データサイエンティストのための
スキルチェックリスト/タスクリスト概説

Skill Checklist & Task list Overview

一般社団法人
データサイエンティスト協会
スキル定義委員会

独立行政法人
情報処理推進機構
ITSS ＋（データサイエンス領域）

はじめに —— Preface

世の中は、産業革命以来の歴史的な局面を迎えようとしています。特に、OpenAI の ChatGPT のような大規模言語モデルが登場したことは、自然言語生成技術の新たな時代の幕開けを象徴しています。これらの生成 AI は、人間のように自然で流暢なテキストを生成する能力を持ち、特定の領域で人間の作業を補完、あるいはそれを超越する可能性を秘めており、AI（Artificial Intelligence：人工知能）の民主化時代が到来しつつあるともいえます。

昨今のインターネットやスマートフォンをはじめとするモバイルデバイスの急速な普及に伴い、データの爆発的な増加と情報処理技術の革新が起き、これまで不可能とされてきた規模のデータ活用が可能となりました。さらに 2020 年 4 月にサービスを開始した 5G 回線の登場によってモバイルデータ通信の高速・大容量化が加速し、収集できるデータも拡大しました。

技術進歩だけでなく、データ活用はすでに私たちの生活に身近な存在となっています。ICT 産業に限らず、自動車、健康医療、インフラ、エンターテイメント、農業……あらゆる産業において実用化が始まっています。こうした膨大なデータのリアルタイム処理も可能になることから、データ活用の重要性がより一層高まっていくことは明らかです。

私たちは今、「データを活用して新しい価値を生み出す」力が必要とされる世界に生きています。これまでは既存の市場やルールのなかで切磋琢磨し活動することがモノ・カネを生み出すことにつながってきましたが、これからは、枠組みを超えた新しい価値を創出・創造することが富へとつながる時代を迎えます。

その新しい価値を生む鍵となるのがデータや AI です。私たちはデータをどのように扱い、どのような価値を、どのような社会を生み出していくのかをデザインし、実現させるスキルを磨かなければなりません。

ひとりでも多くの人がデータを味方につけ、新たな社会を描いてほしい。データサイエンティストスキルチェックリスト／タスクリストはそんな思いを込めて、データ利活用のために必要とされるスキルを体系化しています。学生、社会人、データ活用の課題に直面している人、そして今は縁がないと思っている人にも、本書で取り上げているスキルが新時代を生き抜く一助になれば幸いです。

目次　CONTENTS

はじめに Preface …… 1

Part 01

WHY データサイエンス？ …… 5

社会背景 …… 6

身近にあるデータサイエンス …… 8

Part 02

データサイエンティストに必要なスキルセットとタスク …… 11

スキルチェックリストとは？ …… 12

モデルカリキュラムとは？ …… 18

スキルチェックリスト の変遷 …… 19

タスクリスト …… 23

Part 03

スキルチェックリストの読み方 …… 25

全体像を理解する …… 26

スキルチェックリストの見方 …… 32

Part 04

タスクリストの読み方 …… 35

全体の構造を理解する …… 36

タスクリストのカラム説明 …… 38

タスクリストにも流れがある …… 40

Part 05

AI 利活用スキル、AI 利活用タスクリストの読み方 …… 45
AI 利活用スキルとタスクリストの誕生の背景 …… 46
データサイエンティストの本質は変わらない …… 47
AI 利活用スキルは主に2つに類型化される …… 48
タスクリストは分析タスクの前後でのタスクが特徴的 …… 49

Part 06

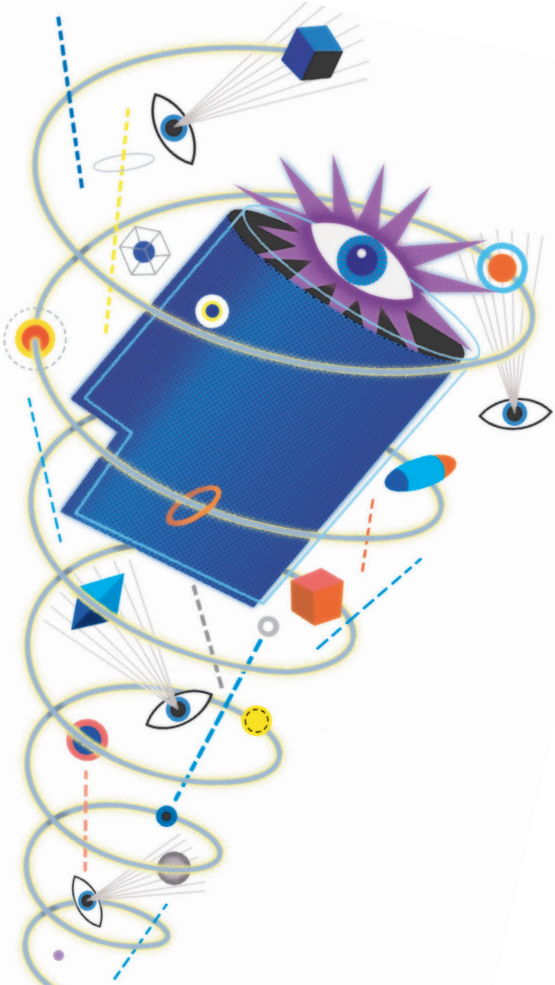
スキルチェックリスト・タスクリストの使い方 …… 51
データサイエンティスト初心者・目指している人 …… 52
データサイエンティスト実務家 …… 53
データサイエンティストチームリーダー　プロジェクトマネージャー …… 54
外注計画担当者 …… 55
人事採用担当者・エージェント …… 56
学校の先生 …… 57
学生（学ぶ側の人） …… 58
育成制度を作ろう …… 59

Part 07

リストにない定義や考え方 …… 63
ドメインスキル …… 64
文章に込められた意味 …… 65
進化の速い分野 …… 68
ビジネスや事業への関わり方 …… 72
お役立ち情報 Appendix …… 74

Part 01

WHY
データサイエンス？



●社会背景

社会的ニーズの高まりから生まれた

近年の技術革新によって、私たちが手にできるデータは爆発的に増大しました。そうしたデータを応用したいという世の中のニーズの高まりに呼応して活用できる人材が広く求められるようになり、登場したのが「データサイエンティスト」という職業です。**データサイエンスとは、アルゴリズムや統計などといった情報科学系の理論を活用してデータを分析し、有益な知見を見出すことを追究する新しいアプローチ**です。

この 10 年ほどの間にデータサイエンスによるデータ活用の手法や成果は飛躍的に発展し、多くのデータサイエンティストが活躍するようになりました。多くの企業や分野でデータサイエンティストが求められ、人材不足が懸念される状況にもなっています。その一方で、「データサイエンティストの仕事も将来的には AI に取って代わられるのでは……」という声も聞かれるようになりました。もちろん、人間よりも AI が計算・処理した方が速度も速く精度も高いでしょう。よって、識別や予測などの領域は、AI に代替されていくものでしょう。しかし、**どのようにデータを使って社会にどのような価値を生み出すかをデザインし実行していくことは、人間にしかできない仕事**です。

また、昨今の生成 AI の台頭により、新たなビジネス機会が生まれ、データサイエンティストの役割はますます重要となっています。データサイエンティストは、データを分析し、AI の潜在能力を引き出すことで、企業のデジタル変革を推進し、市場ニーズに応じたサービス開発を支援したりなど、事業改革のリーダーとして、企業の競争力向上と持続可能な成長を支える中核的な存在となっていくでしょう。

AI の技術・研究のために、世界各国から優秀な人材が研究機関や企業に集められ、莫大な投資が行われています。日本でも、2021 年度より「数理・データサイエンス・AI 教育プログラム認定制度」が始まり 2019 年 6 月に「AI 戦略 2019 ～人・産業・地域・政府全てに AI ～」が発表され、

一定要件を満たすデジタル社会の基礎知識（いわゆる「読み・書き・そろばん」的な素養）に関する大学や広告の教育プログラムに対して認定を行う事で、AI 教育への取り組みの後押しが始まっています。また、2024 年度には「AI 事業者ガイドライン」の第一版が公開され、国内の AI 関連事業のさらなる推進が期待されています。

データや AI を活用することで何ができるのか、どうすればより社会に役立てられるのかを考え、**新しいデジタル社会を実現していくためには、データサイエンスに関するスキルがとても重要**になります。同時に、スキルレベルを測る手段の重要性も徐々に増してきました。

データサイエンティスト協会では、そうした社会的ニーズに応えるために、「データサイエンティスト検定　リテラシーレベル」という検定試験を実施しています。

●身近にあるデータサイエンス

すべての人に関係するデータサイエンス

データサイエンスと聞くと、「難しい」「理系にしかわからないもの」と感じる人も多いかもしれません。しかし**データサイエンスは、すでに私たちの生活の中でとても身近な存在**となっています。

インターネット通販で商品を選んでいるとき、「あなたへのおすすめ商品」が表示されることがありませんか。これは、過去に同じ通販サイトに訪れた人たちの行動履歴や商品データから、あなたと似たパターンのショッピングをしている人や、似た商品を探し出すことで、あなたが気に入ってくれそうな商品を探して表示しています。

また、近年よく耳にする**自動運転も、データサイエンスの代表的な実用事例**の一つです。自動運転は、走行中の周辺の様子をカメラやセンサーで捉え、そのデータを分析することで周辺の危険を予測し、ハンドルやブレーキを操作してくれます。

最近ではスポーツの領域でも、データサイエンスの活用事例に触れることが多くなりました。メジャーリーグの中継では、打席に立つバッターが映し出されている場面で、ストライクゾーンが画面上に枠となって表示されています。これもリアルタイムで映像を解析することで、どこがストライクゾーンになるのかを判定しています。

また、コロナ禍の際は、日々の感染者数・死者数・重症者数が報道されていました。これらのデータの多くはオープンデータとして公開、利用されました。こうしたデータと GPS などのデータを組み合わせて行っていた外出自粛率の算出や、接触頻度の見える化なども、データサイエンスの活用事例の一つです。

生成 AI の登場により、AI はより身近な存在となりました。これらの技術により、自然言語を用いて AI を活用することが可能になり、情報の取得方法に革新をもたらしました。従来は Web ページから情報を「探す」

ことが一般的でしたが、生成 AI によって情報を「生成する」という新たなアプローチが選択肢に加わりました。この進化は、私たちの情報アクセス方法を根本から変える可能性を秘めています。

一方で「フェイクニュース」などの言葉を最近では聞かれたことがあると思います。機械学習の技術を利用し、SNS 上で実在しない人物や偽の記事を氾濫させ、社会的な混乱やプライバシーの侵害などの問題を招いています。データサイエンティストとしてデータや技術をどう扱うのか？どのように社会に還元するのか？が真摯に問われている一面といえます。

このように、**データサイエンスのスキルが活用される業界やシーンは非常に幅広く、その裏では多くのデータサイエンティストが活躍**しています。また活用シーンが拡大することで、今は縁のない業界や職種の方でも、当たり前のようにデータを活用する世界がきています。

データサイエンスのスキルは、あらゆる業種や職種の方にとっても重要な基礎素養となっていくことでしょう。

データサイエンティストとは？

データサイエンティストとは、「データサイエンス力、データエンジニアリング力をベースにデータから価値を創出し、ビジネス課題に答えを出すプロフェッショナル」です。

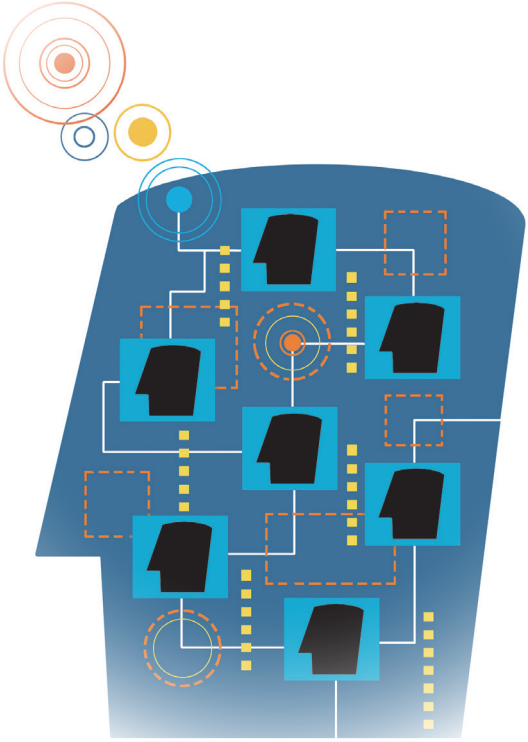
ここでの「ビジネス」とは企業の営利活動だけではなく、社会の役に立つ意味のある活動全般を指します。「プロフェッショナル」とは、体系的にトレーニングされた専門的なスキルをベースに顧客（お客様、クライアント）に価値を提供し、その対価として報酬を得る人です。

図1　データサイエンティストとは、裏打ちされたスキルを駆使してデータを分析し、ビジネス課題の回答を導き出すプロフェッショナルを指します

スペシャリスト	• 何らかの分野に特化した人
エキスパート	• 何らかの分野について、体系的で秀でた知識とスキルを持っている人
プロフェッショナル	• 体系的にトレーニングされた専門性の高いスキルを持ち、 • それをベースに顧客（お客様、クライアント）にコミットした価値を提供し、 • 結果に対して認識された価値の対価として報酬を得る人

Part 02

データサイエンティストに必要なスキルセットとタスク



●スキルチェックリストとは？

必要なスキルを分野別、体系的に一望・俯瞰

Part01 で紹介したとおり、データサイエンスを駆使するデータサイエンティストは、これからの時代にとって欠かせない存在となり、データサイエンティストを志す人やプロフェッショナルとしてのデータサイエンティストを育成する企業が増えています。

しかし、データサイエンティストになりたい！ 当社でもデータサイエンティストを育成したい！ と考えたとき、まず何から始めればよいのでしょう。

データサイエンティストのスキル（＝できること）を一望、俯瞰できれば、データサイエンティストを目指す人の一助になるに違いないという意識のもとで作成したのが「スキルチェックリスト」です。

スキルチェックリストには、データサイエンティストとして活躍するための必要なスキル（スキル項目）を、分野ごとに体系的にまとめています。スキルチェックリストを見れば、データサイエンティストを目指す人やデータサイエンティストを育成したい企業にとって、勉強や教育の目標とすべきスキルが一目でわかります。

それぞれのスキル項目には、「スキルレベル」を定義されていますので、基礎から高難易度まで、必要とされるスキルレベルが確認できます。

スキルチェックリストのねらい

「データサイエンティストにできること（＝スキル）」をスキルチェックリストとしてまとめることに至った背景には、一般社団法人データサイエンティスト協会・スキル定義委員会が共有してきた以下のような問題意識がありました。

・「データサイエンティスト」という言葉だけがひとり歩きし、本質が十分に理解されていない（または誤解されている）。

・「データサイエンティスト」に求められるスキルの幅の広さやレベルの高さが十分に理解されていない。

・「データサイエンティスト」を育成するための体系的な指標がない。

これらを踏まえて、「データサイエンティスト」の仕事やスキルについてもっとくわしく、より多くの方々に知ってもらうためにまとめたのがスキルチェックリストです。

データサイエンティストのできることを単に紹介するのではなく、企業や組織のなかで活躍し、これからの社会、これからの日本を変えていけるような真のデータサイエンティストを数多く生み出したいという願いも込められています。

3つのスキルセット

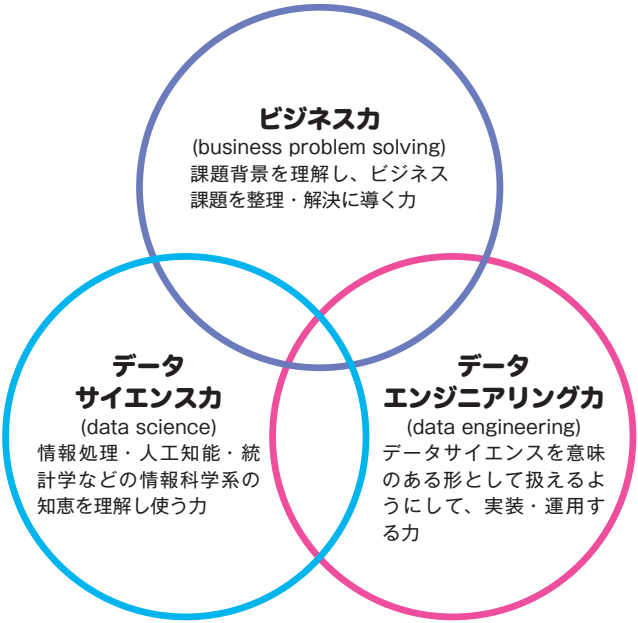
データサイエンティストに求められているスキルセットは、以下の3つです。

・ビジネス力（business problem solving）：
課題背景を理解し、ビジネス課題を整理・解決に導く力

・データサイエンス力（data science）：
情報処理・人工知能・統計学などの情報科学系の知恵を理解し使う力

・データエンジニアリング力（data engineering）：
データサイエンスを意味のある形として扱えるようにして、実装・運用する力

図2　データサイエンティストに求められるスキルセット



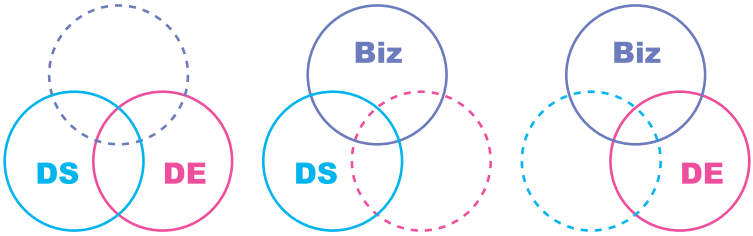
3つのスキルセットはいずれも欠けてはいけない

3つのスキルセット（ビジネス力、データサイエンス力、データエンジニアリング力）はとても重要なファクターで、どれか一つが欠けてもデータサイエンティストとして十分な力を発揮できません。

ビジネス力に欠ければ、そもそも解決すべき問題が定義、整理することができません。データエンジニアリング力が乏しければ、ビジネス課題についてのサイエンスの活用はわかっても、肝心の実装（実際に利用できるシステムを仕立て上げること）ができません。データサイエンス力がなければ、ビジネス課題の上で実装を用意できても、要となるサイエンスの知恵が足りないため、データの力を活かすことができません。

スキルセットは、いわば三位一体なのです（**図3**）。

図3　いずれのスキルセットが欠けてもデータスペシャリストとしての力は発揮できない



データスペシャリストではあるが、そもそも解決すべき問題が定義・整理できない

＞プロフェッショナルではない

ビジネス課題もサイエンスの活用もわかるが、実装できない

＞必要な変化を起こせない

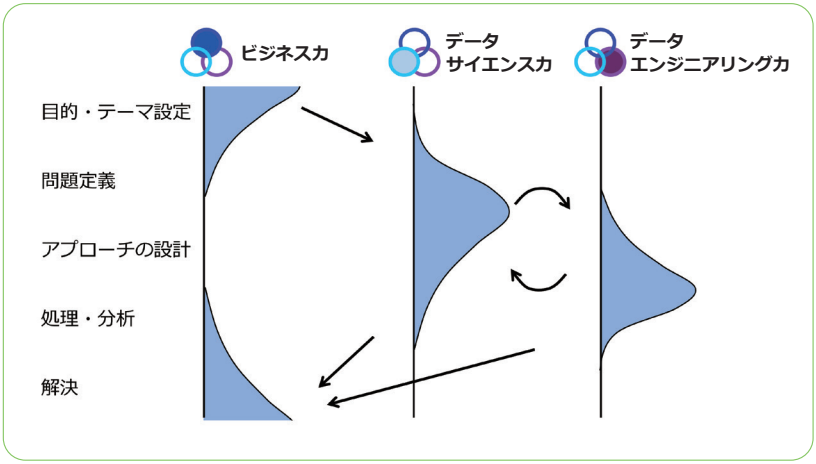
ビジネス課題のうえで実装を用意できるが、要となるサイエンスの知恵が足りない

＞賢いやり方を提供できない

3つのスキルセットはデータサイエンティストにとって欠かせないファクターですが、それぞれの重要性は課題解決のフェーズによって変わってきます（**図4**）。

目的・テーマ設定や最後の解決などのフェーズでは、関係者を説得・調整したり、説明したりするようなビジネス力が強く求められます。定義された課題に対してどのようなアプローチで立ち向かうのかを計画するには確かなデータサイエンス力が必要です。そして、実際の分析を進めるためには、データエンジニアリング力が要求されます。

図4　問題解決のフェーズによって要求されるスキルセットは変わる



よくある誤解　一人でできたらスーパーマン !?

データサイエンティストに要求されるスキル定義は、必ずしも一人ですべてのスキルをこなすことを想定したものではありません。スキルセットごとに得手不得手もあるでしょう。最初からすべてを淀みなくこなせたらそれこそスーパーマン !? かもしれませんね。

得意なスキルを伸ばすことで問題解決に貢献できます。全体を俯瞰してコーディネートする人材は必要ですが、個々のスキルに秀でたメンバーでチームを組み、実務にあたることも現実には多くなるでしょう。こうして経験を積むことは、一人前のデータサイエンティストになるための糧になるのです。

スキルレベルの４段階

データサイエンティストにとって必須であるスキルセットに対して、それぞれに４段階のスキルレベルがあります。

１．業界を代表するレベル ：Senior Data Scientist
２．棟梁レベル ：Full Data Scientist
３．独り立ちレベル ：Associate Data Scientist
４．見習いレベル ：Assistant Data Scientist

スキルチェックリストは、見習いレベル (Assistant Data Scientist) ～棟梁レベル (Full Data Scientist) までの判定に用います。

また、近年、高校や大学でデータサイエンス授業の必修化の動きがありますが、数理・データサイエンス教育強化拠点コンソーシアムが発表した「数理・データサイエンス・AI　モデルカリキュラム」（以下、モデルカリキュラム）で示されたリテラシーレベル、応用基礎レベルとスキルレベルの対比については、**図５**の右に示すように、モデルカリキュラムのリテラシーレベル、応用基礎レベルはそれぞれ見習いレベルとその下のレベルに相当するスキルであると位置づけています。

図５　データサイエンティストのスキルレベル

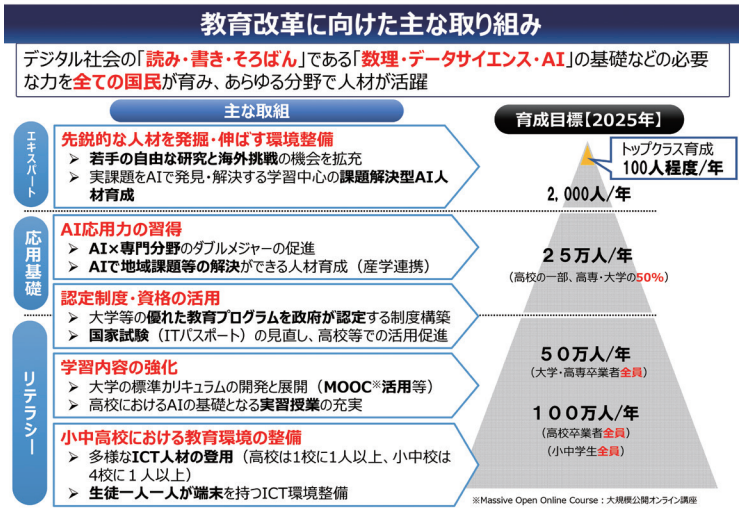
モデルカリキュラムとは？

AI 戦略 2019 に基づき、数理・データサイエンス教育強化拠点コンソーシアムでは、数理・データサイエンス・AI 教育の全国展開を目指し、リテラシーレベル（50 万人 / 年目標）および応用基礎レベル（25 万人 / 年目標）のモデルカリキュラムを公開しています。

数理・データサイエンス・AI（リテラシーレベル）モデルカリキュラム ～データ思考の涵養～（2024 年 2 月 22 日改訂）
http://www.mi.u-tokyo.ac.jp/consortium/pdf/model_literacy_20240222.pdf

数理・データサイエンス・AI（応用基礎レベル）モデルカリキュラム ～ AI× データ活用の実践～（2024 年 2 月 22 日改訂）
http://www.mi.u-tokyo.ac.jp/consortium/pdf/model_ouyoukiso_20240222.pdf

図 6　AI 戦略 2019（抜粋）



出典：統合イノベーション戦略推進会議決定 AI 戦略 2019【概要】
https://www.maff.go.jp/j/kanbo/tizai/brand/attach/pdf/ai-15.pdf

スキルチェックリストの変遷（Ver.1 ～ Ver.3）

データサイエンティストに求められるスキルは、技術や産業の進化に伴い変化していきます。新たな技術が生まれたり、重要な技術が新技術に置き換えられて必要性がなくなったり、かつては高難度だったものが低難度に変化したりすることもあります。

特にデータサイエンス領域は進化・深化が早いため、これまでは全スキルチェック項目の更新を2年に1回実施して最新のスキルを反映してきました。

2017年（Ver.2）に見直したおもな内容は以下の通りです。

- 機械学習、深層学習時代に対応したスキル項目の見直し・追加
- 「画像・動画処理」「音声 / 音楽処理」のカテゴリ・スキル項目の見直し
- 「知財」スキルカテゴリの追加

2019年（Ver.3）に見直したおもな内容は以下の通りです。

【データサイエンス力】

- 「統計数理基礎」を「基礎数学」とスキルカテゴリ名変更し、線形代数や微積分に関するスキル項目を追加
- 「予測」関連のスキルカテゴリを「回帰 / 分類」「前処理」「モデル設計」「評価」へと再分類し、精緻化

【データエンジニアリング力】

- クラウド環境を用いた分析構築スキル項目の追加
- セキュリティ関連スキル項目の強化
  （例：個人情報や GDPR 等を考慮したデータ管理スキル）
- 近年の技術動向反映
  （例：p 値の限界、特徴量エンジニアリング DL ライブラリ活用）

【ビジネス力】

- 分析プロジェクトの初期段階で実務に関するスキル項目を強化
  （例：「着想・デザイン」「契約権利保護」等）
- 事業実装に関連するスキルカテゴリ「事業への実装」「分析評価」の見直し・追加

スキルチェックリストの変遷（Ver.4）

2021 年（Ver.4）に見直したおもな内容は以下の通りです。

【全体】

- モデルカリキュラムとの整合性の調整
- AI 活用や機械学習のシステム運用、アルゴリズム、それに伴う倫理課題に関するスキル追加
- ライブラリ活用やクラウドサービスの一般化に伴うスキルの更新

【データサイエンス力】

- 基礎数学の重要性見直し
- 学習と予測カテゴリの明確化、潮流への対応
- 近年、重要度の増したスキル項目追加

【データエンジニアリング力】

- 新規カテゴリ「AI システム運用」を追加
- 環境構築技術等、技術トレンドへの対応
- リテラシーレベルのプログラミングスキル追加
- Hadoop や Scala、GPU など一般化技術の一部削除

【ビジネス力】

- AI 活用時のビジネス視点を強化
- プロジェクトの多様なフェーズに対応するスキル
- DS から「分析アプローチ」スキル移設
- マネジメントスキルをプロジェクトと組織に 2 分割

スキルチェックリストの変遷（Ver.5）

2023 年（Ver.5）に見直したおもな内容は以下の通りです。

【全体】

●生成 AI 活用のスキルをそれぞれに追加

●カバー範囲の広いスキルが多く「またがり」スキル増加

●モデル性能を向上させるチューニング・リソース管理のスキル等追加

【データサイエンス力】

●カテゴリを目的軸、サブカテゴリを技術軸として再編成

- 生成 AI に関するスキル群を追加し束ねる「生成」カテゴリを追加
- 解析の科学的実践に必要な基礎スキルを「科学的解析の基礎」に統合
- モデリング関連スキルを「モデル化」「モデル利活用」に統合

【データエンジニアリング力】

●クラウド環境で提供される新機能の利活用スキルの追加

●生成 AI 環境の利用や構築に必要なスキルの追加

●学習済モデル / サービス利用環境整備スキルを DS スキルから移動

【ビジネス力】

●生成 AI 活用スキル多数の新規追加

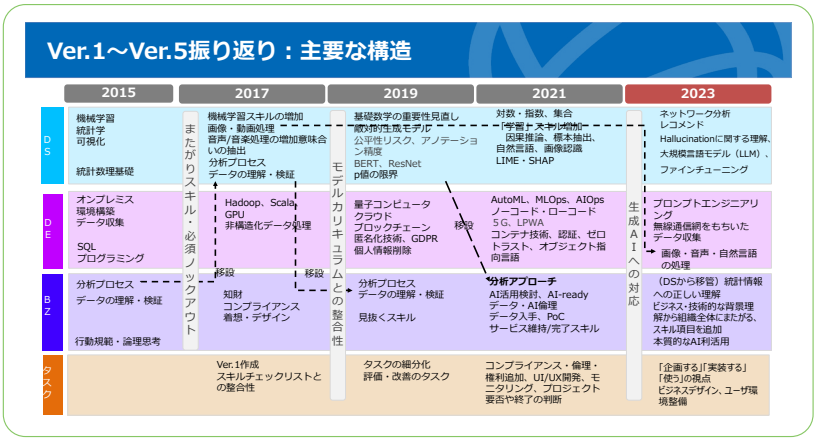
- 行動規範：革新的技術への“わくわく”マインドセット
- 着想・デザイン：生成 AI を用いた新規ビジネス着想やビジネス変革
- アプローチ設計：生成 AI を活用した課題解決の設計
- 分析評価：生成データの評価、基準の作成
- 契約権利保護：権利保護のガイドライン
- PJ マネジメント、組織マネジメント：環境負荷や推進課題への対応

● DS スキルから「データ理解」の移管

スキルチェックリストの主要な変遷（Ver.1 ～ 5）

2017 年（Ver.1）から 2023 年（Ver.5）の各領域の主要な変遷を**図 7** に示します。

図 7　データサイエンティスト協会 10th シンポジウム スキル定義委員会発表資料（抜粋）



●タスクリスト

「データサイエンティストにできること＝スキル」をまとめたスキルチェックリストに対して、「データサイエンティストが行うこと＝タスク」をまとめたものが「タスクリスト」です。

タスクリストは、スキルチェックリストを補完するもので、実際にどのようなプロセス（業務）においてスキルチェックリストに記載されているスキルが必要とされるかを整理したものです。

データサイエンスを駆使して価値を生み出す活動の流れそのものがよくわかっていない、自分はデータサイエンスのプロジェクトのなかでどのような役割を果たしたいのかがはっきりしない……といったときに役立つのがタスクリストです。そのような状況になったら、まずタスクリストを見て業務のなかで自身が担当するタスクは何かを確認します。

その後でスキルチェックリストを参照して、担当するタスクを行うには、どのスキル項目を身に付ける必要があるのかを確認するといった使い方を想定しています。

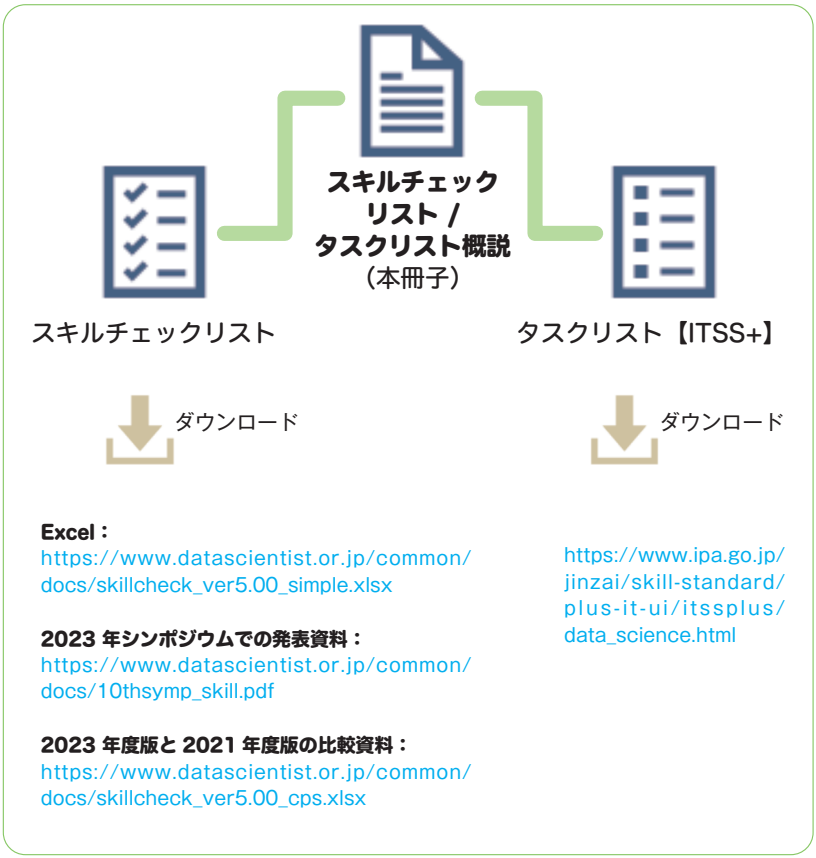
また、生成 AI の登場によってデータサイエンティストの業務は大きく変化します。この変化に対応し、生成 AI をビジネスや実務に活用するためにデータサイエンティストが発揮できるスキルを把握できるよう 2023 年のタスクリスト改訂で「AI 利活用タスクリスト」を追加しています。詳細は「Part4. タスクリストの読み方」で解説します。

●スキルチェックリスト / タスクリストの構造

本書も含めたドキュメントの構造を下図に示しました（**図8**）。

独立行政法人情報処理推進機構（IPA）は、スキル標準の新領域「ITSS+（データサイエンス領域）」整備を一般社団法人データサイエンティスト協会スキル定義委員会とともに取り組んでいます。スキルチェックリスト、タスクリストは、相互のウェブサイトで公開しています。

図8　関連するドキュメントの構造



Part 03

スキルチェックリストの読み方

●全体像を理解する

スキルチェックリストの構成

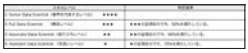
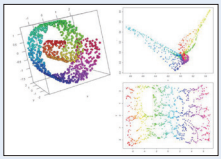
スキルチェックリストは、以下のパートで構成されています（**表 1**）。

- **チェックリストの説明**
- **スキルレベル定義**
- **スキル一覧（各スキルセット）**
- **参考資料**

まず「チェックリストの説明」シートと「スキルレベル定義」シートでスキルレベルの判定基準やスキルカテゴリの一覧、スキルレベルの人物像などを確認します。続いて、各分野のスキル一覧を眺めることで、全体感が把握しやすくなります。

なお、スキルチェックリストは適宜（今までは 2 年ごと）更新されるため、手元にあるものが最新版かどうか必ず確認しましょう。

表 1　スキルチェックリストの全体構成

シート名	記載内容
チェックリストの説明	・スキルレベルの**判定基準** ・**スキルカテゴリ**の一覧 
スキルレベル定義	・分野ごとのスキルレベルの**人物像**
ビジネス力 データサイエンス力 データエンジニアリング力	・スキルセットごとのスキル一覧
参考資料 （データ可視化）	データサイエンス力の「データ可視化」に関する参考資料 （2015 年 11 月 20 日時点） 右図は株式会社 ALBERT WEB サイト https://blog.albert2005.co.jp/2014/12/11/ 

スキルレベルの判定基準

「チェックリストの説明」シートには、スキルレベルの判定基準が記載されています。各スキルレベル達成のためには、必須スキルをすべて満たしたうえで判定基準を達成する必要があります（**図9**）。

図9　スキルレベルの判定基準

「データサイエンティスト スキルチェックリスト」は、データサイエンティストに必要とされるスキルをチェックリスト化したものです。チェックリスト内の各項目のスキルレベルは★の数によって示し、見習いレベル（Assistant Data Scientist）～ 棟梁レベル（Full Data Scientist）までを判定します。

スキルレベル		判定基準
① Senior Data Scientist（業界を代表するレベル）	★★★★	–
② Full Data Scientist（棟梁レベル）	★★★	★★★の全項目のうち、50% を満たしている。
③ Associate Data Scientist（独り立ちレベル）	★★	★★の全項目のうち、60% を満たしている。
④ Assistant Data Scientist（見習いレベル）	★	★の全項目のうち、70% を満たしている。

※「必須スキル」に◯がついている項目は、判定基準を満たしていても、この項目が達成されていないとそのレベルとは認められない項目として設定しています。
※独り立ちレベル以上のレベルは、下位のレベルを満たしていることが前提となります。

図9の「独り立ちレベル」になるための条件を例にあげると、**表2**で示した条件をすべて満たす必要があります。判定によっては、データサイエンス力は「独り立ちレベル」で、ビジネス力は「見習いレベル」ということもあります。

表2　独り立ちレベルに達するための条件（データサイエンス力の例）

条件1	データサイエンス力の★の全項目のうち、70% を満たしている
条件2	データサイエンス力の★★の全項目のうち、60% を満たしている
条件3	データサイエンス力の★の「必須スキル」をすべて満たしている
条件4	データサイエンス力の★★の「必須スキル」をすべて満たしている

スキルレベルの人物像

「スキルレベル定義」シートには、各スキルレベルで想定される人物像が記載されています（**図 10**）。各スキルレベルで想定される人物像が、ビジネス力、サイエンス力、データエンジニアリング力において、それぞれどのようなことができれば望ましいかを記載しています。「スキルチェックリスト」のバージョンアップごとに、世の中で求められるデータサイエンティストのスキルセットも変化してきているため、何が変わったのかを念頭に置きながら確認すると良いでしょう。

図 10　各スキルレベルの解説

	ビジネス（business problem solving）力	データサイエンス（data science）力	データエンジニアリング（data engineering）力
①Senior Data Scientist **業界を代表するレベル**	・業界を代表するデータプロフェッショナルとして、組織全体や市場全体レベルでのインパクトを生み出すことができる -対象とする事業全体、産業領域における課題の切り分け、論点の明確化・構造化 -新たなデータ分析、解析、利活用領域の開拓 -事業や産業全体に対するデータ分析・AI技術を核としたバリューチェーン創出 -技術や環境の変化に応じた業界・社会変革の構想・提言・働きかけ	・業界を代表するデータプロフェッショナルとして、データサイエンスにおける既存手法の限界を打ち破り、新たに課題解決可能な領域を 切り拓いている -既存手法では対応困難な課題に対する新規の分析アプローチの開発・実践・横展開 -高難度の分析プロジェクトのアプローチ設計、推進、完遂能力 -時代を牽引する新たなモデル（独自の生成モデルや基盤モデル、ロボティクス等他分野と融合したモデルなど）の企画・開発推進 など	・業界を代表するアーキテクトとして、データサイエンス領域で取り組もうとしている分析アプローチやAI活用戦略を、挑戦的な課題であっても安定的に実現できる -複数のデータソースを統合した例外的規模のシステム、もしくはプロダクトの構築、全体最適化 -技術的限界を熟知し、これまでにない代案の提示・実行 - 生成モデルを進化させるために必要となるシステムインフラ課題やデータ処理課題の技術的解決（独自モデルの開発に伴うエンジニアリング課題解決や、処理性能やコスト・環境負荷の劇的改善など） など
②Full Data Scientist **棟梁レベル**	・生み出す価値にコミットするプロフェッショナルとして、データサイエンティストとは何かを体現したビジネス判断、課題解決ができる ・初見の事業領域に向かい合う場合や、スコープが複数の事業にまたがる場合であっても本質的な課題を見出し、構造化・深掘りができる ・入り組んだステークホルダー構造の中で、Win-Winの形で価値を設計・創造し、そこからの発展を見据えた仕込みと推進するハブとしての役割を担うことができる ・プロフェッショナルからなる複数のチームによるプロジェクトの役割、目標を定義、リスクをマネージしつつ推進し、全体としてのアウトプットにコミットできると共に、メンバーを育成、さらには持続的な育成システムを作り出すことができる ・自身が関わる事業や組織において、新たに登場したAI技術を速やかに導入できる - 技術的ブレイクスルーを利用したビジネス影響の予測と革新的アイデアの生成 - 想定される事業価値に対する不確実性の見極めと環境配慮 - ビジネス戦略とテクノロジーを高度に結合したデータ戦略・企画・チーム編成の推進 - 組織・サービス・インフラ環境をふまえたAI活用事業・サービス企画及び推進 - AIの倫理的、法的課題への予見と対策とガイドラインの策定・推進 -プロジェクトやチームを超えて、データサイエンス・AIに関する正しい理解、知識・スキル向上のための取り組みを企画・推進できる	・予測、グルーピング、機械学習、深層学習、大量データの可視化、自然言語処理、画像・映像認識、音声認識、最適化問題などの応用的なデータサイエンス関連のスキルを活かし、データ分析プロジェクトの技能的主軸を担うことができる ・初見の事業領域に向かい合う場合や、スコープが複数の事業にまたがる場合であっても、適切な分析・解析アプローチの設計、実行、深掘りができる ・複数もしくは高度な分析プロジェクトを持つチームにおいて、Associate Data Scientist（独り立ちレベル）以下のメンバーの技能を育成することができる ・データ品質を高めるための取り組み方針を策定し推進できる ・各種生成モデルの特性や学習方法、計算コスト等の技術的理解にもとづき、主要ステークホルダーに適切に説明できるとともに、目的に合わせた適切な学習・活用を推進できる	・数十億レコード程度の分析環境の要件定義・設計、データ収集/蓄積/加工/共有プロセスやITセキュリティに関するデータエンジニアリング関連のスキルを活かし、データ分析プロジェクトを中核的に推進することができる ・全体を統括するアーキテクトとして、サービス上のそれぞれの機能がどのデータに関連があるか総合的に把握し、新たな技術を理解しつつ設計や開発に活かすことができる ・複数もしくは高度な分析プロジェクトを持つチームにおいて、Associate Data Scientist（独り立ちレベル）以下のメンバーの技能を育成することができる ・新たなサービスやプロダクトを実現するための、生成AIを活用したシステムを設計・開発・実装できる - 生成AIの性能や実現要件の検討 - 非機能要件（処理パフォーマンスやコストなど）に基づくモデルの選択 - ファインチューニングの必要性や方向性の提示 - データ収集方針やデータベースの設計 - 評価指標の検討
③Associate Data Scientist **独り立ちレベル**	・大半のケースで自立したプロフェッショナルとして、ビジネス判断、課題解決ができる -ビジネス要件の整理、プロジェクトの企画・提案・遂行 -分析のアプローチ設計および結果への適切な対応と評価 -AI・機械学習がもたらす倫理課題への対応（バイアスによる人種差別や知的財産権の侵害など） -データや分析結果の開示範囲、知財リスクの確認などの適切な対応 ・既知の領域、テーマであれば、新規課題であっても解くべき問題の見極めや構造化、深掘りができる ・データ、分析結果に対する表面的な意味合いを超えた洞察力を持ち、担当プロジェクトの検討結果を取りまとめ、現場への説明、実装を自律的かつ論理的に行うことができる ・５名前後のプロフェッショナルによるチームでのプロジェクトを推進しアウトプットにコミットできる -タスクの粘り強い完遂 -イシュードリブンでスピード感のある判断 -プロジェクトマネジメントと個別メンバーの育成 -異なるスキル分野の専門家、事業者との協働 など ・新しいデータ・AI利活用時代に即した基礎的な理解と価値創造の基礎的な実践 - AIや生成モデル技術の最新動向理解と活用検討 - AIと機械学習がもたらす倫理的課題についての理解と説明	・単一プロジェクトにおけるデータ分析をFull Data Scientist（棟梁レベル）に相談しつつ推進できる ・Assistant Data Scientist（見習いレベル）の日々の活動に適切な指示ができる ・既知の領域、テーマであれば、新規課題であっても適切な分析・解析アプローチの設計、実行、深掘りができる ・データの取得段階において、品質をできるだけ高めるための取り組みができる ・基礎的なデータ加工については、自律的に実施できる -外れ値・異常値・欠損値の対応や学習データの作成 -特徴量エンジニアリングによる効果的なデータの作成 など ・基礎的な分析活動については、自律的に実施できる -p値の限界の理解とサンプルサイズの判断 -多重共線性を考慮した重回帰分析 -傾向スコアなどによる因果効果の推定 -系列データの特性を理解した時系列分析 -適切なクラスター数による非階層クラスター分析 -ライブラリなどを活用した機械学習や深層学習、および評価 など ・非構造化データに対する基礎的な分析を自律的に実施できる -画像のパターン抽出や音声のノイズ除去	・単一プロジェクトにおけるデータ処理・環境構築をFull Data Scientist（棟梁レベル）に相談しつつ推進できる ・Assistant Data Scientist（見習いレベル）の日々の活動に適切な指示ができる ・数千万レコード程度のデータ処理・環境構築については自律的に実施できる -データの重要性や分析要件に則したシステム要件定義 -適切なデータフロー図、論理データモデル、ER図の作成 -HadoopやSparkでの管理対象データ選定 -SDKやAPI、ライブラリ、コンテナ技術などの適切な活用 -SQLの構文理解と実行 -分析プログラムのロジック理解と分析結果検証 -データ匿名化方法の理解と加工処理の設計 など ・分析要件に合わせたインフラ環境（GPU/CPU、クラウド/オンプレミスなど）を設計・実装できる ・AIシステム運用を推進できる -AIモデルの作成、デプロイ、学習データの更新、モデル再学習といったライフサイクル管理の実践 ・生成モデルを高度に利用するための技術理解と実践力を有している - データベースやファイルを組み合わせたプロンプトエンジニアリングの高度化 - ファインチューニングの実施・評価
④Assistant Data Scientist **見習いレベル**	・ビジネスにおける論理とデータの重要性を理解したデータプロフェッショナルとして行動規範と判断が身についている -引き受けたことは逃げずにやり切るコミットメント -迅速な報告や、報告に対する指摘のすみやかな理解 など ・データドリブンな分析的アプローチの基本が身についており、仮説や既知の問題が与えられた中で、必要なデータを入手し、分析、取りまとめることができる -データや事象のダブりとモレの判断力 -分析前の目的、ゴール設定 -データの出自や情報引用元の信頼性の判断 -目的に即したデータ入手 -分析結果の意味合いの正しい言語化 -モニタリングの重要性理解 など ・担当する検討領域についての基本的な課題の枠組みを理解できる -担当する業界の主要な変数 -基本的なビジネスフレームワーク など ・新しいデータ・AI利活用時代に即した基礎的な理解と価値創造についてのマインド・素養を持つ -新しい技術がもたらしている変革の概観の理解と関心 -データやAIがもたらす新しいリスクと課題への理解を踏まえた倫理的対応	・統計数理や線形代数、微分積分、集合理論の基礎知識を有している ・科学的解析やデータ分析の基礎知識を有している -比較対象の選択やデータが生み出される背景理解の重要性 -推定、検定 -分析用データの整備、標本抽出 -予測、グルーピングなどのモデリングと基本的な評価方法 - 時系列分析に関する周期性、トレンドなどの基本的概念 - データ品質の重要性に関する理解 など ・機械学習の基本的な概念を理解している -教師あり学習と教師なし学習の違い -機械学習における過学習の理解 -深層学習のメリットに関する理解 など ・適切な指示のもとに、データ加工と基礎的な分析を実施できる -基本統計量や分布の確認、および前処理 -クロス集計による偏りの把握や、単回帰分析の実行 など ・データ可視化の基礎知識を有している -軸だし -不適切な表現の理解 -意味合いの導出 など ・生成モデルの学習と活用について基礎的な理解ができている -生成モデルと識別モデルの区別 -利用されるデータの観点から起きうる基本的な問題の理解	・データやデータベースに関する基礎知識を有している -構造化/非構造化データの判別、論理モデル作成 -ER図やテーブル定義書の理解 -SDKやAPIの概要理解 -クラウドストレージにファイルを格納できる など ・数十万件程度のデータ加工技術を有している -ソート、結合、集計、フィルタリングができる -設計書に基づき、プログラム実装できる ・適切な指示のもとに、以下を実施できる -データベースから条件を満たすデータを抽出できる -インポート、レコード挿入、エクスポート - システムや予測モデルのテスト実施 ・セキュリティの基礎知識を有している - 機密性、可用性、完全性の3要素 - 暗号化、認証、認可の理解 - マルウェア、コンピュータウィルス、改ざんの脅威を理解 など ・AI活用を実践するためのITスキルを有している - 分析機能や学習済み予測モデルのAPI活用やプロンプトエンジニアリング - クラウド上の統合開発環境を活用した分析プログラミング
DS以前の方	・ビジネスは勘と経験だけで回すものだと思っている　・課題を解決する際に、そもそも	・基本統計量の意味を正しく理解していない ・線形代数や微分・積分の基本が理解できていない ・指数を指数で割り算したりする ・「平均年収」をそのまま鵜呑みにしたりする ・グラフ・チャートの使い方が不適切 ・大規模言語モデル（LLM）の出力を鵜呑みにする	・スプレッドシートで関数が使用できない ・繰り返しや分岐による処理フローが理解できない ・自分が使っているツールの仕組みに興味がない ・テキストや画像はデータでないと思っている ・データは自動的に整備されていると思っている

スキルカテゴリ

スキルカテゴリは、スキルを数十個に分類したものです（**図 11**）。

重要なスキルカテゴリは、プロジェクトの内容や役割によって異なります。必要に応じてスキルカテゴリを絞り込むことで、効率的にスキル項目を確認できるでしょう。

基礎的な知識としての「数学的理解」「科学的解析の基礎」、データそのものを取り扱うための「データ収集」や「データ加工」、そしてビジネスの課題解決につなげるための「着想・デザイン」と「アプローチ設計」など、複合的にスキルを習得することで実践力を身につけることができます。

図 11　スキルカテゴリの一覧

スキルカテゴリ一覧							
データサイエンス力	1	数学的理解	9	ビジネス力	1	行動規範	18
	2	科学的解析の基礎	51		2	論理的思考	18
	3	データの理解・検証	21		3	着想・デザイン	14
	4	データ準備	19		4	課題の定義	11
	5	データ可視化	38		5	アプローチ設計	25
	6	モデル化	106		6	データ理解	10
	7	モデル利活用	5		7	分析評価	5
	8	非構造化データ処理	32		8	事業への実装	7
	9	生成	14		9	契約・権利保護	9
	10	オペレーションズリサーチ	14		10	PJマネジメント	27
	項目数		**309**		11	組織マネジメント	15
					項目数		**159**
データエンジニアリング力	1	環境構築	33				
	2	データ収集	19				
	3	データ構造	11				
	4	データ蓄積	16				
	5	データ加工	15				
	6	データ共有	15				
	7	プログラミング	29				
	8	ITセキュリティ	20				
	9	AIシステム運用	11				
	10	生成AI	13				
	項目数		**182**				
項 目 数 合 計							**650**

図12　データサイエンス力のスキルカテゴリ

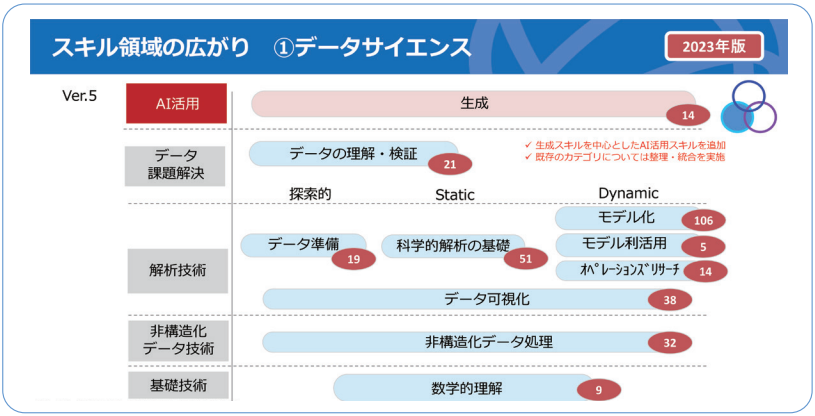


図13 データエンジニアリング力のスキルカテゴリ

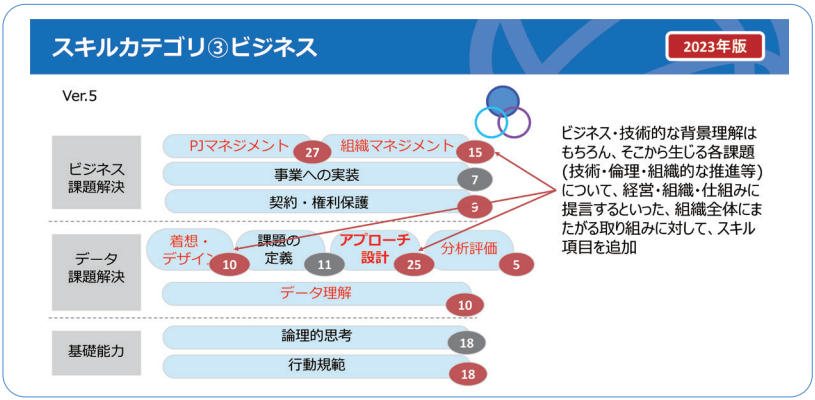
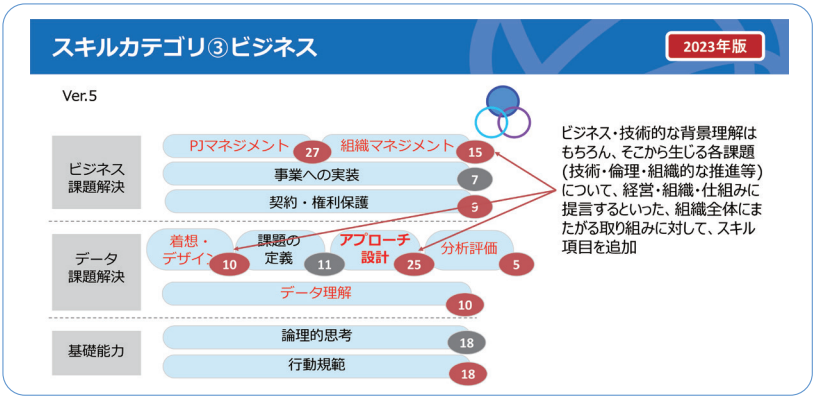


図14 ビジネス力のスキルカテゴリ



●スキルチェックリストの見方

スキルチェックリストの各カラムについて

スキルチェックリストは複数の項目で分類されています（**図 15**）。まずは、「スキルカテゴリ」、「サブカテゴリ」をざっと眺めて、気になる項目をチェックしてみましょう。

図 15　データサイエンス力のスキルリストからの抜粋

スキルの中分類（詳細は Part03「全体像を理解する」P.28）

スキルのレベル（詳細は Part03「全体像を理解する」P.25）

他分野にまたがるスキルには＊を付与

必須スキルには○を付与

AI 活用スキルには○を付与

▼ 他分野寄りのスキル

NO	SubNo	スキルカテゴリ	サブカテゴリ	スキルレベル	チェック項目	BZ	DE	必須スキル	AI活用	AI活用タイプ
282	9	生成	大規模言語モデル	★	大規模言語モデル（LLM）でHallucinationが起こる理由を学習に使われているデータの観点から説明できる（学習用データが誤りや歪みを含んでいる場合や、入力された問いに対応する学習用データが存在しない場合など）				○	⑦技術課題対応
283	5	生成	大規模言語モデル	★★	大規模言語モデル（LLM）を活用することが、独自に言語処理を開発することに対してメリットがある状況を、訓練データの収集コストや目的の汎用性、訓練に必要なコンピューターリソース（GPUやTPU）などの観点から説明できる		＊		○	⑥技術的理解
284	7	生成	大規模言語モデル	★★	目的特化型のモデルと基盤モデルの違いについて汎用性や学習方法（データ量や教師データの有無など）の観点から理解し、目的に合わせて使い分けの判断ができる				○	⑥技術的理解
285	10	生成	大規模言語モデル	★★	大規模言語モデル（LLM）でHallucinationが起こる理由を言語モデルとしての特性（尤度にもとづくワード生成）や学習手法（自己教師あり学習やRLHF等）、汎化（外挿に対する真実味）の観点などから説明できる				○	⑦技術課題対応
286	3	生成	大規模言語モデル	★★	基準となる指標や評価のためのデータセットなどの性能評価方針に従い、大規模言語モデル（LLM）の性能を評価できる				○	④作る-評価
287	11	生成	大規模言語モデル	★★★	Hallucinationの発生の抑制可否に対する現状と、抑制に向けた研究動向について理解している（Alignmentの問題とHallucinationの問題が表裏一体であることに対する理解など）				○	⑦技術課題対応
288	8	生成	大規模言語モデル	★★★	大規模言語モデル（LLM）とそれ以前の自然言語処理技術について、それぞれが得意とするタスクや処理性能、ドメインユースへのカスタマイズ可否などの観点から違いを理解し、自社組織の課題に対して適切な手法の導入を推進できる				○	⑥技術的理解
289	6	生成	大規模言語モデル	★★★	オープンソースの大規模言語モデル（LLM）を活用する上での留意点（入出力データの管理方法や性能面での限界など）と使い方（パラメータ更新可否や代表的なファインチューニング方法など）を理解し導入できる				○	③作る-実装

表3　スキルチェックリストの掲載項目

No	スキルセット全体の通番です。
Sub No	スキルカテゴリ内での通番です。
スキルカテゴリ	スキルを目的や技術のカテゴリに応じて分類したものです。
スキルレベル	各スキルが該当するレベルを見習いレベル（★）、独り立ちレベル（★★）、棟梁レベル（★★★）で表しています。
サブカテゴリ	スキルカテゴリを詳細に分類したものです。
チェック項目	スキルに求められる内容が記載されています。
DS、BZ、DE	他のスキルセットと分野がまたがる傾向にあるスキルの場合、* 印がつきます。たとえばビジネス力中心のキャリアを目指す場合には、ビジネス力のスキルだけではなく、他分野（データサイエンス力、データエンジニアリング力）でビジネス力に＊がついているスキルも確認してみましょう。 DS・・・データサイエンス力 BZ・・・ビジネス力 DE・・・データエンジニアリング力
必須スキル	スキルレベル達成のために必須のスキルです。共通に必要かつ重要なスキルにのみ○がつけられています。必須スキルを一つでも満たさない場合は、当該スキルレベルを達成できません。
AI 活用	生成 AI の普及をはじめとして、AI を活用する場面が増加しました。そのためのスキル「AI 利活用スキル」とまたがるスキルに○がつけられています。
AI 活用タイプ	「AI 利活用スキル」のカテゴリ (タイプ) で何に分類されるかを記載しています。

必須スキル以外にも重要なスキルはある！

たとえば、SQL（データベースの操作で使用する言語）はデータエンジニアリングの基本であり極めて重要なスキルではありますが、分析内容（自然言語処理、画像認識など）によっては使用されないこともあります。また、近年ではクラウドコンピューティングに関する知識や技術も重要なスキルですが、自社の情報セキュリティルールなどにより使用が禁じられている場合もあります。

そうした背景から、SQL やクラウド関連技術など、重要ではあるものの一部のスキルは必須スキルと定めていないものがあります。必須スキルではなくても、近年の技術動向を踏まえ、重要だと思われるスキルは身に付けるようにしましょう。

スキルチェックリストの読み方

スキルチェックリストは、さまざまな読み方ができます。全体の俯瞰やスキルアップに向けた流れを意識した読み方、未知の領域を学ぶための知らない用語の拾い読み……、用途に合わせて読んでみましょう。

●カテゴリ単位、サブカテゴリ単位の俯瞰

カテゴリ単位、サブカテゴリ単位で全体を俯瞰し、自分に関係するスキルを絞り込むこともできるでしょう。

●同一カテゴリのスキルレベルの流れ

同一カテゴリにおいて、「★」→「★★」→「★★★」といったスキルレベルのアップを目指した読み方もできるでしょう。

●同一カテゴリ、同一サブカテゴリ、同一スキルレベルの流れ

同一カテゴリ、同一サブカテゴリ、同一スキルレベルで、どのようなスキルがあるか、スキルの幅を広げる読み方もできるでしょう。

●未知の用語のチェック

スキルチェックリストは、見習いレベル〜棟梁レベルまで必要とされるスキルに関する用語の宝石箱です。関連するカテゴリ、サブカテゴリ内で知らない用語を見つけたら、スキルを獲得するチャンスと考えましょう。

Part 04

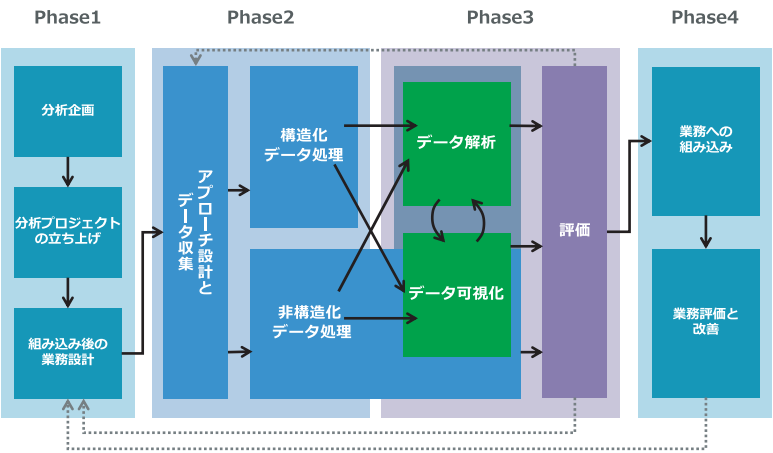
タスクリストの読み方

●全体の構造を理解する

データサイエンティストの仕事・業務

データサイエンティストが実際に行う業務は、大別すると４つのフェーズで構成されます（**図 16**）。Phase 1 では、企画立案からプロジェクトの立ち上げを行います。Phase 2 は、プロジェクトについてのアプローチの設計から、それにまつわるデータを収集し、処理するフェーズとなります。Phase 3 では、前フェーズで下ごしらえしたデータの解析および可視化を行います。そして最終段階となる Phase 4 では、業務への組み込み（実装）とともに、業務自体の評価・改善などを実施します。

図 16　タスクリストにおける４つのフェーズ



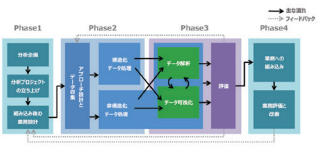
Phase 1	企画立案～プロジェクト立ち上げ
Phase 2	アプローチの設計～データ収集・処理
Phase 3	データの解析～データ可視化
Phase 4	業務への組み込み～業務の評価・改善

業務タスクはタスクリストとタスク構造図で確認

データサイエンティスト業務のタスクを確認する際には、「タスクリスト」と「タスク構造図」の両方を用います。

タスクリストには、データサイエンティストが担う一連のタスクが記載されています。タスク構造図には、データサイエンティスト業務の主要な流れと改善のために行うタスクの戻りについてが記載されています。

図 17　タスクリストとタスク構造図

資料	内容	イメージ
タスクリスト	データサイエンティストが担う一連のタスクを記載	
タスク構造図（中分類）	データサイエンティスト業務の主要な流れ（タスク中分類レベル）とタスクの戻りについて記載	
タスク構造図（小分類）	データサイエンティスト業務の主要な流れ（タスク小分類レベル）とタスクの戻りについて記載	

●タスクリストのカラム説明

タスクリストを確認しよう

タスクリストは、カテゴリー別に分類し、140 項目のタスクで構成されます（**表4**、**図 18**）。まず、「タスク中分類」と「タスク小分類」を眺めながら、データサイエンティスト業務にどのようなタスクがあるか確認してみましょう。

興味を持ったタスク分類があれば、「評価項目」欄でどのようなタスクを行うのかということを確認できます。

表4　スキルチェックテストの各項目

タスク大分類	データサイエンティスト業務におけるフェーズを示す
タスク中分類	データサイエンティスト業務におけるタスクの分類を示す
タスク小分類	タスク中分類の中のさらに詳細な分類を示す
No	タスクリスト上の「評価項目」（下行参照）の通し番号
評価項目	各々のタスクを記述した項目

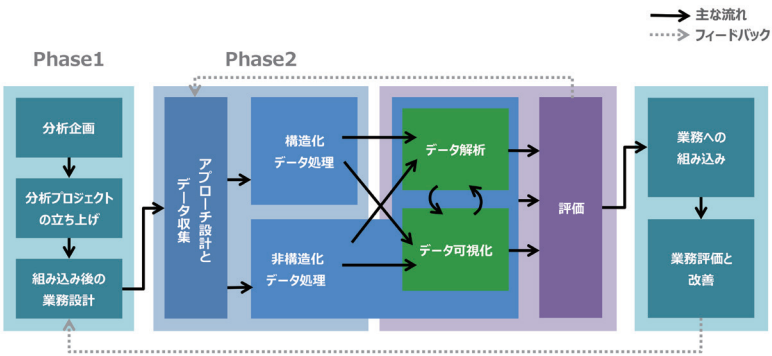
図 18　タスクリストの抜粋

タスク大分類	タスク中分類	タスク小分類	No	評価項目
DS_Phase1	分析企画	分析テーマの検討	1	事業環境（内部環境、外部環境）の整理を行い、データサイエンスを活用し、ビジネス課題を解決するためのビジョン・コンセプトを作成する
			2	ビジョン・コンセプトに基づき、ビジネス課題を解決するための分析テーマを検討する
			3	分析テーマを優先順位づけするための軸を定義し、取り組むべき分析プロジェクト（データサイエンスを活用し、課題解決を行う一連の取組、以下、プロジェクト）を決定する
			4	分析企画を外部に委託する場合は、秘密保持契約および業務委託契約を結ぶ
	分析プロジェクトの立ち上げ	前提条件の明確化	5	プロジェクトのステークホルダーを明らかにする
			6	プロジェクトの背景や問題意識を踏まえて目的と到達目標を明らかにし、ステークホルダーと共有する
		目標の明確化	7	プロジェクトの目標を設定する
			8	プロジェクトの目標に対する評価方法を具体化し、定量的な成功基準を設定するとともに、成功基準の判定時期・判定者を決定する
		推進体制設計	9	プロジェクトの難易度、メンバーのスキル、チームの総合的なバランスを考慮し、実施体制と役割分担を決定する
			10	プロジェクトに必要なコストとプロジェクトの実施によって得られる価値（コスト削減効果を含む）を算出する
			11	プロジェクトの実現性について評価・検討する（技術動向、費用対効果、業務負荷などを考慮）
			12	プロジェクトの実施計画を作成し、必要に応じてパートナーとの連携方針を具体化する
		コンプライアンス・倫理・権利の確認	13	プロジェクトで取り扱うデータや成果物が、最新の個人情報の扱いに関する法令や、その他のAI倫理やプライバシーの問題、プロジェクトに関連する業界の法令/ガイドラインを遵守しているか確認する
			14	プロジェクトで生み出された成果物の知的財産の扱いを整理する（独自手法の場合は、特許申請を検討するなど）
		計画の承認	15	プロジェクトの実施計画について、合意形成のための調整や説明を行い、ステークホルダーの承認を得る
			16	分析業務を外部に委託する場合は、業務委託契約を結ぶ
	組み込み後の業務設	業務設計	17	データ分析結果を利用・適用する対象業務のプロセスなどを把握/設計する
			18	取扱データ別にデータの利活用及び開示のガイドラインと管理・アクセス方法をステークホルダー間で設定する
			19	対象業務の運用体制や運用方法を決定する
			20	対象業務の目的や目標を確認し、モニタリング方針・方法やモニタリング時のKPIを決定する
		アーキテクチャ設計	21	対象業務をすでに実施している場合は、現状業務のシステム構成およびデータ仕様、業務諸元を確認する
			22	対象業務を実現するためのアーキテクチャを検討し、課題解決に向けたITシステム導入のロードマップを作成する
		環境整備	23	アーキテクチャ設計に基づき、プロジェクトに必要なクラウド環境・ハードウェア環境を整備する
			24	アーキテクチャ設計に基づき、プロジェクトに必要な通信環境を整備する
			25	アーキテクチャ設計に基づき、プロジェクトに必要なソフトウェア環境を整備する
			26	不必要な情報の漏れがないように暗号化を行い、防御態勢を解析開始前に整える
			27	プロジェクトに必要な情報共有/プロジェクト管理環境を整備し、必要に応じてSLAを締結する

タスク構造図を確認しよう

タスク構造図は、タスク分類を表す図形と、各図形をつなぐ矢印が記載されています（**図 19**）。タスク構造図には、「タスク構造図（中分類）」と「タスク構造図（小分類）」の 2 種類があります。データサイエンティスト業務の全体像を知りたい場合は、「タスク構造図（中分類）」を確認しましょう。データサイエンス業務の詳細なタスクの流れを知りたい場合は、「タスク構造図（小分類）」で確認できます。

図 19　タスク構造図（中分類）



図形	タスク分類を表す （中分類タスク構造図の場合：タスク中分類を表す 小分類タスク構造図の場合：タスク小分類を表す）
実線矢印	データサイエンティスト業務の主要な流れを示す
点線矢印	データサイエンティスト業務におけるタスクの戻りを示す

●タスクリストにも流れがある

タスクの矢印をたどってみよう

タスク構造図（**図 19**）には、多くの矢印が記されています。実はこの矢印が、データサイエンティスト業務を理解するうえで重要なポイントです。

データサイエンティストの業務は、一つ一つのタスクを順に進めるウォーターフォール型では進行しません。タスクを行ったり来たりするため、同じタスクを何度も繰り返し実施することもあります。また、途中でうまく行かなくなった場合は、かなり前のタスクにまで戻ることもあります。

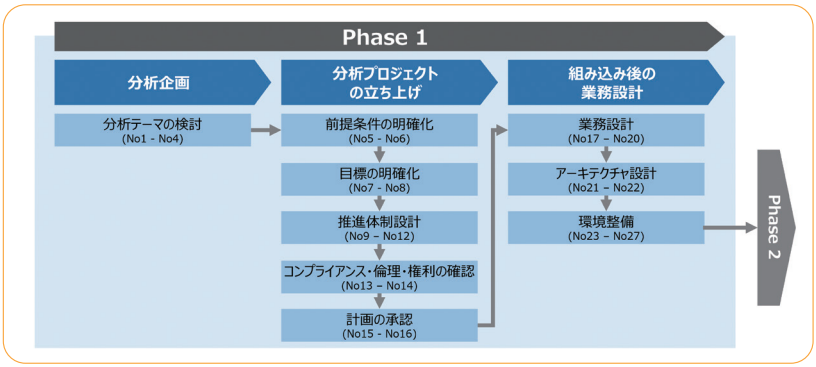
実際にタスクの流れを確認しながら、どのようにデータサイエンティスト業務が進められていくのか確認してみましょう。

Phase1：企画およびプロジェクトの立ち上げフェーズ

データサイエンティストの業務は「分析企画」からスタートします。取り組むべき分析テーマを決定し、分析プロジェクトを立ち上げます（**図 20**）。

データ分析結果を業務に組み込むための「業務設計」、「アーキテクチャ設計」、「環境整備」を行った後、次のフェーズに進みます。

図 20　Phase1 のタスクの流れ

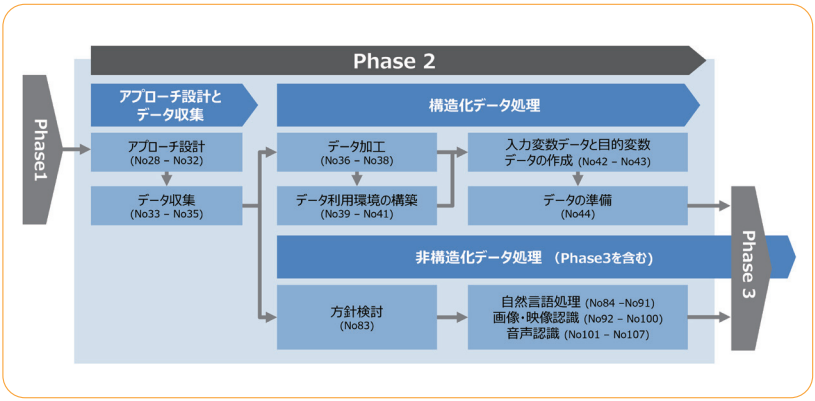


Phase2：アプローチ設計およびデータの収集 / 処理フェーズ

データ分析を始める前に、どのようにデータ分析を進めるのか「アプローチ設計」を行います。このアプローチ設計に基づいてデータを収集します。

次の「データ処理」に関しては、収集したデータの種類に応じてタスクが分岐します（**図 21**）。収集したデータが構造化データの場合は、「データ加工（タスク No36）」に進みます。収集したデータが、テキストや画像・映像、音声のような非構造化データの場合は、「方針検討（タスク No84）」に進みます。必要なデータ処理を終えたら、次のフェーズに進みます。

図 21　Phase2 のタスクの流れ



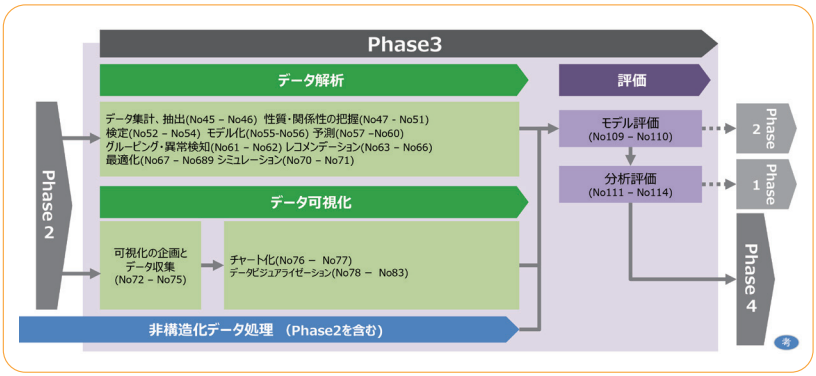
Phase3：データ解析およびデータ可視化フェーズ

それぞれの「データ処理」を行った後は、「データ解析」および「データ可視化」に進みます。「データ解析」タスクと「データ可視化」タスクは表裏一体で相互に行き来しながらタスクを進めるため、双方向の矢印になっています（**図 22**）。

非構造化データに関しては、「データ処理」と「データ解析」および「データ可視化」が密に連携しているため、「非構造化データ処理」の一部として「データ解析」、「データ可視化」を位置付けています。

データ解析・データ可視化後、「モデル評価」と「分析評価」を行います。「モデル評価」の結果が思わしくない場合は、Phase2 の「アプローチ設計とデータ収集」に戻り実施タスクを見直します。また、「分析評価」の結果が思わしくない場合は、Phase1 の「組み込み後の業務設計」まで戻って実施タスクを見直します。「モデル評価」および「分析評価」に問題がなければ、次のフェーズに進みます。

図 22　Phase3 のタスクの流れ

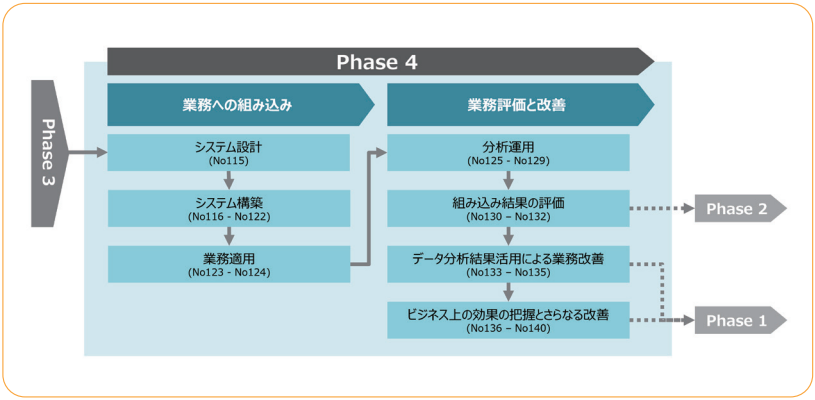


Phase4：業務への組み込みおよび業務評価 / 改善フェーズ

分析結果を業務で活用するために、ソリューションや分析システムを構築します。システム構築後は分析システムを運用しながら業務成果を評価し、必要に応じて改善を行います（**図 23**）。

「組み込み結果の評価」が思わしくなかった場合は、Phase2 の「アプローチ設計とデータ収集」まで戻り、再度タスクを実行します。また、「データ分析結果活用による業務改善」が思わしくなかった場合は、Phase1 の「組み込み後の業務設計」まで戻り、再度タスクを実行します。

図 23　Phase4 のタスクの流れ



すべてのタスクを実施する必要はない

タスクリストには数多くのタスクが記載されていますが、すべてを実施する必要はありません。それぞれの分析プロジェクトによって実施すべきタスクは異なるため、タスクリストを基にタスクを取捨選択します。

たとえば一部のタスクを外部に委託する場合は、「契約」に関するタスクを実施する必要があります（**表 5**）。一方で、外部に委託しない場合は、「契約」に関するタスクを実施する必要はありません。

また、Phase3 の「データ解析」や「データ可視化」に関するタスクは、分析目的に応じて適切なタスクをいくつか選択して実施します。Phase4 の「業務への組み込み」においても、すでに分析システムが存在しているのであれば、システム構築の一部タスクをスキップすることができます。

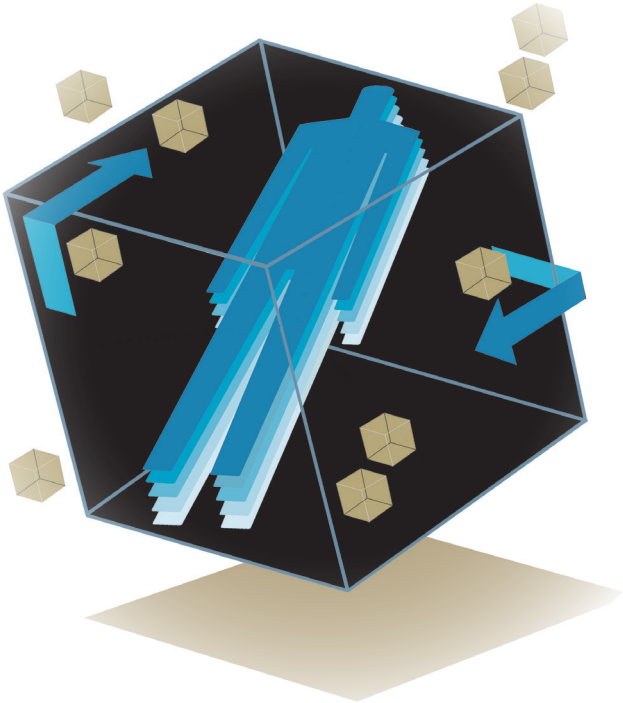
このようにタスクリストは、分析プロジェクトの目的に応じて、取捨選択して使うことを覚えておきましょう。

表 5　契約に関するタスク

No	評価項目
4	【分析企画契約】分析企画を外部に委託する場合は、秘密保持契約および業務委託契約を結ぶ
16	【分析 PJ 契約】分析業務を外部に委託する場合は、業務委託契約を結ぶ（データの取り扱い、知的財産の権利、成果物に対する責任の所在、情報開示、利用範囲など）
122	【開発契約】業務への組み込みを外部に委託する場合は、システム開発のための開発契約書を結ぶ（知的財産、データの取り扱いなど）
129	【運用契約】分析システム運用やモデル見直し業務を外部に委託する場合は、分析運用のための運用契約書を結ぶ（SLA など）

Part 05

AI 利活用スキルと
タスクリストの読み方



AI 利活用スキルとタスクリストの誕生の背景

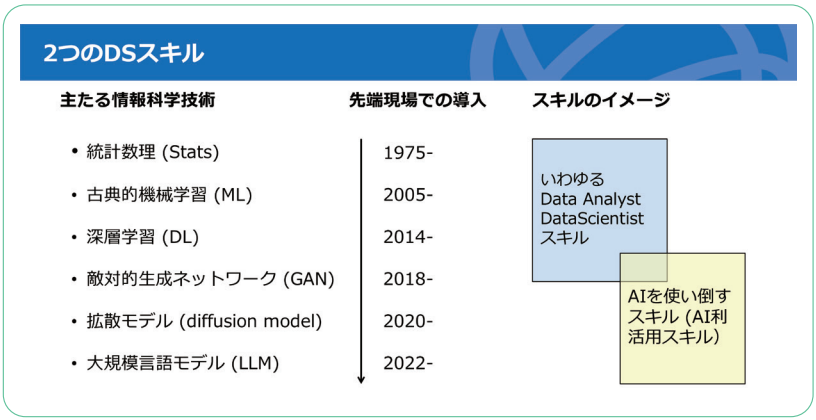
2023 年のスキルチェックリスト Ver.5、タスクリスト Ver.4 から AI 利活用に関するスキルとタスクを発表しました。生成 AI の登場により、アップデートをしました。

技術的な背景について少し触れると、AI の基礎となる情報科学技術はこれまでも著しく技術変化が起きています。2010 年代には深層学習（Deep Learning）や敵対的生成ネットワーク（GAN：Generative Adversarial Network）などにより、現実社会で実用可能な高度な AI が生まれうる時代に入りました。

ただ、これらは当初、高度なエンジニアリングやアルゴリズム理解が必要で、データサイエンティストやデータエンジニアでないと習得が困難でした。その後、2022 年に登場した大規模言語モデル（LLM）によって、AI 利活用が誰にでも可能な新たな局面に突入しました。（**図 1**）

今後は従来のデータサイエンティストに必要なスキルに加え、AI を使いこなすスキル、すなわち AI 利活用スキルが必要になります。また、生成 AI を業務で活用する際のタスクの考え方も今までのデータ分析

図 1　データサイエンティストスキルと AI 利活用スキル



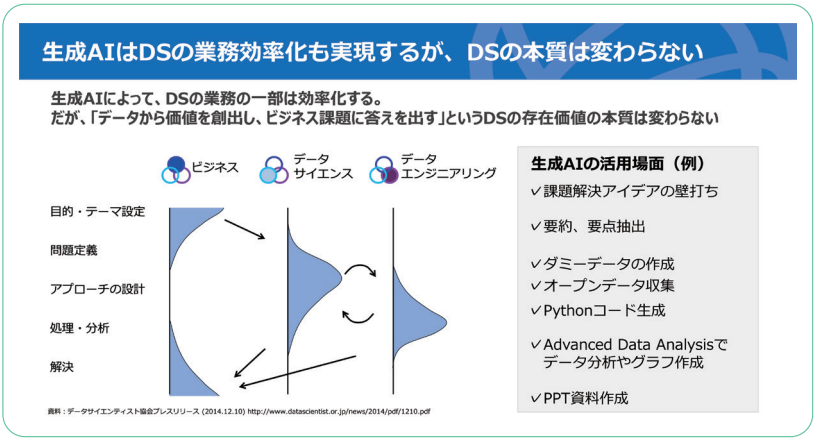
プロジェクトとは違うことが見えています。これらを明らかにするためにAI 利活用スキルとタスクリストを策定しています。

●データサイエンティストの本質は変わらない

ビジネス課題に向き合い、データサイエンスの社会実装を担うデータサイエンティストにとっては、実務で AI を使う / 使わないは手段の一つでしかない、と言えます。実際に一連のデータ分析プロジェクトの中での生成 AI の活用場面は**図2**のようにアイデア出しや要点整理からプログラミングや分析作業の補佐的な役割であることを示唆しています。超強力な補佐となり得る一方、データから価値を創出するのは人間の力です。よって、生成 AI はそれを補助する役割を果たすものの、業務効率化の範疇に留まる、というのが 2023 年発表時点での見解です。

一方、生成 AI を活用したサービス開発や業務効率化プロジェクトなど、AI 利活用に向けた様々な試みや実装が始まっており、データサイエンティストの業務も大きく変化するケースがあります。生成 AI の関連プロジェクトを企画・開発し、実際の業務やサービスの中でどのように AI を

図2　データ分析プロジェクトにおける生成 AI の活用場面例



利用するか？についてのスキルとタスクを明確にするために、AI 利活用スキルやタスクリストを参考にしてください。

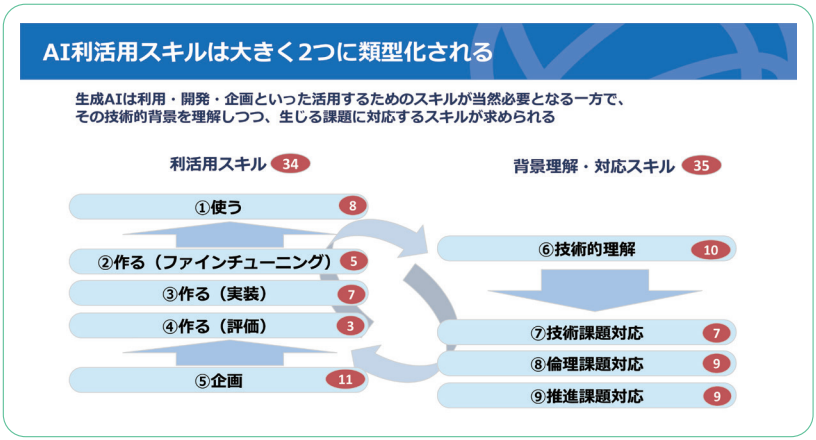
● AI 利活用スキルは主に２つに類型化される

生成 AI を業務に適用する際、単に、企画・開発・利用といった「利活用スキル」だけでは足りません。例えば、OpenAI 社の ChatGPT を企業として契約、利用するケースで考えると、

- どのようにデータを入力し、プロンプトを扱えば、業務で適用可能なレベルになるか？
- OpenAI 社に企業情報が流出するなどの事業リスクは？
- 今後の発展可能性に対し、変更可能なデータ基盤を整備できるか？
- 利用者への適切な教育や開発ガイドラインの策定など、開発時、利用時の対応は？

など、生成 AI の技術的理解の上、派生する様々な課題に対応するスキルも重要になると言えます。よって、**図３**のように AI 利活用スキルは

図３　AI 利活用スキルの分類

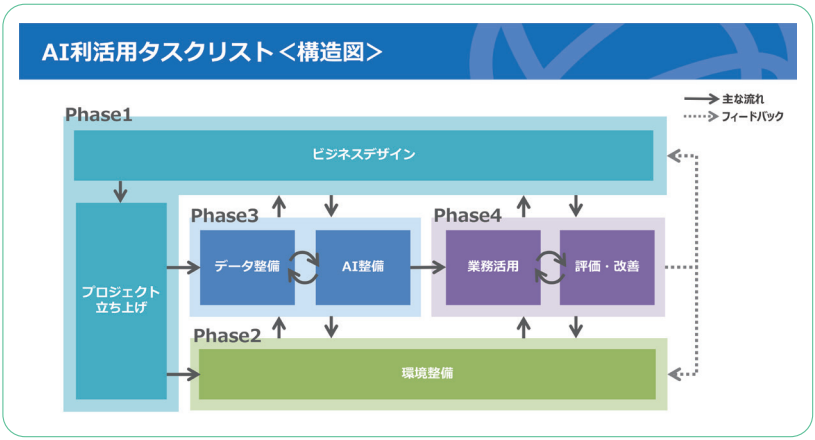


大きく２つに類型化され、全体では９つのスキル体系でまとめています。また、スキルの中ではビジネス（BZ）/ データサイエンス（DS）/ データエンジニアリング（DE）の何に近いスキルといえるか？というフラグがついています。加えて、大規模言語モデル（LLM）と画像生成モデル（Diffusion）では求められるスキルが違う場合があります。その見極めのためのフラグを設けていますので参考にしてください。

●タスクリストは分析タスクの前後でのタスクが特徴的

次に AI 利活用タスクリストです。**図4**のようになります。データサイエンスのタスクリストは Part4 で示したように左から右に、企画→設計→分析→評価→事業実装と流れ、必要に応じてフィードバックされる回路のような構造でした。それに対し、AI 利活用タスクリストは Phase1 のビジネスデザイン、Phase2 の環境整備が上下に大きく横たわり、いつでもこのタスクに戻る構造になっています。これも例で示すと、

図4　AI 利活用タスクリストの構造図



・既存の LLM で開発した独自 AI の評価段階で、より安全性の高い LLM の新機能が発表された。念のため、技術調査とビジネス検討を行った
・業務で AI を活用してみたが想定よりレスポンスが遅く、実行環境の見直しが必要

などのケースが発生します。結果、ビジネスデザインや環境整備タスクまで戻る態勢を整備する必要があります。これには、生成 AI は発展途上技術であること、「生成」であるがゆえに、実際に業務活用に移らないと評価や改善がしにくいこと、などが起因しています。

一方、一つ一つのタスクは従来のデータサイエンスタスクに比べ、大きく異なるものではないので、この AI 利活用タスクリストの構造の特徴を捉え、アジャイル型の推進体制を確保し、柔軟に推進することが重要です。

Part 06

スキルチェックリスト・タスクリストの使い方

●データサイエンティスト初心者・目指している人

データサイエンティストは、実はとても多様性のある職業の総称です。駆け出しのデータサイエンティストだけでなく、デジタルリテラシーを高めたい方や DX に関わる方も、スキルチェックリストを活用することで、データサイエンティストの全体像を把握し、自身に必要な取り組みを考えることに役立ちます。

データサイエンティストの全体像の把握

データサイエンティストのスキルセットは、①ビジネス力、②データサイエンス力、③データエンジニアリング力――の 3 つの分野に大別できます。それぞれにどのような違いがあるかをみてみましょう。スキルセットごとの特徴や役割を知ることは、自分自身が目指そうとしているデータサイエンティスト像を考え、より具体的に感じられるきっかけにもなります。

勉強計画を立てる

いずれかの分野のスキルセットにしぼってみても、具体的なスキルは多岐にわたっています。スキルも膨大にありますが、現在は分業も進んでいるためすべてを一人で遂行しないとならないというわけではありません。

スキルチェックリストには、データサイエンス業界で使われる用語がちりばめられています。そのなかから興味があるものを探して、まず何に着手するかを考えてみるとよいでしょう。

また、Part1 で紹介したデータサイエンティスト検定を受験することで自分の理解度を測ることができ、今後の進路やキャリアの実現においてサポートとなることでしょう。

タスクリストの使い方

タスクリストは、データサイエンティストが担う仕事の全体像を知ることができる、いわばガイドラインです。まだ自分が経験したことのない業務だとしても、タスクリストをどのスキルを学ぶべきかの目安として役立ててください。

●データサイエンティスト実務家

すでにデータサイエンティストとして活動している方にとっても、スキルチェックリストは、自分の守備範囲を把握し、強みを自覚しつつ、次のスキルアップへの道筋づくりに役立ちます。

自分のスキルのアセスメント

実務家のみなさんは、スキルチェックリストを自身のスキルのアセスメント（客観的な評価）として使うことができます。実際に自分がどのようなスキルを持っているかを確認することで、強みと弱みをはっきりと把握できます。

スキルアップに活かす

自身のスキルの強みと弱みを知ることは、スキルアップに活かすことにつながります。たとえば、

- 「見習い」レベルをクリアしていないスキルセットがあれば、まずその分野のスキル習得を急ぐ
- 主軸とする分野のスキルアップを狙う
- 全スキルセットの「★★」の必須スキルを埋める
- チームメイトのスキルの穴や、業務で活用できそうだがまた持っていないスキルを伸ばす

など、さまざまな計画が考えられます。

スキルチェックリストを参考に、自分なりの計画を立てるとよいでしょう。

タスクリストの使い方

タスクリストでは、自分が業務で担当してきた範囲や、仲間の担当範囲についてより理解を深めることができます。またプロジェクトマネージャーがどのような視点でプロジェクトを見ているかが擬似体験できるでしょう。

普段の業務における自分の担当範囲のみならず、挑戦できそうな領域があれば、そこに挑戦すべく計画を立ててみてもいいでしょう。

●データサイエンティストチームリーダー
●プロジェクトマネージャー

チームやプロジェクトを率いる立場の方にとっても、スキルチェックリストとタスクリストが役に立ちます。たとえばスキルチェックリストは、チームメンバーのスキルアセスメント（スキルの客観的な評価）に活用でき、そこからプロジェクトへの配置、育成計画にも役立ちます。またタスクリストは、プロジェクトの計画策定から失敗の原因特定、人員の配置計画にも役立てることができます。

チーム全体のスキルアセスメントと育成・採用

スキルチェックリストは、データサイエンス業務のスキルの一覧表です。メンバーにスキルのアセスメントを行ってもらうことで、今チームにどのようなスキルを持っている人がいるのかが把握できます。

またメンバーがデータサイエンティストとしてどのようなキャリアを歩みたいと考えているかを把握する際にも、スキルチェックリストの３つの分野のスキルセット（ビジネス力、データサイエンス力、データエンジニアリング力）は共通言語として大いに役立ちます。

各メンバーの現状のスキルと望んでいるキャリアパスに合わせて、適切なプロジェクトへの配置、育成計画への応用、採用計画の策定にも役立てることができます。

プロジェクト運営

タスクリストには、データサイエンティストが取り組むプロジェクトで行われるタスクが網羅されています。このなかから自身のプロジェクトに合うものを選び出せば、プロジェクトの計画策定にも大いに活用できます。

スキルチェックリストのなかには、プロジェクト運営に必要となるスキルも記載されています。データサイエンティストのプロジェクトは、技術的な課題解決はもちろん、多様な立場の人が合意形成（コンセンサス）を

図りながら進めるチームワークも重要であることを忘れないようにしてください。

さらにスキルチェックリストを利用すると、誰をどの工程に配置するか、いつまでにどういうスキルを持った人間が必要になるかなどを考えることができます。これらの事実を踏まえて、戦略的なプロジェクト運営が可能です。

また、過去の失敗や、想定より長い時間がかかったプロジェクトの原因究明にも利用できます。たとえばタスクに過不足があった、または実施の順序に問題があったということも特定できるかもしれません。

●外注計画担当者

データサイエンティストの業務領域は多岐にわたっており、専門性も求められるため、一部の業務を外注するケースもあります。その際、企業が求めるスキルと候補者が提供するスキルのミスマッチが起こることもあります。

スキルチェックリストを活用すれば、こうしたミスマッチを未然に防ぐことができます。

必要なスキルの理解と候補者のスキルのアセスメント

データサイエンティストに外注をする際には、まず全体の工程のどの部分を外注し、どの部分を内製するかを決める必要があります。タスクリストには、今後発生するタスクが系統的に記されているので、これを参考に外注工程を決めましょう。

その次に、どのようなスキルを持った人材が必要なのかを決めることになります。スキルチェックリストから必要なものを選び出すことで、求めている人材像をはっきりさせることができます。

次に、候補者となる方のスキルのアセスメントをしましょう。候補者にスキルチェックリストを直接使ってもらったり、履歴書や面談・面接の中で使ってアセスメントしたりすることもできるでしょう。また、人材エージェントにあらかじめスキル要件を伝えることで、効率的に人材を紹介してもらうことができます。

処遇・待遇を決める

データサイエンティストは新しい職種であり、業務内容も多岐にわたることから、適切な評価と報酬の設定がむずかしいこともあります。プロジェクトが要求するスキル水準がわかれば、市場のなかでどの程度の待遇を用意すべきかを検討する際の参考にできます。

●人事採用担当者・エージェント

人事採用担当者の候補者理解とスキルのアセスメント

ひとことでデータサイエンティストと言っても、実際には領域は幅広く、タスクも多様です。そのため、過去の経歴でデータサイエンティストとして活躍していたとしても、自社業務で活躍できるとはかぎりません。

そこで、タスクリストを参考に自社におけるタスクの全体像を確認し、強化したいタスクを実行できるスキルを持つ人材要件を定めましょう。また、役割が違うと給与も変わるので、適切な給与設定をするためにもスキルやタスクに基づいた職務内容の定義が必要になります。

転職支援エージェントにおける候補者アセスメント

求職者のスキルを、チェックリストを用いて確認します。また、求人要件の業務範囲であるタスクの経験があるかどうかを見極める際にもタスクリストが役立ちます。

●学校の先生

スキルチェックリストの使い方

対象となる学生が持つ専門性やこれまでの習熟度の違いを踏まえて、まずは「★」に絞ってスキルチェックリストを押さえていきましょう（高度な訓練や教育を受けている学生には、「★★」や「★★★」に進んでもらってください）。

社会経験が乏しい学生には、ビジネス力の項目では実感しにくい部分があるかもしれません。PBL* 型の授業などを通じてスキルを習得できるように心掛けてください。実際の課題を解決していくなかで、スキルチェックリストが持つ意味を実感できるでしょう。

スキルチェックリストがカバーする範囲はとても広いので、学生が気後れしてモチベーションを落とさないように気を配ってください。そして、スキルチェックリストを見ながら、学生自身がなりたいデータサイエンティスト像を膨らませることができるように指導していきましょう。

タスクリストの使い方

授業などでデータ分析を実践した際には、タスクリストを参照して現状がどのタスクに該当するかを理解してもらえるようにしましょう。

＊ PBL (Project Based Learning)：問題解決型学習。または課題解決型学習。

●学生（学ぶ側の人）

スキルチェックリストの使い方

自分自身のスキルをチェックしてみましょう。いきなり全部をチェックしようとせず、まず「★」からスタートするとよいでしょう。

ビジネス経験のない人にとって、ビジネス力の内容はハードルが高いと感じるかもしれません。「★★」や「★★★」を見ても、気後れしないでください。将来の自分はどのようなスキルを持ったデータサイエンティストでありたいか、どのような分野で活躍したいかを思い描きながら、あわてず焦らず、スキルを少しずつ身に付けていきましょう。

大学や高専の数理・データサイエンス・AI 教育で学ぶ内容（リテラシーレベル・応用基礎レベル）は、「★」のスキルに対応しています。授業の中で学んだ内容を身につけることができているか、チェックリストを見ながら確認しましょう。

また、新しく加わったスキルを見ることで、最近のトレンドを確認しましょう。

タスクリストの使い方

まずタスク構造図を眺めて、全体の流れとフィードバックループ（タスクの戻り）について理解しましょう。次にタスクリスト全体をどのようなステップでタスクが進行していくのかを考えながら見ていきましょう。

タスクリストのなかで「何を言っているのか、まったく想像つかない」という項目もあるかもしれません。しかしここで大切なのは、各フェーズの役割とタスク中分類を踏まえて目的に応じたタスクを実行することであり、タスクの流れとフィードバックループの関係を理解することです。

いきなりすべてのタスクが実行できるようにはなりません。ローマは一日してならず、着実に前に進んでいきましょう。

分析プロジェクトを担当する際にはタスクリストを参照し、現状のタスクをしっかり把握しましょう。

●育成制度を作ろう

重要スキルを定義する

スキルレベル列を「★」または「★★」または「★★★」でフィルタし、レベルごとに検討しましょう（**図 23**）。

図 23　スキルレベルのフィルタを設定する

「自社必須スキル」列を作り、スキルチェックリストの定義では必須スキルになっていないものでも、自社にとっては重要なスキル項目について「○」を追加しましょう（**図 24**）。

スキルチェックリストのスキル項目は、特定のドメインへの依存度が高いものや個別の手法については、利用頻度が高かったり重要であったりしても、必須の「○」がついていない場合があります。

図 24 「自社必須スキル」列を作成する

	A	B	C	D	E	F	G	H	I	J
1	スキルチェックリスト　2019年　改訂版 ＜ビジネスカ＞									
2							▼ 他分野寄りのスキル			
3	NO.	Sub N	スキルカテゴリ	スキルレベル	サブカテゴリ	チェック項目	DE	DS	必須スキル	自社必須スキル
4	1	1	行動規範	★	ビジネスマインド	ビジネスにおける論理とデータの重要性を認識し、分析的でデータドリブンな考え方に基づき行動できる			○	○
5	2	2	行動規範	★	ビジネスマインド	[illegible]			○	○
6	3	3	行動規範	★	ビジネスマインド	課題や仮説を言語化することの重要性を理解している			○	○
7	4	4	行動規範	★	ビジネスマインド	現場に出向いてヒアリングするなど、一次情報に接することの重要性を理解している			○	○
8	5	5	行動規範	★★	ビジネスマインド	ビジネスではスピード感がより重要であることを認識し、時間と情報が限られた状況下でも、言わば「ザックリ感」を持って素早く意思決定を行うことができる			○	○
9	6	6	行動規範	★★	ビジネスマインド	作業ありきではなく、本質的な問題（イシュー）ありきで行動できる			○	○
10	7	7	行動規範	★★	ビジネスマインド	分析で価値ある結果を出すためにはしばしば仮説検証の繰り返しが必要であることを理解し、粘り強くタスクを完遂できる			○	○
11	8	8	行動規範	★★★	ビジネスマインド	プロフェッショナルとして、作業量ではなく、生み出す価値視点で常に判断、行動でき、依頼元にとって真に価値あるアウトプットを生み出すことをコミットできる			○	○
12	9	9	行動規範	★	データ倫理	データを取り扱う人間として相応しい倫理を身に着けている（データのねつ造、改ざん、盗用を行わないなど）			○	○
13	10	10	行動規範	★★	データ倫理	チーム全員がデータを取り扱う人間として相応しい倫理を持てるよう、適切にチームを管理できる				
14	11	11	行動規範	★★★	データ倫理	データの取り扱いに関する、会社や組織全体の倫理を維持、向上させるために、必要な制度や仕組みを策定し、その運営を主導することができる				
15	12	12	行動規範	★	コンプライアンス	個人情報に関する法令（個人情報保護法、EU一般データ保護規則：GDPRなど）や、匿名加工情報の概要を理解し、守るべきポイントを説明できる			○	○
16	13	13	行動規範	★★	コンプライアンス	担当するビジネスや業界に関係する法令を理解しており、データの保持期間や運用ルールに活かすことができる	*			
17	14	14	行動規範	★★	コンプライアンス	個人情報の扱いに関する法令、その他のプライバシーの問題、依頼元との契約約款に基づき、匿名化すべきデータを選別できる（名寄せにより個人を特定できるもの、依頼元がデータ処理の結果をどのように保持し利用するのかなども考慮して）				
18	15	15	行動規範	★★	コンプライアンス	仮名化と匿名化の違いを理解しており、適切な方法で匿名加工情報を扱うことができる	*	*		
19	16	1	契約・権利保護	★	契約	請負契約と準委任契約の違いを説明できる				

提供する学習方法を決める

「自社必須スキル」列を「○」でフィルタしましょう（**図 25**）。

図 25　「自社必須スキル」列でフィルタを設定

スキルチェックリスト　2019年　改訂版 <ビジネスカ>

▼ 他分野寄りのスキル

NO	Sub No	スキルカテゴリ	スキルレベル	サブカテゴリ	チェック項目	DE	DS	必須 スキル	自社必須スキル
1	1	行動規範	★	ビジネスマインド	ビジネスにおける論理とデータの重要性を認識し、分析的でデータドリ				
2	2	行動規範	★	ビジネスマインド					
3	3	行動規範	★	ビジネスマインド	課題や仮説を言語化することの重要性を理解している				
4	4	行動規範	★	ビジネスマインド	現場に出向いてヒアリングするなど、一次情報に接することの重要性を				
5	5	行動規範	★★	ビジネスマインド	ビジネスではスピード感がより重要であることを認識し、時間と情報が も、言わば「ザックリ感」を持って素早く意思決定を行うことができる				
6	6	行動規範	★★	ビジネスマインド	作業ありきではなく、本質的な問題（イシュー）ありきで行動できる				
7	7	行動規範	★★	ビジネスマインド	分析で価値ある結果を出すためにはしばしば仮説検証の繰り返しが必要 し、粘り強くタスクを完遂できる				
8	8	行動規範	★★★	ビジネスマインド	プロフェッショナルとして、作業量ではなく、生み出す価値視点で常に 頼元にとって真に価値あるアウトプットを生み出すことをコミットでき				
9	9	行動規範	★	データ倫理	データを取り扱う人間として相応しい倫理を身に着けている（データの 用を行わないなど）				
12	12	行動規範	★	コンプライアンス	個人情報に関する法令（個人情報保護法、EU一般データ保護規則：GD 工情報の概要を理解し、守るべきポイントを説明できる				
21	6	契約・権利保護	★★	権利保護	AI・モデル開発において、既存の適切なガイドラインを参照・確認でき 「AI・データの利用に関する契約ガイドライン」など）				
25	1	論理的思考	★	MECE	データや事象の重複に気づくことができる				
28	4	論理的思考	★★	構造化能力	様々なデータや事象を、無層やグルーピングによって、構造化できる				
29	5	論理的思考	★	言語化能力	通常見受けられる現象の場合において、分析結果の意味合いを正しく言				
38	14	論理的思考	★	説明能力	報告に対する論拠不足や論理破綻を指摘された際に、相手の主張をすみ				
39	15	論理的思考	★★	説明能力	論理的なプレゼンテーションができる				
40	16	論理的思考	★★★	説明能力	プレゼンテーションの相手からの質問や反論に対して、説得力のある説				
52	5	課題の定義	★	スコーピング	主に担当する事業領域であれば、取り扱う課題領域に対して基本的な課 る（調査活動の5フォースでの整理、CRM課題のRFMでの整理など）				
58	11	課題の定義	★★★	スコーピング	仮説や可視化された問題がなくとも、解くべき課題を構造的に整理でき を特定できる				

自社必須 スキル

並べ替え

昇順　降順

色別: なし

フィルター

色別: なし

1 つ選択してください

および　または

1 つ選択してください

検索

(すべて選択)

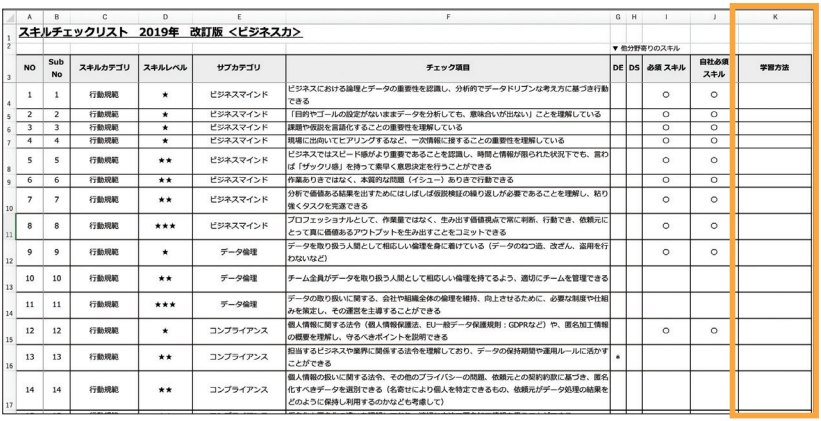
○

(空白セル)

フィルターのクリア

「学習方法」列を作り、学習方法を記載しましょう（**図 26**）。学習方法の種類としては書籍、ウェブサイト、研修、OJT などがあります。詳細な書籍名、サイト URL、研修名まで書きましょう。

図 26　「学習方法」列を作る



スキルチェックリスト　2019年　改訂版 <ビジネスカ>

▼ 他分野寄りのスキル

NO	Sub No	スキルカテゴリ	スキルレベル	サブカテゴリ	チェック項目	DE	DS	必須 スキル	自社必須スキル	学習方法
1	1	行動規範	★	ビジネスマインド	ビジネスにおける論理とデータの重要性を認識し、分析的でデータドリブンな考え方に基づき行動できる			○	○	
2	2	行動規範	★	ビジネスマインド	「目的やゴールの設定がないままデータを分析しても、意味合いが出ない」ことを理解している			○	○	
3	3	行動規範	★	ビジネスマインド	課題や仮説を言語化することの重要性を理解している			○	○	
4	4	行動規範	★	ビジネスマインド	現場に出向いてヒアリングするなど、一次情報に接することの重要性を理解している			○	○	
5	5	行動規範	★★	ビジネスマインド	ビジネスではスピード感がより重要であることを認識し、時間と情報が限られた状況下でも、言わば「ザックリ感」を持って素早く意思決定を行うことができる			○	○	
6	6	行動規範	★★	ビジネスマインド	作業ありきではなく、本質的な問題（イシュー）ありきで行動できる			○	○	
7	7	行動規範	★★	ビジネスマインド	分析で価値ある結果を出すためにはしばしば仮説検証の繰り返しが必要であることを理解し、粘り強くタスクを完遂できる			○	○	
8	8	行動規範	★★★	ビジネスマインド	プロフェッショナルとして、作業量ではなく、生み出す価値視点で常に判断、行動でき、依頼元にとって真に価値あるアウトプットを生み出すことをコミットできる			○	○	
9	9	行動規範	★	データ倫理	データを取り扱う人間として相応しい倫理を身に着けている（データのねつ造、改ざん、盗用を行わないなど）			○	○	
10	10	行動規範	★★	データ倫理	チーム全員がデータを取り扱う人間として相応しい倫理を持てるよう、適切にチームを管理できる					
11	11	行動規範	★★★	データ倫理	データの取り扱いに関する、会社や組織全体の倫理を維持、向上させるために、必要な制度や仕組みを策定し、その運営を主導することができる					
12	12	行動規範	★	コンプライアンス	個人情報に関する法令（個人情報保護法、EU一般データ保護規則：GDPRなど）や、匿名加工情報の概要を理解し、守るべきポイントを説明できる			○	○	
13	13	行動規範	★★	コンプライアンス	担当するビジネスや業界に関係する法令を理解しており、データの保持期間や運用ルールに活かすことができる		★			
14	14	行動規範	★★	コンプライアンス	個人情報の扱いに関する法令、その他のプライバシーの問題、依頼元との契約約款に基づき、匿名化すべきデータを選別できる（名寄せにより個人を特定できるもの、依頼元がデータ処理の結果をどのように保持し利用するのかなども考慮して）					

確認方法を定める

育成施策は多くの場合、成果として育成者数の報告を求められます。必要に応じて「自社必須スキル」を満たした人数がカウントできるように、「確認方法」を定めておきましょう。

図 27　確認方法を記載する

スキ改訂版 <ビジネス力>

▼ 他分野寄りのスキル

NO	サブカテゴリ	チェック項目	DE	DS	必須 スキル	自社必須スキル	学習方法	確認方法
1	ビジネスマインド	ビジネスにおける論理とデータの重要性を認識し、分析的でデータドリブンな考え方に基づき行動できる			○	○		
2	ビジネスマインド	「目的やゴールの設定がないままデータを分析しても、意味合いが出ない」ことを理解している			○	○		
3	ビジネスマインド	課題や仮説を言語化することの重要性を理解している			○	○		
4	ビジネスマインド	現場に出向いてヒアリングするなど、一次情報に接することの重要性を理解している			○	○		
5	ビジネスマインド	ビジネスではスピード感がより重要であることを認識し、時間と情報が限られた状況下でも、言わば「ザックリ感」を持って素早く意思決定を行うことができる			○	○		
6	ビジネスマインド	作業ありきではなく、本質的な問題（イシュー）ありきで行動できる			○	○		
7	ビジネスマインド	分析で価値ある結果を出すためにはしばしば仮説検証の繰り返しが必要であることを理解し、粘り強くタスクを完遂できる			○	○		
8	ビジネスマインド	プロフェッショナルとして、作業量ではなく、生み出す価値視点で常に判断、行動でき、依頼元にとって真に価値あるアウトプットを生み出すことをコミットできる			○	○		
9	データ倫理	データを取り扱う人間として相応しい倫理を身に着けている（データのねつ造、改ざん、盗用を行わないなど）			○	○		
10	データ倫理	チーム全員がデータを取り扱う人間として相応しい倫理を持てるよう、適切にチームを管理できる						
11	データ倫理	データの取り扱いに関する、会社や組織全体の倫理を維持、向上させるために、必要な制度や仕組みを策定し、その運営を主導することができる						
12	コンプライアンス	個人情報に関する法令（個人情報保護法、EU一般データ保護規則：GDPRなど）や、匿名加工情報の概要を理解し、守るべきポイントを説明できる			○	○		

確認方法については、以下を参考にしてください。

- 多くのスキルに該当する研修は、研修受講を確認の条件とする
- 書籍やウェブサイトが該当するスキルは、知識確認のテストを用意する
- OJT しか該当しないスキルは、経験を確認する記述式の設問を用意する
- 可能であれば「★」は研修とテストのみが望ましい
- 「★★」以降は経験を確認する必要性が高く、読む人（審査員）を必要とする
- 「★★」以降は研修受講を振るい落としの条件として使える
- 「★★」以降は、知識確認よりも、ケーススタディ、ハッカソン、コンテスト、アクションラーニング等での成果物で評価する

Part 07

リストにない定義や考え方

●ドメインスキル

汎用性の高いスキルを中心に定義

データサイエンティストが関わる業務は、多岐にわたります。データ活用の方向性を決めたり分析結果を読み解いたりするには、その業務に特有の知識が必要不可欠になります。しかし本スキルチェックリストでは、特定の業務で求められるスキルについては取り上げていません。本スキルチェックリストでは、なるべく特定のドメイン（業務の領域）に寄らない、汎用性の高いスキルを中心に定義しています。

たとえば、デジタルマーケティングに携わるデータサイエンティストにとってマーケティングの知識は必要不可欠ですが、品質管理に携わるデータサイエンティストにマーケティングの知識は必要ありません。逆に品質管理に携わるデータサイエンティストには、QC（Quality Control）の知識が要求されますが、デジタルマーケティングに携わるデータサイエンティストには、QC の知識は不要です。

データサイエンティストはあらゆる事業ドメインにおいて活躍しており、今後もさらなる活躍が期待されています。そのため、ドメイン知識を漏らさずスキルとして定義することは現実的には不可能に近く、また一般性のあるスキル定義も困難と考えられることから、スキルチェックリストでは定義していません。

スキルチェックリストを活用する際は、対象ドメイン固有のスキルを加えると、より有効に活用できます。

●文章に込められた意味

ビジネス力

ビジネスマインドに関するすべてのスキル項目（No.1 ～ No.10）は、多くが必須となっています。これには、データサイエンティストは価値創出にコミットし、実際に価値を生み出せる人材でなければならないという意志が込められています。

必須にはなっていませんが、スキル項目 No.5 の「わくわく」とは、未知の領域を探求する好奇心を感じる前向きな感情を指しています。データサイエンティストは、データサイエンスや AI の技術進化に対応するために、新技術やトレンドへの好奇心が重要で、好奇心がないと最新の進歩を見逃し、データの深い理解と斬新な解決策の提案能力を損なって競争力を失うリスクがあります。2023 年から生成 AI に関連するスキルが多く追加されているように、データサイエンティストは最新の技術も活用しながらデータから価値を生み出す、最適な案を提案できる人材であることが求められるのです。

また、スキル項目 No.6 の「ザックリ感」という表現は、時間と情報が限られているビジネスの現場において、迅速かつ最善の意思決定をして前へ進んでいくという姿勢を示しています。「仮説検証思考」、「イシュードリブン」といった思考・姿勢と関連付けて理解していただきたいスキル項目です。

その他、スキルカテゴリにある「着想・デザイン」のデザインとは、製品のデザインといった造形ではなく、「系」のデザインを意味します。予測モデルが勝手に価値を生み出すことはなく、予測モデルが業務フローのどの個所でどのような役割を果たすのか、全体の関連をデザインすることの重要性を表現しています。

データサイエンス力

データサイエンスを実践するにあたっては、「知識として身に付いている」こと、そして「実装できる」ことが必要です。

「説明できる」という言葉で表現されているスキルの場合、「★」では知識として「説明できる」レベルであることを意味します。これに対して「★★」における「説明できる」は、「実装できる」レベルに達しているスキルと知識が身に付いていることを意味します。そのようにデータサイエンス力では、次々に登場する新しく難解な技術を理解し、そして利活用することをスキル項目として整理しています。

たとえばデータサイエンス力 No.163 では、「★」レベルとして「『教師あり学習』『教師なし学習』の違いを理解している」というスキルを求めています。さらに、「★★」レベルになると「教師あり学習と強化学習の違いを、前提、定義、応用先といった観点で説明できる」となり、実装に必要なレベルのスキルが求められます。

また、No.184「不均衡データ（Imbalanced Data）がモデルに与える影響を理解し、サンプリングや評価／損失関数のチューニングなどを適切に対処できる」のように、「★★」レベルでは、特定の条件下において必要となるスキルが求められます。

なお、2023 年版では、カテゴリ・サブカテゴリをそれぞれ目的軸と技術軸として再編成した上で、実際のデータ分析の手順に沿うようスキルの順番を変更しています。

データエンジニアリング力

データエンジニアリング力では、「設計できる」「構築できる」といった表現を多用しています。

スキル項目に記載されている内容と同一の経験や実績がなくても、作業を行うために必要なスキルや知識があり、設計や構築の手順を説明できる状態にあれば「○」と評価してよいでしょう。

データエンジニアリング力のスキル項目では、記載する技術やアーキテ

クチャを具体的にイメージしてもらうために、比較的知名度の高い製品名やソフトウェア名を例示していることがあります。これらは理解促進のための例示であり、取り上げている製品の利用を推奨する意図はありません。

データエンジニアリングの分野において、一般的なシステムエンジニアとデータサイエンティストのスキルに違いはあるかということを耳にします。実際、システムエンジニアとデータサイエンティストの間では、多くのスキルが重複します。両者の違いは、システムエンジニアにとっては与えられた要件を元に高品質なシステムを実装する設計力・開発力が重視されるのに対し、データサイエンティストにとっては大規模で複合的なデータ分析環境構築に必要な要件を定義するスキルが重視されることです。

また、自然言語処理や画像・映像認識、音声認識などに代表されるような高度で専門性が高いエンジニアには、データサイエンティストとしての基礎的な部分を押さえることでプロジェクト全体を俯瞰し、チームとしての開発力を確実なものにする役割が求められます。

つまりデータサイエンティストには、より広い知識と上流の視点でシステムを俯瞰する能力が求められるのです。

タスク

タスクリスト Phase3 の「評価」タスクは、モデル評価 (No.109 ～ No.110) と分析評価 (No.111 ～ No.114) に分けています。

データ解析の評価というと、モデルの精度を評価する「モデル評価」を想像しがちですが、一連のデータサイエンティスト業務のなかでは、「モデル評価」と「分析評価」の両方を実施する必要があります。

活動目的に対する有効性・妥当性を判断する「分析評価」を忘れてほしくないため、明示的に「評価」を二つの項目に分けてタスクを定義しています。

●進化の速い分野

データサイエンティストに必要な技術・知識の広がりと進化・深化の速さに対応するため、スキルチェックリストやタスクリストの改訂のたびに、すべての分野において、スキル項目の追加・削除やレベルを精査・見直しをしています。

2023 年版では、生成 AI 時代における「AI 利活用スキル」としてハルシネーションやファインチューニング、マルチモーダル AI などに関する実践的なスキルをそれぞれレベル分けした上で 69 項目を定義しています。

ビジネス力

ビジネス力のなかでも、マインド、論理思考、課題定義といった本質的な問題解決での思考力に関する部分は、変化が少ないスキルカテゴリです。一方で、法令・契約・権利保護に関しては、常に最新のルールや世の中の動きを理解しておくことが求められます。たとえば、GDPR（General Data Protection Regulation：一般データ保護規制）が、2018 年 5 月 25 日施行されたことを受けて、2019 年版のスキルチェックリストでは、新たに関連するスキル項目を記載しました。

また、2021 年版では、AI 活用にまつわる倫理スキル（No.11,No.12）や着想・デザインスキル（No.45 ～ 51 など）を新規に追加し、ビジネス観点での AI 活用スキルを大きく変更しました。

さらに、2023 年版では、生成 AI 時代における AI 利活用スキルとして多くのスキルを追加しています。たとえば、倫理・法整備状況を踏まえた利活用に関する行動規範（No.13、14）、生成 AI を用いた新規ビジネス着想や課題解決設計のための着想・デザインスキル（No. 38、42 ～ 48）、アプローチ設計スキル（No.72、73、75 ～ 77、82 ～ 86）、データサイエンス・AI の推進課題に対応するための組織マネジメントスキル（No.148、154、156 ～ 159）などを追加しています。法令・契約・権利保護については、常に最新のルールを捉えておく姿勢が重要であると共

に、倫理的、実装的な側面での AI を扱うスキルの重要性が増しています。

データサイエンス力

データサイエンスの分野では、日々新たな技術や理論が誕生し実践されています。ほんの数年前までは理論の提唱と一部先進企業による実践にとどまっていた技術や理論も、活用に広がりを見せています。例えば、2019 年版のスキルチェックリストで追加されたディープラーニング技術の高度化による生成モデルの発展（No.183）や 2021 年版で追加した 自然言語処理モデルの Attention 機構に関する理解（No.207）については、今日では ChatGPT などの大規模言語モデルを使った生成 AI が日常的に活用され、スキルそのものもより身近なものになっています。

この生成 AI に関しては、そして 2023 年版では以降特に進化と利用拡大が早かったため、2023 年版ではこれらに関するスキルが多く追加されています（No.282 ～ 295）。特に、今後発展が予想されるマルチモーダルに関するスキル（No.202、213、216）や、生成 AI を活用する上で解決すべきハルシネーションに関するスキル（No.282、285、287）は、レベルごとに必要となるスキルが追加されています。

既存の技術や理論においても、データ活用が広がる中で活用・実践に進展が見られます。施策などの介入効果において、比較するグループ間での効果に対する影響の偏り（セレクションバイアス）をできる限り排除し、介入効果をより正確に推定するための因果推論などはその一例でしょう。従来、医療や教育の効果分析などを中心に研究や実践が進んでいましたが、今日では広告出稿などランダム化比較試験が難しいマーケティング施策の効果推定においても、様々な企業で実践されています。このような流れを受け、2021 年版では因果推論についてサブカテゴリを新設し、2023 年版ではさらにスキルを追加しました。（No.53、No.56）。

また、AI の進化や利用範囲が広がるなかで、ネットワーク、グラフデータに対する理解や実践の重要度が増しており、高等教育の情報Ⅰ・情報Ⅱの科目でも一部で取り上げられています。2023 年版ではグラフデータを扱うネットワーク分析をサブカテゴリとして新設し、関連するスキルも新たに定義しました。（No.240 ～ No.244）。

社会のあらゆるデータがAIの学習対象（訓練データ）として広がっていくなかでデータ品質の重要性がますます高まっていることを踏まえ、2023年版ではデータ品質に関する対策やその限界（No.190、191）についても追加しました。

一方、データ活用の広がりにより技術・理論の普及や新たなライブラリの登場などが進むことで、既存技術の難易度は低下することもあります。ロジスティック回帰や最尤法（No.147）に関する知識などは判別問題で広く使われるようになり、2019年版のスキルチェックリストでは「★★★」から「★★」にスキルレベルが変化しました。

このように、必要とされるデータサイエンス力のスキルは数学的理解など基礎となるスキルを除いて変化が早いため、スキルを高めていくには、スキルチェックリストの内容だけでなく、最新の技術動向や活用動向を学んでいくことが重要です。

データエンジニアリング力

現在、2023年より広まった生成AIに関連する取り組みが活発に行われているため、プロンプトエンジニアリング（No.170～173）、コーディング支援（No.174、175）、ファインチューニング（No.176～178）、生成AIの技術活用（No.179、180）、生成AI開発（No.181、182）という、生成AIに関連する多くのスキルを新たに定義しました。また、生成AIなどの大規模なモデルは、モデルを1から作るのではなく、学習済みのモデルをAPI経由で利用することが一般的なため、APIを通じて学習済みのモデルやサービスを利用するスキルをデータサイエンス力のスキルから移動、再定義して追加しました（No.123～126）。

さらに、2021年版のスキル定義時は、クラウドベンダーの提供するサービスメニューから要件に合うサービスを選択して分析環境を構築するスキルが重要になりつつありましたが、その後ますますクラウドサービスが発展し、分析環境を構築するだけでなく、クラウドが提供する機能を使いこなすスキル（No.25、No.117、No.132、No.165）も重要になっています。データサイエンティストは、新しい技術動向には常に目を配り、将来そうした最新技術が普及段階に入った際に、それを採用するか否かを見極める

眼力・能力が求められます。

タスク

タスクは、本質的な流れは変わらないため、細分化、および、新しい情報による更新されることはがはあっても骨格はあまり変化しません。

しかし、生成 AI の登場により、AI モデルを自ら構築するのではなく、基盤モデルと呼ばれる大規模な AI モデルを利活用する、という、これまでのデータサイエンスタスクとは働き方のことなるタスクが、新たにビジネス上重要になりました。

このため、従来のデータサイエンス タスクリストはそのままに、AI を利活用するためのタスクリストを、別途新たに定義しました（2023 年 10 月）。

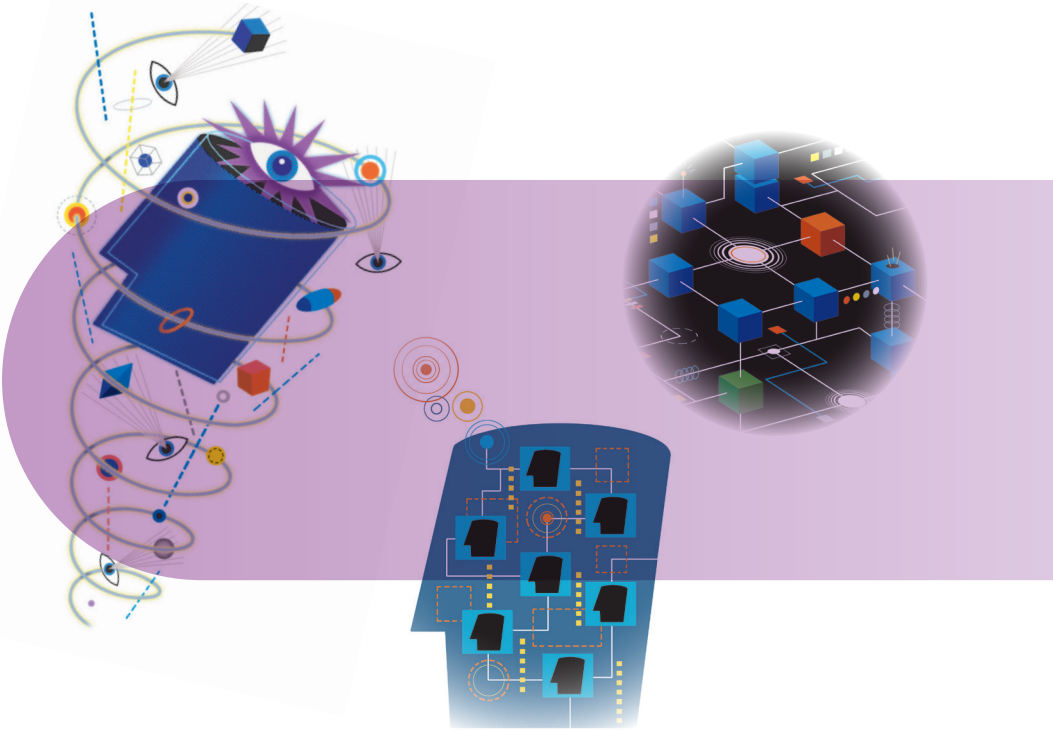
●ビジネスや事業への関わり方

データサイエンティストはビジネスパーソンたれ！

現在の日本社会では、データドリブンな経営推進や事業推進、デジタルトランスフォーメーションのためのデータ活用など、データサイエンティストが活躍できる場が飛躍的に広がっています。

その一方で、データサイエンティストが事業に加わっても、なかなか結果が出ないということもよく聞かれます。なぜなのでしょう？

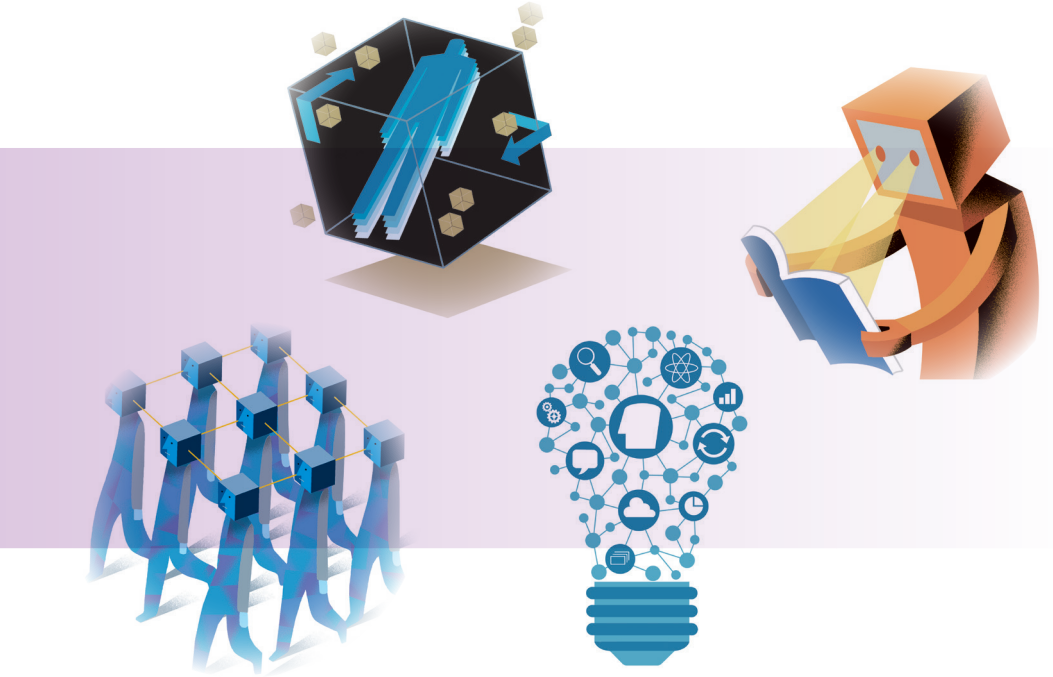
ビジネスシーンにおいては、「実践力」や「現場力」が重んじられることがよくあります。データサイエンティストはもちろん科学的に、論理的に物事を捉え、データの価値を最大化する役割を担っているわけですが、「現場」ではなかなか理解を得られにくいものです。



「現場」や「事業」、あるいは「経営」を背負っている人、チーム、組織のメンバーとの相互理解を深め、協働し、データの力で新しい事業価値を生み出すために、データサイエンティストは「現場」にもっと積極的に入り込むしかないのです。

それにはいろいろな方法があるでしょう。企業風土や組織の立ち位置によっても変わるかもしれません。たとえば、現場で通用するくらいの「現場実践力」を高めていくということもあります。それを「言語化」し、「ロジカルに説明」すること。情熱を持って数字や分析結果を武器にして、現場同士をネットワークし導くこと。現場や事業の中期目標に合わせ、目利き力を生かして先回りのデータ活用戦略を立てること……これらを体得することは、これからのビジネスパーソンにとって当たり前のスキルなのかもしれません。

ビジネス力にもそのエッセンスは含まれています。データサイエンティストは、常にビジネスパーソンでもありたいものです。



●お役立ち情報

一般社団法人データサイエンティスト協会では、さまざまなお役立ち情報を公開しています。ぜひご活用ください。

▼データサイエンティスト検定　リテラシーレベル

2021年9月からデータサイエンティスト協会で始まった検定試験です。スキルチェックリストで定義しているアシスタント・データサイエンティスト（見習いレベル:★）と数理・データサイエンス教育強化拠点コンソーシアムが公開している数理・データサイエンス・AI（リテラシーレベル）におけるモデルカリキュラムを総合し、実務能力と知識を有することを証明する試験となっています。データサイエンティスト初学者やこれから目指す学生、社会人の方を対象にしています。

毎年、3月、6月、11月前後に実施し、参考図書や対策講座も出ています。学んだことや実践していることの腕試しにもなりますが、これを学ぶことで実務的な能力を身につけ、社会の役に立てるデータサイエンティストを目指すきっかけとなることでしょう。詳細は協会HPに記載されています。
https://www.datascientist.or.jp/dskentei/

▼データサイエンス100本ノック（構造化データ加工編）

構造化データについて、無料で利用できるオープンデータが多数ありますが、データ分析力の実践という観点で環境整備がなされたものは、あまり多くありません。

分析実務においては構造化データの活用が多くを占めるという実態を鑑み、データと実行環境構築スクリプト、そして演習問題をワンセットにしてGitHubにて公開しています。
https://github.com/The-Japan-DataScientist-Society/100knocks-preprocess

▼スキルチェックリスト（GitHub版）

スキルチェックリストのさらなる普及を目的として、csv版や検索を行えるビューワーをGitHubにて公開しています。

・csv版スキルチェックリスト
https://github.com/The-Japan-DataScientist-Society/skills-checklist

・スキルチェックリストビューワー
https://the-japan-datascientist-society.github.io/skills-checklist-viewer/

▼一般社団法人データサイエンティスト協会

一般社団法人データサイエンティスト協会は、新しい職種であるデータサイエンティストに必要となるスキル・知識を定義し、育成支援や評価制度の構築など、高度 IT 人材の育成と業界の健全な発展への貢献、啓蒙活動を行っています。

また、所属を超えてデータ分析に関わる人材が開かれた環境で交流や議論をして自由に情報共有したり、意見を発信できる場を提供しています。

https://www.datascientist.or.jp/

▼デジタルスキル標準

また、各企業・組織の DX 推進等においてデータサイエンティストとしてどのような役割やスキルを担うことが求められているか、協力して推進する他の人材・役割はどのようなものがあるのか等については、情報処理推進機構（IPA）が公開しているデジタルスキル標準もあわせて活用してください。

https://www.ipa.go.jp/jinzai/skill-standard/dss/about.html

データサイエンティストのための
スキルチェックリスト / タスクリスト概説
Skill checklist & Task list Overview

2020 年 7 月 7 日　初版発行
2022 年 4 月 1 日　第二版発行
2025 年 7 月 1 日　改訂三版発行
監　修：一般社団法人 データサイエンティスト協会
発　行：独立行政法人 情報処理推進機構（IPA）
社会基盤センター人材プラットフォーム部
〒 113-6591 東京都文京区本駒込二丁目 28 番 8 号
文京グリーンコートセンターオフィス
e-mail：ikc-sx@ipa.go.jp
Web：　https://www.ipa.go.jp
制作 / デザインレイアウト：ACMEX INC. デザイン室
イラストレーション：ネモト円筆

引用・転載の際は、出典を明記してください。

監修：一般社団法人データサイエンティスト協会

監修 / 発行：独立行政法人情報処理推進機構（IPA）