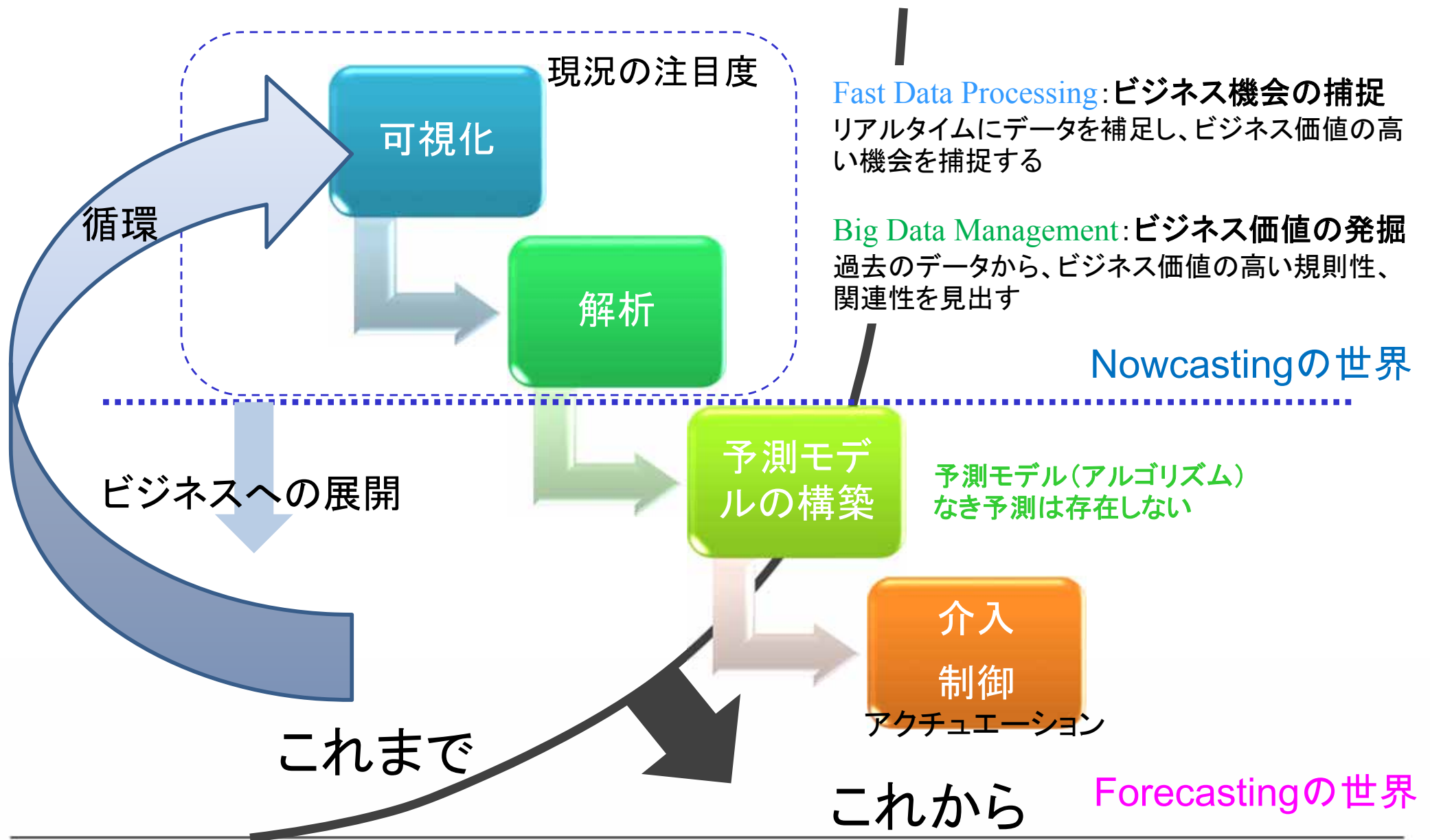


# ビッグデータ利活用の4つのステージ



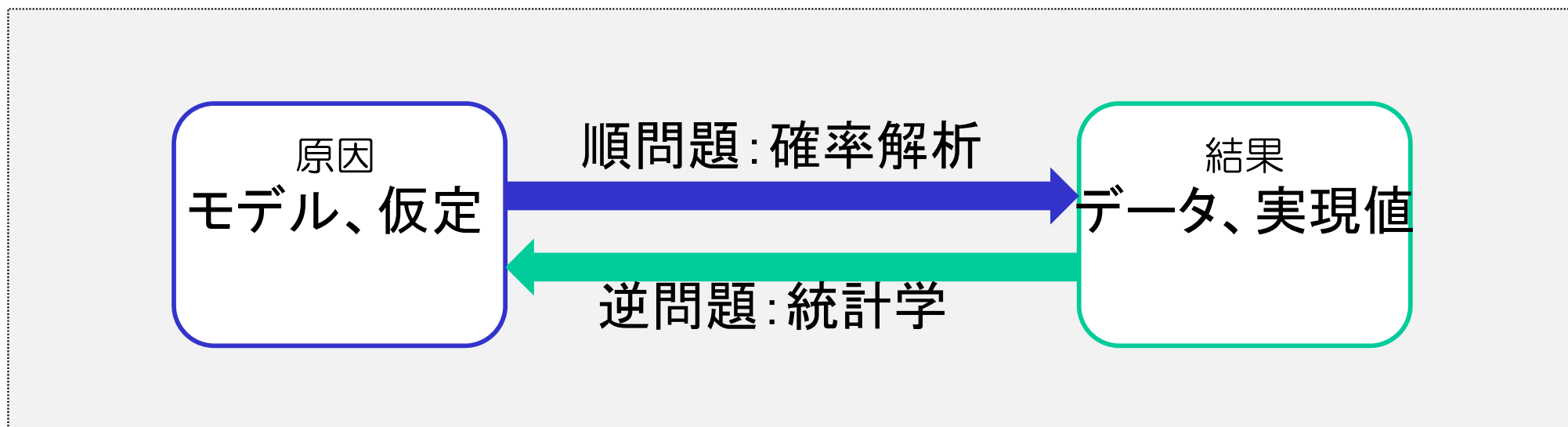
# 求められる人材

- 順解析と逆解析  
実世界の問題のほとんどが逆解析
- ニーズをくみとれるコミュニケーション力。技術シーズが出発点では期待感に答えにくい

# 確率論と統計学：さいころ

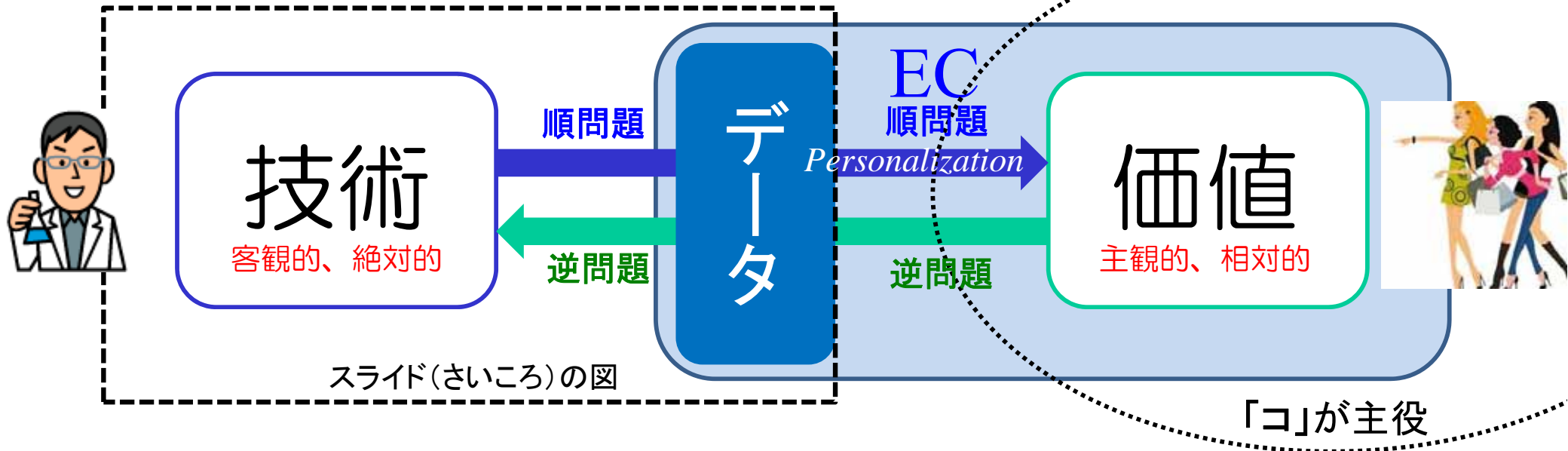
$$p(X = 1) = \frac{1}{6}$$


5, 4, 1, 3, 3, ...



# 帰納と演繹

## ■ 研究開発上、重要な視点とは？

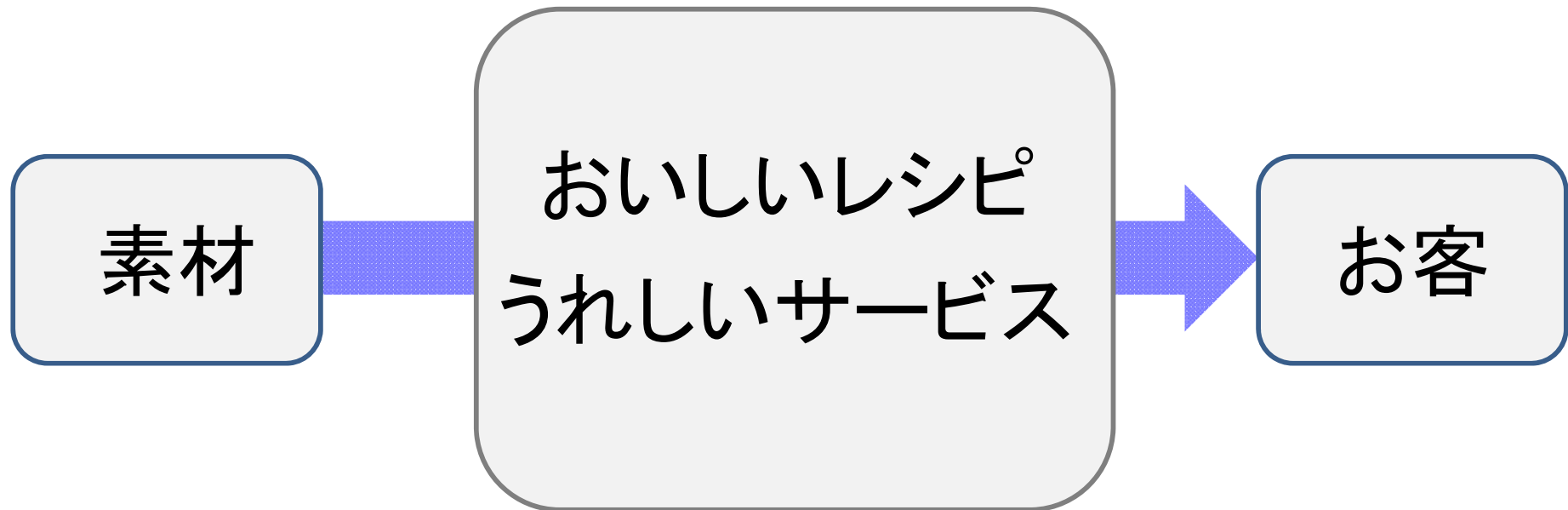


理論と仮定から結果を導く

VS.

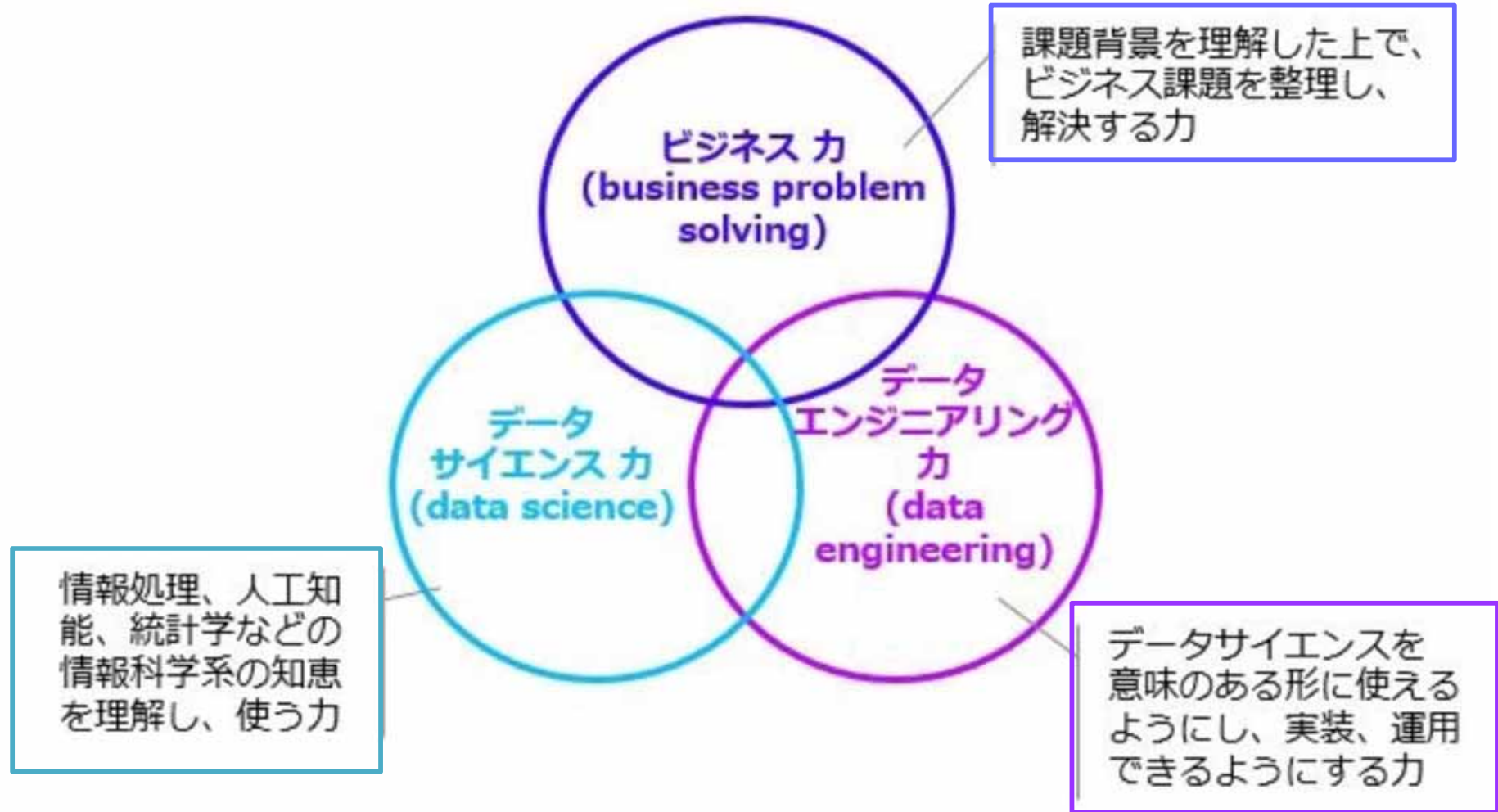
結果から原因を探る

# 不足している人材のタイプ



分析シナリオをみつける  
モデリング

# データサイエンティストに求められるスキルセット



# データサイエンス人材育成のあるべき姿と現実に向けた仮説

## 育成レベルと、毎年の育成目標人数

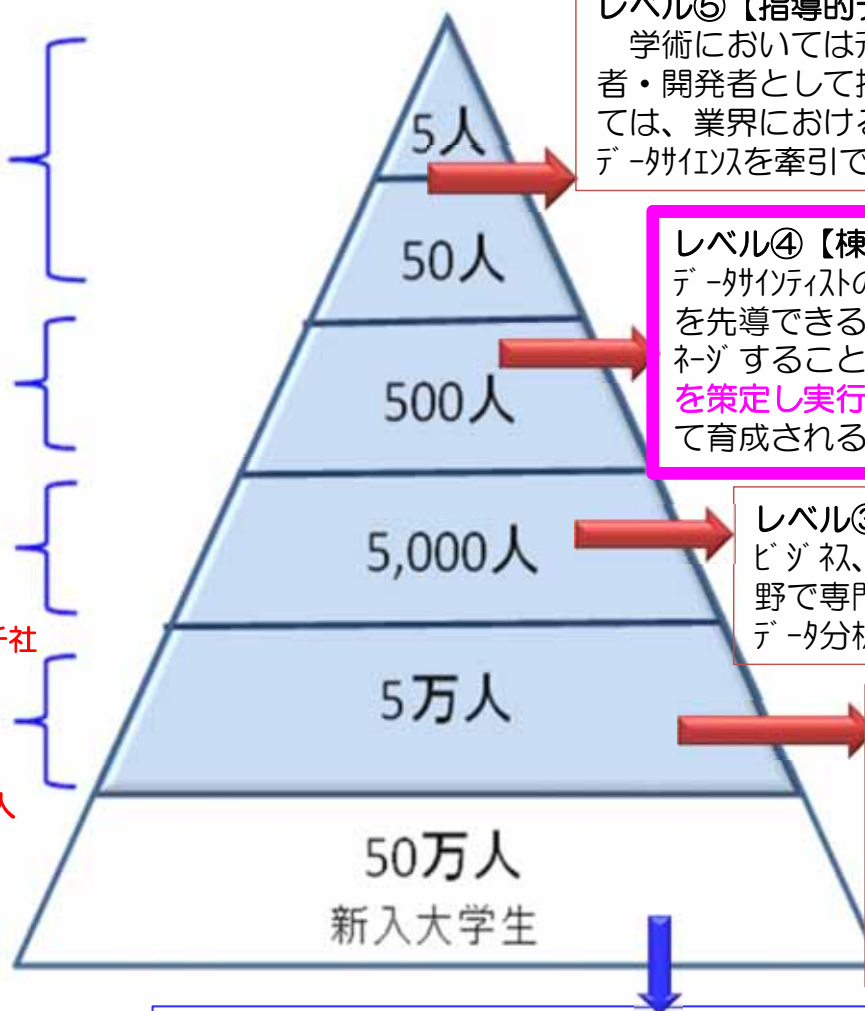
データサイエンティスト協会が定めたスキルレベル (2014年12月)

業界を代表するレベル  
Senior Data Scientist  
トップタレント  
年間数名～数十名

棟梁レベル  
(full) Data Scientist  
毎年5千人の「独り立ち」を指導統括  
「独り立ち」6～15人につき1人程度

独り立ちレベル  
Associate Data Scientist  
資本金10億以上の会社はおよそ6千社  
産業界向きだけでも5千人程度

見習いレベル  
Assistant Data Scientist  
理系修士入学者は年間およそ5万人



レベル⑤【指導的データサイエンティスト】  
学術においてはデータサイエンスの最先端を切り開くワールドクラスの研究者・開発者として指導的な能力を発揮する者、また産業界においては、業界におけるビッグデータ・データサイエンスに基づくビッグデータ・データサイエンスを牽引できるトップタレント (Googleのシリ・ペイ級の者)

レベル④【棟梁】  
データサイエンティストのチームを率いて、組織におけるビッグデータ利活用を先導できる能力をもった人。複数の応用分野を俯瞰的にマネージすることができ、データサイエンスの観点から全体最適の戦略を策定し実行するリーダシップが求められる。主に実務を通して育成される能力

レベル③【独り立ち】  
ビジネス、データサイエンス、データアナリティクス のいずれかの分野で専門的な能力を持ち、自らのノウハウで高度なデータ分析・問題解決能力を発揮する。実務経験が必須

レベル②【見習い (基礎能力)】  
全てのデータサイエンティスト (実務家・研究者問わず) が持つべき能力。ミドルクラスのマネージャにも必須。理系の修士は全て、文系でも社会科学系・言語学・心理学等の専攻で身に着けるべき。適切な指導の下、ビッグデータ利活用プロジェクトの一部を担当可

レベル①【データリテラシー】  
文系・理系を問わず全ての学生が持つべき高校から大学学部レベルの素養 (リテラシー) ベースになる統計的概念、データに基づく思考や問題解決の基礎理念、ITリテラシー他

平成27年7月30日  
大学共同利用機関法人 情報・システム研究機構  
ビッグデータの利活用に係る専門人材育成に向けた  
産学官懇談会



# まとめ

- インフラ維持・管理のためのビッグデータ活用には、信号処理、統計数理、機械学習の高度な分析能力をもつ人材の確保が必須
- 棟梁レベルのデータサイエンティストの育成が急務