
記事

[Tomoko Furuzono](#) · 2020年9月7日 6m read

フリーテキスト検索：SQL開発者が秘密にしているテキストフィールドの検索方法*

アプリケーションに、効率的に検索したいフリーテキストを含むフィールドがありますか？これまで複数の方法を試してみたものの、顧客が要求するパフォーマンスを満たせなかった経験はありませんか？私は変わった手段を使ってあらゆる問題を解決できると思っていませんか。もうご存じですよ。私ができるのは、パフォーマンス低下に対処する優れたソリューションを提供することです。

いつものように、要約版が必要な場合は記事の最後まで飛ばしてください。ただ、それだと私はがっかりしてしまいますが。

最近の（2015.1以降の）バージョンのCaché/Ensemble/HealthShareのSAMPLESネームスペースでSample.Companyのバージョンを開くと、擬似ランダムに生成されたテキストであるMissionフィールドが表示されます。このテキストフィールドを検索してみましょう。私はこの演習のために約256,246社データを生成しましたが、ご自身で必要な数の会社を生成してから同じ手順に従ってください。例えば、次のクエリを実行するとしましょう。

```
SELECT * FROM Sample.Company WHERE Mission LIKE ' % agile % '
```

これはかなり合理的なクエリですが、どのように実行されるのでしょうか？もし、インデックスがない場合は間違いなく各エントリを読み取る必要があるため、7854レコードを取得するのにグローバル参照が256,277回発生してしまいます。これはあまりよくありません！Sample.Companyにインデックスを追加し、より合理的に実行できるかどうかを見てみましょう。次の1行を追加します。

```
Index MissionIndex on Mission;
```

では、インデックスを構築して同じクエリを実行してみましょう。どんな結果になるのでしょうか？グローバル参照は279,088回でした。

皆さんの「でもKyle、グローバル参照が増えているじゃないか！これはひどくないか？インデックスが効果的だと思っていたのに！！！」という声が聞こえてくるようです。

ちょっと落ち着いてください。(それに、感嘆符3つは多すぎます。) また、不規則な挙動を直面した場合は、少し時間をかけて考えることをお勧めします。インデックスを読み込むコストと、全レコードを読み込むコストはどちらが大きいのでしょうか？インデックスの方が小さいため、テーブル全体をスキャンするよりはMissionIndex全体を読み込むほうがコストは小さくなるでしょう。その後はテーブル全体の一部を読み取って表示すればいいだけのことです。そのため、グローバル参照は増えますが、作業は少なくなります（全データがディスク上にあると仮定した場合）。もちろん、この挙動についてはさらに多くを語ることができますが、そのためにはCachéブロックの構造について広範囲に説明しなければならないでしょうが、それは私がここで説明しようとしていることからはずれてしまいます。

最初の対策とクエリの削減には成功したかもしれませんが、どう見ても私たちが望んでいたほどの結果は得られていません。何か即効性のある解決策が必要です。グローバル参照を減らし、しかも簡単にそれを実現したいのです。どうすれば良いのでしょうか？その答えは、iFind Indexです。

さて、皆さんはiKnowのことも、iKnowに響きが妙に似ているiFindのこともご存じないかと思います。皆さんが求めているのは必要なパフォーマンスを得るためのSQLソリューションであり、新しいテクノロジーの詳細を学びたいと思っているわけではありません。ご心配なく。iFindはiKnowエンジンを利用していますが、iKnowの知識が全くゼロでも利用することができます。では、確かめましょう。次のようなインデックスを定義するとします。

```
Index MissioniFind on (Mission) as %iFind.Index.Basic;
```

%iFind.Index.Basic

とはどういう意味なのでしょう？まあ、それは置いておきましょう！美しいと思いませんか！？これを入れていつもと同じようにインデックスを構築してください（%BuildIndices()）。このインデックスによってクエリの挙動が若干変わります。この新しく素敵なインデックスを使用することを次のようにクエリに伝える必要があります。

```
SELECT * FROM Sample.Company WHERE %ID %FIND searchindex(MissioniFind, ' agile ')
```

入力する必要があるのは、インデックス名と検索対象の単語だけです。インデックスを追加し、構築し、以前よりも若干変わったSQLを使用するのは大変です。果たして、このクエリにそれだけの価値はあるのでしょうか？このクエリでも7854レコードが返ってきましたが、グローバル参照は7928回でした。

グローバル参照が7928回ですよ！グローバル参照数はレコード数よりもわずかに多いだけです。これは素晴らしい結果になりました。きっと皆さんは質問したいことがあるはずです。どんな質問なのかは予想できますし、その質問に回答する用意があります！

このインデックスを他のインデックスと組み合わせることはできますか？

いい質問ですね！はい、できます！このテクノロジーでは、インデックスを組み合わせるのが非常に効果的です。

制限はありますか？

残念ながら、制限はあります。ビットマップ対応のIDが必要です。つまり、テーブルのIDが正の整数である必要があります。言い換えれば、複合IDや文字列IDなどではありません。

これで本当にうまく行きますか？

うまく行きます！フリーテキストフィールドがあり、少なくともiFindインデックスを試したことがないのであれば、本当にもったいないことをしています！

では、iKnowはどうなのでしょう？iFind

インデックスはiKnowエンジンを利用していますが、既存のアプリケーションでiFindを使用するのにiKnowの知識を身に付ける必要はありません。定義して構築し、使用するだけです！

詳細はどこで確認できますか？もちろん、ドキュメントも用意しています。こちらでご確認ください：

<http://docs.intersystems.com/latest/csp/docbook/DocBook.UI.Page.cls?KEY=Giknowifind>

こちらにもドキュメントがあります。

<http://docs.intersystems.com/latest/csp/documatic/%25CSP.Documatic.cls?PAGE=CLASS&LIBRARY=%25SYS&CLASSNAME=%25iFind.Index.Basic>

iFindには他にも優れた機能がありますか？もちろんです！iFindはiKnowを使用しているため、多彩な能力を持っています！あいまい検索、ステミング/組み入れ、ランキングを実行でき、iKnowエンティティを検索できます。ご興味があれば、%iFind.Index.Semanticおよび%iFind.Index.Analyticインデックスクラスをご覧ください。これらは、上記で使用した基本的なものよりも多くのiKnowの付加機能を使用します。または、こちらの開発者コミュニティの投稿でiFindイ

ンデックスの上で動作するデモインターフ

ェイスを確認してください：<https://community.intersystems.com/post/iknow-demo-apps-part-5-ifind-search-portal>

もはやこの投稿でカバーできる範囲を超えています。これらのテーマについて詳細を確認したい場合はご質問ください！私はいつでも皆さんに必要な情報を伝え、ご説明します！

概要：フリーテキストフィールドで検索を行う場合は、iFindを次のように定義して追加してください：

```
Index <IndexName> on <FreeTextField> as %iFind.Index.Basic
```

%BuildIndicesを使用してインデックスを構築します（通常通り）

次のようにクエリを書き換えます：

...WHERE ID %FIND search_index(<IndexName>,<Search Value>) AND ...

そして、超高速なフリーテキスト検索のメリットを享受してください！

*その方法は秘密でも何でもありません！

[#iFind](#) [#オブジェクトデータモデル](#) [#ObjectScript](#) [#SQL](#) [#インデックス付け](#) [#Caché](#) [#InterSystems IRIS](#)

ソースURL:

<https://jp.community.intersystems.com/post/%E3%83%95%E3%83%AA%E3%83%BC%E3%83%86%E3%82%AD%E3%82%B9%E3%83%88%E6%A4%9C%E7%B4%A2%EF%BC%9A%96%8B%E7%99%BA%E8%80%85%E3%81%8C%E7%A7%98%E5%AF%86%E3%81%AB%E3%81%97%E3%81%A6%E3%81%84%E3%82%8B%E3%83%86%E3%82%AD%E3%82%B9%E3%83%88%E3%83%95%E3%82%A3%E3%83%BC%E3%83%AB%E3%83%89%E3%81%AE%E6%A4%9C%E7%B4%A2%E6%96%B9%E6%B3%95>