

# IBM StoredIQ et RGPD

Accélérez votre évaluation des sources de données non structurées



## Qu'est-ce que le RGPD ?

Le Règlement Général sur la Protection des Données (RGPD) fait partie de la législation de l'Union européenne (UE) en matière de protection des données et de la vie privée. Depuis son entrée en vigueur le 25 mai 2018, les organisations qui collectent, conservent ou traitent des données à caractère personnel des personnes concernées de l'UE doivent respecter les exigences du RGPD. Le RGPD a pour objectif de simplifier les réglementations en matière de protection des données des pays de l'UE et de ceux de l'Espace économique européen, tout en traitant la question de l'exportation mondiale de données à caractère personnel relatives aux personnes concernées. Le règlement pourrait présenter un changement majeur dans votre traitement des données personnelles.

En vertu du RGPD, les individus ont le **droit de déterminer** ce qu'il advient de leurs données, y compris quelles informations sont utilisées, par qui et comment. Vos clients peuvent demander à votre organisation de modifier, supprimer, exporter ou s'abstenir de traiter leurs données personnelles et celle-ci est tenue de satisfaire ces demandes.

Les organisations ont besoin d'outils d'analyse de contenu pour transformer leurs données, car le non-respect du RGPD peut avoir de lourdes conséquences. Certaines ont du mal à identifier et à développer des solutions pour s'y préparer, mais il y a certaines étapes élémentaires que vous pouvez suivre pour commencer le processus.

## Préparation au RGPD

### Que peut faire votre organisation pour se préparer ?

Premièrement, évaluer les sources de données de votre organisation. Une fois que toutes les sources de données personnelles sont identifiées et inventoriées, les responsables des données peuvent évaluer l'écart entre comment sont gérées les données à caractère personnel et comment elles devraient être gérées conformément aux normes du RGPD. Pour combler cet écart, votre organisation peut avoir besoin de modifier l'emplacement de stockage des données, la manière dont elles sont transférées dans votre système ou la manière dont les employés réagissent aux demandes des clients et aux violations de données. Ces tâches peuvent paraître insurmontables compte tenu du nombre considérable de données conservées par les entreprises. Heureusement, **les solutions IBM® sont là pour vous aider.**

## Accélérez votre évaluation RGPD avec IBM StoredIQ

IBM® StoredIQ® résout les problèmes pouvant affecter la gestion de vos enregistrements grâce à l'analyse des sources de données non structurées. Cette analyse approfondie vous donne une meilleure visibilité sur l'emplacement où résident les données et vous permet de prendre des décisions opérationnelles mieux informées. Voici un aperçu plus détaillé du fonctionnement de StoredIQ :

### Obtenez une meilleure compréhension des données qui se trouvent dans vos sources.

- Découvrez et triez les données à l'aide de rapports et de cartes reposant sur les métadonnées
- Identifiez les aspects sensibles des données personnelles ainsi que les problèmes potentiels de conformité et de violations
- **Localisez les informations personnelles et sensibles** à l'aide de puissantes fonctionnalités de recherche avec des modèles prédéfinis et personnalisables.

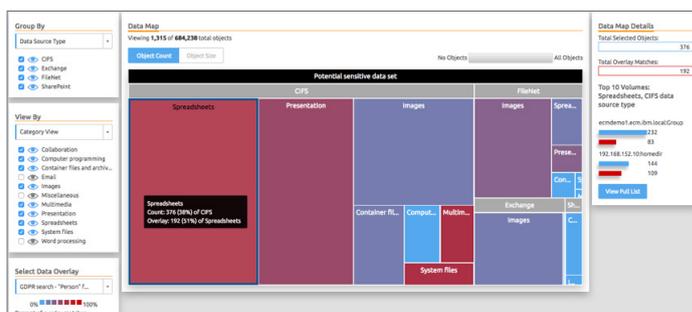


Figure 1. La carte de données IBM StoredIQ vous aide à visualiser les aspects sensibles des données personnelles.

### Simplifiez la gestion et l'analyse des données

- Localisez les informations personnelles et sensibles à l'aide de fonctionnalités de recherche puissantes ainsi que de modèles prédéfinis et personnalisables
- Prévisualisez les documents à partir des résultats de recherche en utilisant la fonction de surbrillance pour valider les correspondances
- Créez des rapports par terme pour les fichiers correspondant à des modèles ou à des règles de distribution spécifiques
- Vérifiez que les documents dont le contenu est sensible n'ont pas été stockés dans des endroits inattendus

### Créez un catalogue central avec des règles et politiques de gestion en matière de RGPD

- Intégrez StoredIQ à [IBM InfoSphere® Information Governance Catalog](#) (IGC) pour faire concorder les conditions et règles commerciales avec votre analyse
- Utilisez IGC pour les sources de données non structurées et l’Outil d’analyse des informations IBM InfoSphere pour les sources de données structurées.
- Intégrez les résultats de votre analyse dans votre catalogue central

### Générez de nombreux rapports pour éclairer la prise de décision

- Créez des rapports CSV permettant de consulter les détails des fichiers qui correspondent aux phrases recherchées
- Créez des rapports de gestion pour consulter à haut niveau la distribution et la quantité de données personnelles
- Configurez les niveaux de sensibilité pour chaque catégorie de données et partagez les rapports dans toute l’organisation.

### Soutenir les demandes des personnes concernées

- Utilisez des filtres prédéfinis dans StoredIQ pour identifier les données personnelles liées à une personne dans des sources connectées de données non structurées.
- Utilisez des filtres prédéfinis pour identifier et supprimer les données sélectionnées lorsque des personnes demandent que leurs données soient effacées.

StoredIQ prend en charge de nombreuses sources de données non structurées, notamment SharePoint, les serveurs de messagerie et les partages de fichiers. Pour obtenir une liste complète, [visitez le centre de connaissances IBM](#).

### Comment StoredIQ identifie les données personnelles

StoredIQ lit les données à partir de sources de données non structurées sélectionnées et crée un index répertoriant les métadonnées et le texte plein. Les documents restent en place : aucune donnée n’est transférée dans un entrepôt. L’index peut être enrichi de balises pour identifier des données personnelles.

Les plug-ins d’analyse, également appelés [cartouches](#), identifient des données personnelles et enrichissent l’index avec les résultats. Ces cartouches peuvent contenir une logique d’analyse basée sur différentes technologies allant de simples expressions régulières aux approches cognitives complètes, telles que le traitement du langage naturel (NLP) et l’apprentissage automatique. Une cartouche dans StoredIQ est conçue pour déterminer des informations telles que des noms, des numéros ou des adresses et permet de rechercher ces informations.

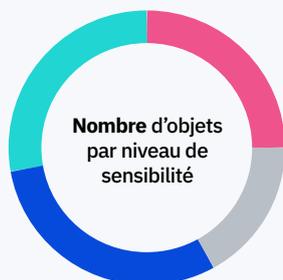
### IBM fournit des cartouches-échantillons pour le RGPD

Les cartouches de découverte de données axées sur le RGPD sont conçues pour aider les organisations dans leurs efforts pour se préparer au RGPD et aux autres réglementations en matière de confidentialité des données. IBM fournit deux cartouches StoredIQ axées sur le RGPD.

#### Distribution de données à caractère sensible :

##### Niveau de sensibilité:

- Élevé (47 449)
- Modéré (50 542)
- Faible (56 450)
- Inexistant (31 555)



##### Niveau de sensibilité:

- Élevé (214,57)
- Modéré (219,21)
- Faible (117,24)
- Inexistant (10977,09)



Figure 2. Représentation du rapport de confidentialité des données

**Cartouche de base axée sur le RGPD**

Basée sur des expressions régulières pour tous les pays européens dans lesquels les types de données à caractère personnel suivants sont pris en charge par défaut :

- Numéros de compte bancaire, IBAN international et versions nationales
- Numéros d'identification nationaux, tels que le numéro d'identification fiscale, de carte d'identité nationale ou le numéro de passeport
- Numéros de téléphone, adresses e-mail et adresses IP

**Cartouche avancée axée sur le RGPD**

Basée sur des modèles NLP ou SystemT, dans lesquels les types de données personnels suivants sont pris en charge :

- Noms de personnes
- Localisations
- Organisations
- Adresses
- Dates.

Les deux cartouches peuvent être personnalisées et étendues en fonction de vos besoins.

Pour plus d'informations sur le concept de cartouche et sur la création de cartouches en fonction de vos besoins, consultez ce [didacticiel IBM developerWorks](#).

StoredIQ a également des macros prédéfinies avec des types de données personnelles supplémentaires afin de soutenir la préparation à d'autres réglementations relatives à la confidentialité des données :

- Numéros de carte de crédit
- Numéros de sécurité sociale des États-Unis, qui peuvent inclure des tirets et des chiffres
- Numéros de téléphone des États-Unis, qui peuvent inclure des tirets, des points, des parenthèses, des chiffres et des lettres ; cette macro permet de filtrer les mnémoniques qui utilisent des lettres pour commercialiser des numéros de téléphone
- Adresses IP, qui peuvent inclure des points, des chiffres et des lettres
- Adresses e-mail, qui peuvent inclure des symboles tels que @, des points, des chiffres et des lettres.

De plus, vous pouvez créer un filtre en utilisant des listes de mots. Cette approche est utile lors de la recherche de données personnelles sensibles.

**StoredIQ Insights**

StoredIQ Insights permet une recherche interactive avec des nombres et des pourcentages mis à jour instantanément dans les composantes du filtre. La liste résultante fait l'objet d'une mise à jour dynamique lorsque vous saisissez une chaîne de recherche ou lorsque vous sélectionnez une composante dans la bibliothèque. Les critères de recherche sont surlignés dans l'aperçu de chaque document. Des annotations indexées créées avec des cartouches RGPD axées sur le RGPD sont disponibles dans le volet des composantes et peuvent être utilisées ou non dans le filtrage. Cette fonctionnalité peut vous aider à rechercher des données personnelles et à valider les résultats. Pour plus d'informations sur la production d'IBM StoredIQ, consultez la page [IBM StoredIQ Marketplace](#).

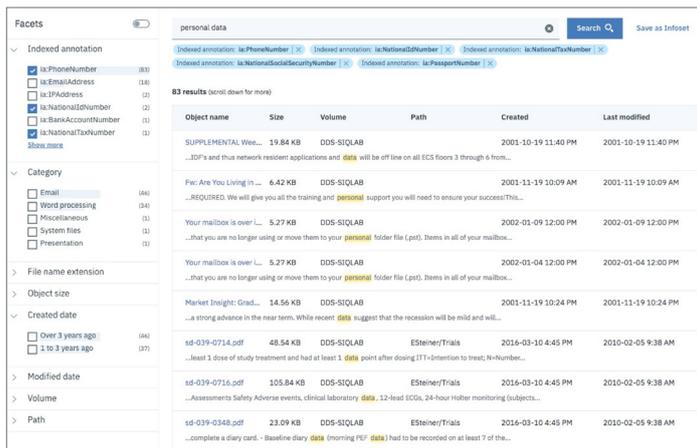


Figure 3. Interface utilisateur de StoredIQ Insights



## Pour en savoir plus

IBM StoredIQ fait partie des solutions de gouvernance de données IBM. Ces solutions vous permettent d'obtenir une meilleure compréhension de vos données, qu'elles soient structurées ou non, et de les gérer pour vous aider à réduire vos risques de non-conformité et à responsabiliser vos intervenants en matière d'informations. Pour en savoir plus sur les solutions de gouvernance des données IBM, visitez [ibm.com/analytics/data-governance](https://www.ibm.com/analytics/data-governance).

De plus, IBM Global Financing propose de multiples solutions pour vous aider à acquérir la technologie dont vous avez besoin pour développer votre activité. Pour en savoir plus, visitez [ibm.com/financing](https://www.ibm.com/financing).

IBM France  
17 Avenue de l'Europe  
92275 Bois Colombes Cedex

IBM, le logo IBM, InforSphere, StoredIQ et **ibm.com** sont des marques d'International Business Machines Corp., déposées dans de nombreux pays du monde. Les autres noms de produits et de services peuvent être des marques d'IBM ou d'autres sociétés. Une liste actualisée des marques déposées IBM est accessible sur le Web sous la mention « Copyright and trademark information » à l'adresse [www.ibm.com/legal/copytrade.shtml](https://www.ibm.com/legal/copytrade.shtml).

Ce document est considéré comme à jour à sa date initiale de publication et peut être modifié par IBM à tout moment. Toutes les offres ne sont pas disponibles dans chaque pays dans lequel opère IBM.

Les exemples de clients fournis ne sont mentionnés qu'à titre d'information. Les performances réelles peuvent varier selon les configurations et les conditions de fonctionnement spécifiques. Il appartient à l'utilisateur d'évaluer et de vérifier le fonctionnement d'autres produits ou programmes avec les produits et programmes d'IBM. LES INFORMATIONS CONTENUES DANS CE DOCUMENT SONT FOURNIES « EN L'ÉTAT », SANS AUCUNE GARANTIE EXPRESSE OU TACITE, NOTAMMENT SANS AUCUNE GARANTIE DE QUALITÉ MARCHANDE OU D'ADAPTATION À UN EMPLOI SPÉCIFIQUE, ET SANS AUCUNE GARANTIE OU CONDITION DE NON-INFRACTION VIS-À-VIS DES LOIS. Les produits IBM bénéficient d'une garantie, conformément aux conditions générales des contrats dans le cadre desquels ils sont fournis.

© Copyright IBM Corporation 2019



Veillez recycler