas toujours nécessaire d'exécuter une analyse d'impact, il peut être intéressant de remarquer que le RGPD considère une définition particulière du risque dans sa réflexion. Elle est définie plus en détails par l'APD dans la FAQ en rapport avec l'AIPD. Elle explique que la définition du risque a été précisé dans le considérant 75 du RGPD[46] et par le groupe de travail "Article 29". L'APD présente ainsi dans sa réponse à cette question que " *Ce qui doit être considéré comme un risque élevé devra toujours se baser sur une évaluation concrète du responsable du traitement. Ce dernier peut dans ce cadre utiliser une combinaison d'une part des lignes directrices des contrôleurs (la Commission vie privée et les autres DPA européennes) et d'autre part des listes de risques et des méthodes connues pour réaliser une évaluation des risques appropriée, à la lumière des traitements qu'il effectue.*"[56]
La CNIL[7], quant à elle, définit la manière de présenter un risque comme : *"Un 'risque sur la vie privée' est un scénario décrivant : un événement redouté (atteinte à la confidentialité, la disponibilité ou l'intégrité des données, et ses impacts potentiels sur les droits et libertés des personnes) ; toutes les menaces qui permettraient qu'il survienne. Il est estimé en termes de gravité et de vraisemblance. La gravité doit être évaluée pour les personnes concernées, et non pour l'organisme."*[57]

Les situations qui sont porteuses d'un risque élevé sont définies dans l'article 35.3 du RGPD[46] et dans une liste fournie par L'APD[58] auxquelles peuvent s'ajouter les situations considérées comme risquées par le ou la responsable du traitement. Dans ce cas-là, une AIPD est obligatoire. Le but de l'étape ici présente n'est pas de réaliser une AIPD mais de simplement décider quelles sont les Zones de risque qui seront traitées en priorité. Si une AIPD est nécessaire, elle ne concerne pas seulement l'informaticien et doit être faite, conformément au RGPD, en impliquant d'autres membres de

[7] La CNIL est la "Comission Nationale de l'Informatique et des Libertés". Il s'agit d'un organisme français. Cependant, comme les articles et ressources repris ici se concentrent principalement sur le RGPD et non pas sur les précisions que la France lui a ajouté, la citer comme source de ce mémoire qui s'intéresse à l'application du RGPD dans un contexte belge, reste cohérent.

l'organisation que les seuls informaticiens. A cela s'ajoute le fait qu'il est par contre possible que l'informaticien soit impliqué par son organisation lors de la réalisation de l'AIPD car, comme l'APD l'écrit elle-même, lorsqu'une AIPD est réalisée, elle fait intervenir *"les bonnes personnes au sein de l'entreprise"* comme par exemple, toujours selon l'APD les *"concepteurs de nouvelles applications"* et *"ceux qui prennent des décisions stratégiques en matière de développement de projets"*.[59] On peut aussi citer le fait qu'il existe des logiciels guidant la réalisation de l'AIPD, par exemple *"l'Outil PIA"* de la CNIL.[60] Il y a dès lors peu d'intérêt à inclure cette procédure particulière dans une première extension de LINDDUN dans le cadre du RGPD.

La méthodologie ne concernera donc pas la réalisation de cette AIPD. Cependant, il est potentiellement possible de réutiliser la définition du risque pour prioriser les Zones de risque de potentielle non-conformité et donc pouvoir déterminer lesquelles d'entre elles doivent être corrigées en priorité. Les précisions apportée par l'APD[59] et la CNIL[60] rappellent néanmoins que cette apréciation du risque et du degré de risque est très liée à l'analyse propre au responsable du traitement.

Gardons aussi à l'esprit qu'une priorisation des zones de risque de non-conformité doit quand même arriver à un respect complet du RGPD après résolution de ces Zones de risque. Si cette méthodologie veut aider à identifier des Zones de risque de potentiel non-respect qui doivent ensuite être laissées à l'appréciation des personnes utilisant cette méthodologie, elle ne peut pas garantir une conformité correcte et n'a pas vocation à le faire.
Le seul moyen de s'assurer une conformité correcte est de consulter le délégué à la protection des données ou quelqu'un ayant les compétences nécessaires pour certifier ce respect correct selon le RGPD. D'ailleurs, dans l'article 35 du RGPD détaillant l'analyse d'impact , il est rappelé que le DPO doit être consulté si un tel délégué a été désigné.

4.6.2 Utiliser un autre type de méthode de priorisation ?

Pour cette étape, il est donc possible d'appliquer une priorisation :
- soit liée à la définition des risques développée dans le RGPD,
- soit de considérer comme prioritaire toute Zone de risque de non-conformité au RGPD et comme moins prioritaire, toute autre menace pour la vie privée qui va au-delà des demandes du RGPD mais que l'organisation souhaiterait quand même mettre en place. Le problème d'une telle méthode de priorisation est que le RGPD insiste déjà sur une limitation stricte des données traitées en fonction des finalités poursuivies et que tout choix d'une organisation d'utiliser par exemple des statistiques anonymes au lieu d'utiliser des données personnelles aura théoriquement été déterminé dans le cahier des charges de la création du système d'information. Cependant, dans le cadre d'une adaptation d'un système existant, cette dernière objection sera moins pertinente car les choix faits dans le cadre de l'adaptation peuvent potentiellement différer des choix réalisés précédemment lors de la création du-dit système,
- soit de réaliser la priorisation avec le délégué à la protection des données qui pourrait déterminer quelles sont les Zones de risque potentielles représentant un réel risque de non-conformité et leur donner la priorité maximale. Les Zones de risque de potentielle non-conformité pourraient dès lors recevoir une priorité plus basse. Cette dernière méthode se rapproche de celle présentée en premier et permet à un informaticien d'avoir confirmation par le délégué à la protection des données des zones de risque de non-conformité. Le travail réalisé par cette méthodologie donnant alors l'occasion aux 2 parties d'utiliser leurs compétences respectives pour assurer un meilleur respect du RGPD.

4.6.3 Conclusion de l'étape 4

D'une manière comparable à ce que fait LINDDUN, le choix de la méthode de priorisation va donc être dépendante de qui fait l'analyse, du contexte de l'organisation dans laquelle est faite l'analyse et du temps qui peut être alloué pour la résolution des Zones de risque découvertes au fur et à mesure de l'analyse. Il est donc plus logique de laisser l'apréciation des différentes stratégies possibles de priorisation à celui qui a toutes les informations relatives au contexte de l'organisation.

4.7 Etape 5 : Identifier des stratégies d'atténuation

4.7.1 Introduction

Les stratégies d'atténuation ont pour but de réduire les risques qu'une menace se réalise. Il s'agit de choisir comment agir pour éviter que la menace ne puisse se réaliser. LINDDUN s'appuie sur 2 stratégies principales, l'une "proactive" qui cherche à ce que l'utilisateur ne fournisse pas plus d'informations que nécessaire et l'autre "réactive" qui tend à s'assurer que l'identité de celui qui se connecte reste confidentielle et/ou à s'assurer que les données qui passent par le système ne soient pas compromises. Tout en s'assurant que les données soient aussi correctes que possible et en évitant qu'elles soient exposées.[48]

Pour aider à sélectionner ces stratégies, LINDDUN met à la dispositions des analystes un document présentant la taxonomie des stratégies d'atténuation des risques et faisant le lien entre les différentes stratégies et les arbres de menaces rendus accessibles par LINDDUN.[48]

Dans un premier temps, si on s'intéresse au cas du respect du RGPD, il est évident que le but est de réussir à respecter l'intégralité de la loi. Cet élément en tête, il est alors intéressant de se demander quelles sont les stratégies informatiques qui peuvent être mises en place pour arriver à ce but. Le RGPD fait mention de certaines d'entre elles comme la pseudonymisation au sens du RGPD (détaillée dans l'article 4.5 et mentionnée à plusieurs autres endroits du texte comme dans l'article 25 ou encore 32.1.a)) qui concerne la *"Sécurité du traitement"*)[46]) ou le chiffrement (aussi mentionné dans l'article 32.1.a)) ou encore des *"mesures techniques"* (termes se retrouvant dans de nombreux articles comme les articles 4, 5, 17, 24, 25[46], ...) qui peuvent recouvrir de multiples stratégies et implémentations différentes.

Des décisions de nature technique doivent donc être prises et auront un impact sur la manière dont le RGPD sera mis en oeuvre. C'est pour cela que la taxonomie des stratégies d'atténuation de LINDDUN peut être utilisée dans le cadre de la réflexion propre au RGPD car celui-ci laisse à l'informaticien le choix de certains moyens techniques.

La suite de ce paragraphe va examiner comment les éléments introduits dans les étapes précédentes peuvent interagir avec ce que LINDDUN a déjà défini, c'est-à-dire le diagramme présentant la taxonomie des stratégies d'atténuation, d'une part, et le tableau faisant le lien entre les différents arbres de menaces et les différentes stratégies d'atténuation possibles, d'autre part.[42]

4.7.2 Les stratégies déjà établies par LINDDUN

La taxonomie proposée par LINDDUN est une vue générale de certaines stratégies d'atténuation qui sont couramment utilisées. LINDDUN précise que d'autres stratégies existent.[42]

Le but de ce point est de découvrir si les feuilles qui ont été greffées aux stratégies de menaces peuvent aussi bénéficier de la stratégie d'atténuation qui a été assignée à l'arbre de menaces adapté au RGPD. Dans un premier temps, les stratégies mises en avant par LINDDUN seront examinées. Nous proposerons ensuite les adaptations propres au RGPD.

- **Identifiability**

Au niveau de l'*"Identifiability of entity"*, la stratégie mise en avant est *"protect ID"*, le détail de l'arbre de taxonomie dans ce cas particulier est repris ci-dessous dans la figure 4.9[42].

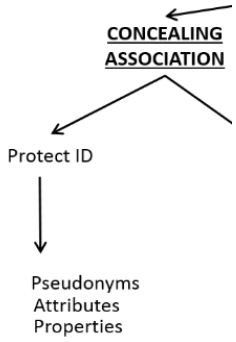


Figure 4.9: "Identifiability : Protect ID"

• **Unawareness**

L'*"Unawareness of entity"* est détaillée en 2 sous-parties. La première est *"Providing too much information"*[55]. Dans ce cas, LINDDUN préconise la stratégie *"Awareness"*. La partie de l'arbre exposant cela est présenté dans la figure 4.10.[42]

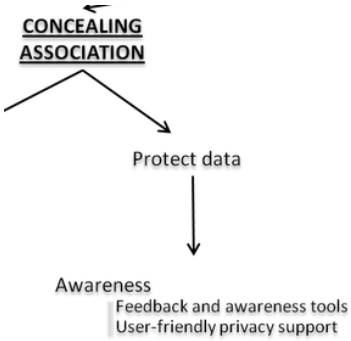


Figure 4.10: "Unawareness : Awareness"

Quant à la seconde, *"Data accuracy"*[55], c'est la stratégie *"Review Data"* qui est mise en avant, la partie précise de l'arbre de taxonomie est présentée dans la figure 4.11. [42]

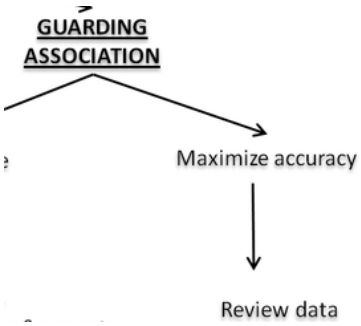


Figure 4.11: "Unawareness : Review Data"

• **Non-compliance**

Pour la menace de *"Non-compliance"* LINDDUN privilégie l'approche *"Gard Exposure"* qui cherche à éviter que les données ne soient révélées à quelqu'un n'étant pas censé y avoir accès, le détail de l'arbre présentant la taxonomie des stratégies d'atténuation dans ce cadre particulier est présenté dans la figure 4.12[42].

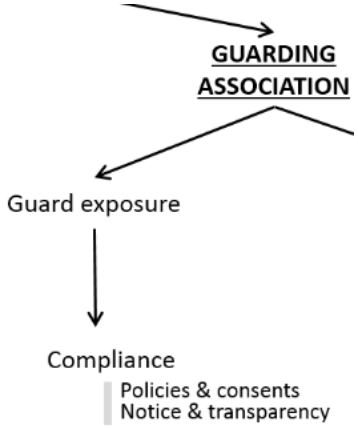


Figure 4.12: "Non-compliance : Gard exposure compliance"

4.7.3 Qu'en est-il des éléments du RGPD placés dans des arbres de menaces à l'étape précédente ?

Les éléments rajoutés en rapport avec le RGPD peuvent bénéficier des stratégies préétablies par LINDDUN. Cependant, certains éléments tirés d'articles particuliers demandent une stratégie plus précise. Par exemple, si l'article 13 (qui détaille les informations qui doivent être précisées à la personne concernée lorsqu'on récolte les données auprès d'elle) n'est pas encore respecté, la seule stratégie acceptable est de faire ce que l'article demande, c'est-à-dire donner les informations détaillées dans cet article à la personne concernée.

En effet, si certains articles parlent de "mesures techniques" ou laissent les détails de l'implémentation au soin de l'informaticien, bon nombre de ces détails ne se placent pas au niveau des spécificités de l'implémentation et déterminent ce qui doit être fait quelles que soient les technologies utilisées ou les stratégies d'implémentation que l'informaticien peut choisir de mettre en place lors de la conception ou de la modification d'une architecture logicielle.

Dans le cas où le détail technique peut potentiellement porter à conséquence sur la manière dont le système d'information répond au RGPD, les stratégies privilégiées par LINDDUN prennent tout leur sens. En effet, elles peuvent influencer positivement le niveau de *privacy* que l'implémentation peut permettre.

Dans le deuxième cas, celui où le RGPD impose un objectif non-détaillé, la démarche reste importante parce que, bien qu'il peut parraître "facile" de "simplement" faire ce qui est demandé, une réflexion technique préalable plus poussée peut être nécessaire pour l'implémentation d'autres articles. Ce qui peut avoir des conséquences sur l'architecture globale du système. L'article 20, qui impose que le système doit pouvoir ressortir les données de n'importe quel utilisateur si celui-ci en fait la demande, en est un bon exemple.

Il est quand même possible de rapprocher les 2 nouveaux arbres de menaces que sont les "Traitements de données et finalités" et les "Droits des personnes concernées" de certaines stratégies générales, à l'image de ce que LINDDUN a fait pour ses propres menaces. La stratégie à appliquer ici se rapproche de celle choisie par LINDDUN vis-à-vis de la "non-compliance". Cela est dû au fait que le RGPD, par son caractère prescriptif, guide la mise en conformité.

4.8 Etape 6 : Sélectionner les PETs correspondantes

Cette dernière étape est principalement l'occasion de transformer les stratégies sélectionnées à l'étape précédente en des exigences qui peuvent être utilisées dans une étape de développement futur.[48]

Pour faciliter la démarche, LINDDUN met à disposition une table de solutions.[42] Pour cela les **Privacy Enhancing Technologies** (PETs), un ensemble de technologies qui a pour but d'aider à améliorer la gestion de la protection de la vie privée, peut s'avérer utile. Le terme "PETs" peut recouvrir de nombreux types d'outils dont les technologies qui cherchent à protéger les données personnelles et l'identité numérique d'une personne.[61].

Dans le cadre des mesures techniques à prendre en fonction du RGPD et des conclusions de l'étape 5, il est possible de garder un raisonnement similaire. Dans le cadre des articles pointant vers des "mesures techniques", les recommandations de LINDDUN sont cohérentes également pour le RGPD. Dans le cadre des articles détaillant (pour reprendre l'exemple de l'article 13) précisément ce qui doit être fait, les mesures techniques à prendre découlent directement de l'objectif à atteindre.

LINDDUN cite dans sa table de solutions toute une série de technologies existantes qui aident à améliorer le degré de vie privée. Dans le cadre du RGPD, certaines de ces technologies sont très intéressantes pour aider à répondre à certains articles. Elles ne vont pas toutes être rappelées ici mais sont disponibles sur le site officiel de LINDDUN, dans la partie traitant des stratégies d'atténuation.[42]

La suite va s'intéresser au cas des PETs spécifiques au RGPD.

4.8.1 PETs spécifiques au RGPD

Plusieurs recherches et articles traitent des liens qui peuvent être explicitement établis entre le RGPD et des PETs. Certaines technologies pouvant aider à implémenter la conformité au RGPD pendant le développement ou la modification d'un système d'information y sont présentées. La suite de ce paragraphe va en détailler quelques-unes qui sont liées au sous-ensemble du RGPD examiné.

Rappelons que le fait d'utiliser des PETs peut aider à réaliser la stratégie choisie mais ne peut pas garantir que le RGPD est respecté. N'oublions pas non plus que les conditions d'application ne sont pas gérées par l'informaticien lui-même. Ces conditions relèvent du domaine juridique.

• **Identifiability**
- *Nécessité de conserver séparément les données pseudonymisées et les données permettant de les relier à quelqu'un* (I_e21)
Une recherche de l'Université de Leicester est allée plus en détails sur le sujet de la pseudonymisation au sens du RGPD et des moyens qui pouvaient être mis en place pour l'appliquer. [62] Ce mémoire n'a pas retenu le processus de pseudonymisation comme l'un des éléments examinés. Il ne s'intéresse ici qu'au fait que de telles données doivent être conservées de manière à ce que les données pseudonymisées ne puissent pas être identifiées comme appartenant à une personne en particulier. Les mesures informatiques peuvent être une séparation physique des données, une politique d'*access control* spécifiques, ... Cela dépendra du contexte technique et des choix de l'analyste menant l'analyse du système.

• **Unawareness**
- *Transparence* (U_6)
La transparence a un statut particulier. Si elle peut être implémentée, elle peut aussi être considérée comme un PETs en elle-même. [63]. LINDDUN la place même dans le détail d'une de ses stratégies d'atténuation et ne lui assigne pas de PETs.[42]

- *Obligation de notification aux destinataires* (U_7)
Cette obligation sera développée en plus de détails dans le point concernant les "Droits des personnes concernées".

- *Informations qui doivent être données à la personne concernée* (U_8, U_9, U_12, U_15) et *Informations quant au consentement* (U_10, U_11)
Les informations à donner à l'utilisateur sont détaillées dans le tableau de synthèse des figures 3.2 et

3.3 et dans le détail des éléments de l'arbre de menaces qui sont rappelés dans le titre de ce point. Ces informations doivent être données à l'utilisateur selon les modalités précisées dans l'article. Utiliser une PETs serait peu pertinent car cela ajouterait des étapes supplémentaires qui prendraient plus de temps que de faire ce que l'article demande en affichant les informations selon les modalités prescrites.

- *Droit qu'a une personne de savoir si ses données sont oui ou non traitées* (U_13, U_14)
Ce droit est traité dans son ensemble dans la menace "Rights of Data Subject".

• **Non-compliance**
- *Nécessité de pouvoir prouver que le consentement a effectivement été donné* (NC_5)
La gestion du consentement dans son ensemble est détaillée dans la menace ci-après.

• **Traitements de données et Finalités**
- *Conditions que le traitement doit respecter* (TF_2)
Si toutes les conditions que doivent respecter les traitements ont été rappelées dans cet élément de l'arbre de menaces, ce mémoire s'intéresse principalement à la transparence qui est détaillée ci-dessus dans le point sur l'Unawareness.

- *Gestion des finalités* (TF_3)
La gestion des finalités ne dépend pas entièrement de l'informaticien. Cela dépend de l'organisation pour laquelle il travaille. Ce que l'informaticien peut faire est de s'assurer que chacun des traitements qu'il veut implémenter dans son sytème possède une finalité (par exemple par l'intermédiaire du Dataflow Diagram présenté dans l'étape 1). Si un traitement n'en a pas, l'informaticien peut alors en informer les personnes se chargeant de les définir (dont il peut éventuellement faire partie).
Les finalités peuvent aussi être gérées et présentées à l'utilisateur par l'intermédiaire d'un "Privacy dashboard" comme celui développé par l'Université de Vienne dans son article "Designing a GDPR-compliant and Usable Privacy Dashboard" qui informe l'utilisateur, entres autres, des finalités de chaque traitement. [64]

- *Base juridique : consentement* (TF_1, TF_4 à TF_8)
L'élément de l'arbre de menaces concernant les bases juridiques est principalement adressé à ceux qui déterminent les aspects fondamentaux qui concernent les traitements. Ceci signifie que cela dépasse la sphère de l'informaticien étant donné qu'il travaille dans une entreprise ou une ASBL qui a sa propre vision dont l'informaticien doit tenir compte pour faire son travail.
Le consentement est un sujet complexe. La recherche qui s'est attachée au consentement et aux challenges qu'il présentait s'est attardée sur sa révocation. Elle mentionne aussi un nombre important de recherches sur ce sujet ainsi que quelques technologies pouvant aider à mettre en place le respect de ce consentement. Parmi les pistes citées se trouve la procédure User-Managment Access (UMA) qui est un web standard. Elle consiste en une architecture OAuth qui permet aux applications compatibles de mieux gérer le consentement.[65]
Il existe aussi des langages d'access control comme E-P3P qui peuvent aider à gérer des polices d'utilisation lisibles par machine et les choix d'"opt-in" ou d'"opt-out" des utilisateurs. [66]
Un autre moyen de gérer le consentement est d'utiliser un "Privacy Dashboard" auquel l'utilisateur a accès, qui peut lui permettre d'avoir accès aux informations concernant les traitements faits sur ses données et qui lui donne l'occasion d'ajuster son (absence de) consentement pour chaque traitement. Un tel "privacy dashboard" a été présenté par des chercheurs de l'Université de Vienne.[64]

•**Droits des personnes concernées**
- *Droit d'accès* (D_1 à D_3) et *Droit de rectification* (D_17)
L'utilisation de "Privacy dashboard" a aussi été réfléchie vis-à-vis du RGPD comme dans cette recherche de l'Université de Vienne. Il peut permettre à l'utilisateur d'accéder à l'intégralité de ses données et d'introduire des demandes de rectification et d'effacement des données.[64]

- *Droit à l'effacement* (D_4 et D_5) et *Obligation de notification aux destinataires* (D_6 et D_7)
Ce droit demande que les ordinateurs apprennent à "oublier" certains utilisateurs. La recherche de E. Politou et al. relève le fait que les responsables du traitement doivent à la fois trouver le moyen de

suivre les données et prouver que leur suppression a bien eu lieu. Elle explicite ensuite le fait que si le tracking technologique des informations ne représente pas un problème technique de grande envergure, la capacité de prouver que la suppresion a bien eu lieu peut s'avérer plus complexe, ce qui rend potentiellement plus difficile de s'assurer que tous les destinataires ont été informés du changement. La recherche cite notamment une autre approche : l'Information Flow Control (IFC) qui permet de tagger les données en fonction de leur niveau de confidentialité et leur intégrité (liée au fait que leur source est fiable). IFC peut alors aider à implémenter ces exigences du RGPD.[65]

- *Droit à la portabilité des données* (D_8 et D_9)
La recherche de Lachlan Urquhart, N. Venkata Sailaja et Derek McAuley[67] s'est intéressée à la réalisation de ce droit et privilégie l'utilisation de Personal Information Managment System (PIMS). Cette recherche cite notamment Mydex (qui permet à l'utilisateur d'accéder à un data store privé avec de nombreuses fonctions de gestion de ses informations personnelles et la possibilité d'autoriser des applications-tiers) ou OwnCloud (qui met à disposition une interface pour que l'utilisateur accède à ses informations et des possibilités de synchronisation client. Il donne aussi accès à une API et à la possibilité de créer des plugins permettant de traiter les informations en fonction des choix de l'utilisateur). Cette recherche traite aussi des limitations de ces PIMS.

- *Droit d'opposition* (D_10 à D_14) et *Décision individuelle automatisée dont profilage* (D_15 à D_16)
Comme ce mémoire n'examine que la base juridique qu'est le consentement, la gestion de ces droits est directement liée au consentement. Les détails de la gestion du consentement sont présentés dans le point sur les "Traitements de données et Finalités".

4.8.2 Conclusion de l'étape 6

Il existe une multitude de PETs qui peuvent parfois s'appliquer dans des contextes différents. L'analyste reste donc maître de l'analyse car il connaît le contexte informatique et celui de son organisation. Cela lui permet de prendre les meilleures décisions face à l'application du RGPD.
Cette étape cite certains des PETs existantes mais n'a pas vocation à être exhaustive. Des recherches sur les meilleurs moyens techniques pour appliquer le RGPD sont toujours en cours et le nombre de PETs existantes pourraient encore grandir dans les années à venir.

4.9 Amélioration de la méthodologie et automatisation

Une manière efficace de trouver les premières pistes d'amélioration de la méthodologie est de la confronter à un scénario pratique pour trouver certains éléments qui auraient pu être omis ou sous-estimés lors de l'élaboration théorique. Le détail des améliorations possibles sera repris après chaque étape de la méthodologie appliquée sur le scénario concret.

En ce qui concerne l'automatisation de la méthodologie, il est indéniable qu'une intervention humaine sera toujours nécessaire étant donné que des choix portant sur le sens doivent être faits. Cependant, certaines parties de la méthodologie pourraient vraissemblablement être automatisées.

Chapitre 5

Application à un scénario concret

5.1 Introduction

Si les chapitres précédents présentaient l'état de l'art, le RGPD et l'élaboration théorique de la méthodologie, ce chapitre va s'intéresser à l'application pratique de la méthodologie. Pour cela, nous avons rencontré 3 responsables de la gestion de données d'organismes différents afin d'identifier les procédures mises en oeuvre et les difficultés face au RGPD. Nous avons ensuite appliqué notre méthodologie à ces situations de manière à identifier les améliorations qui pourraient être apportées à notre méthodologie. Ces situations devant demeurer confidentielles, nous avons choisi de créer un cas anonyme permettant de ne présenter que les conclusions les plus intéressantes pour la réflexion. Le cas se place au moment où un système d'information devrait être modifié pour entrer en conformité avec le RGPD. Il est à noter que pour créer un cas regroupant les éléments principaux des autres scénarios, certains détails ont dû être adaptés pour que le cas reste cohérent. Le cas anonyme n'est donc pas une description exacte des cas pratiques examinés lors de la réflexion précédant l'écriture de ce chapitre.

Ce chapitre va donc commencer par présenter le scénario anonyme avant d'examiner comment la méthodologie peut être appliquée en pratique et comment elle pourrait être améliorée par la suite.

5.2 Présentation du scénario

Le musée "Histoire d'ici" est un musée qui se spécialise dans l'histoire locale. Il propose 2 types d'activités principales : la visite de la collection permanente qui est gérée par le musée lui-même et des visites guidées historiques de la ville et de la région où il se trouve. Ce musée existe depuis longtemps et opère grâce à son équipe d'une cinquantaine de personnes.

Le scénario présenté ici va s'intéresser plus précisément aux données des participants aux visites guidées. Les données personnelles du personnel du musée ne seront pas abordées ici car ces données sont fortement soumises à d'autres lois que le RGPD, comme par exemple les lois sur le travail.

Le musée gère des données personnelles lors des réservations en ligne des visites touristiques. Ces visites, qui sont connues pour regorger de détails historiques et de précisions architecturales s'adressent à un public composé principalement d'adultes et d'adolescents de plus de 16 ans. S'il arrive que des enfants visitent le musée en lui-même ou les expositions temporaires, l'achat de tickets doit se faire sur place et aucune donnée personnelle n'est récoltée sur les enfants, le système se souvient juste avoir vendu une carte "- de 16 ans" à un jour et une date particulière. Le musée collabore aussi avec d'autres musées pour proposer un abonnement aux sorties touristiques qui permet d'avoir accès à toute une série d'activités culturelles organisées par les organisations participantes. Les clients qui décident d'en acquérir donnent leurs informations au système du musée "Histoire d'Ici" qui les traite localement et les envoie à un système partagé avec les organisations donnant accès aux autres activités culturelles.

Pour ce qui est du système du musée, les informations sont soit récoltées grâce à un formulaire papier, soit par un formulaire en ligne. Il existe un règlement qui concerne la participation aux activités culturelles et qui aborde la manière dont les données sont utilisées. Le traitement au niveau du musée permet de donner l'occasion aux visiteurs ayant l'abonnement d'accéder à toutes les activités gérées par le musée (visites des expositions permanentes et temporaires, visites guidées de la ville, ...). L'existence du rêglement est signalée lors de l'inscription.

Le musée garde en mémoire l'historique des activités auxquelles un visiteur participe pendant un an. Cela lui permet de faire des statistiques quant à ses activités et de prévoir si nécessaire, plus de places pendant les visites guidées. Le système du musée peut aussi, à tout moment, sortir la liste des clients présents dans chacune des activités. Les données présentes dans cette base de données de clients ne sont en théorie jamais effacées pour permettre à un client de facilement recréer un abonnement même si cela fait des années qu'il n'en a plus eu un. Dans ce dernier cas, les données seront quand même remises à jour sur base de la carte d'identité lors de la création du nouvel abonnement. Les données stockées peuvent néanmoins être effacées ou rectifiées sur demande. L'équipe technique qui gère le système informatique s'est déjà penché sur ces données et chacune d'entre elles a été jugée nécessaire pour l'exécution de la tâche demandée.
Tous les membres du personnel du musée qui s'occupent de la gestions des visiteurs ainsi que l'équipe technique gérant le système informatique ont accès à ces données, chacun d'entre eux à néamoins accès à elles par l'intermédiaire de son compte personnel. L'historique des activités du système est gardé. Les membres suivent aussi une formation sur le thème de la sécurité chaque année qui leur rappelle les bonnes pratiques pour s'assurer que les données restent accessibles seulement à ceux qui peuvent les voir.
Les données sont collectées pour 3 finalités principales qui sont présentées à l'utilisateur lors de son inscription :
• Les statistiques aidant à la gestion et à l'amélioration des conditions pratiques des activités proposées aux visiteurs.
• La gestion des activités et des visiteurs : réduction du prix de l'abonnement en fonction du statut du visiteur (étudiant, professeur, personne âgée, ...), caution supplémentaire en cas de prêt de matériel (écouteurs pour visite guidée pré-enregistrée, matériel spécifique à certaines visites guidées (casque si visite de grotte, ...)), courrier de rappel si le matériel prêté n'a pas été rendu, ...
• La gestion des situations d'urgence comme les évacuations : connaître le nombre de personnes présentes dans le musée permet de savoir si tout le monde a bien réussi à en sortir en cas d'urgence.
Les données permettant de savoir quand une personne est présente dans le musée sont considérées comme sensibles. Bien qu'elles soient gardées par le système, l'analyse suivante ne va pas se concentrer sur elles étant données qu'elles sortent du cadre de la réflexion de ce mémoire tel qu'il a été défini dans les chapitre précédents de ce mémoire.
Les données collectées sont : prénom, nom, date de naissance, code identifiant de l'abonnement, titre (pour adresser les communications), date de naissance (pour s'assurer que le client a bien plus de 16 ans), nationalité (pour les statistiques), adresse légale pour courrier de retard, numéro de téléphone et courriel pour contact, par exemple pour préciser les activités auxquelles l'abonnement donne droit, photo pour le contrôle d'accès et la création de la carte d'abonnement.
La carte d'abonnement est valable pour une période de 3 ans.

Les utilisateurs ont accès à un formulaire leur permettant de rectifier leurs données et peuvent demander à ce que leurs données soient effacées.

Le système partagé est sous la responsabilité de l'une des autres organisations proposant des activités culturelles auxquelles l'abonnement donne accès. Il est donc vu ici comme un sous-traitant. Le détail n'est donc pas retenu ici comme ce mémoire ne traite pas le cas de la sous-traitance.

5.3 Etape 1 : construction du diagramme de flux de données

5.3.1 Application de l'étape 1

Le Data Flow Diagram s'intéresse à la gestion des données des clients dans le cadre de l'abonnement leur permettant de participer aux activitées proposées.

Le type des données est inscrit sur le graphique s'il s'agit de données pseudonymisées (il n'y en a pas dans ce scénario concret), anonymisées ou sensibles. Les données sensibles sont indiquées sur le graphique dans un souci d'exhaustivité mais le détail les concernant ne sera pas détaillé dans la suite comme cela ne fait pas partie du sous-ensemble du RGPD examiné ici.

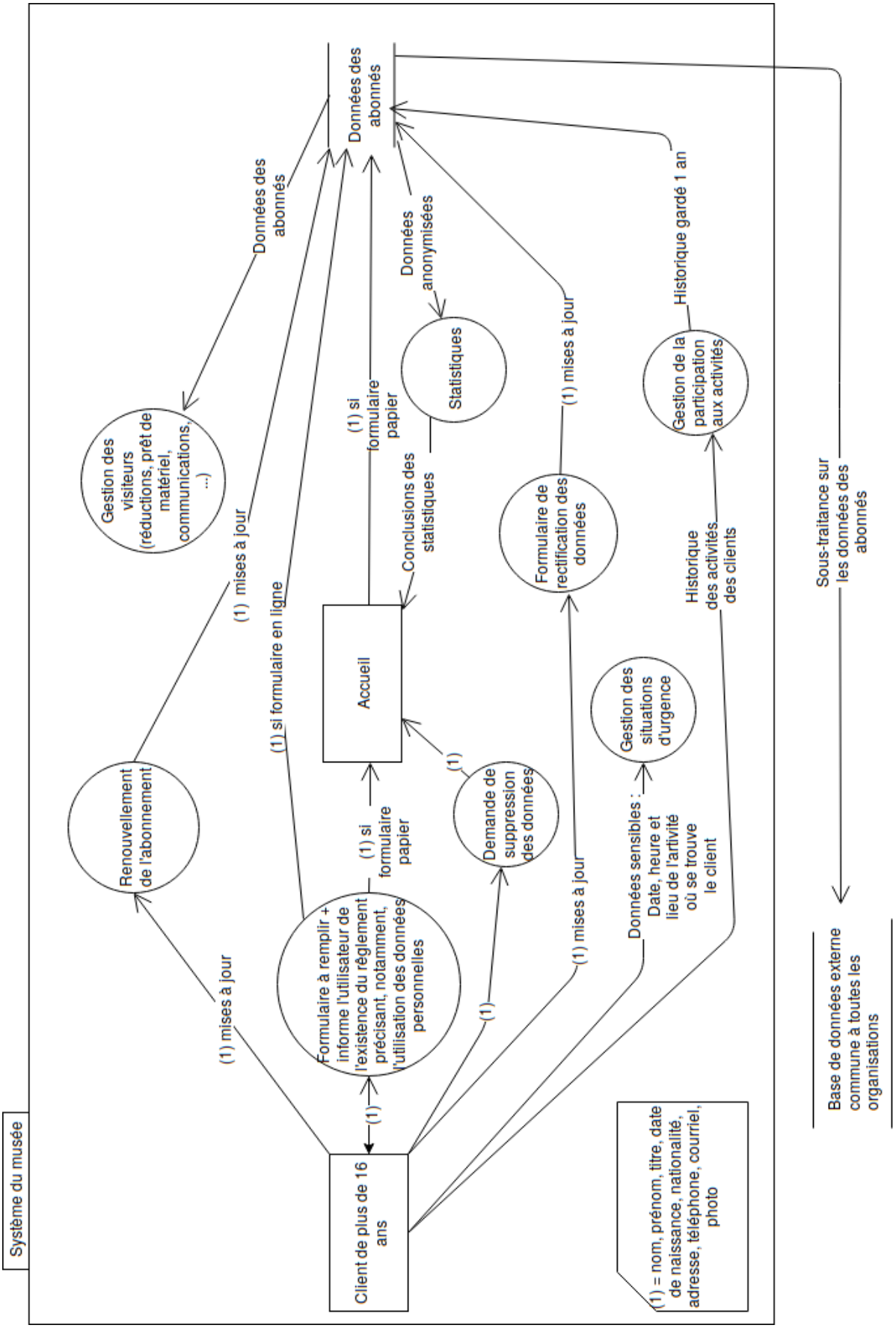


Figure 5.1: Dataflow diagram qui concerne les clients

5.3.2 Conclusions de l'application de l'étape 1

L'application de cette première étape peut donc présenter des pistes quant aux futures améliorations possibles pour la méthodologie.

Sur les cas pratiques de grande taille, les diagrammes de flux de données peuvent rapidement devenir très chargés. L'analyse propre au RGPD demande de rajouter un certain nombre d'informations qui peuvent rapidement compliquer le graphique.
Un moyen de régler le problème est de créer plusieurs dataflow diagrams qui peuvent chacun représenter une partie du système. Par exemple en prenant dans un premier temps une vue très générale avant de la préciser dans d'autres diagrammes s'intéressant chacun à un point plus particulier.

Un autre point est soulevé, celui de la sous-traitance. Ce mémoire a choisi de se concentrer sur d'autres éléments que le cas de la sous-traitance. Il est cependant intéressant de remarquer qu'il est quand même possible d'indiquer sur le graphique la présence de sous-traitants, tout comme les destinataires peuvent être signalés. Si cette information ne permet que de compléter la modélisation, elle pourrait être utile dans une future version de la méthodologie qui prendrait en compte plus d'éléments du RGPD.

Notons aussi que le RGPD s'intéresse à la fois aux mesures techniques et organisationnelles. Or, le diagramme de flux de données ne représente pas systématiquement l'ensemble des mesures organisationnelles qui existent.
Cela met en lumière le fait que le dataflow diagram, s'il est idéal pour observer la manière dont les données bougent dans le système d'information, ne permet pas à lui seul d'avoir une vue d'ensemble suffisante pour assurer un respect total du RGPD. Par exemple, il ne remplace pas une analyse de menaces comme STRIDE lorsque l'on veut examiner de la sécurité du système nécessaire au respect du RGPD, il ne détaille pas les politiques d'accès aux données et reste principalement concentré sur le système informatique.
Cela ne veut pas dire qu'il ne peut pas jouer un rôle clé dans la compréhension du trajet des données dans le système et la prise de décisions quant au RGPD. Garder à l'esprit l'existence des limites peut être particulièrement utile dans le cas où un informaticien voudrait pousser la réflexion plus loin que ce que met en place cette méthodologie.[1]
Pour créer une méthodologie plus complète, il pourrait donc être intéressant de se demander si d'autres types de modélisations ne pourraient pas ajouter des compléments d'informations intéressants à ce que le diagramme de flux de données propose déjà ici.

5.4 Etape 2 : Relier les menaces concernant la vie privée aux éléments du diagramme de flux de données

5.4.1 Application de l'étape 2

L'étape 2 permet de faire le lien entre les éléments du diagramme de flux de données et les menaces pour la vie privée abordées par la méthodologie.

Chacune des croix faites dans le tableau ("x") a été numérotée pour qu'il soit plus facile de faire référence à la case du tableau dans laquelle elle se trouve. Chacune de ces croix sera détaillée par la suite et sa présence dans le tableau sera ainsi justifiée.

Comme il n'y a pas de donnés pseudonymisées dans le scénario concret auquel on s'intéresse ici, la menace "identifiability" n'a pas été sélectionnée. Toutes les autres menaces interviennnent par contre

[1] La définition des limites technologiques doit également être prise en considération dans les choix de politique de l'entreprise face au RGPD. Il est de la responsabilité de l'informaticien d'avoir et de transmettre au DPO une information claire sur ces limites technologiques, même si le choix de l'entreprise sur sa politique d'application du RGPD dépasse sa seule responsabilité.

à un ou plusieurs moments du scénario.

Aucune croix n'a été faite pour les entités (client de plus de 16 ans et Accueil) car le client de plus de 16 ans est l'entité que le RGPD vise à protéger et l'accueil a ici un rôle de transfert des données vers les systèmes informatiques qui se chargent du traitement. Comme l'analyse s'intéresse au point de vue du Responsable du traitement, on se concentre ici sur les flux de données tels qu'ils sont censés se produire. Les cas où une malveillance ou une erreur pourraient se produire pourront être traités par après ou, par exemple, à l'aide d'une méthode d'analyse de sécurité prenant le point de vue d'un attaquant.

Il n'y a pas non plus de croix pour les lignes concernant les "données anonymisées", les "statistiques" (qui ont lieu sur les données anonymisées) et les "résultats des statistiques", cela est dû au fait que le RGPD ne s'applique pas au traitement des données anonymisées (comme cela est expliqué dans le considérant 26 du RGPD[46]).

Le dernier élément du graphique qui n'a pas été retenu comme faisant l'objet d'au moins une menace est le "Renouvellement de l'abonnement", cela est dû au fait que la procédure de renouvellement de l'abonnement est fortement similaire à celle de la création de l'abonnement. L'utilisateur met ses données à jour lorsqu'il renouvelle son abonnement, au lieu de les donner pour la première fois lorsqu'il crée un abonnement. Les menaces qui peuvent s'appliquer au "Renouvellement de l'abonnement" sont donc les mêmes que celles qui peuvent s'appliquer au "Formulaire pour la création de l'abonnement". Pour ne pas répéter une croix et un raisonnement, le choix a donc été fait de sélectionner les menaces qui peuvent s'appliquer au "Formulaire pour la création de l'abonnement.

	I	U	N	P&P	R
ENTITIES :					
Client de plus de 16 ans					
Accueil					
DATA STORE :					
Base de données des abonnés					x8
PROCESSES :					
Formulaire pour la création de l'abonnement		x1			
Renouvellement de l'abonnement					
Gestion de la participation aux activités		x2		x7	
Gestion des situations d'urgence	Ce cas n'est pas traité ici car il fait usage de données sensibles				
Statistiques					
Gestion des visiteurs		x3			
DATA FLOWS :					
(3) = nom, prénom, deuxième prénom, date de naissance, nationalité, adresse, téléphone, courriel, photo					
Historique des activités des clients			x6		
Données anonymisées					
Conclusions des statistiques					
Demande de suppression des données		x4			
Formulaire de rectification des données		x5			

Figure 5.2: Tableau reliant le diagramme de l'étape 1 concernant les clients et les types de menaces pour la vie privée

Le tableau est explicité en détails dans la partie qui suit.

5.4.2 Conclusions de l'application de l'étape 2

La méthodologie se révèle intéressante pour examiner les traitements et les flux de données qui existent déjà au moment de l'analyse.
En ce qui concerne les éléments qui devraient être rajoutés au système, ils seront plus facilement identifiés si l'analyste a pris connaissance des threat trees précédemment.

Il serait aussi possible dans de futurs travaux de s'intéresser à une partie plus grande du RGPD et même d'agrandir le nombre de types de menaces existant.

5.5 Etape 3 : Identifier les scénarios de menace

5.5.1 Application de l'étape 3

Cette étape va préciser ce qui a été déterminé précédemment. Elle va aussi faire appel au détail des arbres de menaces qui ont été définis dans le point *"4.5 Etape 3 : Identifier les scénarios de menace"*. Les éléments de ces arbres seront appelés grâce à leurs identifiants pour garder le texte concis.
Reprenons la structure de l'étape 3 proposée par LINDDUN[48] :

"Refine threat via threat pattern"
• Menace : Unawareness
x1 : Les informations qui sont données à l'utilisateur le sont par l'intermédiaire d'un règlement dont l'existence est signalée à l'utilisateur lors de la création de son compte. Les éléments d'information qui se trouvent dans le règlement ne sont pas repris dans le formulaire de création. Il y a donc un risque que l'utilisateur n'ait pas reçu les informations nécessaires lors de la collecte de l'information (U_8)

x2 : L'utilisateur doit être prévenu de la manière dont ses données vont être utilisées. Cela est fait par l'intermédiaire du règlement, cela rejoint ce qui a été dit dans le x1. De plus, un problème de transmission de l'information vis-à-vis de l'utilisateur pourrait aussi être un problème de transparence. (U_6)

x3 : L'utilisateur doit être prévenu de la manière dont ses données vont être utilisées. Cela est fait par l'intermédiaire du règlement, cela rejoint ce qui a été dit dans le x1 et le x2.

x4 : La demande de suppression des données respecte le fait que l'utilisateur doit avoir la possibilité de demander la suppression de ses données. Cependant, comme cette possibilité d'action n'a pas été signalée à l'utilisateur, il est donc possible qu'il n'ait pas conscience de son existence. Cela rejoint ce qui a été dit pour le x1 au sujet de la transmission d'informations.

x5 : Le formulaire permettant à l'utilisateur de rectifier ses données permet de respecter la nécessité que l'utilisateur ait l'occasion de rectifier/corriger ses données. Cependant, comme cette possibilité d'action n'a pas été signalée à l'utilisateur, il est donc possible qu'il n'ait pas conscience de son existence. Cela rejoint ce qui a été dit pour le x1 au sujet de la transmission d'informations.

• Menace : Non-compliance

x6 : NC_5, le responsable ne peut pas prouver que le consentement a été donné. Dans le cas présent, si l'utilisateur pouvait se douter de l'existence d'un historique des activités auquelles il a participé, la demande de consentement ne semble pas avoir été demandée de manière claire. Pourtant la visite du musée par l'utilisateur n'a *a priori* pas d'incidence sur sa visite d'un autre lieu touristique.

• Menace : Processing & Purpose

x7 : TF_4, la demande de consentement n'a pas été mise en évidence séparément des autres questions lors de la récolte des données. L'utilisateur n'a peut-être pas été correctement informé et n'a pas pu choisir quels seront les traitements auxquels il consent.

• Menace : Right of data subject

x8 : Les droits auxquels l'utilisateur peut faire appel ne sont pas tous représentés dans le cas analysé. Il manque notamment : le droit d'accès (D_1, D_2, D_3), le droit d'opposition (D_10, D_11, D_12 , D_14)et le droit à la portabilité des données (D_9) qui pourtant sont tous susceptibles de s'appliquer dans ce scénario.

"Document assumptions"
Les hypothèses posées à ce moment de la réflexion concernent les éléments à modifier pour respecter le RGPD. Ces hypothèses pourraient ensuite être levées, par exemple en consultant le délégué à la protection des données. Dans le scénario présent, une hypothèse pourrait être posée sur la marche à suivre pour prouver que le consentement de l'utilisateur a bien été donné (en référence à x6). Une autre hypothèse pourrait être avancée sur le fait que devoir envoyer un mail au musée pour faire supprimer ses données ne serait pas suffisant pour respecter le droit à l'effacement même si l'utilisateur a conscience de cette option (en référence à x4). Il pourrait alors être décidé qu'il faut un moyen technique pour que ce dernier puisse effacer ses données de manière plus directe, par exemple.

"Document threats using threat template"
Cette étape n'a pas été reprise ici étant donné que les menaces sont la version non-respectée des articles. Il est possible de documenter plus en détails la manière dont le système actuel risque de ne pas respecter l'article en question. Cela peut donner des détails sur la manière de changer le système pour atteindre la conformité attendue.
Comme le scénario concret présent est le résultat de l'amalgame de plusieurs autres scénarios, décrire les détails de chacune des architectures concernées risquerait de s'éloigner du propos tenu ici.

5.5.2 Conclusions de l'application de l'étape 3

La sous-partie du RGPD examinée dans ce mémoire permet de s'intéresser aux bases de ce qui pourrait être une méthodologie utilisable. Pour continuer la réflexion au delà de ce mémoire, il serait intéressant de compléter les arbres de menaces avec plus d'informations provenant du RGPD.

5.6 Etape 4 : Prioriser les menaces

5.6.1 Application de l'étape 4

Cette étape va prioriser les menaces relevées par l'étape précédente.
Comme détaillé dans le point présentant l'étape 4 de manière théorique dans le chapitre précédent, il existe de multiples méthodes de priorisation possibles. Celle qui est adoptée lors d'une analyse est laissée à l'appréciation de l'informaticien se chargeant de l'analyse.

• Eléments x1, x2, x3, x4, x5 et x7 du tableau de l'étape 2 :
Pour le scénario nous occupant ici, il est possible d'utiliser l'une des méthodes de priorisation proposées. On peut faire l'hypothèse que dans un cas comme celui-ci, les menaces auxquelles donner la plus haute priorité soient celles des articles du RGPD qui explicitent ce qui doit être fait sans faire mention de la nécessité d'utiliser des moyens organisationnels ou techniques particuliers. Par exemple, le fait que l'utilisateur doit être informé de son droit à effacer les données. Déterminer si oui ou non l'utilisateur a été informé de cela sera probablement sans équivoque.

• Eléments x6 et x8 du tableau de l'étape 2 :
Les menaces moins prioritaires pourraient être celles qui demandent un plus grand temps de réflexion, celles qui demandent des mesures techniques ou organisationnelles particulières comme par exemple la nécessité de sécuriser un système. Déterminer si des mesures ont été mises en place pour sécuriser le système peut être fait sans équivoque. Déterminer s'il s'agit des mesures "*appropriées*" peut demander plus de discussions.

5.6.2 Conclusions de l'application de l'étape 4

Le choix de la méthode de priorisation va conditionner les actions concrètes exécutées pour mieux tenir compte du RGPD. Comme le temps est une ressource finie et qu'il faut viser une conformité avec la totalité du RGPD, cette étape est cruciale car elle va permettre de décider quelles sont les menaces qui doivent être traitées le plus rapidement.
La méthodologie présentée ici, tout comme LINDDUN, laisse à l'analyste le soin de choisir la méthode de priorisation la plus adaptée à sa situation. Soulignons encore une fois que l'aide d'une personne compétente dans le domaine du droit est ici un atout très précieux car cette personne peut déterminer si les conclusions des étapes précédentes sont correctes et quels sont les points qui doivent être traités en priorité.

5.7 Etape 5 : Identifier les stratégies d'atténuation

5.7.1 Application de l'étape 5

Les stratégies peuvent être de 2 types particuliers :

Dans le cas des menaces ayant reçu la plus grande priorité à l'étape précédente, c'est-à-dire celles qui contiennent la stratégie dans leur formulation, la stratégie à suivre consiste à appliquer l'article (comme par exemple, si l'article dit qu'il faut donner telles et telles informations à l'utilisateur dans tel contexte) et les stratégies d'atténuation découlent directement de ces exigences.
Il s'agit ici de :
• **x1 à x5 :** Pour atteindre une meilleure conformité, il serait avisé de donner les informations demandées par le RGPD dans le cadre de la récolte d'informations auprès de l'utilisateur directement à l'utilisateur. Le détail de ces données a été repris dans les tableaux 3.2 et 3.3.
• **x7 :** La stratégie adoptée pourrait être de présenter les demandes de consentement de manière séparée des autres informations lors de la récolte des données.

Dans le cas des menaces demandant une réflexion plus poussée, les articles peuvent donner un cadre (comme la nécessité de sécuriser des données grâce à des moyens techniques appropriés), il s'agit alors d'une stratégie d'atténuation qui va demander à l'analyste une réflexion technique plus approfondie et même d'utiliser d'autres outils d'analyse, comme par exemple STRIDE qui permettra d'avoir une meilleure idée quant à la sécurité du système analysé.

Il s'agit ici de :
• **x6 :** Prouver que le consentement a bien été donné peut demander de préciser une marche à suivre pour atteindre ce but. Une stratégie pourrait être de garder en mémoire les informations nécessaires à prouver cette demande de consentement.
• **x8 :** La mise en oeuvre de chacun des droits de l'utilisateur demande des mesures spécifiques qui peuvent nécessiter une réflexion technique supplémentaire. Chacun des articles donne des indications sur la direction à suivre et peuvent donc aider à trouver une stratégie adéquate : - le droit d'accès demande que l'utilisateur puisse recevoir confirmation ou infirmation que ses données sont traitées par l'entreprise et qu'il puisse y accéder si c'est le cas.
- le droit d'opposition exige que l'utilisateur soit mis au courant de ce droit et qu'il puisse choisir de s'opposer aux traitements faits sur ses données. Ce droit s'applique ici étant donné que le traitement repose sur le consentement.

- le droit à la portabilité des données donne à l'utilisateur la possibilité de récupérer ses données sous un format tel que csv, json ou xml et puisse les donner à un autre responsable du traitement.
Ces indications permettent de se diriger vers des stratégies cohérentes avec les demandes.

5.7.2 Conclusions de l'application de l'étape 5

L'identification des stratégies d'atténuation dépend des choix effectués précédemment, du contexte dans lequel se trouve le système analysé ainsi que de l'analyste et des choix qu'il va faire.

5.8 Etape 6 : Sélectionner les PETs correspondantes

5.8.1 Application de l'étape 6

Il existe de nombreuses Privacy Enhancing Technologies. Si on se réfère à la définitions théorique de l'étape 6 dans le point 4.8 du chapitre précédent, il est possible d'en trouver quelques unes qui sont adaptées à la situation.
• **x1 à x5 :** Il a été déterminé qu'utiliser des PETs en plus de faire ce que demande l'article demanderait plus de temps.
• **x6 et x7 :** Il existe une PET qui pourrait permettre d'aider à gérer les droits des utilisateurs (en leur permettant notamment d'avoir accès à leurs données personnelles, de les rectifier, de pouvoir choisir comment ils veulent qu'elles soient utilisées, ...). Il s'agit du privacy dashboard créé par l'Université de Vienne.[64] Il pourrait s'avérer très efficace pour gérer une grande partie des modifications qui doivent être faites au système déjà existant et pourrait se greffer sur le système déjà existant comme une interface qui s'adresse au client.

• **x8 :**
- Droit d'accès et Droit d'opposition : l'outil présenté dans les éléments x6 et x7 pourrait être principalement intéressant pour implémenter ces 2 droits particuliers.
- Droit à la portabilité des données : L'utilisation d'un Personal Information Managment system tel que Mydex a été recommandé par une recherche s'intéressant à la mise en place de ce droit particulier.[67]

5.8.2 Conclusions de l'application de l'étape 6

Il existe toute une série de technologies qui peuvent aider à améliorer le respect de la vie privée dans le contexte du RGPD. Le choix de la technologie la plus adéquate est fortement dépendante du contexte technologique de l'organisation faisant l'analyse, ainsi que des fonds disponibles pour se procurer une technologie.
Il est aussi intéressant de se concentrer sur des PETs capables de gérer un maximum d'aspects du RGPD pour réduire le nombre de technologies à utiliser lors de la mise en conformité du système.

5.9 Synthèse des réflexions sur la méthodologie

Ce chapitre a cherché à appliquer la méthodologie de manière à tirer des conclusions sur ce qui peut être amélioré.

La méthodologie a donc réalisé un premier travail important, celui d'étendre LINDDUN à des concepts spécifiques au RGPD en adaptant chacune des étapes, en adaptant ou en créant des types de menaces et les arbres de menaces associés, et en proposant une revue de la littérature sur les *Privacy Enhancing Technologies* en rapport direct avec le RGPD.
La méthodologie représente un point de départ et pose les bases de ce qui pourrait être une méthodologie plus complète.

L'utilisation d'un dataflow diagram est idéale pour observer la manière dont les données se déplacent dans un SI. Cependant, il ne représente pas toujours toutes les mesures organisationnelles qui ont lieu en

dehors du système d'information. Il peut donc être intéressant d'utiliser d'autres outils de modélisation qui peuvent aider à avoir une meilleure vision globale de ce qui se passe dans le système informatique. Déterminer quels sont les outils complémentaires les plus efficaces pourraient être l'objet d'un travail futur.

La méthodologie telle qu'elle est présentée dans ce travail propose un outil d'analyse intéressant, mais elle demande de l'analyste qu'il prenne le temps de se familiariser avec les différents types de menaces et avec les idées principales traitées dans les arbres de menaces. Cela permet que l'analyse puisse tenir compte de tous les éléments possibles. La spécificité de notre méthodologie est le fait qu'elle s'intéresse à un texte de loi qui doit être respecté dans son ensemble. L'analyste doit donc s'être fait une première idée des concepts qui doivent apparaître dans le système sur lequel il travaille. Par exemple, l'utilisateur a des droits que le système doit pouvoir prendre en compte. Cependant, si par exemple, le système ne fait nulle mention de la possibilité pour l'utilisateur d'effacer ses données, l'analyste doit se rendre compte qu'il manque quelque chose. Cela n'est possible que s'il a une connaissance basique des arbres de menaces de la méthodologie.

Il est aussi important de remarquer que si la méthodologie s'adresse à un public d'informaticiens, les résultats obtenus en l'utilisant doivent être présentés et discutés avec les responsables juridiques de la PME ou de l'ASBL afin qu'ils puissent confirmer que les résultats se conforment bien au RGPD. L'application du RGPD fait intervenir plusieurs disciplines (informatique, droit, ...) et est donc susceptible de demander l'intervention de plusieurs personnes provenant de ces différentes disciplines.

La méthodologie telle qu'elle est présentée ici n'est pas encore adaptée à une utilisation en conditions réelles même si elle offre déjà des résultats concluants à son échelle. Plus de tests devraient être effectués pour prouver une efficacité en conditions réelles et dans des contextes aussi différents que ceux des entreprises auxquelles le RGPD s'applique.

Pour qu'elle soit adaptée à une utilisation en conditions réelles, la première étape serait de prendre en compte une plus grande partie du RGPD pour compléter les types de menaces et les arbres de menaces qui interviennent dans l'analyse du système. Cela permettrait aussi d'offrir une meilleur vision quant à la sous-traitance ou quant à la gestion des données sensibles dans les cas autorisés.

Une deuxième étape pourrait travailler sur l'automatisation (partielle) de la méthodologie. Cependant, cette méthodologie laisse une certaine liberté à l'analyste, ce qui le rend indispensable à la réflexion. Ce besoin d'une intervention humaine est aussi dû au fait que le sujet traité est en rapport direct avec un texte de loi. Il est donc intéressant de faire intervenir un juriste dans la réflexion pour s'assurer que les conclusions sont logiques. La méthodologie s'intéresse pour le moment à une vue très haut niveau des différentes entreprises. Chacune des entreprises qui peut être analysée possède son propre contexte et ses propres spécificités. Créer une solution unique convenant à toute entreprise et parfaitement automatisable est donc particulièrement difficile si pas impossible.

Ce mémoire n'a pas pu s'intéresser à toutes les possibles améliorations de la méthodologie mais offre des bases qui, à plus long terme, pourraient éventuellement aboutir à une méthodologie applicable à des cas réels.

Chapitre 6

Conclusion

Le RGPD et les changements qu'il amène sont un sujet complexe qui dépasse les limites de l'informatique elle-même. Ce mémoire présente une tentative de méthodologie visant à aider à appliquer le RGPD dans le cadre des PME et ASBL à une échelle nationale. Ce règlement transforme en langage juridique des choix éthiques qui doivent être implémentés dans des technologies informatiques. N'étant pas juriste de formation, ce mémoire se penche sur les difficultés d'application technologique et non pas sur les interprétations juridiques.

La méthodologie présentée ici s'est basée sur LINDDUN, cette dernière étant utilisée dans le threat modelling qui se concentre sur la vie privée. Le challenge était de réussir à étendre LINDDUN sur la base du RGPD pour mettre en place une méthodologie permettant d'aider un informaticien à identifier certaines Zones de risque de non-conformité et de pouvoir choisir le plan d'action à mettre en place pour y répondre.

Comme ce travail est un mémoire, il a posé les bases de ce qui pourrait, peut-être, faire l'objet d'une recherche plus poussée. L'application à trois cas pratiques en un cas anonyme, n'est bien sûr pas suffisante pour prouver l'efficacité de la méthodologie présentée. Néanmoins, il a permis d'étendre la méthodologie LINDDUN à un sous-ensemble du RGPD en adaptant les types de menaces LINDDUN et en ajoutant deux autres, spécifiques au RGPD. Il serait nécessaire de la mettre à l'épreuve de manière plus exhaustive avant de pouvoir la considérer comme applicable plus généralement.

Il faut aussi noter que le RGPD s'applique dans de nombreux contextes techniques et organisationnels différents. Cette méthodologie a dû rester à un haut niveau qui permet une première approche de la situation. Elle doit prendre en compte les spécificités du contexte de chacun des systèmes concernés. La méthodologie n'a pas pour but de créer une solution unique pertinente pour chacun d'entre eux. L'analyste reste celui qui décide de l'orientation de l'analyse.

Cette méthodologie a également été confrontée au fait que le RGPD n'est pas la seule législation applicable et appliquée aux traitements de données par les PME et ASBL. De plus, les législations peuvent varier selon les pays.

Un dernier point qui peut être rappelé est que cette méthodologie n'est pas abilitée à confirmer un respect du RGPD, mais peut aider à identifier des zones de risque de non-conformité qui devront être traités techniquement.

Plus concrètement, cette méthodologie pourrait aussi se révéler plus rapidement applicable en conditions réelles si elle pouvait être (partiellement) automatisée. Cependant, comme il s'agit d'une méthodologie qui existe dans un cadre juridique particulier, elle nécéssiterait une intervention humaine, même minimale afin de s'assurer que le RGPD et les autres lois nationales sur la vie privée dans le contexte de l'utilisation de l'outil, soient bien respectées.

Afin que la méthodologie présentée puisse être opérationnelle en conditions réelles, il serait impor-

tant de sélectionner des cas concrets provenant de PME et d'ASBL qui travaillent dans des secteurs très différents pour pouvoir apprécier à quel point cette méthodologie est capable de s'adapter à des contextes multiples. En effet, la nature des données et leur traitement peuvent différer de manière importante et donc avoir une répercussion sur la manière dont le RGPD va pouvoir s'appliquer en situations diverses. Il est à noter que comme l'APD l'avait écrit dans son vade-mecum, créer une solution unique pour l'application du RGPD à toutes les PME s'avèrerait extrêmement difficile, si pas impossible.[44] Néanmoins, la méthodologie pourrait peut-être dans le futur être capable de pouvoir s'appliquer à un plus grand nombre de types de PME et d'ASBL.

En conditions réelles, les résultats obtenus devraient être présentés et discutés avec les responsables juridiques de la PME ou de l'ASBL afin qu'ils puissent confirmer qu'ils se conforment bien au RGPD.

La définition des limites technologiques doit également être prise en considération dans les choix de politique de l'entreprise face au RGPD. Il est de la responsabilité de l'informaticien d'avoir et de transmettre au DPO une information claire sur ces limites technologiques, même si le choix de l'entreprise sur sa politique d'application du RGPD dépasse sa seule responsabilité.

Dans une perspective plus générale, l'arrivée du RGPD demande parfois aux informaticiens de réinterroger toute l'architecture d'un système préexistant et conditionne la création d'un nouveau système. En ce sens, il réoriente les architectures informatiques. Le RGPD nous rappelle ainsi que la technologie n'est pas neutre.

Bibliographie

[1] A. De Preter, "Strategy & Architecture Framework", http://www.labnaf.one/guidance/index.html, consulté le 7 feb. 2019.

[2] LINDDUN, "LINDDUN in a nutshell", https://linddun.org/linddun.php, 5 dec. 2018, consulté le 4 feb. 2019.

[3] "Protection des données personnelles: plus de 95.000 plaintes déposées dans l'Union européenne en huit mois," *Le Soir*.

[4] CNIL, "RGPD : quel bilan 6 mois après son entrée en application ?", https://www.cnil.fr/fr/rgpd-quel-bilan-6-mois-apres-son-entree-en-application, consulté le 12 fév. 2019.

[5] F. Pasquale, "Mettre fin au trafic des données personnelles," *Le Monde diplomatique*, May 2018.

[6] S. Zuboff, "Un capitalisme de surveillance," *Le Monde diplomatique*, Jan. 2019.

[7] J. Thorel, "Privés de vie privée," *Le Monde diplomatique*, Jan. 2015.

[8] "RGPD : présentation et conséquences pour les entreprises", https://www.act.com/fr-fr/old/act-blog-old/blog/2017/09/27/rgpd-presentation-et-consequences-pour-les-entreprises, 31 dec. 2018.

[9] W. Ashford, "GDPR will require 28,000 DPOs in Europe and US, study shows", https://www.computerweekly.com/news/450283253/GDPR-will-require-28000-DPOs-in-Europe-study-shows, 20 avr. 2016, consulté le 3 dec. 2018.

[10] Commission Européenne, "À quoi correspondent les données à caractère personnel?", https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_fr, consulté le 31 dec. 2018.

[11] Comission Européenne, "Que régit le règlement général sur la protection des données (GPDR)?", https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-does-general-data-protection-regulation-gdpr-govern_fr, consulté le 31 dec. 2018.

[12] Comission Européenne, "Sept étapes pour aider les entreprises à se préparer à l'entrée en vigueur du règlement général sur la protection des données", https://ec.europa.eu/commission/sites/beta-political/files/ds-02-18-544-fr-n.pdf, consulté le 29 dec. 2018.

[13] APD, "Pourquoi ne pouvons-nous pas encore répondre à toutes les questions ?", https://www.autoriteprotectiondonnees.be/pourquoi-ne-pouvons-nous-pas-encore-repondre-a-toutes-les-questions, consulté le 28 jan. 2019.

[14] D. B. E. Søren Debois and T. H. (KU), "On Purpose and by Necessity: Compliance under the GDPR," *IT University of Copenhagen*, p. 20, Feb.

[15] K. Winter and S. Rinderle-Ma, "Untangling the GDPR Using ConRelMiner," *arXiv:1811.03399 [cs]*, Nov. 2018. arXiv: 1811.03399.

[16] CISCO, "Next-Generation Firewalls (NGFW)", https://www.cisco.com/c/en/us/products/security/firewalls/index.html, consulté le 7 jan. 2019.

[17] Juniper, "Next-Generation Firewalls - Juniper Networks", https://www.juniper.net/us/en/solutions/security/next-gen-firewall, consulté le 7 jan. 2019.

[18] Palo Alto, "Next-Generation Firewalls Palo Alto Networks", https://www.paloaltonetworks.com/products/secure-the-network/next-generation-firewall, consulté le 7 jan. 2019.

[19] C. Brook, "What is a Next Generation Firewall ?", https://digitalguardian.com/blog/what-next-generation-firewall-learn-about-differences-between-ngfw-and-traditional-firewalls, 5 dec. 2018, consulté le 9 jan. 2019.

[20] SNORT, "Snort - Network Intrusion Detection & Prevention System", https://www.snort.org/, consulté le 7 jan. 2019.

[21] AlgoSec, "AlgoSec Delivers Complete End-to-End Security Management for Cloud Security Controls", https://www.algosec.com/press_release/algosec-delivers-complete-end-to-end-security-management-for-cloud-security-controls/, consulté le 26 jan. 2019.

[22] Juniper, "GDPR is here, where are you ?", https://www.juniperemea.net/gdpr/, consulté le 7 jan. 2019.

[23] P. Ferrara and F. Spoto, "Static analysis for gdpr compliance," in *ITASEC*, 2018.

[24] O. Tripp, M. Pistoia, S. J. Fink, M. Sridharan, and O. Weisman, "Taj: effective taint analysis of web applications," in *PLDI*, 2009.

[25] S. Horwitz, T. Reps, and D. Binkley, "Interprocedural Slicing Using Dependence Graphs," vol. 12, no. 1, p. 35, 1990.

[26] ISACA, "Prévention contre la fuite de données," *Livre blanc de l'ISACA*, 2010.

[27] Open Group, "ArchiMate® 3.0.1 Specification", http://pubs.opengroup.org/architecture/archimate3-doc/chap03.html, consulté le 1 jan. 2019.

[28] Archi, "Archi – Open Source ArchiMate Modelling", https://www.archimatetool.com/, consulté le 1 jan. 2019.

[29] Archi, "Resources – Archi", https://www.archimatetool.com/resources/, consulté le 26 jan. 2019.

[30] W3, "ODRL Information Model 2.2", https://www.w3.org/TR/odrl-model/, consulté le 29 avr. 2019.

[31] H. J. Pandit, D. O'Sullivan, and D. Lewis, "GDPR-driven Change Detection in Consent and Activity Metadata," p. 5.

[32] M. Pankowska, "Research Worker Tasks Modeling for Hospital Processes' Accreditation," p. 12.

[33] C. Palmér, "Modelling EU DIRECTIVE 2016/680 using Enterprise Architecture," p. 58.

[34] Obus Software, "GDPR Software", https://www.orbussoftware.com/governance-risk-and-compliance/gdpr/, consulté le 28 jan. 2019.

[35] A. De Preter, "Labnaf Strategy and Architecture Framework – A Language Built in its Native Framework", http://www.labnaf.one/, consulté le 6 feb. 2019.

[36] A. De Preter, "Labnaf – All-in-one Strategy & Architecture Framework," p. 16.

[37] A. De Preter, "The Labnaf Architecture Framework", http://www.labnaf.one/ln-content/events/Labnaf%20at%20Enterprise%20Architect%20User%20Group.pdf, consulté le 7 feb. 2019.

[38] M. Robol, M. Salnitri, and P. Giorgini, "Toward gdpr-compliant socio-technical systems: Modeling language and reasoning framework," in *PoEM*, 2017.

[39] S. Hernan, S. Lambert, T. Ostwald et A. Shostack, "Uncover Security Design Flaws Using The STRIDE Approach", https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxzZWN1cmVwcm9ncmFtbWluZ3xneDo0MTY1MmM0ZDI0ZjQ4ZDMy, nov. 2006, consulté le 29 avr. 2019.

[40] Chun Yu Cheung, "Threat Modeling Techniques", Delft University of Technology, 2016.

[41] K. Wuyts, "Linddun: a privacy threat analysis framework,"

[42] LINDDUN, "LINDDUN Mitigation Strategies & Solutions", https://linddun.org/solutions.php, consulté le 2 avr. 2019.

[43] Y. S. Martín and J. M. del Álamo, "A metamodel for privacy engineering methods," in *IWPE@SP*, 2017.

[44] APD, "RGPD : Vade-mecum pour les pme", https://www.autoriteprotectiondonnees.be/sites/privacycommission/files/documents/PME_FR_0.pdf, consulté le 5 fév. 2019.

[45] "La personnalité juridique et les deux catégories de personne", http://www.maxicours.com/soutien-scolaire/droit/1re-stg/152194.html, consulté le 2 mar. 2019.

[46] Commission Européenne, "Règlement Général sur la Protection des Données", https://eur-lex.europa.eu/eli/reg/2016/679/oj, consulté le 21 mar. 2018.

[47] Avocats Mathias, "Droit à la portabilité : conseils pratiques", https://www.avocats-mathias.com/donnees-personnelles/droit-a-la-portabilite, mar. 2018, consulté le 12 mai 2019.

[48] Kim Wuyts and Wouter Joosen, "LINDDUN privacy threat modeling: a tutorial", Technical Report (CW Reports), volume CW685, Department of Computer Science, KU Leuven, July 2015.

[49] LINDDUN, "LINDDUN Privacy threats catalog", https://distrinet.cs.kuleuven.be/software/linddun/catalog.php, consulté le 22 fév. 2019.

[50] LINDDUN, "LINDDUN catalog - Non-compliance", https://linddun.org/noncompliance_tree.php, consulté le 25 fév. 2019.

[51] LINDDUN, "LINDDUN catalog - Identifiability of entity", https://linddun.org/identifiability_E.php, consulté le 18 mars. 2019.

[52] LINDDUN, "LINDDUN catalog - Identifiability of process", https://linddun.org/identifiability_P.php, consulté le 12 mars. 2019.

[53] LINDDUN, "LINDDUN catalog - Identifiability of dataflow", https://linddun.org/identifiability_DF.php, consulté le 18 mars. 2019.

[54] LINDDUN, "LINDDUN catalog - Identifiability of datastore", https://linddun.org/identifiability_DS.php, consulté le 18 mars. 2019.

[55] LINDDUN, "LINDDUN catalog - Unawareness of entity", https://linddun.org/contentunawareness_E.php, consulté le 12 mar. 2019.

[56] APD, "Analyse d'impact relative à la protection des données - FAQ", https://www.autoriteprotectiondonnees.be/faq-themas/analyse-dimpact-relative-%C3%A0-la-protection-des-donn%C3%A9es#t10180n20726, consulté le 3 nov. 2019.

[57] CNIL, "Ce qu'il faut savoir sur l'analyse d'impact relative à la protection des données (AIPD)", https://www.cnil.fr/fr/ce-quil-faut-savoir-sur-lanalyse-dimpact-relative-la-protection-des-donnees-aipd, 11 juin 2018, consulté le 3 nov. 2019.

[58] APD, "Liste des types d'opérations de traitement pour lesquelles une AIPD est requise", https://www.autoriteprotectiondonnees.be/sites/privacycommission/files/documents/Liste_des_traitements_AIPD.pdf, consulté le 3 mar. 2019.

[59] APD, "Analyse d'impact relative à la protection des données", https://www.autoriteprotectiondonnees.be/analyse-dimpact-relative-a-la-protection-des-donnees, consulté le 3 nov. 2019.

[60] CNIL, "Outil PIA : téléchargez et installez le logiciel de la CNIL", https://www.cnil.fr/fr/outil-pia-telechargez-et-installez-le-logiciel-de-la-cnil, 12 juin 2018, consulté le 3 nov. 2019.

[61] E. D. Cristofaro and M. K. Wright, "Privacy enhancing technologies," in *Lecture Notes in Computer Science*, 2013.

[62] M. Mourby, E. Mackey, M. Elliot, H. Gowans, S. Wallace, J. Bell, H. Smith, S. Aidinlis, and J. Kaye, "Are 'pseudonymised' data always personal data? implications of the gdpr for administrative data research in the uk," 2018.

[63] GDPR Informer, "PET's and the GDPR : adoption time ?", https://gdprinformer.com/gdpr-articles/pets-gdpr-adoption-time, consulté le 17 avr. 2019.

[64] P. Raschke, A. Küpper, O. Drozd, and S. Kirrane, "Designing a gdpr-compliant and usable privacy dashboard," in *Privacy and Identity Management*, 2017.

[65] E. A. Politou, E. Alepis, Constantinos, and C. Patsakis, "Forgetting personal data and revoking consent under the gdpr: Challenges and proposed solutions," *J. Cybersecurity*, vol. 4, p. tyy001, 2018.

[66] GDPR.be, "PET's and GDPR", https://gdpr.be/uncategorized/summary-of-privacy-enhancing-technologies-a-survey-of-tools-and-techniques/, consulté le 17 avr. 2019.

[67] L. Urquhart, N. V. Sailaja, and D. McAuley, "Realising the right to data portability for the domestic internet of things," *Personal and Ubiquitous Computing*, vol. 22, pp. 317–332, 2017.