

生々しい本音が書かれていて面白い。

分析の報告を受ける側から詰め寄られたり、一緒に同行したビジネス側の人間にハシゴを外されたり、さまざまな人間ドラマが垣間見れるのが分析結果報告会議です。会議に参加した人たちを「全員悪人」だと感じた経験は一度や二度ではありません。

松：ポジショントークは一切無しこの本持って、いろんな会社で革命起こしましょう！怒りをぶつけるんです。上司に叩固め、経営者に延髄斬り、マスコミにジャーマン・スープレックス！

マ：本気かい…ええ？俺たちは炎上を恐れちゃいけねえ。

松：本気の…つもりです。この本は、夢も理想も語りません！極めて実用的で、**現場目線**だけど、一番しっかりしています。上司だけじゃない、経営者やマスコミにこそ読ませたいです！

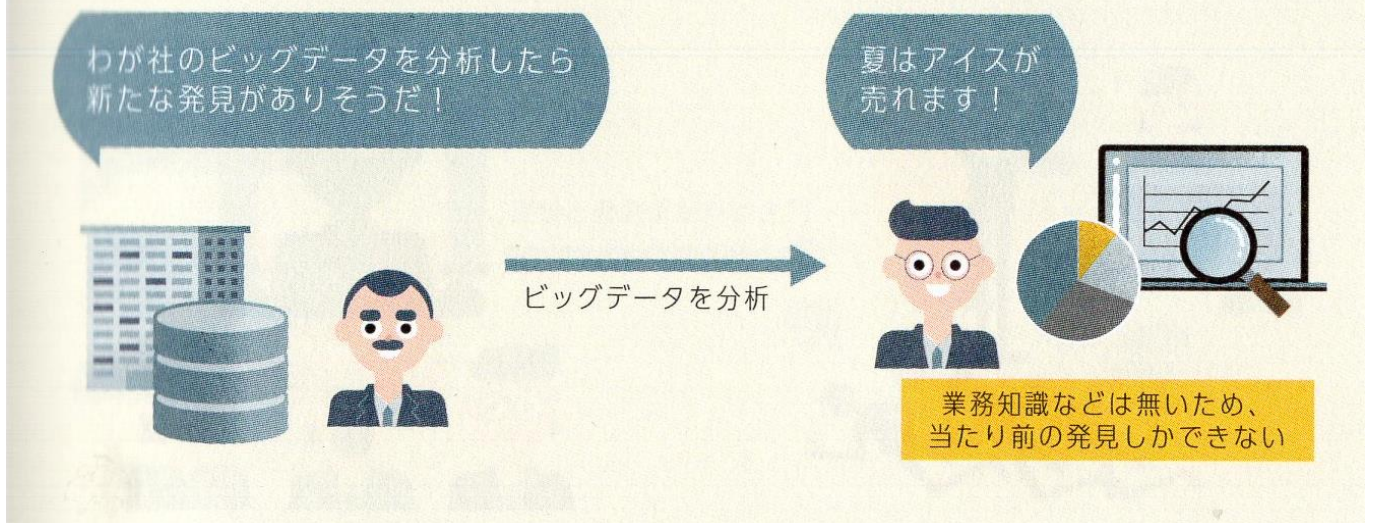
マ：命をかけたのか、命を。上司も経営者もマスコミも**既得権益**なんだけ、お前。

松：これまでの ERP、BI ツール、ユビキタス、Web2. 0、ビッグデータ、IoT、ブロックチェーン、そして今は RPA で、次は量子コンピューターなんですよ。

マ：だったらどうして Excel で仕事してるんだっ！

データサイエンティストは必ずしも業務知識やビジネスセンスを持っているわけではありません[01]。

[01] データサイエンティストが分析すると…



研究開発において短期的な成果の重視や、前例主義や縦割り組織のカニバリゼーションに囚われては何もできません。作業着を着せて、低スペックなパソコンを支給し、セキリテイルールに縛られ、安い机と椅子で仕事をさせる会社では、永遠にスティーブ・ジョブズは生まれません。

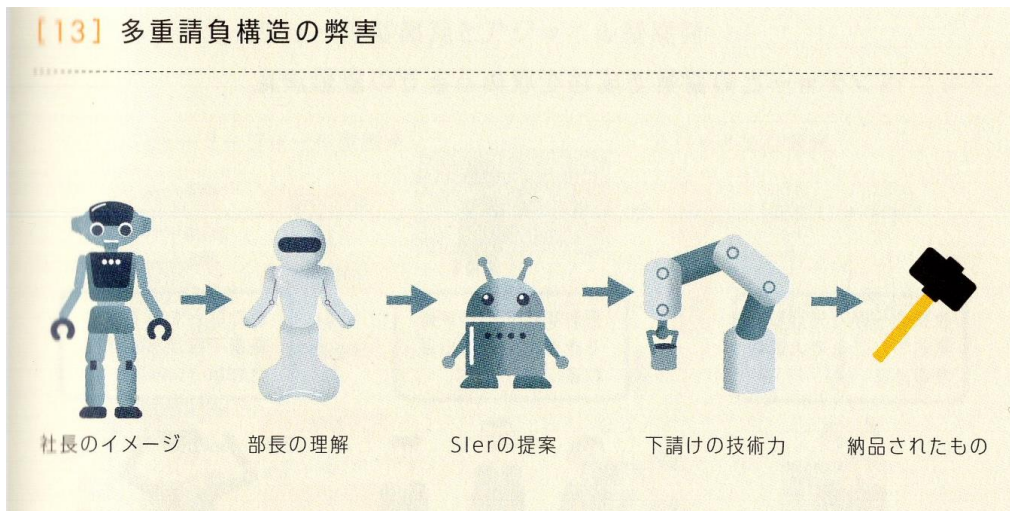
[09] 監視と管理の仕事環境では何も生まれない



◆業務知識という闇

SIer が重視する「業務知識」も、このような状況を踏まえてブラックボックス化が進行しています。長年現場のルールで運用された自社向けの基幹システムなどは、担当者の頭の中が仕様となって「秘伝のタレ」になっているのです。

運用で曖昧さをカバーしたり、「ウチは特別」という思想を持って自社独自のやり方にこだわりすぎたシステムでは、データサイエンスや AI は活用できません。基幹システムを最初から作り直すのは現実的ではない以上、システムを利用する人間から変わらなければなりません。「業務知識」が明文化されていなければ、AI 以前に業務システムさえも役に立たなくなってしまいます [13]。



~~~~~

SIer とはシステム構築する会社のこと

読みは「エスアイアー」です。和製英語ですので日本特有の表現となります。

~~~~~

前提として、AI やデータサイエンスの世界では、これまでの業務システム開発における常識が通用しません。技術力を無視してエンジニアの頭数だけを揃えて、マネジメントでカバーしたところで、いつまで経っても AI は完成しません。

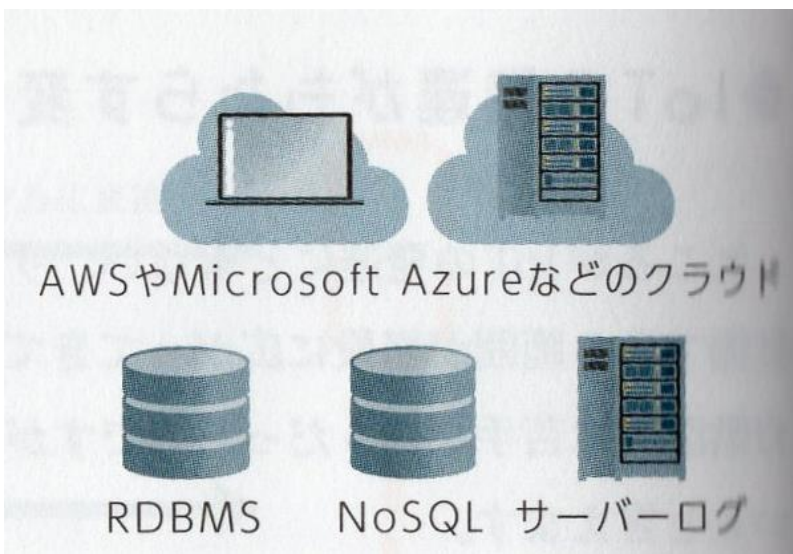
...

今後 AI ・データサイエンス型のサービスを提供するのであれば、製造業をモデルにした従来型の組織・開発・人材・ビジネスモデルから大転換しなければなりません。今後は内製化・クラウド化・SaaS 型への転換が進むため、外注・多重請負・労働集約の従来型システム開発のビジネスは成立しません。

データサイエンスビジネスは、まず何を知りたいのかを定義することが必要です。

		分析したら…	
		わかった	わからなかった
依頼した側が…	知っている	ここが多い！	データに現れない暗黙知
	知らなかった	本来期待される	

データを収集・蓄積する。



~~~~~

### RDBMS 【 Relational DataBase Management System 】

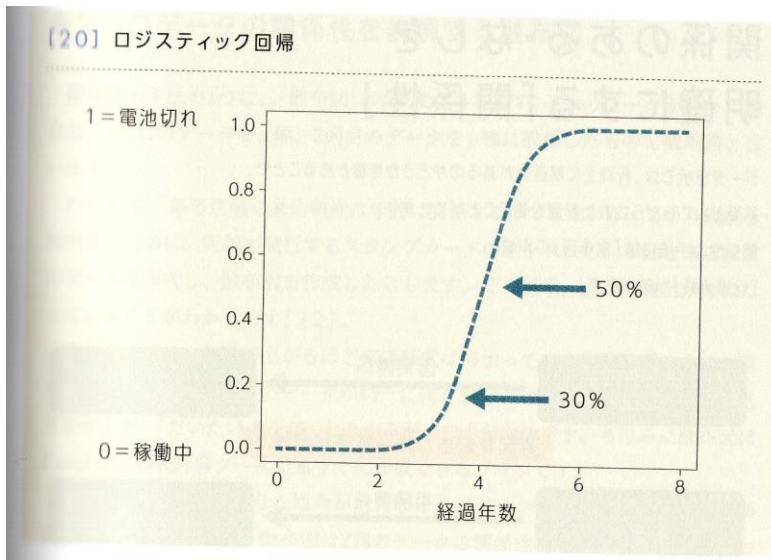
RDBMS とは、リレーショナルデータベースを管理するための専用のソフトウェア。

リレーショナルデータベース（RDB : relational database）は一件のデータを複数の属性の値の組として表現し、組を列挙することでデータを格納していく方式で、属性を列、組を行とする表（テーブル）の形で示されることが多い。

~~~~~

たとえば、乾電池を 2 個使う懐中電灯を何年使い続けると電池切れになるかを集計したとします。集計の結果、右図 [20] のように、だいたい 4 年ぐらいは使えて、それ以降は電池切れの確率が高くなるという結果が出ました。では、4 年半の時点で電池切れが起こる可能性は何%になるのでしょうか？ それを知るのがクラス分類の手法の 1 つである「ロジスティック回帰」です。

[20] ロジスティック回帰



多重共線性 (通称マルチコ)

プログラムを使って一つ一つに指示を出さずとも、データの傾向から機械がルールや判断基準を抽出して、アルゴリズム化してしまう学習手法を「機械学習」と呼びます。

~~~~~

### アルゴリズム 【 algorithm 】

アルゴリズムとは、ある特定の問題を解く手順を、単純な計算や操作の組み合わせとして明確に定義したものの。数学の解法や計算手順なども含まれるが、IT の分野ではコンピュータにプログラムの形で与えて実行させることができるよう定式化された、処理手順の集合のことを指すことが多い。

~~~~~

ディープラーニングとは、従来は人間が与えていた特徴量 (特徴) を自ら探す学習法です。

R は **Python** パイソンに比べて計算が得意で、統計学のアルゴリズムへの対応範囲が広く、可視化もしやすいのが特徴です。一方で、**Python** は **R** のようにデータ分析に特化していません。その代わりに、システムに組み込んで複雑な処理を自動化できるのが特徴です。

BI ビジネス・インテリジェンス

1371 使いやすいBIツール

	特徴	価格
Qlik Sense	セルフサービス型のBIツール。直感的な操作でデータを可視化したり分析したりできる。マルチデバイスに対応しているため、場所を問わず作業ができる。	Cloud Basic : 無料 Cloud Business : \$15 / 月 ※価格は 1 ユーザーあたり
Microsoft Power BI	セルフサービス型のBIツールで、ノンプログラミングでデータ分析が可能。各種データ処理のほか、可視化されたデータを組織で共有することができる。	Standard Edition : 30,000 円 / 月 Professional Edition : 60,000 円 / 月 IoT Edition : 90,000 円 / 月 ※価格は 10 ユーザーあたり
Tableau	低コストで導入できるノンプログラミングのBIツール。グラフなどの多彩な表現によって、誰もが理解しやすい形でデータを可視化し、共有することができる。	Tableau Creator : 102,000 円 / 年 Tableau Explorer : 51,000 円 / 年 Tableau Viewer : 18,000 円 / 年 ※価格は 1 ユーザーあたり
MotionBoard	データの可視化に加え、ビジネスの状況をリアルタイムに把握できる。専門知識が不要で直感的に操作でき、さまざまなチャートで表現できる。	Power BI Pro : \$9.99 / 月 Power BI Premium : \$4,995 / 月 ※価格は 1 ユーザーあたり

分析の報告を受ける側から詰め寄られたり、一緒に同行したビジネス側の人間にハシゴを外されたり、さまざまな人間ドラマが垣間見れるのが分析結果報告会議です。会議に参加した人たちを「全員悪人」だと感じた経験は一度や二度ではありません。

アドエビスは年間 120 億以上の国内のアクセスログデータおよびサーブドパーティデータと独自開発の AI を使って、デバイスやブラウザーが違っても、同一ユーザーであることを類推する技術を開発しました。

TKO (低姿勢、こまめ、おだてる)

【PoC】「Proof of Concept」の略称で、「概念実証」を意味しています。本格的にプロジェクトを運用する前に、目的に対して実現可能かどうかや、効果が得られるのかどうかを検証するための工程です。さまざまな開発プロジェクトで PoC が行われていますが、とくに IT や新技術を活用したプロジェクトにおいては、PoC の実施は不可欠であると言えます。

【アノテーション】「注釈」を意味する言葉で、IT の分野においては、あるデータに関連する情報を注釈として付与する作業のことを意味します。とくに機械学習においては

欠かせない処理です。たとえば、猫の画像であれば「猫」というラベルを、犬であれば「犬」というラベルを付けることで、パターンを認識できるようになり、正しい判断や予測が可能になります。