

JISA招待論文

法人審査業務におけるWebおよびSNSデータの活用—法人審査ソリューション—

宮本 拓也¹

¹ (株) NTTデータ

本稿は、昨今のICT技術の指数関数的向上による従来実現できていなかった、身近で自然に存在する多種多様なデータを活用している事例として、活用いただいている法人審査ソリューションについて述べたものである。法人審査ソリューションの概要、実際に法人審査ソリューションを適用した事例、導入にあたってのノウハウ、今後の展望について述べる。

1. はじめに

現在のITの世界ではICT技術の指数関数的な性能向上により、これまで構想としては存在していたものの実現できていなかったことが実現できてきている。

その1つが身近で自然に存在する多種多様なデータの活用であり、従来のリレーショナルデータベース（RDB）で扱ってきた構造化データに加えて、非構造化データと呼ばれるデータも扱えるようになってきている。具体的には、文字、音声、画像、動画などありとあらゆるさまざまな種類のデータを同列で扱うことができる（図1）。

身近で自然に存在する多種多様なデータの活用 ⇒ デジタル社会において社会・顧客の変化を捉える源泉



図1 多種多様なデータの例

その技術的背景には、以下が挙げられる。

- 多様なデータをやり取りするためのネットワーク性能の向上
- ストレージの技術革新による大容量データ格納のためのコスト低廉化
- 多様、大量データの高次活用のためのCPU処理能力の向上

こうした進歩を背景に、従来は構造化（定型化）されていなかった情報は人手で扱う必要があったが、システムによって人手を介さず扱えるようになってきている。その恩恵は多様な業務、分野で享受できるものであるが、とりわけこれまでシステム化されていなかった分野ではなお得られる果実が大きくなる。

そういった分野の1つが法人（加盟店）審査業務である。そもそも法人（加盟店）審査業務の分野では個人審査などの領域に比べてシステム化が進んでおらず、投資も限定的になっていた背景がある。その理由は、個人に比べて法人はその数が圧倒的に少ないことにあり、故に人手での対応でもまかなえていたことにある。従来はそうした法人の情報については調査会社のレポートを参照したり、社員による実地調査などを行うことで収集を行っていた。ところが昨今はベンチャー企業やWeb上での起業、個人事業主の増大など働き方や企業形態の多様化と柔軟化によってその数が爆発的に増加しており、人手でカバーするのは質的にも量的にも不可能になってきている。

加えてインターネットの普及により、Webというソースを審査業務へデータ活用することも一般的になっている。特にインターネット上のみで商売をしている企業についてはソースがWebのみに限られるためほかのソースではカバーが困難なこともあり、その必要性、重要性がますます顕著になっている。同様にその舞台がユーザ動態に合わせて単純なWebサイトだけでなくSNSにも拡大しており、データ活用への期待が高まってきている。

ここでは法人（加盟店）の審査において今まで人手でも成し得なかったことの実現をデータの活用という観点から紹介する。

2章では、法人（加盟店）審査ソリューションの概要と課題について述べる。続く3章では、法人（加盟店）審査ソリューションそのものの構成とそれぞれのポイントについて述べる。続く4章では、ソリューション導入の効果について述べる。続く5章では、審査業務へのインターネット情報の活用可能性について述べ、最後の6章では、まとめと今後の展望について述べる。

2. 法人（加盟店）審査ソリューションの概要と課題

2.1 法人（加盟店）審査業務の概要

法人（加盟店）審査ソリューションは、インターネット上のWebサイト・SNS等の社外情報や社内に蓄積されている過去取引履歴、ネガティブ情報等の社内情報を当該法人単位で収集することを自動化し、NoSQLデータベースに格納し、以下に記する法人審査業務の高度化・効率化を実現するソリューションである。

・法人審査業務の種類

法人審査には次の2種類の審査がある。

初期審査：新規に申込のあった（契約意思のある）法人について、契約／取引対象として契約元法人が個別に定める水準を満たすかの審査を行うこと

途上審査：初期審査を通過した契約中の法人について、契約元法人が定めた一定周期もしくは契機のタイミングで個別に定める水準を満たすかの審査を行うこと

・初期審査業務の流れ

1. 申込受付：紙やFAX， Web， メール等を受領することで受付とする。基本的に申請書がベースである。
2. 情報登録：受領した申込について手入力で内容をシステムに入力（登録）する（Web経由以外が対象となる）
3. 受付処理：入力された情報に従い申込の受付処理を実施する。
4. 情報収集：社内にある基本情報， インターネット情報， 外部データ（社外のデータベース等）の情報を検索し， 収集する。
5. 情報整理：収集した情報を法人ごとに整理する。
6. 判定：審査担当者に判定依頼を行い， 情報を元に判断する。必要に応じて追加情報の収集を行う。場合によっては上席に判定依頼を実施する。
7. 決裁：判定に応じた決裁を実施する。
8. 通知：判定結果を通知する。

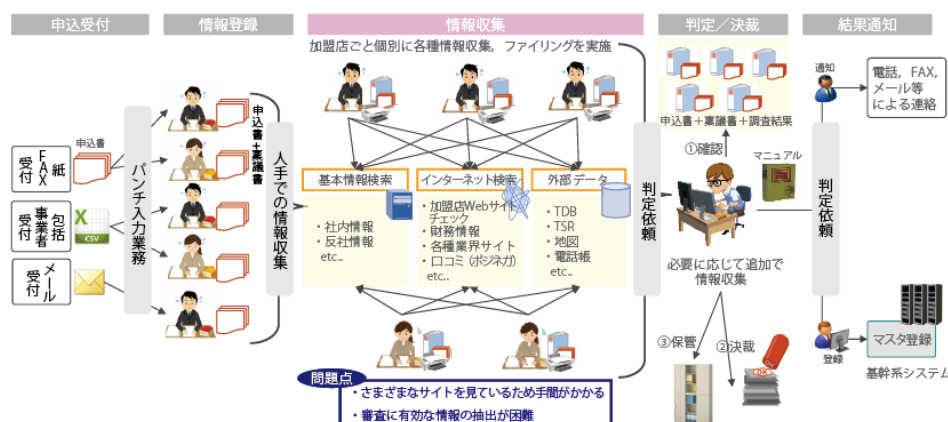


図2 法人（加盟店）審査業務フロー

・途上審査業務の流れ

契機：途上審査委業務の契機には，大きく2つがある。1つ目が「既存の契約条件に影響を与えそうな事案の発生を契機とするものである。法人の倒産， 経済産業省といった当局からの問合せや， 不正検知などが確認され当該法人の再調査および契約交渉を行う際に実施する」。2つ目は，周期的な実施である。昨今では当局から，既存の契約に対し定期的なモニタリングの実施を求められている。

上記の契機により，対象加盟店が抽出され，途上審査が開始される。途上審査の流れは，初期審査業務の流れ 4. 情報収集～6. 判定 部分と同じである。

判定結果に基づき，当該法人の調査や交渉が実施される。

2.2 法人（加盟店）審査業務の課題

初期審査業務の課題は以下の3点である。

- 申込から結果通知までのプロセスの期間を短縮する。
- 判断根拠となる収集情報の高準化と判定精度を向上させる。
- 業務プロセスにおけるリソース配分の適正化を行う。具体的には従来は情報収集プロセスにかかっていたリソースを判定プロセスに振り分けることを目指す。

上記初期審査業務での課題について具体例を挙げて説明する。

あるクレジットカード会社における審査業務の流れは以下であった。

- 加盟のための審査を受け付けた後、申込書の情報を人が手打ちで入力を行う。
- 入力された内容を元に次に必要な情報の収集を社内のシステムと社外のインターネットから人手で収集する。
- 社外のインターネットからはいわゆる検索エンジンを活用して、法人名の検索から企業情報や評判情報、実在チェックとしての地図上での検索を実施する。
- 必要な情報をすべてプリントアウトして法人ごとに1冊のバインダに綴じた上で審査担当へ渡す。
- 審査担当はバインダを見ながら審査を行い、結果を判定する。

しかし、この方式では手間もコストもかかっていた。申込書の受付から判定まで一法人の審査にかかる所要日数は10日前後に及んでいた。

一方でいかに早く申請書に対する回答判断を正しく下して利用ユーザ（加盟店）となってもらうかは、クレジットカード会社からすると競争力にも直結するため、その短縮化・自動化は課題でもあった。

また、情報収集については担当者によって検索手法もレベルも異なることから審査に必要な情報がきちんと取れてるかにはバラつきもあった。

業務としての肝はこの判断部分にあるため、情報収集に限られた人的リソースをかけるのではなく、いかに判定にかかわる部分にリソースをシフトさせるのかも課題として上がっていた。

また、途上審査業務の課題は以下の4点であった。

- 定期的な契約済み法人の全量審査を実現（可能に）する
- 途上審査につながる2.1で述べた契機を洩らさずに把握する
- 捉えた契機を元に該当する法人の途上審査を実施する仕組みの確立
- 当局が定める審査水準や審査にあたっての要請への柔軟な対応をいかに実現するか

途上審査業務の課題についても具体例を挙げて説明する。

初期審査以上にこれまで実施できていなかったものが途上審査である。実施ができない理由は量である。申込を契機とする新規審査は一時期に集中することは少ないが、途上審査は、審査を通過している全法人を対象とするため数千、数万を対象として実施する必要がある。定期的に全量検査を行うというのはこれまでは非現実的でありどこも実施できていなかった。

一方で昨今の消費者保護の観点から不正加盟店のチェックや取り締まりがクレジットカード会社に課せられており、全量チェックも早晚必須となる風潮にある。

3. 法人（加盟店）審査業務ソリューション

3.1 法人（加盟店）審査業務ソリューション概要

人海戦術となっている法人（加盟店）審査業務のうち、システム化がされていない情報収集プロセスおよび判定プロセスに着目して、この部分の新たな仕組みとして本ソリューションを適用することで法人（加盟店）の審査業務において手間がかかるネット情報収集の業務負荷軽減や収集プロセスの自動化および判定の自動化の実現を通じて、審査期間の短縮および収集情報の高度平準化を実現している。また、途上審査についてはいかにタイムリーに契機を把握して、審査プロセスに入るのかが重要となり、その自動化も実現している。

ソリューションとしてのポイントは、図3の通り、インターネット情報収集（クローリング）を元にして、企業・店舗Webサイトという情報に加えてソーシャル情報も収集することにある。ポイントを整理すると、構造化されていないデータの活用、初期審査の実現、途上審査の実現、の3点となり、そのコンセプトは以下の通りである。

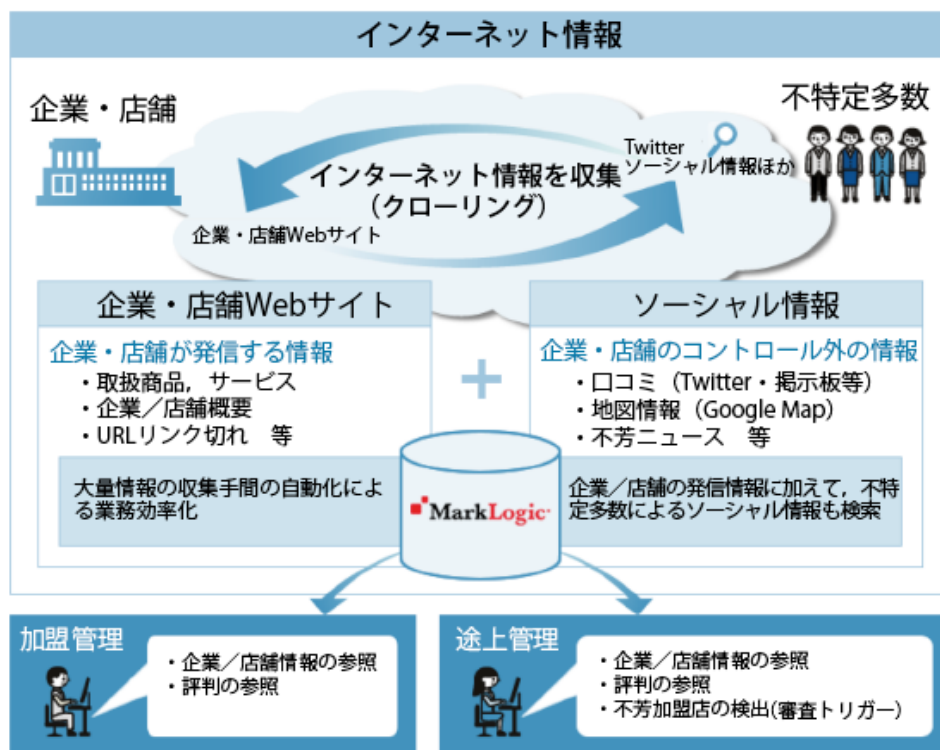


図3 法人（加盟店）審査ソリューションコンセプト[1]

・審査情報収集の自動化

従来の法人（加盟店）審査業務でその必要期間の多くを占めていた、審査根拠情報の収集の自動化により法人（加盟店）審査業務の期間短縮と収集情報の収集範囲と抽出内容の高準化を実現する。

・審査情報について

法人（加盟店）審査業務ソリューションでは、次の4種類の情報を審査根拠情報として活用する。

-社内情報（審査実施企業の過去実施情報等）

対象法人の過去履歴有無や代表者の確認等に使用

-当該法人オフィシャル情報（公式Webサイト等）

情報の整合性確認に使用

-社外企業チェック情報（企業調査情報等）

客観的評価情報として使用

-評判情報（SNSや口コミサイト等）

客観的評価情報として使用，社外企業チェック情報に比べてカバー率が広く，かつ生の率直な評価の傾向

・法人（加盟店）審査ソリューションの構成

法人（加盟店）審査ソリューションは以下の構成からなる。

-審査情報収集

--クローリング

--解析

-判定

-途上審査（契機把握，全量チェック）

本ソリューションの構成の特徴として，審査情報収集部分については，インターネット上の情報は不定形かつどこに存在するかも分からないため，クローリングで見つける機能部（クローリング）と見つけた情報を解析する機能部（解析）に分けていることが挙げられる。以下，構成の各部の詳細について記述する。

3.2 審査情報収集 クローリング

審査情報収集 クローリングを実現するにあたり，本ソリューションのために独自にクローラを開発した。開発にあたっては以下の点に留意，工夫を行った。

・情報源に合わせた取得方法

インターネット上の情報をクローリングするにあたっては法人が正式に公開しているオフィシャル情報と，外部サイトにおける評価などの評判情報を区別した。それぞれの情報について取得の方法やロジックが異なるため，別個の開発としている。

たとえばオフィシャル情報については，社名や住所などの組合せから公式WebサイトのURL特定を行う一方で，評判情報はそもそも一個所にしかないわけではないため，社名+ほかのキーワードで広くクローリングを実施する必要がある。

・多様なWebサイトに対応するための汎用化とチューニング

情報を見つける機能部ではクローリング技術が鍵となる。極力汎用化を行い、多様なサイトに対応できるようにしておくことがまず重要となる。一方ですべてのサイトに対応できるクローラの実現は事実上不可能なため、いかにチューニングと呼ばれる適応作業を簡易に行えるようにするかということと、適応作業を行うための体制整備とともに技術力確保と継続的なモニタリングが必要となる。

チューニングについては、Webサイト分析のための仕組みが元々備わっている機能であるサイトの型の類型化に加えて変更への対応パターンを用意しておくことで実際に変更が発生した際の対応を簡易にしている。

・変化する情報源への対応

インターネット上の外部情報は、検索する日、時間、タイミングによって結果が変わるという特性がある。真に正しい姿（結果）というのは存在しないし、定義もできない。検索結果＝収集する情報が常に変化し続けるという前提のもと、ソリューションについてユーザにも特性を理解してもらう必要がある。収集する情報が常に変化し続けるということは、収集した内容とその内容をベースにした判定結果がクローリングのタイミングに影響を受けるということである。結果だけでなくサイトに書かれている内容も更新などで日々変化していくため、あるときは正常に取得できても次のときには取得に失敗することもある。上述の通りこうした変化に強いスキームを構築しておくことが重要となる。一例として正常に取得できない場合にはアラームを出す、類似性から候補を提示する、などの手当ても補足機能として具備している。

3.3 審査情報収集 解析

審査情報収集 解析を実現するにあたり、以下の点に留意、工夫を行った。

・情報源ごとの抽出情報

企業の公式サイトやSNSといった情報源ごとに抽出する情報は異なる（図4）。たとえば、企業公開情報であれば企業の公式Webサイトから、評判情報はSNSから、としている。また、それぞれの特定

を行った上で抽出を行っている。必要な情報はWebサイトごとに表現が異なるケースもあるため、NTT研究所の自然言語解析を活用してその吸収を行っている。

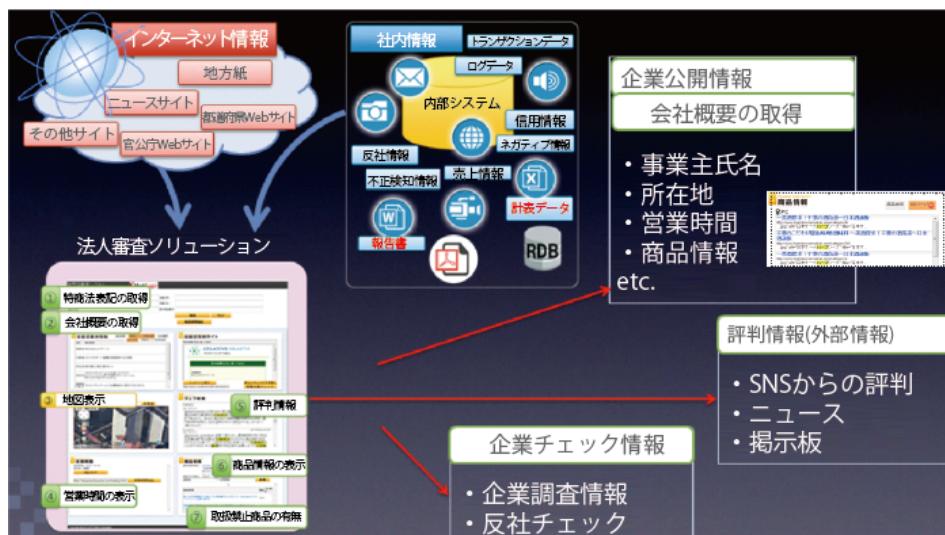


図4 情報源ごとの抽出情報

クローリングにより収集した情報は、インデックスを自動で付与するとともにエンリッチ技術を用いて自動でタグ付与を行うことで後からの検索をきわめて効率的に実施可能とした。インデックスは後から当該情報にたどり着くためのカタログ情報として活用する。タグはXML形式でデータを保持する際に各データ項目一つひとつにそのデータの意味とともに付与されており、データ項目の値そのもの以外にタグによる抽出（グルーピングや絞り込み）を実現している。

・エンリッチ技術

クローリングにより収集した情報の解析にはNTT研究所の自然言語解析技術をベースに当社独自にブラッシュアップを行ったエンリッチ技術を開発・採用している。

エンリッチ技術とは、日本語独自の文法や言い回しへの対応をベースとして、**図5**の通りその文章に書かれている内容を形態素解析で分解した上で、個々の構成要素の意味を解析し、人名であれば「名前」、日付であれば「申請日」といった解釈に基づいた結果をタグに反映させる技術の総称である。

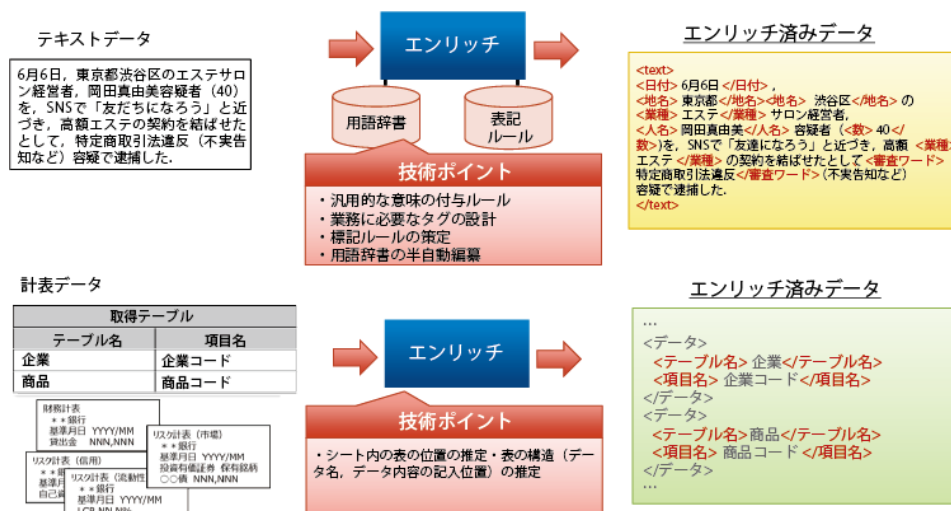


図5 エンリッチの概要

エンリッチの精度については、特に形態素解析後の個々の構成要素解釈のパートでいかに正しく単語の意味を理解し、タグに反映させることができるかが重要になる。そのために各用途に応じた辞書などをあらかじめ鍛えて用意しておくことと、そのための準備期間を用意する必要がある。一朝一夕に蓄積できるものでもないため、ソリューションとしては試行導入期間を用意し、ユーザの知見を吸い上げて反映させることが非常に重要になってくる。一言でユーザの知見といってもユーザによって特性や得意領域などが変わるため、この辞書は同じ業界、業務のユーザ同士でも最終形は大きく変わってくることを念頭に、ユーザとの意識合わせを進めることが必要になる。

・言語非依存なアーキテクチャ

今のところは日本でのみ提供しているソリューションであるため、エンリッチについても日本語解析技術としての、日本語独自の文法や言い回しへの対応に主眼を置いている。一方でグローバルなソリューションとして展開するには日本語以外の言語への対応が必要となるが、その鍵を握るのがエンリッチとなる。

多言語対応を行うには、形態素解析部分とエンリッチ部分をそれぞれ多言語でも適用可能とすることが必要になる。現状後者は各言語ごとにエンジンとその学習や辞書の用意が必要となる。その入れ

替えて機能するようなアーキテクチャを採用することで、日本語でなければ動かない、という仕組みから脱却させている。

3.4 判定

判定を実現するにあたり、以下の点に留意、工夫を行った。

- 判定ルールについては、ユーザごとに審査や判定の観点やルールは異なるため、独自に作りこめる余地を残している。
- 審査プロセスフローについては、共通的には一連の流れ（フロー）部分を用意し、それぞれ各所（収集した情報の見せ方、判定に必要とする情報、必須情報／参照情報の区分け、等）はユーザごとに設定できるようにしている。
- 審査／判定の自動化を行った。ただし、審査／判定自体の最終判断については人が行えるようにし、多段階の承認プロセス実施にも対応している。

基本思想としては極力人手は排除し、情報収集から審査／判定まで自動化を行うことで、従来の手作業による審査業務と比較し、大幅な期間短縮を実現する。一方で、最終判断についてはまだシステムのみで白黒つけられるケースは少なく、収集した情報が100%の精度でないことからくるミスジャッジなどの可能性もある。より精密な審査／判定を行うために従来業務のプロセスと同様に、本ソリューションとしても取り込み、人が最終判断を行うことを前提としている。

3.5 途上審査

途上審査を実現するにあたり、以下の点に留意、工夫を行った。

• 判定ルール

初期審査ルールと同様の判定に加え、過去の取引実績との突合による差異検出なども判定基準に入れられる。

4. ソリューション導入の効果

4.1 初期審査業務における効果

法人（加盟店）審査ソリューションを導入したことにより、初期審査業務において以下の効果が期待できるようになった。

- **申込から結果通知までの期間短縮**

情報のインプットから情報収集，システム判定までの審査業務が
一気通貫でシステム化されたことで，従来10日ほどかかっていた審
査が2～3分と大幅に短縮化された。

- **判断の根拠となる収集情報の高準化**

情報収集については一律システムが自動で実施するため，担当者
ごとのばらつきなどがなくなるとともに，インターネット上の細か
いところ隅々までカバー可能となり，高度に平準化が行えた。

- **情報収集プロセスから判定プロセスへと人的リソースの配分をシフトの実現**

効率化されたことで人手を審査する部分そのものに人的資源を割
り振ることができるようになった。

- **審査の判定結果をその根拠の記録と検索の実現**

申込書の内容や収集した情報，審査を行った結果などを証跡とし
て保存することで，いつでも振り返り可能となった。

特に情報収集についてはオフィシャル情報の収集に加えて評判情
報も探すにあたり幅広に，かつ人手ではボリュームに負けて探しき
れないものもシステムで行うことにより短時間での取得が可能とな
り，また取りこぼしの可能性が減っている。

4.2 途上審査業務における効果

途上審査業務において以下の効果が期待できるようになった。

- **定期的な全量チェックの実現**

従来は全量チェックはまったく行えておらず抽出検査（ランダム
抽出，社内アラートやニュースで確認できた特定法人の抽出）にと
どまっていたものを，スケジューリングを組んだ上で全量実施可能
とした。

- **複数世代の審査情報の保持**

途上審査で収集した情報については，初期審査で収集した情報に
加えて複数世代の審査情報を一法人あたり保持することが可能であ
る。これにより過去の審査状況の閲覧や状況変化の有無なども容易
に確認が可能とした。

5. 審査業務へのインターネット情報の活用可能性

本稿では、法人審査業務におけるインターネット情報の活用について述べた。法人審査に限らず、インターネットから収集される情報を審査の根拠として活用するには以下のような課題がある。

・変化する情報源へのユーザの理解

3章で前述した通りインターネット上の情報は常に変化し続けており、情報源としても不定であることへの理解は必要不可欠である。だからこそ審査というある断面では適切に情報を捉えることが必要であり、その瞬間の真実がそこにあることも同様に理解しておくことが重要となる。加えてその変化を追うことで不正の芽を見つけたり、ビジネスチャンスの胎動を感じることも可能であり、単にシステム観点でなく、ビジネスとして変化を追うことの必要性とチャレンジについて説明することが重要である。

・情報解析精度向上のためのユーザの知見活用

同じ業界のユーザでも得意とする領域や重点顧客、分野があり、当該分野における情報解析の観点が異なるのは一般的である。判定結果もリスクを取るユーザと取らないユーザでは同じ情報をベースにしても自ずと異なってくる。必要なのはそういったユーザごとの嗜好や特徴を知見という形で引き出し、システム上でいかに反映させるか、そのための柔軟な仕組みを作っておくかである。

6. まとめと今後の展望

法人（加盟店）審査では不定型なさまざまな情報を扱う必要があるが故にシステム化（自動化）が行えず、審査／判定に多大なる時間、人的コストがかかっていた。これに対し、法人（加盟店）審査ソリューションとして、情報収集に始まり情報解析ならびに審査／判定に至るまでの自動化とユーザごとの要件反映を行える柔軟な仕組みを実現した。その結果従来10日程度かかっていた審査／判定が2～3分で実施可能に、また途上審査の全量チェックを可能にした。

法人の審査という場面で活用が広がっている法人審査ソリューションだが適用領域が広がってきている。

金融分野に限らず法人の審査というのは一般法人でも取引先審査などどこでも当たり前に行われている業務である。

審査だけでなく法人についての情報収集と判断を行うソリューションでもあるため、営業先開拓やCRMに近い使い方でもニーズがある。

最終的には法人にまつわるデータベースに昇華するため多様なニーズへの対応を機能面でも拡充して行っていきたい。

また、日本国内だけでなく、国外への法人審査ソリューションの展開も進めている。実現には、見つける機能部では外国語サイトへの対応が、解析機能部では同様に外国語の言語解析が必要となるが、英語を皮切りに準備を進めており、グローバルに通用するソリューションとして展開を図っていきたいと考えている。

参考文献

- 1) NTTデータ：JISAアワード2017において、NTTデータの「法人（加盟店）審査ソリューション」がWinnerを受賞、
<http://www.nttdata.com/jp/ja/news/information/2017/2017061201.html>（2017年9月22日時点）

宮本 拓也（非会員）miyamotoky@nttdata.co.jp

2000年筑波大学第三学群社会工学類修了。現在、（株）NTTデータ勤務。法人審査ソリューションを含むソリューションの新規企画、営業に従事。

採録決定：2017年9月26日

編集担当：飯村結香子（日本電信電話（株））