

MarkLogicのセマンティック

セマンティックは、関係性やコンテキストに基づくデータモデル手法に新しいアプローチを提供します。セマンティックデータである「トリプル」は相互にリンクされ、事実と関係の数千億件のファクト(事実)と関係性のグラフを構成します。MarkLogic®では、ドキュメント、データ、トリプルを密接に関連させて利用できます。トリプルはデータをリンクし、メタデータを記述し、コンテキストを提供します。この新しいモデルにより、組織はよりスマートなアプリケーションを構築し、散在するデータをより簡単に統合できます。

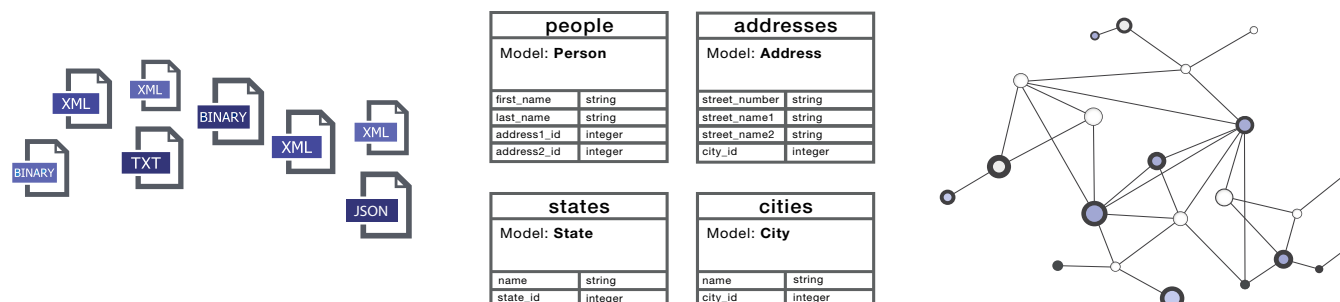
基本的にセマンティックとは、2つのエンティティ(ユーザー、場所、モノ)を関係性に基づいてリンクし、トリプル(3つ組)を構成するデータモデルです。リンクすることでトリプルはグラフを形成します。グラフには階層構造は無く、コンピューターにも理解でき、新たなファクトを推論できます。トリプル記述の標準言語はRDF(Resource Description Framework)であり、標準クエリ言語はSPARQLです。

データはトリプルに格納されます。以下のようになります

Subject	Predicate	Object
John Smith	livesIn	London
London	isIn	England

クエリはSPARQLで行います。単純なルックアップ以上のことができます。イングラウンド(内の場所)に住んでいる人を探せます。

W3C標準に基づき、トリプルはRDF(リソース記述フレームワーク)で表現され、SPARQLでクエリされます。



エンタープライズNoSQLデータベース+トリプルストア

- ・ **ドキュメント+データ+トリプル** :MarkLogicは、ドキュメント、データ、トリプルをネイティブに保存し、クエリを実行できる唯一のデータベースです。トリプル、位置情報データ、レンジ値、フリーテキストを含むクエリの組み合わせも可能です。
- ・ **エンタープライズ仕様** :ACIDトランザクション、スケーラビリティと弾力的な拡張性、政府レベルのセキュリティ、高可用性/災害復旧への対応。MarkLogicは、これらすべての機能を備えた唯一のトリプルストアです。
- ・ **セマンティック+検索** :MarkLogicに組み込まれた検索機能では、読み込まれたトリプルがすぐに検索の対象となります。特殊なトリプルインデックスが検索性を向上させます。
- ・ **大規模なスケールとハイパフォーマンス** :MarkLogicは、数千億に及ぶ膨大な数のトリプルを、シェアードナッシングアーキテクチャに格納します。大規模なクラスターに拡張した場合でも、パフォーマンスは低下しません。

提供される機能

- ・ 数千億個のRDFトリプルを格納・管理
- ・ ドキュメント/データ/トリプルを対象としたクエリ
- ・ トリプルインデックスで1秒以内に検索結果を表示
- ・ 大規模クラスタ全体を活用して高いパフォーマンスを出すためにトリプルをキャッシュ
- ・ MarkLogic Content Pump (mlcp) によるトリプルの一括読み込み
- ・ メタデータによる出自情報ならびに具体化 (reification)
- ・ シリアライゼーションおよび推移閉包のためのXQueryヘルパーモジュール
- ・ MarkLogic APIによる更新と集計
- ・ グラフトラバーサル(プロパティパスと推移閉包)
- ・ クエリ時にルールセットを使用したセマンティック推論
 - RDF、RDFS+、OWL Horst用のルールセット
 - ユーザー定義のルールセット
- ・ SPARQL 1.1の完全サポート
- ・ SPARQLエンドポイントとグラフストア プロトコルのサポート
- ・ サーバーサイドJavaScript、Node.jsによるSPARQL
- ・ JenaとSesame APIのサポート
- ・ セマンティックのテクノロジーパートナー (Smartlogic、Pool Party、Cambridge Semantics)との完全統合
- ・ MarkLogicの基幹業務に対応したエンタープライズ機能: ACID トランザクション、認定済みのセキュリティ (ドキュメントとトリプルレベル)、高可用性と災害復旧、スケーラビリティと弾力的な拡張性

セマンティックの実践

- ・ **より柔軟なデータモデリング** : JSON、XML、RDFなど、データモデル化の選択肢が広がります。MarkLogicでは、トリプルでドキュメントに注釈を付けたり、ドキュメントでトリプルに注釈を付けたり、またはドキュメントにトリプルを埋め込むことも可能です。
- ・ **業界・分野ごとにナレッジを管理** : オントロジーを使用して、データに対して特定ビジネス分野のコンテキストを提供します。例えば、金融サービス業界ならではの「法人」の意味を厳密に定義できます。
- ・ **検索と検出の性能を向上** : セマンティックグラフを活用することで、標準的なドキュメント検索が強化されます。トリプルでドキュメントを関係づけられるほか、トリプルをメタデータとして挿入することで多くの情報を提供できます。
- ・ **コンテンツを動的にパブリッシュ** : セマンティックな関係性に基づき、大規模なコンテンツパブリッシングを自動的に、リアルタイムに実行します (例えば、試合の得点、チームのプロフィール、ニュース記事のようなイベントデータ)。
- ・ **データ統合を迅速かつ容易に** : セマンティックはデータを互いにマッピングする「接着剤」のような役割を果たし、ETLを迅速化します。共通のセマンティック定義により、リレーショナルテーブルのエンティティ同士がリンクされます。
- ・ **エンティティの抽出とエンリッチメント** : 非構造化データから意味のある情報を抽出し、エンリッチメントを行うセマンティック技術を提供する組織と、MarkLogicはパートナーになっています。非構造化データから得られたトリプルは、MarkLogicに格納されます。

MarkLogicについて

MarkLogicが提供する、強力、アジャイルで信頼性の高いエンタープライズNoSQLデータベースのプラットフォームは、10年以上にわたる実績があります。米国政府や大企業をはじめさまざまな組織においてあらゆる種類のデータの価値を高め、実際の活動に繋がる情報をもたらしています。詳しくは、jp.marklogic.comをご覧ください。

© 2016 MARKLOGIC CORPORATION. ALL RIGHTS RESERVED. このテクノロジーは米国特許番号7,127,469B2、米国特許番号7,171,404B2、米国特許番号7,756,858 B2、および米国特許番号7,962,474 B2で保護されています。MarkLogicは米国およびその他の国におけるMarkLogic Corporationの商標または登録商標です。ここに記載されているその他すべての商標または登録商標は各社の所有物です。

マークロジック株式会社 MARKLOGIC K.K. 150-0043 東京都渋谷区道玄坂 1-12-1 渋谷マークシティウエスト 22 階
+81 3 4360 5354 | jp.marklogic.com | MarkLogic-JP@marklogic.com