



DX(Digital Transformation)の実現が要請される中、 情報経済におけるデータ活用のあり方、その考察(1)

YOUniteフェデレイテッド・データ・ユニフィケーションで実現

目次

- 1. はじめに 01
- 2. DX を推進しデータ・ドリブン経営を実現 05
- 3. DQ やガバナンスの課題 05
- 4. DX を推進する上での制約事項 06
- 5. 実現に向けた施策 07
- 6. コスト削減効果も 10
- 7. 最後に 12

1. はじめに

企業が直面しているビジネス背景

全国規模でビジネス展開し100社以上の子会社を持つある企業は、情報経済の真ただ中で競合優位(TCA: Total Competitive Advantage)を勝ち取り、企業グループを存続させるために経産省が推し進めるDXを実現しデータ・ドリブン経営を行うための取組みを始めています。そのためにゴールを設定し、現状を分析し課題を洗い出して検討を行っています。

以下ではどのように実現にむけて課題を克服し、データ・ドリブン経営を成し遂げようとしているかを解説していきます。

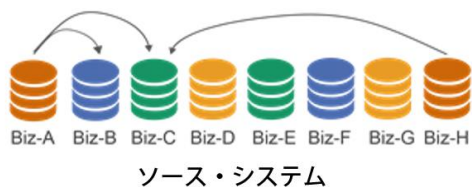
ビジネスの可視化と運営

経営管理を可視化するためにはデータの統一、意思決定者が適切なデータをいつも即座に利用するなどを通して組織はデータ・ドリブンになります。

⇒ さまざまな売り上げ情報や在庫管理情報などの数値データの精度向上と分析。核となる情報を全社統一してシェアし迅速な経営判断につなげる。

⇒ 既存の投資(インテグレーション)を無駄にせず、以下のような機能を持つデータ・ユニフィケーション・プラットフォームを導入する。

a. 従来のバッチによるインテグレーションを行っている

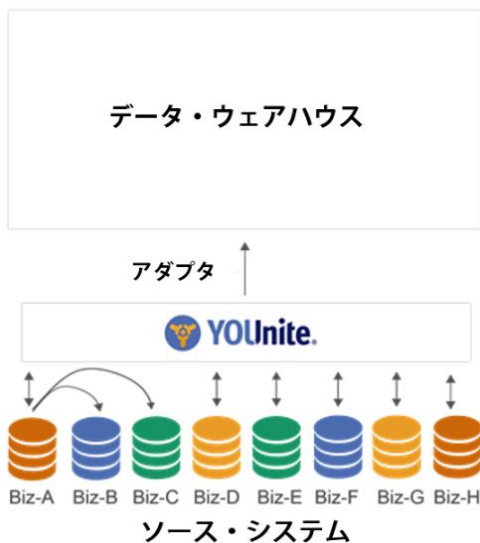


b. プリミティブなパスでは、主要なソース・システムからデータを取り出し、既存のインテグレーションを残したままデータ・ウェアハウスにロードすることができる

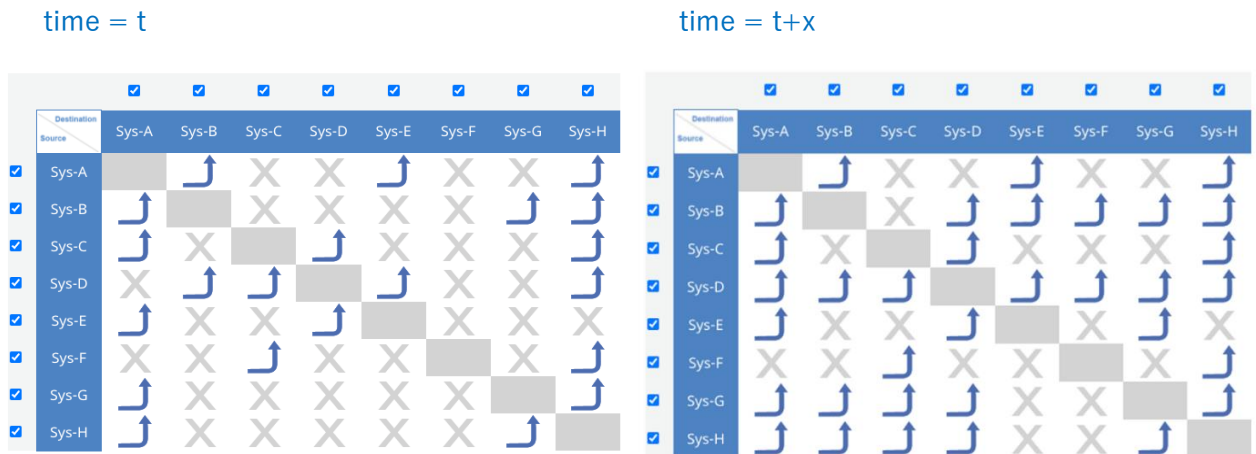


c. 従来のインテグレーションを超えてシステム間のリアルタイムなイベント・ドリブン同期を実現する。

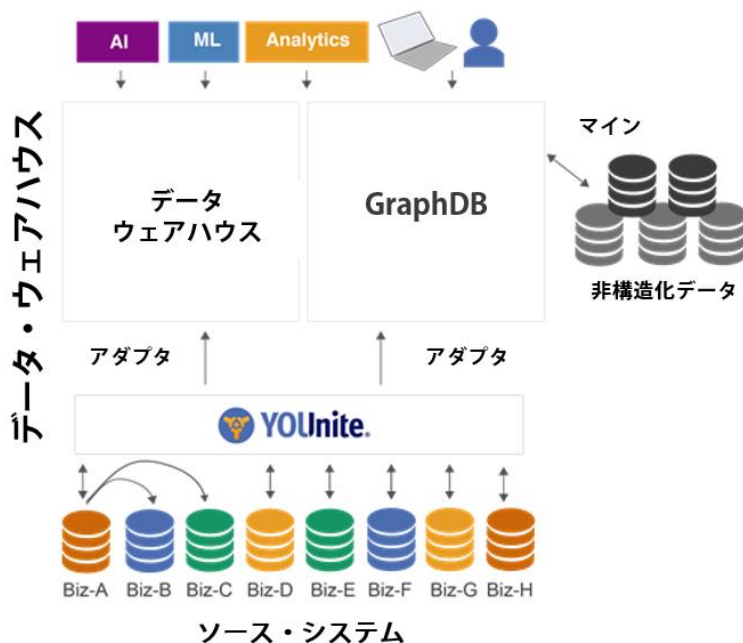
c-1. これまでの統合機能を壊さずに(投資を無駄にしない)、フェデレイテッド・データ・ユニフィケーションで新しいインテグレーションを実装・展開すること可能とする。アップデートが必要な既存のインテグレーションや、リアルタイム同期が必要なインテグレーションは、フェデレイテッド・データ・ユニフィケーションに移行させることができる



それぞれのシステムが時間とともにインテグレーションされていくようすを以下の図で説明



c-2. 新旧のシステムがデータ・ウェアハウスやGraphDBとあいまって顧客の売上や在庫管理などの最新のデータを経営陣が把握し、最適な経営判断を行うことが可能となる



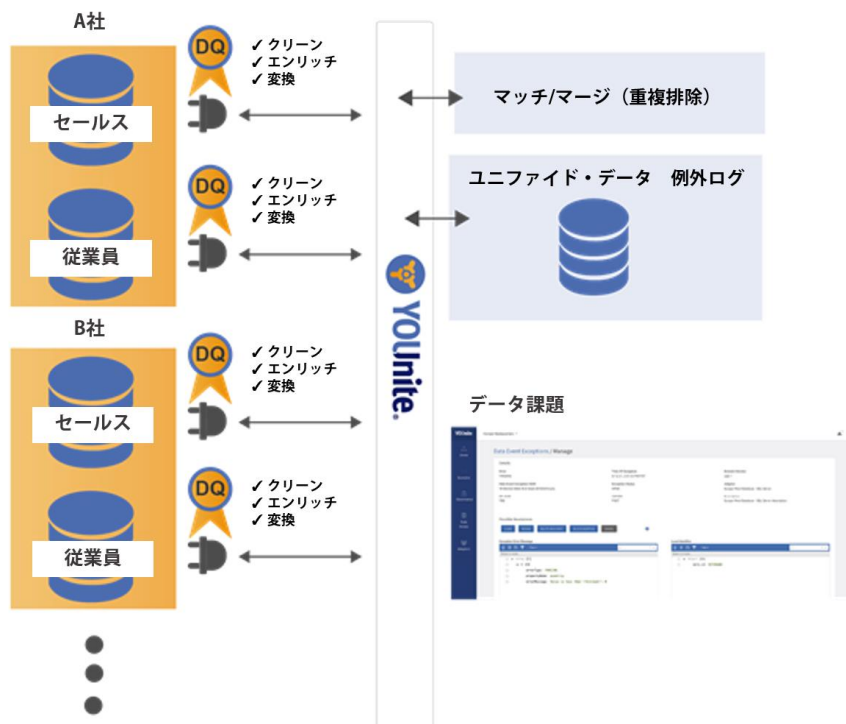
グループ内での情報とノウハウの共有から新たなインサイトを生む

⇒ 各会社が独自で蓄積・管理しビジネス展開している顧客情報を、グループ全体として活用することで顧客へのつながりをより強固なものとする。

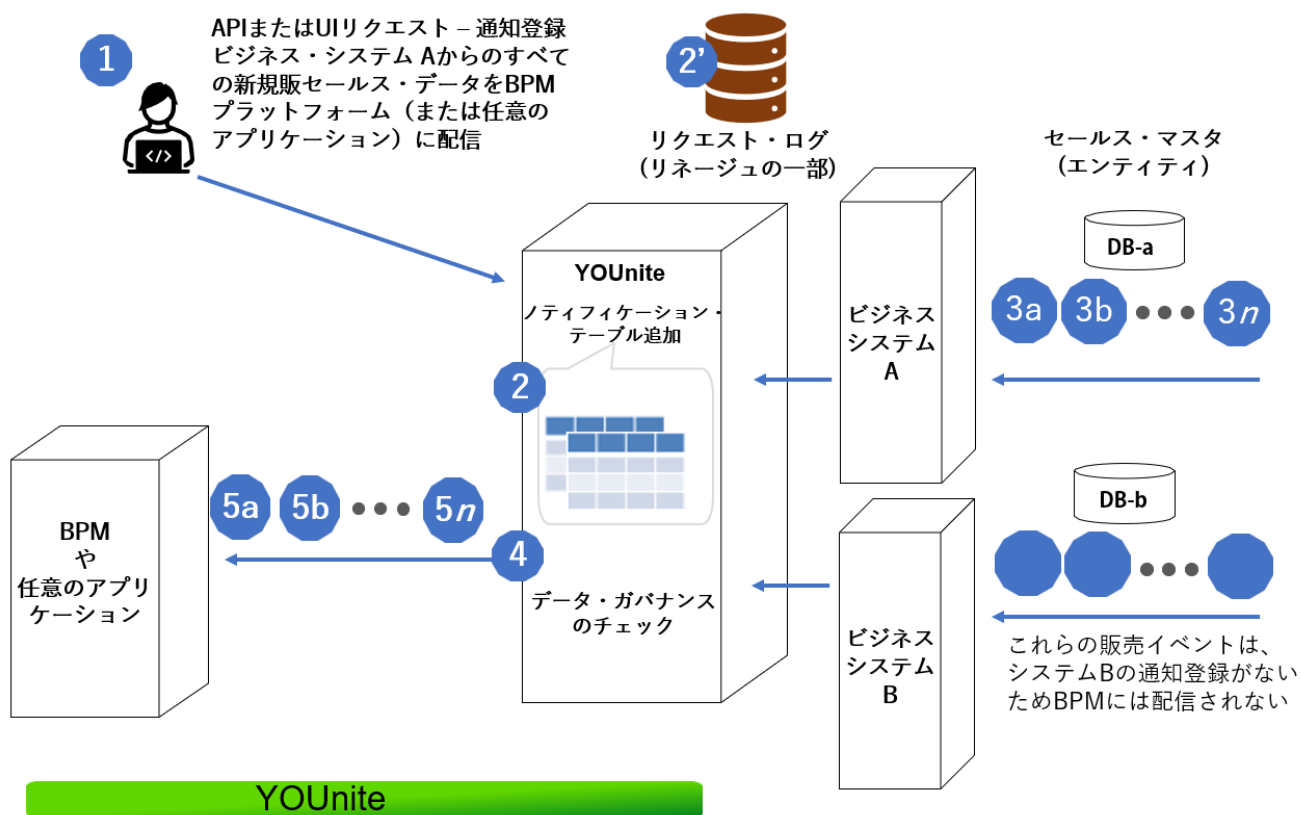
ノウハウやビジネス環境の共有から相互に連携した進め方が可能となる

⇒ YOUnitéフェデレイテッド・データ・ユニフィケーションは、望みの構成が可能なデータ・ガバナンスにより、個々のビジネス・ユニットが、会社グループや会社間で共有されるデータをコントロールできるようなフレームワークを提供する。

⇒ データ・クォリティ(DQ)の変換/ルール、正規化(マッチ/マージ)ルールを会社間で共有できる。
データの更新/アクセスログ、データの課題が一元化される。



⇒ ビジネス・プロセスの管理フローを会社間でも共有される。



マーケット浸透と新規ビジネスの開拓

⇒ 明確なデータ・インテグレーションのプランを持つことで、多くの買収や合併の際に問題となるデータ・インテグレーションの憶測や意思決定の不備を解消することができる。フェデレイテッド・データ・ユニフィケーションでは、企業は既存のガバナンス、同期、データ仮想化のフレームワークを共有することが可能。

2. DXを推進しデータ・ドリブン経営を実現

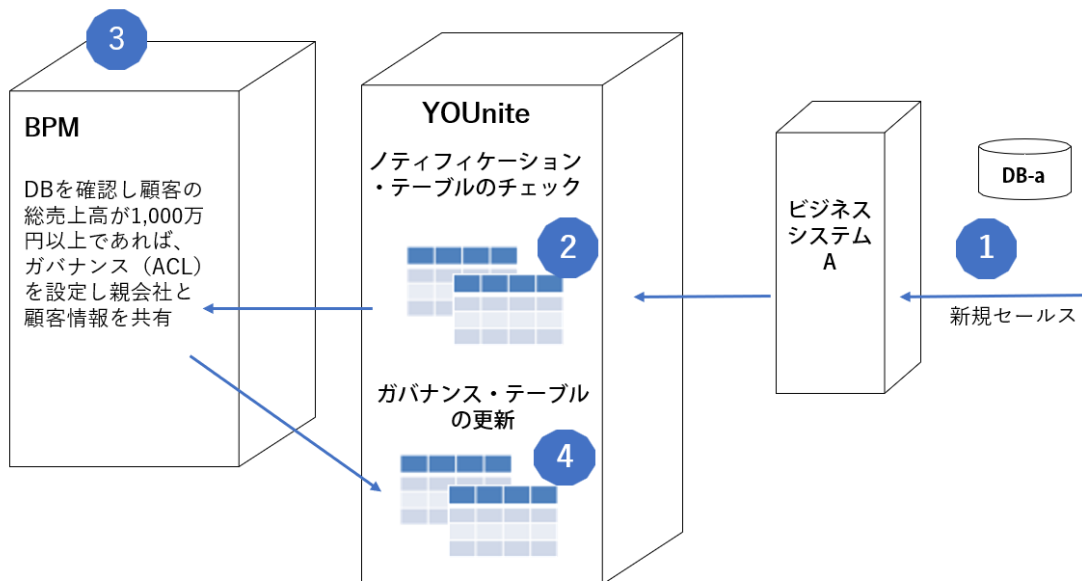
全国各社が保有する顧客情報を、締結中の個人情報規約の範囲内でグループ共有する。

各システムの顧客マスタ・データの更新を、リアルタイムまたはバッチで同期

⇒ YOUniteは、リアルタイムとバッチの両方で、全社のソース・システム間でデータを同期させる。教訓 - バッチシステムは、他のバッチシステムまたはデータ・ウェアハウス/データ・レイクのみを矛盾が生じないため更新すべき。

顧客の業務情報(トランザクション)を自システム以外でも分析に利用

⇒ ガバナンスは、会社やソース・システム間のデータの流に影響を与える。ビジネス・トランザクション情報を分析して共有する必要がある場合、BPMツールを使用してガバナンスを変更したり、ワークフローを作成したりする。例えば、ビジネス・システムAからのすべての売上に対して通知登録が行われているとする。



3. DQ やガバナンスの課題

各社のサービスでは顧客との契約時に利用範囲等を明示し、データの利用範囲を決めている。

⇒ ガバナンス・ルールは事前に定義され、レコード作成時にガバナンスが適用され、顧客との契約時に制限が適用される。プライバシーに関する規制やお客様ごとのオプトアウト(Opt-out)/オプトイン(Opt-in)の選択が可能。

各社は個人情報を、そのビジネスに必要な情報のみ収集している(氏名のみ、住所のみといった具合)。

各社のシステムは、通常各社それぞれに構築しているため、データベースはそれぞれに異なる。

⇒ 既成のアダプタとして、リレーショナル・データベース用とKafka用、加えてアダプタ開発用SDK、この種のシステムに付随するもの：

- 1) 既成のアダプタとのインテグレーション
- 2) アダプタ開発
- 3) バッチのみ (選択可能)
- 4) インテグレーションが難しいケース - ガイダンスを提供

基幹業務には顧客マスタ・レコードが3,000万ある場合。

⇒ 扱える顧客マスタ・レコード数(トランザクション数/日)に制限はない。

データ・クウォリティの現状

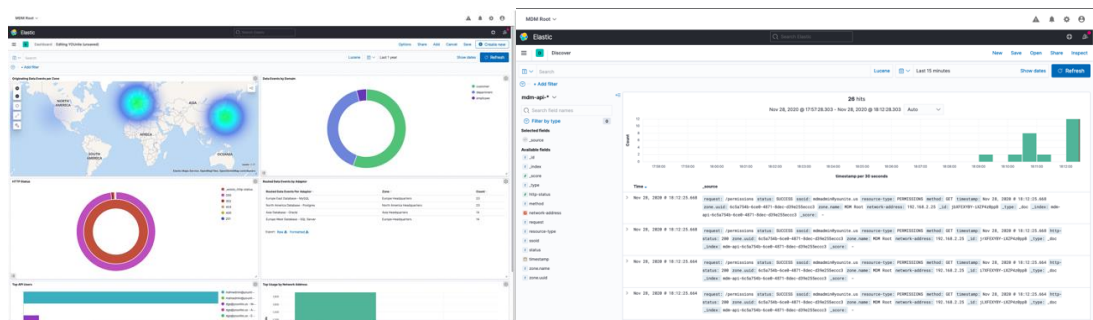
基幹となる業務(システム)がいくつか存在するが、それぞれのシステムが持つ顧客情報が正規化されていない。勿論、これら以外のシステムも同じような状況。

⇒ データを正規化し、ひとつDQベンダを選択しアダプタを構築するか、あるいはすべてのデータに一貫したDQを適用し、ユニフィケーションを開始する準備が整うまで同期しない状態を続けるか。

各システムの情報管理は担当者依存のため、顧客マスタの品質が維持できない。

⇒ DQの現状で述べたように、ビジネス間でも基準に沿ったDQを再利用/共有するのがよい。

データ・ユニフィケーションにより、正常なデータ・クウォリティ変換と例外が記録され上級管理者向けの会社全体のレポートが作成され、それぞれのビジネス・グループに行動指針を与えることができる。



4. DX を推進する上での制約事項

基幹業務システムの更新は1年足らず。このようなタイミングでの業務システムの改修は、基本は行われぬ。仮にあって最低限の範囲で実施となる。

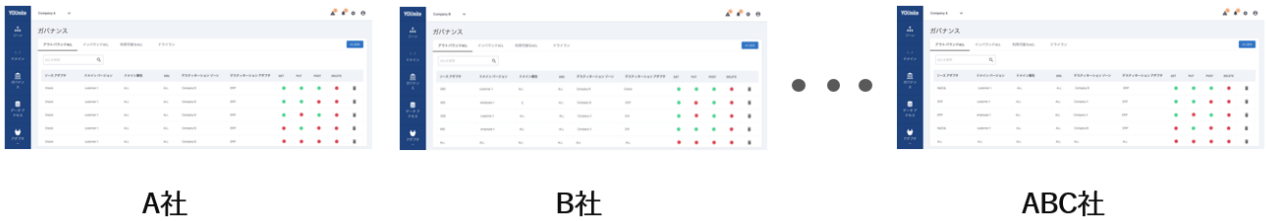
⇒ 既存のシステムに手を加えることなく、段階的に導入していくことができる。

各社の個人情報取り扱いの制約及び、これまでに収集済み情報が異なるため、自社業務範囲外のデータを扱うことはできない。

例) B社はA社が持つ顧客の電話番号を以下のプロパティの違いからわかるように利用できない

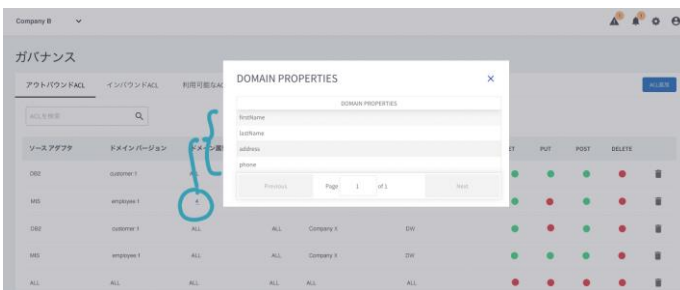
A社収集済情報 = 氏名、住所、電話番号

B社収集済情報 = 氏名、性別、メールアドレス



⇒ 各企業は、ネットワークのファイアウォールのルールと同様に、独自のインバウンドおよびアウトバウンドのデータ共有ルール（データ・ガバナンス）を管理している。データ・ガバナンスが適用される範囲は同期とデータ仮想化。

⇒ データ・ガバナンスのコントロールは、企業間での個々のデータ・プロパティ(氏名、住所、性別、電話番号、電子メールアドレスなど)の共有を制限/許可する。



5. 実現に向けた施策

顧客情報の正規化

各業務システムのデータの洗い出し

- ・各システムでの名寄せにより、各システムでのデータ正規化
- ・各社で持つ顧客属性情報（氏名など）のうち、他システムの顧客情報と紐づけられる属性を抽出

データ・マッピング

紐づけ実施(例)

システム	カスタマID	氏名	生年月日	住所	電話番号	メール
A	◎	◎	○	◎	◎	○
B	◎	◎	×	×	◎	×
ABC	◎	◎	◎	×	◎	◎

顧客情報管理
 ◎：必須
 ○：任意
 ×：対象外

カスタマID	氏名	生年月日	住所	電話番号	メール
C001113	ラリー・ネルソン	1971/1/10	千葉県松戸市・・・	047-123-5678	larry@xxx...
CID-003	ラリー・ネルソン			047-123-5678	lnelson@xxx...
CID-009	マーク・ザッカーバーグ	1986/11/10		044-987-6543	mark@xxx....
CID-666	スティーブ・ジョブズ	1962/4/1	東京都港区・・・	03-3210-9875	
CD123	Steve Jobs	1962/4/1		03-3210-9875	
CD234	Mark Elliot Zuckerberg	1986/11/10		03-2222-4444	mark@xxx...

⇒ スコアリング・ルールとマッチング・ルールは、顧客を含むすべての情報を正規化するためのフレームワークとなる。マッチング・アルゴリズムには、2つのレコードの正規化ルール(DR Key Properties)が一致を示しているか、または2つ別々のレコードであるかを判断するSQLのようなルール・セットを含んでいる。

⇒ デフォルトでは、マッチング・アルゴリズムが指定されていない場合、レコードはDRキーのプロパティのすべての値が同一である場合にのみ一致すると判断される。カスタム・ルールやスコアベースのマッチング・アルゴリズムは、データが完全に一致するかしない場合に異なるシステム内の重複レコードを特定するのに便利。

⇒ スコアベース・マッチングとルールベース・マッチングの両方で、曖昧マッチングが可能である。曖昧マッチングはデータ課題となり、適切にデータ・スチュワードが、新しいレコードが既存のレコードとのマッチングなのか、それとも新しいレコードなのかを判断する。

⇒ 例えば、次の従業員のデータ・ドメインには、スコアリング・アルゴリズムがあり、新しい従業員のレコードのスコアが（既存のすべての従業員と比較して）30未満であれば、新しいレコードとみなされる。他のすべての従業員レコードと比較して、スコアが70より大きい場合には、一致しているとみなされる。スコアが30よりも大きく70よりも小さい場合、入力されたレコードは曖昧であるとみなされデータ課題が発生する。マッチング/スコアリングのデータ課題では、データ・スチュワードが新しいレコードなのか、既存の従業員レコードと一致するのかを手動で判断する必要がある。

1.MATCH TYPE > 2.MODEL SCHEMA > 3.SCORING GROUP > MATCHING GROUP (OPTIONAL) > DR KEYS (OPTIONAL) [VALIDATE] [SAVE VERSION]

POSSIBLE MATCH SCORE
30

DEFINITE MATCH SCORE
70

Description
Employee matching algorithm

Scoring Algorithm

```
1. IF (employeeid EQ CASE SENSITIVE ) THEN 20 ELSE -10 WHEN[ (employeeid SRC_NOT_NULL CASE SENSITIVE AND employeeid DEST_NOT_NULL CASE SENSITIVE ) ]
2. IF (firstName EQ CASE INSENSITIVE ) THEN 20 ELSE -10
3. IF (lastName EQ CASE INSENSITIVE ) THEN 20 ELSE -10
4. IF (email EQ CASE INSENSITIVE ) THEN 20 ELSE -10
5. IF (birthdate EQ CASE SENSITIVE ) THEN 20 ELSE -10
```

[+ SCORING GROUP]

⇒ 部分的に一致する文字列の値用に、いくつかのファジー・マッチング・アルゴリズムがある。

MATCHING CRITERIA

AND [x] [+ MATCHING STEP] [+ CHILD GROUP]

[x] address [x] EQ [x] FALSE [x]

[x] cit [x] EQ [x] FALSE [x]

[x] sta [x] EQ [x] FALSE [x]

LEVENSCHTEIN [x]

TRIGRAM

SOUNDEX

METAPHONE

DMETAPHONE

<https://younite.us/resources/Matching-Algorithms.html> に詳細があります。

フェデレイテッド・データ・ユニフィケーション

データ・ユニフィケーション

YOUniteサーバで各業務システムが持つ顧客メタデータと顧客IDをマッピング。

⇒ 各ビジネス・ユニットのニーズに合わせて、マッチ/スコアのアルゴリズムを企業が定義。各社のマスタ担当者(データ・スチュワード) が連携し、顧客データをグループとして正規化(名寄せ等)。

⇒ 同上。

マスタ更新

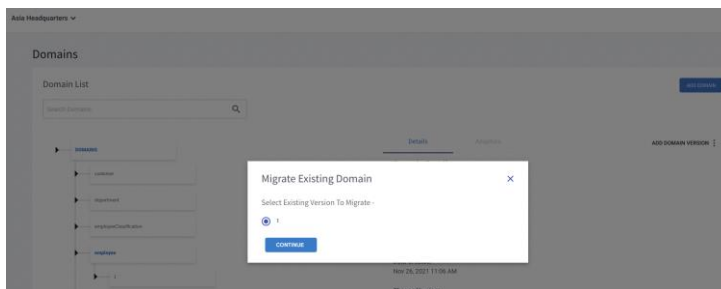
各システムが更新、新規追加、削除した情報はリアルタイムで他システムへ通知。

属性追加には、どのような作業が必要か？

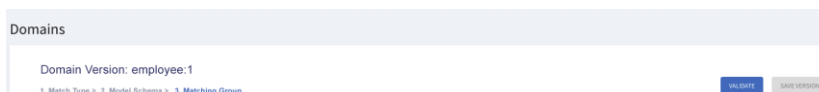
⇒ データ・ドメイン(例: customer)に属性を追加するには、既存のドメイン(1.0)をサポートするすべてのアダプタに互換性のあるマイナー・バージョン(例: 1.1)を作成する簡単な作業。

⇒ 属性の削除、属性タイプの変更、正規化ルールの変更(マッチング/スコアリングを含む)には、メジャー・バージョン(バージョン)の更新が必要であり、更新されたドメイン・バージョンをサポートする必要があるすべてのアダプタを更新しなければならない。バージョン更新を管理するための確立されたワークフローがあり、通常 1 つのドメイン・バージョン(例: employee v1)が新しいドメイン・バージョン(employee v2)に移行される。

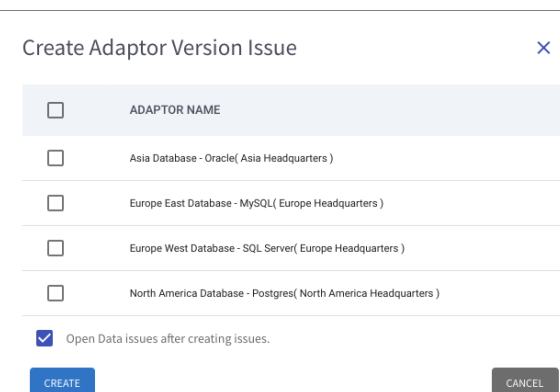
既存のデータ・ドメインの移行をYOUnite UIデータ・スチュワードが要求



データ・スチュワードはウィザードでデータ・ドメインを移行



データ・スチュワードは、以前のドメイン・バージョンをサポートするすべてのアダプタのデータ・スチュワードに作業項目(データ課題)を追加するよう促される



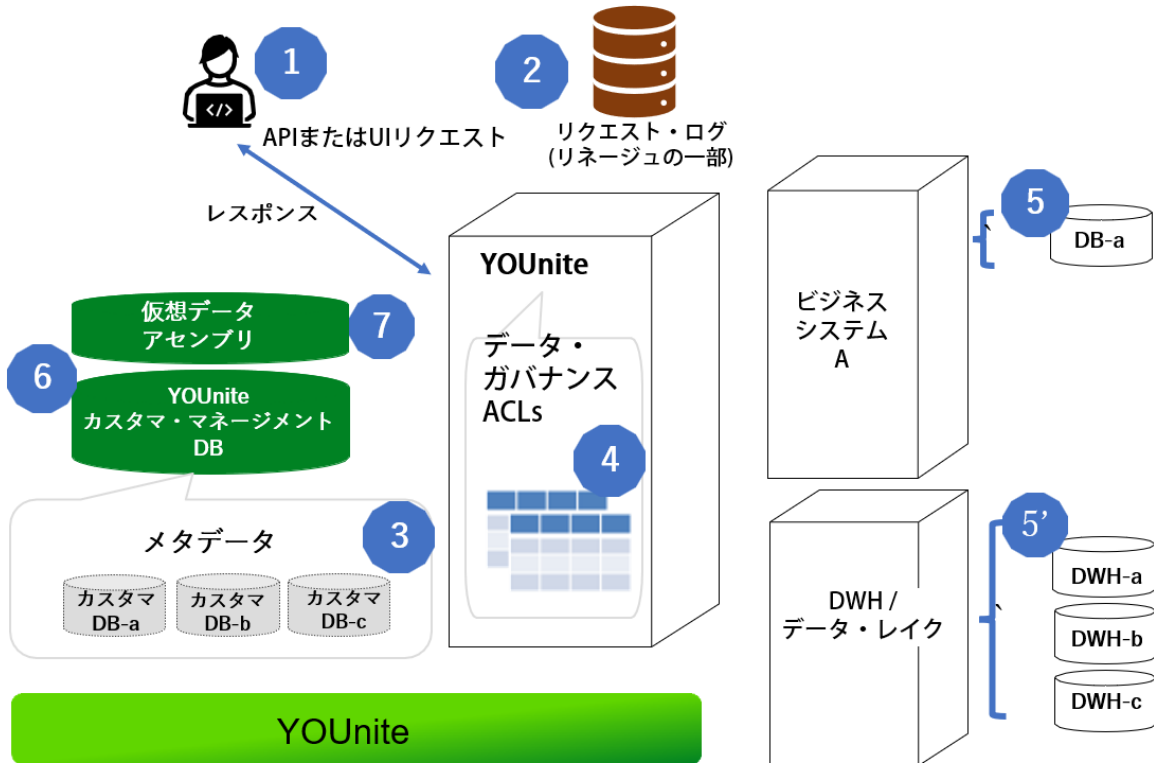
注意：アダプタは、1 つ以上のバージョンのデータ・ドメインをサポートすることが可能

データの利活用

DWH連携

既存のDWHと連携

A業務担当者は、A業務システムの顧客の情報をA業務だけではなくこの顧客に関わる B業務+C業務の情報も抽出可能。



6. コスト削減効果も

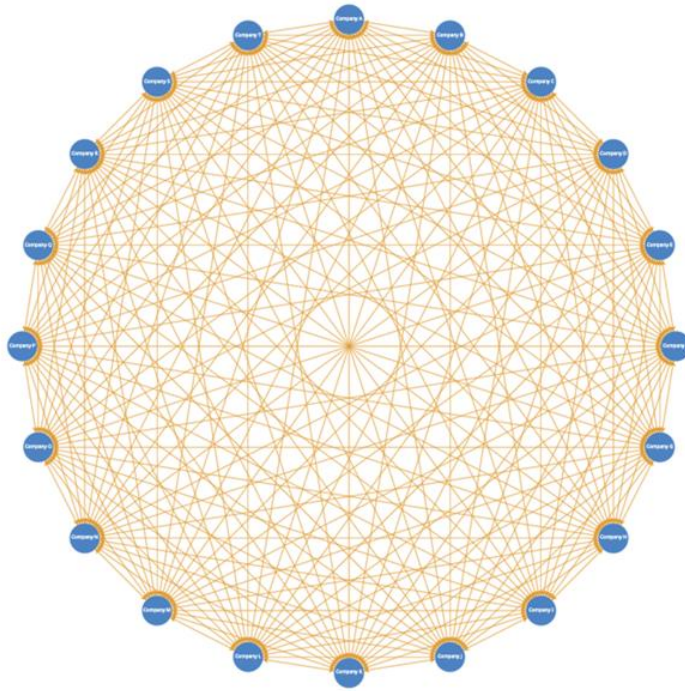
30社すべてをその場しのぎの統合すると手に負えなくなるのは必然。その結果、システムが適切にインテグレートされず、データの質が低下する。また、メンテナンスや再コーディングなどの年間コストが激増する。

以下は、20のシステムのフェデレイテッド・データ・ユニフィケーションについてのコスト試算。

YOUnite フェデレイテッド・データ・ユニフィケーション・モデル



アドホック・コネクション・モデル



総インテグレーション・コストの見積り例

Total Integration Cost Estimates

Total ad-hoc integrations needed : 380

Total YOUnite integrations needed: 40

Estimated cost per integration : \$25,000

Estimated annual integration cost: \$25,000

	Ad-Hoc	YOUnite	Savings
Cost to Implement and Deploy	\$9,500,000	\$1,000,000	\$8,500,000
Annual Costs	\$9,500,000	\$1,000,000	\$8,500,000
Total costs over 5 years	\$57,000,000	\$6,000,000	\$51,000,000

機能比較表

Feature Comparison

It's possible that your enterprise has partially implemented some of these features on an ad-hoc basis with your existing integrations. With YOUnite you get unified federated data routing, data access and single pane views across the entire data tier for the following:

Feature	Ad-Hoc	YOUnite
Event-Drive, Real-Time Data Unification/Router	?	Yes
Fine-Grained Data Governance/Firewall	?	Yes
Data Lineage	?	Yes
Enterprise-Wide Federated Data Access (REST API, GraphQL, Lucene, JDBC)	?	Yes
Enterprise Wide Data Quality	?	Yes
Global Match/Merge/Data Exceptions	?	Yes

7. 最後に

YOUniteは、ネクスト・ジェネレーション・データ・マネジメントであるデータ・ファブリック(Data Fabric)の先駆者です。

YOUniteをお試しになりたい場合は、お客様の環境要望に沿ったかたちでのMock-upデモを行うこともでき、次いでPoC(Proof of Concept)へとステップを踏んで実施することができます。

是非弊社までお問合せください(Email : sales@younite.us)。

詳細をお知りになりたい方は弊社サイトにある日本語ドキュメントや日本語吹替えビデオもご参照ください。

<http://younite.us/articles-and-videos/>