

令和2年度「不動産経済分析研究会」議事概要

テーマ 「コロナショック後の人出変動と企業倒産：ロケーションデータと倒産データを用いた実証分析」

1. 一橋大学 准教授 宮川大介氏の説明概要は以下の通り。

[大きなテーマ]

- ・ ビックデータを使って、いかに経済状態を計測していくか。

[要点]

- ・ 非伝統的データであるビックデータを用いた計測、予測、因果推論を踏まえて意思決定を高度化する。
- ・ 本報告では、理論のガイドの下で複数の適切なデータを用いた分析の事例として、コロナショックによって企業の事業継続や市場からの退出（企業ダイナミクス）にどのような影響が生じているかを検討する

[セットアップ]

- ・ COVID-19の感染拡大がある中で企業が事業継続に関する意思決定をどのように行うのかを考えたい。
- ・ アメリカ、イギリス、日本に関して、横軸に時間、縦軸に倒産企業の件数をプロットしたのを見ると、1つ不思議なことに気づく。COVID-19が企業の継続に何らかの影響があるのであれば、倒産増加の動きがあつてしかるべきだと思われる。ところが、倒産件数で見ると、むしろ減っているということがいろいろな国で見ても分かる。日本に関しては、一旦下に下がってもう一回元に戻っているが、それ以降、倒産件数は減少しており、いろいろな視点から見ていく必要があるだろうということに気づく。

[二つの視点]

- ・ COVID-19の感染拡大の動向そのものを見ることに代えて、付随して生じた人流の変動（Mobility）を捉えることで、感染拡大の影響を捉える上での代理変数（Proxy）として用いる。
- ・ Mobilityを計測した1つの例として、グーグルのモビリティレポートというものが公開されている。これは、世界各国である程度の行政区分で区切られたエリアにお

ける 2020 年 1 月の後半対比での人出の変動を数値で表しており、日本の場合では、都道府県のメッシュと 6 カテゴリーに関する混雑度合の変動を計測している。

例えば、**Retail&recreation** というのがあり、この中には小売店とかレストランといったものが入っており、**Grocery&pharmacy** はスーパーマーケットとか、ドラッグストアといったところが含まれ、このほかにも住宅とかオフィス、公園等、いろいろなカテゴリーがある。COVID-19 の様相を、経済的にどのようなことが起きているのかという視点から把握するには使いやすいデータとなっている。

- COVID-19 のショックを計測した代理変数がグーグルモビリティから取得できたので、ショックが行き着く先の企業の様子として倒産を見ることにする。
- 使用する倒産データは、東京商工リサーチの情報。
- 人流が変化したときに、倒産の確率がどの程度変動するのかを見てみると、コロナショック直後の時期に関して言えば、ショックにひもづけられた人流の変動が倒産とリンクしていたことが分かる。
- この点を確認するために、別種の適切なデータとして、延滞にも注目する。同様に、人流の滞りが延滞の増加につながっていることが分かる。
- ここまでは、データを使ってショックのダイレクトな影響をみてきたが、もう複雑な経路についても検討してみる。
- COVID-19 に対しては、行動制限政策などの影響を受けた人流の変動が企業の業績へ影響するという経路とは別に、例えば、取引先の業績が滞ることが自社に対して波及するという経路も想定される。この点については、自社所在地だけではなくて、取引先が所在するエリアのアクティビティが自社の業績に影響するという形で確認されている。
- 企業の退出に関するセレクションメカニズムについて、コロナ前の時点で業績が悪かった会社が、コロナ禍初期においては退出したという様子が見られる。
ただし、その後については、様相がだんだん変わってきており、セレクションが弱くなってきている。
- このように個別のメカニズムに関する分析結果をまとめると、全体の把握ができ、本来的なメカニズムを理解できると考えられる。

- ・ 予測と現状

理論モデルとデータに基づいて、コロナショックに伴う業績の悪化だけが生じた場合にどの程度の企業退出が生じたかをシミュレートしてみると、廃業はほとんど予測どおりの件数に増加している。7月に公表したベンチマークモデルでは、年度の退出件数が前年対比で2割増という数字をシミュレーションから得ていた。現状では休業に関してはほぼ2割増。一方で、倒産は前年より少ないという状況にある。つまり、借金の負担が小さい企業は、自社の業績見通しが芳しくないことを踏まえて休業の形で退出しているが、負債がある企業は潤沢に借入れをして生き延びている。

[まとめ]

- ・ 計測・予測・因果推論というのがデータに向き合うときの1つのスタイルである。これらをうまく使うためには、適切でかつ複数のデータを使い、かつ何らかの理論的ガイドを基にして検討することが重要である。データの整備については、すぐにオルタナティブ・データを使いたがる向きもあるが、構造化されてストックされているデータが使われずに残っているケースが多く存在している。典型例が税務のデータと通関のデータであるが、そういうものを発掘してほかのデータと組み合わせていくことが、効果的な利用につながるのではないかと考えている。

テーマ 「コロナショックが消費行動に与えた影響：店舗・観光に与えた影響を中心として」

2. 東京都市大学 准教授 秋山祐樹氏の説明概要は以下の通り。

[研究背景と目的]

- ・新型コロナウイルスの蔓延とそれに伴う外出自粛要請等より、外出する人出が各地で減少した。
- ・その結果、全国的に消費地における消費（小売・飲食等）の落ち込み、それに伴う企業の倒産などが明らかとなりつつある。
- ・一方、どこでいつどの程度、消費の増減が起きているか、ということについてはまだ明らかにされていない。
→その増減の空間的分布を高精細に把握することができれば、国・地方自治体としてどの地域が特にコロナウイルスによる影響を大きく受けているのか、どの地域にどの程度の経済的支援を実施する必要があるのか、などを検討することが可能になる。
- ・しかし、以上の現象を把握するためには、膨大な数の人々の動きを連続的（時系列的）に把握し続けることで、人出の増減を把握し、人々の外出に伴う消費額の増減を推定する技術を開発する必要がある。
- ・そこで、LocationMind 株式会社の人流ビッグデータ（LocationMind xPop）を活用し以上の目的達成を目指す。

[手法の概要：全体の流れ]

- ・GPS データや各種統計を用い、LocationMind xPop 滞留点データ（2019 年、2020 年（自宅・勤務地以外））を組み合わせることで、全国 500m メッシュごとの推定消費額（2019 年、2020 年）を作成する。

[手法の概要：各ユーザの年収推定と消費地における推定消費額]

- 1) xPop の元データから各人の居住地、勤務地、その他の滞留点を算出する（NTT ドコモにて統計処理）。
→このうち、外出に伴う消費地での消費は「その他の滞留点」で実施されると仮定。
- 2) 各ユーザの居住地をベースに各ユーザの年収を推定する。
→総務省統計局「労働力調査」（2019 年）より、2019 年の市区町村ごとの平均年収を収集可能。
※なお、町丁目や地域メッシュ単位の推定平均所得のデータは有償だが利用可能。

例：ESRI (<https://www.gisdata-store.biz/product/1965/>) 全国メッシュで 70 万円。

- 3) 以上の年収を全て消費地での消費に費やすわけではない。以下の消費・資金の流れが発生するはずなので、それらを年収に応じて推定し年収から差し引くことで、各人の消費地での年間消費額を推定する。
 - ・ 年収別 e コーマス（電子商取引）平均消費額→家計消費状況調査平成 29 年改定（2015 年 1 月～）
 - ・ 年収別貯蓄率
→家計の金融行動に関する世論調査（2 人以上世帯）（金融広報中央委員会）（2019 年）：年間手取り収入（臨時収入を含む）からの貯蓄割合（金融資産保有世帯）
 - ・ 年収別平均年間返済額
→家計の金融行動に関する世論調査（2 人以上世帯）（金融広報中央委員会）（2019 年）：年間手取り収入（臨時収入を含む）からの貯蓄割合（金融資産保有世帯）
- 4) 以上の年収を全て消費地での消費に費やすわけではない。以下の消費・資金の流れが発生するはずなので、それらを年収に応じて推定し年収から差し引くことで、各人の消費地での年間消費額を推定する。
 - ・ 年収ごと金融資産に回す割合
→日本証券業協会・平成 30 年度証券投資に関する全国調査 月々の収入から金融商品にまわす割合
 - ・ 家賃 or 住宅ローン返済→一括 25%と設定
 - ・ 光熱費→家計調査 2019 可処分所得に対する光熱費の占める割合
- 5) さらに以上で得られた消費額を月ごとに按分する。
 - ・ 月ごとの消費額は「地域別消費総合指数 2019 月別（総務省）」で都道府県ごとに把握可能。

以上の処理により、各ユーザの消費地における推定消費額を月ごとに推定できる。

[メッシュ単位（500m）年間推定消費額（2019 年）]

- ・ 関東、近畿、中京圏とか、いかにも人が多くいそうな場所で多額のお金が使われているというのが分かる。

[週毎の推定消費額]

- ・ 全国の総計では、1 月、2 月ぐらいまでは 2020 年のほうが高くなっているが、3 月ぐらいからどんどん減ってきて、特に 4 月、5 月辺りが相当落ち込んでいる。大体昨年度の 2 割弱ぐらい減ってしまっているような傾向が出ている。一回 6 月ぐらいで持

ち直すが、また7月、8月と落ちていくといったような結果になっている。

- ・日経新聞掲載の消費支出のグラフと、傾向が似ているというような結果も得られた。

[メッシュ単位 (500m) 消費額の比較 (2020年/2019年)]

- ・全国的には、1月はまだ、2020年のほうが2019年の消費額よりも僅に大きいが、そんなに傾向は変わっていない。3月になると全体的に消費額が下落し始め、徐々に4月ぐらいになるとさらに落ち込み、特に東京の都心辺りが顕著である。5月ぐらいになると一旦消費額は上昇し、特に都市の郊外辺りで上昇してくるが、中心市街地は相変わらず下落傾向で、郊外でやたらと上昇傾向が見られる。これはまだ検証中ではあるものの、恐らく自粛でみんなゴールデンウィークに帰省しなくなって、都市の東京とか大阪の郊外にとどまっているので、例年であればそこで使われないお金が使われていると推測される。しかしながら、総額で言うと5月はかなり減少している。あとは6月、7月と見ていくと面白いが、また8月は5月と結構同じような傾向が出ており、これも恐らく旅行控えとか、帰省控えで、例年だとお金が使われていないような都市の近郊でお金が使われているのではないかという結果が出ている。

[主な繁華街・観光地における月別消費額の推移 (2019年・2020年1~8月)]

- ・有楽町～銀座

1月、2月はほとんど変わらないが、3月から減って行って、4月、5月辺りは相当落ちているような傾向が見られる。これが7月、8月ぐらいになってもまだ全然回復してきてないということになっている。

- ・梅田駅周辺

- ・新宿歌舞伎町周辺

有楽町～銀座と同様の傾向で、4月、5月辺りの落ち込みは相当激しい。

- ・札幌市すすきの周辺

有楽町～銀座と同様の傾向で、4月、特に5月の回復が銀座辺りに比べてもあまりよくないということも出ていおり、夏、7月、8月になっても全然戻ってこない。

- ・仙台市国分町周辺

4月はかなり落ちているが、意外と7月、8月ぐらいになるとだんだん戻ってきているというような結果が出た。

- ・福岡市中州周辺
仙台市国分町周辺と同様の傾向。
- ・盛岡市大通周辺（暫く感染者が発生しなかった地域）
他の都市の繁華街とほぼ同じような傾向。盛岡でもしっかり自粛がなされていたという結果が出ている。
- ・東京ディズニーシー周辺
2月いっぱい営業自粛により閉鎖。6月まで閉鎖していたので、3月、4月、5月はほぼゼロに近い。6月から少しずつ復活し始めているという結果が出ている。
- ・伊勢神宮（内宮）周辺
伊勢神宮の前のおはらい町とかおかげ横丁と言われる場所であるが、こちらも1月、2月辺りは昨年と変わらないが、3月ぐらいから落ち始めて、5月はほぼゼロに近いぐらいに落ちている。6月で一回持ち直してくるが、例年の夏のレベルまではまだ戻ってきていないという結果が出ている。

[今後の課題]

- (1) 2019年の年間合計値が統計値（商業統計）とずれる理由の検討
 - ・POIによる消費額の違いや、場所による消費額の違いが考慮されていないため。
 - ・インバウンドによる消費が考慮されていないため。
 - ・統計値が古い（2014年）ため。
 - ・統計値は年間販売額であり、店頭での購入以外の販売額も含む可能性があるため。などが考えられる。店頭での小売販売額を把握できる統計や、より新しい時期の統計（メッシュ集計）があると良い。
- (2) 2020年の消費額の確からしさの検証方法の検討
何か良い方法はないか。マクロ的には消費動向調査などとの比較が可能か。
- (3) 消費行動の細分化の検討
 - ・買い物、飲食、宿泊、観光別の消費額の推定
→移動履歴、移動距離、消費地の属性（業種構成等）などを用いて推定する手法を開発中。
- (4) 年収推定方法の高精細化
 - ・現状は市区町村単位だが、有償のデータを使えば町丁目毎での推定が可能になる。

3. 清水主査 総評

- ・データ解析の意義と限界を知ることは非常に重要なことである。これは政策当局にとっても、業界や企業の経営でも今後において一層重要性は増していくものと考えられる。そのときに我々が考慮しておくべきことは、データで見えることと見えないことがあるわけで、その見えることは一体何なのか、その裏側にある見えないこと、実はそれが本質なのですが、それが何なのかを類推していくことの価値を知っておくべきであるということである。
- ・世の中ある意味いい時代になってきたと言われる。ビッグデータの定義としては、データが大きい、ボリュームが非常にあること、多様であること、そして即時性が高く備わっていることだとよく言われているが、不動産の分野も実はデータがかなり蓄積されてきた産業であった。近年においては、そのデータの空間の粒度とか時間の粒度みたいなものがかなり細くなってきた。その中でそのようなデータが可視化されることによって、行動の変容がおこったり、意思決定が変わったり、消費者の行動が変質することによって、不動産市場が高度化・効率化してきているわけで、我々はそれをどういうふうに一層進化させていくのかということを考えているところまで来ている。
- ・データが資源化してきている。今まで私自身 30 年ぐらいこういうデータを扱ってきたが、企業に眠るデータは、実は大量にあるということを学んできた。そういうデータをまさに発掘するということを示されたが、例えば某会社にも 1990 年代にビッグデータがあったが、磁気テープという古いデータ保管媒体で記録されて倉庫に眠っていた。その倉庫のどこにあるんだと言っても分からないと言うので、都内の倉庫にあるのを発掘した。さらに、その磁気テープを読み込むことができる機器がないということで、大学の研究室に持って行って、500 本の磁気テープのデータを一本ずつ落としてデータを作ったのが 1996 年だったと思う。
- ・国の統計でも、旧・国土庁では、1986 年からの不動産の移転登記情報をずっと集めていた。土地利用計画法に基づくもので、監視区域を指定する、排除するため使われてきたわけであるが、世の中の取引を全部データ化していたというようなことはきっと日本ぐらいしかやっていなくて、それも座標がついて、細かい空間粒度で、全国で作られていたという歴史を持つ。実はそういうデータが眠ってしまっているわけである。企業、国とか自治体にもたくさんそういうデータがあって、そういうのを発掘して精製して、さらにはないものは生産していくということをやらないと、実は研究ができない、または分析ができないわけであるが、そういうところに対し

での評価とか価値が非常に低かった時代がある。

- 1990年初頭から、今でいう機械学習、つまりニューラルネットワークとか決定木について勉強を始めていた。ビッグデータ解析という言葉がなかった時代であるが、そのような研究をしていたら、ある著名な不動産分野の教授から「大きなデータを使っていると、時間ばかり取られているから研究をやった気になるだけで、そんなものは研究ではない」ということを随分言われた。それが、データサイエンスという言葉が登場し、研究としてきちんと評価されるようになってきたというのは、ある意味いい時代になってきた。何よりも、きちんと実社会で実装されるようになってきている。
- 不動産市場のなかでのデータサイエンスを考えると、データから見えることというのは、不動産価格や家賃といった結果だけでなく、その裏側で短期的に起こっていることを実は正しく理解できていない状況であることを知らないといけない。我々は未知な状態というのはリスクと考えるので、それに対して臆病になりがちである。そうすると、ある意味、未知な状態がデータによって既知に変えることができれば、どれぐらい行動してもいいのかとか、不動産市場の足元で今何が起こっていて、どのような収益のインパクトがどういう原因によって変動しているのかということを理解することができるわけであるので、そのようなことをきちんと業界全体に示していくことによって、各不動産業界の経営の意思決定とか、中長期的な視点をお示しできるのではないかと思っている。
- まず我々は、企業の行動、消費者の行動、家計の行動というのが、オフィス市場とか商業施設市場、そして住宅市場それぞれにどういうことが起こるのかを今考えていかなければいけないわけであるが、人の動きがどうなるかということは、例えば生産現場がどうなっているかということを考えると、オフィスに通勤している人が少なくなってますねということは言われているわけだが、どれぐらい減っているのか、オフィスでの滞留時間が、そこに来ている人の数が、どれぐらい変わっているのかというようなことが今見えるようになってきているわけである。
- かつては朝9時から例えば6時、7時まで、8時間、9時間をオフィスビルで過ごして、それを週5回過ごして、家賃が3万5,000円です、4万円ですと言っていた時代が、今こういうデータを見てくると、滞在時間が4時間から6時間ぐらいに減って、週2回ぐらいしか来ていませんというようなことになったときに、そうするとオフィスマーケットにどんなインパクトを与えるのかということを我々は類推できるようになるわけである。そうすると、単純に考えればその分利用率が下がっているの

で、今までの賃貸借契約において、坪3万5,000円を2年というような契約をしていたものが、その稼働状態が変わることによって市場にどういうふうインパクトが出てくるのか。または働く場所がオフィスともう一つ違う場所に、例えばサテライトオフィスみたいなところでの滞在時間が変わってきているとか、そういうことによって働き方が変わって、お金の配分がどう企業を変えてくるのかということ、考えなければいけない話だろうと思われる。

- ・工場、オフィス、店舗と考えると、工場はもうかなりオートメーション化が進んできましたので、実は人がいなくても工場が動くが、無理やり人をそこに張りつけていたことによって工場の生産をしていた。もともと動かすのにオートメーション化が進んでいるところにこのようなことが来ていますから、人の動きもそれほど変わっていません。でもオフィスというところは、今、ものすごく人の動きが変わってきているわけである。さらに店舗も変わってきています。何でそれが起こっているかという、消費の量は実は変わってないかもしれませんがねと考えると、消費する場所と質が変わってきていますねということを考えることになるわけである。例えば、ボーナス日には、ごはんをどうしようかという話をすると、久しぶりに牛肉でも食べようかという話になるとする。前だったら、レストランに行ってすき焼きを食べていた家計が、今はスーパーでお肉を買ってきて家で食べましょうということになる。そうすると、消費の場所が今まで週1回でも外食でしていたものが内食に変わるという変化がある。消費の量は変わらないけれども場所が変わってくるということになる。そうすると店舗での売上げは下がってしまう。一方で牛肉を買う、食べることの量は変わらないが、スーパーマーケットで買うから、そんな高いお肉は買わず、質が変わるか、または質が同じでも価格が変わるかとなって、そこが変化として生まれてくる。そうすると、そのときにスーパーの中での売上げが伸びていて、外食のお金の量が下がってきているというような形になってくる。その中でリテールというようなものの中でも、商業施設でも伸びているところと伸びていないところが生まれてきていて、これがどこまで続くのかというような議論が今後出てくる話になるかと思われる。

- ・住宅の市場のところは生活の場所の変化になるわけだが、今までは、忙しい方々が多かったので、家で過ごす時間というのは、月から金までは夜11時ぐらいに帰って夜寝て、そして7時、8時に出勤するまでの時間を過ごしていたというのが、急激に家で過ごす時間が長くなってきている。コロナの前からそういうことが起こることとは言っていた。高齢化社会が進むことは家の中にいる時間が長くなる人が増えてくるということの意味するので、そうすると住宅の価値はそれほど下がらないかもしれないという話を、高齢化社会の影響という研究会を当機構で開催したと

きに一度紹介したことがあった。コロナショックは、高齢化が本格化する前に、行動制約をされることによって家の中の時間が長くなった。そうすると家に必要とする面積がこう変わりましたねとか、快適に過ごすためには、または家の機能がそこで生産の場が出てくるわけで、家の中での、例えば食事を生産するとか、またはそこで働くとかというようなことは想定されていなかったわけだが、家の機能が少し変わってきた。そうすると、家の本質的な価値は上がっていることになる。

- こういうふうに考えてくると、今、人流とかそのようなデータを眺めていくだけでも、オフィス市場にどういうインパクトが出てきているのか、商業施設市場にどのようなことが起こってきているのか、または住宅市場にどういう影響が出てきているのかということは見えてくるということになる。これが今、短期的に見えていることという話になる。でも見えてないものは何なのかと考える必要がある。最近、鎌倉とか逗子とかいう郊外で家を買う人が増えていますねとか、地方へ動く人たちが増えてきましたねというようなことも言われ始めている。

しかし実際、家を買う人たちの検討者の数というのは、例えば今ストックが100あるとすると、そのうち5%ぐらいの人たちが動くと考えたら、僅か5%の方々の消費行動の変化に対してそういう報道が出てくるわけである。しかし95%以上の人たちは今までの家に住み続けているわけであるから、そういう動かない市場も大切に見ていかないといけない。データとしてはまだ見えてない人たちの動き方がどうなってくるのかということまでは、今のデータから見えていないということになる。

- そうすると、データで見えていることと見えていないことはきちんと分けて考えていかないと、実際データで見えていることだけが真のように語られてしまうわけであるが、そのような目に見えているものだけで物事を眺めてしまうと、間違った判断が出てきてしまうのではないだろうか。いわゆるステルスという、本当は見えてない状態のところ動いていたものが、ビッグデータを使うことによって見え始めているものもあるが、本当は見えてないものがまだ不動産市場の場合には裏側に大きなストックを抱えているので、これから何が起こってくるかというのは、理論とか因果推論みたいなことをしっかりやらないと、データだけで物事を見てしまうと間違った判断になるだろうということを感じている。

4. 質疑応答

[宮川准教授の説明関連]

- ・人流と延滞の関係において、東京が全体的な傾向から外れていることについては、どのような理由が考えられるか。
→ (宮川准教授) はっきりとは確認できていないが、金融機関へのアクセスというのが、もしかしたら東京の企業は円滑に行われていて、手元の流動性が増えて最初の段階からコロナショックが遮断されていたという可能性もあるのかと思われる。

[秋山准教授の説明関連]

- ・メッシュ単位 (500m) 年間推定消費額 (2019年) に関して、駅周辺のメッシュの場合、乗り換え客は考慮されているか。
→ (秋山准教授) 乗換えなどで一時的に滞留しているような人は加味されていない。ある程度の時間滞留している人でないとおそらく消費行動はしていないので、基本的に短時間の滞留時間は入れていない。
- ・メッシュ単位 (500m) 年間推定消費額 (2019年) に関して、メッシュの切り方は、研究者において判断されているのか。
→ (秋山准教授) データ元というか、日本の統計データのメッシュの切り方によってやっている。
- ・データ分析をする際、サンプルデータに特異値がある場合、それを除いたり、何か操作することはあるか。
→ (秋山准教授) 異様な数字が出ているところは、理由を調べてはいるが、特段操作はしない。
- ・統計値が古いので推計値とずれるという点に関し、商業統計の最新のデータは何年ぐらいのものがあるのか。
→ (秋山准教授) 今は商業統計というのはなくなっており、経済センサスに変わっている。たしか 2018 版があるが、メッシュにはなっておらず、市町村とかになってしまっている。

- ・場所による消費額の違いが考慮されていないというのはどういうことか。
- (秋山准教授) 要するにこれはあくまでも各人の滞在時間が重みになって、そこで幾らお金を使っているかが決まっている。各人の年収プラス滞在時間になっているが、もう一つは同じ滞在でもお金の使い方が違うと思われる。観光の行動に伴って例えば銀座に来ている人と、日常的に銀座を使っている人だとお金の使い方がもう全然違う。日常的に使う人はそこでそんなにバリバリお金は使わないですけれども、観光で地方から銀座に来た人は、多分、せっかくなのでいろいろ買物をしようという行動にするとと思われる。そういった行動によるお金の使い方の重みづけとかを入れないと駄目というところで、そういう意味では場所とか、行動とか、目的とかいうことかもしれない。

[清水教授の説明関連]

- ・不動産分野に限らず、ビッグデータを使った分析の過程においては、政府、行政、大学、民間団体等で、連携を取りながら行われていると思われるが、今以上にマッチングされればよいと考えられる。この点に関しては、どのような状況になっているか。
- (清水教授) 不動産業界は、昔からデータをものすごく大事にして、作ってきた業界だと思われる。1990年ぐらいだったか、例えば都庁と国土交通大臣免許の方々とかに対していろいろな苦情が入った。そのような苦情が記載された文字データが大量に残っているので、そのデータを解析できないかという相談があった。いわゆるテキストマイニングである。また移転登記情報に関し、国土庁で、土地情報課ができ、各法務局に入っていたデータを集めて、その中で個別の取引を見ながら、収集された取引価格情報が残っている。自治体からまた国に報告されてくるような着工統計があるとか、そういうデータが国にもあるし、登記という資源もあるし、また各ポータルサイトみたいなところとか、各不動産団体にも、ばらばらと各ビッグデータがあって、それぞれが個別に蓄積されてしまっていて、一旦それを統合しようとした瞬間、ハレーションが出てしまう。そのようなものをさらに民間と突合することになると、きつともっと大きな問題になってきてしまうと思われる。

綿々と蓄積されてきたそういうデータが使えるルール、例えば先ほどの秋山准教授と共同で研究を行った会社は携帯電話会社のデータを直接には使えないので、東大発ベンチャーという形でそこに会社をつくって、そこで集計していただいて、集計データを使わせていただけるというルールを作っている。もしダイレクトに使えないのであるならば、それぞれの様々な事情の中で、シンクタンクのようなものがあって、そこから加工して、使えるような状態で出すとか考えられる。

今、不動産価格が土地情報ライブラリーで開示されるようになったが、もともとは取引価格情報を世の中に開示しようということから始まった。しかし、アンケ

ートで国が集めるということとなり、さらに開示といっても開示項目は少なかった。でも、その後の審議を通じて、項目をちょっとずつ増やして、今の情報開示のスキームができた。またはレイنزのほうも、東西の統合から始まって、今は4レイنزの統合を始めている。レイنزデータも、個別データは開示はできないけれども、駅別ぐらいに集計したものはずいぶんと前から開示している。このような、様々な取組を政府・業界がやってきた。次の段階では、個別の団体が独自の指標を作り、バラバラと行ってきたものを、どこかで集積していくようなことをやっていかないと、もったいないと思っている。

以 上