

平成 25 年度
データサイエンティスト育成研究会
< 報告書 >

平成 26 年 3 月 26 日

一般社団法人コンピュータソフトウェア協会
データサイエンティスト育成研究会

目次

はじめに	2
1. 目的	3
2. 方法	4
3. 結果	5
1) データサイエンティストの定義	6
2) データサイエンティストの位置付け	7
3) データサイエンティストの職務	8
4) 職務とタスクの定義	9
5) 職務に求められる知識とスキル	10
6) データサイエンティスト育成プログラムの概要	12
7) 検討を要する課題	13
8) データサイエンティストの事例	14
4. 考察	15
おわりに	16

はじめに

データサイエンティスト育成研究会は、一般社団法人コンピュータソフトウェア協会（以降、CSAJ と略す）の研究会として6カ月の予定で2013年10月15日（キックオフ）にスタートいたしました。その後、情報サービス業界視点で「データサイエンティスト」を定義するところから活動を始め、月に1度のペースで研究会を開催いたしました。キックオフを含め5回の研究会によって検討、議論した結果として導き出した「データサイエンティスト」育成のベースとなる定義や枠組み考え方をご報告させていただきます。

研究会の開催、推進にあたりましてはご参加いただいた会員各位に多大なご支援ご協力を賜りました。この場をお借りして深謝いたします。

データサイエンティスト育成研究会 主査 昆 凡子

1. 目的

当研究会は、2013年度研究会設置申請において以下の目的を掲げ、設置された。

大量データの有効活用が大きなテーマとなっている昨今、これを操る「データサイエンティスト」の育成が急務になっている。「データサイエンティスト」は、企業や政府機関、自治体の持っている大量データを分析し、企業の業績を飛躍的に伸ばす役割を担いつつある。そこで、「データサイエンティスト」の定義、資格要素など、国内においてまだまだあいまいな部分を調査研究し、その結果を踏まえ、CSAとして、国に提言していけるような内容作りを目指す。

その後、初期段階で検討し、以下を目的とすることで研究会として合意した。



情報システム（情報サービス）業が、ビジネスにおいてイニシアティブを取るためのデータサイエンティスト像を明確にし、その育成について概要を定める。

特に、情報システム（情報サービス）業がイニシアティブをとるためと強調した背景には、情報サービス業が従来どちらかというと受身でサービスを提供してきていることを打破したいという研究会メンバーの積極的な姿勢がある。ビッグデータの活用など、経営や事業に能動的に参画する方向を提示する良い機会という認識である。

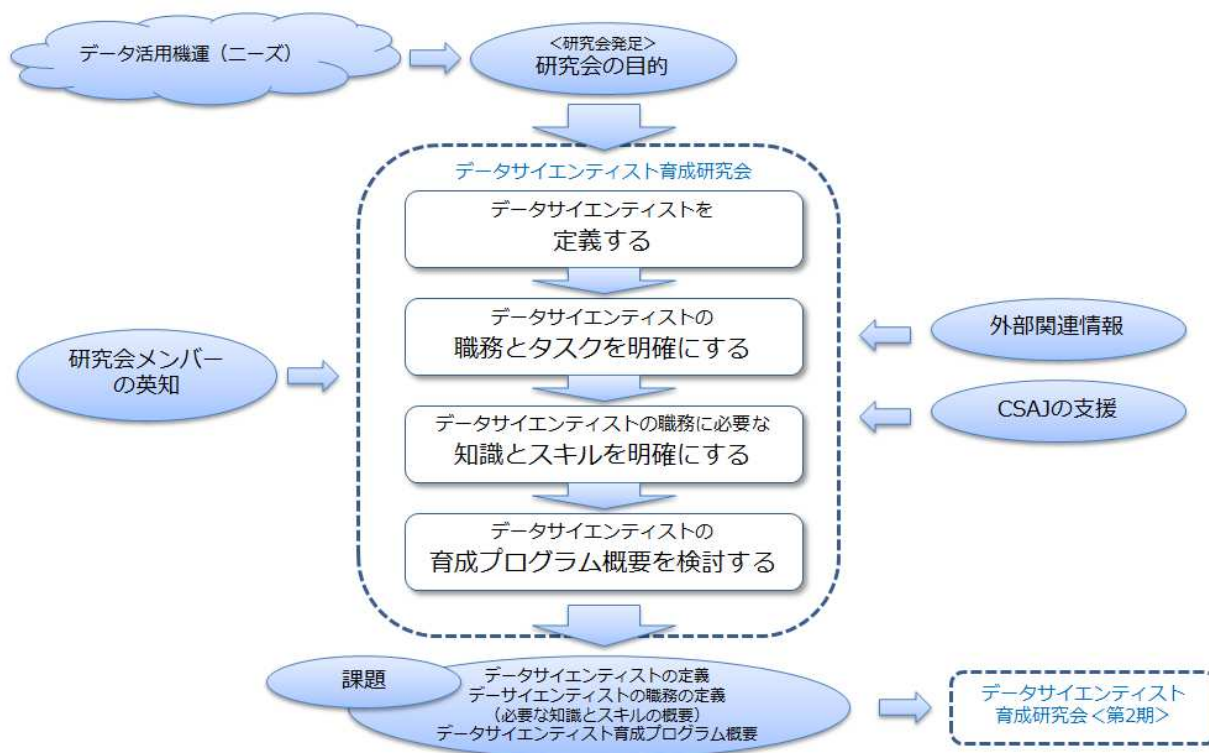
2. 方法

研究会活動は、2013年10月以降第2火曜日を中心にCSAJ会議室を利用して行った。毎回テーマを設定し、参加メンバーによるディスカッション中心の活動を行い、随時テーマに合った知識を有するオブザーバ、講師を招聘した。開催日とその概要は以下の通りである。

第2回～第4回は、メンバー各位に宿題作業をお願いし、提出いただいた内容をコアメンバーがとりまとめ、その資料を研究会にて討議し、合意形成してゆく方法を採用した。

タイトル	開催日	概要	参加者数
第1回 研究会	平成25年10月15日（火）	キックオフ 活動計画の確認	12名
第2回 研究会	平成25年11月19日（火）	データサイエンティストの定義	15名
第3回 研究会	平成25年12月17日（火）	データサイエンティストの職務・職責の考え方およびタスク分析	15名
第4回 研究会	平成26年1月14日（火）	職務に必要な知識・スキルについての検討と研修の枠組み作り	15名
第5回 研究会	平成26年2月12日（水）	課題の確認、事例の補足および報告会用資料の作成	14名
第6回 研究会	平成26年3月26日（水）	研究会報告と講師による講演	29名

研究会の目的を達成するための手順を下図のように設定し、順次研究活動を進めた。



3. 結果

月に一度の研究会を半年間実施し、結果として以下に列挙する内容を得ることができた。それぞれの内容はまだまだ概要であったり骨格であり、検討議論の余地がある。また、研究結果をより具体的なものに落とし込むことで活用範囲が広がる。今後の研究会活動に期待するものである。

- データサイエンティストの定義
- データサイエンティストの位置付け
- 職務とタスクの定義
- 職務に求められる知識とスキル
- データサイエンティスト育成プログラムの概要
- 検討を要する課題
- データサイエンティストの事例

研究結果内容はすべて今後の研究会活動のインプットとなる。

3. 結果

1) データサイエンティストの定義

研究会としてのデータサイエンティストを以下のように定義した。

データサイエンティストとは、

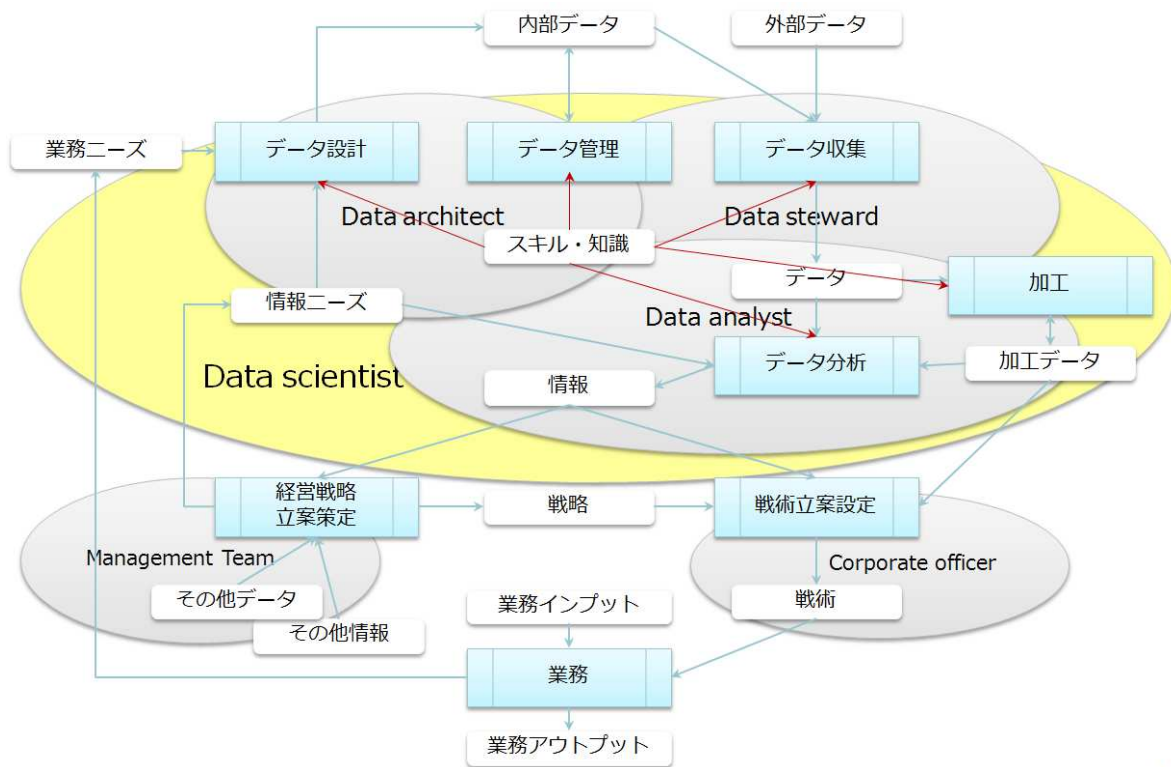
**「ビジネスにおいて、競争優位性をもたらすために、データの
収集・加工・分析に優れた専門性を発揮し、知見を引き出す人材」**
である。

現時点での研究会メンバーが合意した定義であり、今後、研究会活動を進めて行く中で微調整する可能性はある。

3. 結果

2) データサイエンティストの位置付け

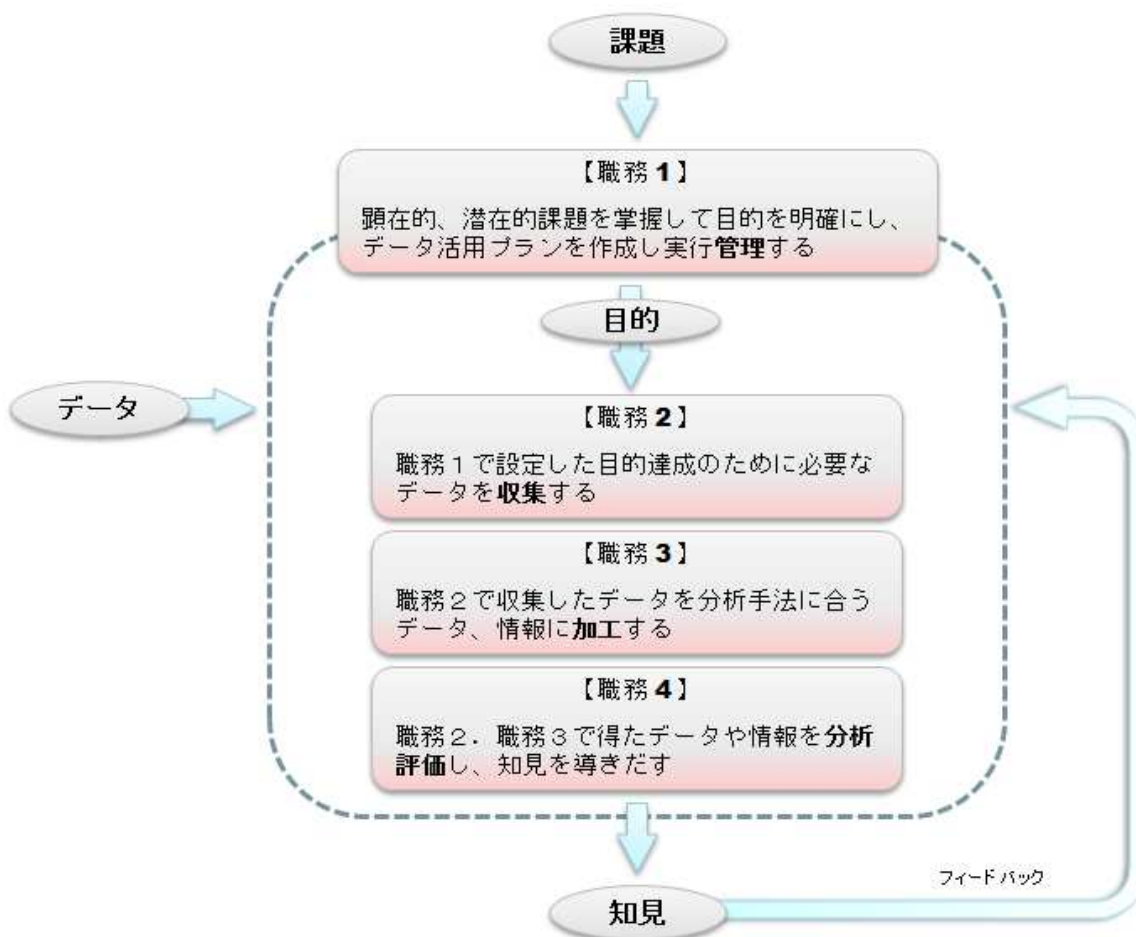
研究会として定義したデータサイエンティストは、これまで様々な局面で定義されて来たデータアーキテクト、データスチュアード、データアナリストなどを包括するような人材像となった。下図は、データを活用する機能情報関連を研究会として示したものである。この図をベースとしてデータサイエンティストがカバーする領域を検討し、職務の定義作業を進めた。



3. 結果

3) データサイエンティストの職務

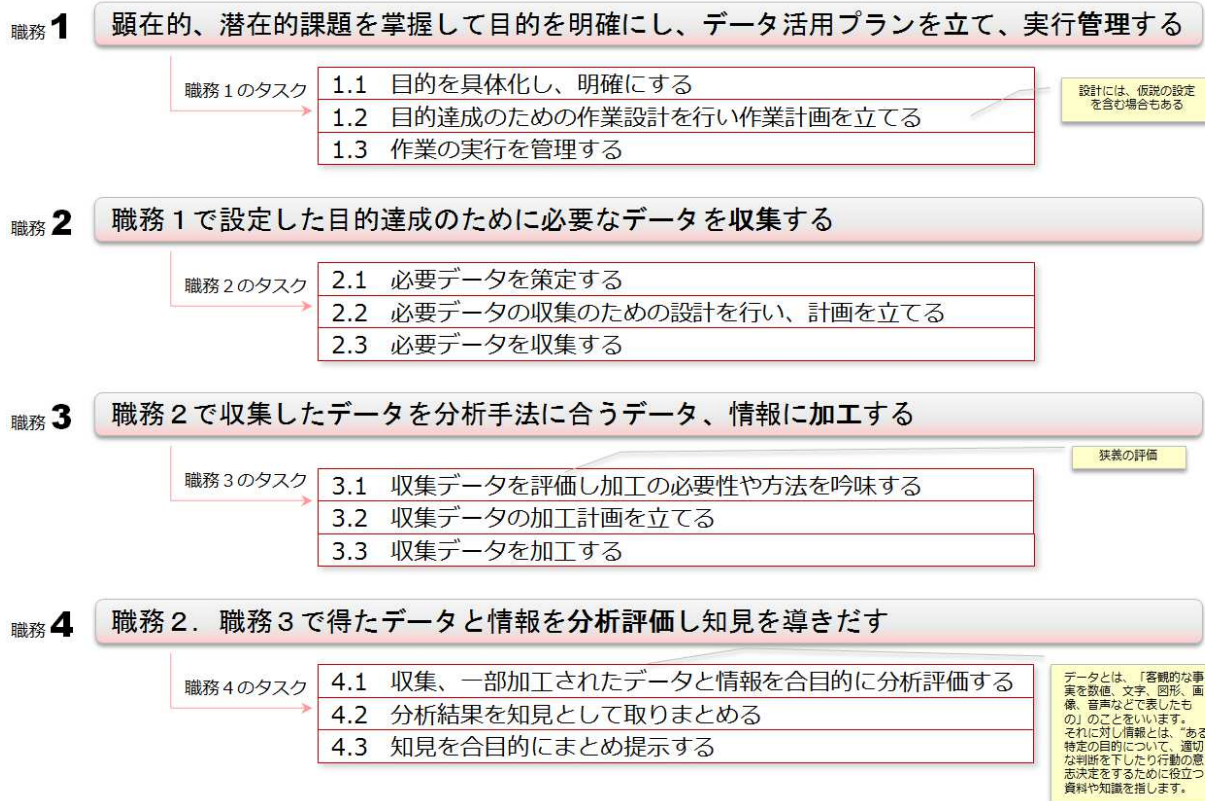
職務とは担当する任務（責任をもって取り扱うつとめ）である。データサイエンティストの職務は、データ活用により解決が見込まれる課題に対し、ビジネス視点で価値ある知見を導き出すことである。（研究会の定義より）その観点でインプット・プロセス・アウトプットを機能的に分解するというシステム設計的アプローチを行い、下図に示すような職務1～4を定義した。



3. 結果

4) 職務とタスクの定義

3) で定義した職務をブレイクダウンし、担当する任務（責任をもって取り扱うつとめ）としてのタスクを以下の通り、洗い出した。タスクへの分解も定義から職務への落とし込みと同様システム設計的アプローチで行った。



3. 結果

5) 職務に求められる知識とスキル <1>

職務とタスクの定義を基に職務単位に必要な知識をスキルの洗い出しを行った結果を以下に列挙する。知識やスキルの表現粒度は大きい育成プログラム概要を設定する上で必要最低限の検討を行った。今後、育成プログラムのガイドラインを作成する過程で再検討の必要性があると認識している。

職務 1 顕在的、潜在的課題を掌握して目的を明確にし、データ活用プランを立て、実行管理する

職務1のタスク	1.1 目的を具体化し、明確にする
	1.2 目的達成のための作業設計を行い作業計画を立てる
	1.3 作業の実行を管理する

■ 知識

- 1K1 経営の知識
- 1K2 業界の知識
- 1K3 業務の知識
- 1K4 マーケティングの知識
- 1K5 ITの知識
- 1K6 分析のための知識
- 1K7 プロジェクト管理の知識

■ スキル

- 1S1 戦略立案スキル
- 1S2 IT活用スキル
- 1S3 プロジェクト管理スキル
- 1S4 コミュニケーションスキル

職務 2 職務1で設定した目的達成のために必要なデータを収集する

職務2のタスク	2.1 必要データを策定する
	2.2 必要データの収集のための設計を行い、計画を立てる
	2.3 必要データを収集する

■ 知識

- 2K1 業務の知識
- 2K2 ITの知識
- 2K3 データの知識

■ スキル

- 2S1 データの収集/整理スキル
- 2S2 大規模データ、特殊データの処理基盤スキル
- 2S3 ITスキル

3. 結果

5) 職務に求められる知識とスキル <2>

職務 3 職務 2 で収集したデータを分析手法に合うデータ、情報に加工する

職務 3 のタスク	3.1 収集データを評価し加工の必要性や方法を吟味する
	3.2 収集データの加工計画を立てる
	3.2 収集データを加工する

■ 知識

- 3K1 業務の知識
- 3K2 ITの知識
- 3K3 データの知識

■ スキル

- 3S1 データの可視化スキル
- 3S2 大規模データ、特殊データの処理基盤スキル
- 3S3 ITスキル

職務 4 職務 2、職務 3 で得たデータと情報を分析評価し知見を導き出す

職務 4 のタスク	4.1 収集、一部加工されたデータと情報を合目的に分析評価する
	4.2 分析結果を知見として取りまとめる
	4.3 知見を合目的にまとめ提示する

■ 知識

- 4K1 経営の知識
- 4K2 業界の知識
- 4K3 業務の知識
- 4K4 マーケティングの知識
- 4K5 ITの知識
- 4K6 分析のための知識

■ スキル

- 4S1 戦略立案スキル
- 4S2 プレゼンテーションスキル
- 4S3 IT活用スキル
- 4S4 分析スキル

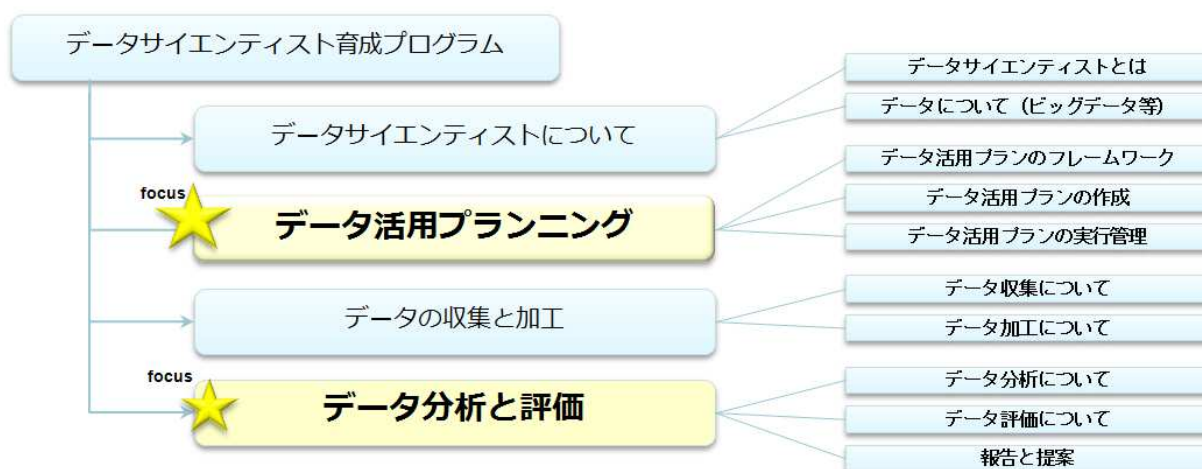
3. 結果

6) データサイエンティスト育成プログラムの概要

本研究会が作成するデータサイエンティスト育成プログラムは、研究会が設定したデータサイエンティストの定義、職務とタスクの定義、職務に必要な知識とスキルから次のように定義した。

「情報サービス関連に携わる人材に対し、データサイエンティスト（当該研究会の定義）として、企業経営や事業に対し積極的にアプローチして利益貢献できるために必要な基礎知識を提供し実践できるようにトレーニングを実施するものである。」

上記定義を基に検討し、下図に示す育成プログラムの構成概要を設定した。特にデータサイエンティストは「データ活用プランニング」を行う人材ということに着目している。育成プログラムのガイドラインの骨格になると想定している。



3. 結果

7) 検討を要する課題

研究会活動の中で、解決していないまたは検討が不十分な内容については課題とした。以下に課題を列挙する。これら課題は、本研究活動の成果物の一部として今後の活動に引き継ぐことを望む。

1. データサイエンティストの活用対象データはビッグデータだけではない

データサイエンティストにスポットライトが当てられた背景にビッグデータがあるためビッグデータを分析し活用する人材として捉えられることが多い。当該研究会において、ビッグデータは重要な活用対象であるが、その他のデータも活用対象としてきた。今後、活用対象データについては明示して種々対応することが必要かもしれない。十分議論されていないので今後の課題としたい。

2. データサイエンティストに求められる知識およびスキルは広範である

データサイエンティストという人材を定義したが、その知識およびスキルは広範となり、それらすべてを一人で網羅し、獲得することは可能かもしれないが現実的ではない。今後の検討の中で組織対応するなどを検討し、その管理と一部の役割をデータサイエンティストに求めるなどその辺り明確にすることを今後の課題としたい。

3. 仮説を立ててデータ活用のアプローチを行うことを明確には組み込めていない

仮説設定は、常に行われるものではないが、有効なアプローチの手段である。その点を考慮してデータ活用プラン作成のアプローチについて検討することを今後の課題としたい。

4. 知見を引き出した後のフィードバックを含めた対応についてはどう扱うか

知見を引き出した後工程について十分議論されていないので今後の課題としたい。

3. 結果

8) データサイエンティストの事例

データサイエンティストは、定義したもののイメージしにくいとの意見があり、当研究会ではデータサイエンティストの職務事例を収集して提示することとした。過去の実績や想定されるものを事例として整理してゆく。

今回は、研究メンバーから事例案をそれぞれ提示していただいている状況で、十数件の事例候補が挙がっている。事例収集整理は、今後も継続し最終的に事例集としてまとめる予定である。

【事例案 1】 広報戦略策定支援（想定）

パブリシティや CSR 活動に対する報道状況や消費者の声をネット上も含め収集分析し、次なる広報戦略の策定を行う。ブログやツイッター、2ちゃんねるなどのデータも対象となる。

【事例案 2】 リスク管理（実例）

お客様相談室および営業部門で収集した問い合わせやクレームデータの関連性やリスクを分析し、クライシスの発生を抑止するしくみを構築。データの内容および関連性を、設定したルールに基づき分析しリスクレベルにより警告メール を危機管理部門に自動通知するシステムを開発。

4. 考察

当初、データサイエンティストという名称に対し研究会として定義するのか、それともおぼろげながらではあるが求めている人材像を明らかにし名称を与えるのか、研究会として迷った。しかし世の中は、まだデータサイエンティストの定義が明確になされていない状況にあったため、データサイエンティストという名称の理解はまちまちではあるが、その知名度の高さからデータサイエンティストという名称に対し定義を行うこととした。定義を行うに際しては、情報サービス業がビジネスにおいてイニシアティブをとるための人材という視点を明確に打ち出した。このことにより、定義されたものはCSAJの研究会として、その目的を含め相応しいものとなっていると研究会としては思っている。

データサイエンティストの定義から育成プログラムの概要作成に持ち込んだ手順は、研究会の時間的な制約などから演繹的なアプローチを選択した結果である。

データサイエンティストの定義から次への落とし込みは、ビジネスにおける人材ということからすれば職務の明確化となる。職務は単純明快に、その「人材が担当する任務」とした。職務が明確化したので次なる落とし込みはタスクとした。ここでタスクの認識も多様化しているため「課せられた任務」として落とし込みを行った。職務およびタスクの落とし込みの際にも、常に目的を念頭においた。

今回の研究テーマは、データサイエンティスト育成であることを改めて確認し、次なる検討作業として、職務とそのタスクに必要な知識とスキルを整理することにとりかかった。その際、タスク毎の知識およびスキルを洗い出すことも考えたが、重複も多くかえって解りづらいと判断し職務に対して知識とスキルを整理した。今回は、育成プログラム概要を設定するために必要なレベルの整理に留まっている。

今回の最終段階としデータサイエンティストの定義、職務とそのタスクの定義、職務に必要な知識とスキルの概要をインプットとし育成プログラム概要についての検討を行った。その際、大きな気づきがあった。それは、検討してきたデータサイエンティストを別の視点で表現すれば、シンプルにデータをビジネスに活用するためのプランを立てられる人材だということに対する気づきである。そこで次なる研究に繋げるための方針として、育成プログラムを構築する上で「データ活用プランニング」に焦点を当てることとした。このことは研究会が定義したデータサイエンティストの育成を短期間で費用対効果高く行う可能性を開いたものと捉えている。

おわりに

本研究会は、メンバーの総意をベースとし、すべてメンバーの経験と英知により結果を導き出した。その結果、既成概念に囚われることなく自由な発想で議論を行うことができた。今後、より具体的な落としこみを行うにあたっては、多面的な摺り合わせを行うため、文献や関連情報が必要になるとことを想定している。

研究会メンバー（敬称略）

（主査）

昆 凡子 株式会社シムコス

（メンバー：五十音順）

青木 明彦 株式会社算法
井口 一世 株式会社井口一世
宇野 和彦 株式会社スキルメイト
小倉 健一 日本システム開発株式会社
鹿島 史恵 株式会社井口一世
加藤 秀治 リコー I T ソリューションズ株式会社
木村 雅晴 株式会社シムコス
国井 勝則 オデッセイヒューマンシステム株式会社
昆 充生 株式会社シムコス
新志 有裕 株式会社MM総研
武田 大輔 リコージャパン株式会社
武智 秀昭 株式会社パスカリア
谷川 智久 日本システム開発株式会社
中島 洋 株式会社MM総研
中原 崇 株式会社パスカリア
西村 紀彦 株式会社ミック
根岸 邦彦 根岸事務所
本田 義久 リコージャパン株式会社
松倉 泉 LLPハーネス有限責任組合
丸山 啓之 株式会社富士通マーケティング
山田 晃嗣 株式会社シムコス

（オブザーバ）

丹下 博詞 株式会社アグラ

（事務局）

小林 哲 一般社団法人コンピュータソフトウェア協会

平成 25 年度データサイエンティスト育成研究会 報告書
平成 26 年 3 月 26 日 第 1 版

一般社団法人コンピュータソフトウェア協会 (CSAJ)
〒107-0052
東京都港区赤坂 1-3-6 赤坂グレースビル 4F
TEL : 03-3560-8440 FAX : 03-3560-8441 <http://www.csaj.jp.jp/>