

기업에서의 데이터 사이언티스트의 더 큰 활용을 위하여

미야자키 치히로(宮崎 地洋)
DX컨설팅부 부주임 컨설턴트

I. 시작하며

2019년에 미국의 대형 구인정보 사이트가 발표한 'The Best Job in America 2019'에 따르면¹⁾, 2019년에 미국에서 가장 매력적인 직업은 '마케팅 매니저'와 '프로덕트 매니저'를 누르고 '데이터 사이언티스트'로 나타났다. 이 랭킹은 연수입, 일의 만족도, 구인 수 등을 종합하여 판단한 것으로, 이 사이트에서는 2019년 미국에서의 '데이터 사이언티스트'의 연수입 중앙값을 \$108,000로 밝히고 있다. 이 정도까지 '데이터 사이언티스트'라는 직업의 중요성이 커진 배경에는 GAFFA²⁾를 비롯한 테크놀로지 기업의 급속한 발전과 이에 따라 세상 속의 온갖 제품과 우리들의 행동 등이 데이터화됨에 따라 축적된 데이터를 분석하고 어떻게 비즈니스상의 가치로 전환해 갈 것인가가 문제시되기 시작했기 때문이다. 게다가 데이터 사이언스에 관한 기술영역에 있어 AI와 기기학습, 딥러닝이라는 단어로 대표되는 연구가 최근 십수년간 급속도로 진보하고 있는 것도 그 배경 중 한 요인이라고 생각된다.

일본에서도 요즘 데이터 사이언스와 그에 관련된 AI, 빅데이터와 같은 키워드에 대한 주목도는 매우 높아졌다. 하지만, 단어들의

유명세(Buzzword)에 비해, 실제 비즈니스와 기업 활동에 있어 데이터 사이언스나 데이터 사이언티스트를 유효하게 활용하고 있는 사례는 아직 많지 않다.

본고에서는 데이터 사이언티스트의 필요 스킬에 대한 정의와 기업에서 이들이 필요하게 된 배경에 대해 언급하면서, 왜 일본에서 데이터 사이언티스트의 활용이 충분히 진전되지 않는지에 대해 분석하고, 향후 본격적인 데이터 사이언티스트의 활용을 추진하기 위해 어떠한 연구가 필요한지를 제안하고자 한다.

II. 데이터 사이언티스트의 필요 스킬과 이들이 필요하게 된 배경

1. 데이터 사이언티스트의 필요 스킬

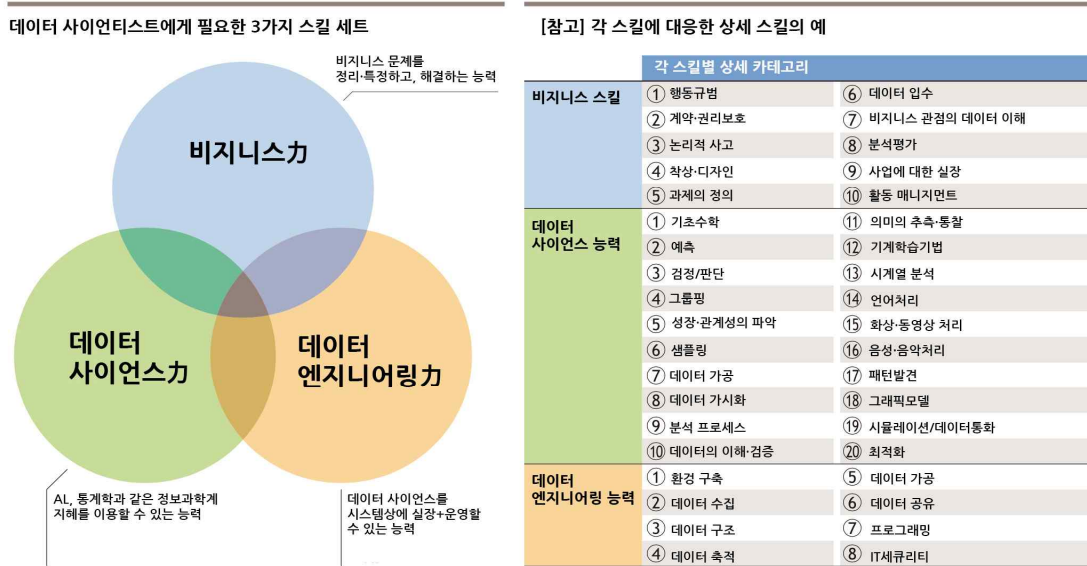
데이터 사이언티스트란 구체적으로 어떤 역할 및 업무를 담당하는 인재인 것일까? 2014년에 '(사)데이터 사이언티스트 협회'가 데이터 사이언티스트에게 필요한 스킬에 대해 발표한 자료를 살펴보면, 데이터 사이언티스트는 '비즈니스력', '데이터 사이언스력', '데이터 엔지니어링력'의 3가지 스킬이 요구되고 있다[그림 1].

각각에 대해 살펴보면, '비즈니스력'은 "(현상에서의) 과제의 배경을 이해한 뒤, 비즈니스 과제를 정리하고 해결하는 능력"이라고 정의할 수 있으며, 이는 현상에서 비즈니스상의 과제를 정의하는 능력과 데이터 사이언스를 현장·실무에 적용시키는 능

1) 미국 대형 구인정보 사이트 Glassdoor가 2019년 1월에 발표한 직업 랭킹. 1위는 '데이터 사이언티스트', 2위는 '수간호사', 3위는 '마케팅 매니저', 4위는 '작업치료사', 5위는 '프로덕트 매니저'.

2) Google, Apple, Facebook, Amazon

[그림 1] 데이터 사이언티스트에게 필요한 3가지 스킬에 대한 정의



출처)일반사단법인 데이터 사이언티스트 협회 '데이터 사이언티스트 스킬 체크 리스트'에서 NRI 작성

력에 더해 프로젝트 매니지먼트에 관한 스킬(계약 및 권리보호, 데이터 윤리 등을 포함) 등이 포함된 스킬이라 할 수 있다. '데이터 사이언스력'은 “정보처리, 인공지능, 통계학 등의 정보과학계의 지혜를 이해하고 사용하는 능력”이라고 정의할 수 있으며, 이는 기초수학 및 통계학에 관한 이해에 더해, 데이터의 이해/가시화 및 기계학습을 비롯한 각종 알고리즘/모델링에 관한 이해 및 프로그래밍에서의 실행력 등이 포함된 스킬이다. 마지막으로 '데이터 엔지니어링력'은 “데이터 사이언스를 의미 있는 형태로 사용할 수 있도록 만들고 실행 및 운용할 수 있도록 하는 능력”이라고 정의할 수 있으며, 이는 기초적인 프로그래밍 스킬을 바탕으로 현장 실행을 지원하는 시스템 환경 구축 스킬 및 데이터와 시스템의 보안에 관한 스킬 등이다. 데이터 사이언티스트는 이들 3가지 스킬 중 어느 한 가지도 빠져서는 안 된다고 여겨지고 있으며, 동시에 과제 해결 단계에 따라 중심이 되는 스킬이 변화해 가는 것이라고 하고 있다[그림 2]. 그러나 이러한 스킬에 관한 정의와 가이드라인은 여전히 개선 중에 있으며, 또한 정의된 '비즈니스력', '데이터 사이언스력', '데이터 엔지니

어링력'의 3가지를 갖춘 인재 또는 조직을 어떻게 구성할 것인지, 어떤 기업 안에서 3가지 능력을 합쳐서 발휘하도록 할 것인지에 대해서는 논의가 진행 중인 상황이다.

2. 데이터 사이언티스트가 필요해진 배경

데이터 사이언티스트가 필요해진 배경을 고찰해 보면, 크게 3가지의 이유가 있다고 생각된다. 첫 번째는 정보통신 기술의 발전이다. 데이터 사이언스 영역에 관한 큰 기술적 진보로서는 ①네트워크·통신기술의 발전에 따른 Web서비스와 SNS의 보급에 따라 세상 속의 데이터가 폭발적으로 증가한 점, ②다양한 분석 툴의 등장에 따라 분석 자체의 장벽이 낮아진 점, 그리고 ③방대한 데이터 처리에 필요한 하드웨어 기술의 가격이 저렴해진 점의 3가지를 꼽을 수 있다. ①에 있어서는 네트워크·통신기술의 발전에 따라 Web 서비스와 SNS, IoT 센서 등이 급속하게 보급됨으로서 낱알이 인터넷 상에서 주고받는 데이터의 총량 그 자체가 폭발적으로 증가하게 되었다. 이에 따라 이전에는 일부 기업 및 연구소·대학 등에서만 데

[그림 2] 데이터 사이언티스트(DS)를 활용한 문제해결 프로세스에 따른 각 스킬의 필요도 · 발휘도



출처) 일반사단법인 데이터 사이언티스트 협회 '데이터 사이언티스트 스킬 체크 리스크'에서 NRI 작성

이터 사이언스 분야의 데이터를 처리할 수 있던 것이 누구라도 대량의 데이터에 간단히 접근하여 분석할 수 있게 됨으로서 데이터 사이언스의 분석 대상이 크게 확대되었다고 볼 수 있다. 또한 ②에 있어서는 지금까지 분석을 진행하기 위해서는 통계학적 지식과 고도의 프로그래밍 언어 능력이 필수였지만, 현재는 초보적인 프로그래밍 지식만 있으면 분석에 필요한 알고리즘 자체는 다른 유저가 작성한 것을 무상으로 활용할 수 있는 시스템이 갖추어져 있어 누구라도 손쉽게 고도의 분석을 시행할 수 있게 되었다. 또한 유상 툴일 경우에는 Tableau³⁾와 DataRobot⁴⁾ 등과 같이 프로그래밍조차 필요 없이 직관적인 조작으로 데이터의 분석 및 가시화를 수행할 수 있게 되어 데이터 분석의 장벽도 매우 낮아지고 있다. 게다가 ③에서 꼽은 하드웨어 기술의 진보로 인한 영향도 커서, 이전의 온프레미스형(on-premise)⁵⁾이 주류였던 데이터 분석기

반의 구축에 있어 Amazon Web Services(AWS)⁶⁾로 대표되는 클라우드형으로 분석기반을 구축할 수 있는 시스템이 보급되면서, 그때마다 분석에 필요한 만큼을 과금하기만 한다면 누구라도 분석에 충분한 시스템 환경을 얻을 수 있게 된 것이다.

두 번째는 기업에서의 디지털 트랜스포메이션(DX)⁷⁾ 의식의 고조이다. 앞에서 논의한 정보통신 기술의 발전을 배경으로, 특히 대기업들은 디지털을 활용한 기존 업무의 효율화와 더불어 새로운 제품·서비스·비즈니스 모델의 창출을 중요 경영과제로 선정되고 있다. 그리고 기업 내외에 존재하는 대량의 데이터를 분석·활용하기 위한 데이터 사이언티스트들을 DX 실현의 핵심인재로 보고 있어, 그 필요성이 급속하게 높아지게 되었다.

세 번째는 특히 일본에서 심각화되고 있는 인구감소 문제가 일본에서의 데이터 사

3) 미국 Tableau Software사가 제공하는 데이터의 가시화/그래프화 등을 직관적으로 수행할 수 있게 하는 툴
 4) NTT Data가 제공하는 기계학습을 자동적으로 수행할 수 있게 하는 툴

5) 자사에서 보유하는 설비로 시스템을 구축하고 관리하는 운용 형태
 6) Amazon이 제공하는 클라우드 컴퓨팅 서비스
 7) 일본의 경우, 디지털 트랜스포메이션(Digital Transformation)를 DX로 사용하고 있음

[그림 3] 일본기업의 데이터 사이언스 활용에 관한 대처 사례

기업명	업계·분야	데이터 사이언스 활용에 관한 대처 사례
다이킨 공업	제조업	<ul style="list-style-type: none"> IT부분을 중심으로 인프라 정비, 설계 책정, 테크놀로지 활용을 강화하기 위해 신입 사원을 다수 채용하여 업계지식과 AI 최신 기술 스킬을 폭넓게 육성 IT부에서 시범적으로 적용한 RPA 및 챗봇을 사업부에 제안하는 등 연계를 강화하고 있음 상기의 연계 강화를 통해 업무의 효율화·고도화로 인한 사업 공헌으로 이어지고 있음
트러스트 나카야마	유통업	<ul style="list-style-type: none"> 상품을 즉납할 수 있는 체재정비를 위해 10년간 100억엔 이상을 서비스 제공·정보기반의 시스템 정비·강화를 위해 투자 하여 이익 확대를 실현함 기간제 시스템을 재구축한 다음에 이용 데이터를 시가 자동으로 수집·분석·제안하는 시스템의 개발·도입을 실현함 최종적으로 업무의 효율화와 클라우드화, 업계 플랫폼 창출을 목표로 함
다이와 증권	금융업	<ul style="list-style-type: none"> AI를 활용한 경제지표 예측, 마켓정보·상표 정보의 제공을 서비스로 전개함 So과 SoE와의연계 강화를 위해 자사에서 API기반을 개발하여서비스 실현까지의 기간 단축을 달성함 전사원을 대상으로 Python자격 취득의 비용 보조, 승진 제도에 기능 습득을 적용시켜, IT 인재를 강화하고 있음
KDDI	정보 통신업	<ul style="list-style-type: none"> 실현을 위해 그룹 회사에서 DX인재를 모집하여 [KDDI DIGITAL GATE]를 개설, 유저와 자사독자의 IoT인프라를 활용한 제품 및 서비스의 공동 연구를 추진 IoT외에 AI, VR, 5G등을 활용한 최신기술 개발에 매진, 자동 운전에서의 부드러운 자선 합류를 실현하는 AI기술, 활용 영상을 바탕으로 카메라가 없는 장소를 포함한 전시점에서의 VR 시청을 가능하게 하는 기술 등이 있음
에자이	의료·간호업	<ul style="list-style-type: none"> 소유하고 있는 다양한 헬스케어 데이터를 접근 가능한 형태로 하나의 데이터 베이스로 관리하고 동시에 그 데이터를 바탕으로 개별 솔루션을 제공하기 위해 AI, IoT,클라우드와 같은 최신기술 활용을 추진하고 있음 새로운 이념실현, 이를 위한 업무 프로세스에 적합한 형태의 기간 시스템 재구축을 실시하여 시스템 운용 비용의 70%를 삭감하고 서비스 창출에 대한 투자로 전용하여 부가가치 창출을 실현함
Japan Taxi	교통·운수업	<ul style="list-style-type: none"> 자사에서 개발·제공하는 택시 배차 앱에 도착 전 지를 절차·인터넷 결제를 진행하는 추가 및 상업 시설용 배차앱을 개발하는 등 디지털화에 매진 중 각 택시에 태블릿을 제공하여 승승 서비스 중 고객에게 기업 광고를 노출하는 서비스 'Tokyo Prime' 또는 방영 후에 상품을 제공하는 서비스 '프리택' 등 승승에 그치지 않고 새로운 비즈니스 모델을 정출하고 있음

출처) 日経 BP「DXサーベイ」、富士キメラ総研「2018 デジタルトランスフォーメーション市場の将来展望」、각종 HP에서 NRI작성
 닛케이 BP(DX서베이), 후지키메라총합연구소(2018 디지털 트랜스포메이션 시장의 장래 전망)

[그림 4] 유럽과 미국기업의 데이터 사이언스 활용에 관한 대처 사례

기업명	국가명	업계·분야	기업명
Amazon	미국	소매·유통업	<ul style="list-style-type: none"> 클라우드 서비스를 제공하는「AWS」에서 컨택트 센터 서비스와 데이터 베이스 제공 서비스를 확충하고 있음. 이러한 프로젝트를 만들 수 있는 기술 분야의 채용이 MBA채용을 대신하여 매년 증가하고 있음
Volkswagen	독일	제조업	<ul style="list-style-type: none"> 클라우드·서비스 회사 지멘스와 아마존과 산업을 클라우드 플랫폼의 개발, 이에 따른 공장 기능 통합을 실시 중. 대부분의 공장 생산 설비의 일일적 데이터 관리에 따른 공장 최적화·업무 효율화에 힘쓰고 있음
Johnson & Johnson	미국	의료	<ul style="list-style-type: none"> 헬스케어 관리를 위해 웨어러블 트랙 기구와 연계하여 심박수·수면시간·관절통의 레벨 등 의료·건강 데이터를 관리하여 약의 효과 상태를 파악하는 모바일 앱을 개발함 외과 수술용 로봇을 개발하여 테스트하고 있는 단계이며, 앞으로 로봇을 통해 데이터 수집·분석·가시화·제안까지 가능한 누구나 접속 가능한 플랫폼으로 가능하도록 개발 예정
JPMorgan Chase	미국	금융업	<ul style="list-style-type: none"> 일반 소비자용 온라인 금융 플랫폼을 개시, 모바일 뱅킹 앱을 통한 오픈 플랫폼의 정비 등 일상 생활 금융 서비스의 디지털화에 매진 중 핀테크 기업의 자사내 소프트 사업 및 테크놀로지 영역의 투자액 확대, 블록체인·가상 화폐 연관 인재 확보에 따른 디지털 결제 화폐 개발등 최첨단 테크놀로지를 활용한 DX추진에 매진 중
Apple	미국	정보통신업	<ul style="list-style-type: none"> 자사 독자의 iPhone 연동형 신용카드를 발행하여 소비자의 구입이력 관리·청구지불 기능을 기존 ApplePay에 추가하여 금융 서비스의 일면을 강화, 생활의 모든 장면에서 Apple 제품 침투를 목표로 함
Uber Technologies	미국	교통·운수업	<ul style="list-style-type: none"> 손쉬운 이용과 동시에 고객 불만 등을 유연하고 빠르게 서비스에 반영시키기 위해 독자적으로 시스템을 구축하여 고객 만족도의 향상으로 이어가는 중 배차서비스르 비롯해 주요 4가지 서비스 등록자의 플랫폼을 형성하여 디지털화의 이익을 누릴 수 있는 체제를 정비, 최근에는 고용자에 대한 디지털 승급 서비스 제공도 도입함

출처) Forbes「Top 100 Digital Companies」、FINANCIAL TIMES「Top 100 global brands 2019」、각종 HP에서 NRI 작성

이언티스트 활용의 필요성을 높이고 있다. 국립사회보장·인구문제 연구소가 발표한 통계에 따르면 일본의 노동인구는 1995년을 피크로 감소일로로 걷고 있으며, 2020년 시점에서 7,341만 명까지 감소한 노동인구는 2060년에는 4,418만 명으로 2020년의 약 60%까지 감소할 것으로 전망되고 있다. 이러한 상황 속에서 사회시스템을 유지하기

위해서는 지금까지 사람의 손으로 진행해 왔던 업무 효율화 및 노동력 절약의 추진과 더불어, 1인당 생산성이 보다 높은 업무로 전환해 나가는 작업이 필수가 되고 있다. 게다가 세상의 가치가 기존의 제품 중심에서 정보·서비스 중심으로 전환해 가고 있으며, 기존의 비즈니스와 업무에 대해서는 가능한 한 사람 손을 통하지 않고 효율적으로

진행해 나가는 것이 기업에서도 요구되고 있다. 이로 인해 기업내의 인적 리소스도 데이터 사이언스를 활용하여 사람의 손과 사람의 감, 경험에 의존해 왔던 영역을 데이터 분석을 통해 보다 생산성이 높은 업무로 전환해 나가야만 하는 것이다.

III. 데이터 사이언티스트의 활용 사례

데이터 사이언티스트는 업종·업계를 불문하고 다양한 영역에서의 활약이 기대되고 있으며, 또한 데이터 사이언티스트의 능력을 최대한 발휘하기 위해서는 각 업종·업계에 맞춘 스킬을 가진 인재를 배치할 필요가 있다. [그림 3]과 [그림 4]에는 국내외의 각 업종·업계의 대표적인 데이터 사이언스 활용사례에 대해 게재하였다. 데이터 사이언스는 제조업부터 유통, 금융, 정보통신, 의료·간호, 교통·운송 등 폭넓은 업계에서 활용이 추진되고 있으며, 앞으로도 모든 업계에서 데이터 사이언티스트가 활약할 수 있는 영역이 확대될 것이라 상정된다. 특히 미국은 서두에서의 직업 랭킹에서도 언급한 바와 같이 데이터 사이언티스트의 필요성과 평가가 매우 높아지고 있으며, 그 배경에는

GAFAs를 중심으로 한 데이터 활용을 사업의 근간으로 하는 기업의 급성장과 그러한 기업에서의 활약을 지향하는 학생·연구자의 증가, 나아가 이러한 우수한 인재의 유동성·획득전쟁을 촉진하는 기업·대학간의 장기인턴십과 공동 연수 시스템의 존재 등이 크게 영향을 주고 있다고 생각된다. 또한 [그림 5]는 기업에서의 대표적인 분석 테마에 대해서 정리한 것으로 각 분석 테마에 따라 다루는 데이터의 종류, 크기, 형식은 완전 다르며 또한 데이터의 성질이 다르면 분석에 필요한 알고리즘과 모델에 관한 지식·스킬도 달라진다.

예를 들면 마케팅 분석 중 하나인 '이탈/해약 징조 분석'에서는 축적된 고객 데이터를 이용하여 과거에 이탈/해약했던 고객이 이탈/해약에 이르는 과정에서 어떠한 행동(서비스 신청, 문의 이력 등)을 취했는지, 또 어떠한 세그먼트의 고객일수록 이탈/해약 리스크가 큰지 등을 데이터를 통해 분석한다. 이러한 분석 결과를 현재 고객의 데이터에 대해 적용함으로써 이탈/해약 이전에 그 징조를 알아차릴 수 있어 영업과 메일 매거진 등을 통해 이탈/해약 방지를 위한 액션을 취할 수 있다.

[그림 5] 기업의 대표적인 분석테마

마케팅 분석 <ul style="list-style-type: none"> ● One to One/ 옴니채널 마케팅 ● 이반(離反)/해약(解約)징조 분석 ● 크로스 셀 · 업 셀 ● 고객경험(CX) 향상 ● 유인 접객 채널의 고도화 ● AI를 이용한 시책 타겟팅 최적화 등 	서플라이체인 분석 <ul style="list-style-type: none"> ● IoT에 의한 설비 관리 최적화 ● 서플라이체인의 효율화 ● 재고의 최적화 ● AI를 이용한 수요 예측/시뮬레이션 자동화 	오퍼레이션 분석 <ul style="list-style-type: none"> ● IoT로 업무 프로세스 가시화·효율화 ● 컨택트 센터 운영의 효율화 ● 음성인식·문자 해석을 통한 접객 고도화 등
휴먼 리소스 분석 <ul style="list-style-type: none"> ● 인재 평가·배치의 최적화 ● 인재 채용의 효율화 ● 워크 스타일의 고도화 ● 휴직·퇴직 징조 분석 등 	휴먼 리소스 분석 <ul style="list-style-type: none"> ● 여신 모델 구축 ● 독촉 모델 구축 ● 부정 검지 모델 구축 ● 인시텐트 예측 모델 구축 등 	파이낸스 분석 <ul style="list-style-type: none"> ● 투자 포트폴리오 최적화 ● 투자 판단 기준 고도화 ● 투자 안전/사업 모니터링 고도화 등

출처)NRI 작성

또한 서플라이체인 분석 중 '재고의 최적화'의 경우, 제조공장에서 하루에 제조가능한 생산량, 각 물류창고에서 보관 가능한 재고량, 그리고 고객으로부터의 수요량 등을 가미하여 고객의 수요를 최대한 만족시키면서도 각 서플라이체인 상에서 비용 최적화가 되는 재고의 보유방법, 보관방법, 보관기한 등을 결정하기 위한 분석에 사용할 데이터·분석 모델이 필요하다.

이러한 서플라이체인 분석과 마케팅 분석은 분석 목적부터 다루는 데이터와 알고리즘까지 달라지지만, 모두 데이터 사이언스를 활용한 전형적인 사례 중 하나이며 데이터 사이언티스트가 다루는 테마의 광범위한 넓이를 이해할 수 있을 것이다.

IV. 왜 기업에서 데이터 사이언티스트 활용이 진행되지 않는 것인가

지금까지 데이터 사이언티스트의 스킬에 대한 정의 및 데이터 사이언티스트의 수요가 높아지고 있는 배경, 그리고 구체적으로 데이터 사이언티스트가 다루는 분석 테마에 대해서도 정리하였다. 본장에서는 왜 일본의 기업에 있어 데이터 사이언티스트의 활용이 진행되지 않는가에 대해 고찰하고자 한다.

1. 데이터 사이언스의 기초지식 부족 및 불명확한 활용목적

기업의 데이터 사이언티스트 활용이 진행되지 않는 첫 번째 요인으로 원래 활용주체인 기업이 데이터 사이언스를 활용하여 어떤 일이 가능한지, 어떤 임팩트·효과를 내고 싶은지를 명확하게 이해하거나 정의하고 있지 않은 상황을 들 수 있다. 서두에서 데이터 사이언티스트에 관련된 키워드로서 AI와 기계학습과 같은 키워드를 소개했지만, 기업 안에서 이들 용어에 대한 의미의 차이를 이해하고 바르게 구분하여 사용하고 있

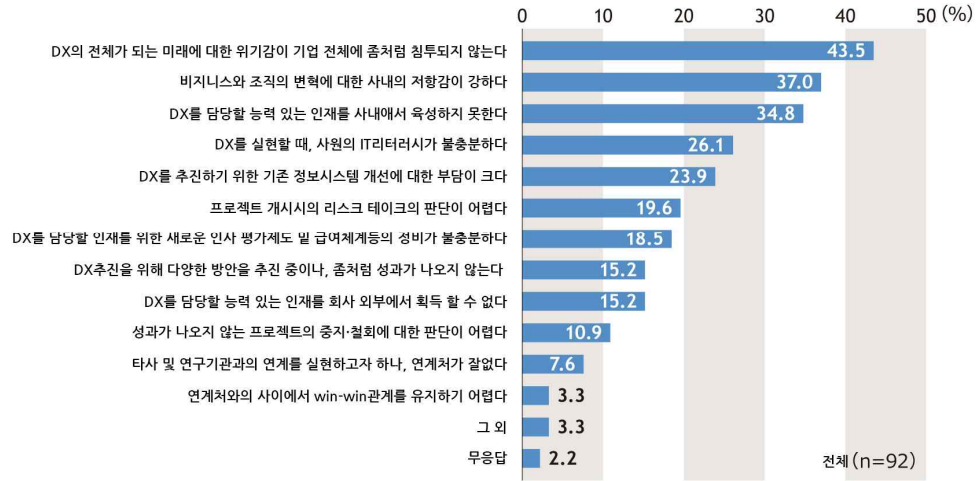
는 사람들은 아직 얼마 되지 않으며, 또한 자사의 비즈니스 모델 및 사업 환경 등을 비교해 보았을 때, DX 및 데이터 사이언스 활용에 있어 자사가 어떤 방향성(기존 사업의 효율화를 목표로 할 것인지, 새로운 비즈니스 모델로의 전환을 목표로 할 것인지 등)을 지향할 것인가가 정의되어 있지 않은 케이스도 많이 존재한다. 이 점을 등한시한다면 자사에 데이터 사이언티스트를 영입하더라도, 데이터 사이언티스트에 대한 기대와 요구가 자사의 실태와 데이터 사이언티스트의 스킬과 크게 괴리되어, 데이터 사이언스의 활용이 전혀 진행되지 못하게 되어 버리고 말게 된다고 생각된다.

2. 사내 전체의 데이터 사이언스 활용에 대한 의식의 희박성·저항감

두 번째 과제는 기업이 데이터 사이언스 활용의 필요성과 목적을 인식한 후에도 현장을 중심으로 한 사내 전체에서 데이터의 활용에 대한 희박한 의식과 저항감을 들 수 있다. 2019년에 독립행정법인 정보처리 촉진기구(IPA)가 실시한 '디지털 트랜스포메이션 추진 인재의 기능과 역할방식에 관한 조사' [그림 6]에 따르면, DX를 추진할 때의 과제들 중 'DX의 전제가 되는 미래에 대한 위기감이 좀처럼 기업 전체에 침투되지 못하고 있다(43.5%)', '비즈니스 및 조직 변혁에 대한 사내의 저항감이 강하다(37.0%)', 'DX를 담당할 능력 있는 인재를 사내에서 육성하지 못하고 있다(34.8%)'가 상위 3개 과제로 꼽히고 있으며, 데이터 사이언티스트의 활용이 진행되지 않는 배경도 이와 같은 이유일 것이라고 추측된다.

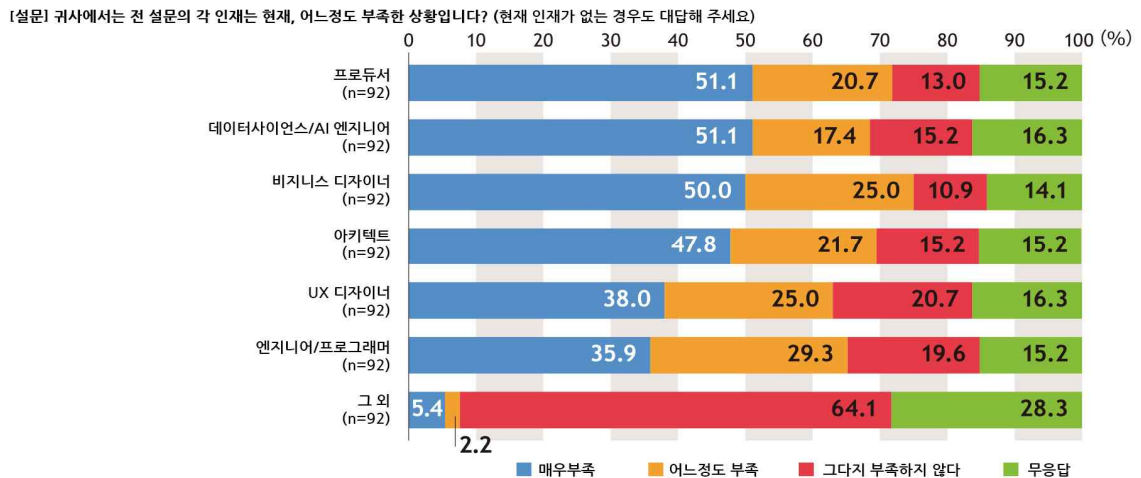
IPA의 조사결과에서 상위로 지적된 사원의 희박한 위기의식과 변혁에 대한 저항감은 기업의 데이터 사이언스 활용 현장에서 필자도 자주 목격하였으며, 조직 내부에 DX·데이터 사이언스 전문부서를 설치하여 우수한 데이터 사이언티스트를 배치하였음에도 다른 조직(특히 현장 부서)측의 DX

[그림 6] DX를 추진할 때의 자사 과제



출처) 독립행정법인 정보처리 추진기구 '디지털·트랜스포메이션 추진 인재의 기능과 역할 방식에 관한 조사' (2019년 5월 17일) NRI 작성

[그림 7] DX추진에 있어 인재의 부족감



출처) 독립행정법인 정보처리 추진기구 '디지털·트랜스포메이션 추진 인재의 기능과 역할 방식에 관한 조사' (2019년 5월 17일) NRI 작성

및 데이터 활용 의식이 낮고 수동적인 자세를 취하게 되어 DX 전문부서가 가져오는 분석결과 및 기획 등을 일방적으로 받기만 하는 입장이 되고 마는 케이스나, 현장이 외부에게 영향을 받는 것을 꺼려해 적극적으로 활용하지 못하는 케이스도 드문드문 보인다. 또한 앞의 앙케이트에서 상위 3위에 이어 지적된 'DX를 실현하는 데에 사원의 IT 리터러시가 불충분하다(26.1%)'에 대해서도 현장에서 목격되는 경우가 종종 있다. 어떤 기업에서는 현장 부서에 대해

DX·데이터 사이언스를 활용하여 해결해야만 하는 비즈니스 과제에 관한 의견을 수집한 결과, 기술적인 제약과 사내의 데이터 활용상황 등에 대한 이해가 부족하여 전혀 비현실적인 요망이나 아이디어가 튀어나와 DX·데이터 사이언스 부서가 곤경에 처하는 사례도 보이는 등 어떻게 DX 전문조직과 그 외 조직과의 갭을 채울 수 있을지도 중요한 과제가 된다.

3. 데이터 사이언티스트의 획득·육성

의 어려움

세번째로 정작 데이터 사이언스 활용을 추진하고자 하는 타이밍에서 기업 내의 데이터 사이언티스트의 부족은 큰 과제 중 하나로 등장한다. [그림 7]은 DX 추진을 담당하는 인재 부족에 대해 조사한 결과를 나타낸 것으로 '매우 부족'이라고 응답한 비율이 가장 높았던 인재부문 중 하나가 '데이터 사이언티스트(또는 AI 엔지니어)'였다. 인재부족에 대한 과제로는 사내에서의 육성의 어려움과 회사 외부에서의 채용·정착율이 낮다는 양면이 존재하고 있다. 사내에서의 육성이 어려운 배경에는 앞서 언급한 데이터 사이언티스트 협회의 스킬 체크리스트에도 나와 있듯이 비즈니스력에 더해 데이터 사이언스력과 데이터 엔지니어링력이 각각 요구되는데, 이들 스킬은 전문성이 높은 동시에 기술 발전의 속도가 매우 빨라 본래 이들 영역에서 초심자가 배우는 데에는 일정 이상의 육성기간이 필요하다는 점이 과제로서 지적되고 있다. 이러한 인재는 시장에서 크게 부족한 상황이며 만일 채용에 성공해도 사내에 정착하지 못하는 경우가 다음 과제로 지적되고 있다.

데이터 사이언티스트가 정착하기 어려운 요인 중 하나로, 대부분의 일본기업에서는 아직까지도 데이터 사이언티스트의 활용 목적이 명확히 정해지지 않았고 동시에 사내의 커리어패스 설계에 관한 지식·경험이 부족한 부분이 있다. 본고의 서두에서 미국의 데이터 사이언티스트의 연수입 중앙치(약 1,188만 엔)에 대해 소개했지만, 일본에서는 약 710만 엔⁸⁾ 정도로 미국 시세의 약 60% 수준에 머물러 있어, 우수한 데이터 사이언티스트에게 일본 기업은 전혀 매력적인 취업처로 비춰지지 않는 것이 실태이다. 또한 데이터 사이언티스트로 채용했음에도 불구하고 사내 육성의 시스템이 정비되지 않아 관습적인 현장연수 등으로 피폐해지거

나, 애초에 분석해야 할 데이터가 존재하지 않고 아날로그적인 정보(종이 장부 등)가 대부분으로 데이터를 정비하는 부분부터 시작해야 하는 등 고도의 데이터 사이언스 스킬을 발휘하기 이전 단계에서 퇴직해 버리는 케이스도 자주 들려오고 있는 상황인 것이다.

V. 데이터 사이언스의 활용을 위한 방향성

본장에서는 4장에서 정리한 데이터 사이언티스트의 활용에 관련된 과제에 대한 해결 방향성에 대해서 서술하고자 한다.

1. 자사의 데이터 활용에 관한 현상이해의 심층 분석

데이터 사이언스의 활용해 대한 이해를 배양하기 위해서는 우선은 자사의 현상을 이해하고 어떤 데이터 사이언스를 활용하는 것이 가능한지, 어떤 것부터 시작해야 하는지를 판단할 필요가 있다. 대부분의 기존 일본기업에서는 아직도 데이터를 사용하지 않는 업무 영역이 다수 존재하며, 수집되는 데이터의 양과 종류도 매우 한정되어 있는 것이 현재 상황이다. 이러한 기업에서는 우선 데이터 사이언스를 활용하여 무엇을 실현하고자 하는지, 이를 위해 자사의 데이터와 인재가 어느 정도 부족한 상황인지를 정확하게 인식한 다음에 어디서부터 데이터 사이언스의 활용을 추진할 것인지를 판단할 필요가 있다. 예를 들면 자사와 유사한 기업 및 업계에서의 데이터 사이언스의 활용사례 등을 연구함으로써 자사에서 어떤 데이터 사이언스 활용의 여지가 있는지, 반대로 현시점에서 어느 부분에 한계가 존재하고 있는지를 이해하고 보다 현실에 입각한 목적설정을 진행하는 것이 중요하다. 또한 AI와 기계학습과 같은 키워드에 대해서도 기초적인 알고리즘의 구조와 원리를 이해한 뒤 상용 사례 등을 참조

8) 국내 대형 구인정보 사이트 Indeed에 2019년 11월~2020년 10월까지 12개월간 게재된 구인광고 7,666건으로부터의 추계치

한다면, 어떤 업무에 이들 기술을 적용시킬 수 있을지 활용하는 데에 어떤 제약이 존재하는지 등을 적절하게 인식할 수 있을 것이다.

이러한 이해는 데이터 사이언스 활용을 추진하고자 하는 경영층과 실제 활용주체가 되는 현장 부서의 양자 간에 확실히 공유해 두는 것도 중요하다. 양자 간에 이해에 있어 차이점이 발생하는 경우에는 현장의 활동이 경영층에게 적절하게 평가되지 않는 부분도 있으며, 또는 반대로 현장이 목적을 잃고 헛되게 리소스를 소비하는 사태에 이를지도 모르기 때문에 유의할 필요가 있다.

2. 최고경영자의 주도과 현장의 참여

다음으로 필요한 것이 '최고 경영자의 주도·헌신(Commitment)'과 '현장을 참여시킨 데이터 사이언스 활용 사이클의 철저화'를 통한 '사내 전체의 데이터 사이언스에 대한 희박한 의식과 저항감의 불식'이다.

DX 및 데이터 사이언스를 활용한 대처를 진행하기 위해서는 이를 추진하는 전제로서 데이터를 수집하기 위한 사내 시스템·구조의 조성부터 관계자의 조정, 분석 결과의 현장 적용까지 성과를 얻기 위해 매우 긴 시간을 필요로 하기 때문에 단기적인 리턴을 얻을 수 없는 케이스가 많다. 이 때문에 경영층을 비롯한 최고 경영자는 장기적인 시점에서 이들의 영역에 투자를 해 나가지 않으면 성과로 이어지는 일은 한층 더 곤란해진다. 또한 이렇게 장기적인 시점이 필요한 요인으로 인해 현장 부서에서는 DX·데이터 사이언스 부서에 대해 '돈 먹는 벌레'라는 인상을 더욱 강하게 만들어 버릴 우려도 있기 때문에 최고 경영자는 항상 이러한 리스크를 의식하고 현장 부서에 DX·데이터 사이언스 활용의 중요성을 계속 발신할 필요가 있다. 구체적으로는 경영 메시지와 중기 경영계획 등에서 데이터 사이언스 활용의 중요성을 계속적으로 발신하거나, 데이

터 사이언스 전문 부서를 경영조직 밑에 배치하고 명시적으로 데이터 사이언스 활용에 투자하는 예산을 확보하는 등의 대처가 필요하다.

다음으로 요구되는 것이 사업부문에 있는 현장 부서의 참여이다. 전장에서도 지적한 바와 같이 현장부문이 수동적이면 데이터 사이언스의 활용이 진전되지 않을 뿐만 아니라 처음부터 분석결과와 아웃풋의 정밀도 자체도 전혀 높아지지 않는 결과가 되고 만다. 데이터 사이언스 활용의 프로세스는 비즈니스 과제의 발견에서 시작하여 이를 해결하는 데이터 사이언스 방법의 검토, 그리고 최종적으로는 비즈니스 과제를 해결하기 위한 현장 부서로의 구현·적용이 필요해지기 때문에 사업부문의 이해·협력이 필요불가결하다. 때문에 데이터 사이언스 활용으로 성과를 내기 위해서는 조기에 현장 부서를 참여시키고 가능하다면 비즈니스 과제를 발견하는 단계부터 일체가 되어 검토·반주하더라도 상기의 일련의 활동 프로세스를 몇 번이라도 반복하면서 분석의 정밀도를 한층 더 향상시켜 나가는 것이 필요하다. 또한 데이터 사이언스 활용에 대한 책임부서는 현장 부서로 하고, 어디까지나 데이터 사이언스 부서는 이를 지원하는 형태로 만드는 등 현장 부서의 주체성을 끌어내기 위한 조직·체제 만들기도 의식할 필요가 있을 것이다.

3. 필요인재의 정의와 팀 빌딩

마지막으로 '데이터 사이언티스트의 확보·육성의 어려움'을 해결하기 위해서는 우선 자사가 현재 상황에서 필요로 하는 데이터 사이언티스트의 역할·스킬을 정의한 다음에 필요한 인재를 확보하는 것이 필요하다. 3장에서도 기술한 바와 같이 각 기업에서 실시하는 분석 테마는 그 기업의 업종·업계의 특성 및 취급하는 상품·서비스의 특성에 따라 다를 뿐만 아니라 기업 자체의 데이터

활용의 성숙도에 따라서도 크고 작은 차이가 발생한다. 때문에 각 기업은 스스로의 데이터 활용의 성숙도를 파악한 다음에, 그에 맞는 인재를 확보·육성해 나갈 필요가 있다.

예를 들면 E-Commerce, SNS, 온라인 콘텐츠 배포 등을 다루는 기업이라면, 처음부터 데이터 그 자체가 상품이며, 동 기업에서 활약하는 인재에게는 일정 수준 이상의 분석 스킬이 요구될 것이다. 또한 정보통신 분야 및 금융 분야 등도 고객 데이터의 활용 등을 빠른 단계부터 추진해 온 배경이 있기 때문에 타 업계에 비해 데이터 분석은 발전되어 있으며, 보다 실천적인 분석 스킬을 보유한 인재가 필요하다고 생각된다. 한편 아직 수집한 데이터의 양과 종류가 한정되어 있는 기업에서는 데이터 사이언스 및 데이터 엔지니어링에 관련된 스킬보다도 기존 비즈니스에 대한 이해와 현장을 참여시켜 나가는 능력 등이 데이터 사이언티스트에게는 요구될 것이다. 또한 이러한 인재를 무리하게 외부의 데이터 사이언티스트를 채용하지 않더라도 자사 비즈니스에 대한 이해가 깊은 사내의 인재를 데이터 사이언스 활용 담당으로 임명하는 등 우선은 기존의 인재들 속에서 작게 시작한다는 의식도 잊어서는 안된다.

자사의 데이터 활용 스테이지에 맞는 필요 인재가 정의된 뒤 다음으로 필요한 것이, 실은 앞에서도 서술한 바와 같이 현장을 참여시키고 필요한 팀을 구성하는 것이다. 대부분의 기업에서는 한 명의 데이터 사이언티스트에게 '비즈니스력', '데이터 사이언스력', '데이터 엔지니어링력'의 모든 능력을 갖추고 있는 것을 과잉적으로 기대하는 경향이 있다고 필자는 생각하고 있으며, 본래 이 3가지 능력은 팀 또는 조직 전체로 발휘해 나가야만 한다고 생각하고 있다. 비즈니스력에 대해서는 현장부문과의 연계 및 현장부문 출신자를 팀에 참여시킴으로서 보충할 수 있고, 데이터 사이언스력과 데이터 엔지니어링

력에 대해서는 회사 외부의 파트너 기업 및 대학 등의 우수한 인재를 채용(또는 외주)하여 보충함으로써 팀·조직으로서 데이터 사이언스의 활용을 최대화하는 것이 가능해지는 것은 아닐까라고 생각한다.

또한 이러한 조직·팀의 핵심을 담당하는 고도의 데이터 사이언티스트를 적극적으로 획득하기 위해, 또 외국계 기업 등과의 인재 획득 경쟁에 대항하기 위해서도 급여·대우면에서의 재검토도 필요할 것이다. 예를 들면 IT 대기업인 DeNA는 신규 졸업자인 AI 스페셜리스트에게 최고 1,000만 엔의 연봉을 제시하는 등 채용을 적극화하고 있다. 인재 획득을 추진할 때에는 종래의 직종 수준에 업매이지 않고 적절하게 마켓 니즈를 파악한 채용을 실시할 필요가 있다. 또한 채용에 있어서는 고용 형태의 재검토도 함께 진행할 필요가 있다고 생각 된다. 앞서 기술한 바와 같이 데이터 사이언티스트의 활용이 선행적으로 추진되고 있는 유럽과 미국의 기업 등에서는 잡(Job) 방식⁹⁾의 고용 형태가 주류이기에, 각 팀, 각 안건 단위 등으로 필요한 능력·스킬을 가진 인재를 보완할 수 있다. 반면, 대부분의 일본 기업에서 채용하고 있는 멤버십 방식¹⁰⁾의 고용 형태에서는 데이터 사이언티스트의 스킬에 특화된 전문성만으로는 채용이 어려워, 결과적으로는 우수한 인재의 획득기회를 놓치는 상황으로도 이어질 우려가 있다고 생각된다.

VI. 마치며

본고에서는 왜 일본에서 데이터 사이언티스트의 활용이 충분히 진전되지 않는 것인가를 필자의 실제 경험도 참고하면서 고찰하였고, 앞으로 일본 기업에서 데이터 사

9) 직무 및 근무지, 노동시간 등이 명확하게 정해진 고용형태

10) 종합직 등의 형태로 채용하고 전근 및 이동, 로테이션을 복수회 실시하면서 장기적인 인재육성을 진행하는 것을 전제로 한 고용형태

이언티스트의 활용을 추진해 나가기 위해서 필요한 노력에 대해서 정리하였다.

언스의 활용을 고민하는 기업들에게 도움이 되기를 기대한다. (完)

특히 우수한 데이터 사이언티스트의 육성은 일본의 국제적인 경쟁력을 유지·향상시키기 위해서도 중요한 과제이며, 이를 위해서도 데이터 사이언스를 활용하여 기업에서의 활약을 목표로 하는 학생·연구자를 더욱 늘려가는 방안을 연구해야만 한다. 예를 들면 데이터 사이언티스트용 장기 인턴십의 개최 및 데이터 사이언티스트계 연구실과의 공동연구, 또는 기업과 공동의 데이터 사이언티스트 강좌의 개강 등 보다 적극적인 산학 연계기회의 창출 등을 통해, 기업과 학술계 인재와의 접점확대를 도모하는 것이 앞으로 더 한층 중요해 질 것이라고 생각된다.

데이터 사이언티스트의 활용 및 나아가 더 큰 의미에서의 DX의 추진은, 아직도 개발도상의 영역이자 명확한 정의 및 추진을 위한 해결방법도 계속 모색해야만 하는 단계에 있다. 본고의 내용이 DX·데이터 사이

필자소개

미야자키 치히로 (宮崎 地洋)
노무라종합연구소 DX컨설팅부 부주임 컨설턴트
전문분야는 기업의 디지털 트랜스포메이션 지원 등

본 칼럼은 NRI Public Management Review 2020년 12월호에서 발췌하여 한국어로 번역하였습니다. 문의사항은 노무라종합연구소 서울로 연락 바랍니다.

문의처: nri-seoul@nri.com

홈페이지 www.nri-seoul.co.kr의 insight메뉴에서 더 많은 기사를 볼 수 있습니다.

또한, 2020년 12월호에 대한 전문 및 기사(일본어)는 www.nri.com에서 열람 가능합니다.

본 칼럼의 무단 전재 및 복제를 엄격히 금합니다. 모든 내용은 일본의 저작권법 및 국제조약에 따라 보호받고 있습니다.

Copyright © by Nomura Research Institute, Ltd. All rights reserved.