

ZL GDPRソリューション

個人情報アクセス要求への対応

一般データ保護規則が2018年5月25日に発効したため、EU居住者のデータを管する組織は、個人データへのアクセスまたは削除を求める個人からの要求である個人情報アクセス要求に迅速に対応することが求められています。

現在、組織が個人情報アクセス要求を効果的に満たすことを阻む2つの基本的な技術的障害があり、これに対処できるソリューションはほとんどありません。

1. スケーラビリティ
2. 複雑さ

スケーラビリティ

大規模な組織の多くは、数十年にわたってさまざまなデータサイロを蓄積しており、それらを橋渡しする能は、現代の最大のIT課題の1つになりつつあります。データリポジトリの数が増え続けるにつれて、個人データを検出する場所の数も増えます。もし明日、個人情報アクセス要求が来たら、組織はいくつのリポジトリを用意する必要があるのでしょうか？大規模な組織では、次のような数十または数百のサイロにまたがるデータが存在する場合があります。これには次のものが含まれます：

- » 形式ベースのサイロ（電子メール、ファイル共有、SharePoint、IM、ECM など）
- » 地理的な場所ベースのサイロ
- » 管理ベースのサイロ（インプレース vs. アーカイブ、クラウド vs. ハイブリッド）
- » 機能ベースのサイロ（eDiscovery、規制コンプライアンス、レコード管理）

これらのリポジトリを個別に検索することは不可能であり、各検索エンジンには独自の制限がある可能性が高いため、一元化されたプラットフォームからすべてのサイロを検索する機能が不可欠です。

複雑さ

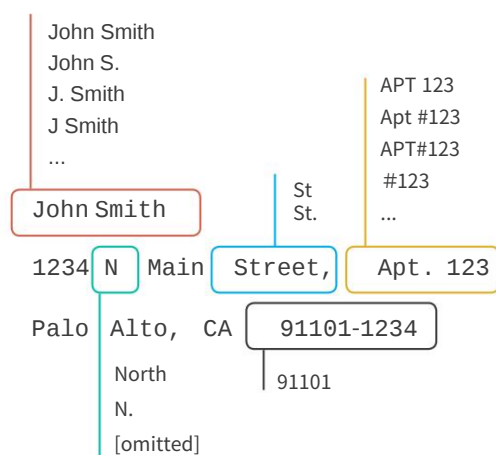
サイロ問題以外にも、個人情報アクセス要求を満たすには依然として複雑な問題が伴います。つまり、個人情報を含む文書を正確かつ防御的に特定することができない一方で、それがそれぞれの主題に関連していることを保証できないのです。検索エンジンが制限されていると、主に次の2つの由から、個人情報アクセス要求に対する防御的な対応ができなくなります。

1. 大量の無関係な非個人情報が入るため、公開または削除前に文書を確認するのに必要な時間が倍増します。各文書を確認するには、次の点を確認する必要があります。

- 対象者に関連し、個人データが含まれます。
- 非公開または編集すべきその他の機密データは含まれていません。
- ビジネス、記録管理、法律、規制目的などの「正当な利益」により削除できない重要な情報は含まれていません。

2. 対象者の個人データを含むすべての文書を見つけることができない可能性があり、その結果、検索が不完全になる可能性があります。

この単純な名前と住所の組み合わせだけで、検索する必要があるバリエーションを考えてみましょう。



完全かつ正確な結果を得るためにこの検索に必要な操作は簡単ではありません。ユーザーは、複数のバリエーションやスペルを持つ複数のキーワード、主題の名前から特定の距離内に表示されるキーワード（主題に関連していることを保証するため）、複数の形式で表示される数値、およびほとんどの検索エンジンでは見逃されるその他の多数の基準を検索する必要がある場合があります。たとえば、「Street」は、「St.」や「St」などの他の形式で表示されることがあります。同じことは「North」という単語にも当てはまり、「N」として表示されるか、完全に省略される可能性があります。また、主題の完全な住所がドキュメントに表示されるが、フルネームが表示されない場合や、その逆の場合もあります。

ほとんどの検索エンジンは、検索する必要がある個人データの種類ごとに考えられるすべてのバリエーションは言うまでもなく、住所が表示される可能性のあるすべてのバリエーションを考慮していません。さらに、時間制限があるため、ユーザーはすべての変数を同時に検索する必要があります。そうしないと、各変数の組み合わせには多くの個別の検索が必要になり、各リポジトリ間でその検索が倍増します。現在のアプローチでは、組織は1つのサブジェクトアクセス要求に対して何百もの個別の検索を実行しなければならない可能性があります。

スケーラビリティと複雑性の解決

多くのソリューションでは、組織が500万のドキュメントに対して複雑な検索を実行できますが、この数を50億に変更すると、市場に出回っている多くのプラットフォームは機能しなくなります。

同様に、企業全体の検索を提供しようとしている製品もありますが、膨大なデータ量により、そのアーキテクチャではGDPRに必要な詳細な検索と分析を提供することができません。

結局のところ、一部のソリューションは検索の複雑さを提供し、他のソリューションはスケーラビリティに近づいているかもしれませんが、これらの両方のコンポーネントを同時に提供することは、ほとんどのソリューションでは解決できない基本的な問題です。

eDiscoveryの世界を見れば、一方は提供していてももう一方は提供していないソリューションの例が数多くあります。そのため、eDiscoveryのより複雑な処理および分析フェーズは、データ量が削減された後のEDRMプロセスの後半に取っておく必要があります。

類似点があるにもかかわらず、GDPRのサブジェクトアクセスリクエストはいくつかの重要な領域でeDiscoveryから逸脱しており、難易度が上がっています。

» 個人情報アクセス要求は、eDiscoveryケースで通常発生するよりもはるかに頻繁に、また大量に発生することが予想されます。そのため、組織は、eDiscovery検索よりもサブジェクトアクセスリクエストを実行する時間が短くなります。

» GDPRの個人情報アクセス要求では、特定の管理者に属するデータ全体ではなく、膨大な量のデータ全体にわたって複雑な検索操作を実行する必要があります。

» 検索に「ベストエフォート」基準に依存するeDiscoveryケースとは異なり、GDPRでは、個人データが見つからない場合に要求元のデータ主体からの異議申し立てを回避するために、組織は場所に関係なくすべての個人データを検索する必要があります。

» GDPRでは、文書の単一のコピーだけでなく、すべての複製を管理することが求められます。

» GDPRで要求される検索の規模はより大きく、すべてのエンタープライズリポジトリに及びます。

すべてのサイロを検索し、数十億のドキュメントにわたるデータエントリのすべてのバリエーションをタイムリーに把握する機能がなければ、サブジェクトアクセスリクエストは防御不可能になり、組織は金銭的制裁を受けることになります。

ZL Unified Search

ZL GDPR ソリューションは、スケーラビリティと複雑さの両方の課題を1つのプラットフォームで解決します。ZLは、スケーラブルで柔軟なBigDBアーキテクチャを活用して、数十億のドキュメントと数百の固有のデータソースを含む世界最大規模のエンタープライズデータ環境を検索および分析できます。これには次のものが含まれます。

- | | |
|---------------|-------------|
| » ファイル共有 | » SMSとMMS |
| » 電子メール | » ソーシャルメディア |
| » SharePoint | » IM |
| » ECM Systems | |

ZLの並列クエリエンジンを使用すると、組織はカスタマイズされた企業全体の検索を実行して、関連する個人データを識別し、修正することができます。名前と住所の検索サンプルをもう一度見てみると、ユーザーはZLの検索エンジン内で次のクエリを入力して、エンタープライズリポジトリでJohn Smithの名前と住所を含むドキュメントを検索できます。

```
(John) w/2 (Smith) w/3  
(1234 w/2 (Main w/2 (Street OR St. OR St))  
w/3 (Ap* w/2 123) w/2 Palo Alto  
w/2 CA w/291101*)
```

このクエリの条件は、次のコンポーネントに分けられます。

1. ワイルドカード拡張「*」: 複数のバリエーションを持つキーワードを識別します。「Ap*」は「Apt,」「Apartment,」などを含む文書を返します。

2. 近接ルール「w/[数値]」: これらの条件の対象となるドキュメントには、互いに指定された距離内にキーワードが含まれている必要があります。この条件は、キーワードと主題の関連性、または2つのキーワードが互いに関連していることを確認するために使用できます。

3. ファジーロジック: スペルミスのあるキーワードや類似の単語を含むドキュメントを返します。

4. ブール論理「AND/OR」: 最大限の検索柔軟性を実現します。

示されているように、キーワードと条件の小さなサンプルでは、完全な検索を行うために高度な複雑さが必要です。この複雑さは、一般的なサブジェクトアクセス要求の場合と同様に、キーワードと条件の数が増えるにつれて指数関数的に増加します。これに、数十億に及ぶ可能性のあるドキュメントを検索するという要件を組み合わせると、負荷を処理できるソリューションはほとんどありません。

ZLの並列クエリエンジンは、クエリとなる条件ごとに複数の検索を並に実することで機能します。各条件に回答するドキュメントは相互に参照され、必要な各条件に回答するドキュメントが表示されます。ユーザーは、これらのドキュメントにタグを付けたり、保持ポリシーを適用したり、エクスポートしたり、削除したりすることを選択できます。

地球上の最大規模の企業のエンタープライズ データをサポートするためにゼロから構築された ZL の非常にスケーラブルなアーキテクチャは、グローバル規模で要求される複雑な検索を提供できる点で他に類を見ません。

ZLGDPR ソリューションは、最終的に、個人データがどこにあっても、GDPR コンプライアンスのために個人データを検して対処するための真の道筋を提供します。

受賞歴



2018年 Gold Stevie® 賞
最優秀コンテンツ分析
ソリューション部門
International Business
Awards



KMWorld Trend-Setting
Products of 2018



2017年 Gold Stevie® 賞
最優秀 ビッグデータ
ソリューション部門
American Business
Awards