



# PoCで終わらせない！ データ活用を成功させる秘訣

BULB株式会社

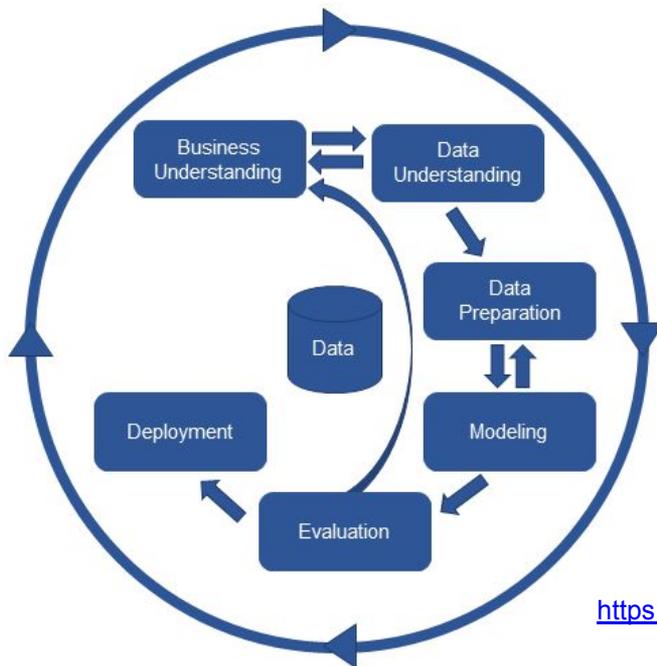
足立 悠

# 目次

- データ分析あるある
- **まずはPoCから始める**
- **知っておきたいキーワード**
- 分析力を手にするために
- 参考情報

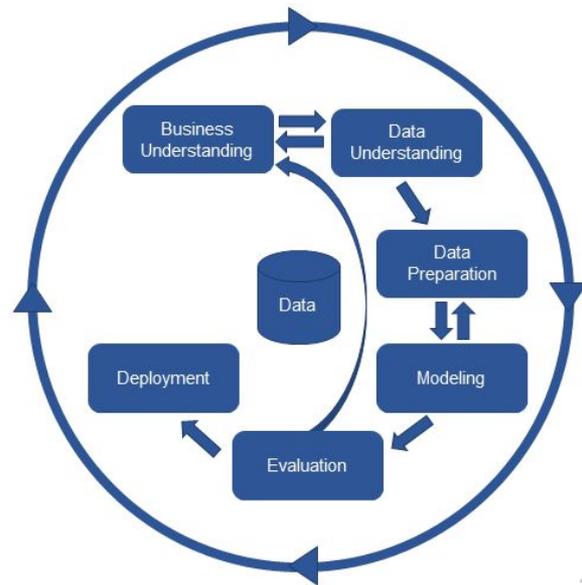
データ分析の  
フレームワーク  
**CRISP-DM**  
(**C**ross Industry  
**S**tandard Process  
for **D**ata **M**ining)

- データ分析はROIを計算しにくい
- 試行錯誤を繰り返し小さい成果を積み重ねていく進め方がおススメ
- CRISP-DMに沿って進めると成果を得やすい



# 各フェーズの作業

1. **ビジネス理解**
2. データ理解
3. データ準備(前処理)
4. モデル作成
5. 評価
6. 共有／展開



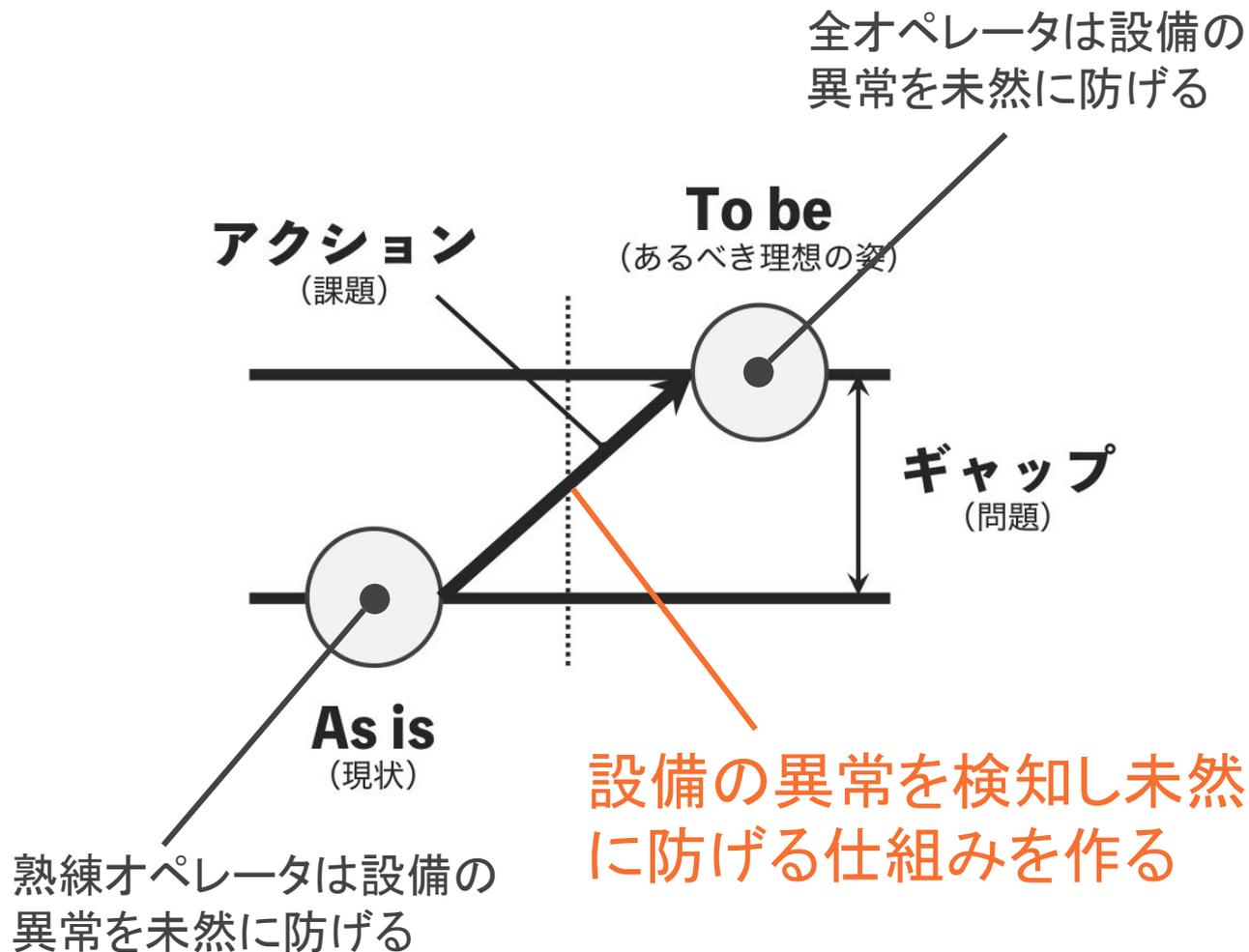
## 製造業のよくある悩みごと



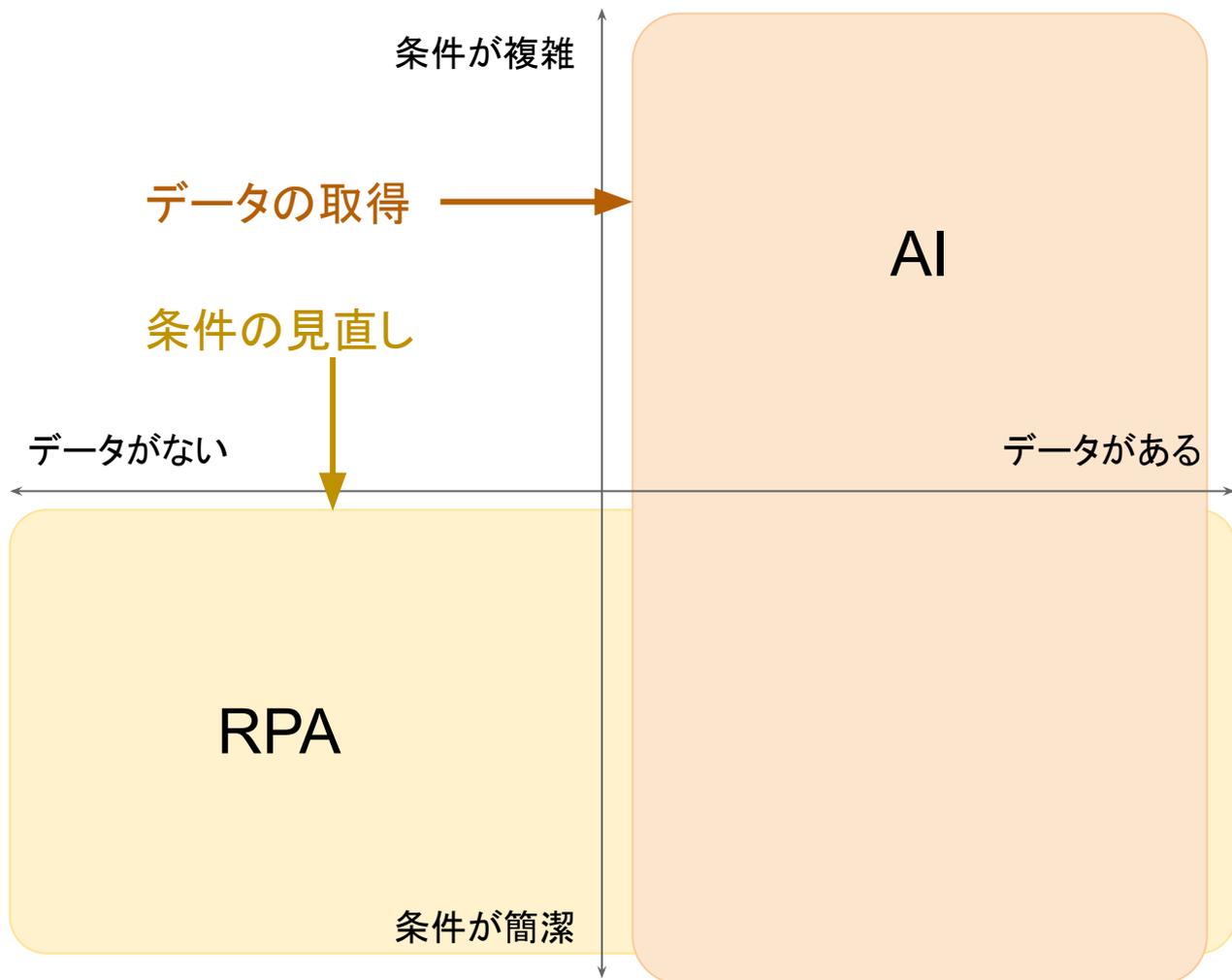
- 計画外の設備停止が多く製造ラインの稼働率が頭打ち。少しでも**稼働率を向上**させたい。
- 設備の定期メンテを状態に応じて最適化し**コストを削減**したい。
- 熟練エンジニアの定年退職など、今後人手不足になるのは間違いない。**誰でも対応できるような仕組み**がほしい。

# 課題の定義

## 製造業を例に

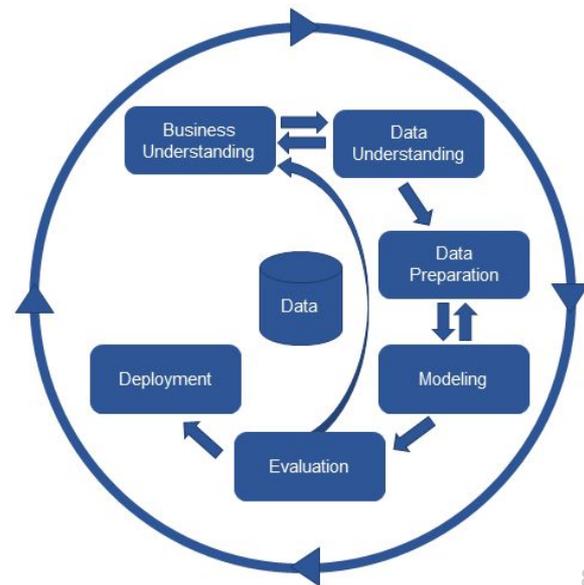


# AIで解くか RPAで解くか



# 各フェーズの作業

1. ビジネス理解
2. データ理解
3. データ準備(前処理)
4. モデル作成
5. 評価
6. 共有／展開



前から一人の人間(大人)が歩いてきます。  
男性でしょうか？女性でしょうか？



ただし **データだけで判断** して下さい

## 前から歩いてくる人の情報(データ)

年齢	43歳
出身地	東京都
身長	162cm
体重	58kg
右視力	1.2
左視力	1.0
好きな番組は	バラエティ
よく聞く音楽は	サザンオールスターズ

## 前から歩いてくる人の情報(データ)

髪の毛の長さ	38cm
スカートを履いている	YES
ハイヒールを履いている	YES
喉仏の凹凸	小



The image features a magnifying glass with a black handle and frame, positioned over a blue-tinted globe. The globe shows the continents of North and South America. In the background, there is a blurred document with Latin placeholder text. A semi-transparent white horizontal band is overlaid across the middle of the image, containing the main text.

**特徴のあるデータを見つけられるかどうか  
が成否を左右する**

# 特徴量の例

- 構造化データ(表形式で表せるもの)

例: 顧客データ

各項目が特徴量

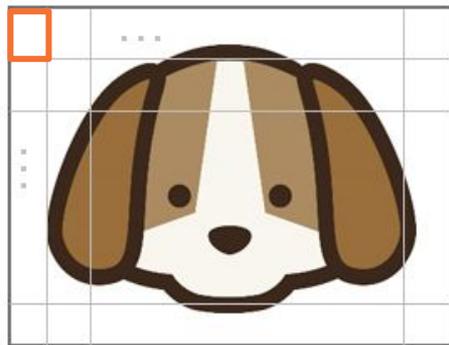
age	job	marital	education	default	balance	campaign	pdays	previous	outcome	y
30	unemployed	married	primary	no	1787	1	-1	0	unknown	no
33	services	married	secondary	no	4789	1	339	4	failure	no
35	management	single	tertiary	no	1350	1	330	1	failure	no
30	management	married	tertiary	no	1476	4	-1	0	unknown	no
59	blue-collar	married	secondary	no	0	1	-1	0	unknown	no
35	management	single	tertiary	no	747	2	176	3	failure	no
36	self-employed	married	tertiary	no	307	1	330	2	other	no
39	technician	married	secondary	no	147	2	-1	0	unknown	no

# 特徴量の例

- 非構造化データ(表形式で表せないもの)

例: 画像や文書、時系列ログなど

各ピクセルが特徴量



各単語が特徴量

親譲の無鉄砲で小供の時から損ばかりしている。小学校に居る時分学校の二階から飛び降りて一週間ほど腰こしを抜ぬかした・・・

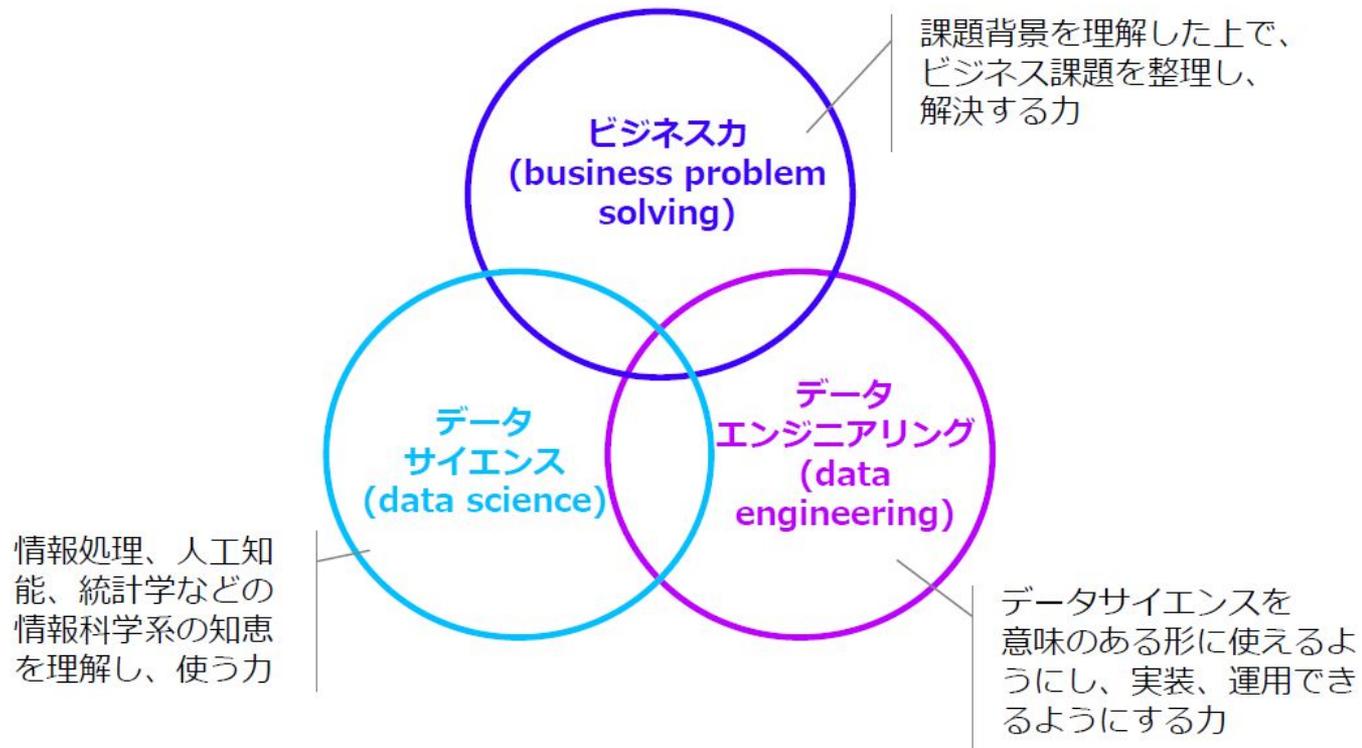
質の良い・特徴ある  
データを集める  
ために

- ★ 設備の故障予測・異常検知を例に
  - 熟練者は**何の挙動に着目**して、故障・異常と判断するのか
  - 予測・検知したい時間間隔に対し、**データの取得間隔**は適切ななのか
  - **めったに故障・異常が発生しない**理由は何か、そもそも自動化する意味があるのか

# 目次

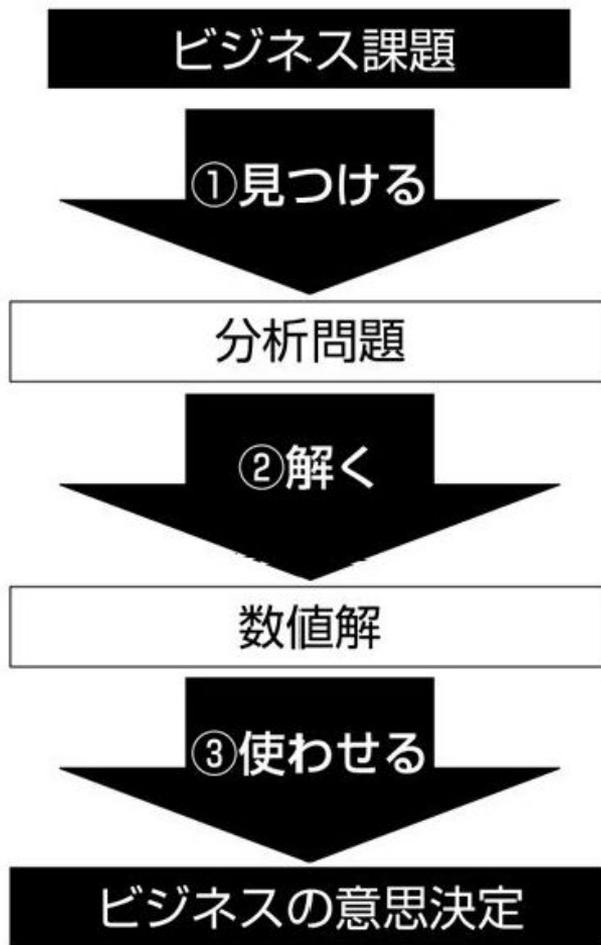
- データ分析あるある
- まずはPoCから始める
- 知っておきたいキーワード
- **分析力を手にするために**
- 参考情報

# AI人材に必要なスキル



各スキルに長けた人を集めて**チーム**を作る

# 外注の力を 借りる



外注化してもよいか？



# 参考書籍と サイト

- 会社を変える分析の力  
<http://bookclub.kodansha.co.jp/product?item=0000210704>
- データサイエンティスト養成読本 ビジネス活用編  
<https://gihyo.jp/book/2018/978-4-297-10108-4>
- マスクド・アナライズのAIベンチャー場外乱闘！  
<https://www.itmedia.co.jp/news/series/11565/>
- If文から機械学習への道  
<https://www.slideshare.net/nishio/if-80195170>
- Aidemy : 機械学習概論  
<https://aidemy.net/courses/2010>

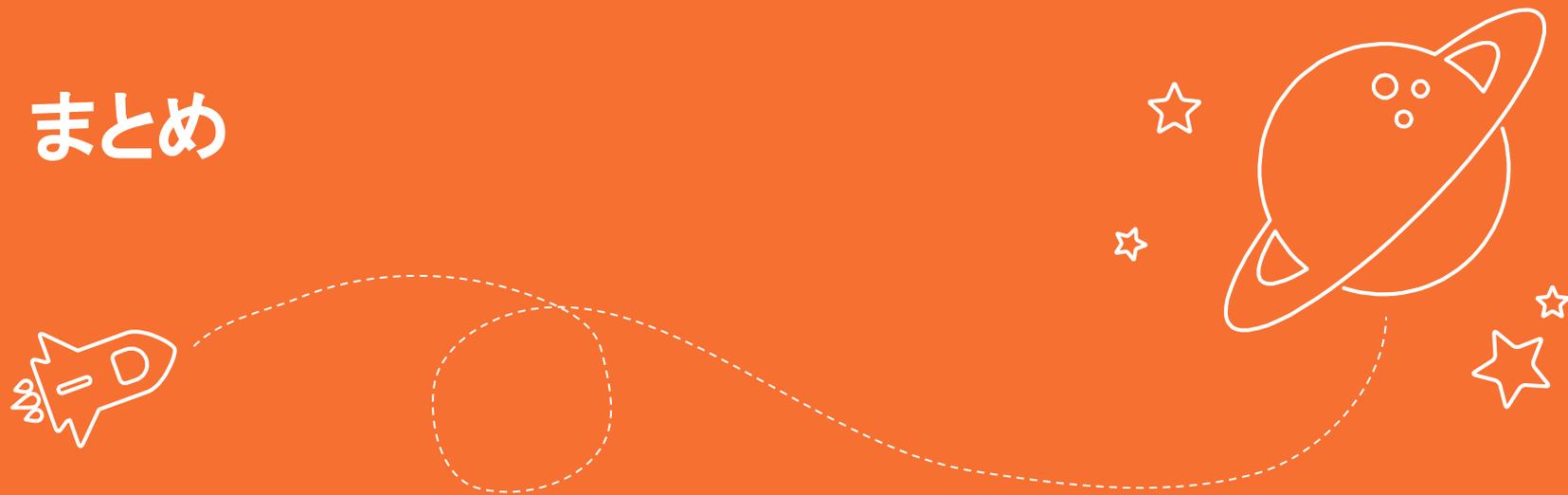
# 参考書籍と サイト

機械学習のための前処理」入門

[http://www.ric.co.jp/book/contents/book\\_1196.html](http://www.ric.co.jp/book/contents/book_1196.html)



まとめ



本日の内容をもとに、データ活用を小さく始めていきましょう！