

日本語仮訳:

# データ集約型科学に必要な デジタル人材の育成

経済協力開発機構(OECD)

科学技術イノベーションポリシーペーパー (90 号)

Original:

BUILDING DIGITAL WORKFORCE CAPACITY AND SKILLS FOR  
DATA-INTENSIVE SCIENCE  
OECD SCIENCE, TECHNOLOGY AND INDUSTRY  
POLICY PAPERS, No.90, OECD Publishing

原著は、OECD 科学技術産業局 ポリシーペーパー 第 90 号として、" BUILDING DIGITAL WORKFORCE CAPACITY AND SKILLS FOR DATA-INTENSIVE SCIENCE ", の題目で英語で発表されたものである。

© OECD 2020, <https://doi.org/10.1787/e08aa3bb-en>.

本翻訳は OECD が作成したものではなく、OECD の公式翻訳ではない。翻訳の品質および著作物の原文との整合性は、翻訳の著者に責任がある。原著と翻訳との間に齟齬がある場合は、原著の本文のみを有効とする。

© 2020 文部科学省による翻訳

Originally published by the OECD in English under the title: OECD (2020), "Building digital workforce capacity and skills for data-intensive science", OECD Science, Technology and Industry Policy Papers, No. 90. © OECD 2020, <https://doi.org/10.1787/e08aa3bb-en>.

This translation was not created by the OECD and should not be considered an official OECD translation. The quality of the translation and its coherence with the original language text of the work are the sole responsibility of the author or authors of the translation. In the event of any discrepancy between the original work and the translation, only the text of original work shall be considered valid.

© 2020 Ministry of Education, Culture, Sports, Science and Technology, Japan, for this translation

## 目次

アブストラクト.....	2
序文.....	3
謝辞.....	5
エグゼクティブ・サマリー .....	6
データ集約型科学に必要なデジタル人材・スキルの構築 .....	9
1. はじめに.....	9
2. データ集約型科学に対するデジタル人材のニーズについて何がわかっているのか? .....	10
2.1. デジタルスキルを備えた研究人材に対するニーズ.....	11
2.2. データ集約型科学におけるデジタルスキルの需要と供給.....	13
3. デジタルスキルを備えた社会と科学のためのデジタルスキル人材 .....	15
4. 科学のエコシステムとケーススタディ .....	17
4.1. 国際的な取組.....	19
4.2. 国内向けの取組.....	20
4.3. 分野別の取組.....	21
4.4. 大学（個別あるいは複数大学）の取組.....	22
5. デジタルスキルを備えた研究人材を育成するには何が必要なのか .....	22
5.1. 定義化の必要性：デジタルスキル、フレームワーク、役割 .....	23
5.2. トレーニングの提供.....	26
5.3. コミュニティ形成 .....	29
5.4. キャリアパスと報酬の構造.....	31
5.5. デジタル人材に関する外的要因.....	33
5.6. 目標に向けた各アクション範囲のリンクづけ.....	34
6. それぞれのアクターに対する提言 .....	35
6.1. 国または地方政府 .....	37
6.2. 研究ファンディングエージェンシー .....	38
6.3. 専門的な学会やアカデミア .....	40
6.4. 研究機関と基盤.....	41
6.5. 国際協力.....	42
7. 大学への提言 .....	43
8. 結び：一貫した政策的措置の必要性.....	47
参考文献 .....	49
添付資料 .....	56
専門家グループメンバー .....	56
ケーススタディ質問票 .....	57
ワークショップ参加者 .....	58
用語集 .....	59

## アブストラクト

本報告書では、データ集約型科学に必要な人的資源について検討を行い、特に公共部門で実施された研究や関連する課題、およびトレーニングへのニーズに焦点を当てている。デジタル化は、科学によって推進されると同時に、科学的実践のあらゆる側面に影響を与えるものである。データへのアクセスを含むオープンサイエンスは広く推進され、サイバーインフラストラクチャーやデジタルプラットフォームへの投資が増えているが、これらのツールを完全に活用するために必要な研究者や研究支援の専門家が備えるべきスキルに十分に注意が払われているとは言えない。本報告書作成の最終段階の際に起こった新型コロナウイルス感染症拡大の事態は、データ集約型科学が極めて重要であることを浮き彫りにし、また、科学界全体としてデジタルの能力とスキルの強化に向けて戦略的なアプローチが必要であることを認識させた。本報告書では様々な関係者に対する政策提言のほか、優れた取組の具体例を示している。

## 序文

過去10年にわたるオープンサイエンスに向けた全体的な動きの一部として、データへのアクセスやデータ集約型研究の促進に向けた政治的関心が高まり民間の投資が増えてきた。この動きは、科学界全体にわたって適用されてきた情報とコミュニケーションテクノロジーの急速な発展と並行している。これにより、新たな科学的ブレークスルー、科学的結果の透明性と再現性の向上、そして科学と社会に全体的な利益をもたらす革新性の向上が約束される。この約束は実現し始めているが、適切なスキルを持つ科学人材が不足していることにより、多くの分野で制約が生じている。

本報告書は、OECDグローバルサイエンスフォーラムから委託され、次のことを特定した。

- データ集約型科学に必要なデジタルスキルへのニーズ
- これらニーズの進展に応じて持続可能な能力を構築するための課題
- これらのニーズに対処するためにさまざまなアクターが実行できる政策的アクション

本報告書の主な焦点は、公共部門/学術研究の要件にあり、データ集約型の研究に必要なスキルの多くは、ビジネスや業界を含むさまざまな部門間で移転できることが認識されている。本報告書では、特に公共部門あるいは学術研究で必要なスキルを対象としたが、データ集約型科学に必要なスキルの多くは、民間や産業界といった異なるセクター間で移転可能であることが認識されている。また、一般的なデジタルスキルや学部教育や学校教育の役割よりも、卒業後の科学研究で必要とされる専門的な能力やトレーニングについて検討している。デジタルスキルに関連する教育の役割はOECDやその他機関のレポートで言及されているため、本報告書では対象としない。

本報告書が最終段階に入った際、新型コロナウイルス感染症が世界中の何百万もの人々の命や暮らしを脅かし、世界は劇的な変化を迎えることとなった。これによって人々、ビジネス、国の諸機関、科学研究がデジタルツールやデータへの依存度を増しているかが明らかになった。これらツールを使うためのスキルや能力を備えている人々は、パンデミックが引き起こす大規模な混乱にうまく対処することができると言える。パンデミックの経過を把握し、政策決定に情報を与えるために、データ集約型科学が重要であることが明らかになった。様々な科学分野からのデータや情報の統合や新しいソフトウェアの開発は、長期的な社会経済的な効果のモデルを作り評価するのに必要である。ウイルスの機能を理解し診断や治療法、ワクチン開発に向けた基礎研究や国際協力において、デジタルツールは重要な役割を果たしている。同時に、個人データを回収し、リンクづけ分析するツールの能力は、新たな倫理や法的な問題を引き起こし科学や社会に重要な示唆を与えている。

研究者だけではなく、政府、産業、その他幅広いコミュニティ向けに、研究方法や成果へのアクセスを容易にするデジタルプラットフォームを国際的な研究コミュニティが開発し、新型コロナウイルス感染症に迅速に対応した。これらの取組は新しいものではあったが、重要な研究へのオープンアクセスを提供するインフラが未整備であることを示すことにもなった。新型コロナウイルス感染症に対する科学の対応が成功するかどうかは、適切に整理され、相互運用可能かつ保存された状態のリソースの有無にかかっている。これらのリソースを作り、活用するためにはスキルを有する人材が必要であり、研究支援の専門家と研究者の両方が必要である。本報告書に記載のあるいくつかのケーススタディ（バイオインフォマティックのネットワークであるELIXIRを含む）は新型コロナウイルスへの科学研究の最先端であり、デジタルスキルと能力の強化においてそれらの経験は重要な教訓となりうる。

新型コロナウイルス感染症の拡大はデータ集約型科学の重要性と可能性を浮き彫りにしている。いずれの国も科学のためのデジタルスキルと能力を最優先にし、国際的に協力して行うことが必要である。その意味において、2019年に本報告書の初稿を作成した時よりも現在のほうが、本報告書の政策提言の重要性が増している。

## 謝辞

このプロジェクトを監督および実施するために、GSF加盟国からの指名により、国際的な専門家グループ（EG）が設立された。EGはミシェル・バーカー（オーストラリア）が議長を務め、本報告書はEGの成果物である。本報告書は議長によって起草され、他のEGメンバー、OECD-GSF事務局であるカセージ・スミス、田村嘉章が作成に携わった。

EGメンバーに加え、多くの外部専門家が重要な貢献をした。これには、本報告書に含まれる13のケーススタディのプロジェクトリーダーが含まれ、彼らはインタビューを通じて彼らの経験を提供するとともに、最終報告書についても意見を述べた。これらのプロジェクトリーダーの多くは、2019年10月ドイツのケルンで開催されたワークショップにも参加した。このワークショップは、クリストフ・ウルフ（GESIS会長）とイングビル・モックマン（EGメンバー）の協力により、ライプニッツ社会科学研究所（GESIS）で開催された。

## エグゼクティブ・サマリー

データ集約型科学は新しい時代に移りつつあり、その期待は大きい。ビッグデータ分析と新しい形式のデータへのアクセスは、すでに重要で斬新な科学的洞察を提供し、イノベーションの推進力となっている。異なる分野のデータを結びつけ分析することにより、国連の持続可能な開発目標でもある複雑な社会的課題を科学的に分析することが可能となりつつある。データ集約型科学はより持続可能で豊かな未来への変換に必要な新しい知識を生み出す可能性を秘めている。

デジタル化は、科学のすべての分野だけでなく、学術研究の他の分野にも変革をもたらしている。研究に重要かつ新しい方向性を与えるだけでなく、研究の透明性、厳格性、公正性の向上にも貢献している。デジタルテクノロジーやデータへのアクセスは変革を促し、人々のデジタル能力やスキルは将来の科学的な成功を左右する重要な決定要因となってきている。

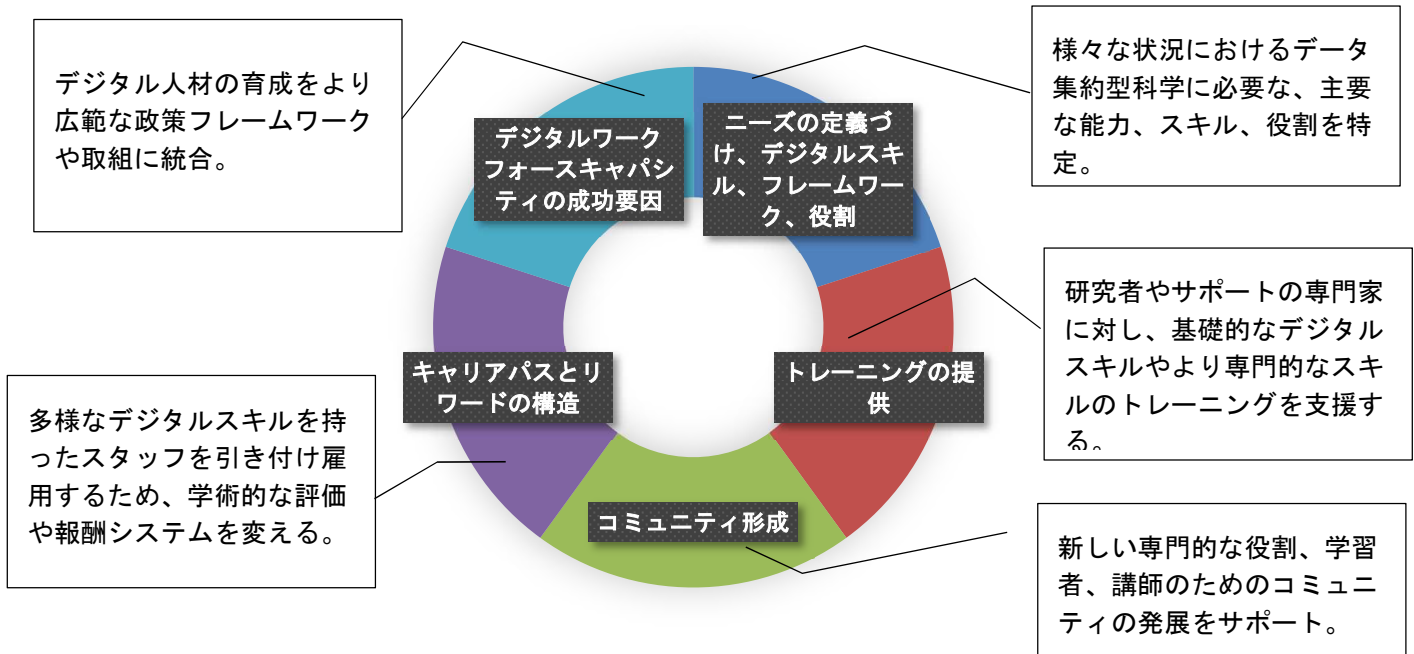
OECDやその他機関は様々な産業部門で必要となるデジタル人材・能力について分析し、データ集約型科学に特に必要な要因を検討しており、本報告書はそれらを参照し、人文学・社会科学を含むすべての科学分野を対象としている。

本報告書では13のケーススタディを掘り下げて分析し、データ集約型研究に必要なデジタル人材・能力の改善に向けて何をしてきたのか、また、今後何が必要なのかを調査した。また、本報告書では、科学において必要とされるデジタル人材が社会経済の他のセクターとはいかに異なるのかを分析しており、科学の特殊性が必要とされるスキルへ与える影響について述べている。これら必要とされるスキルは、科学全体としていえるものと、ある分野や研究領域に特化するものがある。

研究領域固有の専門スキルと基本的なデジタルスキル、といった共通セットを備えたデジタルスキルのある研究者と、データ管理者や研究のソフトウェアエンジニアを含むさまざまな専門的な研究サポートスタッフの両方が必要である。研究はますますチームで行われるようになり、チーム内での能力の分担が非常に重要になってきており、研究者に何が期待され、サポートサービスで何が提供されるべきなのかを規定することは困難であり、かつ、研究領域によって異なる。しかしながら、データ集約型科学は技術的なスキルだけではなく、コミュニケーションやチームワークといった人間中心のスキルであることが認識されつつある。多くの分野においてはまた、特に機密データを取り扱う際、倫理や法律の専門家が必要である。

科学のためのデジタル人材・能力を育成し維持するためには5つの主要領域を同時並行で扱う必要がある（図表ES1参照）。政府、研究ファンド機関、学会、研究機関、大学といった多様な関係者が協力することが必要である。

図表ES1：5つの主要なアクションと目標



政府が行うべき取組として重要なものは以下のとおり。

- 研究にはデジタル人材・能力が必要であり、5つの主要領域を統合するには戦略的な計画が重要であることを政策レベルで認識する。5つの領域とは人材を構築し維持するために必要な以下を指す；ニーズの明示、トレーニングの提供、コミュニティの構築、キャリアパスと報酬、幅広い外的要因。
- ニーズに見合うトレーニング等を提供するため、国のデジタル人材・能力に何が必要で、研究エコシステムがどのような状況にあるか（現在の状況）を分析する。この際、海外あるいは地域での取組を考慮し、それらを政府がいかに効果的に活用できるかを検討する。
- 変化する状況をモニタリング・評価しつつ、データ集約型科学のメリットを最適化するために必要な速度と規模で人材を育成する取組を促進する。

ファンディングエージェンシーが行うべき取組は以下のとおり。

- デジタル人材を育成するためのファンディングメカニズムを確保する。
- 物理的研究インフラの整備とデジタル人材を育成するための戦略と投資を連携させる。



その他にも、データ集約型科学を可能とする環境を与え、デジタルスキルを備えた研究人材の強化に向けた努力を強固にするため、教育、オープンサイエンス、研究公正、評価といった点で政府やファンディングエージェンシーがとりうる政策は幅広い。

多くの国において、大学は高等教育、トレーニング、公的研究の中心的な役割を担っており、持続可能なデジタル研究人材の育成に中心的な役割を担っている。大学は自治権を有する一方、政府やファンディングエージェンシーが与える方針や政策的誘導に対応している。大学はデータ集約型科学に必要なデジタル人材やスキルの強化に必要な以下の取組を前述の5つの主要領域において実施することができる。

- 研究者や研究サポートスタッフへのトレーニングの提供
- 適切な評価、認定、報酬メカニズムを備えた新しいキャリアパスの開発

第7章に詳細を記したが、上記の取組のいくつかは既存の体制に組み込むことが可能である。例えば、大学の図書館はデータ管理スキルの育成を促進するセンターとなることが可能であり、計算機部門はソフトウェアやプログラミングスキルを研究活動全体に広げることにも貢献することが可能である。その他の取組は、さらに系統的、組織的、文化的な変化が必要となる。

コミュニティ育成やトレーニングの提供に関し、学会、学术界、研究機関、研究インフラを含め、他の様々な関係者が重要な役割を担っており、より詳細な提言は、第6章、7章に記載している。また、デジタル研究キャパシティのニーズを公共部門と協力して定義するとともに、トレーニングを提供する者として、民間セクターの役割も大きい。

多くの国や研究機関ではいくつかの提言をすでに実施しており、互いに学ぶことが可能である。この分野で国際的な協力を行い、リソースや経験を共有することをすべての関係者に提言する。

## データ集約型科学に必要なデジタル人材・スキルの構築

### 1. はじめに

デジタルの時代は科学<sup>1</sup>の手法に変化を起し、あらゆる研究分野はますますデータに依存している。同時にデジタル化により可能となったオープンサイエンスに向かう潮流が台頭し、科学の厳格性や誠実さの確保のための新しい課題が現れている。このような変化のスピードは速く、特に、これまでデータが豊富でなかった科学的分野において人材育成に関する課題が生じている。デジタル人材は研究者、研究チーム、データサービスのプロバイダー、研究インフラや研究機関、といった複数のレベルで必要とされる。異なる専門的な役割が生まれており、例えば、「データサイエンティスト」にも様々なタイプがあり、研究を支援する立場の者もいれば、自ら積極的に研究に関与する者もいる。また、研究分野によって、必要とされる専門のタイプやレベルも異なってくる。

科学研究予算の5%まではFAIR原則（見つけられる、アクセスできる、相互運用できる、再利用できる）に則ったデータ管理に充てるべきであり、研究現場の20名のスタッフのうち1名はデジタルのスキルを身に付けて研究をサポートできる専門性を備えるべきとの試算がある（Mons, 2020[1]）。この試算は、実験計画、データの取り込み、データ管理、保存、分析、公表、再利用など研究を支援する様々な種類の専門家が、ヨーロッパだけでも50万人必要であることを意味する。

認識されているデジタル人材の不足に対応するために、さまざまな研究分野や国で多くの取組が実施されている。しかしながら、それらは、バラバラでほとんど調整されていない活動であり、より戦略的なアプローチが有益となりうる。多くがデジタルスキルやフレームワーク、役割に焦点を当てているが、他方、トレーニングの提供、コミュニティ形成、キャリアパスや報酬制度も重要である。データ集約型科学はチームワークが重要であり対人スキルが将来の成功への重要な要因であるということが次第に認識され始めているものの、多くの取組が技術的なスキルを対象としている傾向にある。同様に、オープンサイエンスの新しい幕開けにおいて、これまで研究の外周にあった倫理や法的な課題が中核のものとなりつつあり、それに対応する能力の必要性が増している。

オープンサイエンスの時代に研究者にとって必要なデジタルスキルやトレーニングに関する議論、様々なタイプのスタッフに必要とされるインセンティブシステムやキャリアパスに関する議論が個別に行われている。既に、様々な官民の関係者によって多くのトレーニング施策が提供されており、専門的役割と新しいキャリアパスを定義するための措置が講じられていますが、これらが現在および将来のニーズを十分満たしていない可能性がある。不足箇所や優れた取組を特定し、効果的な政策を策定するためには、科学における人

---

<sup>1</sup> 「科学」には自然科学、社会科学、人文学が含まれている。また、本報告書では「科学」と「研究」は互換性をもって使用している。

間のデジタルスキルをより体系的に検討する必要がある。例えば、アカデミアと産業界が適度に競争・協力することの必要性や、データ倫理の重要性などは、科学分野において幅広く普及する可能性が高い人工知能（AI）の出現によって明らかになった緊急的な課題の一部である。

本報告書は、OECDグローバル・サイエンス・フォーラム（GSF）のデータ集約型科学のためのデジタルスキルに関する専門家グループ（メンバーはAnnex1を参照）によって作成され、科学のためのデジタルスキルを備えた人材の育成強化に関する情報と選択肢を政策立案者に提供することが目的である。まず、本プロジェクトでは、本報告書において使用する重要な用語の定義と共通言語について合意した（報告書巻末の用語集を参照）。本報告書では、まず、デジタルサイエンスエコシステムの簡単な説明とともに、データ集約型サイエンスのためのデジタルスキル人材の重要性、デジタルスキル人材を育成して維持するために必要な取組について述べる。その後、分析の基礎となった13のケーススタディについて詳細に説明する。次の5つに焦点を当て分析を行った。

1. 定義づけの必要性：デジタルスキル、フレームワーク、役割
2. トレーニングの提供
3. コミュニティ形成
4. キャリアパスと報酬の構造
5. デジタル人材の育成・確保に有効な外的要因

この分析を通して、それぞれに対応する5つのゴールを特定した。多数の関係者がこれらのゴールの達成に向けて役割を担っており、政府やファンディング機関、専門的な学会や学术界、研究機関や研究インフラ、大学や図書館それぞれに対する提言を行う。

## 2. データ集約型科学に対するデジタル人材のニーズについて何がわかっているのか？

近年、労働市場の需要とデジタルスキルに関して様々な研究がおこなわれている。しかし、これらは主に一般的な社会ニーズを対象とするものであり、科学分野におけるニーズに特化していない。第2章ではデジタルスキルを備えた研究人材の経済的重要性を含めた幅広い分析を検証し、研究セクターにおける需要と供給に関する分析の詳細について述べる。テクノロジーや研究の急速な進化により、将来のニーズを予測することが困難になる可能性があることを考慮しても、現在人材が不足していることは明らかである。

## 2.1. デジタルスキルを備えた研究人材に対するニーズ

多くの研究において、教育システムの役割と結び付けつつ、労働市場の中でのデジタルスキルが分析されている。多くのセクターの労働市場において、デジタルリテラシーが重要となっており、技術的なデジタルスキルと対人的なスキルの両方が必要であることが指摘されている。例えば、Piatetskyは2018年に求人広告から人材の需要拡大について分析しており、IBMが「2020年までにData Science and Analyticsの求人数が、36万4千件近く増加し、約272万件になると予測」していたことを指摘している（Piatetsky, 2018[2]）。

特に科学に目を向けると、英国がDynamics of data science skills (The Royal Society, 2019[3])の中で、「主要な産業界が最も経験豊富なデータサイエンティストとAI研究者の多くを雇用していることから、メディアの報告では、学術界から産業界への研究者の自然な流れが持続不可能なレベルに達している可能性がある」と示唆していることを報告した。また、求人広告を分析したところ、2013年にデジタルスキルに関連する求人は英国で670万人だったが、2017年には920万人となっていた（36%の増加）。

OECD/GSFでもまた、デジタルスキルを備えた専門家に対するニーズに焦点を当てたオープンサイエンスに関する報告書を近年発表している。そこでは研究データレポジトリ、国際的な研究データネットワーク、科学におけるオープンで包括的な協力が含まれている（OECD, 2017[4]; Dai, Shin and Smith, 2018[5]; OECD, 2017[6]）。これらの調査では、データ収集と管理、データ分析やソフトウェア開発、リーダーシップ、チームワーク、コミュニケーションや問題解決など、科学のあらゆるドメインで専門家が重要であることを認識する結果となった。同様に、ヨーロッパでは、「データサイエンティストを専門家として育成し、データシステムにより多くのユーザーを可能にする幅広いデータリテラシーを養うトレーニングにかつてないほど取り組む」ことが求められている（European Strategy Forum on Research Infrastructures, 2018[7]）。

*Charting the digital transformation of science* (Bello and Galindo-Rueda, 2020[8])では科学研究におけるデジタル化の本質と効果について詳細に分析している。この調査では2018年のOECDによる研究者を対象とした国際的調査（ISSA2）に回答した約1万人の研究者を分析し、科学研究のデジタル化において次の4つ側面を明らかにしている。

- 科学的プロセスのすべての段階でのデジタルツールとデジタル手法の適用
- デジタル化によるデータやプログラムの共有
- 洞察を得て予測を可能とするためのデータ集約型デジタルツールの使用
- デジタルIDの開発及び科学研究におけるオンラインコミュニケーション

これらの側面は、異なるコミュニティ、分野、セクター、さらには国々が研究のデジタル化にいかに対応しているかを比較する際に用いられる。この結果は、デジタル人材に求められる能力が分野（図1）と国（図2）によって大きく異なることを示している。また、この調査分析では、年齢や性別などの要因にも大きく左右されることが明らかになった。

図1：各分野におけるデジタルスキルの要件

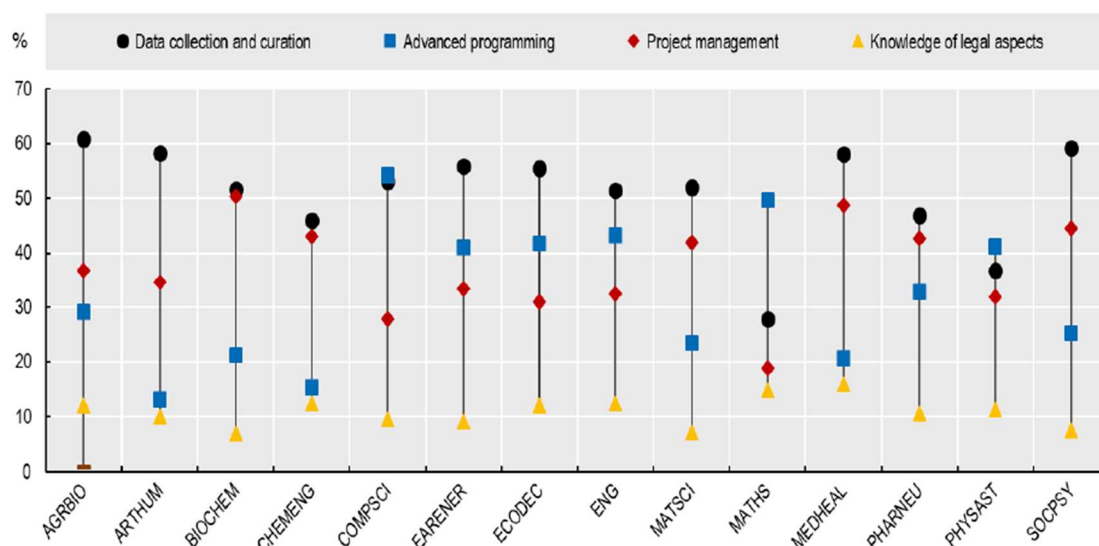


図2：データ集約型研究へ向けた各国における課題



（デジタルスキルのある研究人材が必要となる）研究におけるデータアクセスと利用の拡大から得られる潜在的な経済的利益の推定値は多数ある。近年では、*Cost-benefit analysis for FAIR research data* (European Commission, 2019[9]) において、FAIR原則に準拠しない研究データがヨーロッパ経済に与える

コストは全体として、年間約102億ユーロとしている。オーストラリアの公的資金による2014年の研究では、研究成果を最大化するためのスキルの価値は、少なくとも約19億米ドルであり、最大年間60億米ドルに上る可能性がある」と試算されている（Houghton and Gruen, 2014[10]）。

## 2.2. データ集約型科学におけるデジタルスキルの需要と供給

高等教育部門には、ある種のスキルの需要と供給を分析するためのプロセスが十分に開発されているが、科学研究部門の人々のデジタルスキルに特化した取組はほとんどない。科学部門における需要と供給の統計を正確に予測することは容易でない（Ashley, 2016[11]）が、いくつかの研究はいずれも同じ方向性を示している。例えば、デジタルトレーニングとスキルは需要が供給をはるかに超えている。2016年から2018年にかけて行われた研究は次のことを示唆している。

- 米NSFの生物化学の研究者の分析では、トレーニングが最も不足していることが指摘されている：「これだけニーズがあるということは、NSF、大学、および他の機関がコンピューターのリソースを物理的に提供することでは素晴らしい仕事をしたが、効果的な使用を促すのに必要なものは提供していないことを示唆している」（Barone, Williams and Micklos, 2017[12]）。
- オーストラリアでの生物科学インフラの分析では、データと技術の観点から研究者を分類し、さらに多くの生物科学の研究者がデータ集約型およびバイオインフォマティクス集約型のスキルを必要とするため、今後5年間で大幅な変化が予想されることがわかった（Lonie and Francis, 2017[13]）。
- 日本では研究図書館のデータ管理スキルの分析を行い、41%のスタッフがデータ管理計画に精通していないことが分かった（Kurata, Matsubayashi and Takeda, 2017[14]）。

科学における研究支援の専門家がどれだけデジタルスキルを備えているかのデータはほとんどなく、これらの役割に対する需要に関しても信頼のおけるデータがない。特に、異なる研究ドメインにわたって必要な役割や特殊性を明確に定義付けることがなされていない（詳細は第5章1節を参照）。2019年にアメリカで行われたR1大学（博士課程があり特に高いレベルの研究がなされている大学）のデータ司書に関する分析では、「R1大学の約4分の1にはデータ司書に特化したスタッフがおらず、次の4分の1は1名のみおり、約3分の1は2、3名の少人数のチームが配置されている。残りのごくわずかのR1大学では4~10名の研究データチームがある。R1大学あたりのデータ司書の数平均すると2名をわずかに超える程度である」（Springer, 2019[15]）。近年のオーストラリアの研究IT支援ワークフォースの分析で、研究者をデジタルの側面でサポートする職員のフルタイム当量（EFT）を換算したところ次の結果を得た。

- 60~70名の研究者の研究データ収集と分析に対し、1EFT

- 90名の研究者の研究データ管理とスチュワードシップに対し、1EFT
- 250名の研究者のデータに関するトレーニングやアドバイスに対し、1EFT
- 100名の研究者が関与する様々なソフトウェア工学における役割に対し、1EFT
- 250名の研究者へのソフトウェア工学におけるトレーニングやアドバイスに対し、1EFT

ヨーロッパでは、研究の場においては20名のうち1名のスタッフはデジタルのスキルを備えた研究支援の専門家であるべきとされ、現在の比率が不十分であるという証拠を裏付けている。ヨーロッパだけでも少なくとも50万人の研究支援の専門家が必要であり、科学研究予算の5%まではFAIRデータの管理に充てるべきであることが指摘されている（Mons, 2020[1]）。

科学分野全体と特定の分野両方について、デジタルスキルを有する研究スタッフの需要と供給の情報が必要である。本報告書に記載のあるケーススタディにも含まれているいくつかのトレーニングの取組の主要な推進力となったことから、それら需要と供給に関する情報が重要であることは明らかである。

- カナダの企業であるElement AIは、カナダが2018～19年にかけて大学において500名を超える大学院生にAIのトレーニングを実施（Kiser and Mantha, n. d. [17]）したものの、さらに数千人がトレーニングを必要としていると試算した。現在のレベルを引き上げようとするカナダの目的は、やがてカナダのAI戦略へとつながり、トレーニング不足と国外への人材流出に対処することとなった。現在、カナダ先端研究機構（CIFAR、表1を参照）がこの戦略の実装を担当している。
- 2014年にはSoftware Sustainability Institute（SSI）が英国のラッセルグループに属する15大学から1万5千人の研究者をランダムに選び調査を実施した。それによると、回答者の90%以上は自身の研究にソフトウェアが重要になっていることを認識し、70%がソフトウェアなしでは研究することが不可能であると回答した。これにより、研究に必要なソフトウェアを開発するために専門的なスキルが必要であることが明らかになり、SSIの重要な課題の一つとなった。
- Carpentriesのトレーニングは、2013年のオーストラリアの生物科学コミュニティにおける調査などにおいて特定されたニーズに基づいて構築されている。「調査対象となった研究者の60%以上にとって最もニーズが高いのが追加のトレーニングであるのに対し、追加の計算能力へのアクセスが必要なのはわずか5%である…」これは他の国の生物科学研究者によって共有され、ヨーロッパのELIXIRによって明確に認識されている（Teal et al., 2015[18]）。

結論として、データ集約型科学におけるデジタルスキルの需要と供給の問題について、より詳細な分析を行うことが重要である。この文脈で、*デジタル変革の測定：将来へのロードマップ*（OECD, 2019 [19]）は、タスクとスキルに関する詳細調査を活用し、また、ビジネスコミュニティと協力して、スキル不足の新しい



指標を定義することにより、需要と供給の問題に関する新しい洞察を得ることが可能であることを提案している。

### 3. デジタルスキルを備えた社会と科学のためのデジタルスキル人材

このプロジェクトではデジタルスキル人材に関連する問題が科学に固有のものなのか、という点が重大な問題であった。第3章では研究セクターの特殊性を検証するとともに、その特殊性によって、一般的に説明できるデジタルスキルの枠を超えたスキルが必要とされるのかについて検証する。また、科学研究が多種多様であることを考慮し、研究分野間の違いについても検証する。

公的な研究システムは、他のセクターでの研究とは異なった形で発展してきており、その特殊性を生み出している。ドイツにおけるデジタル人材に関する問題の最近の分析（ドイツ科学情報インフラストラクチャー会議（RfII）、2019 [20]）によれば、「科学労働市場はデジタル経済、あるいは産業における研究開発とは異なる特性を持っている。これらの特性の中には、デジタル変革に対する変化と革新的な対応の促進に役立つものもあれば、それを妨げる傾向があるものもある」。これらの特性については第5章にて検討するが、学术界と産業界の間で人材が移動する際の障壁や、民間セクターに比べて劣っている学术界の給料やキャリアパスが含まれる。

データ集約型研究における整合性や再現性の確保、もしくは、アルゴリズムやデータの倫理的な使用を確保しつつ、オープンサイエンスなど新しいニーズに対応するため、デジタル人材が必要であると科学分野がとらえていることが、次章で述べる13のケーススタディで明らかになった。これらの要件は、科学に固有であると認識されることもあれば、少なくとも他のセクターよりも科学分野では重要であると認識されることもある。しかし、実際には必ずしもそうとは限らない。たとえば、オープンサイエンスの要件は公共部門の研究に固有の場合があるが、再現性と倫理的な実施は、公共部門と民間部門の両方で共通であると認識されている。

このプロジェクトでは、社会全体に対して科学が異なるスキルを必要とするかどうかを分析した。欧州委員会の「市民向けデジタルコンピテンスフレームワーク」（DigComp）<sup>2</sup>では科学における能力の関連性と妥当性に関する評価を行った。DigCompは、研究やデータ集約型のドメインに焦点を当てたものよりも、かなり広くより一般的なフレームワークの一例である。デジタル能力は次の5つの主要な要素から成り立っており、それぞれが研究に応用できる一般的な能力を含んでいる。

---

<sup>2</sup> <https://ec.europa.eu/jrc/en/digcomp> を参照。



- 情報およびデータリテラシー：データの参照、検索、およびフィルタリング。データソースの信頼性と信頼性の評価。データの整理と保存。
- コミュニケーションとコラボレーション：データの共有。参照や属性の把握。協調過程に向けたデジタルツールと技術の使用。レピュテーションの保護。
- デジタルコンテンツの作成：オリジナルに関連した新規コンテンツ・ナレッジの作成。コピーライトとライセンスの理解。プログラミングとソフトウェア開発。
- 安全性：個人データの保護。
- 課題解決：デジタル環境を個別のニーズに合わせてカスタマイズ。ナレッジ創出やプロセス導入に向けたデジタルツールの活用。デジタル能力の不足を特定し、自己改善の機会を確保する。

DigCompのフレームワークは研究に従事する人々の全体的なスキル式を評価するには健全な基盤となっている。ただし、公的資金による研究では、以下の能力が追加で必要な可能性がある（公的資金以外での研究においても必要な場合がある。）。

- 情報およびデータリテラシー：データ評価と分析をするための統計学の理解。再現性の要件に関する理解。
- コミュニケーションとコラボレーション：データや情報、コンテンツの共有、デジタル市民権への参画、協調の改善に向けたオープンサイエンスの原則の順守。参照・引用の手法に関する知見を研究データやソフトウェアの参照・引用に拡大する。自身の所属機関だけではなく、アカデミック研究全般のレピュテーションを保護する。
- デジタルのコンテンツ作成：知識を伝えるためのデータや情報の可視化。
- 安全性：データの切り離し、匿名化、隔離といったツールや技術を理解し、機微データを保護する。

また、研究に必要なスキルの組み合わせにも違いがあり、研究が行われる分野、およびそれらが生成して消費するデータのサイズ、タイプ、および多様性によってある程度影響を受ける（図1）。何十年にもわたり強力な計算機の基盤を築いてきた分野では、関連する高いデジタルスキルを持つ研究者を育成する学部課程を構築した。例えば、高エネルギー物理学、天文学、バイオインフォマティクス、医学情報学が一般的であるが、これらの分野でさえも、追加のトレーニングを受けたり、デジタルスキルを持つ研究支援スタッフを必要とする場合がある。対照的に、一部の分野は計算機を使用する初期段階にあるため、その研究者は基本的なデジタル研究スキルすら不足している場合がある。科学分野によって大規模なデータセットを取り扱うレベルが異なり、その成熟度がデジタルに対する能力不足の規模や性質に影響を与える場合がある。特に、計算機の使用が最近増加した分野では、基本的なデジタルスキルに対する投資を増やす必要がある。すべ

ての研究者の一般的なトレーニング（または基礎スキルのトレーニング）へのアプローチに影響を与えるため、科学分野の違いを理解することが重要である。

また、個々の研究プロジェクトによって生成されるデータの性質と規模、研究の質問に答えるデータセットを生成するために必要なデータ統合の程度により、必要なスキルは異なる。高エネルギー物理学は、通常、配信と分析に関して技術的な課題があるが、他の多くのソースからのデータと統合することは不要である。一方で臨床研究は他の臨床試験のデータ（10～100）と統合し分析する必要があり、また、それぞれの臨床試験においても、様々な臨床の設定におけるデータを統合しなければならない。バイオインフォマティクスは、数千人の研究者が行った数十万回の実験結果を統合しており、天文学もデジタルに移行することで、望遠鏡の画像が写真であった時代には不可能だった大規模な統合研究が可能となった。後者の分野では、統合したデータコレクションの抽出が必要だが、それは高エネルギー物理学ではあまり必要とされないスキルである。このように、スキルに必要な要素が専門的な役割と研究チームの構成にどのように対応するかは非常に多様であり、5章1節にて説明する。

要約すると、科学に必要な独自のデジタルスキルまたは能力（または能力のレベル）があるが、これらは主に社会や産業のすべてのセクターで必要とされる基本的なデジタルスキルセットの延長である。これらのスキルは、データ集約型科学のパフォーマンスだけでなく、オープンサイエンスへのシフトや研究の再現性確保の必要性など、その他の推進力にも関連している。学术界や公共部門の研究に固有のものではなく、ほとんどが民間部門にも等しく関連する。ただし、必要なスキルの組み合わせは、科学の分野やドメインによって異なる。

#### 4. 科学のエコシステムとケーススタディ

デジタル人材の育成に関する科学分野での課題は、科学エコシステムの性質と関連している。国はデジタル研究に必要な能力を評価および育成するためのさまざまな取組を展開しているが、それは通常、その国の当事者や歴史、文化の独自の組み合わせが影響しており、国としてのデジタル戦略やデジタルスキル戦略、インターネットへの接続性などの要素が含まれており、より広範な研究、社会、経済の背景が影響している。独自システムの長所と短所や不足を特定する方法、および最も関連性の高い施策を支援または奨励するための最良の方法を検討するに際し、科学政策立案者は、自らの科学エコシステムの複雑な要素を考慮に入れる必要がある。

第4章ではこのプロジェクトの一環として取り扱った13のケーススタディを紹介する。これらケーススタディは、様々な国やデジタルサイエンスエコシステムにおける取組で構成されている。

表1. ケーススタディー一覧

	構造（権限）	対象	研究分野
英国アラン・チューリング研究所（Turing）	研究機関（国内）	博士課程学生	すべての領域、ただしAIに重点
学術データサイエンスアライアンス（ADSA）	大学ネットワーク（国内）	あらゆる研究者および研究ソフトウェアのエンジニア（RSEs）	すべての領域
オーストラリア国立データサービス（ARDC）	政府のプログラム（国内）	研究者と研究支援の専門家	すべての領域
カナダ先端研究機構（CIFAR）	政府のプログラム（国内）	高校生からキャリアの若い研究者（ECRs）	すべての領域、ただしAIに重点
Carpentries	コミュニティプログラム（国際）	あらゆる研究者	すべての領域
科学技術データ委員会（CODATA）と研究データ同盟による研究データサイエンススクール	コミュニティプログラム（国際）	ECRs	すべての領域
デルフト工科大学（TU Delft）	大学（地方）	あらゆる研究者	すべての領域
ELIXIR	政府間組織（国際）	研究者と研究支援の専門家	バイオインフォマティクス
GESIS-ライプニッツ社会科学研究所	研究機関（国内）	中堅以上の研究者	社会科学の応用
オープンアクセスリポジトリ推進協会（JPCOAR）	研究ネットワーク（国内）	司書	すべての領域
Millennium Institute of Astrophysics（MAS）	研究機関（国内）	ECRとRSEs	天文学の応用
RSE協会	専門家（国際）	RSEs	すべての領域
Software Sustainability Institute（SSI）	政府のプログラム（国内）	あらゆる研究者とRSEs	すべての領域、ただしソフトウェアに重点

13のケーススタディは、国際、国内、領域、あるいは大学や複数の大学イニシアチブによって分類されており、詳細は以下の通り。

## 4.1. 国際的な取組

国際的な取組は、複数の国からの専門知識とサポートを活用することができ、デジタル能力の国家的および制度的戦略を開発および実装するための重要な補完・支援になる場合がある。ケーススタディに含まれる国際的な取組には、政府間およびコミュニティ主導の両方が含まれる。

### I. Carpentries

国際プログラムであるCarpentriesは、研究者と研究支援者に向けて基礎的なデータ分析とプログラミングに関するトレーニングを提供している。もともとは資金を助成する形であったものが、会員登録型に移行した。資料はすべて公開されており、事務局は資料の管理とインストラクターの認証が役割となっている。資料（ライブラリ、ソフトウェア、データ）は一般的だが、学習者の専門分野と背景に合わせてイベントを提供している。

### II. 科学技術データ委員会（CODATA）と研究データ同盟（RDA）による研究データサイエンススクール

世界中のどこからでも受講できる2週間のデータサイエンススキルのトレーニングイベントを開催しているほか、オープンで再利用可能なトレーニング資料を作成する国際的な取組。低・中所得国のニーズに焦点を当て、該当する分野の研究者との議論を通じてコースのトピックを絞る。Carpentriesの資料を用いることもあり、授業はチューターと地元のスタッフにより提供され、受講者は将来のインストラクターになることを奨励されている。コンテンツは低中所得国に固有ではなく、高所得国でのニーズにも合致しており、長期的にはOECD諸国でのイベントをサポートすることにより、他地域へも提供することを目標としている。

### III. ELIXIR

ライフサイエンスのデータに関する国際的な分散型研究インフラであるELIXIRは、5つの技術的なプラットフォームの1つを介して調整されるトレーニングを提供している。データ管理や活用に関するスキルのトレーニングを通じ、研究者はELIXIRのサービスやツールを使用できるようになる。また、ポータルサイトやイベント、ツールキット、および講師を育成するための資料を公開している。

### IV. RSE協会

英国でSSIによって設立された専門家による会員制組織であり、現在は独立。RSEが交流する場を作り、知識を交換し、RSEのキャリアパスを改善し役割を認識するよう協働している。このモデルは国際的なRSEネットワーク開発に関し、現在、他国でも採用されている。

## 4.2. 国内向けの取組

ケーススタディの多くは政府の支援による取組であり、政府が直接行う場合の範囲と影響の可能性を分析した。

### I. オーストラリア国立データサービス (ARDC)

ARDCはオーストラリアの研究をサポートするための国家的取組であり、5つの活動分野のうちの1つに熟練した人材の育成が含まれる。ARDCは研究機関でもあるため、研究インフラの利用を最大化するにはスキルが必要であると認識しており、対面およびオンラインでのトレーニングと実践コミュニティの開発を促進している。アクティビティには、講師養成のプログラムや研究ソフトウェアの専門家を育成するためのデジタルスキルトレーニングなどが含まれている。

### II. カナダ先端研究機構 (CIFAR)

CIFARは汎カナダAI戦略のもと、カナダのAIに関する国家の強みを維持および改善することに焦点を当て、トレーニングを提供している。トレーニングは、優秀な大学院生向けの専門的なテクニックの取得から高校生まで幅広く、ワークショップやカンファレンスも開催される。多くのアクティビティが若い女性を含む、少数派グループに関心を払い、少なくとも1つのアクティビティに途上国からの学生を含むこととしている。戦略ではAI機関や研究者の研究活動をサポートすることにより、カナダでのAI人材の魅力、定着、育成を促進するような豊富なトレーニング環境の構築を支援している。

### III. オープンアクセスリポジトリ推進協会 (JPCOAR)

JPCOAR研究データ作業部会の研究データ管理教材開発プロジェクトは、国立情報学研究所 (NII) と連携して、日本での研究データ管理の推進を目的とし、これらの組織は、オープンサイエンスを推進するための幅広い取組の一環として協力している。この活動は、開始当時は図書館員のためのデータスキルに焦点を当てていたが、その後、その範囲を拡大した。プロジェクトでは、開発した資料に基づいてコースを運営しており、これには、他のソースからのオープンソースのトレーニング教材の再利用や翻訳し適用したものを含む。

### IV. Software Sustainability Institute (SSI)

SSIは英国のファンディング機関によって支援を受けており、研究ソフトウェアへの投資を最大限に活用することを目的とし、その活動の一部にRSEと研究者のためのスキル開発がある。スキル開発は、ワークショップやフェローシッププログラムなど、さまざまな活動を通じて実施される。RSEのキャリアパスと専門家としての認識に特に重点が置かれ、トレーニングでは技術的スキルと人間中心のスキル（チームリーダーシップなど）の両方の需要に対応するものとなっている。

### V. 英国アラン・チューリング研究所 (Turing)

データサイエンスに関する英国の全国的なイニシアチブは、研究を実施および適用し、将来のリーダーを育成し、国民の声となって行動することを目的としている。ほとんどの取組は博士課程学生にフォーカスしており、協力機関や大学と協力している（例：博士課程トレーニングセンター）。学際的な観点に重点を置き、TurningはAIなどの特定の研究課題に合わせて専門的なトレーニングを提供しているが、受講生はすでに基礎的なスキルを習得していることが期待されている。

### 4.3. 分野別の取組

分野別の取組は、国際的、地域的、および／または国内的な観点から分類できる。特定の分野に焦点を当てたケーススタディとして次の3つの取組をピックアップした（上記で述べたELIXIRもこの分類に当てはまる）。

#### I. GESIS-ライプニッツ社会科学研究所

GESISは社会調査データのアーカイブ、調査方法論およびコンピューターサイエンスの研究において、ドイツ国内で長年にわたってその役割を果たしてきた。例えば個人用デバイスからスマートホームなど、研究のデータソースが変化していることから、関連するスキルを対象として、一般公開のオンサイトトレーニングや、計算社会科学に関する長期的なワークショップやサマースクールを実施している。それらのプログラムは最も優れたノウハウを持ち、かつ、大きな違いをもたらさう専門的な研究分野に特化していることが特徴である。同時に、GESISは欧州社会科学データアーカイブコンソーシアム（GESSDA）のメンバーでもあり、社会科学の基礎的なデジタルスキルのトレーニングコースも運営している。

#### II. Millennium Institute of Astrophysics (MAS)

MASは天文学における特別な課題に対応するために必要なスキルに特化した、チリの国家的なイニシアチブである。天文情報学の新たな分野で必要とされるスキルには、大量のデータストリームにおいてリアルタイムにイベントを識別するスキルが含まれており、天文学を超えた幅広い応用の可能性を秘めている。特にECRを対象とし、ワークショップや大学院プログラムを含むさまざまな活動を行っており、独自の法的地位を持ちながらも、いくつかの国立大学との強いつながりのなかで、民間部門とのつながりを強化する取組も行っている。

#### 4.4. 大学（個別あるいは複数大学）の取組

大学はデータ集約型科学の人材育成において、重要な役割を果たしており大学職員はすべての（つまり大学独自の取組だけではなく）ケーススタディにおいて中心的に関与していた。次のとおり、2つのケーススタディを紹介する。

##### I. 学術データサイエンスアライアンス（ADSA）

他のケーススタディよりも新しい取組であるが、少数の米国の大学において過去にファンドされた取組に基づいて設立された。ADSAでは研究と教育全体を通じたデータ集約型のツールと技術の教え方と使い方をサポートしている。様々な方法があり、例えば、ハックウィーク、ミニワークショップ、過去のプログラムの卒業生へのグラント（Moore-Sloan Data Science Environments）などがあり、通常は提携したデータサイエンス機関のキャンパスまたは大学の組織で行われる。

##### II. デルフト工科大学（TU Delft）

デルフト工科大学では大学主導で、独自のデータとスタッフに焦点を当てている。データ管理者に資金を提供し、研究者を”データチャンピオン”として指名することで、ネットワークを通してだけでなく、トレーニングイベントやオンライン学習を通じてスキルを広げる。データ分析というよりも、準備、キュレーション、管理などのタスク、および一般的なポリシーを分野固有のものへ変換することに重きを置いている。したがって、開発されたスキルは、一般的なスキルと分野固有のスキルを混合したものとなる。

専門家会合のメンバーの1, 2名がAnnex 2 Bのインタビューの質問を使い、ビデオ会議によって各ケーススタディの担当者にインタビューを行った。インタビューでは、政策的な示唆を持つ問題、課題、優れた取組に関する質問を行った。また、これらのケーススタディはGESISが2019年10月28～29日にドイツケルンで開催した国際ワークショップでも紹介された（Annex 3）。このワークショップでは活発な議論がなされ、経験と今後の見通しに関する意見が交わされたほか、本報告書にも含まれるポリシーメッセージと提言が発表された。

#### 5. デジタルスキルを備えた研究人材を育成するには何が必要なのか

データ集約型の科学インフラへの投資を補完するため、デジタルスキルを備えた人材が必要である。科学分野のみに貢献するのではなく、社会のニーズにより応える人材を育成するため、社会の多様性を考慮に入れなければならない。この人材は、長期にわたって、育成、維持される必要がある。この目的に到達するには次の5つ、ニーズの定義化、ニーズへの対応、規模拡大、持続可能性の確保、変化の促進、が重要である。5章ではそれぞれについて述べる。



## 5.1. 定義化の必要性：デジタルスキル、フレームワーク、役割

2章2節に記載した通り、多くの国や研究分野では、科学に必要なデジタルスキルと全体的なデジタル能力のトレーニングに対するニーズが満たされていない。ここでは、デジタル能力のフレームワークの使用や、これらが新しい役割とどのように関連しているかなど、データ集約型科学に必要なデジタルスキルを特定する方法について説明する。これらを理解することは、政策立案レベルで特に有用であり、より効果的に目標を定め、適切な政策を設計することが可能となる。

科学に必要なデジタルスキルには、専門的な特殊スキルと基礎的なスキルの両方が含まれる。特殊なデジタルスキルのニーズは分野ごとに異なり、分野や新しい技術や政策のニーズに応じて進化する。同時に、一般的かつ基本的なデジタルスキルは、理想的には学部の科学教育に組み込まれるべきだが、現在はそのようになっている。これらの基本的なスキルには、オープンサイエンスと研究の社会的説明責任に関連する幅広い概念とプロセスを組み込む必要がある。

データ集約型科学に必要とされるものとして一般的にリスト化されているデジタルスキルは、複数のカテゴリに分類されるが、使用される用語とカテゴリは統一されていない。たとえば、FAIRデータ専門家グループが作成した「Turning FAIR into reality: FAIRデータに関する欧州委員会専門家グループの最終報告と行動計画 (European Commission, 2018 [21])」によれば、能力の2つの主要な領域をデータサイエンスとデータスチュワードシップとし、以下のように定義づけている。

- データサイエンススキルはコンピューターサイエンス、ソフトウェア開発、統計、可視化、機械学習によって構成されていると理解されている。また、計算インフラや情報モデル、アルゴリズム、情報統合の知識も含んでいる。
- データスチュワードシップはデータが適切に管理、共有、および研究ライフサイクル全体とその後のストレージで保存されるようにする一連のスキルである。これには、アクティブな調査プロセス中に、データセットの不整合を取り除くためのデータクリーニング、データの整理と構造化、メタデータの追加またはチェック、およびデータ管理の問題の解決が含まれる場合もある。

推察するに、データサイエンティストとデータスチュワード（データ管理者）の両方が必要で、それに応じてキャパシティビルディングの取組に焦点を当てるべきと結論づけられるかもしれないが、これは、スキルと役割の間の複雑な関係を単純化しすぎているように思われる。それらは科学コミュニティのさまざまな場所で異なる役割を果たしている。データサイエンス、およびその派生物であるデータサイエンティストは、コンセンサスを得た定義づけが特に難しい用語であり、それらの定義と使用法は分野によって異なる。これらのあいまいさが原因で、データサイエンティストという用語は、政策または計画の観点であまり使用



されていない。これは包括的な用語であるため、科学人材におけるデジタルスキルの状況を評価するには、その中身をさらに分析する必要がある。これに関しては、図3を参照してほしい。

多くの取組において、デジタル能力をリスト化し、またはフレームワークを作成している（OECD, 2019[22]; European Commission, 2017[23]; FAIR4S, 2019[24]; Research Data Alliance, 2015[25]; Demchenko, Belloum and Wiktorski, 2017[26]; Molloy, Gow and Konstantelos, 2014[27]）。

特に研究に関しては、EDISONプロジェクトは「データサイエンティスト」のスキルと能力の包括的なフレームワークを構築する初期の試みであった（Demchenko, Belloum and Wiktorski, 2017[26]）。より最近では、EDISONプロジェクトの作業がさらに強化され、データスチュワードシップに関連する能力が含まれるようになり、その結果、より広い背景での再利用が見られる。

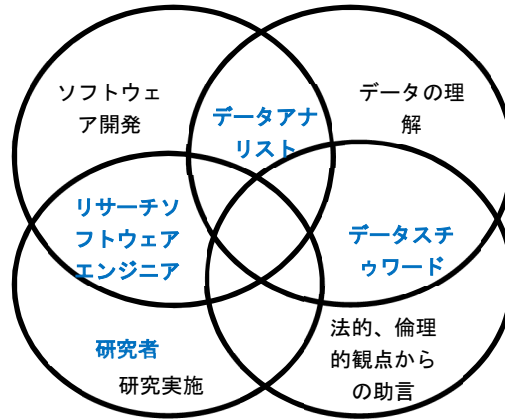
これらのフレームワークは奥行きと柔軟性の点で異なる。FAIR4Sなどのフレームワークでは繊細なアプローチがとられており、特に能力やスキルを様々な成熟度レベルに分類している。FAIR4Sでのレベルとは、初級（認識／理解）、中級（適用できる能力）、エキスパート（このスキルの実践できる能力）である。そして、役割の中にはそこで期待されるスキルと成熟度合いのレベルの両方が含まれる。例えば、ライフサイエンス分野では、データ管理者のタイプを責任や作業ごとに分類した（Scholtens et al., 2019[28]）。

デジタルスキルのフレームワークは、不完全性があるものの、政策立案レベルで特に価値がある。適切なフレームワークは、あるレベルの専門知識を持つ特定のスキルを持つ人々が国や地方で必要とされる人数を特定するのに役立つ。フレームワークの種類によっては、これらのスキルを習得する方法、たとえば、正式な教育、自己学習、メンタリングなどがあり、適切な政策介入をより効果的にすることができる。本プロジェクトの初期段階で、データ集約型研究に必要な主要スキルと役割を簡略化した。この「メタフレームワーク」はAnnex2にあるインタビューの質問事項に含まれており、個々のケーススタディの焦点を明確にするために使用した。このメタフレームワークは、データ主導型研究で必要とされる人材のそれぞれの役割や研究の各プロセスにおいて必要とされるスキルを表している。このフレームワークはケーススタディ全体を比較するのに有用であったが、各ケーススタディのインタビューを進めるうち、個々の研究分野の特異性を反映し、事例ごとにフレームワークを修正する必要性が生じた。このことは、研究のためのデジタルスキルを議論する際、フレームワークを用いることが有用であったことを示している。

スキルを定義することにより、新たな専門職の役割を明確化することが可能となる。図3は、データ集約型科学のために必要な4つの領域を示している（本報告書巻末の用語集を参照）。これらの領域は、データやソフトウェアだけでなく法的・倫理的な要素を含んでいる。また、この図では、必要とされるスキルよりも、役割や責任を表している。したがって、データを担当するデータ管理者はデータ使用の法的および倫理的影響を完全に理解する必要があり、ソフトウェアエンジニアリングを担当するRSEは研究目標を、分析を担当するデータアナリストはソフトウェアを理解する必要がある。研究者の役割は研究の領域と研究チームの規模

と構造によって変化する。大規模なチームでは、研究者とデータアナリストの役割は明確に区別されているが、それ以外の場合では、あまり区別されず同じ役割を担っている場合がある。

図3 役割と責任を表すベン図



出典：Simon Hettrick 氏 (SSI) が本プロジェクトのワークショップを通して独自に作成したもの

専門職の役割を定義付けることは重要であるが、明確に定義することは容易ではない。データ集約型科学のためのデジタルスキルのある人材へのニーズは進化し続けているため、本プロジェクトでは、特定の職務内容を定義することよりも、スキルや能力、チーム内での配置について議論した。実際には、データサイエンティスト、RSE、データアナリスト、データ管理者、データマネージャー、データ司書、デジタルキュレーターなどはさまざまな能力を持っているが、これらの肩書きを持つ人材は、それぞれの専門分野に基づいた異なるスキルセットを備えている。小規模な研究チームではジェネラリストが必要となるであろうが、大規模な場合は専門家の方がより必要だろう。一部の環境では、研究チーム外のグループまたはサービスプロバイダーによって専門スキルが提供される場合があり、例えば、研究機関のデータリポジトリ機能の場合が該当する。したがって、すべての研究グループは、特定のスキルを持つ人々にアクセスする必要があるが、すべてのスキルを保有する必要はない。これは、外部の技術インフラとスキルに依頼して一人で研究を実施する研究者の場合はなおさらである。

必要なスキルと役割を定義することは容易ではないものの、政策立案やトレーニングだけでなく、専門家コミュニティと関連するキャリアパスの確立にとっても重要である。特に学術界は、デジタル人材には合致しない形で組織化されており、必要なスキルと役割を定義することは有用である（5章3節および4節を参照）。データ集約型科学に必要なスキルの評価は、研究ガバナンスにも幅広い影響を与える可能性があり、例えば、倫理と法的問題の重要性は、研究者や研究サポートの専門家だけでなく、治験審査委員会の構成にも影響を与える（Grant and Bouskill, 2019 [29]; OECD, 2016 [30]）。

### Box 1. デジタルスキルのニーズの定義

スキルとトレーニングのニーズを定義するアプローチを以下のケーススタディで例示する。

- アラン・チューリング研究所は英国の大学と協力して、科学へのチームアプローチ、コミュニティ開発の必要性、コラボレーションに焦点を当てた対人スキルの重要性を認識するプログラムを作成。このプログラムは、博士課程学生を研究所のコミュニティに参加させ、他のメンバーと協調する環境で働かせる。このプログラムにより、新しいトレーニングや研究機会を提供するとともに、国内のデータサイエンスやAIスキルの引き上げや将来のデータサイエンスやAIのリーダーを育成することを目指している。
- 科学技術データ委員会（CODATA）と研究データ同盟による研究データサイエンススクールは、低〜中所得国から30〜60名のECRが参加する2週間のコースを毎年増加させている。コースのコンテンツはCarpentriesのものをかなり利用しているが、過去の受講者や現在の受講者と共同して開発を続けている。受講者の一部はコースのチューターになる。また、トップダウンで内容を作るよりも、コミュニティのニーズを強く取り入れるようにしている。
- カナダのAI戦略の一環として、CIFARでは、すべての科学分野における機械学習（ML）とAIを対象としている。この例としてトロント大学ビジネススクールでのMLプログラム<sup>3</sup>などがある。また、AIとMLがある分野でどう適用するかを特定する目的で、研究者と機械学習の専門家間のワークショップやプロジェクトもサポートしている。
- GESISは研究データインフラであり研究実施組織でもあるため、コンテンツを利用する研究者からのニーズを把握し、提供するトレーニングを継続的に改善している。また、スマートデバイスからのデータといった新興領域を対象とする資料の開発に注力するなど、データ環境のより大きな変化に対応している。

## 5.2. トレーニングの提供

デジタル人材の育成には、適切なトレーニングを提供することが不可欠である。第2節では、トレーニングに利用できる様々な教授法や学習法、トレーニングの取組を拡張する際の課題、デジタルスキルを有する人材の多様性促進を目指したトレーニングの必要性、トレーニング提供に関する民間セクターの役割について検討する。

---

<sup>3</sup> 当該プログラム、および類似のプログラムはカナダのベクター研究所（Vector Institute）のAIプログラムのリストに記載されている（Vector Institute, n.d.[94]）。

ケーススタディでは、どの程度需要に見合った利用可能なデジタルスキルのトレーニングを提供できるのか、増え続ける需要を満たすためにどうやって組織を拡張するか、といった課題がよく見受けられた。トレーニングを提供するトレーナーの大半は認定されておらず、トレーニングの専門知識や貢献が認められないことがよくある。いくつかのケーススタディでは、大量のボランティアを動員してトレーニングを提供しているが、これは長期的な持続可能性に懸念が残る。これらの取組は予算が限られており、トレーナーに賃金を支払うことができていない。なお、トレーニングは無料で提供されていることが多い。最近のオーストラリアの世論調査では、さまざまなデジタルスキルに関する79の研究トレーニングイニシアチブに対して、ニーズにどれだけ応えられているかを質問したところ、イニシアチブの75%以上で需要が供給を上回り、うち9%は供給の5倍以上の需要があった<sup>4</sup>。

デジタルスキルの習得とスキルアップには、さまざまな教授法や学習法があり、研究者と研究サポートの専門家の両方がスキルアップすべきスキルの一部は、学部教育に手を入れることによって最終的に達成される可能性がある。ただし、適切な基盤スキルが学部の科学教育に組み込まれたとしても、デジタルツールは急速に進化しているため、大学院のトレーニングが必ず必要となり、修士課程と博士課程の両方に正式に組み込むほか、より即応性の高い臨時的なプログラムが必要となる。伝統的なコースワークのアプローチから、大規模なオープンオンラインコース（MOOC）、実地研修、ハッカソン、サマースクール、会議、ハッキーアワーなどのドロップインセッション（4章のケーススタディを参照）に至るまで、さまざまなタイプのトレーニングがある。デジタルスキルの優れた基盤を提供する正式な認定コースと、テクノロジーとメソッドの進化に伴うスキルアップを保証する非公式のトレーニングが互いに補完し合うことが理想である。デジタルの進化スピードを考えると、デジタルエコノミーに関係する多くのセクターにいわれているように、研究のためのデジタルスキルの習得も生涯学習としてサポートされるべきである。

デジタル研究エコシステムの重要な部分を占めるコミュニティや草の根組織にとって、共通して問題となるのがトレーニングの規模拡大である。主流の組織が対応していない課題や、組織や国を横断する課題に対応するため、例えば、Carpentries、CODATA-RDA研究データサイエンススクール、RSE協会などは設立された。

人材育成では、科学の生産性を向上させるだけでなく、社会ニーズにも対応するため、社会の多様性を反映した人材を育成する必要がある。デジタル世界におけるジェンダー間の不平等については『デジタルジェンダー格差の解消』（OECD、2018 [31]）などの報告書が指摘しており、より多くの女性を取り込むには政策的な介入によって道が開かれる可能性がある。（デジタル）ジェンダー格差を埋める教育とトレーニングの基本的な役割は、ICTの使用、スキル、学習を促進し、教育者に権限を与え、彼らを積極的に巻き込むな

---

<sup>4</sup> アンケート調査は2019年のオーストラリアe研究スキルワークフォースサミットにて行われた（Australian Research Data Commons, n.d.[95]）

ど、体系的なアプローチが必要である。いまや、デジタル（そしてより広義には科学、技術、工学、数学）でのジェンダーの多様性を促す取組はさまざまあり、他の種類の人材の多様性も、データガバナンスのCARE原則（Research Data Alliance International Indigenous Data Sovereignty Interest Group、n. d. [32]）などを通じて対処され始めている。

非営利を含む民間セクターは、この人材の育成を支援する公共部門のパートナーとして、またデジタルスキルのある研究者の雇用主としての重要な役割を担っている。これは、メンター制度やインターンシップなどの実地研修の機会が採用されている人工知能などの「ホット」な分野で特に当てはまる。DataCampやカーンアカデミーなどのデジタル研究スキルトレーニングの商業的な取組もまた、エコシステム全体の重要な部分である。Google、Microsoft、Amazonなどの多国籍企業も、ECRへの財政的支援や就職機会の提供、Google Summer of Codeなどへの学生の参加など、さまざまな方法で関与している。

### Box 2. デジタル人材・能力に関するトレーニング例

- デルフト工科大学は、エジンバラ大学、デジタルキュレーションセンター（DCC）、Research Data Netherlandsと提携して、研究データ管理に関するMOOCを開発している。このコースでは、組織において研究を改善するための効果的なデータ管理サービスの開発と提供について学ぶことができる。
- Carpentriesでは、インストラクター養成プログラムを通じて教育学および包括的な教育実践におけるデジタルスキルを持つ研究者を養成するといった、講師育成モデルを使用している。これにより、トレーナーはCarpentriesのワークショップを効果的に開催し、教えることができるようになる。2019年現在、この訓練を受けたボランティアのインストラクターは世界中で約2,000名以上おり、約600のワークショップを開催している。Carpentriesのワークショップを受講したものは、トレーニングとアプローチを高く評価しており、一部はインストラクターになる場合もある。ELIXIRには、バイオインフォマティクスに焦点を当てた同様の講師育成型のプログラムがある（Morgan et al., 2017[33]）。
- CODATA-RDA研究データサイエンススクールは、低所得国から中所得国のニーズに焦点を当て、パートナー組織や機関がさまざまな場所でスクールを運営しやすくなるよう、国際ネットワークの発展を目指している。年会在イタリア・トリエステで開催されるほか、ブラジル、ルワンダ、エチオピア、アメリカ、オーストラリアの地元のパートナー機関と共同でもイベントが開催されている。
- Carpentriesはまた、デジタルツールへの民主的なアクセスの一環として、多様性にも焦点を当てている。9つの中核的な価値観の中には、すべての人を包括的に、すべての人がアクセス可能に、そして多様性を通じた強靭さ、が含まれている。これらに向けた行動は、『公平、包括、およびアクセシビリティに関するロードマップ（The Carpentries, 2019[34]）に含まれている。
- CIFARIは全体的な目標の1つとして、デジタルスキル人材におけるジェンダーの平等と多様性の促進を掲げている。この目標に向けて、AIの意識を高める女子学生向けのプログラムや若い女性向けのプログラムを支援している。

### 5.3. コミュニティ形成

専門的なコミュニティは、共通の関心を持つ多様な人々が集まり、様々なレベルでの変化を達成するという観点から、デジタル人材の進展のために重要な部分を占めている。コミュニティは、カリキュラムの標準化から、ベストプラクティスと相互サポートの共有、またはキャリアパスに関する政策への影響まで、さまざまな点に関係する。3節では、トレーナー、学習者、リーダーそれぞれのコミュニティを支援する必要性など、コミュニティ形成に関連する問題を検討する。

RSE、データスチュワード、司書など多くの専門家コミュニティが「ボトムアップ」の形で形成され、特定のスキルを持ったグループや役割があるものとしてますます認知度を高めている。政策的な観点から言えば、英国のTechnician Commitmentはコミュニティが科学の外でどのようにサポートできるかを示す良い例である。これは英国研究会議の主導によるセクター全体のイニシアチブであり、研究に携わる技術スタッフが直面する主要な課題への取組を支援するものであり、毎年5%の割合で増え続ける技術者の需要に応えるために創設された（Science Council, n.d. [35]）。英国中の大学や研究機関がTechnician Commitmentの取組を支援しているため、すべての分野で技術者の可視性、認識、キャリア開発、持続可能性が向上している。

トレーナーのコミュニティは、知識の伝達や相互学習、カリキュラムや評価方法の共同開発、トレーニング資料の再活用の促進、といったさまざまな点で価値がある。トレーナーは通常ボランティアであるため、トレーニングの仕事に対する認知や信用という点で課題に直面する。単位認定もまた困難である。各所での講師養成プログラムの一部として、トレーナーのコミュニティ形成や支援を目的とした取組が多くあり、Carpentriesの中核でもある。別の例では、人文・社会科学分野におけるオープンクラウドプロジェクト（SSHOC）に関連する研究者と研究支援の専門家が今まさに講師養成ツールキットを開発中であり、SSHOCトレーニングネットワークを確立しているところである。

ボランティアによる労働力がトレーニングを拡大するにあたっての主要な課題であるものの、十分に認識されていない。ScrogginsとPasquettoはデータサイエンスを支えるこの見えない労働力に焦点を当てた分析を試み、「データ集約型科学の技術的な側面の裏には目が回るほどの労働力が隠されており、その成果の一部はスター研究者が脚光を浴び、残るほとんどは計算機と組み合わせて舞台裏で進められる」と記した。オーサリング、管理、維持、アーカイブ、協力、といった5つの主要な領域で目に見えない労働力を特定し、「データ集約型科学を完全に理解するには、すべての形式からなる科学的な労働を考慮に入れる必要がある」とした（Scroggins and Pasquetto, 2020[36]）。

デジタルサイエンスのリーダーのコミュニティも、特にリーダーシップのスキルを磨くプログラムや、意思決定者としてのフォーマルあるいはインフォーマルなネットワーク形成を通して現れ始めている。ベストプラクティスを共有し、ソリューションに向けた国際的な協力へ主要なインフルエンサーを集めることは重要であり、これらのコミュニティは成長し続ける必要がある。「意思決定者向けのオープンなデータ駆動型



科学」は、生物科学や生物医学分野における意思決定者をターゲットにしたイニシアチブであり、彼らは、オープンサイエンスやデータ集約型科学の新しい発展について広く理解することを必要としている。トレーニングではファンディングエージェンシーやプログラムマネージャー、グラントの評価者、研究者、教育者など、データ統合やソフトウェア開発、大規模計算機の利活用を含む研究申請書やトレーニングカリキュラムを評価する者も注目している。

### Box 3. デジタル人材に関するコミュニティ形成例

- ADSAIはアカデミックなデータサイエンスのリーダーシップを支援しており、データサイエンスを大学の研究と教育の全領域に統合するために必要な、制度的な変化を共有し、促進している。また、それに向けた共通課題に対するアイデアやソリューションをより良く共有するために、コミュニティ形成やネットワークの組織づくりを行っている。プログラムでは白書やデータサイエンスコミュニティサミット（あるいは会議）を通して学んだことの共有や、大学キャンパスのデータサイエンス施設での年次総会、あるいは大学全体にデータサイエンスの実務を広げるためのMoore-Sloan Data Science Environmentsプロジェクト同窓会の支援、といった内容が含まれている。
- SSIのフェローシッププログラムは、自身の研究領域で研究ソフトウェアをより良く活用することを希望する研究者にファンドを提供するものである。ファンドが与えられる期間は「フェロー」を名乗ることができ、その他のすべてのフェローとつながり、シニアになってもコミュニティを形成することができる。SSIフェローの一部はいまやリーダーの位置についている。
- RSE協会の取組はコミュニティ形成の成功例として、また、新たな役割を担う専門家の例として紹介できる。この新しい専門的な社会はヨーロッパ、アメリカ、オーストラリアに見られ、RSEが学界での役割を認められ、適切な恩恵を受け、キャリアの機会をもつ必要性が求められている。
- CODATA-RDAは孤立した環境で作業を行う個人が受講していることが多く、このスクールで初めて同様のスキルを持ち研鑽を積む受講生と作業する機会を持つこととなる。それゆえ、受講生が独自のコミュニティネットワークを作り、さらに成長できるようにしている。スクールが計画したものはなかったが、このコミュニティの価値は受講生や主催者によって認識され支持されている。

コミュニティ主導のイニシアチブは相互サポートや学習を提供し、個人がデジタルスキル人材の成長の中で果たしうる基本的な役割を強化するのに役立つ。様々なステークホルダーによるオープンサイエンスの実施に向けた、実効的な取組を奨励するために、欧州委員会が設立した「オープンサイエンス政策プラットフォーム」もまた、個人の行動を強調している。このプラットフォームは機関や個人がオープンサイエンスを実現するには何をすべきか、自身の権限の中でコミットするにはどうすべきか、を検討することを促している（Méndez, 2019[37]）。

## 5.4. キャリアパスと報酬の構造

データ集約型科学に必要なデジタルスキル人材を構築し維持するには、キャリアと報酬構造に注意を払う必要がある。多様な学歴を持つデジタルスキルを備えた研究者に新たな役割が生まれており、これをサポートするには適切な人的資源構造と方針が必要である。

デジタル技術を備えた研究者だけでなく、特にデータスチュワードやRSE（5章1節を参照）など研究サポートの専門家に対する需要がある。さらに、司書、アーキビスト、キュレーターといった歴史的にも科学的な情報管理の分野で非常に重要な役割を果たしてきた研究サポートの専門家は、デジタル資産の調整と管理において新しい役割を担い、新しいスキルを身に着ける際に「デジタル」という接頭辞を用いている。個々の研究者はこれらのサポートスタッフに求められているデジタルスキルの一部を備えているかもしれないが、程度の違いはあるにせよ、これらのサポート的な役目を研究チームの一部としてみなしている場合がある。

新たに現れたデジタルスキルを有する専門家の経歴は様々である。例えば、一部のRSEは自身の研究を発展させるためにソフトウェアの開発に時間を費やしているような研究者としてキャリアをスタートさせている。彼らはこの作業を楽しんでおり、専門的なスキル開発に時間や労力を投じてきたため、継続的にソフトウェアや研究での使用法に注力している。その他にも、より従来型のソフトウェア開発のバックグラウンドから始まったが、研究を進めるうちにソフトウェアを使った研究に引き込まれる人もいる（Software Sustainability Institute, n.d. [38]）。それぞれのスキルの価値を認識し、成長する機会を与えることが重要である。ただし、これらの新しい専門的な研究支援の職種に必要な長期的なキャリアパスは、非常にゆっくりと出現しているにすぎない。

研究者と研究サポートの専門家両方に対して、デジタルスキルの習得と適用を促し見返りを与えるような、インセンティブを与えるメカニズムが必要である。これには多くの報告書が出されており、学界におけるデジタルスキルを持った専門家へのインセンティブやキャリアパスの欠如が強調されている（Berente et al., 2018[39]; Working Towards Sustainable Software for Science: Practice and Experience, n.d. [40]）。

本報告書で前述したように、特に人工知能などの「ホット」な分野で、デジタルスキルのある人材をめぐって、学術セクターとその他セクターで獲得競争がある。トレーニングの提供に関しては、異なるセクターである程度協力しており、ケーススタディの一部で示しているように、人材交流のスキームや学術界と産業界での併任といった好例がある。学界は給与の面で競争に勝つことは難しいかもしれないが、学界の研究環境はデジタルスキルのある人材にとって、自身のキャリアの様々な段階において魅力的である。しかし、現在のところ、産業界と学界の間、あるいは専門職としてのキャリアパスとアカデミックでのキャリアパスの間を移動することは困難である。繰り返しになるが、商業的な研究環境でいくら成功しても、科学



雑誌への掲載がない限り、学术界でキャリアを得ることは非常に困難であるという、「学術的な履歴書」への期待が主な障壁の一つとなっている。デジタルスキルを念頭に置いた際、この「学術的な履歴書」はあらためて検討の余地がある。

これらの問題に対応するには、従来の伝統的な学術的な昇進の基準を相応に適用させる必要がある。採用、昇進、テニュア期間の見直しにおいて、成果発表と同様にデータやソフトウェアも科学のアウトプットや資産として認識される必要がある。これは何ら新しい問題ではなく、「研究評価に関する宣言」(DORA, 2012[41])においても研究評価にはすべての研究成果のインパクトの価値だけでなく、データセットやソフトウェアを含めて検討することが求められており、一部の国では実際に取り入れ始めている。しかし、学术界で根強く支持されているのは成果発表のアウトプットをもとに、個別の学術研究の価値を測りリワードを与えることであり、データ、ソフトウェア、アルゴリズム、コードといったデジタルのアウトプットの価値を認識するのはまだまだ限られている。

#### Box 4. キャリアパスと報酬構造の改善例

- TeSSIはELIXIRコミュニティの登録システムであり、トレーニングの提供者がイベントや資料を登録し、著者のクレジットと所有権の帰属を円滑に進めるようにしたものである。2019年4月のトレーニング・プラットフォームには1,208の資料、289のイベント開催情報、9,200を超える過去のイベントへのリンクが張られている(Europe's distributed infrastructure for life-science data, n. d. [42])。
- デルフト工科大学は大学内のデータ管理者の育成に特に重点を置いている。2016年以降、すべての学部は研究データ管理が必要な研究にデータ管理者を配置した。これは、2017~2018年に行われた約700名のスタッフを対象とした調査において、研究のデータ管理を改善することが最もインパクトが大きいことが明らかになったことを受けて行われた。データ管理者は一般的なトレーニングの中核を提供することから始め、次に個別分野のトレーニングが続く。各学部にデータ管理者のポストを確立したことにより、このようなスキルを持つ専門家が研究に貢献しているという認識を強め、彼らのキャリアパスを構築することにもつながった。
- CIFARが重点を置くのは学术界とユーザーの間の緊密な協力関係である。AI Research Chairプログラムに選出された80名もの研究者は、産業界とのクロスアポイントメントとなっている。カナダAI戦略によって支援を受けるAI機関は学术界、産業界、イノベーターの間の相互作用、アイデアの交換、協力を促進することが期待されている。
- MASは大学に付属しているものの、法的に独立した権限が与えられており、柔軟に人材を管理することができる。MASは大学と比べて給与に柔軟性があるため、産業界との競争に直面しつつもデジタルスキルを持つ人材を採用・確保することができている。

## 5.5. デジタル人材に関する外的要因

本報告書は科学活動に焦点を当てているが、より幅広く研究や社会的、経済的な状況下に置かれている。デジタルスキルのある研究人材の育成は、オープンサイエンスの支援や研究公正に関するガイドラインといった、幅広い政策と関連している。同時に、これらの幅広い政策の目標を実現するには、強力なデジタル人材が必要である。5節では幅広い政策や、デジタル人材の育成にプラスの影響を与える外的要因について検討する。

科学研究が影響を受ける幅広い社会経済的な環境を理解することは重要であり、これら社会経済的な環境は、デジタル研究人材の育成に影響を及ぼす以下の複数の要素に分類することができる。

- 経済的環境（オープンサイエンスやオープンデータファンディング）
- 民主的な環境（デジタルリテラシー、労働市場）
- 技術的環境（デジタルコネクティビティ）
- 法的かつ政治的な環境（オープンサイエンスポリシー、データベースの法律）
- グローバル・国際的環境（国際関係は国がどの程度デジタル人材育成の取組に協力するかに影響する）

教育システムのすべてのレベルにデジタルスキルのトレーニングを取り込められるかが、主要な成功要因なのであり、これは適切なスキルを持つ学生が研究キャリアに進むためのパイプラインを作りうる。この文脈において、ハンガリーのデジタル教育戦略を引用すると、ここでは教育システムの全レベルでセクターごとの戦略や専門的な目標と調和させつつデジタルリテラシーを高めることを目的としている（EC/OECD, 2020[43]）。同様に、ポルトガルの省庁間合同のデジタルスキルに関する国家イニシアチブe. 2030は、デジタルスキルを未来志向の社会への道を切り開くツールとして発展させることを目標としている（EC/OECD, 2020[43]）。

科学政策における国内的及び国際的な幅広い傾向は、デジタルスキルの習得の根拠となっており、13のケーススタディの成功要因となっている。また、透明性と開放性を確保し再現性を担保することは、重要な推進力となっている。例えば、公開での精査に耐える再現可能な結果を生成することは、デジタルスキルを持つ人材やオープンデータへのニーズを生み出している。オープンな研究手法が支援され、奨励され、報酬を受ける制度が拡大しており、これらは全てデジタルスキルを有する人材を必要としている。

幅広い研究エコシステムのその他の部分でもデジタル人材の展開にプラスの影響を与えうる。いくつかのケーススタディではデータやコードにアクセスできるジャーナルの出版社といった機関が研究コミュニティ

の行動に主要な影響を与えており、関連するデジタルスキルの需要を増やしているとしている。同様に、オープンソースソフトウェアの普及が、デジタルスキルの必要性の認識を高めている。

#### Box 5. 成功要因とデジタル研究人材育成

キャパシティビルディングを直接目的とはしていないものの、デジタル研究人材の育成に影響を与えている取組は以下の通り。

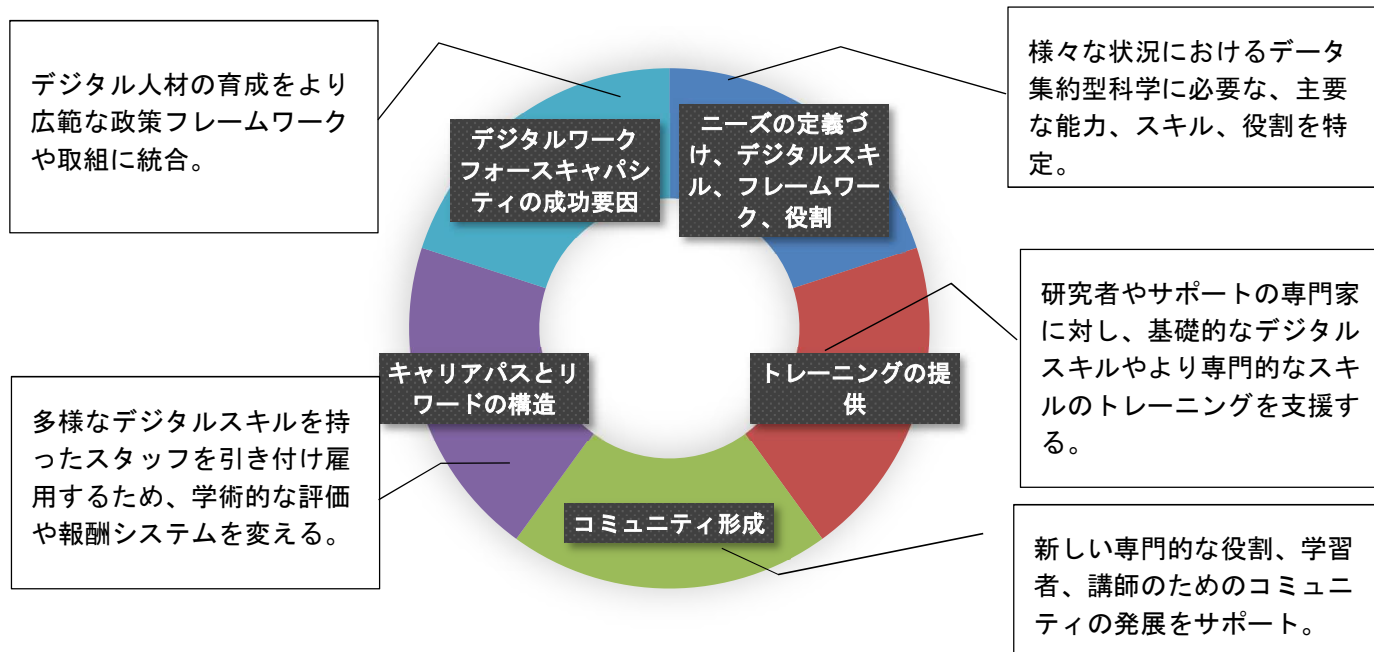
- 国立研究開発法人日本医療研究開発機構（AMED）はファンドの申請書にデータ管理計画に関する情報を記載することを要件とし、誰がその実施に責任を持つのかも含めることとしている。また、データ管理のトレーニング、さらにファンドを受けたスタッフや研究者の再活用に関するファンドも提供している。これは図書館員のための研究データ管理のトレーニングツールの継続的な開発におけるJPCOARの作業を補完しており、このトレーニングは図書館員がデータ管理計画に必要なスキルを身に着けるのに役立つ。
- デルト工科大学はファンドの必要条件にデータやコードの共有やデータ管理計画を加えることは、研究者に新しいデータスキルを身に着けさせるためにプラスの影響を与えているとしている。オランダ研究評議会（NWO）と欧州委員会もこのような要件をつけ、コンプライアンスを強化している。ジャーナルがデータアクセスを要件にしているのもまた同じ効果がある。
- ドイツ科学情報インフラ会議の報告書に基づき、GESISは職業的及び科学的な研究レベルの双方で必要な教育の将来のニーズを計画した。過去にはこの分野で職業的なトレーニングプログラムはなかったが、今や職業コースから生まれた最初の卒業生は、データのライフサイクルを通じて記録及び文書化するスキルを身に着け、社会科学の調査作業に携わっている。また、デジタルスキルの科学教育に加え、GESISが従来型の調査データに補完する新しいデータソースに関するサービスを拡大するというビジョンの達成に貢献している。
- ARDCの取組はオーストラリア政府によるイニシアチブにより強化されている。イニシアチブの一つは共同研究インフラ戦略であり、これによりファンドを得たプロジェクトはFAIRデータ原則の尊重を要件としている。ARDCによって国全体の研究カルチャーがFAIR原則を優先するようシフトし、オンライントレーニングガイドの大規模なレポジトリの提供につながった。

#### 5.6. 目標に向けた各アクション範囲のリンクづけ

上記に記載した、5つの主要なアクションに基づき、デジタル人材を育成するために必要な、それぞれに対応する5つの目標を明らかにすることができる。生産的な科学人材を育成し、確保するには、多数のステークホルダーの間で責任を共有し、様々な関係者が目標に向かって協力することが必要である。

5つの目標を、それぞれに関連する5つの主要なアクションと共に並べて記載したのが図4である。

図4. デジタル研究ワークフォースのキャパシティの発展に向けた5つの主要なアクションと目標



第6章、第7章ではそれぞれのアクターがこれらの目標に向けていかに貢献しうるかを検討する。

## 6. それぞれのアクターに対する提言

上述のようにアクション範囲と目標を定義づけたうえで、科学のエコシステムの議論に立ち返ると（4章参照）、誰が主要な関係者で、これらの目標に向かってとるべき行動は何か、を明らかにすることができる。政策的な観点から5つの主要な関係者は次のとおりである。①国及び地方政府、②ファンディングエージェンシー、③専門的な科学協会、④研究機関と研究インフラ、⑤大学。6章では、最初の4グループへの提言をすでに行われている取組例と共に示す。また、最後に国際協力について可能な取組を簡潔に述べる。大学の役割については7章に続ける。

概要は表2のとおりであり、5つの主要なアクション範囲のどこにどの関係者が一番影響を及ぼしうるかを示している。

表2. 5つの主要なアクションにどのアクターが効果をもたらすのか

	ニーズの定義づけ	トレーニングの提供	コミュニティ形成	キャリアパスとリワード	幅広い成功要因
国及び地方政府	✓	✓	✓	✓	✓
ファンディングエージェンシー	✓	✓	✓	✓	✓
専門的な科学協会	✓	✓	✓	✓	✓
研究機関と研究インフラ	✓	✓	✓	✓	✓
大学	✓	✓	✓	✓	✓

注：

- ・太字のチェックマークはアクターが重大な影響を及ぼしうることを示し、それ以外は小さな影響しか及ぼさないことを示す。
- ・この表は各アクターがどの領域に最も重要な影響を及ぼしうるかを一般化したものであり、実際にはそれぞれの個別の機関によって異なる場合がある。

当然のことながら、大学はほとんどの国で教育、トレーニング、公的研究の中心であることから、表2は大学の役割に集中していることを示している。ただし、政策的な観点からは、トップダウンの国家政策によって命令されるのではなく、自らの意思で行う自律的な組織として広く認識されている。同時に、これらの機関にとってデジタルスキルのあるアカデミックな人材を確保することは重要であり、他のセクターと適切なデジタルのキャパシティビルディングを設計及び実施するために協力し合い、データ集約型科学に向けて環境を整えることは、自らの利益のためでもある。大学がとるべき取組の詳細は次章で述べる。

組織やコミュニティに対する一般的な提言は、それらのデジタル人材戦略の成熟度を評価し改善することである。成熟度のモデルは通常、機関が特定の領域における有効性を評価し、パフォーマンスを改善するには何が必要かを理解するために使われ、戦略的なリーダーシップと紐づけられるときに最も効果的となる。図5では中央政府が国の戦略を評価する際、あるいはあらゆるタイプの機関やコミュニティが内部戦略を改善する際、といった様々なレベルで使用されうるデジタル人材の成熟度モデルを示す。

図5. デジタル人材の成熟度モデル

レベル3 浸透	協調的なアプローチ 実行の浸透 文化的な需要 評価の実施
レベル2 発展段階	デジタルスキルとフレームワーク トレーニングのニーズの査定 コミュニティの発展 キャリアパスの展開 幅広い成功要因
レベル1 基礎	単独のアプローチ コンプライアンスを確保するためのイニシアチブ
レベル0 何もなし	スキル、トレーニング、コミュニティ、キャリアの問題を認識

### 6.1. 国または地方政府

デジタルスキルを備えた研究人材の発展を加速させるために政府がとるべき手段は数多くある。何よりもまず、担当当局がこの問題の緊急性と、行動を行わないことによる研究競争力への悪影響の可能性を理解することが重要である。政府に対する提言は次のとおりである。

- 政策レベルで研究におけるデジタルスキルを備えた人材が必要であることを理解し、5つの主要な領域を取りまとめた戦略的な計画の重要性を認識する。これらの領域はこの人材を育成し維持するために並行して対応されるべきものである。つまり、ニーズの定義づけ、トレーニングの提供、コミュニティ形成、キャリアパスとリワード、デジタル人材育成の外部要因、の5つである。
- 国内のデジタル人材のニーズと、研究エコシステムの状況を分析し、戦略的な計画や投資に反映させ、これらのデジタル人材の不足を解消する。国際的及び分野的な取組と自国の現状を勘案し、それら既存の取組を活用・支援する有効な方策を検討する。
- データ集約型科学の恩恵を最適化するのに必要なスピードと規模感で、人材の育成に必要な取組を円滑に調整する。その際、日々進歩するデータ集約型科学の状況を監視・分析することを継続する。

また、データ集約型科学のための環境を整備するため、以下のような幅広い行動をとることも可能である。

- オープンサイエンスに向けた政策的なサポートや、研究の再現性と研究公正を確保するための取組。
- データとソフトウェアの出力が出版物と同等の価値を報いるような研究評価システムの導入。



- トレーニングの提供、コミュニティ形成、キャリアパスや報酬の構造など、常に変わりゆくデジタルスキルを備えた人材のニーズに対応するために、積極的かつ柔軟な取組を奨励。
- 全ての教育レベルでデジタルスキルを高め、どの省庁がどの教育段階でどのようなトレーニングを担当するのかを明示。

#### Box 6. 政府は何をすべきか？

- フィンランドの教育文化省は2017年にデータ管理と科学的なコンピューターに関するプログラムを立ち上げたが、これは国家イニシアチブである「オープンサイエンスと研究」の主要な部分を占めている。このプログラムは大学で研究や教育をサポートするために、データ管理と研究用の計算機インフラ、サービスの展開と専門家の育成を目指している。
- ドイツの科学情報基盤会議はデジタルスキルを備えた人材に焦点を当てている。今後数年にわたり、研究データにターゲットを絞ったイニシアチブを通してサポートする主な領域は、政策と規制、デジタルスキルおよび基盤的施設の連携といったものが含まれる（German Council for Scientific Information Infrastructures, 2019[45]）。
- 日本の文部科学省によるデータ関連人材育成プログラム（D-DRIVE）は、大学と企業が協力してトレーニングプログラムの開発と実施を支援するものであり、博士課程学生と博士号取得者がデータサイエンスに関するスキルを取得することを目的としている。また、企業データを活用した問題解決型の学習とインターンシップを通じて、実践的なスキルの取得と発展を促している（Ministry of Education, Culture, Sports, Science and Technology, n.d.[46]）。
- 南アフリカの科学イノベーション省は、全国の複数機関による「eサイエンス大学院教育及びトレーニング・プラットフォーム」を支援している。このプラットフォームは、eサイエンスに関連し急速に進化する学際分野における、大学院生のトレーニングを促進するために、適切な資格、カリキュラム、教育的な介入を進めている（National e-Science Postgraduate Teaching and Training Platform, n.d.[47]）。
- 欧州委員会はヨーロッパ・オープンサイエンス・クラウド（EOSC）に関連するデジタルスキルを扱う国を超えた多くの取組（政策、プログラム、プロジェクト）を支援している。EOSCは様々なレベルの異なるスキル開発とトレーニングを、いかにEOSCに組み込むかを特定すべく、スキルとトレーニングに関するワーキンググループを立ち上げた（EOSCsecretariat.eu, n.d.[48]）。

## 6.2. 研究ファンディングエージェンシー

国によって、政府とファンディングエージェンシーの間で戦略的計画とファンディングに関係する責任の分担が異なる。戦略的な計画については、上記のような中央政府と地方政府によるアクションは、ファンデ

イングエージェンシーともたびたび共有される。重要なことは、これらのアクションが効果的に実施されているかである。同様に、研究ファンディングの場合、メカニズムは国によって異なるが、関連する機関は次のアクションを対象としている。

- デジタル人材の育成を支援するファンディングスキームを確保する。これには、例えば、研究者や研究をサポートする専門家向けのトレーニング構築や開発、学術界と産業界の間での交流プログラム、スタッフを採用する際の給与や採用基準の柔軟性、といったものが含まれる。
- デジタル集約型科学を支援するには物理的なリソースと人的なリソースの両方が必要であることを認識し、物理的な研究インフラとデジタル人材に関する計画、政策、投資を連携させる。

ファンディングエージェンシーはまた、政府や他の関係者とも責任を共有し、データ集約型科学を可能とするより広範な環境を生み出さねばならない。政府及びその他関係者は研究がどう行われるかに関する権限とインセンティブに関して大きな役割を果たしているおり、次のような取組が有効である。

- データとソフトウェアの管理計画の策定をファンディングする際の要件にし、計画が順守されていることをチェックするプロセスを構築する。
- データとソフトウェアが成果として適切に評価され、ピアレビューと研究評価過程で考慮されるべきであることを認識する。
- 研究の責任ある行動、研究公正や倫理に関連するガイドラインやプロセスに、デジタル人材の要素を含める。

#### Box 7. 研究ファンディングエージェンシーは何を行うべきか。

能力育成やトレーニングに対するファンディングに加え、ファンディング機関はデジタル人材の育成を促進するための取組を数多く実施している。

- トルコの科学技術研究評議会TUBITAKは、オープンサイエンスポリシーの一部として、研究データトレーニングポータルを開始した（TUBITAK, n. d. [49]）。このポータルサイトはサンプルデータの管理計画を共有するリソースを提供し、この共有を通じて研究者がより効率的かつ生産的な方法で科学研究を実施できるように支援するものである（TUBITAK, n. d. [50]）。
- オランダのNWOでは研究者のリワードシステムを変更し、Veniスキームというキャリアの若い研究者向けのメジャーなファンディングで、文章形式のCV様式を導入した。この新様式は2つのカテゴリ（学術的なプロフィールと主要成果）からなっており、データやソフトウェアを含む多様なタイプの成果を記すことができる（DORA, 2019[51]）。
- 英国芸術・人文科学研究会議（AHRC）は、博士課程学生にデジタルスキルのトレーニングを行うことを求めており、Vitae（高等教育機関、研究機関の博士研究者や研究スタッフ、博士課程に在籍す



る大学院生の自己啓発、専門的能力開発及びキャリア開発を支援する英国の非営利の全国的ネットワーク組織)のフレームワークに基づき、学生は数値、データ管理、統計的手法やソフトウェア、ウェブやソーシャルメディアのコミュニケーションツールなど、既存、かつ、新しい方法論の知識と理解を習得することが奨励されている(Arts and Humanities Research Council, 2014[52])。

- いくつかの国では、個々の研究や機関に関連する科学、技術、イノベーションといった情報を管理するシステムを開発している。例えば、Science Experts Network Curriculum Vitae (SciENCvUSA)、Lattes (ブラジル)、National Academic Research and Collaboration Information System (NARCIS、オランダ)およびResearchfish (英国)など。これらのシステムはキュレーションされたデータやソフトウェア、ワークフローなどを幅広く研究成果として認めている。これにより、データやソフトウェア、それらに関連するスキルの重要性が認識され、研究者や所属機関の行動を変えるような推進力につながっている。

### 6.3. 専門的な学会やアカデミア

アカデミアや専門的な学会は次のような取組を通じて、彼らのコミュニティの中でデジタル人材の育成を重視する文化を醸成する役割を担っている。

- デジタルスキルキャパシティに関する、自身のコミュニティに特化したニーズを定義付け、そのニーズを扱うリソースやメカニズムを提唱する。
- デジタルスキル、フレームワークや役割、トレーニングの提供、コミュニティ形成、キャリアパスと報酬システムに関する会議やイベントで議論やネットワーキングを促進する。
- デジタルトレーニングの資料を開発し広く普及させ、カリキュラムの開発に貢献する。
- データ集約型科学に関し、コミュニティを代表して人材を変革させる計画の策定に貢献する。
- 研究コミュニティや科学政策のレベルに影響を及ぼすような、データ集約型科学とスキルに関する報告書を作成し、各国におけるグッドプラクティスを共有する。

#### Box 8. 専門的な学会は何をすべきか。

- アメリカ天文学会はCarpentriesの取組を支援し、天文学のためのデータCarpentryワークショップのカリキュラム開発を進めている。このワークショップはアメリカ天文学会年次会合にて開催され、カリキュラムの開発と提供の両方で学会と強くつながっている。
- オーストラリア人文学アカデミーは2018年から3年間の未来人文学人材プロジェクトを開始した。これは、オーストラリアの人文学研究人材を包括的に説明し、将来どのような知識やスキルが必要に

なるかの計画できるようにしている。また、データやデジタルリテラシーに重点を置き、将来の研究環境に向けてスキルと知識の優先順位を見出すことを目的としている（Australian Academy of the Humanities, n.d. [53]）。

- 英国の医療科学アカデミーの報告書は健康分野においてAIを利用するための、5つの主要なテーマを特定し、それにはデータとコンピューティング技術、セクターや分野を超えた仕事、トレーニングと能力向上が含まれている。また、報告書では、データ人材が研究から離れる主な理由として、学术界で明確なキャリアパスがなく、給料が低いことが挙げ、これが学問的及び知的な自由を犠牲にしている、と指摘している（The Academy of Medical Sciences, 2019[54]）。

#### 6.4. 研究機関と基盤

研究機関や研究基盤はデータ集約型研究に焦点を置き、大容量データの生成および分析を行っている。ARDC、ELIXIR、GESIS、SSIといった本報告書でケーススタディの対象となった数々の機関が、研究基盤として位置づけられ、幅広いコミュニティにサービスを提供し、MASとTuringは自律的な研究機関あるいは研究拠点に位置付けられる。これらの機関や大学の間には、研究を生み出し、研究者や研究サポートの専門家を雇用するという役割に類似点があり、その結果、研究機関と基盤は次章で詳述する大学や司書に向けた提言の多くを策定するにあたって重要な役割を担っている。研究サービスのプロバイダーであり専門的な研究拠点として、これらの機関はトレーニングで特に重要な役割を果たす。効果的に役割を実行するには、人材育成が全体的なミッションの一部を占めていることが理想であり、大学や国内及び国際的なネットワークとの連携を強め、必要なリソースを備える必要がある。

##### Box 9. 研究基盤ができることは何か

- 芸術・人文学デジタル研究基盤（DARIAH-EU, 2019[55]）はヨーロッパの18の加盟国が共同で支援するコンソーシアムとしての研究インフラであり、データ管理の実践に関する研究者の指導とトレーニングに重きを置いている。DARIAHのミッションは芸術や人文学における研究コミュニティの強化であり、それに向けて文化や社会に関する知識の創成、接続、共有するためのデジタルの手法を使っており、DARIAH ERIC Sustainability Refined (DESIR) ウィンタースクールなどのトレーニングアプローチを行っている。ここでは芸術や人文学のコミュニティで、研究データ管理、キュレーション、共有、保存、再利用のスキルを磨くことを目的としている。
- カナダの研究、産業及び教育の発展のためのネットワークに関する研究ソフトウェアプログラム（CANARIE, 2020[56]）は、2018年以来、研究者と直接作業を行うためにカナダの研究機関にあるソフトウェアチームに資金を提供する研究インフラである。CANARIEはデジタルインフラを計画し実現する非営利法人だが、その資金のほとんどはカナダ政府が提供している。2020年のファンディング

では、分野に関係なく研究者がソフトウェアチームを利用できるようにすることを目的としており、先端的な研究に特化したガイダンス、トレーニング、専門知識、ソフトウェア開発を行っている。

- 欧州社会科学データアーカイブコンソーシアム（CESSDA, n. d. [57]）のトレーニングワーキンググループは社会科学と人文学の幅広い分野の中で活躍するCESSDAサービスプロバイダーのスタッフ、アーキビスト、データプロデューサーあるいは研究者に、トレーニングやアドバイス、あるいは教育的なリソースを提供する場を設けた。これらは研究データマネジメント、データの発見と利用、デジタル保存、データのアーカイブなど様々なトピックが含まれる。

## 6.5. 国際協力

3章で説明したように、本報告書で検討したケーススタディのうちのいくつかは、本質的には国際的なもの、つまり、国際社会に貢献している。いずれのケーススタディも他のデジタルワークフォースに関するプログラムの理解を深めることに強い関心があり、ほとんどが国内外の他機関の取組を採用したり、参考としている。国際的な関わり合いには様々な形態があり、機関同士のパートナーシップ（CODATA や RDA）や、トレーニングの資料やトレーナーの共有といったものも含まれる。国際的なイニシアチブには国や機関のトレーニングと強く結びつくことがしばしば見られるが、正式なパートナーとしては常に認識されているわけではない。機関と国の競争はトレーニングのアジェンダを推進するには役立つが、経験や資料の共有から得られるものは多くある。

デジタル研究人材のニーズやそれがどの程度満たされているか、つまり、全体的なデジタルの準備段階（図5）は国によって異なる（図2）。持続可能なキャパシティの構築に関しては、国レベルと準国レベルでの責任が所在するものの、一方では、研究者がどの国に属していようと、科学と社会経済の発展のために研究できるようにするのは、世界共通の責任である。科学的な研究は本質的には国際的であり、公的な研究データはグローバルな公共財とみなしうるため、持続可能な開発目標への取組に貢献すべきである。本報告書で説明したように、データへのアクセスはデータを活用する能力がある場合にのみ役に立つのであり、グローバルコミュニティには必要なところにはどこであれデジタルスキルを備えた研究キャパシティを構築するという、共通の責任がある。

全てのセクターは次のような機会を利用すべきである。

- 関連する参加可能な国際的な協力にはいつでも従事する。
- 国やコミュニティ全体で、トレーニング資料や優れた取組、経験を共有する。

- データ集約型科学から恩恵を受け可能性があるものの、現在のところは非常に限られた能力しかない国または研究コミュニティにおけるデジタル研究人材育成への取組を支援する。

#### Box 10. どのような国際協力が行われているのか。

- 2018年以降、様々な国際的なデータ共通イニシアチブが定期的に会議を開催しており、協力を進めている。ARDC、EOSC、アフリカオープンサイエンスプラットフォーム、NIHデータコモンズ等が参加しており、連携や効率性、相互運用性を高めるフレームワークへの収束を目指している<sup>5</sup>。
- ELIXIRのトレーニング・プラットフォームには多くの協力者がおり、グローバルなライフサイエンスのデータコミュニティ構築におけるパートナーシップの重要性を示している。ELIXIRはCarpentries、CODATA、RDA、H3ABionet（アフリカの人類遺伝と健康のための汎アフリカバイオインフォマティクスネットワーク）、欧州高度コンピューティングのためのパートナーシップ（PRACE）、米国NIHのBig Data to Knowledge（BD2K、現在はData Commons）、およびバイオインフォマティクスの学習、教育、トレーニングのためのグローバル組織（GOBLET）などを含む多くの機関やイニシアチブと協力している。このうちのいくつかの機関とは正式なトレーニング協力の協定を結んでいる。また、トレーニングやキャパシティビルディングの観点からは、欧州オープンサイエンスクラウドの計画に重要な貢献を果たしている。
- 研究データ連盟（RDA）は技術と社会の架け橋を築くために、データ集約型科学に関心のある様々なコミュニティを結集するグローバルな草の根的な組織である。RDAは多くのワーキンググループで実用的な解決策を開発している。例えば、2019年のRDA第14回大会（Plenary 14）ではデータスチュワードシップを専門化するには何が必要でどのような課題があるのかを探るセッションを組み込んだ（Research Data Alliance（RDA）， n. d. [58]）。
- 世界科学データシステム（WDS、国際科学会議ISCUと協力して進めている科学データの保存・活用のための国際事業）を始めとする、データレポジトリに関する数々の国際的なネットワークが確立され、これらの多くはトレーニングとコミュニティ構築において公式または非公式に役割を果たしている（OECD， 2017[6]）。

## 7. 大学への提言

ほとんどの国では、データ集約型科学に向けたデジタル人材の能力とスキルの育成に関しては、大学が中心的な役割を担っている。大学はアカデミックな研究の支援と実施（内的）、また一般的に社会に向けて科

<sup>5</sup> SciDataCon 2018とRDA総会（Global Open Research Commons Interest Group）にて始まった。

学教育を含む教育の提供（外的）の両方に責任がある。政府や研究機関は大学を支援しインセンティブを高めるのに主要な役割を果たしており、大学は自身のポリシーを設定する自主性と能力を有している。

全てのケーススタディで、デジタルスキルのある研究者や研究サポートの専門家の創出と雇用に重要な役割があることを示している。大学はデジタルスキルを有する科学人材に対するニーズに緊急に対応する必要性があるが、相応の規模のリーダーシップと投資が必要である。しかしながら、このニーズに大学が十分に対応することは今後も難しいと、いくつかのコミュニティは考えている。なお、個々の機関には多くの優れた取組があり、それらを共有するような組織横断的な取組が増えつつある。

大学内でのデジタルスキルの支援と人材育成に目を向けた際、図書館と司書に焦点が集まるのは当然である。司書は自身の仕事でデジタル資産とツールを使用しているため、他の人に対し、データやソフトウェアを活用するためのトレーニングをすることが可能である。その際、特に、基本的なデジタルスキルやデータ管理技術についてトレーニングすることが可能である。「トレーニングは研究司書（例えば情報リテラシーの分野）の典型的なサービスの一つであり、既存スタッフの役割として受け入れられやすい。司書が実施している従来のトレーニングにデジタルスキルのトレーニングを加えることは可能である。」（Cox et al., 2017[44]）。したがって、司書は必要な投資がなされれば、大学がデジタル人材を増強するための重要なリソースとなりうる。

特にコーディングやソフトウェア開発に関連したデジタルスキルを大学でトレーニングするには、コンピュータサイエンスやITの部門が適切である。高性能のコンピューターや専門的な研究データストレージ施設などを設置する機関も、トレーニングリソースとして使用できる。場合によっては、IT企業<sup>6</sup>がデジタル研究のサポートサービスやトレーニングを、大学スタッフや産業界の両方に提供することが可能である。これらのすべてのケースにおいて、図書館と同様に効果的かつ持続可能であるためには、トレーニングが適切に支援され、評価される必要がある。実際には、ほとんどの大規模大学では様々なデジタルスキルのトレーニング機能が混在しており、すべての科学的な分野にわたって日々進化するニーズに対応するため、これらが連携して一体的に実施されることが課題である。

図書館を含む大学への主要な提言は表3のとおりである。これらの提言は5つの主要なアクションエリア（ニーズの定義づけ、デジタルスキルやフレームワーク、役割、トレーニングの提供、コミュニティの育成、キャリアパスと報酬構造、デジタル人材育成を可能とする外部要因）に分けて記載してある。また内的（アカデミックな研究に焦点）な活動と外的な活動に分けて記載している。

---

<sup>6</sup> 例えば、エジンバラパラレルコンピューティングセンターはエジンバラ大学の研究者と地方の中小企業向けの特別なITサービスとトレーニングを提供している（EPCC, n.d.[96]）。

表3. 大学や図書館に向けた主要な提言

注：内的な活動はアカデミックな研究（データ集約型の）に重点が置かれ、外的な活動はより幅広く社会に向けた科学教育においている。例えば、学部生向けの科学教育にデジタルスキルを組み込むなどが含まれる。

内的な活動：学術研究	外的な活動：科学と社会
1. ニーズの定義づけ：デジタルスキル、フレームワーク、役割	
<ul style="list-style-type: none"> <li>▪ デジタル研究のニーズに関する調査を行い、それらに対処するための戦略を展開する。</li> <li>▪ デジタルスキルがオープンサイエンスや研究公正を推進するためのフレームワークに確実に取り込まれるようにする。</li> </ul>	<ul style="list-style-type: none"> <li>▪ デジタルスキルやトレーニングのニーズを図るために、非アカデミックのステークホルダー（例えば、産業界）と協力する。</li> </ul>
2. トレーニングの提供	
<ul style="list-style-type: none"> <li>▪ 研究者や研究サポートの専門家に基礎的あるいは専門的なデジタルスキルのトレーニングを幅広く提供する。</li> <li>▪ スタッフのトレーニングを提供し、ファンドの申請書に明確に定義づけられた十分なデータ管理計画やソフトウェアと同等のもの（例えば、ソフトウェア管理計画）が確実に記載されるように支援する。</li> <li>▪ デジタルスキルに研究を管理するために特別に必要なメンターシップを含める。</li> </ul>	<ul style="list-style-type: none"> <li>▪ すべての分野において、一般的な学部生にも基礎的なデジタルスキルを提供する。</li> <li>▪ 学部レベルから生涯学習に至るまで、包括的で関連性がある最新のデジタルスキルをカリキュラムに取り込む。</li> <li>▪ 教育カリキュラムとトレーニングに柔軟性を確保し、技術的な変化に対応できるようにする。</li> <li>▪ 産業界を含む外部ユーザーと協力し、長期的なデジタル研究トレーニングプログラムを開発し支援する。</li> </ul>
3. コミュニティビルディング	
<ul style="list-style-type: none"> <li>▪ ワークフォースの多様性をデジタルキャパシティの主要要素として積極的に推進する。</li> <li>▪ トレーニングの提供、コミュニティの関与、キャリア評価におけるリーダーシップの重要性を認識する。</li> <li>▪ 新しい役割を持つデータスチュワードやRSEや、デジタルスキルのイニシアチブにおけるトレーナーやリーダーたちの専門的なコミュニティの発展を支援する。</li> <li>▪ データ集約型研究を支援するためにデジタルスキルのサポートデスクを設置し、アシスタンスやそのようなスキルを持ったスタッフを研究チームに取り込む。</li> </ul>	
4. キャリアパスと報酬構造	



<ul style="list-style-type: none"> <li>▪ デジタルの観点で研究サポートができる専門家をキャリアの昇進機会と共に確固たるポジションで雇うようなフレームワークを開発し採用する。</li> <li>▪ キャリアパスと代替的な指標に関して体系的な制度変更を実施し、あらゆる学問レベルで幅広い研究への貢献を認識し、報酬を与えるようにする。</li> <li>▪ データセットやソフトウェアを研究のアウトプットとして価値のあるものとして認める。</li> <li>▪ 研究チームのすべてのメンバーの貢献を認識する指標を使用して、ドメイン研究者（domain researchers）とデジタルサポートスタッフ間の協力を奨励する。</li> </ul>	<ul style="list-style-type: none"> <li>▪ 産業界と学術界の協力や交流を推進し、学術界の外で成功した人々の中にも、再度学術界に戻ることを希望している者がいることを認識する。</li> <li>▪ セクター間でのデジタルスキルのイニシアチブを支援する。</li> </ul>
<p>5. デジタルワークフォースキャパシティに向けた幅広い成功要因</p>	
<ul style="list-style-type: none"> <li>▪ オープンサイエンス、再現性、研究公正の原則を採用する。</li> <li>▪ データ集約型研究における倫理的な問題を評価することのできる評価委員会を設置する。</li> </ul>	<ul style="list-style-type: none"> <li>▪ オープンサイエンス、再現性、研究公正を政策レベルで支援する。</li> <li>▪ デジタル研究人材に求められる能力の定義づけや対処方針について、民間セクターと協力する。</li> </ul>

様々な場所にある大学が連携し、学びあい、専門知識とリソースを組み合わせ、共通のデジタル人材のトレーニングの課題に対処できる可能性は大いにある。非公式のネットワークや交流によってだけでなく、より公式的なパートナーシップを通じて行うことができ、機関全体のレベル、適切な部局レベル、あるいは個人的なネットワークを通じて行うことができる。中でも図書館は大学間の関与のための重要である。大学間協力のいくつかの例はBox 11に示している。

#### Box. 11 大学はどのようにして協働するのか

- ADSAIは2019年に米国でデータサイエンスリーダーシップサミットを開催し、データサイエンスの研究所、センター、プログラム、学部のリーダー、新しいイニシアチブの構築に関心を持つ教員を招集した。当サミットはデータサイエンスのためのアカデミックなコミュニティの構築を目的としており、同様の課題や機会に直面するベスト・プラクティスを共有し、次世代のデータサイエンティストが社会の最善の利益に貢献できるよう準備する際に責任を負うことが必要であることを指摘した（The 2019 Data Science Leadership Summit, 2019[59]）。
- 英国再現性ネットワーク（UKRN）は20以上の大学とそのほかのステークホルダーが参加するコンソーシアムであり、活発な研究活動に貢献する要因を調査し、トレーニング活動を促進し、ベスト・プラクティスを広め、利害関係者と協力して取組を調整している（University of Bristol, n. d. [60]）。SSIはUKRNと様々な共同イニシアチブで提携しており、ソフトウェア開発やデータ管理スキルのトレーニングも含まれている（SSI, 2020[61]）。英国はイングランド北部地域の8大学に



よる研究パートナーシップ（N8 Research Partnership）間で新たな協力関係を築いており、RSEグループの類似点、課題、機会に関して調査を行っている（N8 CIR, 2020[62]）。

- 2020年初頭に、複数の国の8つの大学ネットワークのリーダーが研究データの権利に関するソルボンヌ宣言に署名した。署名者は次のことを宣言する。「オープンな研究データ管理を促進する環境を生み出すトレーニングとスキル開発のプログラムを立ち上げることを大学に促す」（LERU, 2020[63]）。署名したのはアメリカ大学協会（AAU）、アフリカ研究大学同盟（ARUA）、フランス研究型大学（CURIF）、ドイツU15、欧州研究大学リーグ（LERU）、日本学術研究懇談会（RU11）、英国ラッセルグループ、オーストラリアGroup of Eight（Go8）である。

大学図書館は様々な方法でデジタル科学人材の育成に向けて、努力を行っている。

- JPCOARIは、他の機関が公開しているトレーニング資料も使用して、日本の大学図書館員に向けてオンライントレーニングコースを開発している。これらのコースは全国的なMOOCプラットフォームを介して提供され、基礎的な研究データ管理スキルに特化している。
- オーストラリア司書協会（CAUL）は次の二つの目標を設定し、デジタルDexterity（器用）プログラムを導入した。1つめの目標は、自らのミッションを遂行する上でデジタルに”精通している”ことやCAULに自発的に参加することが重要であることをオーストラリアの大学が認識すること。2つ目の目標は、卒業生がデジタルスキルを利用して、グローバルな職場環境で成長し、グローバルな市民として活躍することである（Council of Australian University Librarians, 2019[64]）。
- カナダの研究図書館協会（CARL）は会員向けのデータサービスとスキル開発をサポートし、データ管理計画など、必要に応じた全国的なサービスを提供することを目的とし、複数年の活動に資金を提供している。継続にあたっては連邦政府からのファンドも集めている。
- ヨーロッパの研究図書館ネットワークである”Ligue des Bibliothèques Européennes de Recherche（LIBER）”は、物理的かつバーチャルな研究環境におけるデジタルスキルとサービスのハブとしての研究図書館を展開している。LIBERはトレーニングやデータサービスを会員向けに提供し、データスキルの開発を支援するために政府機関と連携している。

## 8. 結び：一貫した政策的措置の必要性

デジタルテクノロジーはこれまでの科学を変化させており、デジタルスキル人材は科学の可能性やデータ集約型科学の社会経済的な便益を最大化するために必要である。デジタルトランスフォーメーションに関連する人材の課題に対応するには、スピードを加速させ規模を拡大させる必要がある。多くの国の様々な機関が研究のためのデジタル人材に関する課題を解決するイニシアチブに取り組んでいるが、より有益なものとするには協力が必要である。多くの優れたイニシアチブが進行中であり、他の取組から学ぶべきものは多い。

本報告書はデータ集約型科学に向けた人材を育成し、持続させるために必要な要素を検討し、5つの主要なアクション領域、つまり、定義づけ、トレーニングの提供、コミュニティ形成、キャリアパスと報酬、デジ

タル人材育成のための外部要因、に焦点を置いた。様々な分野のあらゆるレベルでこれらを扱う多くの取組があるが、5つすべてを包括する戦略的な施策はほとんどない。

サイエンスエコシステム全体にわたり、様々な関係者がこれらの分野で実践している。本報告書では政策立案者やその他の関係者が注力する主要な領域を特定し、デジタルスキルを備えた人材の加速に向けたアクションについて具体的な提言を記載した。

個人あるいは機関がデータ集約型科学に向けたデジタルスキルを磨くために積極的に取り組んでおり、彼らの努力が次のような共通のビジョンを達成するため、政策立案者は効果を最大化するための十分な規模の支援を実行する必要がある：

*効果的な政策、インセンティブ、投資が未来志向の世界をリードする研究人材を支え、データ集約型科学の可能性を最大限に引き出し、複雑な社会的課題へのソリューションを提供できる。*

## 参考文献

- Apostel, L. et al. (eds.) (1972), *Towards Interdisciplinarity and Transdisciplinarity in Education and Innovation*, OECD Publications Center. [74]
- Arnstein, S. (1969), “A Ladder Of Citizen Participation”, *Journal of the American Institute of Planners*, Vol. 35/4, pp. 216-224, <http://dx.doi.org/10.1080/01944366908977225>. [75]
- Arts and Humanities Research Council (2014), *AHRC Research Training Framework for Doctoral Students*, <https://ahrc.ukri.org/documents/projects-programmes-and-initiatives/ahrc-research-training-framework-for-doctoral-students/> (accessed on 8 June 2020). [52]
- Ashley, K. (2016), *Review: Developing skills for managing research data and software*, Wellcome Trust, <https://doi.org/10.6084/m9.figshare.4133916.v1>. [11]
- Australian Academy of the Humanities (n.d.), *Future Humanities Workforce*, <https://www.humanities.org.au/advice/projects/future-workforce/> (accessed on 8 June 2020). [53]
- Australian Research Data Commons (n.d.), *The Australian eResearch Skilled Workforce Summit*, <https://ardc.edu.au/events/the-australian-eresearch-skilled-workforce-summit/> (accessed on 5 June 2020). [95]
- Babuska, I. and J. Oden (2004), “Verification and validation in computational engineering and science: basic concepts”, *Computer Methods in Applied Mechanics and Engineering*, Vol. 193/36-38, pp. 4057-4066, <http://dx.doi.org/10.1016/j.cma.2004.03.002>. [67]
- Bello, M. and F. Galindo-Rueda (2020), “Charting the digital transformation of science: Findings from the 2018 OECD International Survey of Scientific Authors (ISSA2)”, *OECD Science, Technology and Industry Working Papers*, No. 2020/03, OECD Publishing, Paris, <https://dx.doi.org/10.1787/1b06c47c-en>. [8]
- Berente, N. et al. (2018), *Organizing and the Cyberinfrastructure Workforce*, [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3260715](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3260715). [39]
- Buchhorn, M. (2019), *Surveying the scale of the research-IT support workforce*, <https://ardc.edu.au/wp-content/uploads/2019/07/ARDC-National-Workforce-report-final-v3.pdf> (accessed on 5 June 2020). [16]
- CANARIE (2020), *Local Research Software Support – Call 1*, <https://www.canarie.ca/software/funding/lrss-call1/> (accessed on 8 June 2020). [56]
- CASRAI (n.d.), *Resources*, <https://casrai.org/resources/> (accessed on 5 June 2020). [66]



- CESSDA (n.d.), *Working Groups*, <https://www.cessda.eu/About/Working-Groups> (accessed on 8 June 2020). [57]
- Council of Australian University Librarians (2019), *CAUL Digital Dexterity Position Statement*, <https://www.caul.edu.au/caul-digital-dexterity-position-statement> (accessed on 8 June 2020). [64]
- Cox, A. et al. (2017), “Developments in research data management in academic libraries: Towards an understanding of research data service maturity”, *Journal of the Association for Information Science and Technology*, Vol. 68/9, pp. 2182-2200, <http://dx.doi.org/10.1002/asi.23781>. [44]
- Dai, Q., E. Shin and C. Smith (2018), “Open and inclusive collaboration in science: A framework”, *OECD Science, Technology and Industry Working Papers*, No. 2018/07, OECD Publishing, Paris, <https://dx.doi.org/10.1787/2dbff737-en>. [5]
- DARIAH-EU (2019), *DESIR Winter School: Shaping new approaches to data management in arts and humanities*, <https://www.dariah.eu/2019/09/09/desir-winter-school-shaping-new-approaches-to-data-management-in-arts-and-humanities/> (accessed on 8 June 2020). [55]
- Demchenko, Y., A. Belloum and T. Wiktorski (2017), *EDISON Data Science Framework: Part 1. Data Science Competence Framework (CF-DS) Release 2*, <https://doi.org/10.5281/zenodo.1044346>. [26]
- DORA (2019), *Quality over quantity: How the Dutch Research Council is giving researchers the opportunity to showcase diverse types of talent*, <https://sfdora.org/2019/11/14/quality-over-quantity-how-the-dutch-research-council-is-giving-researchers-the-opportunity-to-showcase-diverse-types-of-talent/> (accessed on 8 June 2020). [51]
- DORA (2012), *San Francisco Declaration on Research Assessment*, <https://sfdora.org/read/>. [41]
- Dutch Techcentre for Life Sciences (n.d.), *About Research Data Management*, <https://www.dtls.nl/fair-data/research-data-management/research-data-management/> (accessed on 5 June 2020). [68]
- EC/OECD (2020), *STIP Compass: International Database on Science, Technology and Innovation Policy (STIP)*, <https://stip.oecd.org> (accessed on 14 May 2020). [93]
- EC/OECD (2020), *STIP Compass: International Database on Science, Technology and Innovation Policy (STIP)*, <https://stip.oecd.org> (accessed on 5 June 2020). [43]
- Enengel, B. et al. (2012), “Co-production of knowledge in transdisciplinary doctoral theses on landscape development—An analysis of actor roles and knowledge types in different research phases”, *Landscape and Urban Planning*, Vol. 105/1-2, pp. 106-117, <http://dx.doi.org/10.1016/j.landurbplan.2011.12.004>. [86]
- EOSCsecretariat.eu (n.d.), *Skills & Training Working Group*, <https://www.eoscsecretariat.eu/working-groups/skills-training-working-group> (accessed on 8 June 2020). [48]
- EPCC (n.d.), *EPCC*, <https://www.epcc.ed.ac.uk/> (accessed on 8 June 2020). [96]
- European Commission (2019), *Cost-benefit analysis for FAIR research data*, <http://dx.doi.org/10.2777/029999>. [9]



- European Commission (2018), *Turning FAIR into reality: Final report and action plan from the European Commission expert group on FAIR data*, Publications Office of the European Union, <http://dx.doi.org/doi:10.2777/1524>. [21]
- European Commission (2017), *Evaluation of Research Careers fully acknowledging Open Science Practices; Rewards, incentives and/or recognition for researchers practicing Open Science*, Publications Office of the European Union, <http://dx.doi.org/doi:10.2777/75255>. [23]
- European Strategy Forum on Research Infrastructures (2018), *Big Data and e-Infrastructure Needs*, ESFRI, <http://roadmap2018.esfri.eu/>. [7]
- Europe's distributed infrastructure for life-science data (n.d.), *Welcome to TeSS: ELLXIR's Training Portal*, <https://tess.elixir-europe.org/> (accessed on 8 June 2020). [42]
- FAIR4S (2019), *EOSC FAIR4S*, <https://eosc-fair4s.github.io/> (accessed on 5 June 2020). [24]
- German Council for Scientific Information Infrastructures (2019), *Empfehlungen zu Berufs- und Ausbildungsperspektiven für den Arbeitsmarkt Wissenschaft*, <http://www.rfii.de/download/digitale-kompetenzen-dringend-gesucht>. [45]
- German Council for Scientific Information Infrastructures (RfII) (2019), *DIGITAL COMPETENCIES – URGENTLY NEEDED!*, <http://www.rfii.de/?p=4015> (accessed on 5 June 2020). [20]
- Grant, S. and K. Bouskill (2019), *Open Science and Institutional Review Boards: Aligning Transparency with Regulatory Protections for Human Research Subjects*, <http://www.metascience2019.org/poster-session/sean-grant/>. [29]
- Gredig, D. (2011), "From research to practice: Research-based Intervention Development in social work: developing practice through cooperative knowledge production", *European Journal of Social Work*, Vol. 14/1, pp. 53-70, <http://dx.doi.org/10.1080/13691457.2010.516624>. [81]
- Hadorn, G. et al. (eds.) (2008), *Handbook of Transdisciplinary Research*, Springer Netherlands, Dordrecht, <http://dx.doi.org/10.1007/978-1-4020-6699-3>. [73]
- Hernandez Montoya, A. (ed.) (2017), "Interdisciplinary Collaboration between Natural and Social Sciences – Status and Trends Exemplified in Groundwater Research", *PLOS ONE*, Vol. 12/1, p. e0170754, <http://dx.doi.org/10.1371/journal.pone.0170754>. [76]
- Houghton, J. and N. Gruen (2014), *Open Research Data report*, <https://www.ands.org.au/working-with-data/articulating-the-value-of-open-data/open-research-data-report>. [10]
- Hulhoven, X. (n.d.), *Develop a sustainable future for Brussels through a co-creation project*, <https://innoviris.brussels/co-creation> (accessed on 27 May 2020). [89]
- Jahn, T., M. Bergmann and F. Keil (2012), "Transdisciplinarity: Between mainstreaming and marginalization", *Ecological Economics*, Vol. 79, pp. 1-10, <http://dx.doi.org/10.1016/j.ecolecon.2012.04.017>. [82]
- Kiser, G. and Y. Mantha (n.d.), *Global AI Talent Report 2019*, <https://jfgagne.ai/talent-2019/> (accessed on 5 June 2020). [17]

- Klein, J. (2008), "Evaluation of Interdisciplinary and Transdisciplinary Research", *American Journal of Preventive Medicine*, Vol. 35/2, pp. S116-S123, <http://dx.doi.org/10.1016/j.amepre.2008.05.010>. [78]
- Kurata, K., M. Matsubayashi and M. Takeda (2017), "Research data management in Japanese universities and research institutions: Status report based on questionnaire survey", *Journal of Information Processing and Management*, Vol. 60/2, pp. 119-127, <https://doi.org/10.1241/johokanri.60.119>. [14]
- Lang, D. et al. (2012), "Transdisciplinary research in sustainability science: practice, principles, and challenges", *Sustainability Science*, Vol. 7/S1, pp. 25-43, <http://dx.doi.org/10.1007/s11625-011-0149-x>. [83]
- LERU (2020), *Data Summit in Paris*, <https://www.leru.org/news/data-summit-in-paris> (accessed on 8 June 2020). [63]
- Lonie, A. and R. Francis (2017), *An Australian Bioscience Data Capability Project Report - Phase 2*, <https://www.embl-abr.org.au/wp-content/uploads/2018/04/ABDC-Project-Report-Phase-2.pdf>. [13]
- Méndez, E. (2019), *Open Science?... Darling, we need to talk*, <https://www.open-science-conference.eu/wp-content/uploads/2019/03/Eva-Mendez.pdf>. [37]
- Ministry of Education, Culture, Sports, Science and Technology (n.d.), *Doctoral program for Data-Related Innovation Expert(D-DRIVE)*, [https://www.mext.go.jp/a\\_menu/jinzai/data/index.htm](https://www.mext.go.jp/a_menu/jinzai/data/index.htm) (accessed on 8 June 2020). [46]
- Molloy, L., A. Gow and L. Konstantelos (2014), "The DigCurV Curriculum Framework for Digital Curation in the Cultural Heritage Sector", *International Journal of Digital Curation*, Vol. 9/1, pp. 231-241, <http://dx.doi.org/10.2218/ijdc.v9i1.314>. [27]
- Mons, B. (2020), "Invest 5% of research funds in ensuring data are reusable", *Nature*, Vol. 578/7796, pp. 491-491, <http://dx.doi.org/10.1038/d41586-020-00505-7>. [1]
- Morgan, S. et al. (2017), "The ELIXIR-EXCELERATE Train-the-Trainer pilot programme: empower researchers to deliver high-quality training", *F1000Research*, Vol. 6, p. 1557, <http://dx.doi.org/10.12688/f1000research.12332.1>. [33]
- Moser, S. (2016), "Can science on transformation transform science? Lessons from co-design", *Current Opinion in Environmental Sustainability*, Vol. 20, pp. 106-115, <http://dx.doi.org/10.1016/j.cosust.2016.10.007>. [69]
- N8 CIR (2020), *N8 RSE Leaders and Aspiring Leaders Meeting*, <https://n8cir.org.uk/events/rse-aspiring-leaders/> (accessed on 8 June 2020). [62]
- National e-Science Postgraduate Teaching and Training Platform (n.d.), *National e-Science Postgraduate Teaching and Training Platform*, <http://www.escience.ac.za/> (accessed on 8 June 2020). [47]
- Noah, W. and M. Jean (1980), *Webster's New Twentieth Century Dictionary of the English Language, Unabridged*, W. Collins. [72]
- OECD (2019), *Measuring the Digital Transformation: A Roadmap for the Future*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264311992-en>. [19]



- OECD (2019), *OECD Skills Outlook 2019 : Thriving in a Digital World*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/df80bc12-en>. [22]
- OECD (2018), *Bridging the Digital Gender Divide: Include, upskill, innovate*, <http://www.oecd.org/internet/bridging-the-digital-gender-divide.pdf>. [31]
- OECD (2017), "Business models for sustainable research data repositories", *OECD Science, Technology and Industry Policy Papers*, No. 47, OECD Publishing, Paris, <https://dx.doi.org/10.1787/302b12bb-en>. [4]
- OECD (2017), "Co-ordination and support of international research data networks", *OECD Science, Technology and Industry Policy Papers*, No. 51, OECD Publishing, Paris, <https://dx.doi.org/10.1787/e92fa89e-en>. [6]
- OECD (2017), "Open research agenda setting", *OECD Science, Technology and Industry Policy Papers*, No. 50, OECD Publishing, Paris, <https://dx.doi.org/10.1787/74edb6a8-en>. [88]
- OECD (2016), "Research Ethics and New Forms of Data for Social and Economic Research", *OECD Science, Technology and Industry Policy Papers*, No. 34, OECD Publishing, Paris, <https://dx.doi.org/10.1787/5jln7vnpxs32-en>. [30]
- Ouellette, F. (ed.) (2017), "Unmet needs for analyzing biological big data: A survey of 704 NSF principal investigators", *PLOS Computational Biology*, Vol. 13/10, p. e1005755, <http://dx.doi.org/10.1371/journal.pcbi.1005755>. [12]
- Piatetsky, G. (2018), *How many data scientists are there and is there a shortage?*, <https://www.kdnuggets.com/2018/09/how-many-data-scientists-are-there.html> (accessed on 5 June 2020). [2]
- Pohl, C., P. Krütli and M. Stauffacher (2017), "Ten Reflective Steps for Rendering Research Societally Relevant", *GALA - Ecological Perspectives for Science and Society*, Vol. 26/1, pp. 43-51, <http://dx.doi.org/10.14512/gaia.26.1.10>. [84]
- Pohl, C. et al. (2011), *Questions to evaluate inter- and transdisciplinary research proposals*, td-net for Transdisciplinary Research. [90]
- Research Data Alliance (2015), *Task Force on Defining data handling related competences and skills for different groups of professions - Working area*, <https://www.rd-alliance.org/group/education-and-training-handling-research-data-ig/wiki/task-force-defining-data-handling> (accessed on 5 June 2020). [25]
- Research Data Alliance (RDA) (n.d.), *RDA 14th Plenary - Programme*, <https://www.rd-alliance.org/rda-14th-plenary-programme> (accessed on 22 June 2020). [58]
- Research Data Alliance International Indigenous Data Sovereignty Interest Group (n.d.), *CARE Principles for Indigenous Data Governance*, <https://www.gida-global.org/care> (accessed on 5 June 2020). [32]
- Roux, D. et al. (2010), "Framework for participative reflection on the accomplishment of transdisciplinary research programs", *Environmental Science & Policy*, Vol. 13/8, pp. 733-741, <http://dx.doi.org/10.1016/j.envsci.2010.08.002>. [77]



- Schneider, F. et al. (2019), "Research funding programmes aiming for societal transformations: Ten key stages", *Science and Public Policy*, Vol. 46/3, pp. 463-478, <http://dx.doi.org/10.1093/scipol/scy074>. [87]
- Scholtens, S. et al. (2019), *Life sciences data steward function matrix*, <https://doi.org/10.5281/zenodo.2554973>. [28]
- Science Council (n.d.), *What is the Technician Commitment?*, <https://sciencecouncil.org/employers/technician-commitment/> (accessed on 5 June 2020). [35]
- Science Europe (2018), *Science Europe Data Glossary*, [http://sedataglossary.shoutwiki.com/wiki/Main\\_Page](http://sedataglossary.shoutwiki.com/wiki/Main_Page) (accessed on 5 June 2020). [65]
- Scroggins, M. and I. Pasquetto (2020), "Labor Out of Place: On the Varieties and Valences of (In)visible Labor in Data-Intensive Science", *Engaging Science, Technology, and Society*, Vol. 6, p. 111, <http://dx.doi.org/10.17351/ests2020.341>. [36]
- Software Sustainability Institute (n.d.), *Research Software Engineers*, <https://www.software.ac.uk/research-software-engineers> (accessed on 9 June 2020). [38]
- Springer, R. (2019), *Counting Data Librarians*, <https://sr.ithaka.org/blog/counting-data-librarians/> (accessed on 5 June 2020). [15]
- SSI (2020), *Software Sustainability Institute joins forces with the UK Reproducibility Network*, <https://www.software.ac.uk/news/software-sustainability-institute-joins-forces-uk-reproducibility-network> (accessed on 8 June 2020). [61]
- Stauffacher, M. et al. (2008), "Analytic and Dynamic Approach to Collaboration: A Transdisciplinary Case Study on Sustainable Landscape Development in a Swiss Prealpine Region", *Systemic Practice and Action Research*, Vol. 21/6, pp. 409-422, <http://dx.doi.org/10.1007/s11213-008-0107-7>. [85]
- Swiss Academies of arts and sciences (n.d.), *td-net Network for Transdisciplinary Research*, <http://www.transdisciplinarity.ch/en/td-net/Transdisziplinarit-t/Definitionen.html> (accessed on 27 May 2020). [91]
- Teal, T. et al. (2015), "Data Carpentry: Workshops to Increase Data Literacy for Researchers", *International Journal of Digital Curation*, Vol. 10/1, pp. 135-143, <http://dx.doi.org/10.2218/ijdc.v10i1.351>. [18]
- The 2019 Data Science Leadership Summit (2019), *About the Summit*, <https://sites.google.com/msdse.org/datascienceleadership2019/home/about-the-summit> (accessed on 8 June 2020). [59]
- The Academy of Medical Sciences (2019), *Artificial intelligence and health: Summary report of a roundtable held on 16 January 2019*, <https://acmedsci.ac.uk/file-download/77652269>. [54]
- The Carpentries (2019), *The Carpentries Equity, Inclusion, and Accessibility Roadmap*, [https://carpentries.org/files/assessment/equity\\_inclusion\\_accessibility\\_roadmap.pdf](https://carpentries.org/files/assessment/equity_inclusion_accessibility_roadmap.pdf). [34]
- The Royal Society (2019), *Dynamics of data science skills*, <https://royalsociety.org/topics-policy/projects/dynamics-of-data-science/> (accessed on 5 June 2020). [3]

- The World Commission on Environment and Development (1987), *Report of the World Commission on Environment and Development - Our Common Future*. [79]
- Tress, B., G. Tress and G. Fry (2006), "Defining concepts and the process of knowledge production in integrative research", in Tress, B. et al. (eds.), *From Landscape Research to Landscape Planning: Aspects of Integration, Education and Application*, Springer, Dordrecht. [71]
- TUBİTAK (n.d.), *Research Data Management Training Portal*, <https://acikveri.ulakbim.gov.tr/> (accessed on 8 June 2020). [49]
- TUBİTAK (n.d.), *TUBİTAK Open Science Policy Accepted*, <https://ulakbim.tubitak.gov.tr/en/haber/tubitak-open-science-policy-accepted> (accessed on 8 June 2020). [50]
- University of Bristol (n.d.), *The UK Reproducibility Network*, <http://www.bristol.ac.uk/psychology/research/ukrn/> (accessed on 8 June 2020). [60]
- Van Noorden, R. (2015), "Interdisciplinary research by the numbers", *Nature*, Vol. 525/7569, pp. 306-307, <http://dx.doi.org/10.1038/525306a>. [70]
- Vector Institute (n.d.), *List of Recognized AI-Related Programs*, <https://vectorinstitute.ai/list-of-recognized-ai-related-programs/> (accessed on 5 June 2020). [94]
- Working Towards Sustainable Software for Science: Practice and Experience (n.d.), *About WSSSPE*, <http://wssspe.researchcomputing.org.uk/about-wssspe/> (accessed on 8 June 2020). [40]
- Wright Morton, L., S. Eigenbrode and T. Martin (2015), "Architectures of adaptive integration in large collaborative projects", *Ecology and Society*, Vol. 20/4, <http://dx.doi.org/10.5751/es-07788-200405>. [80]
- Zinsstag, J. et al. (2015), *One Health: The Theory and Practice of Integrated Health Approaches*, CABI. [92]



## 専門家グループメンバー

Country	Name	Affiliation
Australia	Michelle Barker (Chair of Expert Group)	Director, Skilled Workforce and Partnerships, ARDC
Belgium	Bart Dumolyn*	Department of Economy, Science and Innovation, Flemish Government, Brussels Area
Belgium	Inge van Nieuwerburgh*	Ghent University Library
Canada	David Castle	Vice-President Research, University of Victoria
Chile	Marcelo Arenas	Professor, Dept of Computer Science, Catholic University of Santiago
European Commission	Konstantinos Repanas	Unit for Open Science, Directorate-General Research and Innovation
European Commission	Carlos Casorran	Unit for Open Science, Directorate-General Research and Innovation
France	Nathalie Denos*	Higher Education General Directorate, Ministry for Higher Education and Research
France	Mehdi Gharsallah*	Higher Education General Directorate, Ministry for Higher Education and Research
Germany	Ingvill C. Mochmann	Head of EUROLAB, Knowledge Transfer, GESIS-Leibniz Institute for the Social Sciences, Vice President for Research and Knowledge Transfer, Professor of International Politics, Cologne Business School
Japan	Nobukazu Yoshioka	Associate Professor, National Institute of Informatics
Korea	Seo-Young Noh	Assistant Professor, Chungbuk National University, Korea Institute of Science and Technology Information (KISTI)
Netherlands	Karel Luyben	Rector Magnificus Emeritus, Delft University of Technology
Norway	Gard Thomassen	Assistant Director, University Centre for Information Technology, University of Oslo
United Kingdom	David McAllister	Associate Director, Research and Innovation Talent- UKRI-Biotechnology and Biological Sciences Research Council (BBSRC)
United Kingdom	Kevin Ashley	Director, DCC
United Kingdom	Lauren Clarke	International Policy Manager, UK Research and Innovation
United States	Daniel S. Katz	Assistant Director, National Center for Supercomputing Applications (NCSA), University of Illinois at Urbana-Champaign
United States	Todd K. Leen	Program Director, Computer Information Science and Engineering, Information Intelligent Systems, NSF
The Carpentries	Tracy Teal	Executive Director, The Carpentries
CODATA	Simon Hodson*	Executive Director, CODATA, International Science Council
SSI	Neil Chue Hong	Director, SSI, University of Edinburgh

\* Participated in the 1<sup>st</sup> EG meeting only.

## ケーススタディ質問票

### 1. *General information*

- a) Please give a brief description of this initiative.
- b) How is it funded/supported? What is its business model?
- c) When did it start and what is its expected duration?

### 2. *Overview and focus*

- a) What types of skills does your initiative aim to help develop, for who and through what delivery method/s?
- b) How did you decide which skills to focus on (e.g., generic vs discipline-specific, technical skills versus softer skills like collaboration)?
- c) Can you map highlight the cells in the meta-framework (that is provided as an annex) that are the main focus of your initiative?
- d) Are the roles identified in the meta-framework (see annex) appropriate for your project/policy/initiative? Is the framework missing some key roles?
- e) Are the skills (horizontal axis) identified in the meta-framework appropriate for your project/policy/initiative? Is the framework missing some key skills?
- f) Are there other initiatives with a similar focus to this one and do you cooperate with these?

### 3. *Broader context and drivers*

- a) What was the main driver for your program? Has your program been driven by an assessment of supply and/or demand (or some sort of market analysis), and if so, what?
- b) Does your program look at broader enablers to facilitate the development of a skilled workforce, such as career paths and recognition for different types of research outputs, inclusion of digital expertise in research teams, funder mandates to recognise the importance of digital professionals etc.? How does your initiative fit with this broader context?
- c) Can you identify broader enablers that support the success of your program, e.g. does your program implement more successfully when certain types of support from other organisations is present?

### 4. *Policy requirements*

- a) What are the best examples of actions (e.g., mandates or incentives, including funding) which have led to improvements in digital skills development for data intensive research at scale?
- b) What actions are required to promote digital skills for research by the following actors:
  - National or regional governments
  - Research agencies
  - Professional science associations
  - Research institutes
  - Research community

### 5. *Future perspectives*

- a) What are the most in-demand digital skills for science? What are the skills whose demand is growing fastest? Is there any skill that you think will not be so important in the future?
- b) Does your initiative take into account how research may be different in 10 years (influenced by digitalisation, e.g., AI)?
- c) Do our current skills programs/policies ready us for this?
- d) If not, what new skills or new programs will be needed? What enablers would be needed to support these?



## ワークショップ参加者

The workshop was held on 28-29 October 2019 at the Leibniz Institute for Social Sciences (GESIS), Cologne, Germany.

The workshop was organised in four main sessions that focussed on: 1. The broad context and drivers for digital skills; 2. Defining skills needs; 3. Policy requirements; 4. Future perspectives.

### Attendees

Country	Name	Affiliation
Australia	Ms. Michelle BARKER*	Director, Skilled Workforce and Partnerships, ARDC
	Ms. Rachel WEBSTER	Head of Astrophysics, University of Melbourne
Belgium	Mr. Alexander BOTZKI	Manager BITS Core Facility, Bioinformatics, Flemish life Sciences Institute VIB
Canada	Mr. David CASTLE*	Vice-President Research, University of Victoria
Chile	Mr. Andres JORDAN	Director, MAS
France	Mr. Cedric LOMBION	Open Data and Data Literacy Consultant, School of Data / Open Knowledge France
	Mr. François MICHONNEAU	Infrastructure Team Lead, Curriculum Development Lead, The Carpentries
	Ms. Laura MOLLOY	Committee on Data for Science and Technology, CODATA
Germany	Ms. Helene BRINKEN	State and University Library, Georg-August-University Göttingen
	Mr. Dietmar JANETZKO*	Professor of Information Systems and Business Process Management, Cologne Business School
Japan	Mr. Nobukazu YOSHIOKA*	Associate Professor, Information Systems Architecture Science Research Division, National Institute of Informatics
Netherlands	Mr. Alastair DUNNING	Head of Research Data Services at TU Delft and Head of 4TU, Centre for Research Data, Technische Universiteit Delft
	Ms. Celia VAN GELDER	Head of Training Platform, Dutch Techcentre for Life Sciences (DTL), ELIXIR
New Zealand	Mr. Nick JONES	Director, NeSI, University of Auckland
Norway	Mr. Gard THOMASSEN*	Assistant Director - Research Computing, University of Oslo
South Africa	Mr. Anwar VAHED	Director, Data Intensive Research Initiative of South Africa, Council for scientific and industrial Research (CSIR)
United Kingdom	Mr. Harriet BARNES	Head of Higher Education & Skills, British Academy
	Mr. Simon HETTRICK	Deputy Director, Software Sustainability Institute
	Ms. Corinne MARTIN	External Relations Officer, ELIXIR
	Mr. David MCALLISTER*	Associate Director – Research and Innovation Talent, BBSRC, UK Research and Innovation
	Mr. Ben MURTON	Head of Professional and Academic Development, the Alan Turing Institute
	Mr. Xavier POTAU	Principal Consultant, Technopolis
United States	Mr. Daniel S. KATZ	Assistant Director, National Center for Supercomputing Applications (NCSA), University of Illinois at Urbana-Champaign
	Ms. Tracy TEAL*	Executive Director, The Carpentries
European Union	Mr. Konstantinos REPANAS	Policy Officer, Open Science & EOSCDG Research
GESIS-Leibniz Institute for the Social Sciences	Mr. Fabian FLÖCK	Department Knowledge Transfer, Team EUROLAB, GESIS
	Ms. Ingvill Constanze MOCHMANN*	Department Knowledge Transfer, Team EUROLAB, GESIS
	Ms. Katrin WELLER	Team Leader "Social Analytics and Services, Computational Social Science, GESIS
Project consultant	Mr. Kevin ASHLEY*	Director, DCC, University of Edinburgh
OECD-STI/GSF	Mr. Carthage SMITH	Senior Policy Analyst
	Mr. Yoshiaki TAMURA	Policy Analyst

\* Members of the OECD GSF Expert Group on Digital Skills for Data Intensive Science.

## 用語集

One of the challenges in considering the digital workforce needs of science is reaching a common understanding across different communities, including scientists from different disciplines, and educators and policymakers from different backgrounds. This is further complicated by translation into different languages and cultures. Many detailed glossaries of digital terms in relation to science already exist. The key words that are used regularly in the current report – and which are likely to be useful to science policy makers - are defined below. These definitions have been adopted or adapted from other works in the field (Science Europe, 2018<sup>[65]</sup>; CASRAI, n.d.<sup>[66]</sup>; Babuska and Oden, 2004<sup>[67]</sup>) and are commonly used in the scientific research community.

Artificial Intelligence (AI)	AI is the theory and development of computer systems able to perform tasks normally requiring human intelligence
Big data	Big data is an evolving term that describes any voluminous amount of structured, semi-structured or unstructured data that has the potential to be mined for information. Big data is often dynamic and ensuring its reliability can be a challenge.
Data	Facts, measurements, recordings, records, or observations about the world, with a minimum of contextual interpretation. Data may be in any format or medium, including numbers, symbols, text, images, films, video, sound recordings, drawings, designs or other graphical representations.
Data curation	Data curation covers data selection, storage, preservation, annotation, provenance and other meta-data maintenance, and dissemination, and is needed to increase data interoperability. This includes the required hardware and software support for these tasks.
Data ethics	Data ethics is a new branch of ethics concerned with responsible use of data, algorithms and corresponding practices.
Data governance	The exercise of authority, control and shared decision making (planning, monitoring and enforcement) over the management of data assets.
Data-intensive science	Data-intensive science is considered to be the fourth paradigm of science after the three interrelated paradigms of empirical, theoretical, and computational science. It is seen as a data-driven, exploration-centred style of science, where IT infrastructures and software tools are heavily used to help scientists manage, analyse, and share data.
Data lifecycle	All the stages in the existence of digital information from creation to destruction. A lifecycle view is used to enable active management of the data objects and resource over time, thus maintaining accessibility and usability.
Data literacy	The ability to read, interpret, create and communicate data as information.
Data management	The activities of data policy development, data planning, data element standardisation, information management control, data synchronisation, data sharing, and database development, including practices that acquire, control, protect, deliver and enhance the value of data and information.
Data science	Data science encompasses the processes that deal with the extraction of meaning or knowledge from data.
Data stewardship	Data stewardship is the management and oversight of an organisation's data assets in order to provide professional users with high quality data that is easily accessible in a consistent manner.
Digitalisation	Digitalisation describes the way in which many domains of professional and social life are restructured around digital information and communication technologies.



Digital skills	Digital skills can be defined as a range of abilities to use digital devices, communication applications, and networks to access and manage information. In the context of science, these skills include an understanding of software, tools and data.
FAIR data	FAIR data is data which meet the principles of Findability, Accessibility, Interoperability, and Reusability.
Machine Learning	Machine Learning (ML) is a branch of Artificial Intelligence that is focused on developing systems that can learn from data. An ML algorithm is trained by learning from examples, which normally requires very large datasets.
Metadata	Literally, "data about data"; data that defines and describes the characteristics of other data, often using standardised formats.
Open data	Data that is accessible, freely shared. Open can be freely used, reused, built on and redistributed by anyone and may be subject to the requirement to attribute and share alike.
Open science	Open science promotes openness and early sharing of research ideas, papers, solutions, data and processes. Open science stresses the scientific, economic and societal benefits of increased and open scientific collaboration.
People-focussed skills	People-focussed skills differ from the system-focus of technical skills and include communication, teamwork collaboration, etc.
Research data	Data that are used as primary sources to support technical or scientific enquiry, research, scholarship, or artistic activity, and that are used as evidence in the research process and/or are commonly accepted in the research community as necessary to validate research findings and results.
Software	Software is a set of instructions, data or programs used to operate computers and execute specific tasks.
<b>Roles</b>	
Data analyst	This is someone who knows statistics. They may know programming, or they may be expert in spreadsheets. Either way, they can build models based on low-level data. Most importantly, they know which questions to ask of the data.
Data steward	A person responsible for planning and executing of all actions on digital data before, during and after a research project, with the aim of optimising the usability, reusability and reproducibility of the resulting data (Dutch Techcentre for Life Sciences, n.d. <sup>[68]</sup> ).
Data scientist	A practitioner of data science. It is a generic term that encompasses many fields of specialised expertise. In the current report, data analysts, data stewards and research software engineers are considered as sub-groups of data scientists. In certain contexts, data scientist is also sometimes used in a more limited ways that make it equivalent to either the data analyst or software engineer roles.
Research Software Engineer (RSE)	A growing number of people in academia combine expertise in programming with an intricate understanding of research. These RSEs may start of as researchers who spend time developing software to progress their research or they may come from a more conventional software-development background and are drawn to research by the challenge of using software to further research.
Research support professionals	In the context of digitalisation, these are the people who support scientific researchers conducting data-intensive science. They are not necessarily part of a research team and might be considered as service providers. This is a broad category that can include data stewards, RSEs, data managers, librarians and archivists.