

米国の IT ビジネスにおけるデータをめぐる取り組みの現状

八山 幸司
JETRO/IPA New York

1 はじめに

インターネットの発達とともに新しいビジネスが登場し、人々の社会活動から生みだされるデータは大きく変化してきたが、IoT(モノのインターネット)やビッグデータなど新しい技術の登場によりデータから生みだされる価値はこれまでにないほど大きなものになろうとしている。これまで、IT によって生み出されるデータは、オンラインショッピングといったインターネット上で生成されたデータを利用するという側面が強かったが、近年では、SNS(ソーシャルネットワーキングサービス)やコネクテッドカーなどの登場により、現実世界における社会活動そのものをデータとして利用できるようになってきている。さらに、ビッグデータや人工知能などデータの中から新しい価値を見出す技術によって、現実世界とオンライン上で生み出されるデータが人々の社会活動に大きな価値を持つものとなり、人々の生活・ビジネスを大きく変えつつある。今号では、新しいテクノロジーの登場によってデータをどのように集め、扱い、活用するか、様々な取り組みが進められている米国のデータをめぐる取り組みの現状について紹介する。

最初に、データをめぐる市場動向について紹介する。ビッグデータの市場は 2026 年まで拡大が続き、ハードウェアやアナリティクス関連ソフトウェアが大きく拡大すると見られる。また、ビッグデータの活用分野では、通信・メディアでビッグデータへ投資する企業の割合が大きく、医療、交通、教育、保険、公共インフラなどが続いている。次に、データの活用について取り上げる。人々の社会活動から生み出されるデータは大きく変化しつつあり、これまで企業内のビジネスデータや消費者の購入履歴といった一定の形式を持った構造化データが主流であったが、近年では、ソーシャルメディアや文書記録など整列されていない非構造化データの活用に注目が集まっている。データの活用事例では、企業が持つ膨大なデータを共有する取り組みや、API により様々な形で活用する事例などを取り上げ、また Google 社、Amazon 社、データブローカーによる様々な目的に沿ったデータ収集の事例を紹介する。

連邦政府におけるデータをめぐる動向では、個人データの取り扱い、犯罪捜査とデータ保護の問題、海外への個人データの移転問題について紹介する。米国における個人データの扱いは、統一された法規制がないため各分野において個別に定められている。犯罪捜査とデータ保護の問題では、連邦捜査局(FBI)が Apple 社に対して犯罪捜査のためにスマートフォンのセキュリティの解除を要求している問題について取り上げる。海外へのデータの移転制限では、米国で法規制が定められていない一方で、諸外国では様々な規制が敷かれており、米国と欧州の間ではデータの移転制限の協定について協議が続いていた。

最後に、データセンターの動向について紹介する。米国内のデータセンターの数は約 300 万施設と言われ、多くの企業がデータセンターを活用したビジネスモデルを構築している。また、建設に適した地域にはデータセンターの集積地域が登場している。データセンターには安定性や電力コストなど多くの要件が求められるため、寒冷地への建設や海底データセンターなど様々な取り組みが進められている。

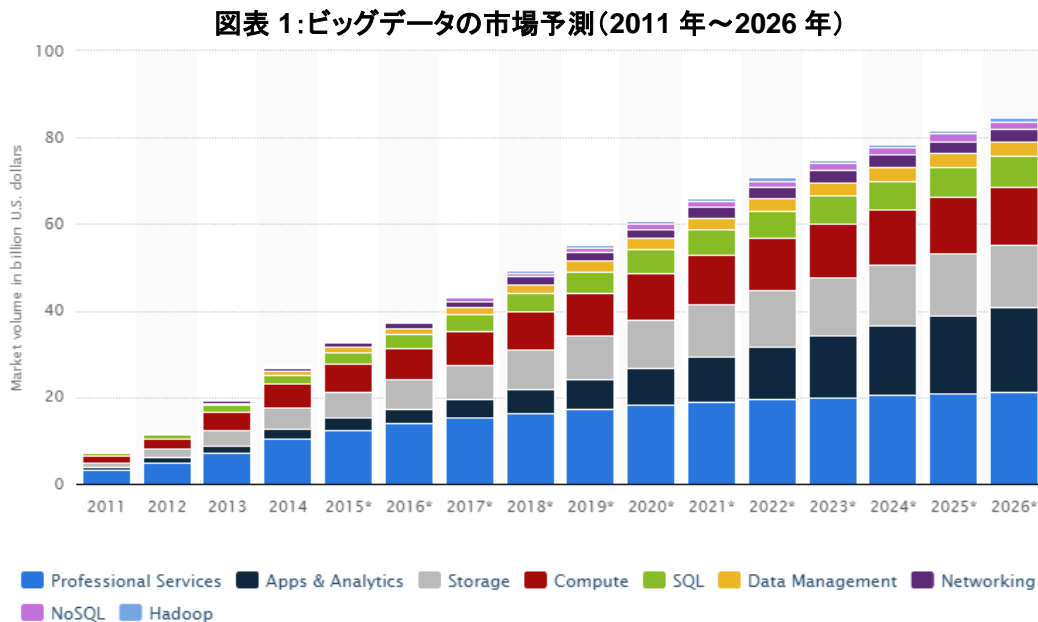
これまで生み出されるデータの多くが利用目的を持って作られていたものの、SNS や IoT の登場によって利用目的にかかわらず人々の行動そのものがデータとなっている。一方で、人々の行動がデータとしてインターネット上に残ることを懸念する声も多く、インターネット上の不利益な個人情報削除する権利「忘れられる権利(Right to be forgotten)」は欧州で認められ、Google 社は「忘れられる権利」による削除要請に基づいて 2014 年 5 月からこれまでに 140 万 9,329 件の URL を削除している。人々の活動から生まれる情報をどのようにして取り扱うのか、日本や欧州と異なる、米国のデータの取り扱いについて紹介する。

2 データをめぐる市場動向

(1) ビッグデータの市場動向

クラウドや IoT の登場によって生み出されるデータは増加の一途を辿り、ビッグデータ市場の拡大や、様々な分野で新しいデータの活用が模索されている。IT 専門家のコミュニティ Wikibon は、ビッグデータの世界市場が 2014 年の 273.6 億ドルから 2026 年には 846.9 億ドルにまで拡大すると予測しており、多くの分野でビッグデータの活用が始まり、大企業を中心にビッグデータに必要な設備への投資や専門サービスの利用が増加している。特に、コンピューターやストレージなどのハードウェアは今後も順調に拡大し、アナリティクスやアプリなどのソフトウェアも 2017 年ごろから急速に拡大すると見られている¹。

図表 1 は、2011 年から 2026 年までのビッグデータの市場予測を示したグラフとなっている。



出典: Statista²

(2) ビッグデータの活用分野

ビッグデータの活用は様々な分野へ広がっており、特に、通信・メディア、医療、交通分野などでの利用が伸びている。米コンサルティング企業 Gartner 社が 2014 年に行った企業への聞き取り調査によると、産業別にビッグデータへ投資する企業数を見た場合、通信・メディアが最も多く、すでに 53%の企業がビッグデータへ投資を行っており、続いて、医療(50%)、交通(50%)、教育(44%)、保険(42%)、公共インフラ(40%)などで多くの企業がビッグデータへ投資している。特に、医療、交通、公共インフラは前年から投資企業が大きく増加しており、これらの分野はセンサーが取り付けられた機械を多く利用するためデータを集めやすく、集めたデータをビッグデータの活用へとつなげやすいためである³。

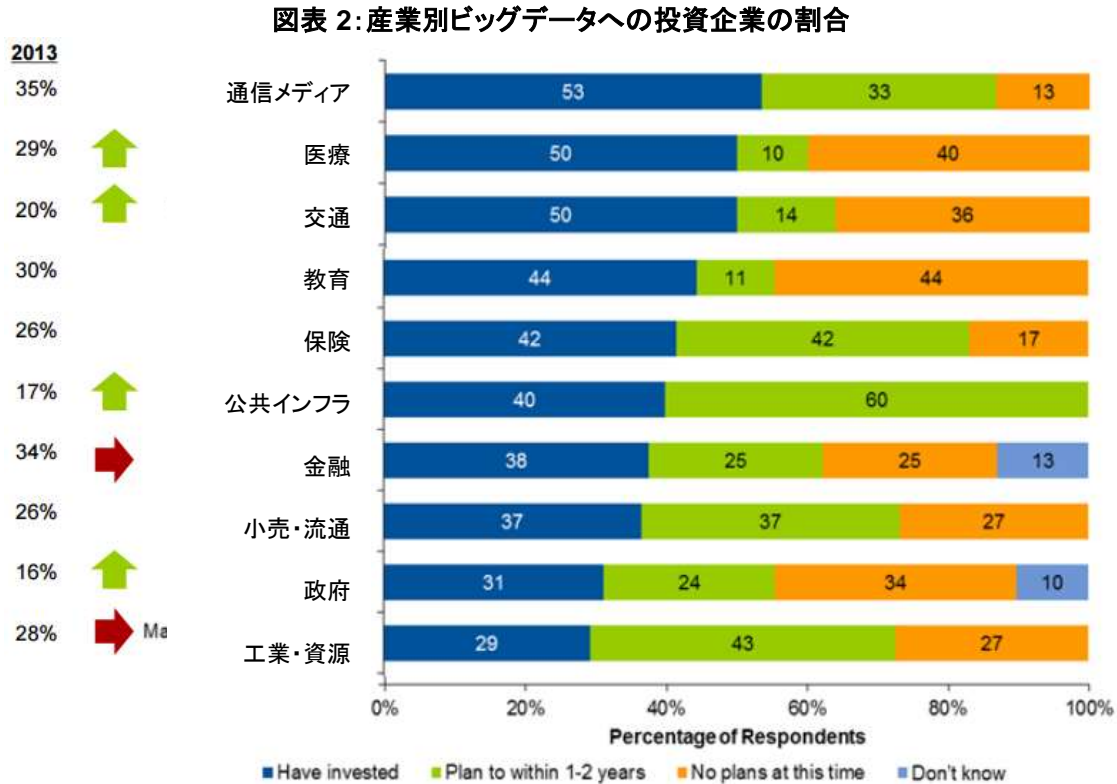
¹ <http://wikibon.com/executive-summary-big-data-vendor-revenue-and-market-forecast-2011-2026/>

² <http://www.statista.com/statistics/255970/global-big-data-market-forecast-by-segment/>

³ <http://www.gartner.com/webinar/2931518/player?commId=136995&channelId=5501&srcId=1-4730952011>

http://public.brighttalk.com/resource/core/58141/january_27_bigdataindustryinsights_lkart_85235.pdf

図表 2 は、ビッグデータへの投資企業の割合を産業別に示したグラフとなっている。



出典: Gartner⁴

ビッグデータへの投資を目的別に見た場合、カスタマーサービスの向上に役立っているのは、「通信・メディア」、「小売・流通」、「保険」、「医療」、「交通」など消費者に直接サービスを提供する産業であり、特に消費者の動向を意識する「通信・メディア」と「小売・流通」ではターゲット広告の利用が大きい。業務効率とコスト削減は目的が近いため、「医療」、「交通」、「公共インフラ」、「工業・資源開発」などで両方の目的に投資が行われている。この他、「保険」、「医療」、「金融」など厳しい法規制が敷かれている分野では、リスクマネジメントやコンプライアンスの管理にビッグデータが利用されている⁵。

図表 3 は、ビッグデータへの投資目的を示した表となっている。

⁴ <http://www.gartner.com/webinar/2931518/player?commlid=136995&channelid=5501&srcid=1-4730952011>
http://public.brighttalk.com/resource/core/58141/january_27_bigdataindustryinsights_lkart_85235.pdf

⁵ <http://www.gartner.com/webinar/2931518/player?commlid=136995&channelid=5501&srcid=1-4730952011>
http://public.brighttalk.com/resource/core/58141/january_27_bigdataindustryinsights_lkart_85235.pdf

図表 3:ビッグデータへの投資目的

	通信 メディア	小売 流通	保険	医療	交通	公共 インフラ	工業 資源開発	教育	政府	金融
カスタマーサービス	77%	68%	80%	83%	89%	60%	51%	70%	44%	65%
ターゲット広告	69%	77%	25%	33%	44%	33%	46%	30%	13%	50%
業務効率	38%	32%	40%	83%	56%	73%	57%	60%	50%	35%
コスト削減	31%	32%	45%	67%	56%	53%	54%	10%	50%	15%
製品開発	62%	23%	40%	-	33%	20%	32%	10%	38%	25%
情報を活用したサービス	31%	9%	30%	17%	22%	27%	41%	-	44%	25%
リスクマネジメント	15%	14%	55%	67%	22%	33%	27%	10%	31%	70%
コンプライアンス	15%	9%	20%	67%	44%	27%	19%	10%	13%	50%
セキュリティ	-	14%	5%	17%	11%	13%	5%	30%	19%	35%
その他	8%	-	-	17%	-	13%	3%	10%	-	5%
目的は決めている	-	9%	5%	-	-	-	5%	20%	6%	-

出典:Gartnerを基に作成⁶

⁶ <http://www.gartner.com/webinar/2931518/player?commlid=136995&channelid=5501&srcId=1-4730952011>
http://public.brighttalk.com/resource/core/58141/january_27_bigdataindustryinsights_lkart_85235.pdf

3 各分野におけるデータ活用

(1) データの種類

我々の日常生活や社会活動の中からは毎日膨大なデータが生まれており、それらのデータを適切に活用するためには利用できるデータの分類が重要となる。人々の生活や社会活動から生み出される様々な形のデータを大きく分類すると、企業内部のデータや医療データといった組織内部の業務の中で蓄積した内部データと、人々の消費活動やソーシャルメディアにおけるユーザーの動向など外部とのやり取りで生まれる外部データに分けることができる。さらにそれらの情報は、元々整理された形で収集される構造化データと、センサーデータなど二次利用を前提としていないことから分類し直す必要のある非構造化データの 2 つに分けられる。また、これらのデータにリアルタイム性などが加わることで、データの利用は目的や産業によって大きく異なっていく⁷。

図表 4 は、データの分類を表した図となっている。

図表 4: データの分類

	構造化データ	非構造化データ
外部データ	<ul style="list-style-type: none"> • 携帯電話 • GPS(位置情報) • クレジットカード利用歴 • 交通機関の利用 • 不動産情報 • 統計情報 	<ul style="list-style-type: none"> • ソーシャルメディア (Facebook、Twitter、Google+など) • ブログ • 企業外の機械から生み出されるデータ
内部データ	<ul style="list-style-type: none"> • 売上データ • 在庫情報 • 金融データ • 医療記録 • 顧客管理システム(CRM) • ウェブサイトのアクセス解析 	<ul style="list-style-type: none"> • デジタル化された文書記録 • 社内ポータル/グループウェア • SharePoint などの文書共有システム • 企業内の機械から生み出されるデータ

出典: PricewaterhouseCoopers を基に作成⁸

(2) 様々なデータの活用

a. データの共有

データの効率的な活用にはデータの種類やボリュームが重要であり、これまで個別に保有されていたデータの共有が不可欠となってきている。企業が持つ膨大なデータを共有する動きはすでに出てきており、2014 年 8 月、国連のビッグデータプロジェクトである国連グローバルパルス(United Nations Global Pulse)とデータ関連研究機関 Data & Society の主催により Responsible Data Forum が開かれ、企業が持つデータ

⁷ http://www.strategyand.pwc.com/media/uploads/Strategyand_Benefitting-from-Big-Data.pdf
http://www.strategyand.pwc.com/media/file/Strategyand_Benefitting-from-Big-Data_A-New-Approach-for-the-Telecom-Industry.pdf

⁸ http://www.strategyand.pwc.com/media/uploads/Strategyand_Benefitting-from-Big-Data.pdf
http://www.strategyand.pwc.com/media/file/Strategyand_Benefitting-from-Big-Data_A-New-Approach-for-the-Telecom-Industry.pdf

を共有し社会へ役立てていくための協議が行われた。協議の中からは、データへのアクセスや共有する企業・公共機関の選定基準について以下のようなアイデアが出された⁹。

- 大学研究機関とのパートナーシップ
- コンペティションによる優秀なビジネスアイデアを持つ企業の選定
- 中間となる機関・企業を介したデータの提供
- API¹⁰の活用
- 未加工のデータではなく、グラフなどの分析済みのデータの提供
- 複数の企業によるデータプールの構築

すでに企業が様々な形で保有するデータの共有は行われており、2015 年 1 月、配車サービスを提供する Uber 社は、マサチューセッツ州ボストンでのサービス開始に伴い、市政府に走行データを提供することを発表した。匿名化された走行データから渋滞状況などを把握し、ボストンの道路整備や交通機関の改善につながると期待されている¹¹。

b. API の活用

様々な形で集めたデータは企業にとって重要な資産であるため、外部へデータを流出させずに広く活用する手段として API が注目を集めている。API (Application Programming Interface) とは、アプリケーションなどから外部のデータやサービスを呼び出すためのプログラムの部品であり、通常はデータやサービスを提供する企業から提供され、特に一般に公開されているものを Open API と呼ぶ¹²。API には様々なものがあり、YouTube 社の動画データを呼び出す YouTube API や Google 社の地図情報を様々な形で利用する Google Maps API に加え、Facebook の「いいね！」ボタンも Facebook とデータを共有するための API となっている¹³。API はウェブ上のサービスだけでなくモバイルアプリにも多く使われており、異なるプラットフォーム間でのデータの共有が容易になるため、IoT への活用やクラウド上のデータへのアクセスを適切に制御するのに有効であるなど様々な利点がある¹⁴。

米ベンチャー企業 Socrata 社では、様々なオープンデータを一元的に提供することで、様々なデータを幅広く活用できるサービスを提供している。通常、企業が提供する API は企業によってコーディングや使い方が異なるが、同社では政府や自治体が提供しているオープンデータをクラウドベースで提供しており、それらのデータは同社の API だけで利用できるという利便性の高いサービスを提供している。同社は、ニューヨーク、サンフランシスコ、デトロイトなどの都市と連携して、各都市のオープンデータのポータルサイトを開設しており、それらの都市のデータは Socrata 社の API で利用可能である¹⁵。また、ホワイトハウスも政府予算を可視化して国民へ伝えるために同社のサービスを利用しており、2017 年度予算案を分野別、政府機関別、予算種別、部局別に加え、予算額や割合など様々な形でデータを見ることができる¹⁶。

⁹ <http://unglobalpulse.org/mapping-corporate-data-sharing>

¹⁰ Application Programming Interface の略。アプリケーションなどから外部のデータやサービスを呼び出すためのプログラムの部品。

¹¹ <http://blogs.wsj.com/digits/2015/01/13/uber-offers-trip-data-to-cities-starting-in-boston/>

¹² <http://searchsoa.techtarget.com/feature/Beat-the-risks-of-managing-public-private-APIs>

¹³ <https://developers.google.com/youtube/v3/>

<https://developers.google.com/maps/>

<https://developers.facebook.com/docs/plugins/like-button>

¹⁴ [https://www-950.ibm.com/events/wwe/grp/grp004.nsf/vLookupPDFs/Accelerate%20your%20Business%20-%20IBM%20API%20Management%2021%20July%202015%20V3%20Alofredo/\\$file/Accelerate%20your%20Business%20-%20IBM%20API%20Management%2021%20July%202015%20V3%20Alofredo.pdf](https://www-950.ibm.com/events/wwe/grp/grp004.nsf/vLookupPDFs/Accelerate%20your%20Business%20-%20IBM%20API%20Management%2021%20July%202015%20V3%20Alofredo/$file/Accelerate%20your%20Business%20-%20IBM%20API%20Management%2021%20July%202015%20V3%20Alofredo.pdf) p.13

¹⁵ <https://nycopendata.socrata.com/>

<https://data.detroitmi.gov/>

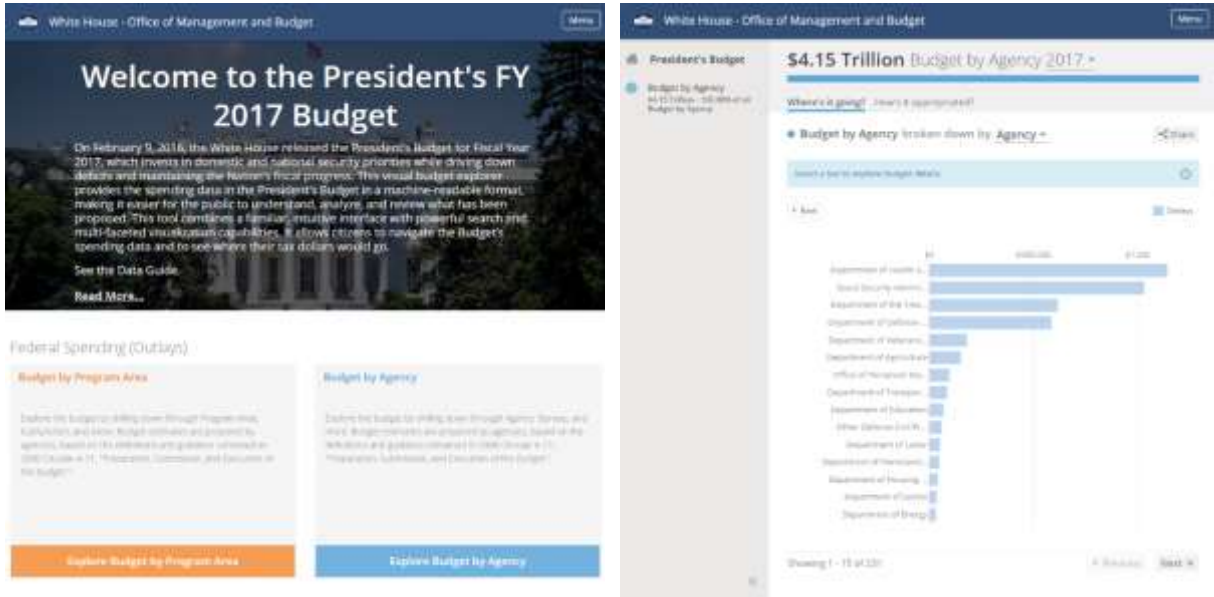
<https://data.sfgov.org/>

¹⁶ <http://whitehousebudget.socrata.com/#!/year/default>

http://whitehousebudget.socrata.com/#!/year/2017/revenue/0/function_title

図表 5 は、Socrata 社の連邦政府予算のページとなっている。

図表 5: Socrata 社の連邦政府予算のページ



出典: Socrata¹⁷

(3) データの収集

a. Google

データを活用したビジネスが本格化するにつれて企業はデータの収集を多様化させており、様々なサービスを無料で提供することでユーザーの情報を得る機会へとつなげている。Google 社では膨大なデータを軸とするビジネスを展開させており、検索サービス、メールサービス、動画配信、Google docs などのクラウドアプリケーション、携帯電話 OS など幅広いサービスを通して、ビジネスの基盤となるデータの収集を行っている。同社は公開しているプライバシーポリシーの中で収集する情報について述べており、検索内容、アクセスしたウェブサイト、視聴した動画、携帯電話端末の情報、位置情報、名前や誕生日などの個人情報に加え、送受信したメールの内容や Google docs の内容など個人が作成したデータが収集されているという¹⁸。同社のサービスの多くが無料で利用可能であり、無料のサービスによって多くのユーザーを集め、様々なデータの収集へとつなげている¹⁹。

他にもデータ収集には様々な無料サービスを介した方法が取られている。2007 年に Google 社が始めた電話番号案内サービス GOOG-411 では、自動音声に従ってビジネスカテゴリを選択し、無料で電話番号案内を受けられるサービスとなっている。ユーザーは音声認識により声で回答できるようになっているが、その際に Google 社は様々な人の音声データを集めていたと言われており、多くの人が特定のキーワードを発音することで良質な音声データを集めたと見られる。2007 年は iPhone が発売された年であり、スマートフォンや音声アシスタントを見据えた戦略のためのデータ収集と見る専門家もいる²⁰。2008 年に同社のスマ

¹⁷ <http://whitehousebudget.socrata.com/#!/year/default>
http://whitehousebudget.socrata.com/#!/year/2017/revenue/0/function_title

¹⁸ <https://privacy.google.com/data-we-collect.html>
<https://moz.com/blog/the-evil-side-of-google-exploring-googles-user-data-collection>

¹⁹ <http://www.pcworld.com/article/2986988/privacy/the-price-of-free-how-apple-facebook-microsoft-and-google-sell-you-to-advertisers.html>

²⁰ <http://www.infoworld.com/article/2642023/database/google-wants-your-phonemes.html>

ートフォン用 OS Android を搭載したスマートフォンが発売され、2010 年 10 月に GOOG-411 はサービスを終了した²¹。

データ収集の範囲の拡大による懸念も出ており、2015 年 12 月、電子フロンティア財団 (Electronic Frontier Foundation: EFF) は、同社が教育現場向けに提供するサービス Google Apps for Education を利用する際に、同社のノートパソコン Chromebook で教育機関向けに販売されているものを使用した場合、初期設定の状態で生徒の検索履歴やパスワードなどが送信される状態になっていると指摘し、問題となった。同社はデータ収集が広告以外の目的であると主張したが²²、Chromebook は米国の授業で使われるノートパソコンの半数を占めており²³、同社のデータ収集の範囲の広さによる懸念が出ている。

図表 6 は、教育現場における Chromebook の利用となっている。

図表 6: 教育現場における Chromebook の利用



出典: CNBC²⁴

b. Facebook

Facebook 社は、様々な方法を使ってインターネット上での個人の行動履歴を収集している。同社では、ソーシャルプラグインと呼ばれる「いいね！」ボタンや共有ボタンのようなウェブサイトに設置が可能な機能を提供しており、それらを通してユーザーがどこのウェブサイトを見たかわかるようになっている。具体的には、ユーザーが Facebook へログイン後に「いいね！」ボタンのある外部ウェブサイトを開覧した場合、クッキー (Cookie) と呼ばれる機能を使いユーザーの識別可能な情報を Facebook へ送る仕組みとなっている。Facebook からログアウト後も情報は送られつづけるため、Facebook はインターネット上での行動を収集することが可能となる²⁵。Facebook 社はソーシャルプラグインを使って同社のユーザー以外のクッキーを収

²¹ <https://googleblog.blogspot.com/2010/10/goodbye-to-old-friend-1-800-goog-411.html>

²² <http://www.pcmag.com/article2/0,2817,2499437,00.asp>

²³ <http://www.cnb.com/2015/12/03/googles-chromebooks-make-up-half-of-us-classroom-devices.html>

²⁴ <http://www.cnb.com/2015/12/03/googles-chromebooks-make-up-half-of-us-classroom-devices.html>

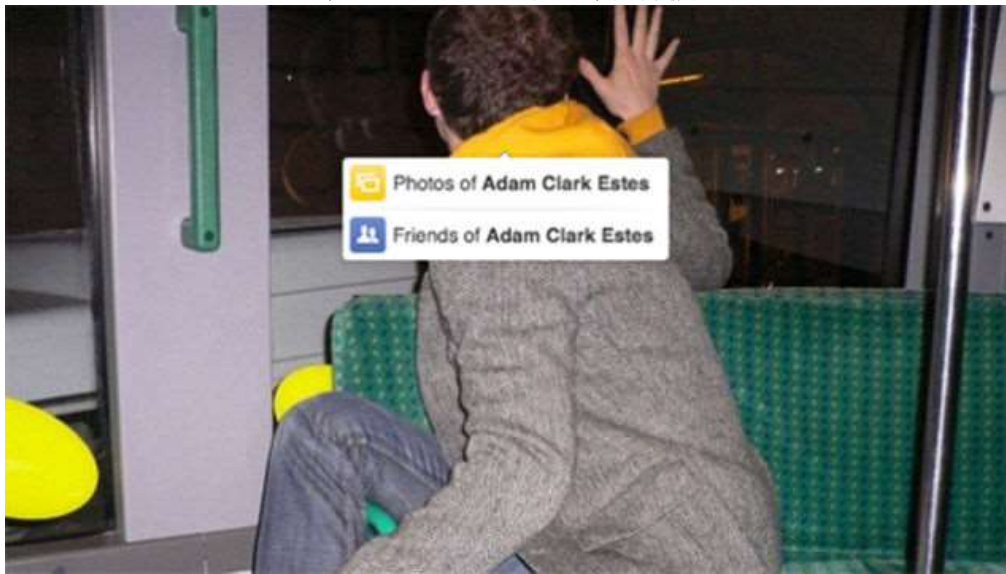
²⁵ <http://liferhacker.com/5843969/facebook-is-tracking-your-every-move-on-the-web-heres-how-to-stop-it>

集しており、2015 年 11 月にベルギーの裁判所がデータ収集の違法性を認め罰金命令を出し²⁶、2016 年 2 月にはフランスのプライバシー監督機関が非ユーザーの行動追跡を中止するように勧告した²⁷。

クッキーの利用には様々な懸念が出ており、モバイル端末に搭載のブラウザではクッキーが有効に働かないなど課題もあるため、近年では他のデータ収集についても模索されている²⁸。Facebook 社が力を入れている顔認識は人間とほぼ同程度の認識率を持っていると言われており、近年では、顔が映っていない状態であっても、ある程度の人物特定が可能であるという。Facebook の AI 部門の Yann LeCun 氏が語ったところによると、人物特定の認識精度を上げるために写真に写った人物の服装、髪の色、ポーズ、体のラインといった特徴も分析対象であるという。実際に 4 万枚におよぶ画像を分析したところ 83%の認識率で人物の特定が出来たという²⁹。

図表 7 は、Facebook 社の顔認識技術となっている。

図表 7: Facebook 社の顔認識技術



出典: Gizmodo³⁰

この他、Facebook 上でユーザーが使うプロフィール写真を、イベントに合わせて加工することが出来るサービスが提供されているが、これが情報収集の一環ではないかとも言われている。2015 年 6 月、米国全州で同性婚が合法化されたことに伴い、Facebook 上のプロフィール写真を LGBT³¹の象徴であるレインボーカラーの 7 色に染める機能を追加した。この機能を使って Facebook が政治的なサポートを表現したことは初めてであり、この機能がデータ収集の一環ではないかという声も上がり、オンラインニュースサイト The Atlantic や Gizmodo が Facebook 社に問い合わせたところ、実験やテストではなくターゲット広告に使用することは無いという回答であった。しかしながら、同社のデータサイエンティストが社内のデータを基にした「オンラインの社会運動における支持の拡散 (The Diffusion of Support in an Online Social Movement)」

²⁶ <http://techcrunch.com/2015/11/11/facebook-faces-privacy-fines/>

<http://www.reuters.com/article/us-facebook-belgium-idUSKCN0SY27220151109#SyLL1uMB2lom0ItK.97>

²⁷ <http://www.reuters.com/article/us-facebook-france-privacy-idUSKCN0VH1U1>

²⁸ <http://venturebeat.com/2014/10/06/the-cookie-is-dead-heres-how-facebook-google-and-apple-are-tracking-you-now/>

²⁹ <https://www.newscientist.com/article/dn27761-facebook-can-recognise-you-in-photos-even-if-youre-not-looking>

³⁰ <http://gizmodo.com/facebook-says-its-facial-recognition-is-so-good-it-doe-1713182436>

³¹ レズビアン、ゲイ、バイセクシュアル、トランスジェンダーの頭文字を取った言葉で、セクシャル・マイノリティ(性的少数者)を指す。

という論文を発表するなど、人々の動向や感情伝播に着目していると見られており³²、同社のデータ収集や利用について注目が集まっている。

図表 8 は、Facebook の同性婚合法化に関連したプロフィール写真の加工機能となっている。

図表 8: Facebook の同性婚合法化に関連したプロフィール写真の加工機能



出典: GIZMODO³³

c. 匿名データの活用

個人の活動に関するデータを匿名化することでデータの取引を可能にし、データの販売や分析により大きな利益を上げるデータブローカーの存在が大きくなっている。データブローカーとは、消費者の活動を収集し他の企業へ販売する業者のことを指し、ビッグデータの活用が大きくなるにつれてデータブローカーの重要性も増してきている³⁴。データブローカーの市場は 2012 年の時点で 1,500 億ドルと見積もられており³⁵、連邦取引委員会 (Federal Trade Commission: FTC) は、データブローカーが 2014 年 5 月の時点で 14 億人分の取引記録と 7,000 億件のデータを所有し、毎月 30 億件のデータが追加されていると見ている³⁶。データブローカーや個人情報情報を再販売する企業は、個人情報情報を匿名化した形で扱っており、匿名化された個人情報情報を組み合わせることで様々な社会活動の分析に使用している³⁷。

データブローカー大手である acxiom 社は、世界 5 億人の消費者に関するデータを保有し、世界中に配置したサーバー 2,300 台を使って消費者の活動を分析している³⁸。同社は消費者の動向を分析する能力に特に優れており、2015 年 3 月には LiveRamp Connect と呼ばれる消費者の様々な情報を統合することがで

³² <http://gizmodo.com/facebook-is-probably-tracking-all-those-rainbow-profile-1714635163>

<http://www.theatlantic.com/technology/archive/2015/06/were-all-those-rainbow-profile-photos-another-facebook-experiment/397088/>

³³ <http://gizmodo.com/facebook-is-probably-tracking-all-those-rainbow-profile-1714635163>

³⁴ <https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf> p.i

³⁵ <http://www.cio.com/article/2377591/data-management/inside-the-shadowy-world-of-data-brokers.html>

³⁶ <https://www.ftc.gov/news-events/press-releases/2014/05/ftc-recommends-congress-require-data-broker-industry-be-more>

³⁷ <http://www.smartdatacollective.com/tamaradull/298401/pii-anonymized-data-and-big-data-privacy>

³⁸ <http://www.webpagefx.com/blog/general/what-are-data-brokers-and-what-is-your-data-worth-infographic/>

きる新しいマーケティングプラットフォームを発表した。ブランド企業などの広告主企業 200 社と、米オンライン音楽配信事業者 Pandora 社などマーケティングに必要な情報を提供するインテグレーションパートナー 130 社が LiveRamp Connect のプラットフォームに参加しており、テレビの視聴率データ、Facebook のデータ、オンラインでの消費者の動向から広告主に広告を出す機会を伝える仕組みとなっている。さらに LiveRamp Connect では、広告主企業の店舗での購買履歴とも連動が可能であり、広告主企業の顧客管理システム (Customer Relationship Management: CRM) とオンライン上の消費者の活動を連動させることができるという点で画期的であり、オフラインからオンラインまで様々なデータを活用したマーケティングが可能となっている³⁹。

図表 9 は、LiveRamp Connect のイメージ図となっている。

図表 9: LiveRamp Connect のイメージ図



出典: axxiom⁴⁰

³⁹ <http://adexchanger.com/data-exchanges/everything-you-wanted-to-know-about-liveramp-the-data-connector-everyone-uses/>

⁴⁰ <http://liveramp.com/why-data-connectivity/>

4 データをめぐる政府の動向

(1) 米国におけるデータの取り扱い

a. 連邦政府や州政府における取り組み

米国ではデータの取り扱いに関する分野横断的な法律が定められておらず、各産業で定められたプライバシー保護に関する法律や州法がベースとなっている。プライバシー保護に関する法律は、政府機関、医療、金融、通信などの分野で個別に定められており、その多くで、個人識別可能なデータの定義、本人の同意のない情報収集や第三者機関による個人データの利用を制限している。また、個人情報の取り扱いに厳格な規定を設けているカリフォルニア州では、2014 年に改正された 2 つの州法によって社会保障番号 (Social security number) や児童の個人情報を取引することを禁じている⁴¹。

図表 10 は、個人データの取り扱いに関する法律を示した表となっている。

図表 10: 個人データの取り扱いに関する法律

法律	個人情報にあるデータ
医療保険の相互運用性及び説明責任に関する法律 (HIPPA)	個人情報および個人の特定につながる医療データ
金融サービス近代化法 (GrammLeach-Bliley Act)	企業が集めた個人情報で、他社のシステムに利用できる可能性のあるデータ
California Security Breach Notification Law	<ul style="list-style-type: none"> • 社会保障番号 • 免許証番号 • 銀行やクレジットカードの情報 • 医療情報 • 保険に関する情報
California Online Privacy Protection Act	<ul style="list-style-type: none"> • 氏名 • 住所 • メールアドレス • 電話番号 • 社会保障番号

出典: Practical Law を基に作成⁴²

b. 犯罪捜査とデータ保護の問題

近年では、犯罪捜査のために捜査機関が IT 企業へデータの引き渡しやアクセスの協力を求めていることが問題となっている。Microsoft 社は、2013 年 12 月、アイルランドのデータセンターに保管されている電子メールの提出を求める米当局の捜査令状が有効であるか判断を求める訴えを連邦地方裁判所に起こした。この捜査令状は、麻薬取引捜査の一環で出されたものだが詳細は明らかにされておらず、対象の人物も米国在住かどうか分かっていない。同社は、アイルランドの同意なしに電子メールの取得をすればアイルランドと欧州の法令に抵触する恐れがあり、米司法当局が司法取り決めを通じてアイルランドから電子メールを取得すべきだと主張した。2014 年 8 月に米連邦地方裁判所はメールの提出を命じる判決を出したが⁴³、2015 年 9 月には同社が捜査令状の指し止めを求め米連邦控訴裁判所へ上訴し論争が続いている⁴⁴。

⁴¹ <http://us.practicallaw.com/6-502-0467#a747828>

⁴² <http://us.practicallaw.com/6-502-0467>

⁴³ <http://www.wsj.com/articles/u-s-judge-rules-microsoft-must-produce-emails-held-abroad-1406826302>

⁴⁴ <http://www.wsj.com/articles/microsoft-email-case-tests-power-of-search-warrant-1441660355>

2016 年 2 月には、カリフォルニア州の連邦地方裁判所が Apple 社に対し、iPhone のパスワード解除を必要とする連邦捜査局 (Federal Bureau of Investigation: FBI) への捜査協力を命じ、iPhone のセキュリティを低下させるツールの開発を迫られた。この問題は、2015 年 12 月にカリフォルニア州で発生した銃乱射事件で、犯人の 1 人が所有していた iPhone 5c のパスワードを FBI が解除できなかったため同社に協力を求めたものである⁴⁵。iPhone は、ユーザーのプライバシーを守るために内部のデータが暗号化されており、自動消去機能を設定している場合、パスワードの入力を 10 回間違えると端末上のデータが削除される仕組みとなっている。このため、FBI はパスワードの無制限入力と自動消去機能を無効化するツールを作成するように求めていたが、同社は、自らセキュリティをすり抜けるバックドアを作ることになると主張して受け入れなかった⁴⁶。

図表 11 は、Apple 社の CEO Tim Cook 氏と FBI 長官 James Comey 氏となっている。

図表 11: Apple 社の CEO Tim Cook 氏と FBI 長官 James Comey 氏



出典: CNBC⁴⁷

また、Apple 社は同社のクラウドサービス iCloud にバックアップされる端末のデータにアクセスできるため、Wi-fi に接続されると iPhone が自動でバックアップする機能を利用できると見ていたが、地元警察が FBI からの要請で事件直後にアップル ID⁴⁸のパスワード(携帯電話を使用するためのパスワードとは別で、Apple 社のウェブサイトで変更が可能)を変更してしまったため、自動バックアップも利用できない状態となってしまった⁴⁹。2016 年 3 月 1 日には米連邦議会で公聴会が開かれ、FBI 長官と Apple 社の代表がお互いの主張を繰り広げ⁵⁰、カリフォルニア州の事件とは別件で Apple 社にパスワード解除を求めていた裁判では、ニューヨーク州の連邦地方裁判所が同社を支持するなど話が拡大している⁵¹。Apple 社は iPhone のセールスポイントの 1 つにセキュリティの高さを謳っているものの、自らの手でセキュリティを弱めるようなこ

⁴⁵ <http://gizmodo.com/why-you-should-care-about-apple-s-fight-with-the-fbi-1759639200>

<http://www.theverge.com/2016/2/17/11031364/apple-encryption-san-bernardino-response>

⁴⁶ <http://techcrunch.com/2016/02/17/tim-cook-apple-wont-create-backdoor-to-unlock-san-bernardino-attackers-iphone/>

⁴⁷ <http://www.cnbc.com/2016/02/18/confused-about-apple-and-the-fbi-what-you-need-to-know.html>

⁴⁸ Apple 社のウェブサービスを利用する際に使用するユーザーID の名称。

⁴⁹ <http://gizmodo.com/apple-this-mess-couldve-been-avoided-if-the-government-1760211382>

<http://gizmodo.com/san-bernardino-county-calls-the-fbi-liars-over-terroris-1760317923>

⁵⁰ <http://gizmodo.com/fbi-director-james-comey-is-a-clown-1762226376>

⁵¹ <http://gizmodo.com/new-york-judge-rules-us-cant-force-apple-to-help-unlock-1762036873>

とをすれば、Google社、Microsoft社、Facebook社などにも同じことを強制される恐れがあるため、前例を作らないことを念頭に置いている⁵²。

本件は、テロ対策と個人情報保護という2つのセキュリティをどう考えるかと言う、データの扱いに関する非常に重要な問題であることから、全米で連日報道されるなど注目を集め、賛成派・反対派・中間派にわかれ大きな議論を巻き起こしており、今後の行方が注目されている。

(2) 海外への個人データの移転制限に関する動向

a. 諸外国における個人データの移転制限

米国ではデータの海外への移転制限に関する法律が無いものの、欧州など様々な国では既に制定されている。米国内から海外へのデータ転送に関する規制や監督機関への届出なども定められていないが⁵³、連邦取引委員会(Federal Trade Commission:FTC)などの監督機関は、米国内から転送されたデータ、委託事業者による海外からの国内のデータへのアクセス、委託事業者によるデータの持ち出しについては米国の法律が適用可能であるという姿勢を示している。また、複数の州では海外へのデータの持ち出しについて州法で規制を設けているものの、その対象は州政府機関や委託事業者に限定されている⁵⁴。

一方で、様々な国ではデータの海外への移転制限が設けられており、それらの国では、取得した個人データをその国のサーバーで保管し処理するデータローカライゼーション(Data localization)が取られている。最も厳しい制限をかけている国としては中国、ロシア、インドネシアなどがあり、国内のサーバーにデータを保管することを定めた規定が出されている。次に制限が厳しいのが欧州であり、統一的なプライバシーフレームワークであるデータ保護指令(Data Protection Directive)によって事実上のデータローカライゼーションが取られている。この他、インド、韓国、マレーシアなどは条件付のデータローカライゼーション規定が定められている⁵⁵。

図表 12 は、データの海外への移転制限を設けている国を示した図となっている。

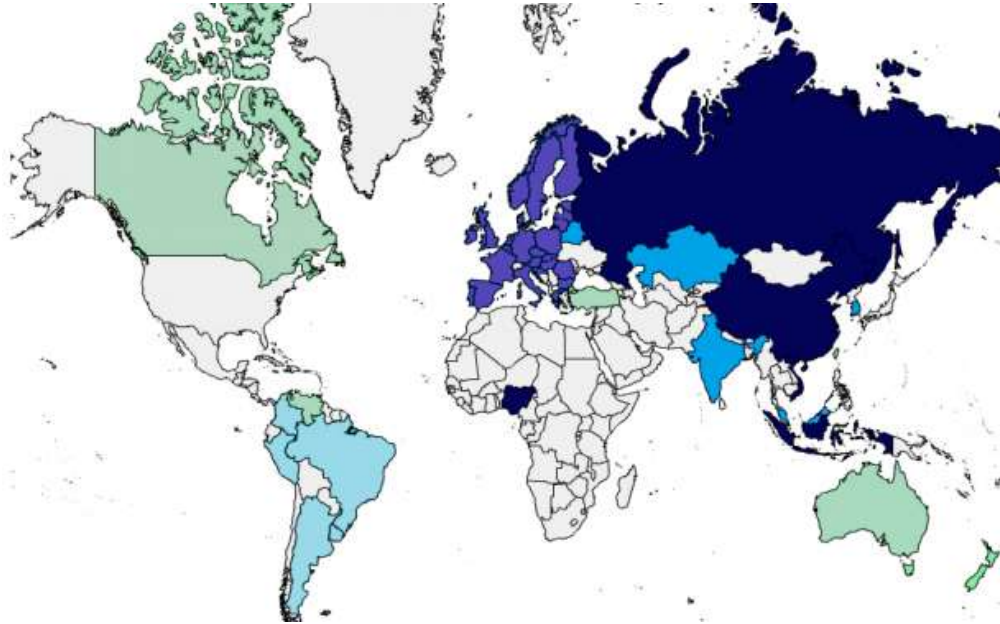
⁵² <http://www.theverge.com/2016/2/16/11027732/apple-iphone-encryption-fbi-san-bernardino-shooters>

⁵³ https://www.hunton.com/files/Publication/1f767bed-fe08-42bf-94e0-0bd03bf8b74b/Presentation/PublicationAttachment/b167028d-1065-4899-87a9-125700da0133/United_States_GTD_T_Data_Protection_and_Privacy_2014.pdf p.197

⁵⁴ <http://us.practicallaw.com/6-502-0467>

⁵⁵ <http://www.albrightstonebridge.com/files/ASG%20Data%20Localization%20Report%20-%20September%202015.pdf> p.15

図表 12: データの海外への移転制限を設けている国



地図上の色	データの海外への移転制限	国名
黒	非常に厳しい移転制限: すべての個人データに対し、データローカライゼーションを行うことを規定している。	ロシア、中国、インドネシア、ブルネイ、ナイジェリア、ベトナム
紫	事実上の厳しい移転制限: 個人プライバシー保護の中でデータローカライゼーションを定めており、データの海外への移転制限を難しくしている。	欧州連合
青	部分的な移転制限: データローカライゼーションを設けており、データの種類によっては海外への移転ができない。	韓国、マレーシア、インド、カザフスタン、ベラルーシ
緑	緩やかな移転制限: 特定の条件でのみデータの海外への移転制限を規定している。	アルゼンチン、ブラジル、コロンビア、ペルー、ウルグアイ
黄緑	特定分野のみの移転制限: 特定の分野でのみデータの海外への移転制限を規定している。	オーストラリア、カナダ、ニュージーランド、台湾、トルコ、ベネズエラ
白	無し: データローカライゼーションを規定していない。	その他の国(日本、米国など)

出典: Albright Stonebridge Group を基に作成⁵⁶

b. 欧州の動向

欧州では、プライバシー保護を重視した様々な規制が打ち出されており、新しい法整備の動きが進んでいる。欧州は、1995 年 10 月に個人のプライバシー保護を目的とした EU データ保護指令を打ち出しており、加盟国に対して個人データの取り扱いに関する指示を出している⁵⁷。この中で、適切な個人情報保護が確保されている場合に限り、欧州域外への個人データの移動が認められている⁵⁸。米国と欧州の間では、個人データ移動についての原則を記した「セーフハーバー協定(Safe Harbor)」を 2000 年に締結し、欧州のプライバシー保護規準に沿ったルールを適用することで欧米間の個人データの移動を可能にしていた。しかしな

⁵⁶ <http://www.albrightstonebridge.com/files/ASG%20Data%20Localization%20Report%20-%20September%202015.pdf>

⁵⁷ <http://searchsecurity.techtarget.co.uk/definition/EU-Data-Protection-Directive>

⁵⁸ https://fra.europa.eu/sites/default/files/fra-2014-handbook-data-protection-law-2nd-ed_en.pdf p.133

がら、スノーデン事件によって大規模な情報収集活動が行われていたことが明らかになったことなどから、2015 年 10 月、欧州司法裁判所 (European Court of Justice) が同協定を無効としたため、米国と欧州は新たな協定を策定し、2016 年 2 月に新たな協定「EU-US プライバシーシールド」に合意した。新しい協定では、欧州の個人データが米国内のサーバーに保管されている場合には情報監視活動の対象外とすることなどが盛り込まれている⁵⁹。

5 データセンターの動向

(1) 重要性の増すデータセンター

多くの企業がデータを活用したビジネスモデルを取り入れて、変化する IT ビジネスや法規制に対応するため、大量のデータを保管することが不可欠となってきており、データセンターの重要性は益々高まっている。データセンターは政府や企業にとって重要な施設であるため、はっきりとした数はわからないものの、エネルギー省は 2014 年 11 月時点で米国に 300 万施設のデータセンターが存在し、およそ 100 人につき 1 つのデータセンターがあると見ている⁶⁰。また、米調査会社 IDC 社は世界全体のデータセンターの数は増加傾向にあり、2017 年には 860 万施設に達すると予測している⁶¹。2016 年 1 月に米シンクタンク Ponemon Institute が公表した企業へ聞き取り調査では、回答した企業の 69% が外部のデータセンターを利用しているという。企業が外部のデータセンターを利用する背景には、企業の IT インフラが複雑で社内でのデータの取り扱いが難しい、社内だとコストが高い、すでに外部のデータセンターを中心としたビジネスモデルを構築しているなどの理由が挙げられている⁶²。

データセンターは様々な面でシステムの持続性が求められるため、建築にも多くの要件が求められる。データセンター内のコンピューターは膨大な電力を必要とするため、十分な電力インフラ、低価格な電力コスト、バックアップ電源などが必要となる。また、数百台のコンピューターから発生する熱を適切に冷却しなければコンピューターへのダメージにつながるため、冷房設備だけでなく、外部の冷気や冷水を使ってデータセンター内を冷却する必要がある。このため、データセンターの建設地には寒冷地であることや水源が近くにあるかなども重要となる。この他、天災の可能性、近隣のビジネス地域、エンジニアなどの人材が集まりやすいかなども要件となる⁶³。これらのデータセンターの建設に必要な条件から、データセンターが集まる地域は限られ、以下のような様々なデータセンターの集積地域が生まれている⁶⁴。

図表 13 は、米国の代表的なデータセンターの集積地域を示した地図となっている。

⁵⁹ <http://www.wsj.com/articles/u-s-eu-reach-deal-on-new-data-transfer-framework-1454429818>

⁶⁰ <http://energy.gov/eere/articles/10-facts-know-about-data-centers>

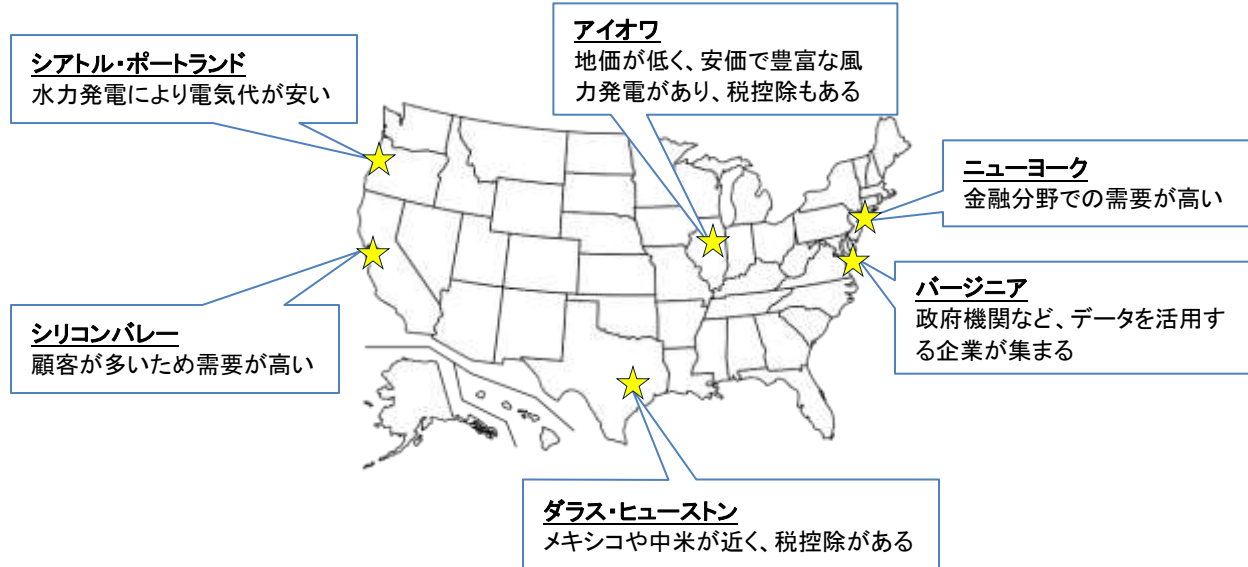
⁶¹ <http://www.idc.com/getdoc.jsp?containerId=prUS25237514>

⁶² <http://www.cyrusone.com/wp-content/uploads/2015/10/the-business-case-for-it-outsourcing.pdf>

⁶³ http://www.expansionsolutionsmagazine.com/industry_articles/view/10789/data_center_site_selection

⁶⁴ <http://www.bloomberg.com/news/articles/2015-11-05/the-best-places-to-build-data-centers>

図表 13: 米国の代表的なデータセンターの集積地域



出典: Bloomberg を基に作成⁶⁵

(2) 新しいデータセンターへの取り組み

a. 再生可能エネルギーの活用

データセンターの運用には多くのエネルギーを必要とするため、様々な新しいアイデアにより運用コストを削減する取り組みが進められている。データセンターは、内部にサーバーなどの IT 機器が高密度で設置されているため電力消費が大きいという問題がある。環境保護団体 Natural Resources Defense Council によると、米国内に設置されたデータセンターの 2013 年の総電力消費量は 910 億キロワット時 (kWh) に達し、これは大型火力発電所 34 基分の年間発電量に匹敵し、ニューヨーク市の一般家庭の年間電力消費量の 2 倍以上だという。さらに 2020 年には 1,400 億キロワット時 (kWh) まで増加すると見られ、エネルギー消費の増加による二酸化炭素排出量の増加が問題となってきている⁶⁶。このため、データセンターを運営する IT 企業では、再生可能エネルギーの使用や外部冷気を使ってデータセンター全体を冷やすことで電力消費量を抑える取り組みを進めている。

再生可能エネルギーの利用で取り組みが進んでいるのが Facebook 社であり、同社が 2013 年に建設したスウェーデンのデータセンターでは、電力のすべてに再生可能エネルギーの利用を実現している。同データセンターでは水力発電による電力を使用し、北極圏に近い外気を施設内の冷却に使用する仕組みとなっている⁶⁷。さらに 2015 年 7 月にはテキサス州 Fort Worth に、2016 年 1 月にはアイルランドに新しいデータセンターを建設することを発表した⁶⁸。この他、Facebook 社では、蓄積したデータセンターの設計をオープン化する Open Compute Project を 2011 年から進めており、同プロジェクトには、米 IT 企業 Intel 社、Microsoft 社、Google 社、米大手通信 AT&T 社などが参加している⁶⁹。

⁶⁵ <http://www.bloomberg.com/news/articles/2015-11-05/the-best-places-to-build-data-centers>

⁶⁶ <http://www.nrdc.org/energy/files/data-center-efficiency-assessment-IP.pdf> p.5

⁶⁷ <http://www.datacenterknowledge.com/the-facebook-data-center-faq-newest-page/>

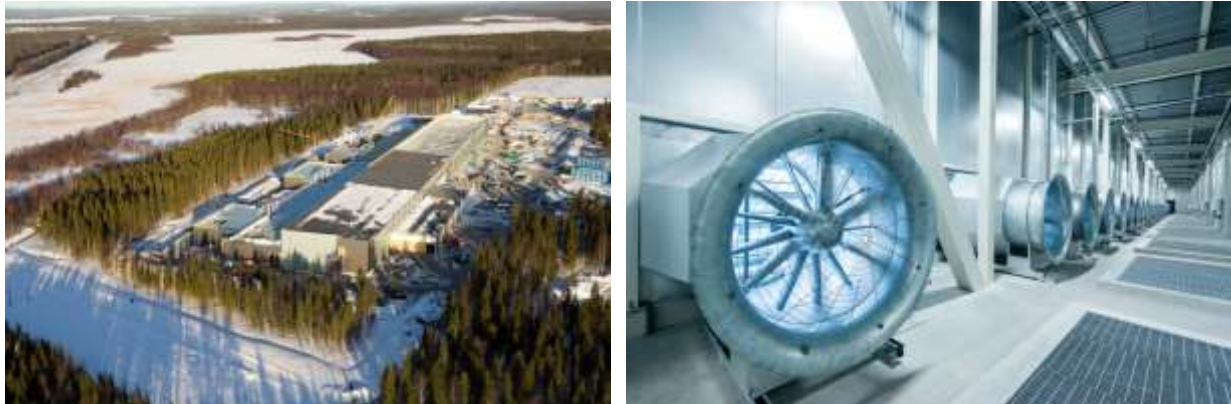
⁶⁸ <http://newsroom.fb.com/news/2015/07/the-newest-addition-to-the-facebook-data-center-fleet-fort-worth/>

⁶⁹ <http://newsroom.fb.com/news/2016/01/county-meath/>

⁶⁹ <http://www.opencompute.org/about/membership-organizational-directory/>

図表 14 は、Facebook 社が所有するスウェーデンのデータセンターとなっている。

図表 14: Facebook 社が所有するスウェーデンのデータセンター



出典: Gizmodo⁷⁰

また Amazon 社では、データセンターへの電力供給に専用の再生可能エネルギーの発電所を建設している。同社はデータセンターへの電力用として、80 メガワットの太陽光発電所をバージニア州に、風力発電所をインディアナ州、ノースカロライナ州、オハイオ州に建設している。将来的にデータセンターの電力をすべて再生可能エネルギーにすることを目標としており、現在は 25%の達成率だが 2016 年末までに 40%にまで引き上げる計画である⁷¹。また、米自動車メーカー Tesla 社が 2015 年 5 月に発表した業務用蓄電システム Powerpack をデータセンターに試験的に導入しており、環境によって発電量が変化する再生可能エネルギーを蓄電することで、安定して電力を利用できる取り組みを進めている⁷²。

b. 様々なデータセンター

データセンターの安定性、拡張性、セキュリティなどを高めるために、新しいタイプのデータセンターの開発が進められている。Microsoft 社では、海水を利用して効率的にサーバーを冷却することを目的とした、海中設置型データセンター計画 Project Natick を進めている。Project Natick は、密閉式の容器に格納したサーバーをおよそ 1km の海底に設置し安定した稼働が可能か確認するというもので、2015 年 8 月から 11 月にかけて行われた実験では成功を収めた。世界人口のおよそ半分が海から 200km の範囲で生活しているため、海中にデータセンターを設置することで多くの都市の近郊にデータセンターを設置したことになるというメリットがある。また、従来であればデータセンターの建物を建設するだけで 1 年以上かかるところを、モジュール型のデータセンターにすることで 90 日以内に設置することが可能になる。エンジニアが海中でメンテナンスすることはできないため、サーバーをメンテナンス無しで稼働させ、5 年ごとに中身のサーバーを交換する計画であるという。

図表 15 は、Microsoft 社の Project Natick となっている。

⁷⁰ <http://gizmodo.com/building-the-part-of-facebook-no-one-ever-sees-1580123711>

⁷¹ <http://gizmodo.com/amazons-building-a-solar-farm-in-virginia-to-green-its-1710563957>
<http://aws.amazon.com/about-aws/sustainability/>

<http://phx.corporate-ir.net/phoenix.zhtml?c=176060&p=irol-newsArticle&ID=2114239>

⁷² <http://www.theverge.com/2015/5/1/8527699/tesla-battery-amazon-target-for-renewable-energy>

図表 15: Microsoft 社の Project Natick



出典: arstechnica⁷³

6 終わりに

SNS や IoT(モノのインターネット)などの広がりによって、人々の行動そのものが大量のデータを生み出すようになってきた現在、ビジネスにおけるビッグデータの活用や、IoT・人工知能の活用が本格化すればするほど、データの重要性は益々高まることとなり、どれだけデータを集められるか、そして所有できるかが、企業の競争力につながってくる。

今後、データの所有に関する競争はさらに激化することが予想される一方で、活用されるデータが広範になるほど個人情報保護の問題も大きくなるであろうし、データの囲い込みをどうするかなど IT ビジネスが健全に発展するために解決しなければならない問題も重要になってくる。また、飛躍的に増大するデータを保持するデータセンターの設置も無視できない問題となる。

このような問題は、ビッグデータ、IoT などデータを活用したビジネスの裾野を支えるものとして、今後の産業発展において重要な要素であり、その動向を注視していきたい。

※ 本レポートは、注記した参考資料等を利用して作成しているものであり、本レポートの内容に関しては、その有用性、正確性、知的財産権の不侵害等の一切について、執筆者及び執筆者が所属する組織が如何なる保証をするものでもありません。また、本レポートの読者が、本レポート内の情報の利用によって損害を被った場合も、執筆者及び執筆者が所属する組織が如何なる責任を負うものでもありません。

⁷³ <http://arstechnica.com/information-technology/2016/02/microsofts-new-way-for-cooling-its-datacenters-throw-them-in-the-sea/>