

データマネジメントについて ～現場で使えるデータマネジメントの実際



2015年1月21日

DAMA日本支部

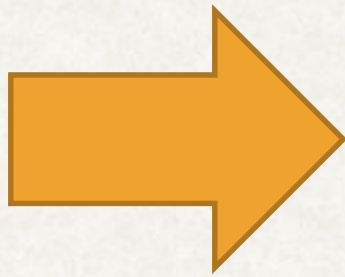
赤 俊也

seki@dama-japan.org/

データ分析大流行・・・でも

★その前に・・・

分析しようとしているデータは
「正しい」の？



- そもそも分析をしているデータの元となる数値が正しく計測できているのか
- 参照するデータが正しいのか
- 予期せぬデータが混じっていないか
- そのデータの定義は「正しい」か

データマネジメントの重要性！

DAMA

Data Management Association International (DAMA-I) は、世界各地に約40の支部を持ち、7,500名を越える会員を擁する全世界のデータ専門家のための国際的な非営利団体。

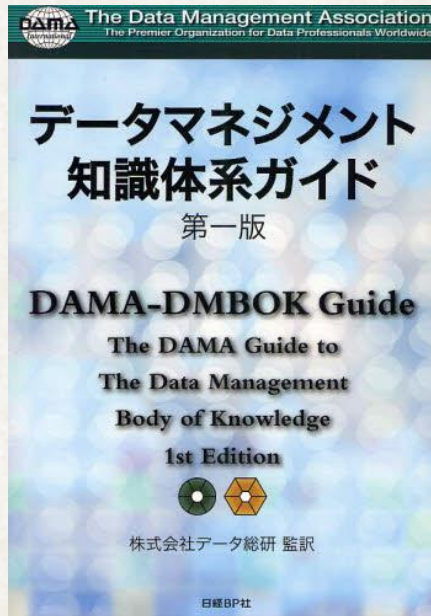
特定のベンダーや技術、手法に依存しないことを大前提として、データや情報、知識をエンタープライズの重要な資産として管理する必要性の理解を促し、この分野の成長を推進している。我が国においても、2010年11月にDAMA-IのJapan Chapterが設立された。

(DAMA日本支部 代表 松本 聡)



DAMA-DMBOK

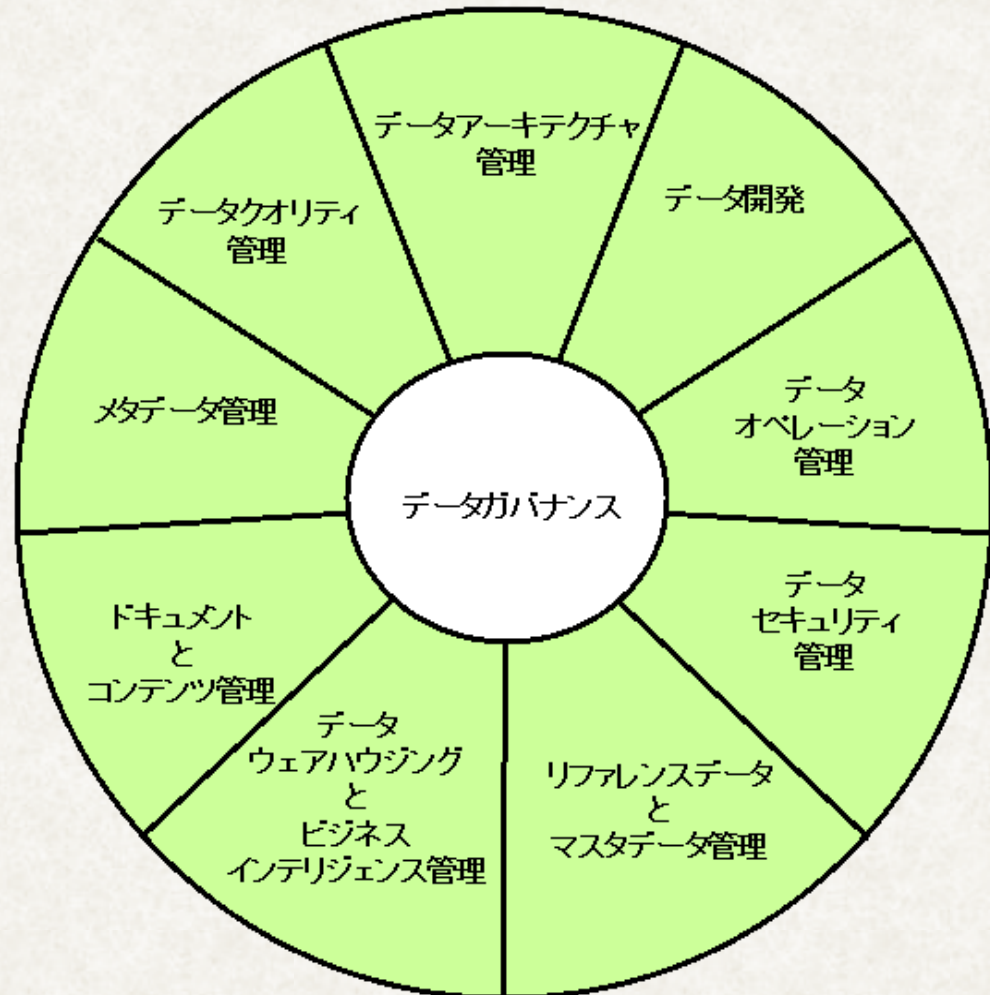
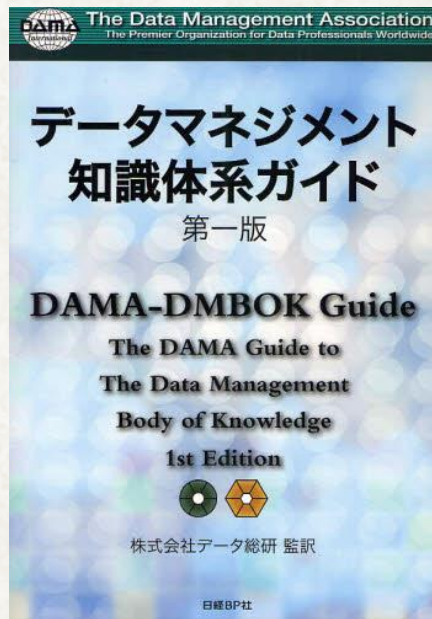
©DMBOK (Data Management Body of Knowledge) Guide



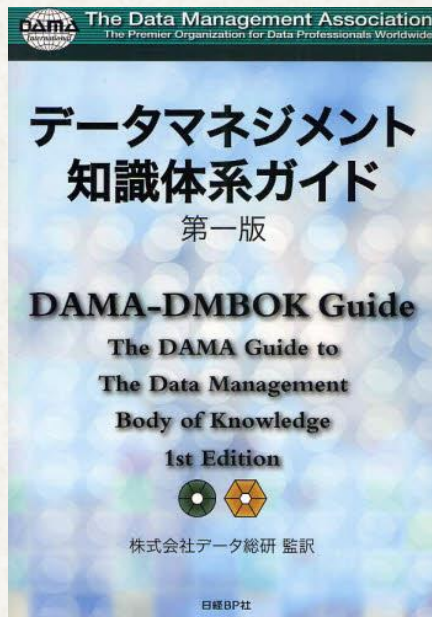
DAMAインターナショナルが編纂したデータマネジメントに関する知識体系のこと。組織が継続的なデータマネジメントを推進するに当たって、必要となる考え方や組織体制、管理機能、用語、ベストプラクティスなどをまとめている。

データマネジメントの要素（１）

★10のデータマネジメント要素★



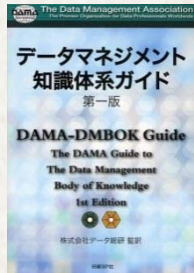
データマネジメントの要素（２）



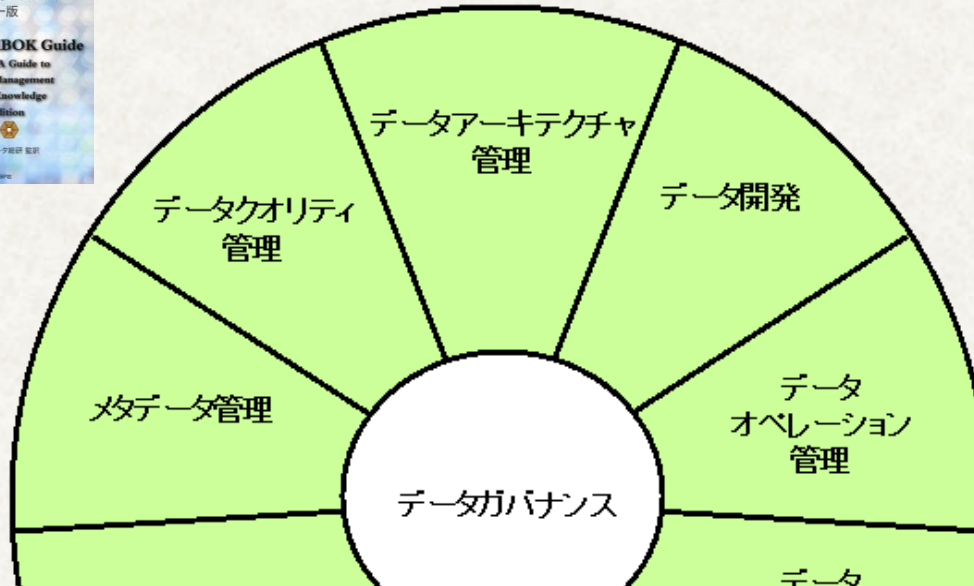
★10のデータマネジメント要素★

- ・ データガバナンス：
データの管理および利用にまつわる計画、監視、統制。
- ・ データアーキテクチャ管理：
データ資産管理のための青写真作成。
- ・ データ開発：
分析、設計、実装、テスト、配備、維持。
- ・ データオペレーション管理：
データ収集から廃棄までのサポート提供。
- ・ データセキュリティ管理：
プライバシー、機密性、適切なアクセスの保証。
- ・ データクオリティ管理：
データクオリティの定義、監視、改善。
- ・ リファレンスデータとマスタデータ管理：
ゴールドバージョンと複製バージョンの管理。
- ・ データウェアハウジングとビジネスインテリジェンス管理：
報告と分析を可能にする。
- ・ ドキュメントとコンテンツ管理：
データベースの外部にあるデータの管理。
- ・ メタデータ管理：
メタデータの統合、統制、供給。

どこから手をつける？



★10のデータマネジメント要素★



- ・データガバナンス：データの管理および利用にまつわる計画、監視、統制。
- ・データアーキテクチャ管理：データ資産管理のための青写真作成。
- ・データ開発：分析、設計、実装、テスト、配備、維持。
- ・データオペレーション管理：データ収集から廃棄までのサポート提供。
- ・データセキュリティ管理：プライバシー、機密性、適切なアクセスの保証。
- ・データクオリティ管理：データクオリティの定義、監視、改善。
- ・リファレンスデータとマスタデータ管理：データの管理。

■どこから手をつければよいか？

⇒まずは「コトバ」の定義から！

報告と分析を可能にする。
データベースの外部にあるデータの管理。
合、統制、供給。

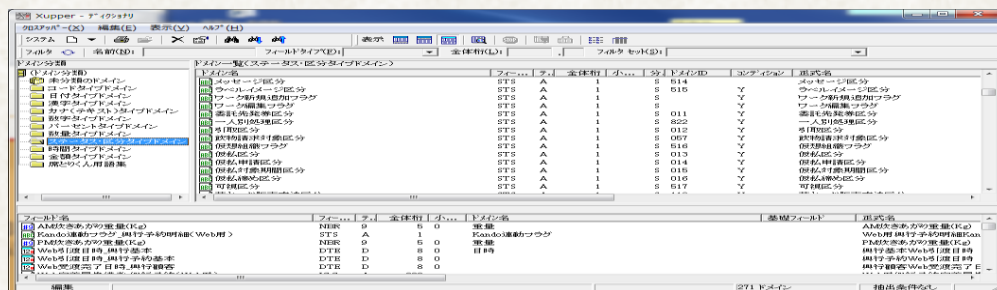
用語集を整備しよう！

■業務改革だろうが、システム開発だろうが
ビジネス「用語集」の整備は重要。

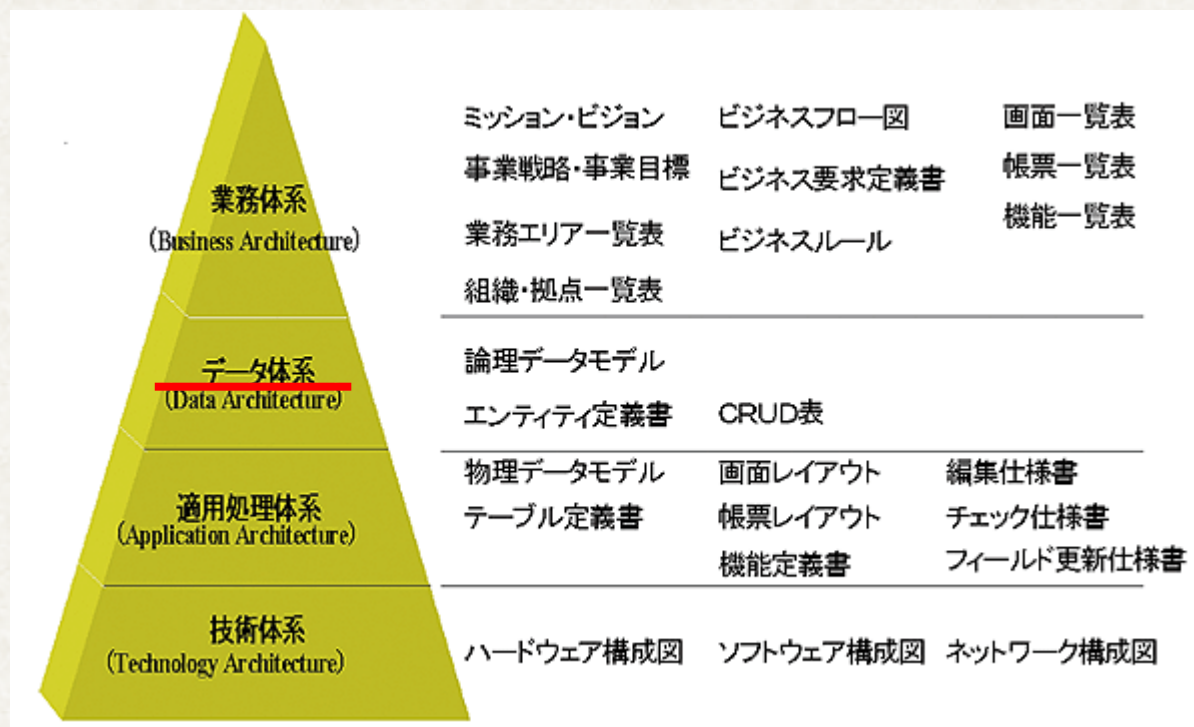
(後回しになりがちだが・・・)

■システムライフサイクルの開発フェーズにおいても保守（稼働後投資回収）フェーズにおいても、
データ・プロセスの品質確保に役立つ。

⇒無駄を省く！

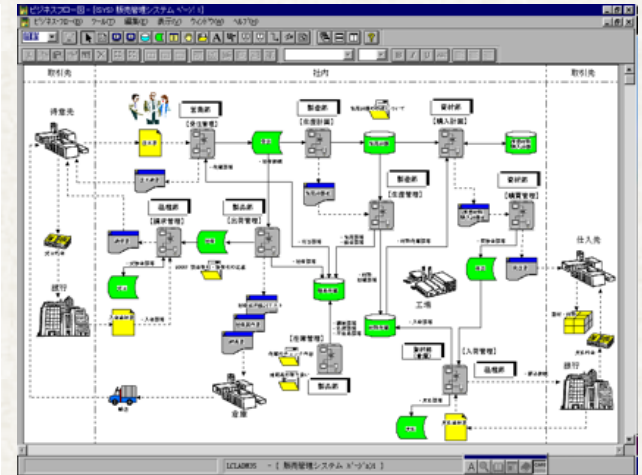
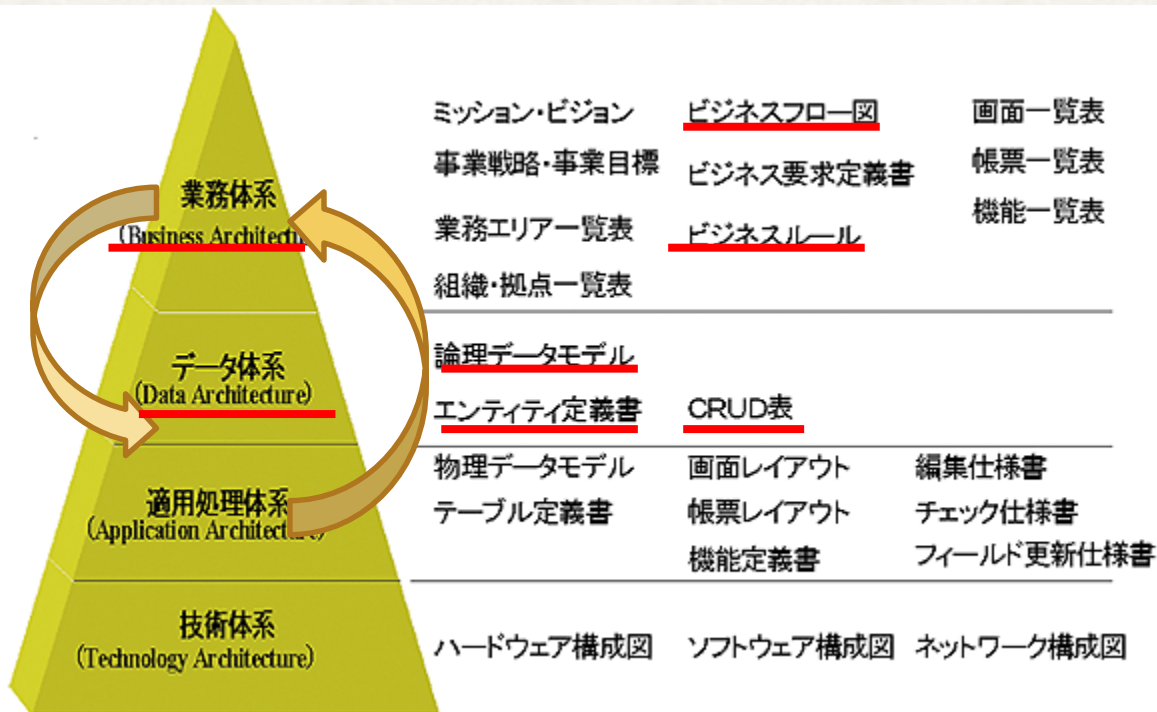


エンタープライズアーキテクチャにおけるデータ



■ システム云々以前にビジネス、
業務を考える上で「データ」は
無視できない！

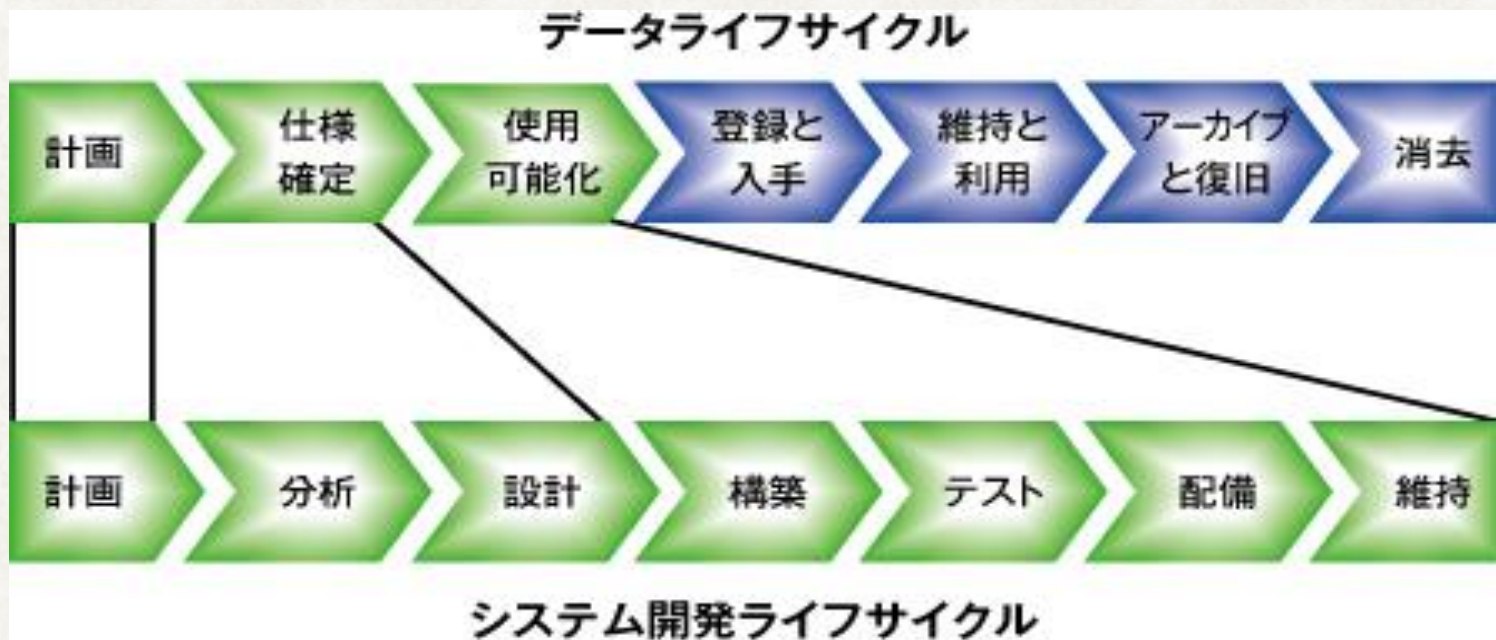
システム開発においても・・・



■システム開発の上流工程において、
「業務」「データ」として抑えるべき
ものがきちんと抑えられていなければ、
結果は・・・

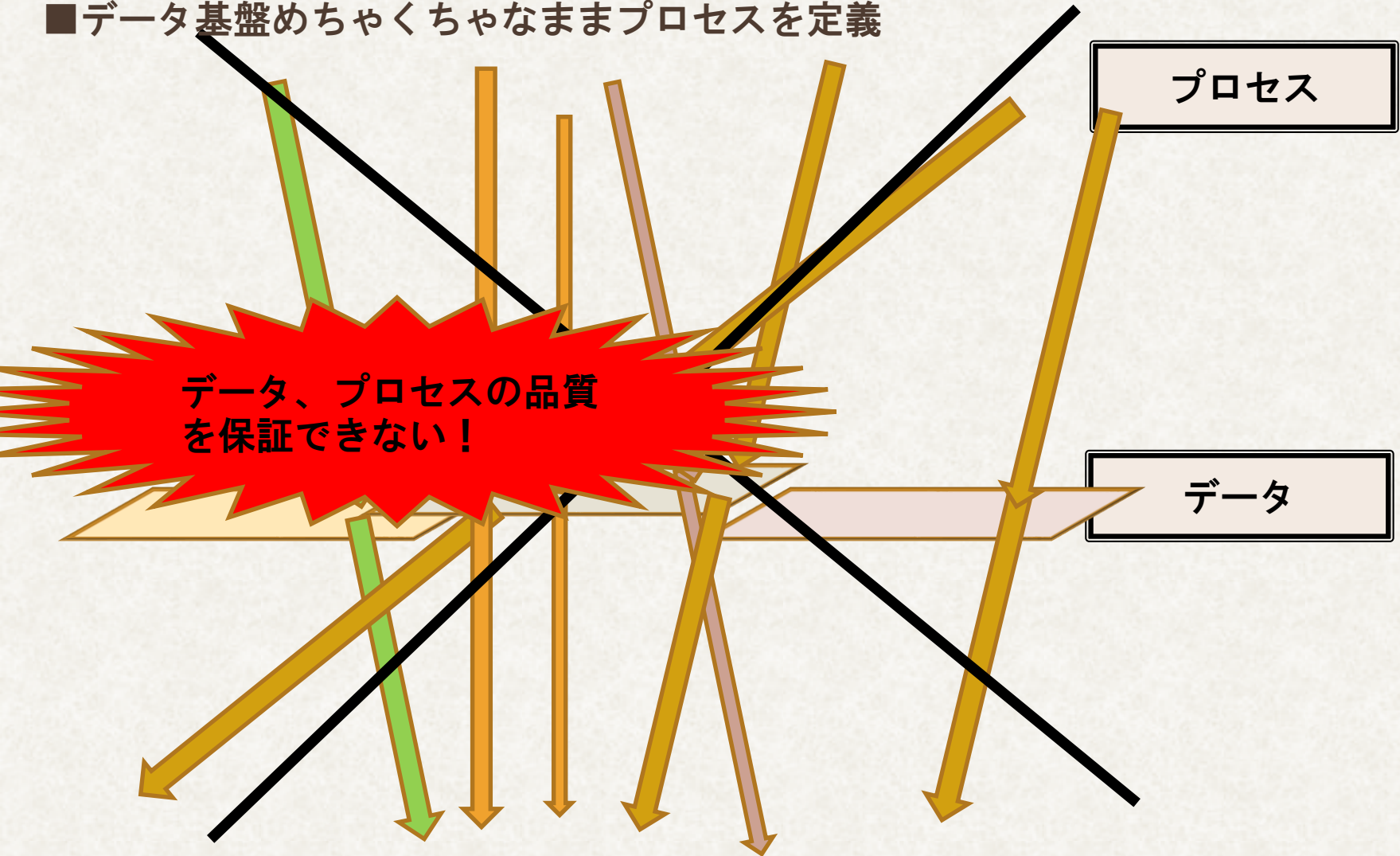
システムライフサイクルとデータライフサイクル

- データライフサイクルはシステムライフサイクルより長い。
データ基盤を整備しつつ開発を進めることにより、
結果的にシステムライフサイクルが長くなる。



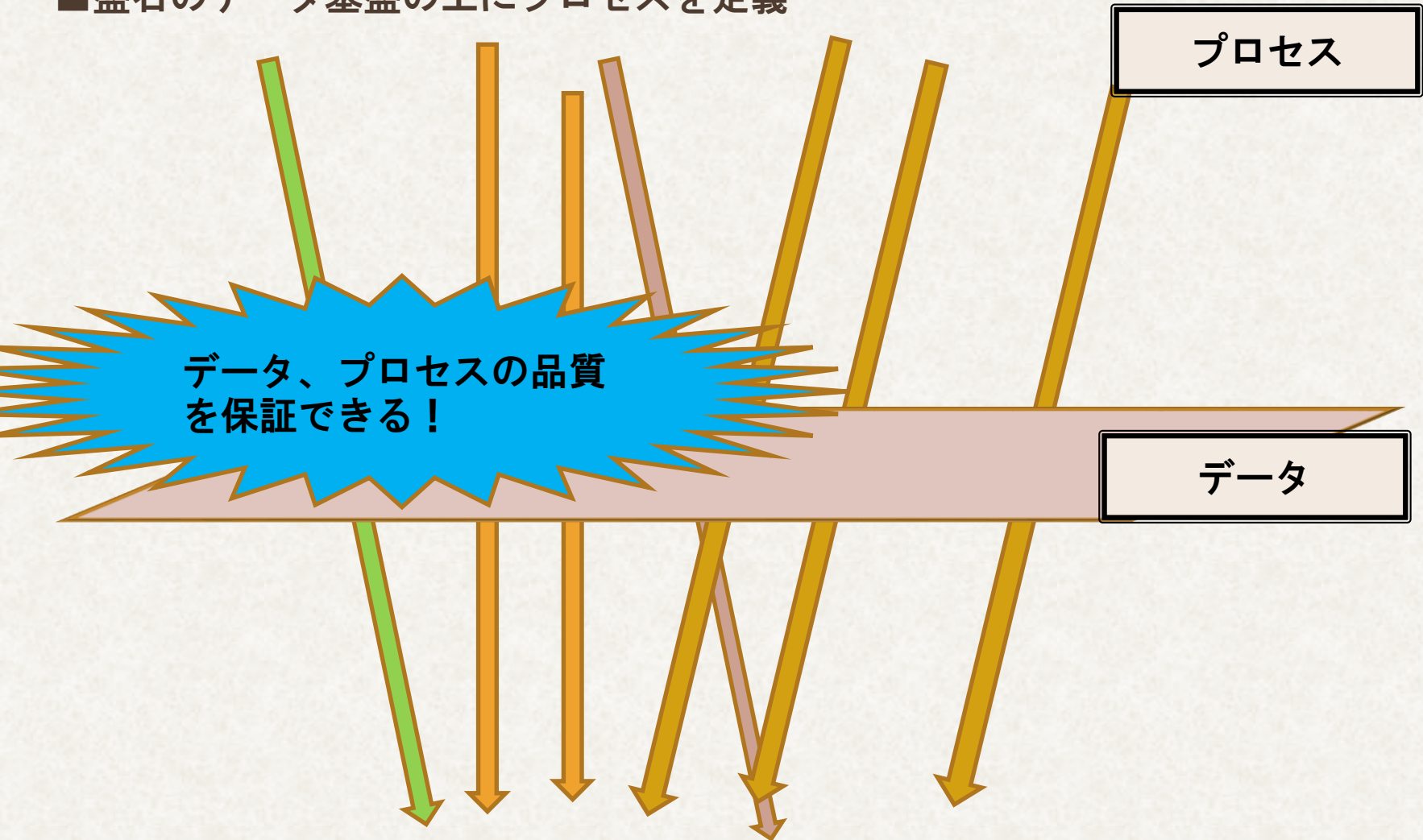
データとプロセスの最悪の関係

■データ基盤めちゃくちゃなままプロセスを定義



データとプロセスの良好な関係

■ 盤石のデータ基盤の上にプロセスを定義

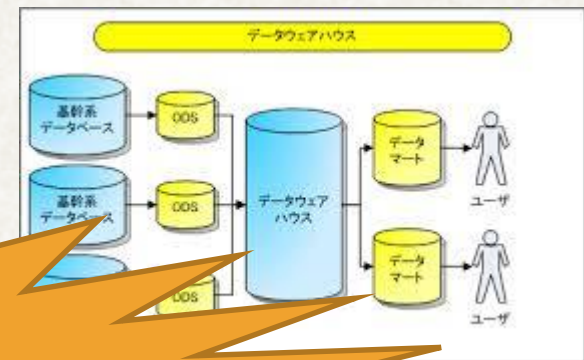


自然に正しいデータが・・・

■自然に正しいデータが溜まる「仕組み」が必要

■じゃあ、どうやって？

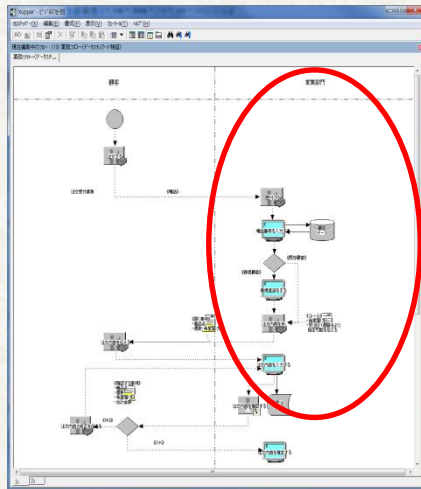
⇒これが言うは易し、行うは・・・



★上流工程から
データマネジメントの
概念をすりこんでおく！

データ品質が大事！

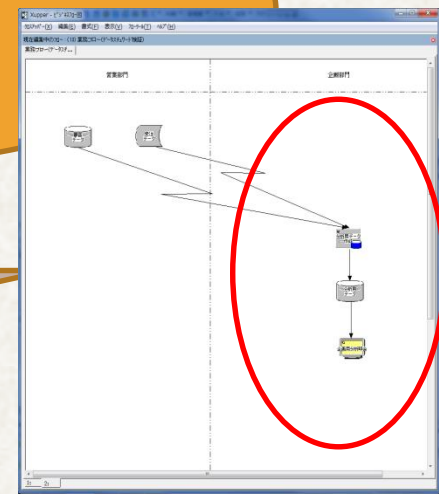
■高品質のプロセスにより生成されたデータは、企業組織の資産になりうる。



生成するプロセス

データ

活用するプロセス



データ品質が引き起こした問題（事例１）

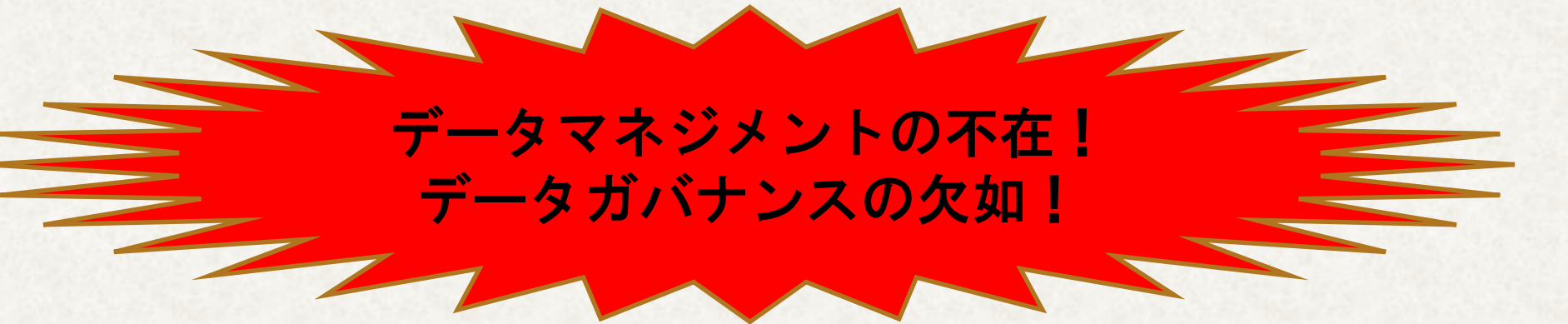
- 火星探知機：1999年に、1億2500万ドルの宇宙船は4億マイルを成功裏に旅してきたが、NASAがヤードポンド法の単位をメートル法に変換するのを忘れたために人為的に失敗した。

メタデータ（データ定義）
の欠如



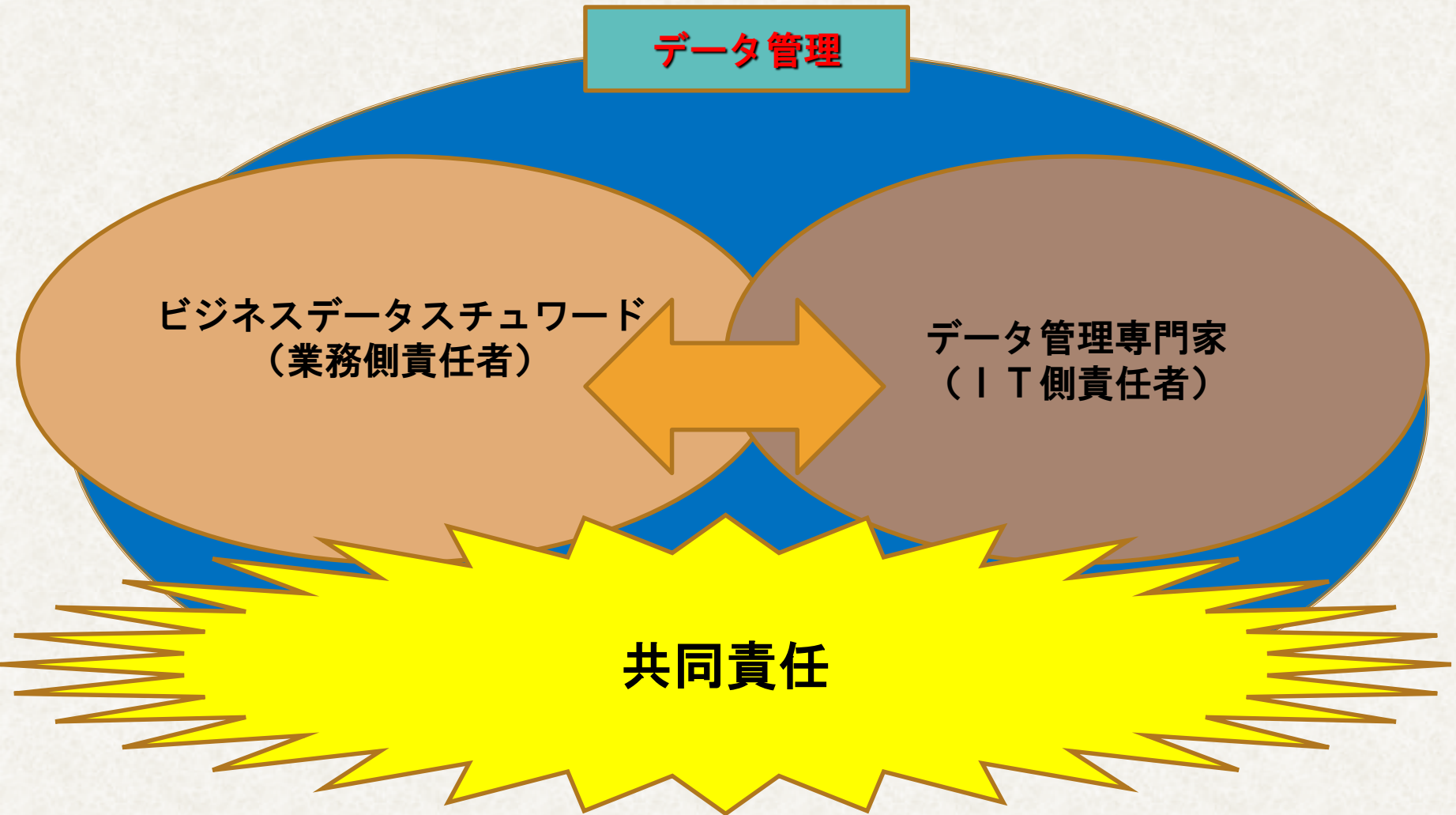
データ品質が引き起こした問題（事例２）

サブプライム問題：信用度の低い借り手に対して住宅ローンを提供し、そのローンを金融機関に販売できるようにした。そして金融機関は住宅ローンを他の債権と組み合わせて適切なリスクとリターンを設定した証券として投資家に販売した。この過程におけるあらゆる段階で不適切なデータと情報が破滅の要因となった結果、数千人が家を失い、多くの企業が破産を申請し、金融機関が多大な資産評価損を計上した。

A large, red, jagged starburst graphic with a gold outline, containing the text.

**データマネジメントの不在！
データガバナンスの欠如！**

現場とIT担当がデータ管理を共同で行う！



DMBOKより

A group of eight young people are captured in mid-air, jumping joyfully on a lush green lawn. They are dressed in casual summer attire like t-shirts, tank tops, and jeans. In the background, a calm body of water stretches towards a distant shoreline under a bright, slightly cloudy sky. One person is crouching on the grass in the foreground, looking up at the group. The overall mood is one of happiness and celebration.

ありがとう
ございました！