

Suica に関するデータの社外への提供について

中間とりまとめ

Suica に関するデータの社外への提供についての有識者会議

2014 年 2 月

目次

1	はじめに	3
2	本会議の設置目的と開催状況	5
2.1	設置目的.....	5
2.2	本会議の開催状況.....	6
3	「7月提供の Suica 分析用データ」の総括	7
3.1	本事業の内容.....	7
3.2	本会議設置までの経緯.....	9
3.3	利用者の声.....	10
3.4	個人情報及びプライバシーの保護の観点からの検討.....	11
4	Suica データの社外への提供について	13
4.1	Suica データの有用性.....	13
4.2	匿名化した Suica 分析用データ.....	17
5	提言	20
5.1	今後について.....	21

1 はじめに

Suica データは、これを分析することにより、利用者による駅の利用状況やその構成を把握することができるので、地域や駅、沿線の活性化に資する、様々な分野で活用されることが期待され、利用者はもとより社会一般にとっても有用な基盤となると考えられる。

一方、Suica データを利活用する際には、特に利用者への配慮が必要になる。東日本旅客鉄道株式会社（以下「JR 東日本」という。）が 2013 年 6 月に、Suica 利用データから氏名、電話番号、物販情報等を除外し、生年月日を生年月に変換した上、さらに、SuicaID 番号を不可逆の別異の番号に変換した（図 3.1-2 参照）データを株式会社日立製作所（以下「日立製作所」という。）に提供（提供は 7 月に実施、以下「7 月提供の Suica 分析用データ」という。）することが明らかになり、多くの利用者から、個人情報の保護、プライバシーの保護や消費者意識に対する配慮に欠けているのではないかとして批判や不安視する声があがった。

そのため、Suica に関するデータの社外への提供に関し、利用者の理解を得ながら実現させることができるかを検討することを目的に、有識者会議（以下「本会議」という。）が設けられた。

本会議ではまず、JR 東日本の今般の対応のどこに問題があったかについて議論・整理した。そして、その整理の下、Suica データについて、有用性の検討と併せて、今後の社外への提供についても検討した。

本会議における検討事項は、次の二点にまとめることができる。

- ①「7 月提供の Suica 分析用データ」提供の問題の整理について
- ②今後の社外への提供について

各事項について十分な検討を行うため、本会議の場以外でも、技術、法律、統計、及び、消費者意識を始めとする社会的責任の観点（以下「CSR」という。）に関する各分野の専門家にヒアリングする等、幅広い専門領域の知見を集めた。ヒアリングの実施状況は以下のとおりである。

ヒアリングした専門家：大学教授・准教授 3 名、弁護士 5 名、その他専門家 4 名

ヒアリング期間：2013 年 10 月～2013 年 12 月

一方、ビッグデータやパーソナルデータの利活用については、政府の高度情報通信ネットワーク社会推進戦略本部（IT総合戦略本部）の下に設けられた「パーソナルデータに関する検討会」でも検討が進められており、2013年12月10日の第5回会合で「パーソナルデータの利活用に関する制度見直し方針（案）」が提示され、それを一部修正した案が、2013年12月20日に同本部で決定された。そこで示されたロードマップでは、この方針に基づいて詳細な制度設計を含めた検討を加速させ、検討結果に応じて、2014年6月までに、法改正の内容を大綱として取りまとめ、2015年通常国会への法案提出を目指すこととされている。

ここで示された、パーソナルデータの利活用に関する制度見直し方針は多岐にわたっており、本会議で検討してきたいくつかの論点への対応は、その動向をも注視して行う必要がある。

そこで、ここでは、本会議での議論を現段階での「中間とりまとめ」とすることにした。

本会議における議論が、Suicaデータの社会・産業に対する新たな価値の創造に繋がることを望むとともに、“くらしの安全・安心”に寄与し、健全なビッグデータの利活用の一助となれば幸いである。

2 本会議の設置目的と開催状況

2.1 設置目的

「7月提供の Suica 分析用データ」の提供について、多くの利用者から、事前の周知がなかったこと等から消費者意識に対する配慮に欠けているという批判の声や法的問題を指摘する声があがり、JR 東日本では、利用者への説明（ホームページへの FAQ 掲載やデータ提供除外（いわゆるオプトアウト）に関する案内等）を開始した。さらに、JR 東日本は、利用者の安心感を高めるために JR 東日本としてどのように取り組むべきかについて、社外の専門家の意見を採り入れることを決定し、2013年9月6日に本会議を設置した。

Suica に関するデータの社外への提供についての有識者会議 構成員（役職は就任当時）

座長：

堀部 政男 一橋大学 名誉教授

（就任期間：2013年9月6日～2013年12月31日）

座長代理：

高芝 利仁 高芝法律事務所 弁護士

委員：（50音順）

青山 理恵子 公益社団法人 日本消費生活アドバイザー・コンサルタント協会
副会長

藤原 静雄 中央大学法科大学院 教授

2.2 本会議の開催状況

本会議は、2013年9月6日より計5回開催された。

第1回会議（2013年9月6日開催）

- ・本会議の進め方について

第2回会議（2013年10月2日開催）

- ・国内外の動向（別紙資料参照）について
- ・専門家ヒアリングの進め方について

第3回会議（2013年11月14日開催）

- ・プライバシーの保護の動向について
- ・専門家ヒアリングの状況について

第4回会議（2013年12月26日開催）

- ・専門家ヒアリングの結果について
- ・「中間とりまとめ（案）」の作成

第5回会議（2014年2月13日開催）

- ・「中間とりまとめ」の承認

3 「7月提供の Suica 分析用データ」の総括

3.1 本事業の内容

JR 東日本では、2008年に中期経営構想「グループ経営ビジョン2020 ―挑む―」の取組の一環として、サービス向上や地域、駅、沿線の活性化を目的に、Suica データを活用した情報分析サービスの検討を開始した。JR 東日本の鉄道部門や生活サービス部門、JR 東日本グループ企業数社とも検討を重ねた結果、具体的には、Suica データから、利用者の駅利用状況や移動、決済等の状況を分析・把握することで、鉄道サービスや駅構内の店舗運営サービス等の改善に活用できることがわかった。

そこで、JR 東日本は、情報分析サービスを本格的に推進するため、2013年4月に「情報ビジネスセンター」を新たに立ち上げた。（「図 3.1-1」参照）

情報ビジネスセンターで分析するデータは、鉄道サービス等を提供する JR 東日本の業務セクションから提供されたものであり、その Suica データは、氏名や電話番号の情報を削除し、生年月日を生年月に変換したデータ（以下「Suica 分析用データ」という。）である。（「図 3.1-2」参照）

一方、Suica データを活用したいという要望を受け、最終成果物が上記の目的に合致することを前提に、日立製作所に対して Suica データを提供する検討を行った。JR 東日本は、日立製作所がビッグデータに関する技術ノウハウを有した様々な業種・業態を擁する企業であり、その技術ノウハウと Suica データを組み合わせることで JR 東日本単独では見出せない大きな価値を創造する可能性を持っていると考えた。そこで、日立製作所において、データが厳格に取扱われることを確認し、個人の特定を行うことを契約で禁止するとともに、物販情報等を削除し、SuicaID 番号を不可逆の別異の番号に変更したうえで、2013年7月に日立製作所に「7月提供の Suica 分析用データ」の提供を行った。（「図 3.1-2」参照）

しかし、多くの利用者から、今回のデータ提供の枠組み・対応について不安や心配の声があがったことから、現在は、データ提供を停止しており、かつ、既に提供した「7月提供の Suica 分析用データ」についても、日立製作所において、これを抹消している。

また、データ提供の再開については、本会議の結果を踏まえながら、JR 東日本が今後決定することとなっている。

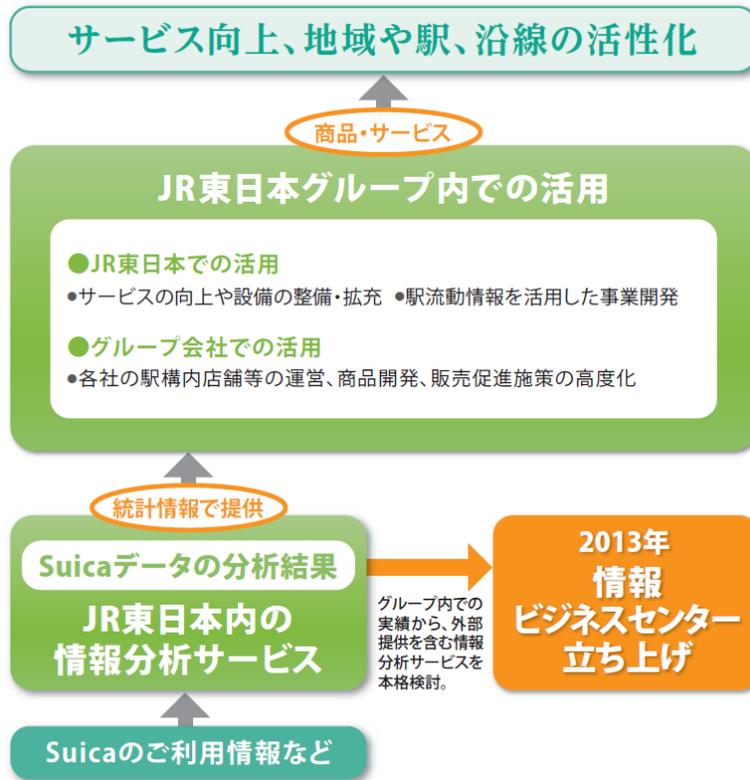
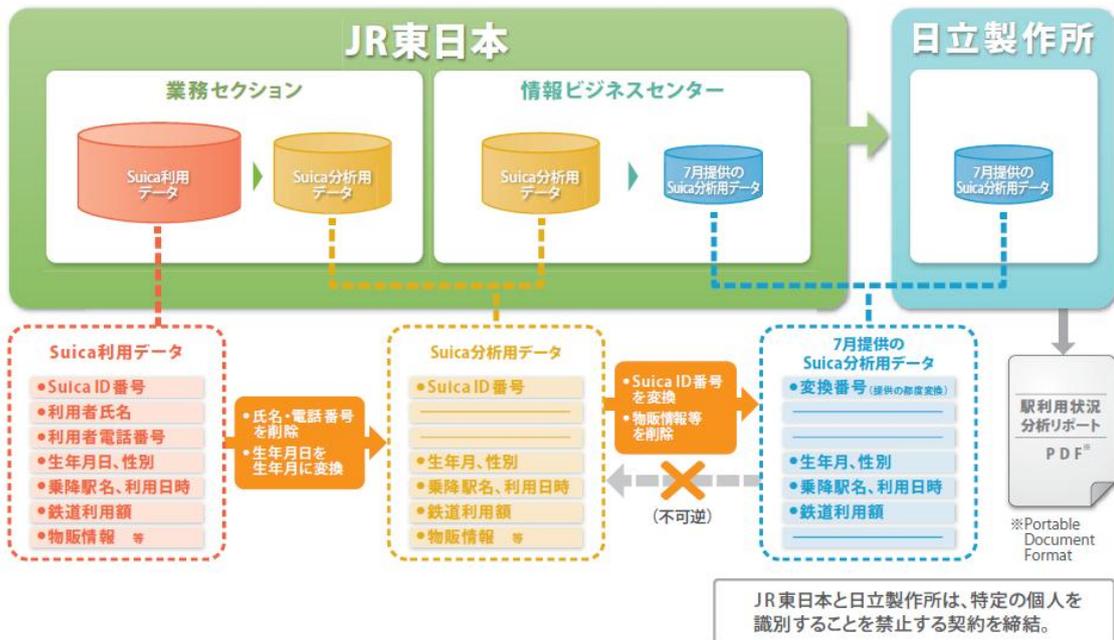


図 3.1-1 (JR 東日本提供)



JR 東日本と日立製作所は、特定の個人を識別することを禁止する契約を締結。

図 3.1-2 (JR 東日本提供)

3.2 本会議設置までの経緯

「7月提供の Suica 分析用データ」を、日立製作所に提供することは、先ず日立製作所の2013年6月27日のプレスリリースによって、公表された。

このプレスリリースに対して、利用者やマスコミから、日立製作所へのデータ提供は許されるのか、本当に個人が特定できないか、周知不足ではないか、といった指摘があり、また、インターネット上でも同様の批判の声が大きくなり、本会議が設置されるに至った。（「図 3.2-1」参照）



図 3.2-1 (JR 東日本提供)

3.3 利用者の声

利用者から JR 東日本への声の件数・主な内容は以下のとおりであった。

- ・利用者の声の受付件数（2014 年 1 月 31 日現在）

問い合わせ先	件数
・除外要望受け付け窓口	2,741 件
・JR 東日本ホームページ	440 件
・ご意見承りセンター	218 件

- ・主な問い合わせの内容について

今回のデータ提供の枠組み・対応について

日立製作所へのデータ提供は許されるのか？
本当に個人が特定されないか？
提供データ項目は適切か？
周知不足ではなかったか？
社外へのデータ提供除外（いわゆるオプトアウト）受付といった対応は妥当か？

JR 東日本のビジネス全体について

JR 東日本がこのようなビジネスを展開すべきか？
利用者にメリットはあるか？
事業に不透明さがあるのでは？

3.4 個人情報及びプライバシーの保護の観点からの検討

本章では、日立製作所に対して提供した、「7月提供の Suica 分析用データ」について、個人情報及びプライバシーの保護の観点から検討する。

3.4.1 検討結果について

JR 東日本は、日立製作所に提供した「7月提供の Suica 分析用データ」が、提供した時点では、一定レベルの匿名化処理がされ、特定の個人を識別することができない状態であると評価できたので、個人情報保護法の「個人情報」に当たらず、同法の適用を受けないと考え、提供した。

すなわち、個人情報保護法における「個人情報」とは、「生存する個人に関する情報であつて、当該情報に含まれる氏名、生年月日その他の記述等により特定の個人を識別することができるもの（他の情報と容易に照合することができ、それにより特定の個人を識別することができることとなるものを含む。）」（法第2条第1項）と規定されているところ、「7月提供の Suica 分析用データ」は、氏名、電話番号、物販情報等の情報を削除し、生年月日を生年月に変換した上、さらに、SuicaID 番号を不可逆の別異の番号に変換して提供していたことから、JR 東日本では、個人情報保護法の「個人情報」に当たらないと解した。

また、プライバシー保護の観点においても「7月提供の Suica 分析用データ」は、上記のとおり一定レベルの匿名化処理がされ、特定の個人を識別することができなくしていることに加え、日立製作所との間で、特定の個人を識別することを禁止する契約を締結していたこと、また、従来、特定の個人を識別できない情報についてプライバシー侵害は問題とならないと解されていることから、法的に問題となることはないと考えていた。

しかし、一部の専門家から Suica データを継続して提供すると個人が識別でき、他の情報との照合によって個人の特定につながるのではないかと等との指摘がなされた。

以上を踏まえ、本会議で確認した結果は、次のとおりである。

(1) JR 東日本が現行法において、合理的と考えられる範囲での法の解釈運用に努めていたことは認められるが、個人情報の定義における特定の個人の識別性の論点については、専門家の間でも解釈に幅があり、また、現在、法改正が検討されていること等の状況にあるため、今後の立法化の動向にも注視していく必要がある。

(2) また、プライバシーの保護については、一定レベルの匿名化処理がされ特定の個人を識別することはできなくしていることに加え、JR 東日本と日立製作所は、特定の個人を識別することを禁止する契約を締結していたことが認められるので、直ちにプライバシーが侵害されるおそれはないと判断されるものの、今後、技術の進展に伴い、特定の個人が識別され新たな問題が生じる可能性も考えられるため、今後とも継続して最善と考えられる配慮を行うべきである。

(3) 加えて、利用者から、プライバシーの保護や消費者意識に対する配慮が十分ではないという批判・不安の声があがったのは、移動履歴という利用者にとって関心の高い情報を提供するにあたって、事前に十分な説明や周知を行わなかったことによると考えられる。

(4) また、データが提供されることを望まない利用者への対応（いわゆるオプトアウト）等についても、2013年7月に日立製作所に「7月提供の Suica 分析用データ」を提供した段階では配慮が不足していたと考えられる。

(5) 「7月提供の Suica 分析用データ」は、オープンデータのようなインターネット上で広く公開する類のデータ提供ではなく、利用制限を定めた契約を締結したうえで提供したものであるため、データが外に出るといったようなことがない限り、プライバシー侵害は発生しにくいと考えられるものの、利用者に不安を与えた事実を重く受け止める必要がある。

4 Suica データの社外への提供について

3章（「7月提供の Suica 分析用データ」の総括）で述べたとおり、「7月提供の Suica 分析用データ」では、直ちに個人のプライバシーが侵害されるおそれはないと判断される一方、今後、技術の進展に伴い、特定の個人が識別され新たな問題を生じる可能性も考えられる。そこで本章では、改めて、さらに安心な Suica データの社外への提供について検討する。

4.1 Suica データの有用性

JR 東日本は、自社サービスの品質向上を目的に、様々な取組を実施してきた。例えば、案内サインの設置や駅構内店舗のサービス改善の効果の検証、あるいはダイヤ改正や駅改良による利用者流動の変化の検証等を「Suica 分析用データ」を活用して行うことで、より詳細・広範に、かつ効率的に行うことが可能になった。

分析活用イメージ ～効果的な案内サインの検討～

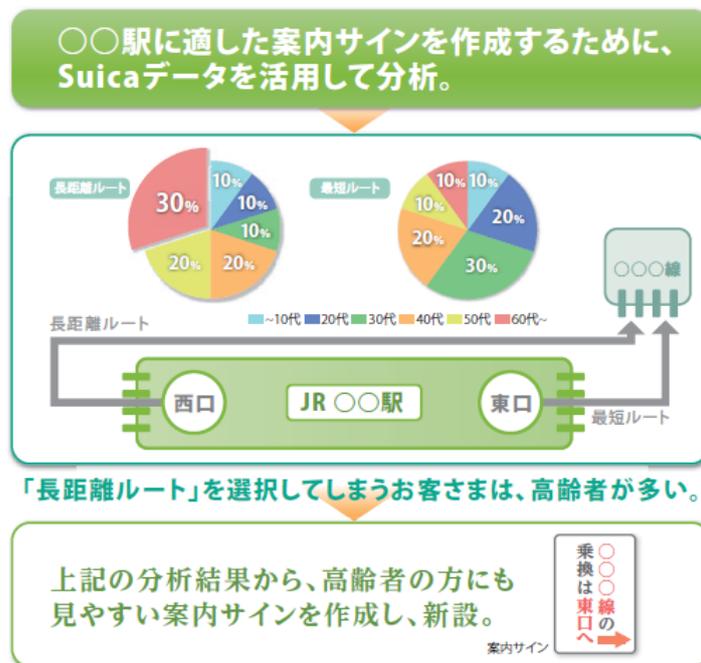


図 4.1-1 (JR 東日本提供)

分析活用イメージ ～駅設備の検討～

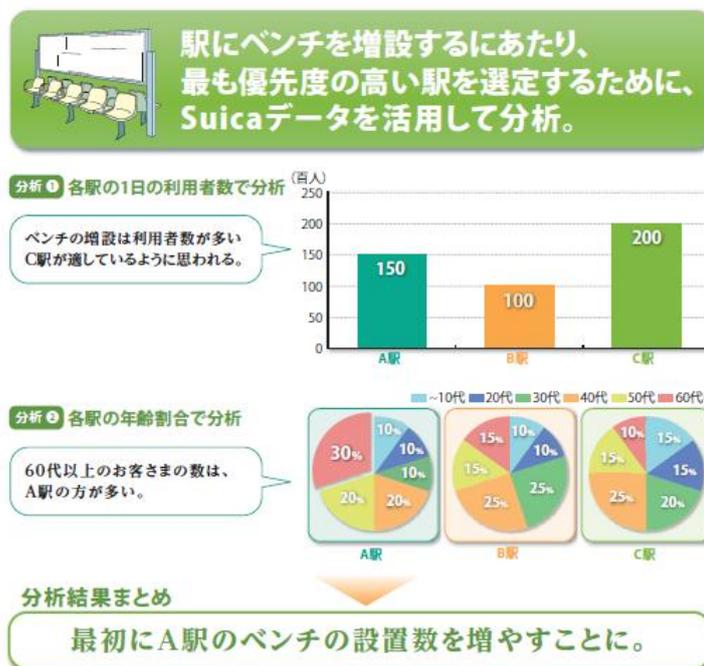


図 4.1-2 (JR 東日本提供)

さらに「Suica 分析用データ」を活用した予測モデルが構築できれば、利用者の潜在的な要望を把握することができ、的確なサービスを創出できるのではないかと考えられる。

また、JR 東日本でこれまで検討してきた自社サービス向け以上の付加価値を生み出すためには、JR 東日本社内にはない、他のノウハウとの融合が不可欠であり、Suica データの社外への提供は重要になってくる。

その際には、JR 東日本において、図 3.1-2 中の「駅利用状況分析レポート」(以下「駅リポート」という。)のような統計情報を作成して提供するか、「匿名化した Suica 分析用データ(4.2「匿名化した Suica 分析用データ」参照)」を提供することになるが、後述のとおり、「匿名化した Suica 分析用データ」の場合、一律に k-匿名化¹処理を実施すると相当数の利用者のデータと特質を消失することも考えられ、この点が社外提供にあたっての課題になると考えられる。

¹ あらゆる観点でデータを絞り込んでも、分解能が k 名までに制限され、特定の個人を識別できないようにする匿名化。詳細は別紙資料 4.1「匿名化、暗号化状態での検索技術」を参照。

4.1.1 Suica データを社外提供する際の課題

JR 東日本が実施した、以下の二つの分析結果を用いて、社外提供に関する課題について検討を行った。

- ・急速に発達する低気圧（いわゆる爆弾低気圧）発生時の利用者の移動の把握
- ・外国人の観光ニーズの把握

4.1.1.1 急速に発達する低気圧（いわゆる爆弾低気圧）発生時の利用者の移動の把握

「Suica 分析用データ」を分析することにより、2012 年 4 月 3 日の爆弾低気圧発生時、利用者の移動の状況がどのようなであったかを、「統計データ」として作成した。具体的には、中央線の高尾・武蔵境駅間の駅から乗車し、23 区内の駅で降車した利用者の動きについて、爆弾低気圧発生時と平常時の動きを把握し、比較（図 4.1.1.1-1 参照）を行った。

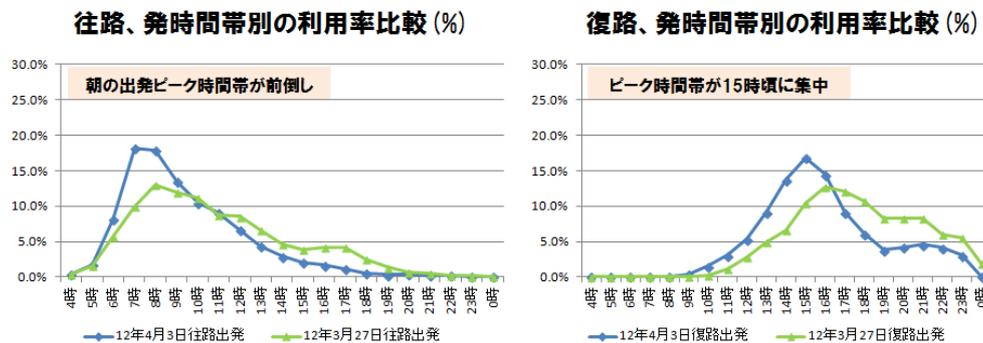


図 4.1.1.1-1 (JR 東日本提供)

図 4.1.1.1-1 からは、朝の出発ピーク時間帯が前倒しになっていたことと、復路について、東京都からの早めの帰宅圏を受けて、ピークとなる移動が 15 時頃に局所的に集中していることがわかる。このような情報を、他交通機関等の関係機関に対して提供することで、今後、同様の事象が発生した際、方面ごとの効果的な利用者案内への活用が期待される。

このように、「匿名化した Suica 分析用データ」を活用すると、利用者の属性等を考慮した高度な分析も可能となるが、一方、k-匿名化処理により、相当数の利用者のデータが使用できなくなり、利用者全体の動きを正確に把握できなくなるという課題も認められた。

4.1.1.2 外国人の観光ニーズの把握

訪日外国人向け商品である「Suica & N'EX」のデータにより、外国人利用者に人気の都内の目的地を分析した「統計データ」を作成した。人気の目的地を把握することで、訪日外国人向けの様々な施策検討が行えるようになると考えられる。(図 4.1.1.2-1 参照)

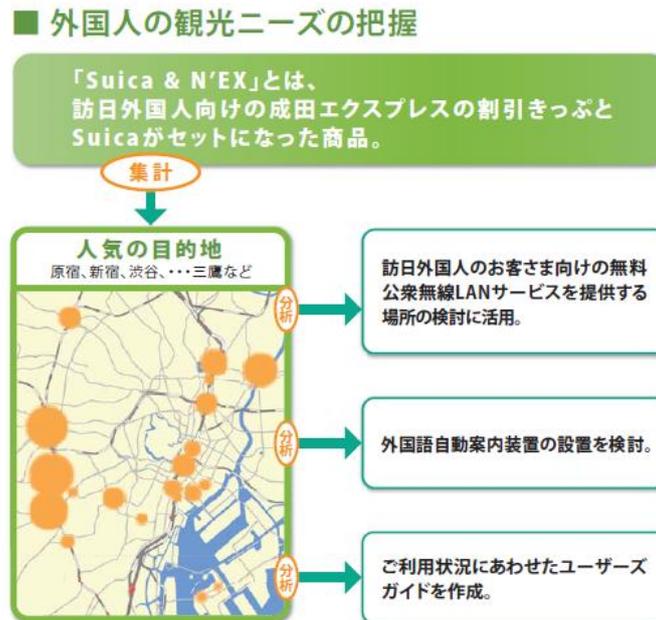


図 4.1.1.2-1 (JR 東日本提供)

オリンピックのような来るべき大規模イベントに備え、このような「統計データ」を活用し、多言語での案内表示や、利用者向けの周辺施設紹介パンフレットの作成等、外国人利用者が迷わず、気持ちよく利用できる駅の実現に向けた施策検討を行っていく必要があると考えられる。

「匿名化した Suica 分析用データ」を作成することにより、駅周辺の商業施設・関連企業に対して、マーケティングデータを提供することが可能になると考えられるが、一方、k-匿名化処理により、中小規模の駅を利用する相当数の利用者のデータが使用できなくなり、特に地域の商業施設・関連企業に対して、価値のあるマーケティングデータを提供できなくなるという課題も認められた。

4.2 匿名化した Suica 分析用データ

「7月提供の Suica 分析用データ」は、氏名・電話番号の識別情報や物販情報等を削除する等の単純匿名化の処理を行っており、日立製作所はこのデータに統計処理を施し、駅の利用者の性別年代構成、滞在時間、乗降時間帯等を、平日・休日別に見える化する駅レポートを作成した。そこで、駅レポートを作成でき、かつ外部へのデータ提供が可能と考えられ、さらに CSR の観点からも、より安心・安全と考えられる集合匿名化を行った場合、どのようなデータを作成できるかを JR 東日本で評価した。

具体的には、政府の「パーソナルデータに関する検討会」や、株式会社 NTT ドコモ（以下「NTT ドコモ」という。）の「モバイル空間統計」の報告で、基本的な集合匿名化の手法として用いられている k-匿名化処理を施した Suica データでの評価を実施した。なお、この評価については、技術専門家へのヒアリングの結果も踏まえている。

4.2.1 匿名化した Suica 分析用データの作成フロー

利用者の乗降データから、氏名・電話番号の識別情報や不要項目の削除、曖昧化、k-匿名化処理(図 4.2.1-1 参照)を行うことにより、「匿名化した Suica 分析用データ」を作成した。

k-匿名化(集合匿名化の方法の一つ)の例



図 4.2.1-1 (JR 東日本提供)

4.2.2 今回作成した、匿名化した Suica 分析用データ

今回の評価で作成した「匿名化した Suica 分析用データ」は、利用乗降駅を内容とする「移動データ」（表 4.2.2-1 参照）と、降車から乗車までの滞在を内容とする「滞在データ」（表 4.2.2-2 参照）の二種類で、この二種類のデータを利用することで、駅の利用者の性別年代構成、滞在時間、乗降時間帯等を、平日・休日別に把握でき、4.2 冒頭で述べたような駅リポートを作成することができる。

「移動データ」（例）

乗車駅	乗車時間	降車駅	降車時間	日時	年齢	性別	データ数	k
A駅	AM8時台	B駅	AM9時台	2013/10/18	50代	女性	15	15
C駅	AM9時台	D駅	AM9時台	2013/10/18	30代	女性	7	7
E駅	PM1時台	F駅	PM1時台	2013/10/18	20代	男性	30	30

表 4.2.2-1（JR 東日本提供）

「滞在データ」（例）

降車駅	降車時間	乗車駅	乗車時間	日時	年齢	性別	データ数	k
G駅	AM9時台	G駅	PM5時台	2013/10/18	30代	女性	16	16
D駅	AM9時台	H駅	AM11時台	2013/10/18	20代	女性	6	6
F駅	PM1時台	I駅	PM3時台	2013/10/18	50代	男性	20	20

表 4.2.2-2（JR 東日本提供）

4.2.3 課題

4.2.1 で駅リポートの作成を目的に「匿名化した Suica 分析用データ」を作成したが、k-匿名化処理によって、相当数の利用者のデータが使用できなくなり、分析結果の精度が低下したり、分析の種類が制限される課題があることがわかった。このような課題を解決するには、データ特性を勘案し、利用者が安心・納得できる土壌づくりを踏まえた、技術の進展あるいは法整備等が望まれる。また、「匿名化した Suica 分析用データ」を外部に提供する形態ではなく、JR 東日本で統計処理を施して駅リポートを作成し、外部に提供することも、課題解決の方法の一つと考えられる。

5 提言

本会議では、本報告書にこれまで記載したとおり、利用者に、自己のプライバシーが侵害されているのではないかという不安を与えた、「7月提供の Suica 分析用データ」の日立製作所への提供の問題の整理と、利用者がより安心と感じる「今後の社外への提供」について検討を行った。

・「7月提供の Suica 分析用データ」提供の問題の整理と今後の社外への提供について

「7月提供の Suica 分析用データ」の日立製作所への提供については、利用者への配慮が欠けていたことが問題であり、JR 東日本という公共性の高い企業の立場からも、その配慮不足・未熟さに対する反省が必要である。

JR 東日本が、現行法において合理的と考えられる範囲で法の解釈運用に努めていたことは認められるが、個人情報定義における個人の識別性の論点については、専門家の間でも解釈に幅があり、また、現在、法改正が検討されていること等の状況にあるため、今後の立法化の動向にも注視していく必要がある。

また、プライバシーの保護についても、直ちに個人のプライバシーが侵害されるおそれはないと判断される一方、今後、技術の進展に伴い、特定の個人が識別され新たな問題が生じる可能性も考えられるため、今後とも継続して、最善と考えられる配慮を行うべきである。

今回の問題で経験したことを活かし、法改正や技術の進展等の社会の動向を注視しながら、利用者が安心・納得できるようなデータ提供のあり方を、多角的な観点から検討すべきである。

・「統計データ」及び「匿名化した Suica 分析用データ(4.2「匿名化した Suica 分析用データ」参照)」について

「統計データ」や、氏名・電話番号の識別情報や不要項目の削除、曖昧化、k-匿名化処理による集合匿名化を施した「匿名化した Suica 分析用データ」は、社会通念上、個人の特定を不可能とし、特定の個人の行動履歴を把握することができないため、これらの提供は、個人情報の保護及びプライバシーの保護の観点から問題となることはないと考えられる。

また、個人の行動を個別に把握できなくすることにより、利用者が、プライバシー侵害についての不安を感じるリスクもより小さくできると考えられるので、CSR の

観点からも望ましいものである。

ただ、「匿名化した Suica 分析用データ」は、k-匿名化処理による集合匿名化によって、移動データや滞在データ等の動線データが本来有するデータ量・特性価値（行動特性の把握等）が相当程度失われることもまた事実である。動線データの特性を活かした匿名化の技術手法については、未だ、公的統計の分野でも確立されていないところがあるが、Suica データは、ビッグデータ時代において、公益的な価値や有力なビジネスを生み出す可能性のある有用なデータである。

・ JR 東日本に期待される役割

JR 東日本としては、今後の法改正や技術の進展等の動向に注視するだけでなく、動線データの特性を活かしたビッグデータの利活用が促進され、また、ビッグデータが生み出す価値が社会に理解されるよう、例えば、公益性の高い統計情報を公表する等の積極的な活動を行っていくことが望まれる。Suica から生み出される巨大な鉄道 IC データは、世界でも類を見ない先進的な分野であるので、日本の優位性を活かす貴重なデータとして、我が国のビッグデータの利活用の競争力を牽引する役割を担うことが期待される。

5.1 今後について

4 章（Suica データの社外への提供について）に記載したとおり、「統計データ」や「匿名化した Suica 分析用データ」は、外部に提供されることによって、その価値を発揮することが期待されるが、「7 月提供の Suica 分析用データ」での反省を踏まえ、今後の提供にあたっては、利用者がより安心・納得できるようにするための積極的な広報活動が必要である。今後、「統計データ」や「匿名化した Suica 分析用データ」の提供を検討するにあたっては、利用者への配慮において、他の民間企業の取組（別紙資料 5「事例）NTT ドコモモバイル空間統計」）等も参考としながら、「統計データ」や「匿名化した Suica 分析用データ」の内容、作成ルール、運用管理体制等を利用者に周知し、利用者が安心・納得できるような配慮を行っていくことが肝要となる。

今後、JR 東日本は、本会議の内容を踏まえ、利用者に配慮したビッグデータの利活用を検討していくことになるが、その際、海外のプライバシーの保護の動向や、我が国における法改正の動向にも注視しながら検討を進める必要があるため、本会議も、必要に応じて JR 東日本に対してアドバイスを行っていくものとする。