

2016年12月

今取り組むべきデータ品質マネジメント

企業のビジネス活動において日々蓄積されるデータは「企業の資産」であるといわれて久しい。あらゆる活動で生まれた情報はデータとして組織間を流れ、蓄積され、加工され、利用される、まさに企業活動における血液ともいえるほど重要な役割を担っている。

そしてデータは、入れ物であるシステムが入れ替わっても、その価値がある限り継続して利用されるため、一般にデータの寿命はシステムの寿命よりも長い。システムの品質は寿命に伴う更改や統合により一新されるが、データの品質は時間が経てば経つほど、入力間違いや重複の発生、更新の漏れ、ルールの逸脱、法的要件や利用目的の変化などにより劣化していく。時間とともに劣化していくデータを長期間にわたり維持し利用し続けるためには、データ品質マネジメントが重要な活動となる。

データ品質マネジメントでは「継続的な品質管理」と「IT部門、業務部門、経営を含めた全社的な取組み」が重要なポイントとなる。本稿では2つのポイントを用いたデータ品質確保の手法について解説する。

1. 品質の高いデータとは

データマネジメントの国際団体であるDAMA (Data Management Association) は2009年にデータ管理の機能や活動の枠組みとしてDAMA-DMBOK (Data Management Body of Knowledge、以下「DMBOK」という)¹を策定した。このDMBOKにおいて品質の高いデータとは、「ビジネスの目的のためにデータが利用に適した状態」であると定義している。

データの品質が低ければ情報は不正確になり、その不正確な情報をもとに判断を行うことで業務効率の低下や誤ったビジネス判断を生むおそれがある。たとえば、マーケティング用の顧客情報として氏名登録する際、「斎藤」を「齊藤」と漢字変換を誤り入力してしまった結果、同一人物に複数のダイレクトメールを送付してしまう可能性がある。また、「0」「1」いずれかの数値のみが許容される項目に、外部データ統合時に「99」というルールを逸脱した数値が入力され、データ処理に異常が発生するなどの事案も考えられる。目的に沿った一定以上の品質が重要である一方で、目的にそぐわない過剰な品質を追求すると、業務プロセスやシステムの見直しを含めた品質維持のコストが増大し、結果としてデータの継続利用が困難となる可能性がある。そのため、「ビジネスの目的のためにデータが利用に適した状態」を達成する、経済的・効率的な水準を目指すことが求められる。

¹ DMBOK (Data Management Body of Knowledge)
<https://www.dama.org/content/body-knowledge>

データ品質は本来システムやデータの設計段階において、データアーキテクチャやシステム設計時に定める入出力のバリデーション等により確保されるべきである。システムが運用フェーズに入り、時間の経過とともに劣化するデータに対しビジネスが求める品質を維持し続けるためには、データ品質マネジメントが必要となる。

2. データ品質マネジメント

データ品質を維持するためには、品質の基準とルール・プロセスを作り、そのルール・プロセスに沿って組織的な取組みを継続して実施していくことが重要である。なかでも、データ品質基準や例外の取扱いについては、データを利用する業務部門が主体的に定めるべきである。

DMBOK²において、データ品質マネジメントは「品質管理技術を応用して、データが用途に適しているかを測定・評価・改善し、保証を行うための各種活動について計画立案、実行、統制を行う」機能として定義されている。データ品質マネジメントに基づく活動により、ビジネスの要求に沿ったデータの品質改善を促すものである。

データ品質マネジメントにおいても、品質管理におけるデミングのサイクル(Plan-Do-Check-Act)と同様に、次のような管理サイクルが提唱されている。

(1) データ品質基準の策定

データの利用目的を確認したうえで、その目的に合わせた品質基準を策定する。まず、データの利用目的を判断できる業務部門を交え、ビジネスの要求に沿った評価軸を定める。次に、評価軸に沿って「データが利用に適した状態であること」を判定するためのビジネスルールを定め、データ品質基準として策定する。DMBOKに紹介されるデータ品質の評価軸には、以下のようなものがある。

【図表1】 データ品質の主要な評価軸

| 評価軸 | 概要 |
|--------|---|
| 正確性 | データが表そうとしている実体が正しく示されていること |
| 完全性 | すべてのデータ要素が揃っていること |
| 一貫性 | 同じ実体を表す2つ以上のデータに不整合がないこと |
| 最新性 | データが期限内の実体を示していること |
| 精度 | データの詳細度（有効桁数など）が十分であること |
| プライバシー | アクセス制御と利用監視がなされていること |
| 妥当性 | 対象の業務内容においてデータの整合性が取れていること |
| 参照整合性 | 参照元のデータが存在すること |
| 適時性 | 必要な時に速やかにデータが利用できること |
| 一意性 | 同じ実体を表すデータが1つだけ存在すること |
| 有効性 | データが定められた属性（型・形式・精度・文字コード等）が有効範囲に収まっていること |

出典：DAMA-DMBOK Functional Framework Version 3.02 (DAMA International, September 2008) を基に KPMGが作成

2 DMBOK には本稿に紹介した活動のほか、データ品質マネジメントに係る「問題の管理」「運用手続きの設計と実行、パフォーマンス監視」「データ品質管理ツール」等に関する説明が記載されている。

(2) データ品質の評価と測定プロセスの策定

定めたデータ品質基準に対して、現状データの形式、内容、構造などがどのような状態であるのかを評価して品質の問題を定量的に認識し、品質測定のプロセスを策定する。

データの評価は、対象データの件数や品質基準の複雑さ等により多くの時間がかかるが、統計的技法を用いた専用ツール（データプロファイリングツール）による時間の短縮も可能である。

この時、データ品質を測定するためのプロセスと、データ修正の手続きを策定する。品質測定のプロセスは、データ品質基準として定めたビジネスルールへの適合状況を測定する手順であり、これと合わせてデータ品質基準に対する許容しきい値を定め、品質監視の指標とする。

(3) データ品質不具合の対処

データ品質測定基準に沿った品質測定の結果をもとに、前述した手続きに沿ってエラーの修正を行い、期待する品質に適合させていく。

データの修正には専用ツール（データクレンジングツール、名寄せツール等）による自動修正と人的作業による手動修正がある。専用ツールを使ったとしても、重要なデータや修正内容の信頼度が低い場合は、人的作業による確認が必要である。データ修正を行う場合は、できる限りデータ発生源に近いポイントで修正を行うことで、修正の対象範囲を小さくすることができる。

また、不具合の対処には個々のエラーの修正だけでなく、新たなエラーの発生を防ぐことも含まれる。そのためには、エラーが発生した原因を特定し、原因に対する対処として業務プロセスや情報システムの見直しを行う仕組みが重要である。

(4) データ品質の監視

定期的に品質測定とデータ修正を行い、データ品質を維持されていることを継続監視する。データ品質を測定し、定めた品質測定基準の範囲内であれば、品質管理のプロセスやデータの品質レベルが利用目的の許容範囲に収まっていると判断できる。一方、定めた品質を下回る場合は、先の「データ品質不具合の対処」を行う。

データの利用目的や外部要件の変化により、データ品質に対する期待が変わる場合は再度、「データ品質基準の策定」からプロセスを繰り返す。

3. まとめ

企業のビジネス活動においてデータは無尽蔵に増加し続け、また、時間とともにそのデータ品質は必ず劣化する。ビジネスに活用できるデータとして取り扱うため、データ品質管理による「品質を作り、品質を維持するための継続的な活動」が不可欠であるが、その活動には時間とコストがかかる。そのため、多少の品質劣化は黙認し、当面の業務を進めていくことが多い。しかし、データ品質の問題を先送りすればするほど劣化は進み、その改善にはより多くのコストや時間が必要となる。

データ品質マネジメントはIT部門の努力だけではなく、データを利用し活用する業務部門の協力や、データのあり方に対する経営層の意識を高めることも必要である。例を挙げると、データ入力を行う業務部門に「正確に入力することの重要性」を理解させることや、経営層のデータ品質マネジメントに係るコストへの理解を求めていくことなどである。データ品質マネジメントの意味と重要性を全社で共有し、一丸となって取り組むことが大切である。

KPMGコンサルティング株式会社
マネジャー 河合 博範

KPMGコンサルティング株式会社

東京本社
〒100-0004
東京都千代田区大手町1丁目9番5号
大手町フィナンシャルシティ ノースタワー
TEL : 03-3548-5305
FAX : 03-3548-5306

大阪事務所
〒541-0048
大阪市中央区瓦町3丁目6番5号 銀泉備後町ビル
TEL : 06-7731-2200

名古屋事務所
〒450-6426
名古屋市中村区名駅3丁目28番12号 大名古屋ビルヂング
TEL : 052-571-5485

kpmg.com/jp/kc

ここに記載されている情報はあくまで一般的なものであり、特定の個人や組織が置かれている状況に対応するものではありません。私たちは、的確な情報をタイムリーに提供するよう努めておりますが、情報を受け取られた時点及びそれ以降における正確さは保証の限りではありません。何らかの行動を取られる場合は、ここにある情報のみを根拠とせず、プロフェッショナルが特定の状況を綿密に調査した上で提案する適切なアドバイスをもとにご判断ください。

© 2016 KPMG Consulting Co., Ltd., a company established under the Japan Company Law and a member firm of the KPMG network of independent member firms affiliated with KPMG International Cooperative ("KPMG International"), a Swiss entity. All rights reserved.

The KPMG name, logo are registered trademarks or trademarks of KPMG International.