

ハイリスクかつ重要な プロジェクトを成功させるために

MARKLOGIC ホワイトペーパー

これまでのやり方を変える時が来ています。自分たちの組織の目標を達成するに当たって、「このプロジェクトも失敗したITプロジェクトの一つとなるだろう」というアナリストの予言が当たらないようにすべきです。すべてのデータを活用し、想定よりも短期間かつ廉価に目標を達成しましょう。ここで重要なのは、今後取り組むハイリスクかつ重要なデータプロジェクトを成功させるには、適切な技術と心構えによってまず完成形を思い描くことが必要だということです。「変化」は素晴らしいことなのです。



目次

| | |
|--|----|
| 概要..... | 1 |
| 認めたくない真実:今や、データ統合プロジェクトについての考えを改めるべき | 2 |
| 技術的なハードル | |
| 心理的な抵抗 | |
| 破滅的な状況 | |
| 既存の技術に基づく意思決定を見直すことでリスクを削減 | 4 |
| 適切な心構えを取り入れる | |
| 適切な技術を取り入れる | |
| 自分の組織に適したものを準備 | 5 |
| リーダーから学ぶ..... | 6 |
| 過去に類を見ない規模の米国全土でのデータ統合プロジェクト | |
| 統合された情報をリアルタイムで提供する拡張性のあるデータハブの構築 | |
| 安全な統合グローバルトレードストアを、6か月という短期間で開発 | |
| 大規模なダイナミックコンテンツ提供の新基準 | |
| 思考のスピードでビジネスが変化する:そのスピードで変化する準備を | 12 |
| さらに詳しく | |

「インタビューや専門家の見解によると、データサイエンティストの時間の最大80%は、データ活用に必要な事前のデータの収集と準備の退屈な作業に費やされています」

ニューヨークタイムズ、スティーブ・ロー氏

概要

大量、多様、複雑なデータは、これまでになかったスピードで拡大しています。日々生成され消費される情報の量には驚くほかありません。実際のところ、2020年には、地球上の一人当たり毎秒1.7MBの情報新規生成されます。¹ こういった世界的「情報肥満」がもたらす問題に加えて、クラウドコンピューティング、モバイルプラットフォーム、IoTの登場により、多様なデータが混ざり合い、ビジネス要件は変わり続けます。組織はこれらを効果的に把握しなくてはなりません。企業の85%が、競争力を獲得するに当たってデータを活用できていないということは、もはや驚くべきことでもないでしょう。²

複雑なエンタープライズプロジェクトは、本来的にリスクが高く、その実装がうまくいかない可能性や実際に失敗することが多々あります。最近『ComputerWorld』は、1000万ドル以上のコストがかかったプロジェクトに関する調査をwebサイトで公表しました。この調査によると、公共/民間のプロジェクト3555件のうち、成功したのはたった6.4%だけです。このうち50%以上では、問題が発生しています（予算/期日超過、ユーザーに不評など）。³

エンタープライズデータ統合プロジェクトにおいて、単なるプランニングの段階だけでも恐怖を感じるのは、あなただけではありません。間違いなくリスクは高いのですから。ガートナーによると、2018年にはほとんどの大規模システムにおいて、実装コストの50%はデータ統合に使用されます。⁴ みなさんも、今後のハイリスクかつ重要なデータプロジェクト実装に当たって、不適切な技術や心構えのせいで大惨事に突き進んでいる可能性はないでしょうか。

ITプロジェクトにおいて旧来のシステムや時代遅れの考え方で現代のシステムを構築しようとした場合、失敗に終わる可能性が高いでしょう。事実、ガートナーによる別のレポートによると、2017年のビッグデータプロジェクトの60%はパイロットや実験の段階よりも先に進むことができません。⁵ 組織行動研究の専門家たちによると、多くのITプロジェクトが失敗する大きな要因として、失敗への恐怖があります。

今後取り組むミッションクリティカルなデータ統合プロジェクトにおいて業務目標を達成するためには、どのような取り組みが必要でしょうか。今後のハイリスクかつ重要なデータプロジェクトは、イカロスのように墜落して失敗する確率が高いかもしれませんが、そのような失敗をどう回避できるでしょうか。これに対する回答を簡単に述べると、成功のためには「適切な心構え」と「適切な技術」が備わった実装チームが必要だ、ということになります。

常に流動的な現実のビジネス状況に対応できる心構えと技術があれば、データ統合プロジェクトの実装と業務目標の達成は可能です。

本ホワイトペーパーでは、失敗のリスクを削減するための知見や実用的なガイドラインを、業界のリーダーたちから学びます。彼らは重要かつ複雑なデータ統合を成功させるに当たって厳しい現実と直面しました。米国保険福祉省のOIG（監察総監室）は、Healthcare.gov（通称オバマケア）の開発について調査し、2016年2月に詳細なケーススタディを発表しました。本ホワイトペーパーは一般公開されたこのレポートに基づき、Healthcare.gov実装の複雑さとその問題解決に関して、より深い知見を提供したいと思います。

1 IDC: Digital Universe Study: Big Data, Bigger Digital Shadows and Biggest Growth in the Far East, December 2012 <http://www.whizpr.be/upload/medialab/21/company/Media_Presentation_2012_DigiUniverseFINAL1.pdf>

2 Gartner: Information Innovation: Innovation Key Initiative Overview, April 2012 <<https://www.gartner.com/doc/1991317/information-innovation-innovation-key-initiative>>

3 Computerworld: Healthcare.gov website 'didn't have a chance in hell,' October 2013 <http://www.computerworld.com.au/article/529631/healthcare_gov_website_didnt_chance_hell_/>

4 Gartner: Predicts 2013 Application Integration, 2013 <<https://www.gartner.com/doc/2238015/predicts--application-integration>>

5 Gartner: Advancing Business with Advanced Analytics <<https://www.gartner.com/user/registration/prospect?resId=3090420&srcId=1-6470978268>>

認めたくない真実：今や、データ統合プロジェクトについての考えを改めるべき

あらゆる種類の企業が、データアセットをIT戦略に適合させ、革新と成長に活用しようと頑張っています。ここで問題となるのは、データは変わり続け、増え続け、溢れかえるようになっていることです。これにより、サイロ間のデータ統合は終わりなきプロセスになってしまいます。現在のプロジェクト要件がようやく固まったと思った瞬間に、さらにデータが追加され、さらに複雑になってしまいます。データアーキテクトが一生懸命になって、エンタープライズデータを管理・活用しようとしているにも関わらず、ほとんどの組織においては、何年もかけて集めたデータの有効活用ができていません。

技術的なハードル

特定のニーズ用に作られた旧来のシステムがあちこちに導入されていることにより、大量のデータサイロが発生し、知見の獲得が困難になっています。このように分断されたサイロにエンタープライズデータが複数回コピーされて分散していることにより、データガバナンスおよび正確さの問題が発生します。またクライアントのニーズにコスト効率よく対応することは極めて困難です。

このように複雑なため、エンタープライズデータ統合プロジェクトの実装は、ハイリスクで時間とお金がかかるものとなります。組織は、この複雑さに複数の技術で対応しようとしますが、これは理想的な方法ではありません。一つの方法として、エンタープライズデータウェアハウス (EDW) があります。これは、別々の業務部門からのデータを統合し、下流の意思決定システムで利用できるようにするものです。しかし旧来の業務システムからのデータ移行に時間がかかりすぎるため、EDWがうまくいかないことが多々あります。また、通常EDWは固定的なリレーショナルデータベース構造に基づいて構築されており、この構造では、ソースシステムや下流の業務要求にあまり対応できません。

別の方法として、エンタープライズアプリケーション統合 (EAI) ならびにサービス指向アーキテクチャ (SOA) があります。これらは「オペレーショナル」なもので業務遂行をサポートします。これらのプレー



図1. 複雑なデータ統合プロジェクトはリスクが高い

ムワークは、企業内に存在するシステムやアプリケーションを統合します。しかし残念ながら、システム内には互いにコミュニケーションできないアプリケーションが数多く存在します。このため同一データが複数の場所に格納されることとなります。この結果、部分的な(「ポイントツーポイント」の)接続が乱雑に発生します。これは管理が難しいです。

最後になりますが、こういった問題克服のために、構造化/非構造化データを大規模格納するデータレイク (Hadoopなど) を試している組織もいくつかあります。こういったデータレイクでは、ガバナンスをきちんと行わずに大量のデータを読み込んだ場合、すぐに混乱を来してしまい、業務上の価値やビジネスインサイトには全く活用できません。⁶

心理的な抵抗

今日、残念ながら多くのITプロジェクトが予算や期間を守れていない、という現実があります。プロジェクトが失敗したり、大きなコストがかかるのは当たり前になっています。マッキンゼー & カンパニーによると予算が1500万ドル以上のITプロジェクトの半数は、以下のようにになっています。

- 予算を45%超過
- 期日を7%超過
- 想定されていた機能の56%を提供

6 [icrunchdatanews.com: How to Keep Your Data Lake from Becoming a Data Swamp, April 2016](https://icrunchdatanews.com/how-to-keep-your-data-lake-from-becoming-a-data-swamp)
<<https://icrunchdatanews.com/how-keep-your-data-lake-from-becoming-data-swamp/>>

「予算が1500万ドル以上のITプロジェクトの半数は、45%の予算超過、7%の遅延となります。また想定されていた機能の56%しか実現されません」

マッキンゼー&カンパニー

もっと悪いことには、失敗したITプロジェクトの17%はあまりに金銭的ダメージが大きく、会社存続の危機までもたらす可能性があるということです⁷。

これほど好ましくない数字が出ているのにも関わらず、なぜこれは他の業務と同じように改善の対象とならないのでしょうか。一つの説として、組織内の人々はその役職を問わず、新しい技術やプロセスの採用による変化を嫌うというものがあります。事実、『フォーブス』誌でのインタビューにおいて、アンドレア・サイモン（文化人類学博士、Simon Associates Management Consultants社CEO）は、経営者や従業員が「成長が阻害されているのは、自分で自分を（ビジネスにおける過去の成功の習慣や方法論によって）制約しているためだ」ということを受け入れるのは難しいと言っています。これは一見、直観に反するように見えますが、核となるのは「過去の認識が現在の見方を制約する」ということです⁸。

このように過去を振り返る傾向により、業務により適した新しいソリューションが無視されてしまいます。現代のデータ統合において、エンタープライズアーキテクトは自分が使ったことがある旧来の技術（具体的にはリレーショナルデータベース）に依存する傾向があり、新規システムの構築にはこの方法しかないと考えてしまいます。あるいは他の新しいやり方を採用しても良いと考えられるようになった場合でも、新しい技術に興奮しすぎて、よくわからないまま先に進んでしまい、必要とされる成熟した機能を見過ごしてしまう可能性もあります。

破滅的な状況

具体的な例を以下に紹介します。これはオレゴン州が独自に健康保険マーケットプレースのwebサイト「カバー・オレゴン」を約250億円をか

けて構築しようとして失敗した例です。これはリレーショナルデータベース技術によるものでした。このwebサイトの閉鎖ならびに州住民を連邦政府の健康保険マーケットプレースwebサイト（Healthcare.gov）へ移行させることが発表された際に、Obamacarefacts.orgは、「この切り替えは数か月にわたる協議の結果、アドバイザーパネルが勧告したものだ。その理由は、このwebサイトの改修には最大78億円かかるが、Healthcare.govへの切り替えは7億円程度しかかからないため」と説明しました⁹。『フォーチュン』誌によると、「オレゴン州は、健康保険加入webサイト『カバー・オレゴン』の失敗は、担当ソフトウェア会社の『仕事に不十分』だったせいだと非難」しています。これに加えて、オレゴン州はこの健康保険マーケットプレースがうまく開始できなかったのは技術的な問題のせいだとし、この技術ベンダーは詐欺を働いたと主張しています。「説明通り動かないような製品とサービスに数百億円支払う」ように、州の担当者を納得させたからだということです。この技術ベンダーはこれに反論し、このデータプロジェクトの失敗にはオレゴン州側にも過失があると主張しています¹⁰。

この訴訟合戦の結果はまだ出ていません。しかしどのような結果になるにせよ、数百億円に及ぶ住民の税金が日の目を見なかった重要なデータプロジェクトに無駄遣いされたということに関して、この計画全体に対していくつか疑問が残ります。極めて多様なデータの統合が必要なプロジェクトの実現が法律で決められた際に、なぜリレーショナルデータベース技術がこれに最適なものとして自動的に考えられたのでしょうか。また技術ベンダーに関してはどうでしょうか。オレゴン州の目標やビジョンを実現するための能力や意欲があるという理由で、このベンダーが選択されたのでしょうか。あるいは

7 McKinsey & Company: Delivering large-scale IT projects on time, on budget and on value, October 2012
<<http://www.mckinsey.com/business-functions/business-technology/our-insights/delivering-large-scale-it-projects-on-time-on-budget-and-on-value>>

8 Forbes: Why We're So Afraid of Change – and Why That Holds Businesses Back, April 2013
<<http://www.forbes.com/sites/womensmedia/2013/04/08/why-were-so-afraid-of-change-and-why-that-holds-businesses-back/#65d63d865cd0>>

9 ObamaCareFacts.com: Cover Oregon Marketplace Changes, April 25, 2015 Update
<<http://obamacarefacts.com/insurance-exchange/oregon-health-insurance-exchange/>>

10 Fortune.com: Oregon sues Oracle over failed Obamacare website, August 2014
<<http://fortune.com/2014/08/22/oregon-oracle-spat-200-million-lawsuit/>>

「みんな知っているでしょう。いつもと同じことをしていたら、いつもと同じ結果しか得られないということ」

トニー・ロビンズ

「過去にこの技術を使ったことがある」というだけの理由だったのでしょか。「なぜそうするのか」「何ができるのか」に関して詳細に検討してみたならば、組織というものの心構えは、より大きな目的の達成には全く適していないということが判明するでしょう。

どの組織も、所有データを活用して革新や市場での勝利を実現しようと考えます。しかし変化に対する根本的な恐怖や、直面する新しい現実に適した技術を認識・選択する能力がないことにより、たとえ完璧なプランであっても実行されないことがあります。このような状況を打開するには、組織の目標実現を促進し成功への準備につながる技術/パートナーを選択することで、リスクを抑制することが必要です。

既存の技術に基づく意思決定を見直すことでリスクを削減

適切な心構えを取り入れる

プロジェクトの成功に必要な人材が配置されていることを、調査・確認したことはあるでしょうか。最初のステップとして、組織ならびに想定パートナーの心構えを査定する必要があります。これはビジョン実現のために適切な人々が選ばれていることを確認するためです。適切な人々というのは、企業の「革新の精神」を喜んで受け入れ、またこれと現在・将来の目標とのバランスを取る人です。またこの際

にリスク抑制も考慮しなくてはなりません。ミッションクリティカルなデータプロジェクトの成功には、担当者ならびにパートナーの柔軟で現代的な心構えが適切な技術の採用と同等に大切だという認識を、最初から共有することが極めて重要です。

適切な技術を取り入れる

現在の問題を解決するのに過去の技術に頼ってはいないでしょうか。現代のシステム設計において、柔軟性とアジャイル性は大切です。というのもこれによってシステムや投資の将来性が確実になるからです。未来を予言することはできません。しかし未知への対応の最善のアプローチとして、要件変更にも柔軟かつ素早く対応できるシステムを作ることができます。

RDBMS

30年以上使用されてきたリレーショナルデータベース管理システム (RDBMS) は、今後すぐになくなることはないでしょう。しかし、長年使われているからと言って、リレーショナル技術に本質的に欠けているアジャイル性が補われるということにはなりません。新しくアジャイルなシステムを開発する場合、柔軟なデータ管理システムが必要です。

複数のデータソースが扱えるデータプラットフォームが必要な場合、リレーショナル技術は十分ではありません。というのもリレーショナル技術は、従来の文書処理 (給与やERPなど) を自動化するため

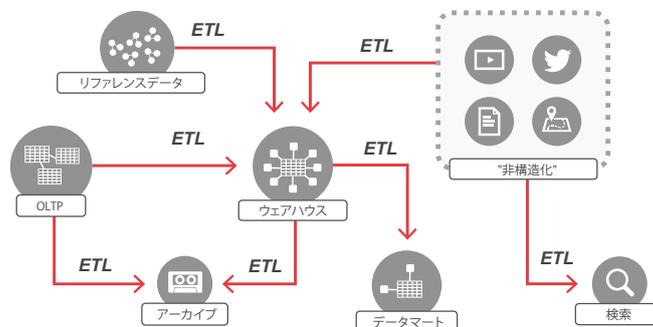


図2. ETLは複雑で時間がかかる

に生み出されたものだからです。しかもはや、処理の自動化は目標ではありません。今日の組織は、複数のデータソースを組み合わせることで素早く変更できるモバイルアプリケーションをクラウドベースで作成できる能力を求めています。RDBMSの基礎である固定的なスキーマでは、素早く変化し続ける今日のニーズに対応できません。固定的なデータモデルならびに変化し続ける下流システムにより、ビジネス要件とプロジェクトのアーキテクチャにはすぐに齟齬が発生します。

オープンソース

お金をかなりかけたけれども実稼働にまで至らないITプロジェクトが当たり前の時代において、コストにうるさい意思決定者であれば、オープンソース技術は魅力的に見えるでしょう。リレーショナル技術とは対照的に、アジャイルなNoSQL(Not Only SQL)データベースは、柔軟なデータ管理ならびに大量データに対するコスト効率に優れた拡張性を提供します。さまざまものがあり、新しいものが何十種類もあります。ここでの問題は、新しいオープンソース技術は、ミッションクリティカルな(基幹業務用の)アプリケーションに必要なエンタープライズ機能(政府レベルのセキュリティ、ACIDトランザクション、HA/DRなど)がないため、組織の業務のサポートや保護ができない、ということです。エンタープライズレベルの保護は極めて重要です。例えばツイートが見つからなかったり、Facebookの投稿の更新ができなくてもそんなに気にならないかもしれませんが、銀行の残高が間違っていたり、株取引の機会を損失したりしたら大問題になるでしょう。このためオープンソースのNoSQLを選択す

るに当たって、こういった根本的なエンタープライズ機能をどのように準備するのかを考慮する必要があります。オープンソース技術を組み合わせる場合、開発費が高くなってしまいます。またできたインフラは扱いにくく信頼性が低いものとなり、今後長い間変化し続けるビジネス要件に応じて変化していくこともできません。このような理由により、オープンソースには注意が必要です。初期費用としてのライセンス料が節約できると考えるかもしれませんが、実はその分が新規開発(とその後の改修)費用に付け替えられているだけなのです。

自分の組織に適したものを準備

固定的なRDBMSやハイリスクのNoSQLとは対照的に、適切なデータベース技術であれば、データサイロの解消、分析と業務の融合、顧客対応時の知見活用などが実現されるはずで、またこれは、複数の種類のデータを扱うことができ、複数の処理、コモディティハードウェアやクラウドによる拡張など、今日の企業の要件に応えることができます。その際に業務を止めることはありません。エンタープライズNoSQLデータベースであるMarkLogic®は、これらすべてを実現し、またデータの完全性と信頼性を保証します。これによりデータプロジェクトに柔軟に取り組むことができ、技術的なマイルストーンにもより容易に到達できます。

組織全体に存在するデータをハーモナイズする柔軟な統合プラットフォームであるMarkLogicは、既存のインフラにスムーズに統合できるように設計されています。このアジャイルなエンタープライズ

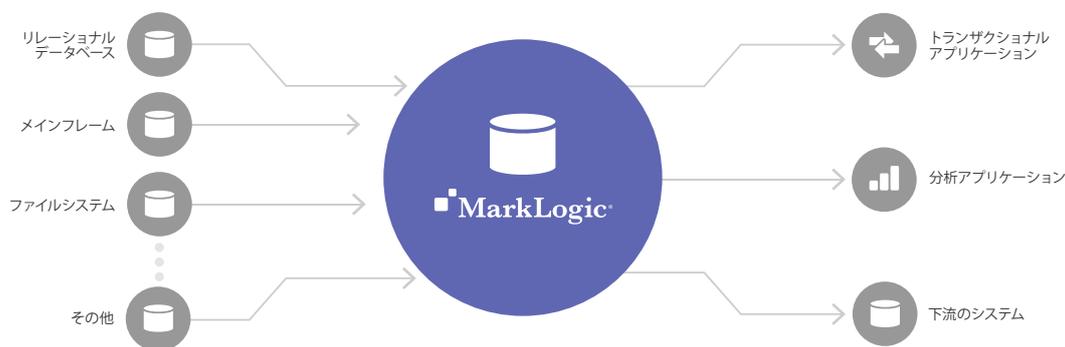


図3. エンタープライズNoSQLによる組織内データのハーモナイズ

NoSQLによって、段階的な実装が可能です。事前にすべてデータソースを特定し、共通のデータモデルを作成する必要はありません。データ統合作業を強化し単純化するために作られたMarkLogicは「データセントリック」で、一つの業務に特化したものではありません。このため同一データを数多くの用途で活用できます。

エンタープライズ用に特化して開発されたMarkLogicは、RDBMSならびにNoSQLが得意とする機能を実装し、企業の要件に応えます。政府レベルのセキュリティ、ACIDトランザクション、高可用性、災害対策、アジャイル性、拡張性、パフォーマンスが備わっています。

これは「すべての」データから価値を引き出す、最高のデータベースプラットフォームです。データの型や、データソース、サイロを問いません。MarkLogicには21世紀のデータ問題を解決する現代的な機能が備わっています。

リーダーから学ぶ

MarkLogicの技術ならびに専門知識は世界中の組織で活用され、これまでに極めて複雑なデータ統合問題を解決してきています。この新世代のデータベース技術によって実装されたハイリスクかつ重要なデータプロジェクトのうち、最近のものをいくつかご紹介いたします。

過去に類を見ない規模の米国全土でのデータ統合プロジェクト

健康保険マーケットプレイス「カバー・オレゴン」が技術的に失敗したのとは際立った対比を見せるのが、Healthcare.govです。これは最初から素晴らしいスタートを見せ、初年度申し込み期間14週間で710万人の米国民に対して健康保険を提供しました。¹¹ Healthcare.govは、最初に適切な技術を選んでは、巨大なプロジェクトの大失敗を回避できるということを、これまでになかったようなやり方で示してくれます。

カスタマー

CMS (Centers for Medicare and Medicaid Services)

背景

過去に類を見ない規模のこのプロジェクトは、新規あるいは変更されるポリシーに対応する必要があり、また法律で決まった施行日に間に合わせることはほぼ不可能に思われました。リスクはかなり高かったのです。ヘンリー・チャオ氏 (CMSの副CIO兼ディレクター、Office of Information Services) は、最初のプロジェクト実施時の技術リーダーでした。彼はプレッシャーがかなりきつかったと述べています。

「ITプロジェクトとしてのチャレンジがいろいろありました。同時接続ユーザー数万人、可用性99.9%、自分では管理できない既存の巨大システム20以上と接続・統合するシステムを作成しなければならなかったのです。数百あるいは数千人の関係者がひっきりなしに意見を言ってくるなかで、このシステムを提供しなくてはならないのです。またこれらの関係者には、最終的な意思決定者がいないのです。このような状況のため、開発中でもシステム要件はずっと変わり続けていました。このシステムの完成日程は動かすことができませんでした。失敗することをみんなが見守っているような状況でした。」¹²

組織の目標

ACA (Affordable Care Act) 法によって決められた、全国民に対する健康保険マーケットプレイスの提供という連邦政府の方針に従い、CMSは各州、さまざまな国家機関、保険会社からのデータを統合し、シームレスなオンライン申し込みサイトを提供する必要がありました。これは数百万人という米国民に健康保険を提供することになります。

課題

CMSではあらゆるシステムの構築、管理、保守を大手のプライムコントラクターに依頼します。これに加えて、CMS内にも極めてスキルが高い従業員やリーダーがおり、長年にわたってアメリカ国民全員に関係する巨大なシステムを問題なく導入してきました。法律により導入を命じられたシステムではすべてリスクが極めて高く、リーダーならびに職員の一人一人がアメリカ国民に対する自分の責任を認識しています。この任務において大事なものはリスクのバランスです。

11 Washington Post: At least 7.1M signed up for 2015 Obamacare plans so far, December 2014
<<https://www.washingtonpost.com/news/wonk/wp/2014/12/30/at-least-7-1m-signed-up-for-2015-obamacare-plans-so-far/>>

12 InformationWeek: Healthcare.gov: Hard-Earned Lessons for CIOs, March 2016
<http://www.informationweek.com/strategic-cio/healthcaregov-hard-earned-lessons-for-cios/d/d-id/1324456?page_number=1>

「HealthCare.govはただのwebサイトではありません。そんなに単純なものではありません。Healthcare.govは、申し込み資格確認と判断を行う複雑なアプリケーションで、データサービスハブと統合されています。これは、各機関からのあらゆるリクエストや回答を扱うブローカーの機能を果たします」

InformationWeek、ヘンリー・チャオ氏

リスクのバランスを取るためのリーズナブルな取り組み戦略としては、過去にうまくいった実績ある技術やプロセスを使用することが考えられます。この組織にはこういった文化(考え方)があることを理解していたプライムコントラクターは、入札の際に、マーケットプレイスならびにDSH(データサービスハブ)で管理する必要があるすべての分断データを、従来の技術(RDBMSなど)で扱うことを提案しました。一方CMSは、将来何が起こるのかを理解しており、リレーショナル技術や従来の開発プロセスではうまくいかなくなると考えていました。第一に、リレーショナル技術では、州、保険会社、他の政府機関からの多様なデータを扱えません。また、このデータベースシステムでは、サービス開始によりユーザーがこのwebサイトを利用し始めた後は、素早く拡張できません。全体的に

考慮した結果、こういった問題点により、既存のリレーショナルデータベース上のアプリケーションではACA(Affordable Care Act)の法的な要件を満たすというCMSのミッションを十分にサポートできません。

CMSならびにヘンリー・チャオ氏(副CIO)は、他に選択肢がなかったため、とりあえずこのプライムコントラクターと契約を結びましたが、その後、別の技術を使うようにと指示を出しました。彼らは2012年1月の技術指示書の中で、データベースプラットフォームとしてMarkLogicを採用するように指定しています。

米国保険福祉省の監察総監室(OIG)によるHealthCare.govに関するケーススタディに、この決定のことが書いてあります。「旧来のものとは異なる

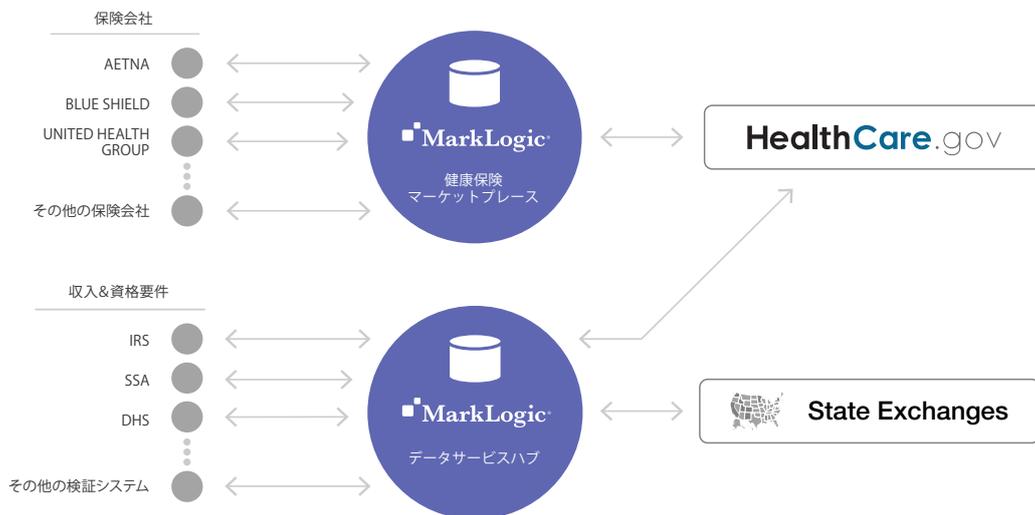


図4. 何百万人というユーザーが新規ヘルスケアプランに申し込める技術プラットフォーム

るNoSQLプラットフォームの大きなメリットとしては、その機能がこれまでのものよりも優れている可能性があります。これにより一度に移動できるデータの量が増え、データやユーザーの増加に応じて容易に拡張できます。」¹³

CMSの要件を満たすことができないRDBMSに代わってMarkLogicが選択されましたが、これは、この規模の複雑なプロジェクト実装に必要な柔軟性やパフォーマンスを提供できるからです。しかし、新しい技術ということで、確固たる実績のある政府系ベンダーからは抵抗がありました。具体的には、NoSQLデータベースの経験がないプライムコントラクターの一つは、期日内の導入に関してMarkLogicがリスクとなる可能性がある、と誤解していました。「従来のデータベースのスキルを持つ人と同じ人数の、MarkLogicのNoSQLプラットフォームのスキルを持つ人を準備することは不可能だ」と報告し、CMSに対して代替案を準備するように推奨しました。¹⁴

しかしCMSは予定通り、MarkLogicを採用しました。最終的には、この新しいデータベース技術にさらに懐疑的だった他の主要コントラクターが、データサービスハブの実装ならびにHealthCare.govの開始に関して、MarkLogicのメリットをすぐに享受できました。

このCMSの経験は、変化することへの恐怖によってもたらされるリスクについて考える良い機会となります。この例では、CMSのコントラクターの一社が、新規技術の採用をリスクと考えました。幸いにして、ヘンリー・チャオ氏は従来のやり方に捕らわれない意思決定を下しました。これは組織の要件や期日を順守できないことの方がリスクが大きく、従来の技術ではうまくいかないことを理解していたからです。これは賢い判断だったということが証明されました。

- 柔軟なMarkLogicプラットフォームは、マーケットプレイスとDSHの両方において、データソースや要件が常に変更されるという状況の中でも、分断データの統合という複雑な作業を行うことができました

- データセンターにおいてネットワーク障害が発生しましたが、MarkLogicデータベースは問題なく機能しました。CMSを守り、データロス是一件も発生しませんでした
- DSHは、新技術に懐疑的なコントラクターによって主導されましたが、サービス開始後、問題なく稼働しています
- ユーザーが直接利用するマーケットプレイスアプリケーションでは、初期の段階で障害が発生しましたが、CMSはシステムの問題を素早く解決しました。ヘンリー・チャオ氏は「うまくいかないときには、変更できるオプションがありました。うまくコーディングされていないアプリケーションを拡張することができたのです」「(MarkLogicは)他の技術では不可能だったオプションをもたらしてくれました」と言っています。¹⁵

主な結果

CMSによると、2016年の健康保険申し込み期間には、HealthCare.govのwebサイトから127万人以上が保険の申し込みを行っています。¹⁶ MarkLogicを採用したことで、CMSは以下のことができるようになりました。

- ACAによって規定された全国規模の健康保険マーケットプレイスを期日までに提供できた
- 拡張性に優れたオペレーショナルデータハブを提供。100%のデータの完全性。同時接続ユーザー28万人をサポート
- 多くのシステムから多様な構造化/非構造化データを安全に統合。シームレスなユーザー体験。毎秒6500トランザクション
- 数百万人の米国民に健康保険を拡大

13 U.S. Department of Health and Human Services: Healthcare.gov Case Study of CMS Management of Federal Marketplace, February 2016 <<http://oig.hhs.gov/oei/reports/oei-06-14-00350.asp>>

14 U.S. Department of Health and Human Services: Healthcare Gov. Case Study of CMS Management of Federal Marketplace, February 2016 <<http://Oig.Hhs.Gov/Oei/Reports/Oei-06-14-00350.Asp>>

15 MarkLogic: The Untold Story of Rescuing Healthcare.gov, December 2015 <http://www.marklogic.com/resources/the-untold-story-of-rescuing-healthcare-gov/resource_download/whitepapers/>

16 CMS.gov. Health Insurance Marketplace Open Enrollment Snapshot – Week 13, February 2016 <<https://www.cms.gov/newsroom/mediareleasedatabase/Fact-sheets/2016-Fact-sheets-items/2016-02-04.html>>

統合された情報をリアルタイムで提供する 拡張性のあるデータハブの構築

カスタマー

エトナ (Aetna)

背景

エトナは、最大規模の健康保険会社の1つで、100万社以上の保険業者から2300万人に保険を提供しています。請求を毎年数十億件処理するだけでなく、自社従業員3万5000人の人事データを管理しなければなりません。これに加えて、2015年にHumanaを買収したことで、新規データならびに要件がさらに大量に追加されました。

組織の目標

エトナのIT部門には「すべての人事データ(従業員データ、給与データなど多くのシステム)を1か所に統合して下流のシステムに渡し、リアルタイムアプリケーションやバッチ分析で使用できるように」というプレッシャーがビジネス部門からかけられていました。ここでの目標は、拡張性のある人事データハブを構築し、統合された人事情報をリアルタイムで提供することでした。

課題

エトナの既存データのほとんどは、従来のリレーショナルデータベースモデルに格納されていました。こういったデータ統合では通常、複雑なポイントツー

ポイントのデータエクステンジソリューションの開発が必要です。その際には、極めて複雑なETLプロセスを事前に把握しておく必要もあります。しかしこのアプローチではお金と時間がかかるだけでなく、リレーショナルデータベースの持つ厳格な制約のせいで、将来何か変更があった場合に大規模な改修が必要な扱いにくいものとなります。

主な結果

MarkLogicを採用したことで、エトナは以下のことができるようになりました。

- 約140のデータフィードから構成される24個の人事データソースの統合に成功。このシステムでは合計数TBのデータを扱い、1日当たりのデータのスループットは50GB
- 柔軟なデータモデルによりアジャイル開発を実現。複雑なETL処理の問題を解決
- 重要なエンタープライズ要件(ACIDトランザクション、セキュリティ、監査対応、透明性など)を犠牲にすることなく、予算と期間内に収めることができた
- 新規データソースを容易に統合し、必要に応じて変更に対応できる将来性のある新規データベースプラットフォーム
- 顧客満足度の向上をより小さなコストで実現



図5.リアルタイムな統合人事データ

安全な統合グローバルトレードストアを、6か月という短期間で開発

カスタマー
ドイツ銀行

背景

70か国に10万人以上の従業員がいるドイツ銀行は、ドイツに本社を持つグローバルな銀行/金融サービス企業です。銀行業務のリーダーであり、法人顧客ならびに個人顧客に金融商品ならびにサービスを提供しています。

組織の目標

複数の取引タイプ全体において、「single source of truth (信頼できる唯一の情報源)」を実現するために、ドイツ銀行ではデータを統合する必要がありました。また、統合取引後処理実現のためにオペレーショナル取引レポジトリ(グローバルトレードストア)を作成する必要がありました。下流の業務部門は、この統合データにアクセスし、必要なさまざまなレポート(規制順守レポートを含む)を作成します。

課題

ドイツ銀行には、取引に関連するデータ入力が多数あります。これはさまざまな業務においてさまざまな方法で利用されます。ここでの課題は、データをハーモナイズする(他と一緒に使えるように調整

する)こと、「single source of truth」を確立すること、データの完全性を実現することでした。彼らはまずRDBMSでこれらを達成しようとしましたが、失敗に終わりました。これは必要とされる共通スキーマを確立できなかったこと、また業務要件が常に変わっていたためです。より柔軟なデータ統合手法が必要とされていました。

主な結果

MarkLogicを採用したことで、ドイツ銀行は以下のことができるようになりました。

- 30以上の取引システムを、柔軟な統合グローバルトレードストアに統合。「single source of truth」を実現
- アジャイルなトレードストアでコスト削減。必要に応じてプロビジョニング容量を簡単に拡張。従来のリレーショナル技術で必要とされた大量のETLは不要
- 一元化されたエンタープライズ仕様のデータベースプラットフォームを銀行の基盤に。安全かつ一貫性のあるトランザクションを実現
- あらゆる構造化/非構造化データを素早く統合することで期間を短縮。最初のシステムは6か月以内で導入

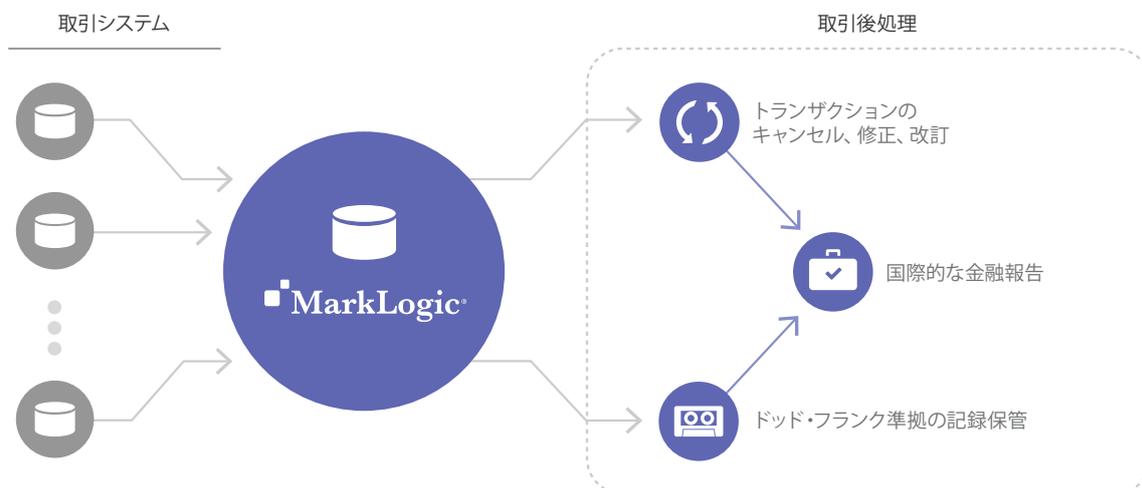


図6. 取引後処理の統合

大規模なダイナミックコンテンツ提供の新基準

カスタマー

BBC (British Broadcasting Company/英国放送協会)

背景

BBCは公共放送の世界的リーダーで、全国放送のラジオ局10局と地方局40局を運営。またデジタルアプリケーションや大規模なwebサイトを運営しています。

組織の目標

BBCは2012年の夏季オリンピックをインターネットを拡充して報道したいと考えました。具体的には、選手情報用のwebサイトを提供し、極めてパーソナライズされたインタラクティブなコンテンツをあらゆるデバイスにリアルタイムで提供するという事です。以前はRDBMSを使って静的なwebサイトを提供していましたが、このような固定的なデータベースプラットフォームでは、改善されたインタラクティブなユーザー体験に必要な、多様なデータの絶え間ない入出力を効果的に管理してオリンピック視聴者やスポーツファンに提供することはできないことを、BBCは理解していました。

課題

多機能かつインタラクティブなwebサイトの提供という目的実現のために、BBCはアジャイルなデータベースプラットフォームを開発する必要がありました。これは毎秒数万件のトランザクションを扱えるように拡張できるものでなければなりません。こういったBBCの努力に対して、さまざまな障害がありました。例えば、分断されたシステムに格納された

大量の複雑なデータや、老朽化してしまってダイナミックなマルチコンテンツ(動画、記事、統計データなど)を視聴者に提供できないITインフラといった問題です。これに加えて、時間を問わずあらゆる場所からwebを閲覧する数百万人のユーザーに対してオリンピック全選手の情報と記事を提供できるアジャイル性とパワーを持ったデータベース技術は存在しませんでした。

主な結果

MarkLogicを採用したことで、BBCは以下のことができるようになりました。

- 提供時間の短縮。従来のリレーショナル技術で要求された作成時間やインフラの制限が削減されたことによる
- ACIDトランザクションを活用し、安定し中断しないデジタル記事を提供。データロスなし。webサイトの安定したパフォーマンス。webサイトの停止なし
- 毎秒2万5000件以上のトランザクションを処理。450億件のリクエスト。1万ページ以上のダイナミックwebページ
- 極めて拡張性が高いパーソナル化されたデジタルユーザー体験を実現。これまでになかったほどの量のインタラクティブなコンテンツを提供(ピーク日2.8TB)。さまざまなデバイスを使用する数百万人のユーザーが対象

MarkLogicによるダイナミックコンテンツ提供システムの成功により、現在BBCでは、MarkLogicの使用領域を拡大し、人気のビデオオンデマンドサービス「iPlayer」のパフォーマンスを改善しています。¹⁷



17 Computerworld UK: BBC iPlayer sees performance uplift after relaunching on NoSQL database, June 2104
<<http://www.computerworlduk.com/news/data/bbc-iplayer-sees-performance-uplift-after-relaunching-on-nosql-database-3524893/>>

思考のスピードでビジネスが変化する:そのスピードで変化する準備を

複雑なエンタープライズプロジェクトの実装においては、いろいろなことがうまくいかない可能性があります。しかし自分たちも失敗する必要はありません。統計的に失敗が多いからと言って自分も失敗していい、ということにはなりません。賢く選択すれば、リスク削減と企業の目標達成のためにデータベースを活用できます。ここでプロジェクト成功のカギとなるのは、「適切な」人を集めて「適切な」技術に基づいて構築するということです。

思考のスピードでビジネスが変化する時代において、競合よりも価値を速く提供できるかどうかで成功が決まってきます。これを実現するために、重要なプロジェクトでは堅牢かつアジャイルなデータプラットフォームが必要です。これは現時点でうまくいくだけでなく、将来も進化していくものである必要があります。安全についての思い込みから、現代のデータに対して過去の技術を使おうとしないでください。旧来の技術はこれを扱うにはできていません。急激かつ継続的に変化するビジネス状況に対応する(無視するのではなく)心構えと技術を受け入れることで、データプロジェクトを成功させ、業務目標を達成できるようになります。

MarkLogicは、オペレーショナルかつトランザクショナルなエンタープライズNoSQLデータベースです。これを使ってあらゆる業界の組織が、サイロのデータを従来よりも短期間かつ低コストで統合しています。ビジネスリーダーたちに、統合データの360ビュー、エンタープライズ対応準備、アジャイル性、拡張性を提供することにより、世界で最も複雑なデータ問題解決のお手伝いをしています。

もはややりたいことを実現できない過去の心構えや技術は忘れてください。現代のデータ要件に特化した適切なチームと技術があれば、ハイリスクかつ重要なプロジェクトに成功をもたらすだけでなく、すべてのエンタープライズデータを活用し競争力を高めることができます。これは歓迎すべき変化なのです。

さらに詳しく

リレーショナルデータベースは、毎日私たちが利用し生み出している、大量の分断された繊細なデータがもたらす困難に対処するにはできていません。しかし、他の方法を使えば、すべてのデータを活用し、想定よりも短期間・廉価で組織の目標を達成することができます。MarkLogicの新世代のデータベース技術がデータプロジェクトを導いていく方法について、ぜひご確認ください。

リレーショナルを超えて

RDBMSでは、ほんの少しの変更でも莫大なコストがかかってしまう可能性があります。このためリーダー組織たちは、リレーショナルを超えて柔軟なデータベース技術へと移行しています。ここではデータモデリングが削減され、数億円規模のコスト削減の可能性もあります。ホワイトペーパーをお読みください。

<http://jp.marklogic.com/resources/beyond-relational-japanese/>

エンタープライズ NoSQL for Dummies日本語版

データベースとセマンティックを統合することで、従来のデータベースでは解決できなかったビッグデータ問題を解決できます。

<http://info.marklogic.com/nosql-for-dummies-jp.html>

データの複雑な課題を解決する

現代のデータタイプのために設計された新世代技術のMarkLogicは、唯一のエンタープライズNoSQLデータベースです。柔軟なデータモデルによりあらゆるデータを格納/管理/検索します。従来のRDBMSが持つデータの堅牢さや一貫性を犠牲にはしていません。弊社webサイトをご覧ください。

<http://jp.marklogic.com/what-is-marklogic/>



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