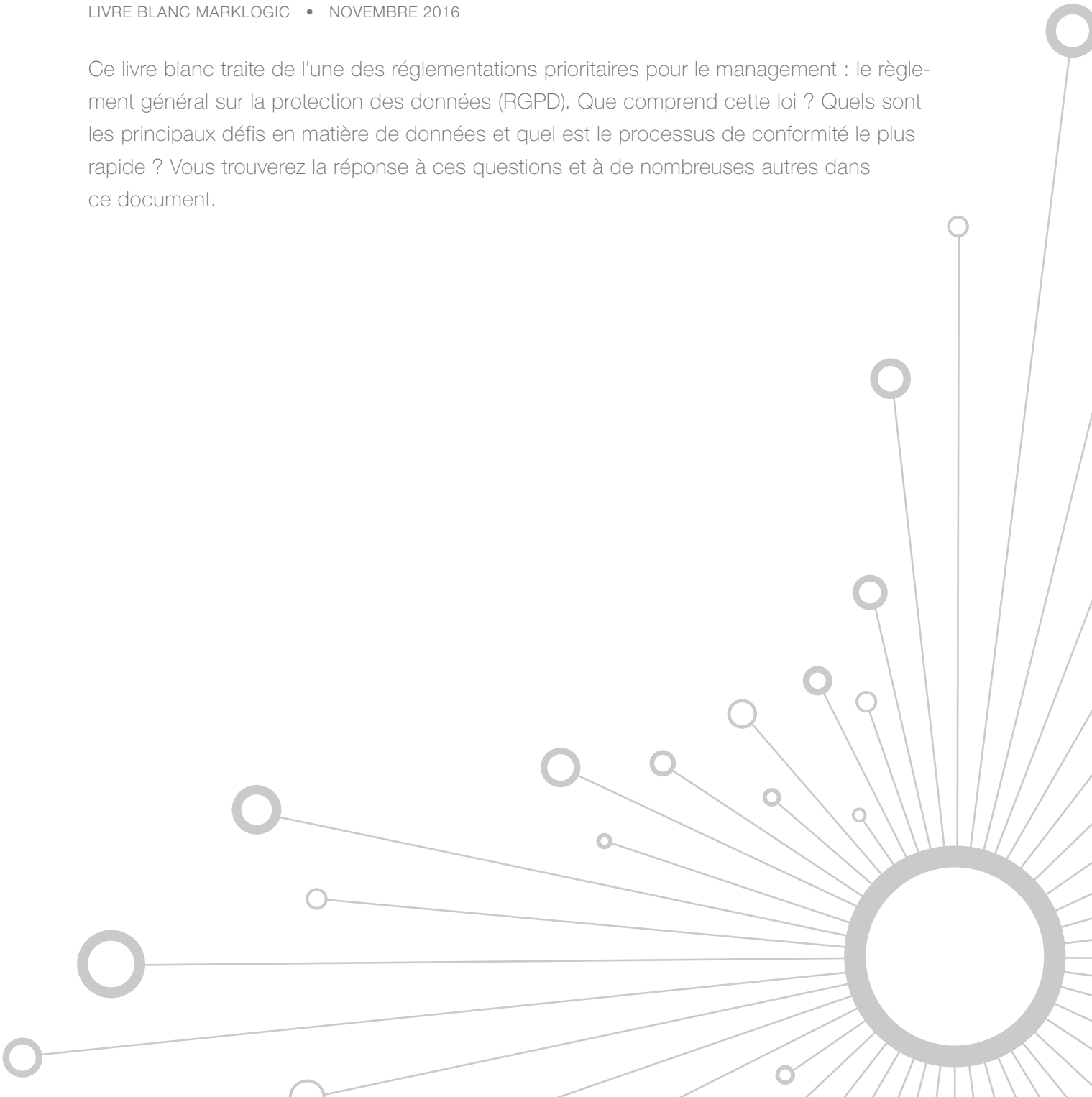
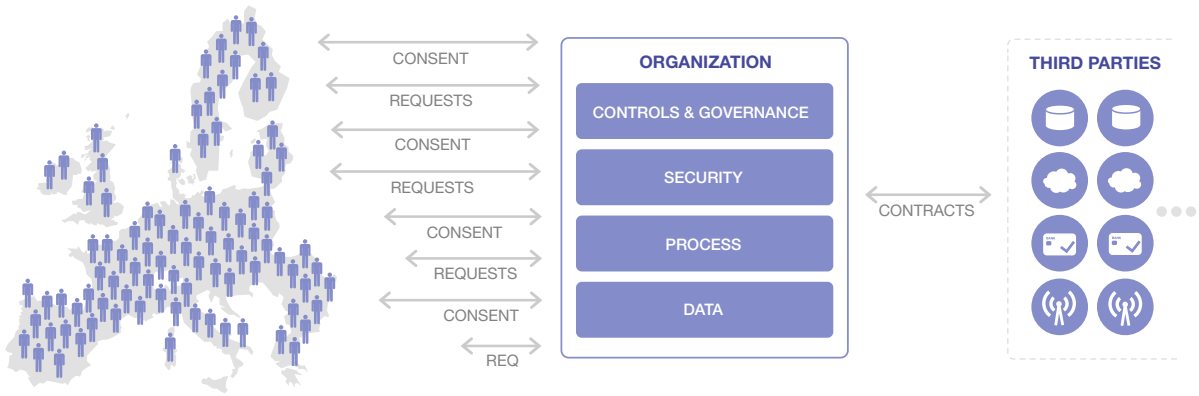
MarkLogic

RÈGLEMENT GÉNÉRAL SUR LA PROTECTION DES DONNÉES - PROCESSUS DE CONFORMITÉ

LIVRE BLANC MARKLOGIC • NOVEMBRE 2016

Ce livre blanc traite de l'une des réglementations prioritaires pour le management : le règlement général sur la protection des données (RGPD). Que comprend cette loi ? Quels sont les principaux défis en matière de données et quel est le processus de conformité le plus rapide ? Vous trouverez la réponse à ces questions et à de nombreuses autres dans ce document.





RÉSUMÉ OPÉRATIONNEL

Le paysage réglementaire se complexifie avec l'introduction de nouveaux règlements dans diverses juridictions. Pour les entreprises détenant des données personnelles de citoyens européens, une des lois les plus importantes est le règlement général sur la protection des données (RGPD) qui entrera en vigueur au mois de mai 2018.

Dans ce livre blanc, nous traitons des principaux défis de cette nouvelle réglementation et fournissons des conseils pratiques pour les relever rapidement et efficacement, avec notamment les points suivants :

- Qu'est-ce que le règlement général sur la protection des données (RGPD) et quel impact a-t-il sur votre entreprise ?
- Liste concernant la préparation des données
- Défis et risques du processus de conformité
- Guide de conformité par étapes : les conseils de MarkLogic
- Comment optimiser les nouvelles technologies au sein de votre processus de gestion des données pour profiter de nouvelles opportunités

QU'EST-CE QUE LE RÈGLEMENT GÉNÉRAL SUR LA PROTECTION DES DONNÉES (RGPD) ET QUEL IMPACT A-T-IL SUR VOTRE ENTREPRISE ?

Le règlement général sur la protection des données (RGPD) définit les droits des citoyens européens (ou Personnes concernées) concernant la confidentialité et la protection de leurs données personnelles. Le règlement signifie également les obligations des entreprises et particuliers chargés de déterminer les finalités du stockage et du traitement de ces données (responsables du traitement ou contrôleurs de données). Les contrôleurs de données seront responsables de toute violation de cette réglementation et les amendes potentielles leur seront adressées.

Toute entreprise responsable des données d'un employé, client, patient, citoyen, etc. devra se conformer à cette réglementation. Le règlement général sur la protection des données (RGPD) entrera en vigueur au mois de mai 2018. *La non-conformité à cette réglementation pourra engendrer des amendes pouvant atteindre jusqu'à 4 % des recettes globales.*

L'impact du règlement est très vaste et la conformité s'avèrera complexe, tout particulièrement pour les grandes entreprises. La réduction des risques liés à ce règlement nécessite une planification proactive ainsi que la capacité à garantir la conformité sur le long terme.

Par exemple, il est important de mettre en place une structure au sein de laquelle l'entreprise pourra répondre aux requêtes des individus (personnes concernées). Ces requêtes peuvent inclure :

- Une déclaration écrite exigeant que vous n'utilisiez pas leurs données personnelles pour prendre des décisions automatisées, interrogeant les raisons d'une décision automatisée ou vous demandant de revoir les résultats d'une décision automatisée
- Une requête concernant les modèles de traitement des données et la communication éventuelle de leurs données à d'autres entreprises ou personnes

LISTE CONCERNANT LA PRÉPARATION DES DONNÉES

Les 9 questions auxquelles vous devez répondre pour garantir la préparation de vos données au règlement général sur la protection des données :

- Savez-vous quelles données de votre entreprise contiennent des informations personnelles, notamment de sources non structurées (notamment des documents) ?
- Savez-vous comment vous utilisez ou traitez les données personnelles et à quelles fins ?
- Savez-vous quel accord a été donné (ou retiré ultérieurement), et à quelles fins, pour ces données personnelles ?
- Savez-vous quels sont les contrôles de sécurité en place pour les données personnelles et si ces données sont supprimées ou anonymisées de manière appropriée lorsque le consentement est annulé ?
- Êtes-vous en mesure de répondre aux requêtes réglementaires concernant l'utilisation et le stockage de données personnelles dans les délais requis par le règlement européen ?
- Pouvez-vous adapter votre capacité à répondre à des milliers de requêtes simultanées potentielles (dans le cadre d'une poursuite judiciaire, par exemple) ?
- Stockez-vous des données personnelles dans des langues multiples ?
- Traitez-vous les données personnelles comme un atout précieux afin d'améliorer l'expérience des clients et de développer potentiellement de nouveaux produits ?

- La demande d'une copie de leurs données personnelles
- Le retrait de consentement pour l'utilisation de leurs données personnelles dans des situations spécifiques (avec certaines restrictions) ou à des fins de marketing direct
- La demande de correction de données personnelles incorrectes

Chacune de ces requêtes doit être traitée dans un certain délai. Par exemple, l'entreprise dispose de 28 jours pour répondre à une requête de retrait de consentement pour une utilisation des données à des fins de marketing direct.

Mais la situation se complexifie encore davantage.

Le règlement est en soi finalisé mais chaque état européen doit l'interpréter en fonction de sa propre législation et décider de recommandations et de mesures d'application. Au Royaume-Uni, par exemple, l'autorité responsable est l'ICO (Bureau du Commissaire à l'information). L'ICO a déclaré qu'il faudra de nombreux mois avant d'appliquer cette recommandation. Jusque là, aucune entreprise ne connaîtra exactement les étapes à suivre pour assurer sa conformité au règlement.

DÉFIS LIÉS AU PROCESSUS DE CONFORMITÉ

La réglementation est constituée de deux mandats majeurs devant être respectés par chaque entreprise :

- Identifier et stocker efficacement les données personnelles conformément aux contrôles de sécurité dans leurs systèmes internes
- Répondre immédiatement aux requêtes des citoyens européens concernant l'utilisation de leurs données personnelles et de leur suppression des systèmes internes

Afin de se conformer aux impératifs cités plus haut, les entreprises doivent identifier les données qu'elles stockent ou traitent, à quelles fins, selon quel type de consentement ainsi que les mesures de sécurité et contrôles en place. De plus, les entreprises doivent savoir comment ces données sont liées les unes aux autres relativement aux personnes concernées afin de répondre aux requêtes individuelles.

Plus la taille de l'entreprise est importante, plus les systèmes traitant les données sont dispersés et plus il est difficile de rassembler ces informations et de répondre aux requêtes.

UTILISATION ET STOCKAGE DES DONNÉES PERSONNELLES

Selon le règlement, les données personnelles doivent uniquement être stockées ou traitées conformément à l'utilisation et aux fins pour lesquelles la personne concernée a donné son accord. Par exemple, l'archivage de ces données à des fins non réglementaires ou légales peut ne pas être autorisé. Citons un autre exemple lié à l'analyse générale : les données contenant des données personnelles peuvent devoir être rendues anonymes avant leur utilisation. En cas de fuite des données, s'il s'avère que des données personnelles étaient stockées ou traitées de manière non-conforme, les entreprises responsables pourront faire l'objet d'une poursuite judiciaire.

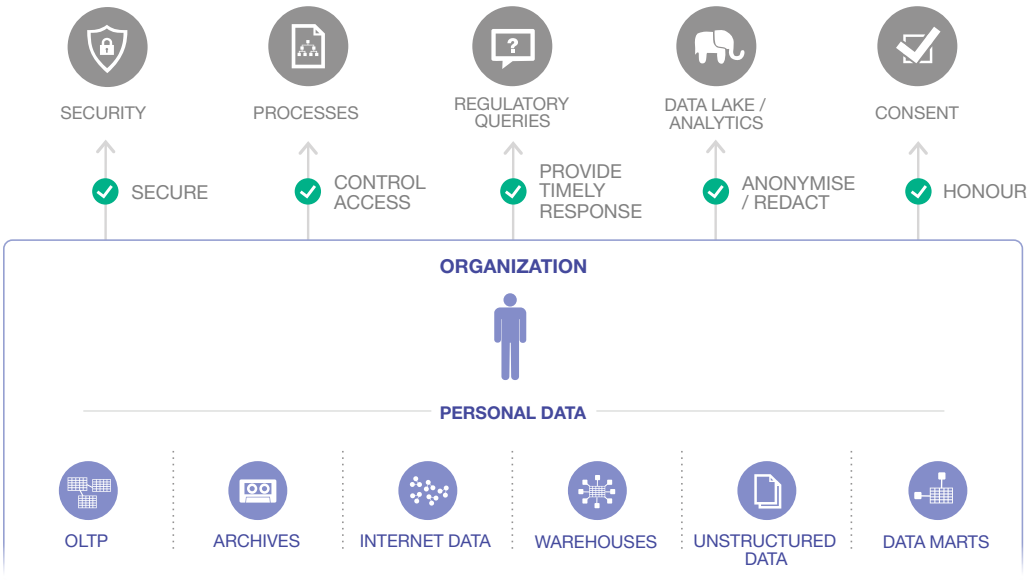
Il est par conséquent nécessaire pour les entreprises d'identifier les données qu'elles stockent et traitent, de prendre des décisions sur leur existence et utilisation à l'avenir et de s'adapter en fonction des changements futurs des réglementations. Cependant, à elle seule, cette simple étape représente un immense volume de traitement des données.

DES RÉPONSES RAPIDES AUX REQUÊTES DES CITOYENS EUROPÉENS CONCERNANT LEURS DONNÉES PERSONNELLES

Afin de répondre aux requêtes concernant l'utilisation des données personnelles ou l'obtention d'un consentement pour une utilisation spécifique de ces données, les entreprises doivent d'abord identifier les données personnelles pertinentes pour la requête. Pour ce faire, ils doivent savoir à quelle personne concernée (citoyen, client, employé, etc.) correspond quelles données. Par exemple, un client peut annuler un contrat avec une entreprise tout en gardant d'autres contrats actifs. Les entreprises doivent comprendre quels éléments de données correspondent à quel client, ainsi que le type de contrat spécifique et les processus associés. Une fois que toutes les données pertinentes ont été identifiées, l'entreprise peut démarrer le processus de retrait ultra-précis de cet ensemble de données.

Les entreprises qui décident de ne pas organiser leurs données selon ce niveau de détail sont confrontées aux risques suivants :

- Les efforts impliqués à la réponse aux requêtes risquent de compliquer, ou de rendre impossible, leur traitement dans les délais requis. En effet, même lorsqu'une entreprise peut identifier les systèmes contenant les données personnelles (et à quelles fins commerciales ces systèmes sont utilisés), les efforts de recherche au sein de chacun de ces systèmes et d'identification d'un individu spécifique peuvent se révéler coûteux.
- Les réponses à ces requêtes ponctuelles se révèlent la plupart du temps urgentes et coûteuses. Elles exercent par conséquent une pression sur les coûts à long terme, bien plus importante que les efforts en amont pour organiser les informations.



> “Le règlement est finalisé mais chaque état européen doit l'interpréter en fonction de sa propre législation et décider des recommandations et mesures d'application.”

- Parce qu'il faut faire face à la coordination de larges groupes d'individus effectuant des requêtes simultanément, même si une requête unique peut être traitée de manière adéquate, le traitement d'une douzaine (ou d'une centaine) de requêtes simultanées devient impossible. De plus, les groupes organisés sont également mieux représentés et peuvent recourir à des poursuites judiciaires collectives.

L'organisation des données personnelles au niveau des personnes concernées peut être effectuée manuellement ou à l'aide d'un support automatisé. Les approches manuelles nécessitent par définition des ressources humaines importantes et des efforts accrus qui se révèlent (même pour les volumes de données les plus modestes) extrêmement coûteux. En outre, quelles que soient les ressources manuelles générées, les résultats statiques produits nécessiteront des efforts continus s'ils veulent être maintenus.

Les entreprises peuvent également faire appel à l'aide de larges fournisseurs d'outils (semi) automatisés. Ces outils, cependant, sont limités sur deux plans en ce qui concerne l'application de la conformité du règlement général sur la protection des données.

- Ils sont adaptés aux données structurées (par exemple, les bases de données relationnelles) ou aux données non structurées (par exemple, les documents) mais pas aux deux types de données. Or, les entreprises disposent de données personnelles dans l'ensemble de leur centre de données et ont besoin d'une solution qui puisse gérer de multiples types de données.

- Ils n'auront pas la capacité d'annoter ces sources de données à l'aide des métadonnées riches requises pour comprendre et analyser les risques et pour répondre aux requêtes réglementaires rapidement.

GUIDE DE PROCESSUS DE CONFORMITÉ PAR ÉTAPES

ÉTAPE 1. IDENTIFICATION DES DONNÉES PERSONNELLES

La première étape consiste à comprendre le type de données personnelles détenues par votre entreprise. Les données peuvent être stockées dans des systèmes de bases de données, des applications (sur site et dans le cloud) ou des documents (Word, PDF, etc.) Quant aux applications, il peut s'agir de systèmes RH, de sites Web (stockant des données de cookies ou autres identifiants en ligne), des systèmes CRM, etc.

Les entreprises doivent tout d'abord identifier tous les systèmes existants (qu'ils soient détenus et gérés en interne ou basés dans le cloud mais pour lesquels l'entreprise est responsable du stockage de données) contenant des données personnelles.

Une fois que les systèmes ont été identifiés, les entreprises peuvent inspecter les données stockées dans ces systèmes et appliquer les règles correspondantes (filtrage par motif, identification de numéros de téléphone, traitement automatique du langage naturel, compréhension de l'utilisation complexe du langage dans différents documents) afin d'identifier toute donnée personnelle. Une fois que les données personnelles ont été identifiées, leur nature, contexte global et provenance doivent être enregistrés dans un système permettant à ces informations d'être facilement récupérées et interrogées.

ÉTAPE 2. DOCUMENTATION DE L'UTILISATION ET DE LA FINALITÉ DES DONNÉES PERSONNELLES

Lorsqu'il a été identifié qu'une source de données contient des données personnelles, la prochaine étape consiste à documenter l'usage et la finalité de ces données.

> “N'oublions pas que le règlement général sur la protection des données évolue toujours en ce qui concerne la compréhension du type de connaissances requises pour les données personnelles. Les entreprises doivent pouvoir enregistrer ces métadonnées personnelles de manière flexible afin de pouvoir ajouter un nouvel attribut plus tard.”

Pour certains ensembles de données, les connaissances concernant, par exemple, le type de processus commercial consommant cet ensemble de données ou le type de politique de rétention appliquée, existent déjà dans d'autres systèmes (système de gestion des processus commerciaux ou système de gestion des ressources, etc.). Ces données peuvent être importées dans une technologie de base de données afin de fournir une vision consolidée unique des données personnelles, annotées avec les informations associées (métadonnées). Interroger cette base de connaissances combinées permet de réduire la complexité et les coûts par rapport aux autres approches (tentative de fédération des interrogations au sein de multiples systèmes différents).

Pour les autres ensembles de données, et pour certains détails, ces connaissances ne sont pas toujours connues ou actuellement disponibles. Si c'est le cas, les entreprises devront faire appel à un processus manuel pour récupérer ces connaissances. L'automatisation de ce processus nécessitera une nouvelle solution technologique dotée de capacités de découverte de données permettant la catégorisation des données pour identifier certains de ces détails.

N'oublions pas que le règlement général sur la protection des données (RGPD) évolue toujours en ce qui concerne la compréhension du type de connaissances requises pour les données personnelles. Les entreprises doivent pouvoir enregistrer ces métadonnées personnelles de manière flexible afin de pouvoir ajouter un nouvel attribut plus tard.

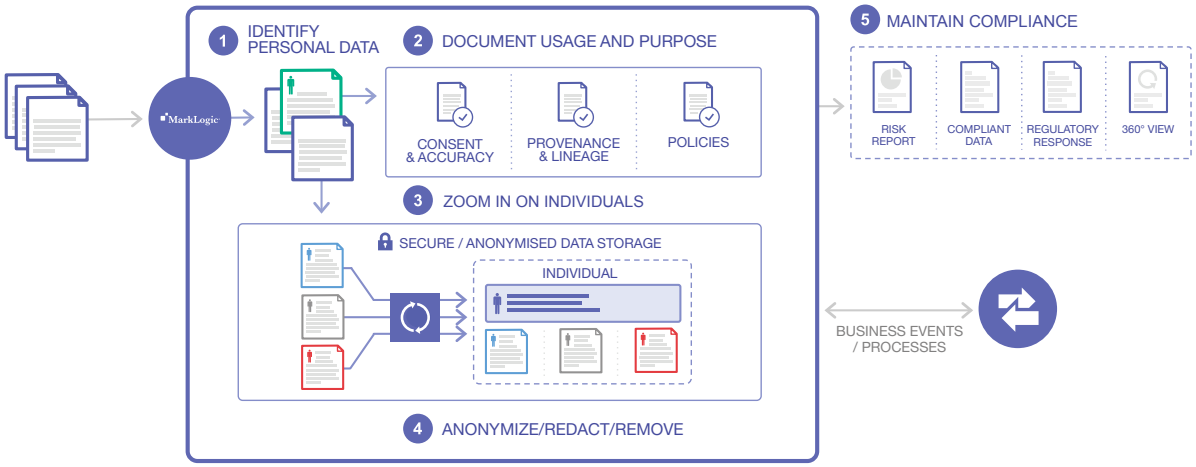
ÉTAPE 3. ZOOM SUR LES INDIVIDUS

Les étapes précédentes illustrent la manière dont les données personnelles peuvent être identifiées et annotées avec des informations concernant leur utilisation et finalité. Cependant, une troisième étape est également très utile pour pouvoir répondre à une requête réglementaire concernant un individu : l'identification et le rassemblement de données concernant cet individu au sein de multiples systèmes. Par exemple, les données d'un client peuvent être enregistrées dans une application de commerce électronique, un système CRM et un système marketing. Et, si une entreprise a fait l'objet d'une fusion ou d'une acquisition, ces données peuvent exister en double dans les systèmes d'origine des deux entreprises.

Les règles de correspondance peuvent être appliquées afin de lier des ensembles de données multiples pour former des entités. Le couplage de données est rarement un exercice précis car, par exemple, des adresses peuvent avoir changé, si bien qu'il est difficile de vérifier que deux organisations distinctes font en fait bien référence à une seule et même entité. Il peut donc être utile d'enregistrer le degré d'exactitude de cette association. De plus, la découverte de cette incohérence au niveau des adresses (par exemple) peut avoir de la valeur pour une entreprise, par exemple pour aider à identifier des données inexactes.

Une fois que les entités ont été identifiées au sein des ensembles de données, les recherches permettant de répondre aux requêtes réglementaires peuvent être précisées pour correspondre à un individu et fournir les résultats les plus précis possibles. Sans ces informations supplémentaires, une interrogation peut uniquement être effectuée au niveau du type de données dans chaque ensemble de données.

Dernier point de cette étape : ce regroupement des ensembles de données en entités peut être effectué de manière répétitive. L'annotation des données doit certes être flexible pour prendre en compte les



exigences futures mais le lien des ensembles de données doit quant à lui être effectué et approfondi afin de respecter les exigences réglementaires rapidement et de manière économique. Par exemple, il peut s'avérer rapide et facile d'identifier 80 % des clients par leur nom, leur adresse ou un autre identifiant. Pour les 20 % restants, l'exercice peut être retardé jusqu'à ce qu'il soit possible de réconcilier les ensembles de données (par exemple, en réponse à une requête réglementaire concernant cet individu).

ÉTAPE 4. ANONYMISATION ET SUPPRESSION DES DONNÉES PERSONNELLES

L'entreprise comprend désormais le lien entre les données personnelles et les personnes concernées, la façon dont les données personnelles sont utilisées ainsi que le type de consentement accordé. Elle doit maintenant interroger ces connaissances pour identifier si les données personnelles sont utilisées ou stockées sans le consentement nécessaire.

Les entreprises peuvent alors choisir d'anonymiser (ou de censurer) les données et/ou de les supprimer au cas où elles n'en n'aient pas l'utilisation légitime. Finalement, dans le cas où les données font l'objet d'un consentement partiel (par exemple, pour un usage mais pas pour un autre), les entreprises peuvent appliquer une censure/anonymisation dynamique ou le fitrage, en fonction de la personne qui requiert les données.

ÉTAPE 5. MAINTIEN DE LA CONFORMITÉ

La dernière étape consiste à interroger les données personnelles afin de répondre à une requête. Des capacités de recherche et d'interrogation riches sont requises pour tous les ensembles de connaissances des données personnelles de l'entreprise afin de rassembler et de présenter les données nécessaires pour répondre à la requête. Les questions peuvent être posées sur les données elles-mêmes, les détails annotés sur ces données ou les deux.

Par exemple, un particulier peut se plaindre parce qu'il reçoit des communiqués de marketing direct. Il est possible que cette personne ait initialement consenti à ces communications marketing mais uniquement pendant une durée limitée (pendant une certaine période ou durant une offre spécifique). L'entreprise devra pouvoir poser des questions sur des détails concernant cette personne (en utilisant uniquement une adresse e-mail par exemple), en ciblant un ensemble de données marketing spécifique et en extrayant le consentement donné, pour quelle campagne ou durée. Une fois que le consentement approprié aura été identifié, l'entreprise devra mettre à jour le système pour indiquer que l'utilisateur a supprimé tacitement son accord. Les processus appropriés devront alors être lancés afin que les systèmes soient mis à jour pour éviter l'envoi de toute nouvelle communication de marketing direct à cette personne.

En outre, dû au caractère évolutif du règlement général sur la protection des données, la solution doit être flexible en ce qui concerne les interrogations et questions qu'elle peut gérer.

Il est important de noter que la possibilité d'accéder à ce niveau d'informations sur vos clients, employés, patients, etc., peut être utilisée à d'autres fins que les seules requêtes réglementaires. Elle permet en effet de mieux connaître ces personnes et de transformer le système en véritable plate-forme de valeur ajoutée.

LES EXIGENCES POUR LA SOLUTION

Les principales exigences pour qu'une solution technologique puisse prendre en charge les étapes mentionnées ci-dessus et assurer la conformité à cette nouvelle réglementation sont les suivantes :

La capacité à extraire des données des systèmes et à les analyser afin d'identifier des données personnelles (que ces données soient hautement structurées ou non structurées, avec peut-être la prise en charge de multiples langues)

- La possibilité d'enregistrer la nature des données personnelles et de les annoter avec des détails riches pouvant changer au fil du temps.
- La possibilité de faire correspondre et de lier les données personnelles dans des entités de personnes concernées
- La possibilité de réagir aux changements de données dans l'ensemble du système afin que cette base de connaissance soit constamment actualisée
- Les capacités :
    - D'assurer la coordination entre les systèmes source afin de supprimer les données ou de les mettre à jour lorsque le consentement est accordé ou retiré.
    - D'enregistrer des données personnelles lorsque le système source n'est pas capable de fournir la censure/l'anonymisation ou les filtres nécessaires (afin de présenter uniquement des données personnelles aux processus et utilisateurs appropriés) ou qu'il ne fournit pas la sécurité nécessaire pour stocker les données personnelles
- La capacité de fournir des recherches et questionnements riches au sein des données personnelles et de leurs métadonnées afin de répondre aux requêtes et événements commerciaux (en cas de retrait de consentement, par exemple) et d'être assez flexible pour prendre en charge de nouvelles questions et interrogations au fur et à mesure que la réglementation évolue et que d'autres utilisations commerciales voient le jour pour la solution.
- La possibilité de fournir des capacités d'entreprise telles que la sécurité renforcée, une haute disponibilité et même la possibilité de fonctionner sur site ou dans le cloud (en fonction des exigences informatiques) puisque la solution sera en contact avec les données les plus importantes de l'entreprise.

CAPACITÉS DE MARKLOGIC

MarkLogic® est une base de données Enterprise NoSQL opérationnelle et transactionnelle, désignée pour intégrer, stocker, gérer et rechercher toutes vos données. En tant que base de données multi-modèle, elle peut gérer tout type de données, structurées ou non. Contrairement aux bases de données traditionnelles, elle ne requiert pas le développement de schéma important en amont ni de processus ETL coûteux pour les entreprises, tant en termes de temps que d'argent.

La plate-forme de base de données MarkLogic fournit une capacité de recherche multilingue de type Google* au sein des données de l'entreprise. Elle permet non seulement de poser des questions de découverte sur les données mais également des questions riches et complexes permettant de rassembler les informations nécessaires pour répondre à une requête réglementaire.

MarkLogic peut identifier les données personnelles au sein de vos données et annoter et enrichir ces informations, permettant l'enregistrement de la façon et de la raison dont elles sont utilisées, des personnes qui en sont responsables et de tout autre attribut pouvant être nécessaire pour répondre aux requêtes.

MarkLogic est une plate-forme à la sécurité certifiée, garantissant l'accès aux informations aux seules personnes autorisées. La censure et l'anonymisation des données sont possibles si nécessaire.

MarkLogic s'intègre aux processus commerciaux afin de s'insérer dans une solution globale pour le règlement général sur la protection des données et peut initier des processus de suppression et de modification des données dans ses systèmes source.

MarkLogic peut être installée sur site ou dans le cloud. La solution est rentable sur le plan opérationnel car elle nécessite seulement une partie des frais administratifs imposés par les autres plates-formes de bases de données.

COMMENT OPTIMISER LES EXIGENCES DE CONFORMITÉ POUR CRÉER DE NOUVELLES OPPORTUNITÉS

Le règlement général sur la protection des données est une loi exhaustive et potentiellement perturbatrice qui pourra entraîner des coûts élevés pour les entreprises qui ne seront pas en mesure de s'y conformer. Cependant, il présente des avantages qui peuvent être réalisés en adoptant la solution mise en avant dans ce livre blanc.

La gestion des données de cette manière, prise en charge par une nouvelle technologie novatrice, permet d'atteindre une visibilité intégrale des principales personnes faisant partie ou interagissant au sein de l'entreprise.

Ces informations peuvent faire la lumière sur les différents aspects de la personne, les processus commerciaux interagissant avec cette personne et ce que l'on sait (et par conséquent ce que l'on devrait savoir) sur elle. Elles offrent également aux entreprises la possibilité d'améliorer la satisfaction et la confiance des clients et des employés. Elles peuvent réduire les coûts en générant des processus commerciaux plus efficaces, améliorer la prise de décisions en utilisant des données plus précises et aligner les unités commerciales via une vision unifiée des clients/employés/citoyens, etc.

Une vision globale des données est un mécanisme puissant pour analyser les modèles de comportement des clients. Elle permet d'améliorer la popularité de l'offre de produit actuelle et de mettre en place un développement de produit novateur.

L'approche MarkLogic permet aux entreprises de transformer leurs efforts de conformité en valeur ajoutée. Elle offre une opportunité d'améliorer la satisfaction/confiance des clients/employés/citoyens, d'aider à réduire les coûts opérationnels et d'offrir des informations plus précises sur les profils de client et les produits qui les intéressent le plus.

PLUS D'INFORMATIONS

Plan de réussite pour les projets de données stratégiques
fr.marklogic.com/resources/plan-success-high-stakes-data-projects

Dépasser le modèle relationnel
fr.marklogic.com/resources/beyond-relational-french

Contactez-nous pour en savoir plus sur nos solutions sur fr.marklogic.com

MARKLOGIC FRANCE SAS
23, rue Balzac, Paris 75008, France
+33 (0) 153 536 784 | fr.marklogic.com | sales@marklogic.com

23, rue Balzac, Paris 75008, France
+33 (0) 153 536 784
fr.marklogic.com | sales@marklogic.com