

数理・データサイエンス・A I 教育プログラム 認定制度（応用基礎レベル）の 対象範囲の検討に向けた一考察

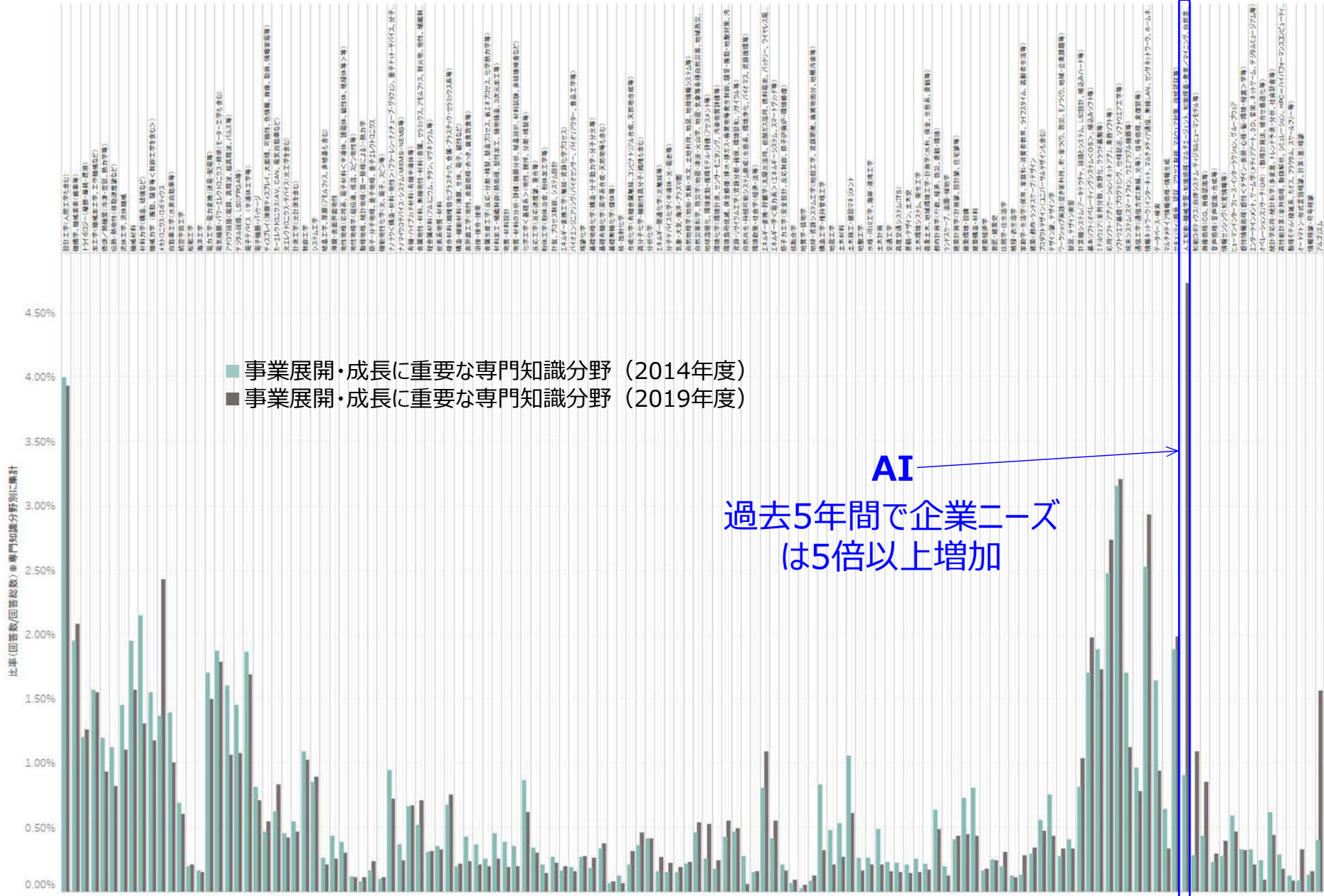
令和2年9月

内閣府政策統括官（科学技術・イノベーション担当）



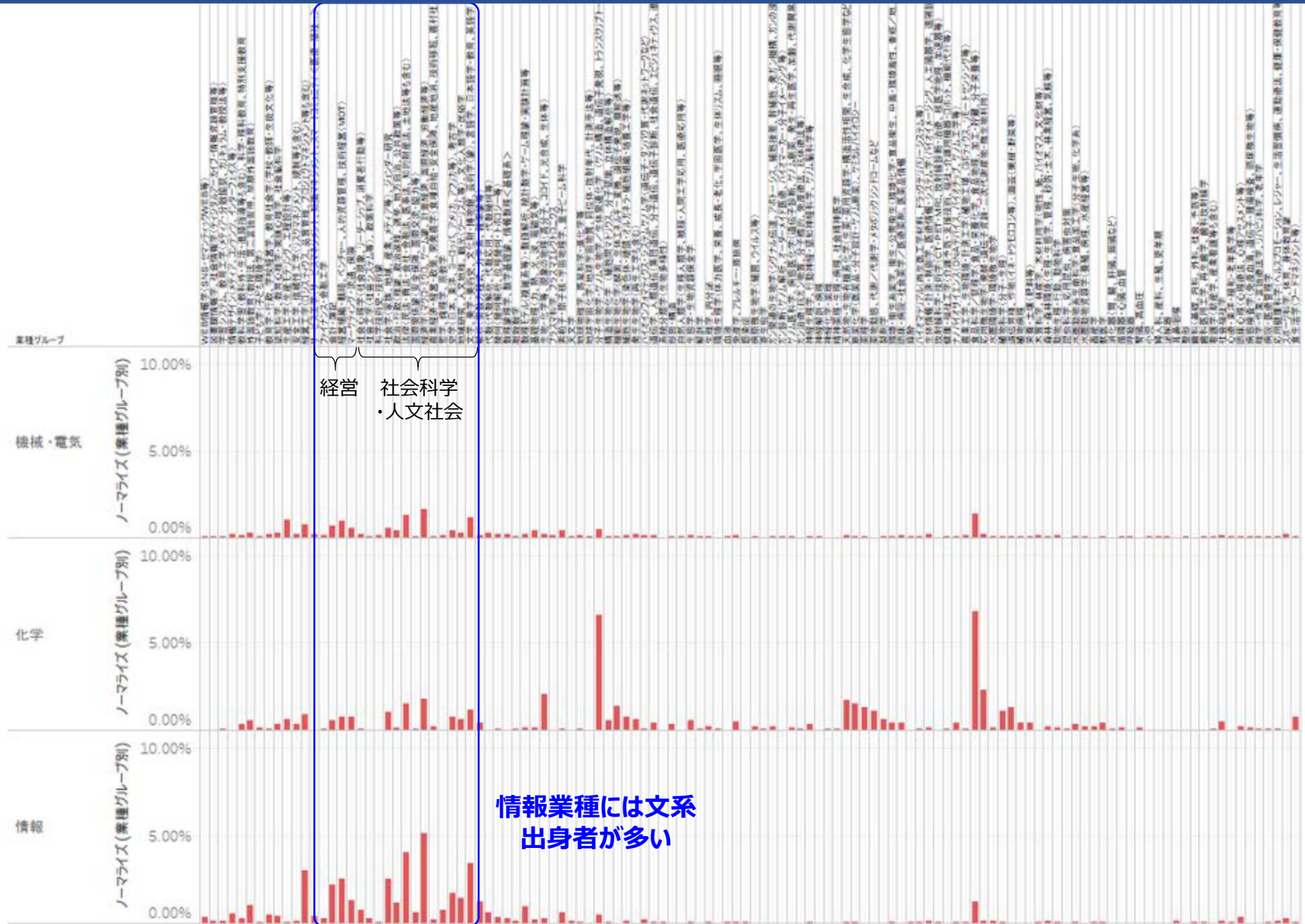
数理・データサイエンス・A I に関する人材確保の必要性 (既存データより)

企業における事業展開・成長に重要な専門知識分野の経年比較



出典：経済産業省 平成26年度（2014年度）産業技術調査事業「産業界と教育機関の人材の質的・量的需給ミスマッチ調査」、内閣府 平成31年度（2019年度）科学技術基礎調査等委託事業「産業界と教育機関の人材の質的・量的需給マッチング状況調査」より一部加工

出身研究室分野と就職先産業分野との関係 (2019年度)

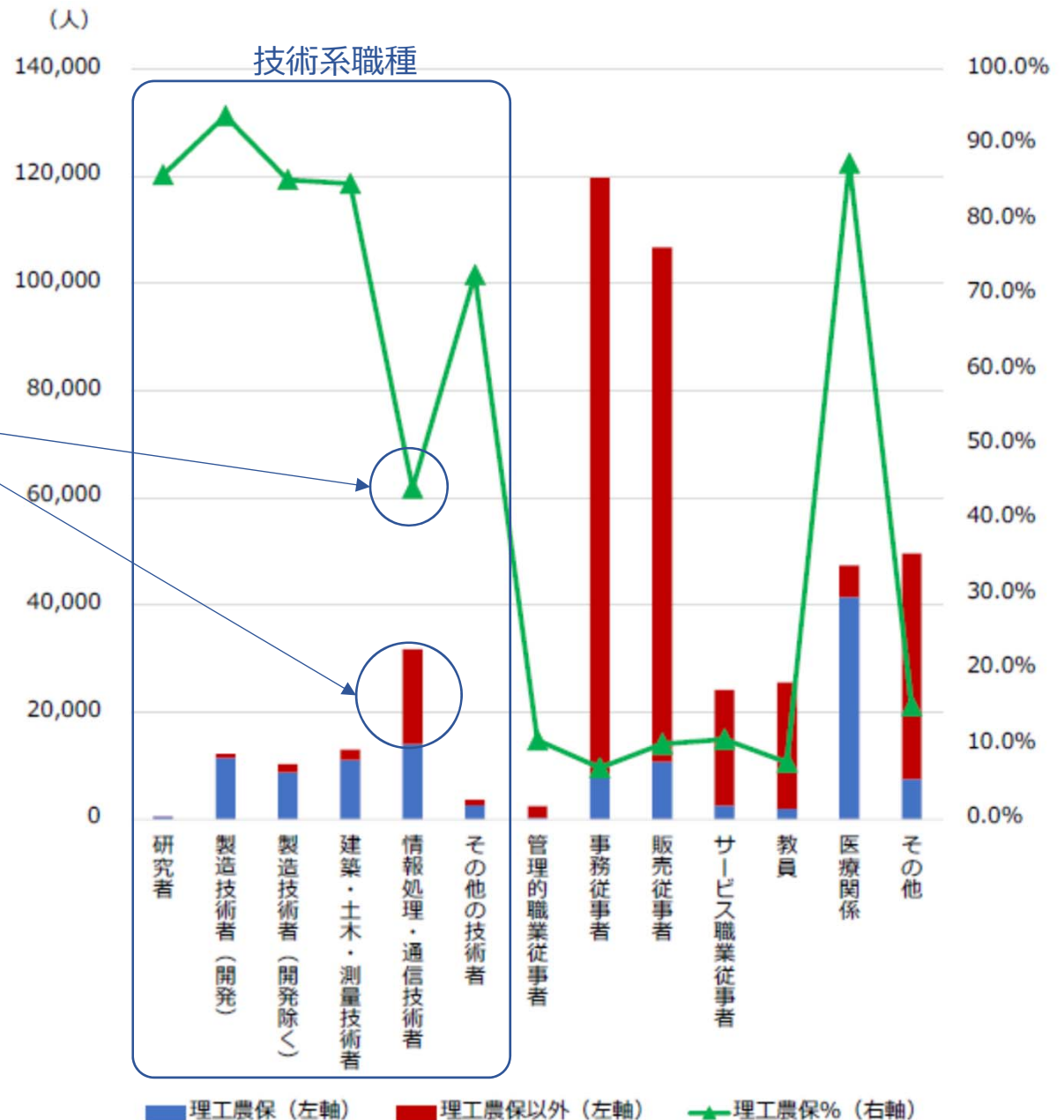


出典：内閣府 平成31年度（2019年度）科学技術基礎調査等委託事業「産業界と教育機関の人材の質的・量的需給マッチング状況調査」より一部加工

情報処理・通信技術者は文系出身者が半数超

情報産業でのエンジニア不足の中、文系学科から多くの人材を調達しているという実態

今後はどの分野もデータを活用した産業になると想像される。この場合、単純に情報分野の学部学科の定員を増やすよりも、各分野で情報を利用できるよう教育していく方向が良いだろう。



応用基礎レベルの対象範囲について

議論の参考とするため、昨年度実施した「AI時代に求められる人材育成プログラム」への応募プログラムをもとに、「応用基礎レベル」の対象範囲を検討。

集計対象とした情報ソース

- 2019年度に内閣府・文科省・経産省の3府省が全大学・高専を対象に行った「AI時代に求められる人材育成プログラム」への応募プログラム96件（73の大学・高専）
- 募集は次の3パターンで実施（ただし、複数パターンが重複した形の申請もあり）
 - a) **リテラシーレベル**：（55プログラム）
主に共通教育等で実施される、全ての学生が専攻に関わらず学ぶべき基礎的なプログラム
 - b) **応用基礎レベル（データサイエンス専門分野タイプ）**：（23プログラム）
主として数理・データサイエンス・AIを学習するプログラム
 - c) **応用基礎レベル（他分野との分野融合タイプ）**：（34プログラム）
主として専門分野を学びながら数理・データサイエンス・AIを学修するプログラム



このうち、パターン分離が可能な以下のデータのみを対象に、次頁の分類を従って科目数をカウント

検討対象データ

- 『リテラシーレベル』：
「a) リテラシーレベル」かつ「文理全体が対象」としているもの（29プログラム、297科目）
- 『応用基礎レベル』：
「c) 応用基礎レベル（他分野との分野融合タイプ）」のもの（29プログラム、532科目）

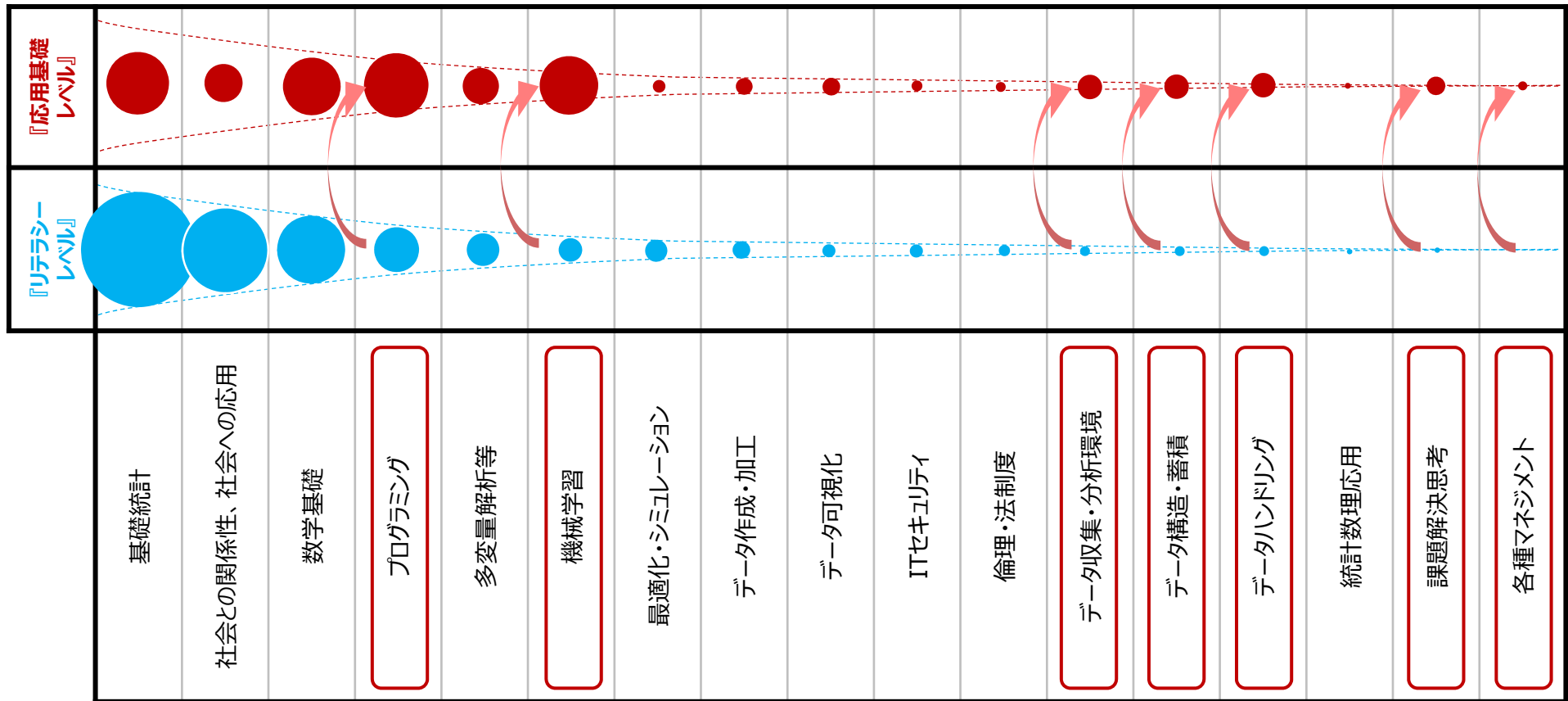
「分類」及び「定義」を、以下のように設定

	分類	定義
データサイエンス力	数学基礎	・線形代数学、微分積分学等。
	基礎統計	・記述統計、相関、回帰分析、推定・検定等の意味や計算。これらに必要な基本的な数学。
	データ作成・加工	・統計分析のための（標本）データの作成・加工処理に必要な知識（標本、実験計画、尺度、標準化、外れ値・異常値・欠損値等）と基本的な処理。
	データ可視化	・データの適切な集計・可視化（集計結果の図表作成）。
	多変量解析等	・各種の多変量解析等（主成分分析、因子分析、コレスポンデンス分析、多次元尺度構成法等、クラスタ分析、時系列分析、パス解析、共分散構造分析等）に関する基本的な知識と活用。
	機械学習	・機械学習の手法・データに関する知識。 ・それらに関連する言語・画像・音声処理に関する基本的な知識。
	統計数理応用	・ベイズ統計を中心とした、より進んだ統計数理全般の知識。
	最適化・シミュレーション	・各種の最適化・シミュレーション手法・アルゴリズムに関する基本的な知識。

	分類	定義
データエンジニアリング力	データ収集・分析環境	・データ収集・分析に必要な計算機環境の設計・管理・運用に関する基本的な知識。 ・データ収集に必要なプラットフォーム（iOS、Android、HEMS等）や通信機能に関する基本的な知識。それらを用いたデータ収集の実施。
	データ構造・蓄積	・データベース、DWHアプライアンス等に関する基本的な知識と活用。
	データハンドリング	・データのクレンジング等の加工、ソート・結合・抽出・集計・四則演算等の操作。 ・種々のフォーマットでのデータのエクスポート、Web API等を用いたデータのインポート。 ・BIツールによるグラフ・レポートの作成。
	プログラミング	・データ処理（抽出・加工・分析等）に必要なプログラミング、APIを使用したプログラミング、SQLの記述・実行。
	ITセキュリティ	・セキュリティの3要素（機密性、可用性、完全性）、プライバシー・個人情報、匿名化、攻撃・防御手法、暗号化に関する知識と活用。
	ビジネス力	倫理・法制度
課題解決思考		・課題の定義・整理、分析枠組みを決定するための論理的思考方法、一般的なフレームワークに関する知識と活用。
社会との関係性、社会への応用		・社会（業務等も含む）との関係の中での、適切なデータ収集・分析、解釈、解決策の提案の実施。
各種マネジメント		・プロジェクトマネジメント、組織マネジメント、企画・提案、人材育成に関する知識・スキル。

（出所）データサイエンティスト協会が定義したスキルカテゴリ（<https://www.datascientist.or.jp/common/docs/skillcheck.pdf>）を参考に、2019年度に三菱総合研究所が作成したもの

●のサイズ：科目比率（レベル毎に科目総数にて規格化）

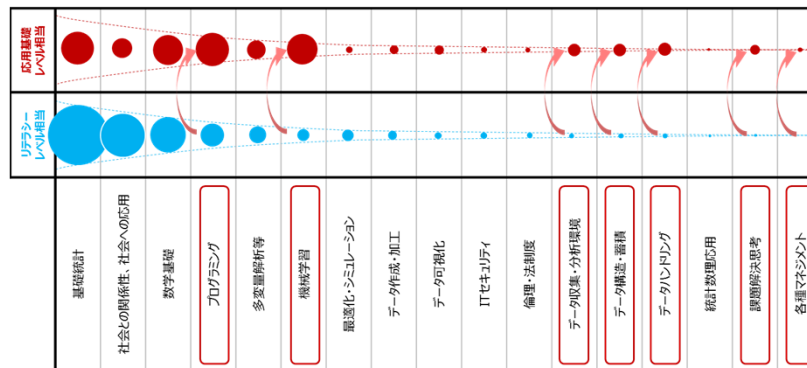


* データの集計は事務局にて実施したものの

集計結果から何か見えるものは無いか？

- ✓ 『リテラシーレベル』、『応用基礎レベル』ともに、全ての分類をカバー
- ✓ □ の分類項目は、『応用基礎レベル』において、『リテラシーレベル』に比べ、より内容が充実した項目であると
考えられるのではないかと。

「モデルカリキュラム（リテラシー）」との関係を、みてみると・・・



『応用基礎レベル』

- ・ 課題解決思考
- ・ 各種マネジメント

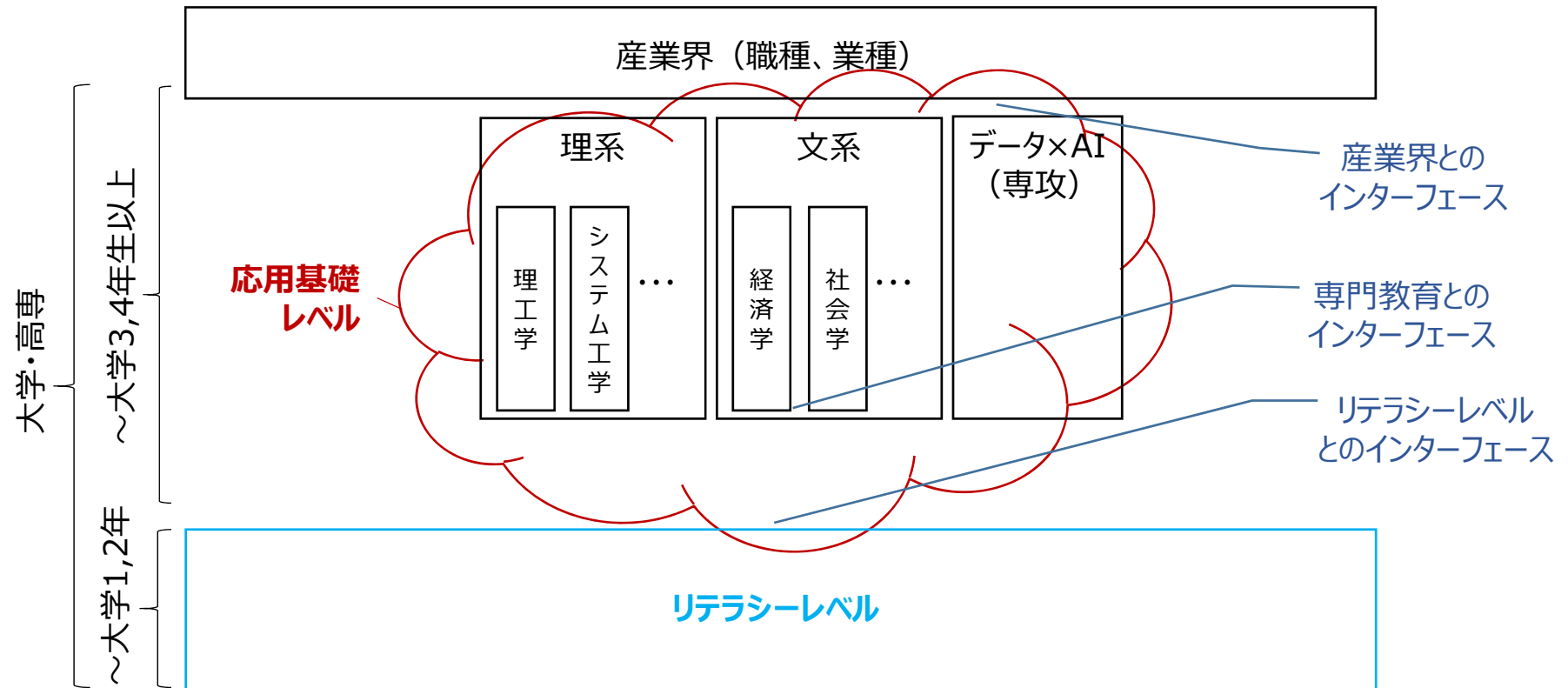
4. オプション

4-1. 統計および数理基礎	4-2. アルゴリズム基礎
4-3. データ構造とプログラミング基礎	4-4. 時系列データ解析
4-5. テキスト解析	4-6. 画像解析
4-7. データハンドリング	4-8. データ活用実践（教師あり学習）
4-9. データ活用実践（教師なし学習）	

データ構造・蓄積+プログラミング
機械学習
データハンドリング
機械学習+データ収集・分析環境

データハンドリング
機械学習
機械学習+データ収集・分析環境

心得	3. データ・AI利活用における留意事項	
	3-1. データ・AIを扱う上での留意事項	3-2. データを守る上での留意事項
基礎	2. データリテラシー	
	2-1. データを読む	2-2. データを説明する
	2-3. データを扱う	
導入	1. 社会におけるデータ・AI利活用	
	1-1. 社会で起きている変化	1-2. 社会で活用されているデータ
	1-3. データ・AIの活用領域	1-4. データ・AI利活用のための技術
	1-5. データ・AI利活用の現場	1-6. データ・AI利活用の最新動向



- 本検討で用いた分類をベースに「応用基礎レベル」の対象範囲を検討することは妥当か。
- 応用基礎レベルのボトム部分（リテラシーレベルとのインターフェース部）は、リテラシーレベルのモデルカリキュラムにて「オプション」設定された学修内容として捉えることは妥当か。
一方で、大学ごとに「リテラシーレベル」のカリキュラムにおいて修学レベルに差が生じていると考えられるが、これを「応用基礎レベル」のカリキュラムの認定において、どのように扱うか。
- 応用基礎レベルのトップ部分（産業界とのインターフェース部）では、各々の大学や学部が目指す産業界ターゲットからのニーズと、期待される輩出人材の品質とのマッチング性を、評価・認定していくということではどうか。