

デジタルリテラシー④（最終） ー数理・データサイエンス領域ー

これまで本コラムにおいて、デジタルリテラシーの3領域のうち「IT・ソフトウェア領域」「AI・ディープラーニング領域」について述べた。残りの1つが「数理・データサイエンス領域」である。

数理・データサイエンス領域は、3領域の中で海上自衛官にとって最も重要であると言える。なぜならば、この領域の知識はDXなどデータの扱いが益々重要視される世の流れにおいて必要不可欠だからである。

IT技術の進展によりビッグデータを扱うためのハードウェア及びソフトウェアが整備されつつある中、ビッグデータに関し海上自衛官にとって重要なことは、収集した多種多様なデータを軍事面において漏れなく有効に利用していくことである。たとえ多くのデータを収集できる環境が整っていたとしても、数理・データサイエンス領域の知識をもってAIや統計学などデータを有効に利用する方法を理解していなければ、宝の山に入りながら手を空しくして帰るという結果になるのである。

前回までのコラムでも述べたとおり、国内全体に係る活動としては、経済産業省がオブザーバーであるデジタルリテラシー協議会が「国民全員が、デジタルを「使う人材」「作る人材」になるためにリテラシーを習得する¹⁾」ことを目指している。

デジタルリテラシーに関するコラムの最終回である今回は、重要性が高くなりつつある数理・データサイエンス領域について述べる。

1 全 般

数理・データサイエンス領域の基礎を理解するために必要な知識の範囲は、一般社団法人データサイエンティスト協会が、データサイエンティストスキルチェックリスト（以下、「スキルチェックリスト」という。）において体系的に定義・作成している²⁾。そして、リテラシーレベルの知識が習得できたか否かについては、アシスタントデータサイエンティストとしての実務能力と知識を有することを証明するための試験として2021年9月から実施されている「データサイエンティスト検定 リテラシーレベル（以下、「DS検定」という。）」を利用することができる。

DS検定はこれを取得することにより、アシスタントデータサイエンティストとしてデータサイエンスプロジェクトの担当レベルに必要な知識や実務能力、また、数理・データサイエンス・AI教育のリテラシーレベルの実力を有していることを示すことができるとされている。

¹ https://www.meti.go.jp/shingikai/mono_info_service/digital_jinzai/digital_skill_hyoujun/pdf/001_03_00.pdf

² https://www.datascientist.or.jp/common/docs/skillcheck_ver4.00_simple.xlsx

すなわち、DS 検定の出題範囲を学習することによって、デジタルリテラシー協議会が言うデジタルを「使う人材」として必要な知識を習得することができるのである。海上自衛官はデジタルを「使う人材」に該当することから、現場で勤務する者のみならず研究開発や防衛力整備に携わる者、更にはそれらを指揮・監督する者まで総員がリテラシーレベルの内容を習得する必要があると言える。

データサイエンティストに求められるスキルは、「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」の3つの分野に分けて定義されている。各分野において自分が何を学習すればよいかはスキルチェックリストの各分野のシートに記載されているチェック項目が参考になる。このうち、スキルレベルの★が1つ（見習いレベル）及び一部の2つ（独り立ちレベル）で示される項目が学習範囲となる。

スキルレベルの各項目の具体的な学習内容は、DS 検定受検対策用の書籍において網羅的に整理されている。ただし、他の領域の検定等と同様、あらかじめ一定の知識を備えていることを前提とした内容もあることから、人工知能や統計学に関する初歩的な内容であっても、専門の図書を用いた方が学習しやすい箇所もある。数多い資料の中から、自分が一番理解しやすいと思える資料を見つけ出すことが肝要であることは言うまでもない。

以下、スキルチェックリストに示されている3つの分野について、知識として有するためのポイントを述べる。

2 ビジネス力

ビジネス力とは「課題背景を理解した上でビジネス課題を整理し解決する力」をいい、アシスタントデータサイエンティスト（見習いレベル）に必要なスキルレベルとして次の3点が揚げられている。

- ① 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。
- ② 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。
- ③ 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。

また、上記スキルレベル習得以前の者は、以下のレベルにあると想定されている。

- I 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、そもそも定量化する意識がない。
 - ・データに付帯する権利や個人情報についての意識がない。
- II とりあえず AI を使えば大抵の課題解決ができると思っている。

この分野で学習する内容は、ビジネス課題に関係するものであるため、一見、海上自衛官にとって必要のない分野だと思えるが必ずしもそうではない。

具体的にスキルチェックリストに示されるチェック項目の内容を見れば、ビジネスの目的を考え、それを達成するための課題・仮説を検討し、原因・解決策を探るためのデータを準備する、といったデータドリブンな考え方に基づいた行動に直結する内容がビジネス力には含まれている。これは、データ分析が目的からぶれていないか、課題・仮説の検討から原因・解決策に至るまで論理的に考えられているか、といった内容を意味しており、これらを確認しつつデータ分析を進め、最終的にはその結果を意思決定につなげる、このような手法が海上自衛官の思考過程と一部共通点があることから、ビジネス力の分野は海上自衛官が習得すべき内容であると言えるだろう。

また、ビジネス課題を整理し解決する過程では、解くべき問題を見極め適切な解を見つけるための「問題解決力」、問いに対する主張と根拠を論理的に構成し筋道を立てて説明するための「論理的思考」、考え行動している自分自身を客観的に見ることとで本質的な課題に気づき問題解決するための「メタ認知思考」、ユーザーの行動を理解したうえで仮説の検証を素早く繰り返し問題解決につなげる「デザイン思考」といった考え方や基本的なフレームワークが内容に含まれ、これらを学習することは自己研さんに努める海上自衛官にとって自己の能力を向上させるための良い機会になると考える。

他にも、一次情報に接することの重要性、データを扱う者にとって相応しい倫理や個人情報に関する法令の改正など、海上自衛官が有すべき知識として必要な内容が多く含まれており、特に、データサイエンスのプロジェクトや分析プロセスなどは、研究開発に関係する配置やその管理者として勤務する者にとって必須の内容であると考ええる。

3 データサイエンス力

データサイエンス力とは、「情報処理、人工知能、統計学などの情報処理系の知恵を理解し使う力」をいい、アシスタントデータサイエンティスト（見習いレベル）に必要なスキルレベルとして次の5点が挙げられている。

- ① 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。
- ② データ分析の基礎知識を有している。
- ③ 機械学習の基本的な概念を理解している。
- ④ 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。
- ⑤ データ可視化の基礎知識を有している。

また、上記スキルレベル習得以前の者は、以下のレベルにあると想定されている。

- I 基本統計量の意味を正しく理解していない。
- II 線形代数や微分・積分の基本が理解できていない。
- III 指数を指数で割り算したりする。
- IV 「平均年収」をそのまま鵜呑みにしたりする。
- V グラフ・チャートの使い方が不適切。

例をとって統計学的な観点から習得以前のレベルを見てみる。例えばIVは、情報発信元が用いたグラフ等の資料を正しく読み取れないことを指摘するものと考えられる。具体的に述べると、国民の年収をヒストグラム³で表した場合、平均値を中心とした左右対称のような分布にはならず、平均値や中央値⁴より大幅に大きな値が存在する右に長い裾⁵を持つ分布で表される。データの分布のパターンと言え、一般的には単峰性⁶で左右対称の分布をイメージしやすく、統計学の知識がない状態で「平均年収」という言葉と数字だけを見た場合、国民の年収におけるデータ全体の分散に意識を配らず平均値 $\asymp$ 中央値 $\asymp$ 最頻値⁷といった誤った認識を持つ可能性があることをここでは指摘しているのだろう。

このようなデータを読み取る力となるデータサイエンス力の分野で学習する内容は広範囲であるものの、機械学習など AI に関する内容は「AI・ディープラーニング領域」の内容と重複する部分があるため、AI を既に学習した者にとっては学習範囲が狭くなる。一方で、文系出身者や数学をあまり学習していない者は、高校数学程度のおさらいが必要になる。線形代数や微分積分などの基礎数学、確率分布などの基礎統計学を中心に学習すれば、データを扱った予測や推定・検定⁸といった内容も理解できるようになる。内容を見れば最初は難しく感じるかもしれないが、国民全員が習得すべき程度であることを踏まえれば、海上自衛官は理解すべきものとして捉える必要があるだろう。

重要な点は、収集したデータからどのように内容を読み取ることができるかであり、そのための手段として数学的・統計学的知識が必要であることから、多くのデータを扱う海上自衛官としてデータサイエンス力の分野は理解しておかなければならない。必要な知識がなければデータを分析するに至らず、学習はそのことに気づかせてくれる。

³ データがとる値を複数の区間に分割し、各区間に入るデータの数を棒グラフで示したものをいう。

⁴ データを大きさ順に並べたときに中央に位置する値をいう。

⁵ 「裾が長い」とは、平均から遠い位置に標本が散らばっていることを意味する。

⁶ グラフの山が一つであることを意味する。

⁷ 最も頻繁に現れる値をいう。

⁸ 母集団（調査の対象となるすべてからなる集合）の特性についての予測（仮説）が正しいか否かを標本（母集団から抽出された一部の集合）データから判断する方法をいう。

4 データエンジニアリング力

データエンジニアリング力とは、「データサイエンスを意味のある形に使えるようにし実装・運用できるようにする力」をいい、アシスタントデータサイエンティスト（見習いレベル）に必要なスキルレベルとして次の4点を挙げている。

- ① データやデータベースに関する基礎知識を有している。
- ② 数十万件程度のデータ加工技術を有している。
- ③ 適切な指示のもとに、データベースから条件を満たすデータの抽出、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ができる。
- ④ セキュリティの基礎知識を有している。

また、上記スキルレベル習得以前の者は、以下のレベルにあると想定されている。

- I スプレッドシートで関数が使用できる。
- II 繰り返しや分岐による処理フローが理解できない。
- III 自分が使っているツールの仕組みに興味がない。
- IV テキストや画像はデータでないと思っている。
- V データは自動的に整備されていると思っている。

データエンジニアリング力の分野においてスキルレベルの内容を見ると、とても専門的な内容であるように思えるが、これらはビジネスパーソンにとって常識的な範囲になりつつある。

海上自衛官は一部の者を除きエンジニアである必要がないため、総員が深く掘り下げた内容まで学習する必要はない。ビッグデータが持つ「データとしての価値」を余らせることなく利用するために必要な、システムの要件定義を見据えたデータの整理やシステムの基礎的なプログラムの設計・実装ができる程度の知識技能を習得しておけばよいのである。

この分野も一部他の領域で学習する内容と重複する部分がある。例えば、セキュリティの基礎的な内容は、IT・ソフトウェア領域の範囲と一部重複し、AIシステムの運用に関する内容は、AI・ディープラーニング領域と一部重複する。既に学習した内容であるならば、おさらい程度の感覚で学習することで足りるであろう。

以上、数理・データサイエンス領域の知識について述べた。

海上自衛官は、戦術面で使用されているものから作戦や戦略における判断に必要なものまで、日常から幅広い範囲においてデータに触れている。任務を達成するためにこれらのデータを正しく読み取り有効に利用するためには、そのツールとして何らかの基礎的知識が必要であることは言うまでもない。

一方で、一般社会に目を向けると、技術の進展によりビッグデータという概念が用いられ、今後もあらゆるセンサーやその他情報源から収集できるデータは質・量ともに拡大していく傾向にある。これらを踏まえると、数理・データサイエンス分野の知識を海上自衛官が習得することは必須なのである。

本コラムを含めこれまで4回にわたり海上自衛官は国民の一員として例外なくデジタルリテラシーの習得が必要だということを述べてきた。データサイエンティストのレベルは4段階で定義されている⁹が、海上自衛官総員が習得すべきレベルは、必要最低限のものとして国民全員の習得を目指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）である。

冒頭でも述べたとおり、デジタルリテラシー協議会は日本のデジタル人材育成を加速させようと努力している。DX 推進という世の流れにおいて、海上自衛官はその職務の性質からデジタルリテラシー習得が必須であるが、リテラシーレベルであれば隙間時間を利用して学習することで必要な知識を身につけることができる。

経歴管理や課程教育に頼ることなく、自己研さんによってすべての海上自衛官がデジタルリテラシーを持つことが、限られた資源の中でその能力を最大限に発揮できる組織の構築につながるのではないだろうか。

（海上自衛隊幹部学校 福山 崇）

（本見解は、海上自衛隊幹部学校における研究の一環として執筆者個人が発表したものであり、防衛省・海上自衛隊の見解を表すものではありません。）

⁹ レベルが高いものから順に、シニアデータサイエンティスト（業界を代表するレベル）、フルデータサイエンティスト（棟梁レベル）、アソシエイトデータサイエンティスト（独り立ちレベル）、アシスタントデータサイエンティスト（見習いレベル）の4段階である。