

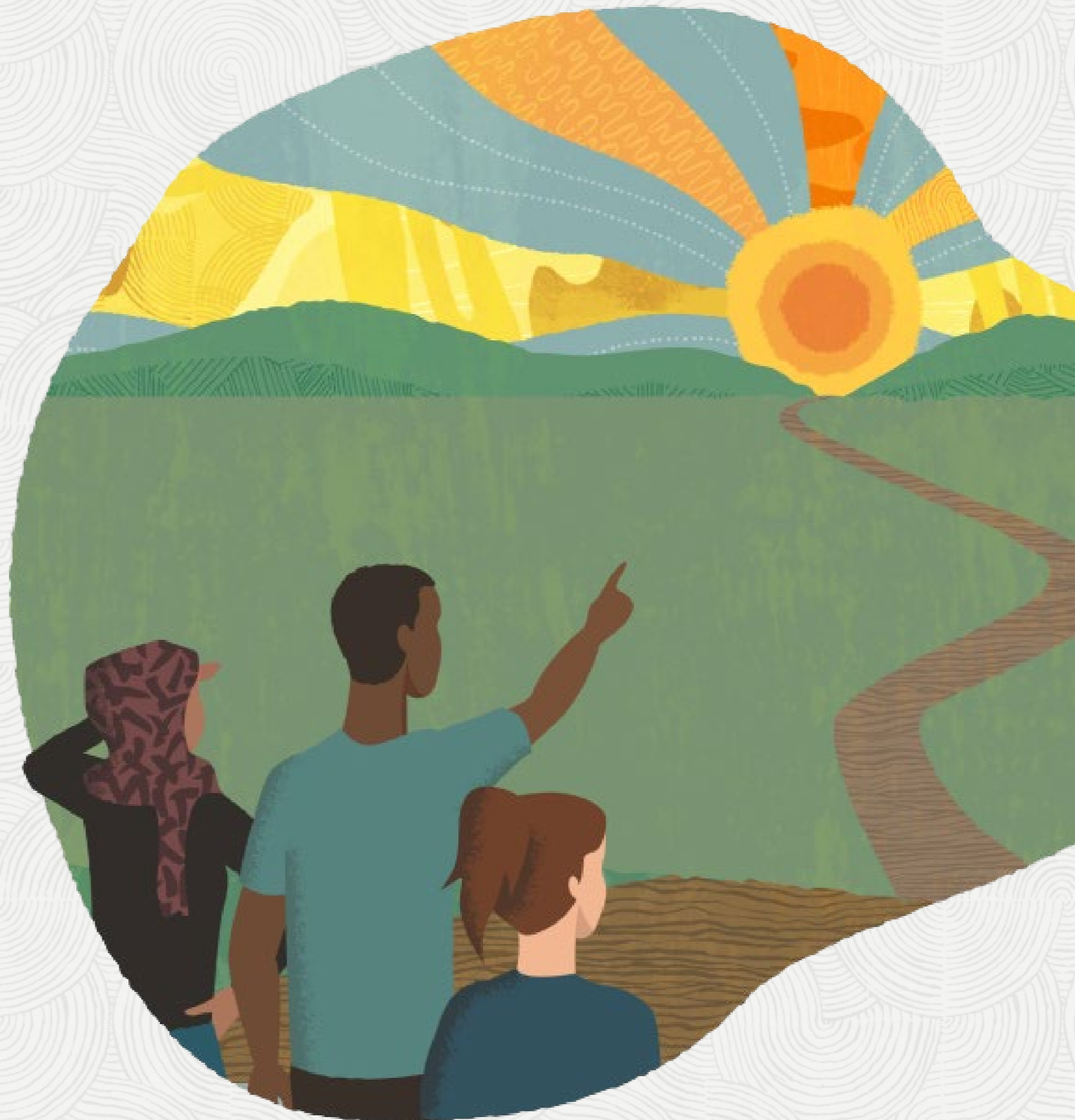
ORACLE

グラフ・データベースとグラフの
分析に関する17の
ユースケース



目次

はじめに	3
グラフが重要である理由	4
グラフ・テクノロジーとは	5
プロパティ・グラフに関する17のユースケース	6
金融サービス	7
製造	10
政府機関	13
データ規則とプライバシー	16
マーケティング	18
AIと機械学習の研究	21
オラクルのグラフ・テクノロジーを使用する理由	23
詳細情報	24





はじめに

ソーシャル・ネットワーク分析の実行、不正な銀行取引の発見、または製品の推奨事項の提供を行う必要があるとしましょう。

たいていは、これらの各質問への回答を見つける作業は複雑で、時間がかかる場合があります。

しかし、グラフ・データベースを使用すると、データの状況を完全に新しい方法で把握できます。新しい知見を発見してください。複雑な問題を解決してください。無限の可能性を解き放ってください。

グラフが 重要である理由

グラフ・テクノロジーは、世界中の組織が、その他の手法ではどうしても効率的に対処できない使用事例に対処する上で革新的な手法となっています。実際のところ、Gartnerは2年連続で、劇的な変化をもたらす可能性が大きいという理由で、分析とデータにおけるトップ・トレンドの1つにグラフを選出しました。現在の世界において、企業は、革新的でなければ破壊されることを分かっています。



グラフでは、データ・エンティティ間の関係とつながりが捕捉されます。これらの関係とつながりは、データ分析で使用できます。多くのデータはつながっており、グラフを使用すると、これらのつながりを調査し、新しい結論を導き出すことが容易になるため、グラフの重要性はますます高まっています。

グラフとグラフ・データベースは、関係を表すグラフ・モデルを提供します。グラフ・モデルを使用すると、パターン認識、分類、統計分析、および機械学習をこれらのモデルに適用できるため、大量のデータに対するより効率的で大規模な分析が可能になります。

グラフの分析に関しては、アルゴリズムにより、頂点間のパスと距離、頂点の重要性、および頂点のクラスタを調査します。多くの場合、アルゴリズムは、受信するエッジ、隣接する頂点の重要性、および他のインジケータを参照して重要性を判断します。

グラフ・データベースでは関係が明示的に保存されるため、頂点間のつながりを使用したクエリーとアルゴリズムを時間単位や日数単位ではなく秒未満で実行できます。ユーザーは数限りない結合を実行する必要がなく、データを分析と機械学習にいつそう簡単に使用して、私たちの周りの世界に関する事実をより多く発見できます。

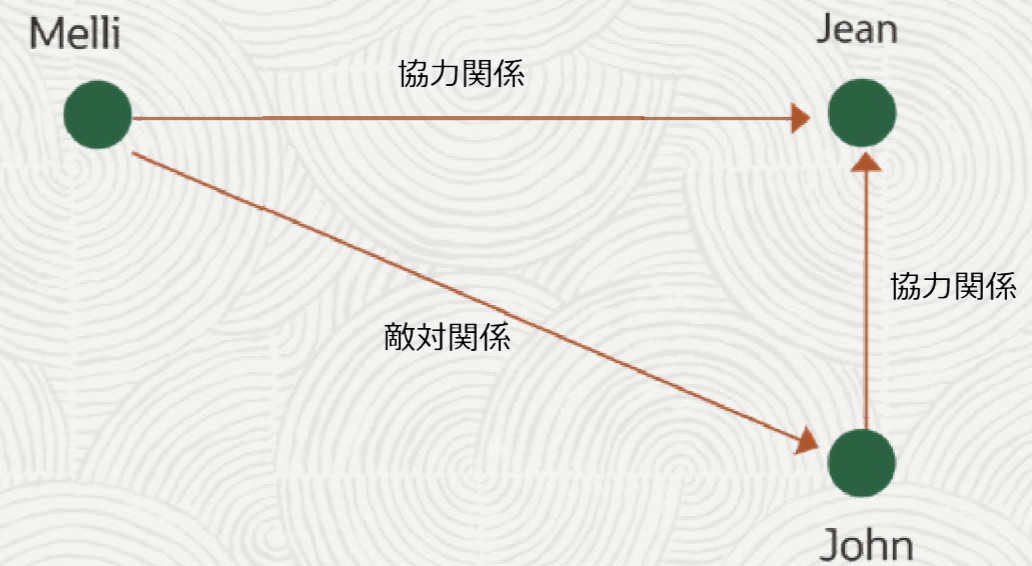
グラフ・テクノロジー とは

グラフには、プロパティ・グラフとRDFグラフの2つのタイプがあります。プロパティ・グラフでは分析とクエリーに焦点を当てますが、RDFグラフではデータ統合を重視します。これらのグラフのタイプは両方とも、点（頂点）の集合とこれらの点間のつながり（エッジ）で構成されています。しかし、違いもあります。

プロパティ・グラフは、データ間の関係をモデル化するために使用され、これらの関係に基づくクエリーとデータ分析を可能にします。プロパティ・グラフには、対象に関する詳細な情報を含めることができる頂点と、頂点間の関係を示すエッジがあります。

以下の例では、Melli、Jean、およびJohnはすべて頂点であり、“協力関係”と“敵対関係”は各頂点間の関係を示すエッジです。各頂点には、Melli、Jean、およびJohnに関する詳細情報を含めることができます（たとえば、彼らがどこに住み、何が好きかなど）。

プロパティ・グラフは非常に用途が広いため、金融、製造、公衆安全、小売、その他多くの幅広い業界と分野で使用されています。



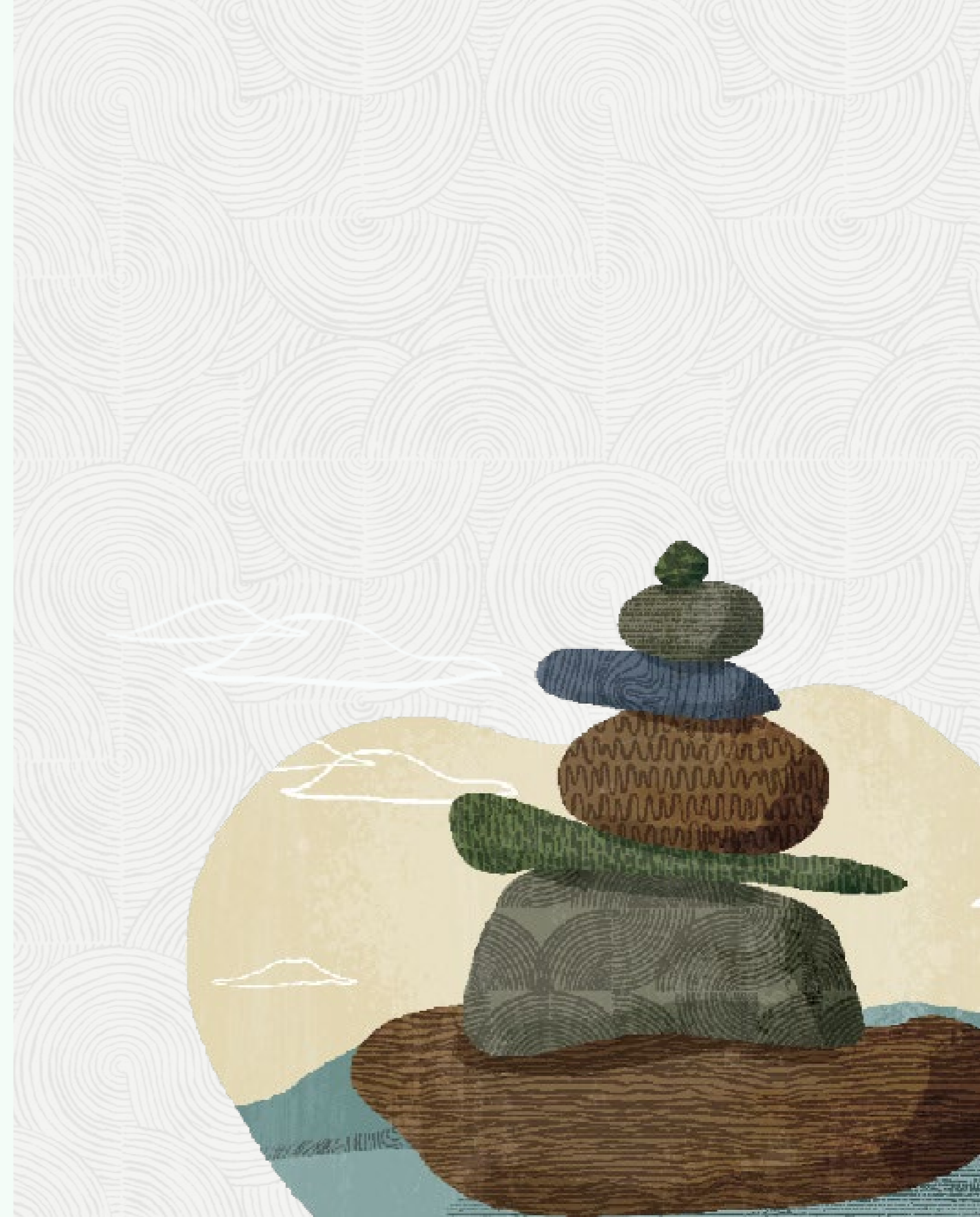
RDFグラフ（RDFはリソース記述フレームワークを表します）は、文を表すために設計されたもので、複雑なメタデータとマスター・データを表すのに最適です。RDFグラフは多くの場合、ドメイン内の複雑な概念や、データに関する豊富なセマンティックと推論を必要とする状況での複雑な概念を表すために使用されます。

RDFモデルでは、文は3つの要素（エッジでつながった2つの頂点）によって表現されます。すべての頂点とエッジが一意的URI（Unique Resource Identifier）によって識別されます。RDFモデルは、明確なセマンティックを使用した標準形式でデータを公開する方法を提供し、情報の交換を可能にします。RDFグラフを採用した組織のタイプには、政府の統計機関、製薬会社、医療会社などがあります。

プロパティ・グラフに関する 17のユースケース

世界中の組織がグラフ・テクノロジーを使用し始めています。このEブックでは、以下の業界とカテゴリで構成されたもっとも一般的なグラフの使用事例のいくつかについて説明します。

- [金融サービス](#)
- [製造](#)
- [政府機関](#)
- [データ規則とプライバシー](#)
- [マーケティング](#)
- [AIと機械学習の研究](#)



金融サービス

金融犯罪者は、どれだけ犯罪行為に熱心に取り組もうとも、その他の犯罪者、場所、そしてもちろん銀行口座のいずれかとの関係であろうと、関係によって結び付けられています。グラフ・テクノロジーでは、この事実を活用して、金融サービスの世界における新しい可能性を明らかにします。

マネー・ロンダリング

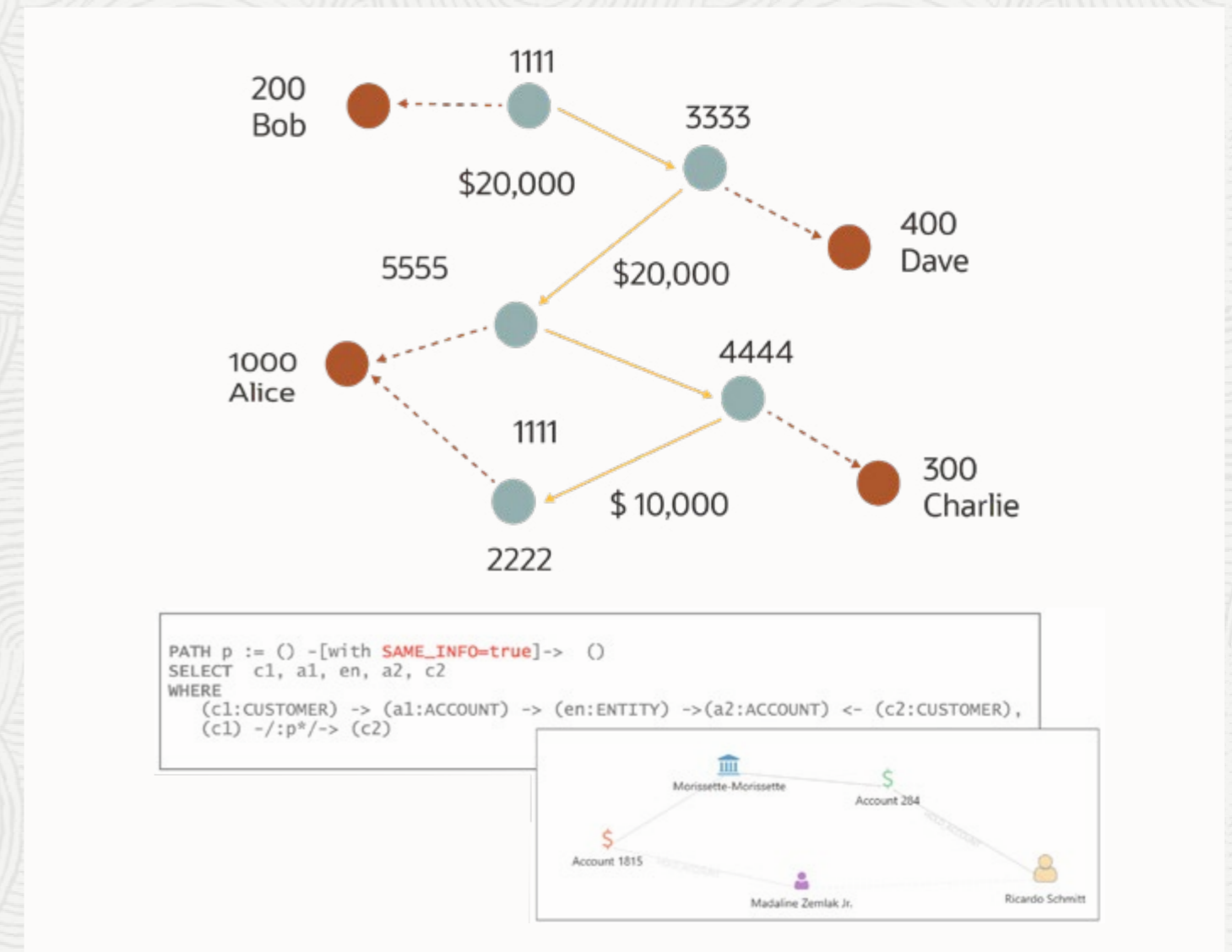
課題

概念上、マネー・ロンダリングはシンプルです。不正な資金をたらい回しにすることで、合法の資金と混ぜ合わせて、有価資産に変換します。これは、[パナマ文書の分析](#)で使用された種類のプロセスです。

より具体的には、循環する資金移転には、不正に得た多額の資金を自分自身に送金するが、“正規”の口座間における長く複雑な一連の有効な振替を介してこの資金を隠す犯罪者が関わっています。これらの“正規”の口座は、実際に合成の個人情報を使用して作成されます。これらは通常、盗まれた個人情報（電子メール・アドレス、住所など）から生成されるため、特定の類似情報を共有しています。そして、このような関連情報により、グラフ分析が不正の起源を明らかにする上で最適なツールになるのです。

グラフ・ソリューション

不正をより簡単に検出するために、ユーザーは、エンティティ間の取引、情報（電子メール・アドレス、パスワード、住所など）を共有するエンティティ、その他からグラフを作成できます。グラフが作成されたら、単純なクエリーを実行し、類似情報が含まれる口座を持つすべての顧客を見つけて、どの口座が互いに送金し合っているかを明らかにします。



マネー・ミールとミール詐欺の検出

課題

ミール詐欺には、違法な物品を転送するマネー・ミールと呼ばれる人物が関わります。これには麻薬が関わる場合がありますが、金融業界に関しては通常、資金が関わります。マネー・ミールは、自分自身の口座に資金を転送します。この資金はその後、通常は別の国にいる別の詐欺師に転送されます。

従来、ルールベースのモデルによってアラートが作成され、疑わしい口座には人間によってフラグが付けられます。また、人間の決定を予測するために機械学習も使用されます。

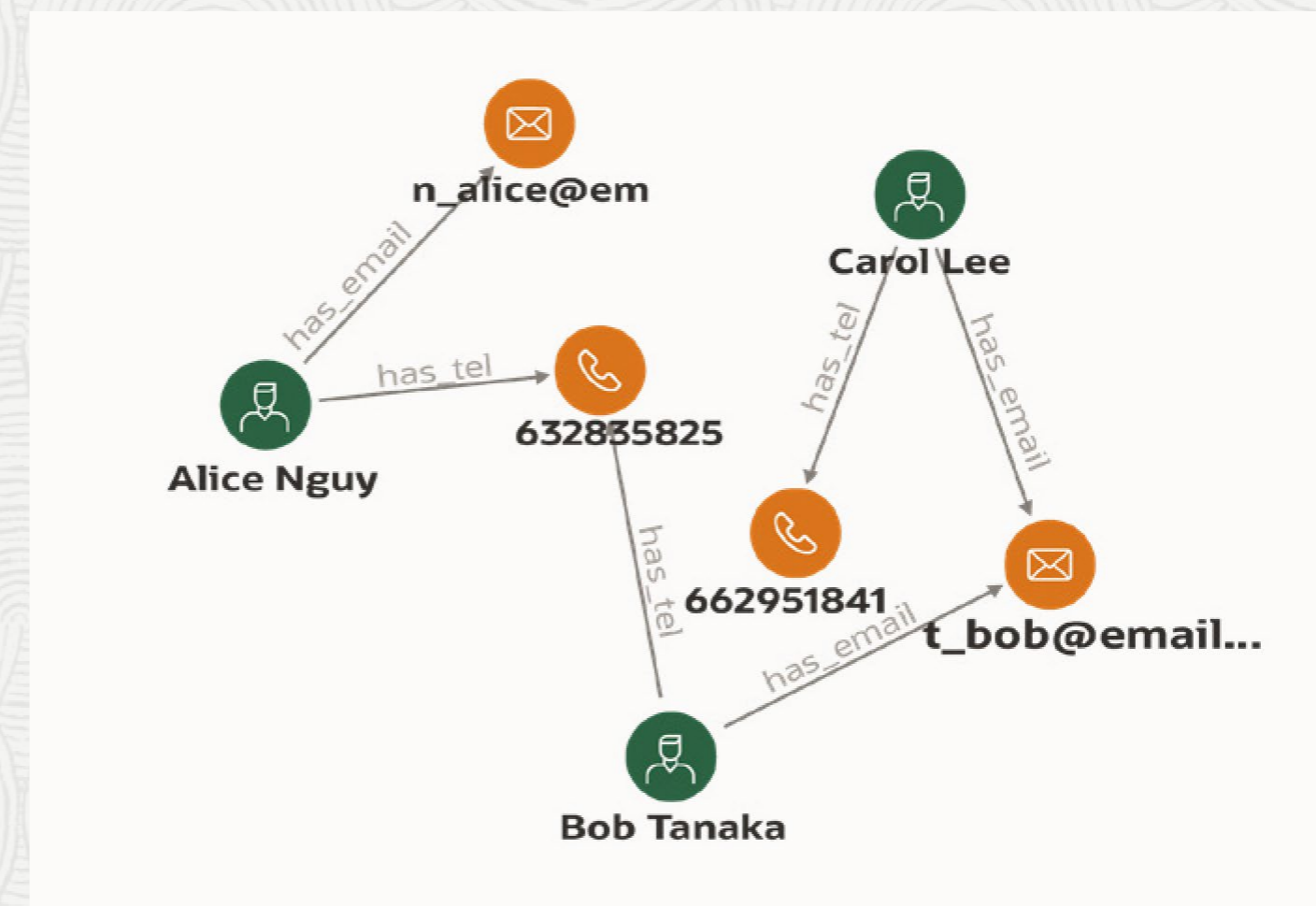
ただし、多くの場合、口座自体に含まれる情報が通常は制限されているため、モデルを改善するのは困難です。

グラフ・ソリューション

ここでグラフが必要になってきます。グラフ・テクノロジーを使用すると、ユーザーは、取引情報をエッジとして取得し、周辺の関係と取引に基づいて口座の特徴をより多く生成できます。たとえば、グラフベースの中心性スコアを使用することで、特定の口座が既知のミール口座にどれくらい近いかを確認できます。

また、これらの偽口座はたいてい、類似情報（住所や電話番号など）を共有しています。なぜなら、口座を登録するためにこのような情報が必要であるからです。そして、犯罪者だけが、情報の取り出し元となる個人情報を非常に多く手に入れています。グラフベースのクエリーを使用することで、グラフのユーザーは、類似関係を持つ口座や、循環などのパターンと関連がある口座を素早く見つけて、これらにフラグを付けてさらに調査できます。

この手法を介して、グラフ・テクノロジーは、マネー・ミールとミール詐欺を発見するようトレーニングされた機械学習モデルを強化できます。



リアルタイムの不正検出

課題

現在の世界では、消費者はサービスや資金転送をすぐに利用することを要求しています。これが犯罪者の好機になります。たとえば、支払サービス・アプリは、資金を有効なユーザーにできるだけ速く送金しようと試みながら、資金が違法な目的のために送金されないようにしたり、遠回りのルートで送金することで実際の受取人を隠したりします。これにより、リアルタイムの不正検出が必要になります。



グラフ・ソリューション

グラフによって、クエリーに対して瞬時に回答を提供できることや、データに対するアクセスが拡張されることが理由で、リアルタイムの不正検出の分野では一般的なテクノロジーとなっています。グラフ・テクノロジーを使用して取引を調査する場合、グラフ内にモデル化できるのは取引のみではありません。

グラフは極めて柔軟です。これは、周辺の異種情報もモデル化できることを意味します。たとえば、クライアントのIPアドレス、ATMのジオロケーション、カード番号、アカウントIDをすべて頂点にすることができ、つながりをすべてエッジにすることができます。

プロパティ・グラフは多くの場合、不正検出に、特にオンライン・バンキングとATM所在地分析に使用されます。なぜなら、ユーザーがデータセットに基づいて不正を検出するルールを設計できるからです。たとえば、以下に関する検出ルールを設定できます。

- ・ 異なる場所に登録されている複数のカードを使用してログインするIP
- ・ 非常に距離が離れた異なる場所で使用されるカード
- ・ さまざまな場所に登録されている他の口座から1回限りのインバウンド取引を受け取る口座

オラクルのグラフ・テクノロジーでは以下が可能であるため、これらのルールはリアルタイムで適用できます。

- ・ グラフを最新の状態に維持し、元のリレーショナル表データセットとの同期を維持する
- ・ 高パフォーマンスのクエリーとアルゴリズムを実行する

製造

製造に関しては、関係と依存関係がすべてです。そのため、より多くの情報を素早く見つける上でグラフ・テクノロジーが最適です。

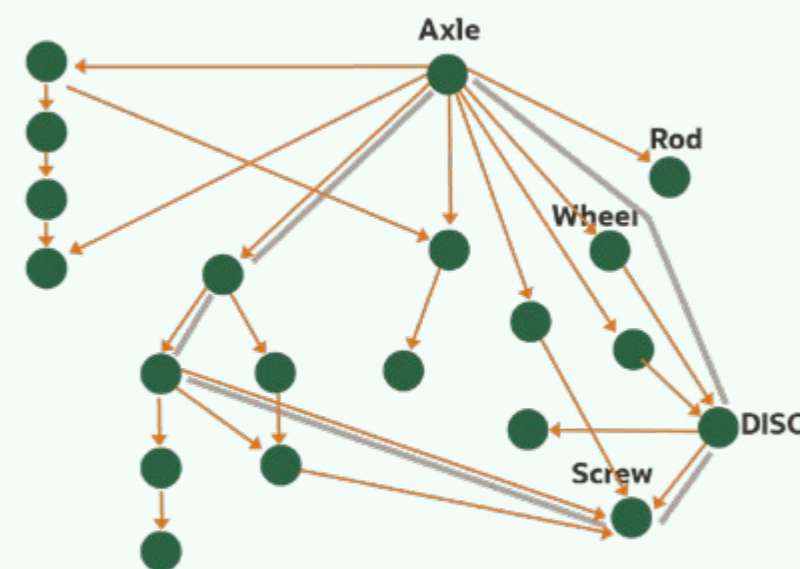
部品構成表

課題

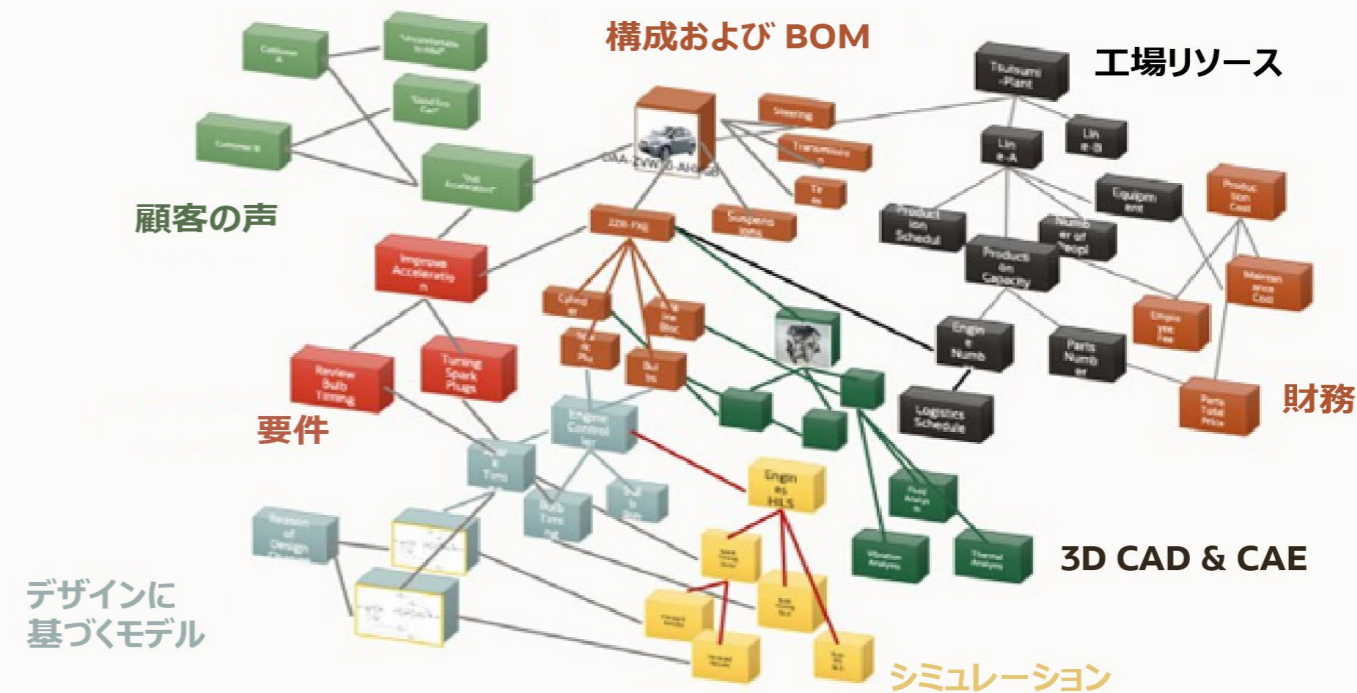
1台の自動車には30,000個の部品があります。それでは、1つの部品を変更するとどのような影響があるでしょうか。いくつかの部品を一度に変更するとどうなるでしょうか。自動車では、部品ごとに数千の依存関係がある可能性があるため、この種の分析は非常に複雑になる場合があります。以前はこのような分析のクエリでは、必要な多段階の表結合すべてが原因で、相当量の時間がかかっていました。

グラフ・ソリューション

グラフ・クエリを使用することで、応答時間を数秒またはそれよりさらに短い時間に短縮できます。これは、リアルタイムの対話型分析が現実のものとなったことを意味します。グラフでは、すべての部品が相互に持つ関係が取り込まれ、これらが明確になるため、欠陥や負の依存関係も明確になります。



部品構成表の分析にグラフを使用することで、製品情報と依存関係を分析するためのモデルを作成できます。また、ベンダー、エンジニア、サプライヤ、材料、材料寿命などの製品に関する詳細情報を追加し、多種多様なモデルを作成することもできます。これにより、部品の信頼性、ベンダーの信頼性、サプライヤのオプションなどを調査できます。



トレーサビリティ

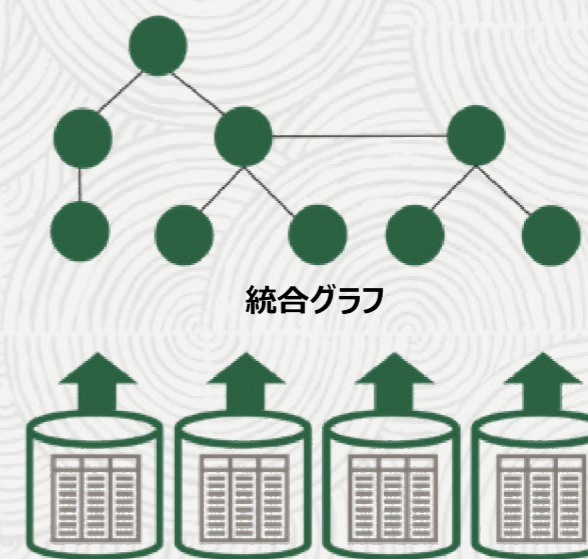
課題

トレーサビリティは、製造の世界では非常に重要です。自動車会社は、限定された時間枠内に工場で製造された部品が特定の自動車モデルに使用されていることが原因で、このモデルに対してリコールを実施することが必要な場合があります。同社は、原因となった部品を追跡してから、市場に出回っているか工場から配送中である自動車を見つける必要があります。この作業は非常に困難である場合があります。

大半の会社には、製品に関するロット情報を管理する本番データベースがあります。しかし、これらの会社には、別の小売データベース、別の販売データベース、および別の出荷データベースもあります。すべての関連情報を探し出して、問題のある自動車、出荷された場所、および販売先を見つけるのは複雑な作業です。

グラフ・ソリューション

グラフ・テクノロジーがなければ、アナリストは、これらすべてのデータベースを組み合わせ、特定の1台の自動車から、製造ラインを管理している工場データベースまでの横断クエリを実行する必要があります。これらすべてに、複雑なデータ・モデル化と多数の結合が必要です。ただし、同社に、すべての関係とグラフ・アルゴリズムを結び付けてつながりと関連情報を強調するためのグラフ・データベースがある場合は別です。



- 部品特定テーブル
- 部品関係テーブル
- 本番、販売、インベントリ...

マスター・データ管理

課題

多くの製造工場で、設計チームが部品に特定の名前を使用している場合があります。製造部門では別の名前を使用している場合があります。そして、その他の部門でも別の名前を使用している場合があります。これらはすべて、同じ品目に対してです。問題が発生した場合、または会社が特定のユースケースに関する詳細情報や、この特定の品目に関与している部品を見つけることを望んでいる場合、一貫性のないこれらすべての異なる名前が原因で、適切な品目と照合し、問題の部品を見つけ出すのが困難になります。

グラフ・ソリューション

RDFグラフは、さまざまな部品をモデル化し、これらの部品が相互に持つ関係とつながりを利用するのに適しています。RDFグラフは、これらすべての情報を使用して、異なる名前が同じ品目を示しているかどうか、複数の品目が関連しているかどうか、さらにはその類似性のために異なる品目を交換して使用できることを示しているかどうかを確認する上で役に立つメタデータ・レイヤーを作成します。これは、製薬業界でも、さまざまな化学薬品、医薬品、および一般名を識別するために使用されます。

RDFグラフがない場合、アプリケーションは通常、適切な品目を見つけるのに役立つロジックを埋め込みます。しかし、多くの場合、各データベースには異なるネーミング規則が採用されているため、このロジックがすべてのデータベースで機能するわけではありません。そして、このアプリケーション・ロジックを作成したDBAが会社を去った場合、このロジックはたいてい失われます。

RDFグラフを使用してこの情報をメタデータ・レイヤーに抽出すると、アプリケーション・ロジックに対するこの依存が解消されるだけでなく、別の有用なレイヤーも作成されます。RDFグラフは、異なる名前が同じ製造品目を指しているかどうかを伝えてくれるだけでなく、これらの品目が他の品目と相互に持つ関係と依存関係も明らかにしてくれます。

これにより、その他の関連品目を見つけて、暗黙的な事実と関係を発見することが容易になります。

要するに、RDFグラフを使用すると、コンテンツとセマンティックを機械で読み取れる有用な方法で捕捉できる自己記述的データを入手する手段を得ることができます。また、アプリケーション・ロジックを最新の状態に維持することも必要なくなります。RDFのコンテンツと品質が向上するにつれて、アプリケーションも自動的に向上します。



政府機関

犯罪行為から接触追跡まで、グラフ・テクノロジーを使用して政府関連の多くの問題に対処できます。

税金詐欺

課題

税金詐欺は、多くの政府にとって深刻さを増す問題です。犯罪者はより巧妙になっているにもかかわらず、政府は多くの場合、リソースがいつそう困窮しています。それだけではなく、現代のテクノロジーは、俊敏性に乏しい政府に新しい問題を突きつけている一方、国境線を越えて資金を移動する簡単な方法を提供しているため、犯罪者の動機をさらに高めています。

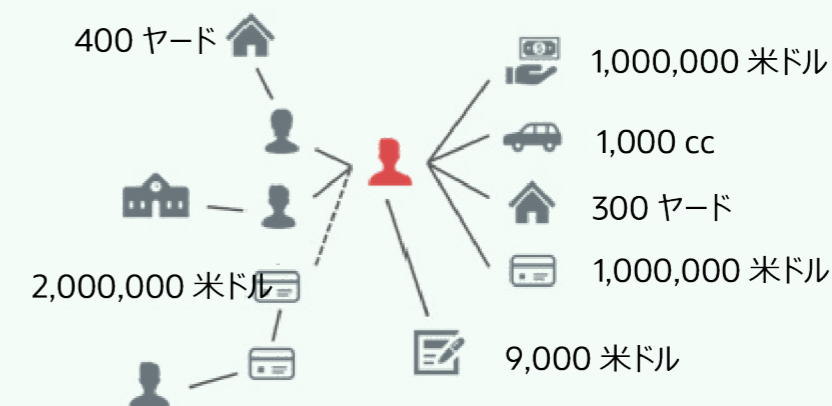
ここで、犯罪者はペーパー・カンパニーを設立し、この会社が正規の法人のように見えるようにすることができます。資金は複数の口座を介して前後に、まんべんなく、意図的に混乱させる遠回りのパスで移動し、最終的には政府の資金が犯罪者の手に渡ります。

グラフ・ソリューション

複数の関係のレイヤーがデータ内に深く隠された、これらの複雑なパスのもつれを解くのは簡単な作業ではありません。関係の各レイヤーを経由してこのパスを追跡するのは困難なタスクですが、グラフ・データベースを使用すると、ペーパー・カンパニー法人の構造を理解し、手動の調査の役に立つ可視化ツールを提供し、複数のホップ内の疑わしいパターンを発見し、曲がりくねっているが最終的には不正な1人の人物または1つの組織へとたどり着くパスを発見できます。

別の税金詐欺のユースケースでは、グラフ・テクノロジーは、人々が隠そうとしていた資産や賃金を明らかにすることもできます。たとえば、個人が複数の企業から賃金を受け取り、その一部を隠そうとする場合があります。または、この人物が、開示されていない別の投資資産を所有している場合もあります。そして、賃貸不動産、使用料、共同事業、地所、信託財産を含む複数のソースからの収入があるときに、特にこれらの資産の所有権に複数の人々が関わっている場合、これらすべてを追跡し、適切な税金が支払われるよう徹底することが困難である場合があります。

グラフ・テクノロジーは、これらの資産とそれに関わっている人々を整理して提示し、これらの関係（および支払うべき金額）をより明確にすることができます。



犯罪捜査

課題

グラフ・データベースは、犯罪行為の分析に大変革をもたらしています。一部の犯罪は小規模で日和見主義的に行われます。しかし、警察が協力して追跡して捜し出し、倒すような種類の犯罪は、人々、グループ、企業、さらには場所の多くが相互につながって大規模に行われる傾向があります。すなわち、このような犯罪はサイロ化せずに行われる傾向があることを意味します。

グラフ・ソリューション

データをグラフに取り込むと、犯罪ネットワークを特定し、パターンを探し出す自然で効率的な手段を入手できます。PageRankや中心性などのグラフベースのアルゴリズムを適用すると、グラフ内の脆弱な人々を見つけ出し、場所に関するより多くのインサイトを発見し、さらには重要な人々や潜在的な犯罪グループを探ることがより容易になります。たとえば、媒介中心性を適用すると、“もっとも弱いリンク”、すなわちグラフが依存する頂点を見つけることができます。この頂点を削除すると、グラフ全体がばらばらに壊れる場合があります。これは、まさに犯罪グループの急所を見つけた可能性があることを意味します。



接触追跡

課題

病気の接触追跡は、世界的に重要なアクティビティです。人々は非常に感染性の高い新しい病気にかかっても、普段の生活を続けて、映画館、満員のジム、混雑した結婚式、聖歌隊の練習に足を運ぶことで、訪れるあらゆる場所にその病気を拡散します。

誰かが病気と診断されると、この病気の人物と接触した人全員を見つけて隔離を求める戦いは、時間との戦いになります。接触追跡者は、病気のさらなる拡散を止めるためにできるだけ迅速に作業を行う必要があります。



グラフ・ソリューション

関係に重点を置いたグラフ・データベースは、病気パターンを分析するために使用するのに理想的です。アナリストは、検査の結果として病気であることが判明した人々、これらの人々が接触した家族や友達、訪れた場所に関する情報を入力し、ホットスポットとつながりを素早く特定できます。この方法により、アナリストは、病気になった人々を隔離し、病気のさらなる感染爆発を防ぐために、より素早く行動できます。

グラフを使用した接触追跡には3つのレベルがあります。

第1に、人々の関係、コミュニティ、および訪れた場所を理解する必要があります。これらは、十分なモバイル・データが提供されれば、グラフを使用して明確にすることができます。

第2に、グラフでは、拡散の可能性を見つける必要があります。これは、病気を拡散する可能性がある人々の中の潜在的なつながりに注目することを意味します。この人物はバスで移動しましたか。バスに乗っていた人を全員特定できますか。

第3に、接触追跡者は、“スーパー・スプレッダー”を見つけ、急いでこれらの人々を最初に隔離する必要があります。これには、広く深く接触した人々、および多くの異なるコミュニティとつながりのある可能性が高い人々を見つけることが含まれます。これには、中心性と媒介性に注意しながらグラフを調査して、つながりの強い人々を見つけることが含まれます。

データ規則とプライバシー

データの価値が高くなるにつれ、企業はより積極的にデータを収集、販売、および利用します。同時に、データに関する法律、規制、および規格も大幅に増えます。しかし、データの量が増え続けるにつれ、そのデータを管理し、データのプライバシーと規制を確保することが、かつてないほど複雑になっています。

GDPR

課題

世界中のデータ管理の専門家は、依然としてGDPRへの対処という問題に取り組んでいます。これらの専門家はどうすれば、特に、人々のプライバシーを維持し、データ・アクセス・リクエストに対応し、忘れられる権利に関するリクエストに応え続けることができるでしょうか。

主な困難の1つは、各データベースに何が保存されているかを確認することにあります。データは移動します。データは変換されます。データはユーザーや他のプロセスによって使用される場合があります。そして、これらすべてのデータで何が起こったのかを追跡してトレースすることは極めて困難である場合があります。

しかし、問題はそれだけではありません。データは当初、表に保存されている場合があります。しかし、このデータからレポートが作成されたとします。

このレポートには、情報が含まれ、アクセス・ルールもあります。誰かが忘れられる権利を行使したい場合、データが当初どこにあったか、どこにコピーされたか、そして表とレポート内のどこで使用されたかに関する電子証跡を追跡することは、すべてが極度に複雑になります。GDPRのこの部分の要件を満たすことは、途方もない作業です。

グラフ・ソリューション

データシステムを追跡する作業は、グラフに最適です。頂点ごとにエッジをたどることで、データのライフサイクル内のさまざまなステップを追跡し、ナビゲートできます。グラフを使用すると、パスをたどり、情報が当初どこにあったか、どこにコピーされたか、どこで利用されたかを確認できるようになります。これらすべての情報がグラフに配置されるため、データの専門家は、GDPRのリクエストに応え、GDPRを遵守し続ける方法をより簡単に確認できるようになります。



データ・プライバシー

課題

組織は、データへのアクセスを制限する必要があります。たとえば、特定のパソコンが特定のファイルを開くことのみを許可したい場合があります。または、特定のチーム、部門、プロジェクトが特定のデータにアクセスできるようにしたい場合があります。アクセス権は管理が複雑です。また、どのチームがアクセス権を持つか、どのチームがこれ以上アクセスすべきではないか、どのチームがジョブを実行するためにより優れたアクセス権を必要とするかを把握することも、非常に複雑な場合があります。

多くの場合、階層構造を動的に変更できるようにするために、このデータ構造を流動的にする必要があります。しかし、これをシームレスに実行することは難しく、何が変更されているか、およびどのように変更されているかを正確に把握する作業は達成が困難です。

グラフ・ソリューション

グラフを使用すると、このような階層構造を非常に動的にすることができ、グラフ・クエリーを使用すると、データ・アクセスの変更の応答時間を短縮できます。

複雑で動的なアクセス制御のため、アプリケーションは、特定の資料のアクセス権を毎回確認する必要があります。しかし、グラフ・クエリーはネットワークを非常に効率的にたどることができるため、アプリケーションはアクセス権をリアルタイムで見つけることができます。

サイバー・セキュリティ

課題

サイバー・セキュリティは、クラウドをめぐる戦いにおいて非常に重要なトピックです。これには、無効なトラフィックの検出、サイバー脅威ハンティング、マルウェア検出などの複雑な領域が含まれます。これらのトピックに対処するためのソリューションの1つは、グラフ・テクノロジーを使用してサイバー・セキュリティを強化することです。

グラフ・ソリューション

グラフ・テクノロジーは、データ・エンティティ間のつながり（コンピュータがITネットワークを介してどのようにつながっているか）を捕捉します。このテクノロジーは、グラフからの特別な信号を利用して異常を検出します。このテクノロジーは、セキュリティ・データの対話型の視覚的な調査を可能にすることで、サイバー脅威の検出を強化できます。これにより、サイバー脅威ハンティングにとって理想的な環境が生まれます。

一例として、オラクルはSaaSセキュリティ・チームと協力し、脅威インテリジェンス・システムを強化しました。このシステムは、アプリケーションの実行を監視し、疑わしいアクティビティを検出し、インシデント・アラートを生成し、調査すべき問題を割り当てます。多くの場合、これらの問題は手動で調査されますが、それには時間がかかり、リソースが非効率に使用される可能性があります。

ただし、特定のインシデントは常にパターンに従います。グラフベースの視覚的な脅威調査を追加し、情報パケットがどこから来たのか、それらがどのように転送されたかというパターンを追跡することで、これらの一貫したパターンを自動的に特定して停止できるため、時間と従業員リソースを節約できます。

マーケティング

マーケティングに関しては関係がすべてであり、マーケティング担当者は、顧客、顧客の相互関係、製品、さまざまな製品間の関係、その他を理解し、顧客が求めるものを効果的に提供する必要があります。

顧客の全方位分析

課題

現在、企業が手にしている、顧客に関する情報はますます増えており、これには、以下が含まれます。

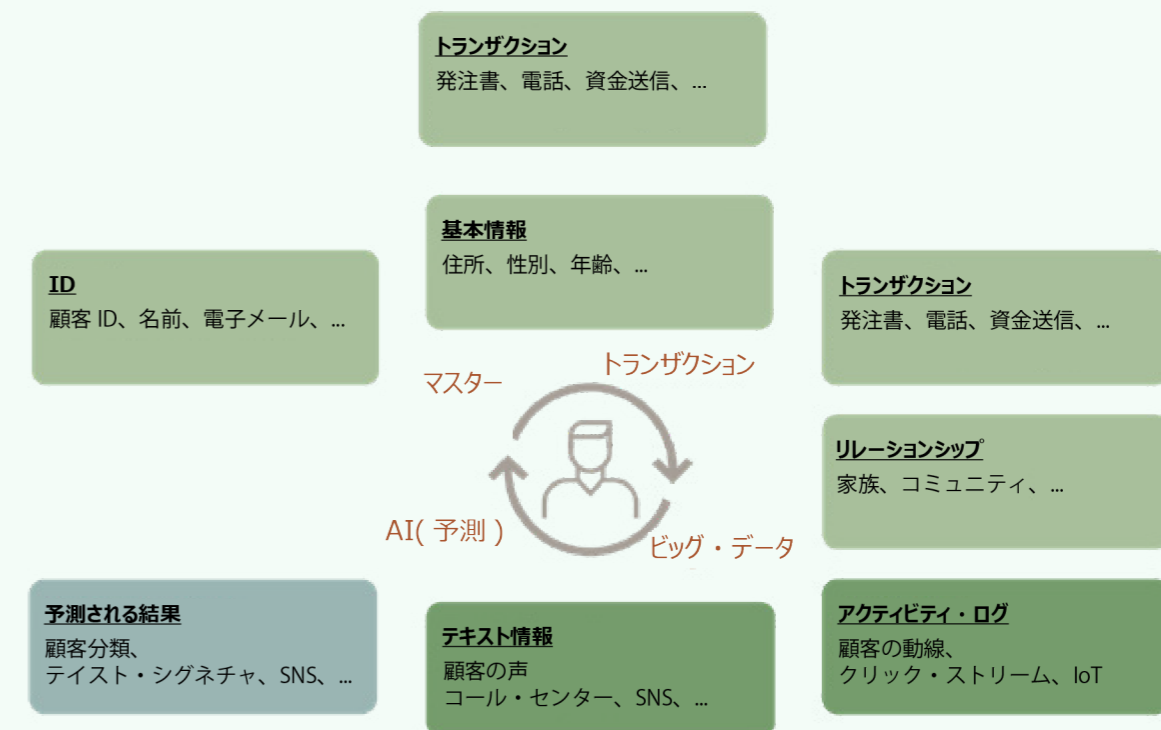
- ・ マスター・データ：名前、年齢、性別、住所
- ・ 取引：購入品、購入した品目のタイプ、購入時間
- ・ ビッグ・データ：コール・センターのログ、動線、Webのクリック・ストリーム、SNSアクティビティ
- ・ 予測：分類、テイスト・シグネチャ
(多くの場合、さまざまなモデルによって作成)

しかし、企業はたいてい、この情報を十分に包括的には使用していません。顧客の全方位分析を適切に行うことは困難です。

グラフ・ソリューション

上記のマーケティング・データがすべて収集されて物理的プラットフォームに統合されている場合、通常はこれらすべてを分析するのは困難です。しかし、これらのデータセットはグラフに論理的に統合でき、グラフ・ユーザーは、1つのエンティティ（顧客）の周辺の情報をすべて簡単に確認できます。グラフを使用すると、マーケティング担当者は、顧客、顧客の相互関係、購入したすべての製品間の関係などをより包括的に把握できます。次に、グラフ・ユーザーはアルゴリズムを実行して、顧客に関するいっそうきめの細かい詳細さえも見つけ出すことができます。

特定の1人の顧客に関するすべての情報を確認することは、顧客を理解し、顧客の全方位分析を実行し、どの予測（通常は機械学習を介して作成）が正確で、その理由が何かを知る上で重要です。



製品の推奨事項

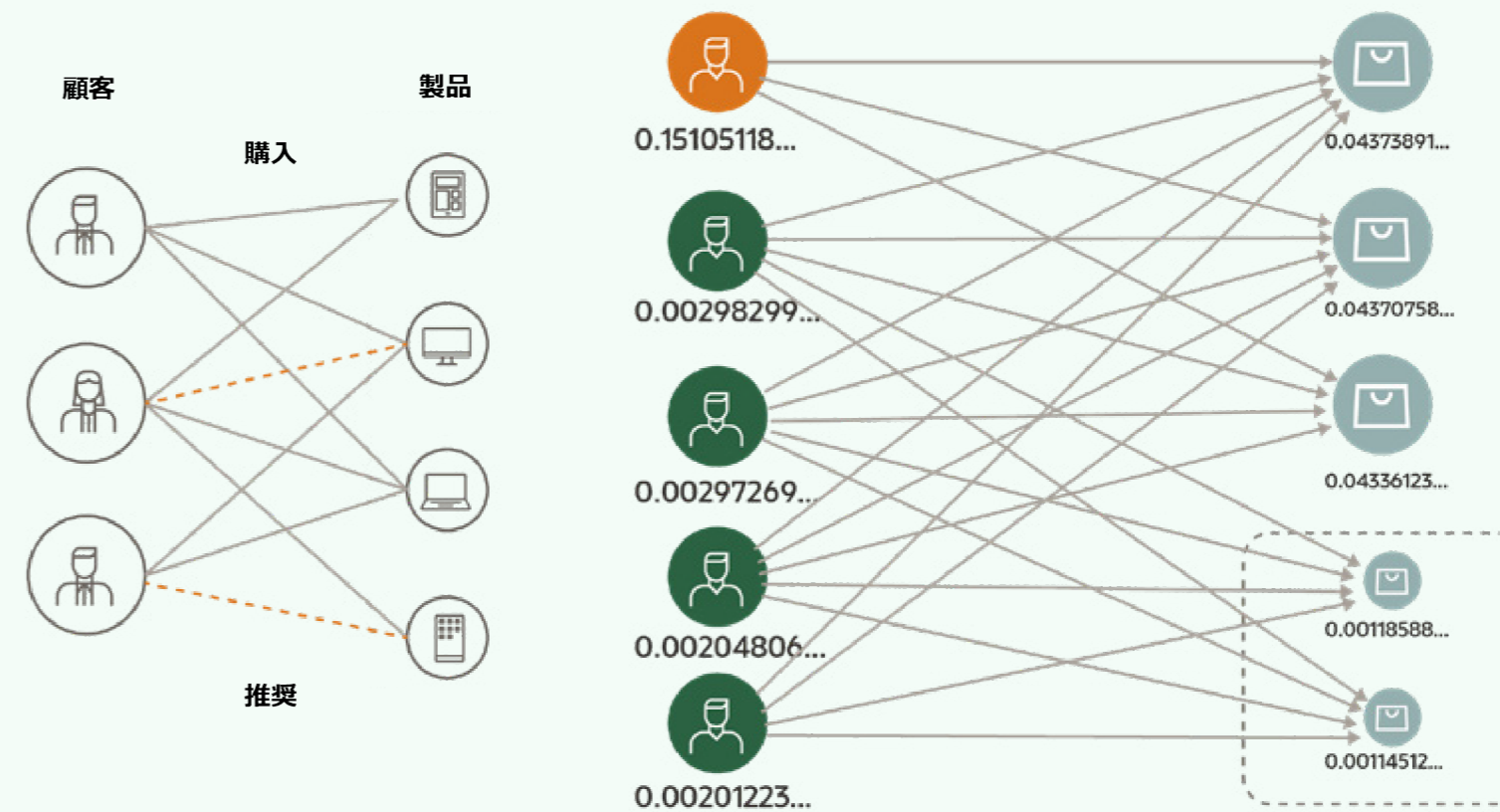
課題

グラフ以外のテクノロジーはリコメンデーション・エンジンをサポートできますが、グラフを使用すると、価値実現までの時間を短縮できます。グラフ・データベースは、顧客と顧客が購入したい製品との間の関係がすでに配置された状態で構築されるため、データを介してアルゴリズムを実行して推奨事項を簡単に素早く見つけることができるようになります。

また、リアルタイムの推奨事項がかつてないほど重要になっています。しかし、これには、製品情報、顧客インベントリ、顧客の過去の行動、現在のサプライヤ情報、物流、さらにはソーシャル・メディアを介してクリックされた広告や調べられた製品などのソーシャル・データを相互に関連付ける能力が必要です。これは、特定の種類のデータベースにとっては極めて困難です。

グラフ・ソリューション

このデータをすべて収集し、つながりを構成し、顧客のニーズと製品の傾向に関する速やかなインサイトを獲得し、リアルタイムの推奨事項を提供するためのテクノロジーが、グラフ・データベースです。実際に、多くの大手企業は、グラフ分析を活用して推奨事項を提供しています。なぜなら、関係はすでに配置されており、これらの関係の分析を使用して推奨事項を非常に素早く提供できるからです。



ソーシャル・メディア

課題

ソーシャル・メディアは、現在の世界で普及し続けており、関係はその主要な一部です。これに関しては、ユーザーをつなげること、およびこれらのユーザーの妥当性を確保することも同様です。ソーシャル・メディアの世界では、“ソックパペット”アカウントが問題になっています。

ソックパペットとは、ボットによって運営される偽のアカウントです。ソックパペットは、特定のトピックやキーワードに「いいね!」をしたりこれらを共有したりして、これらが流行しているように見せかけることで、これらがより重要であるように思わせるよう動作します。

ときには、これはそれほど害を及ぼしませんが、とはいえやはり小売業者や顧客を惑わせるものです。Instagramのインフルエンサーがフォロワーや「いいね!」を購入して自分自身の人気が高いように見せかけることを考えてみてください。またあるときには、これは非常に深刻な場合があります。たとえば、国がボットを使用して他国の政府を不安定化させるためのトレンド・トピックを拡散する場合などです。

グラフ・ソリューション

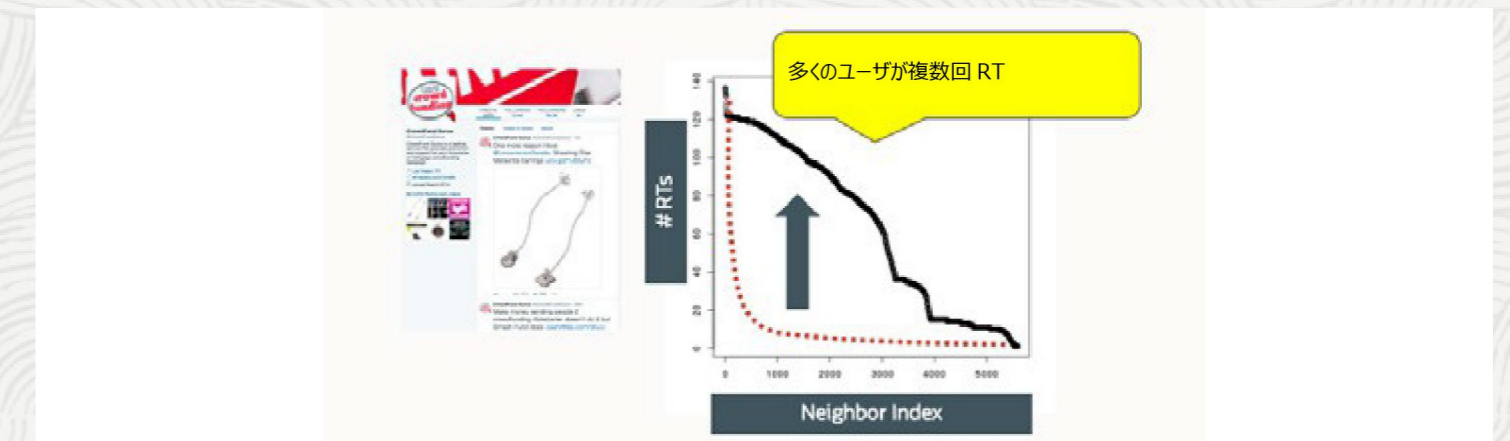
グラフ・データベースは、ソーシャル・ネットワークおよび関連するデータを非常に素早く横断できます。これが、Facebook、LinkedIn、Twitterなどのすべてのソーシャル・メディア企業がプラットフォーム内である種のグラフ処理を活用して世界中の友達や家族を特定している理由です。以前の例ですでに製品の推奨事項について言及しました。同様のプロセスを使用して、ユーザー、イメージ、製品などに関する推奨事項や、不正アクティビティやソックパペット・アカウントを検出するための推奨事項を提供することもできます。

この例では、リツイート回数をエッジの重みとしたアカウント間のグラフを作成し、これらのアカウントが隣接するアカウントを何回リツイートしているかを確認しました。表示されているパターンには、自然に人気のあるアカウントとは非常に異なる特性を示す傾向があります。

自然に人気のあるアカウントのパターンは以下のとおりです。



不自然に人気のあるアカウントのパターンは以下のとおりです。



グラフを使用すると、不自然な行動を素早く見つけ出し、ボットとソックパペット・アカウントを除去できます。

AIと機械学習の研究

AIと機械学習は、ビジネスの結果を改善し、新しい影響力を生み出すことに関する将来性のため、大きな関心の対象分野として一般に認められています。グラフを使用して、いくつかの主要な手段でデータ・サイエンスを強化できます。

特徴量エンジニアリング

課題

機械学習に関しては、機械学習モデルがデータを利用しています。このデータがより優れている（より豊富で、より深く、より完全である）ほど、（通常は）機械学習モデルもより優れています。機械学習モデルを作成するために、データの強化が含まれる特徴量エンジニアリングと呼ばれる終始一貫した手段があります。ここに簡易化した例があります。データ・サイエンティストは個人の自宅住所とオフィス住所を知っている場合がありますが、機械学習モデルにとってはマイル単位の距離がある方が望ましいと言えます。データ・サイエンティストは、特徴量エンジニアリングのこの追加手順を実行してそのマイル単位の距離を算出して、機械学習モデルにとってより優れたデータを作成する必要があります。

ただし、特にデータとの関係に注目してこれらの関係を最前線に持ち込むことに関しては、成し遂げる作業がより複雑になる可能性がある特定の種類の特徴量

エンジニアリングがあります。これを試みるには、必要な結合が多くなりすぎ、成し遂げるのが遅く煩雑な作業になる可能性があります。

グラフ・ソリューション

一般に、機械学習用の特徴量は、グラフ・データベースにロードされたデータセットに対してグラフ・アルゴリズムを実行し、後の機械学習に使用できる豊富なデータを作成することで、グラフを介して作成できます。この特徴量エンジニアリングの手順により、より包括的で有用な情報を備えた機械学習モデルを入手できます。

たとえば、機械学習モデルには、生命保険の購入を検討している新しい顧客に関する情報（この人物がどこに住み、どのような自動車を所有しているかを含む）がすでに含まれる場合があります。このモデルは、既存の顧客を分類し、この情報に基づいて、新しい顧客に関する予測を行うことができます。しかし、このモデルにはグラフのコンポーネントが不足している場合があります。この新しい顧客には、すでに顧客である同僚がいて、これが、新しい顧客が生命保険を実際に契約した理由に関する主要インジケータである場合があります。グラフから導き出された特徴を含めることで、機械学習モデルがさらに強力で正確になる可能性があります。

また、グラフ・アルゴリズムをデータに対して実行することで新たなインサイトを生成できます。たとえば、クラスタ化を使用して、顧客が購入した製品に基づいて同様の顧客を発見する、などです。

グラフ・ニューラル・ネットワーク

課題

これまでに、グラフがどのように推奨事項の役に立つかについて説明してきました。しかし、予測の推奨事項に関してはどうでしょうか。たとえば、オンライン小売店舗が、顧客が品目を使い切ると予測されるタイミングで、顧客に推奨事項を送信したいとしたら

どうでしょうか。推奨事項に予測を追加するのは複雑な作業となる場合がありますが、利益が大幅に増加する可能性があります。これはたいてい、多くの企業にとって手付かずの機会を秘めた領域です。

グラフ・ソリューション

多くのデータ・サイエンティストが、グラフ・ニューラル・ネットワークに興味を持ち始めています。このネットワークは、グラフ自体を機械学習とニューラル・ネットワークの入力として取得できます。グラフは、モデルの柔軟性のため、標準的な表より多くの情報を保持できる可能性があります。多くの場合、グラフから取得した情報を持つ機械学習モデルの方が、表形状の入力に基づく機械学習よりも優れたパフォーマンスを発揮します。

このタイプのニューラル・ネットワークはすでに複数の業界で評価されており、一部の結果では、金融詐欺の検出などで精度が向上することが示されています。このような技術を実行する場合、柔軟性を高めるために元の情報をグラフ形式で保持することが極めて重要です。このため、最先端の機械学習技術を使用するワークフローを構築する上で、グラフ・データベースが主要コンポーネントになります。



Oracleのグラフ・テクノロジーを使用する理由

現在の過度に競争の激しい世界では、すでに所有しているデータを使用する新しい方法が絶え間なく出現しています。

お客様がまだ気付いていなかった可能性がある技術であるグラフにより、分析が変化し強化されています。これは、問題を解決し、価値のあるインサイトをより速く取得するために使用できる強力なツールです。グラフ・テクノロジーがお客様にとって何をしてくれるのかを確認することで、データをより効率的に使用し、結果をより速く見つけ出し、組織が最新の情報を得た状態を確実に維持できます。

Oracleは、お客様の企業がグラフ・テクノロジーを簡単に採用できるようにします。Oracle DatabaseとOracle Autonomous Databaseにはグラフ分析エンジンが組み込まれているため、ユーザーは、グラフ・アルゴリズム、パターン一致クエリ、視覚化の能力を使用することで、データからより多くのインサイトを発見できます。

このEブックに列挙されているユースケースはすべて、実世界の例であり、複数の業界にわたってOracleですでに実装されています。これらのユースケースの多くは、[Oracleのグラフ・テクノロジー](#)とその独自のエンタープライズ・グレードの機能を使用しないと実現できません。

すべてのグラフ・データベースはそれ自体が高パフォーマンスであると主張していますが、Oracleのグラフ・サービスのみが、クエリのパフォーマンスとアルゴリズムの両方において高パフォーマンスであり、業界をリードするデータベースと密接に統合されています。これにより、開発者は簡単に、既存のアプリケーションにグラフ分析を追加し、データベースがデフォルトで備えているスケーラビリティ、整合性、リカバリ、アクセス制御、およびセキュリティを確保できます。

より詳しい情報にご興味がある場合は、今すぐオラクルのステップ・バイ・ステップ・ガイドをご覧ください。OracleのAlways-Free Tierを使用して[Oracle Autonomous Databaseでグラフ・テクノロジーをお試しください](#)。



詳細情報

Webページを見る

グラフを無償で試す

ORACLE

Copyright © 2021, Oracle and/or its affiliates. All rights reserved.
OracleおよびJavaはOracleおよびその子会社、関連会社の登録商標です。
その他の名称はそれぞれの会社の商標です。

