

令和 4 年度民間空き家対策東京モデル支援事業
 (TOKYO Data Highway 等を活用した先端技術を駆使した空き家対策)
 事業報告書

事業者名	マイクロベース株式会社
事業の名称	将来空き家予測による空き家の発生予防施策支援

事業概要	GIS により住民基本台帳や相続登記台帳などの庁内データを統合し、データ連携を可能とすることで、空き家発生の把握に有用なデータセットを構築しました。過去の転出・死亡等のパターンを AI に学習させることによって、現在から中長期の将来までの空き家になるパターンを定量化し、将来確率を戸単位で予測を行いました。これによって将来において空き家が急増する地域を空き家予測マップとして定量化し、ワークショップ形式にて現場担当者と議論を重ねることで、空き家対策に必要な要件・ニーズを明らかにしました。
------	--

1. 事業の背景と目的

空き家は毎年増加の一途を辿っており、市区町村の空き家対策の現場では人員を増やせないなか、年々負担が高まることが予想されます。空き家措置法や財産管理制度をはじめとした空き家管理に関する法制度が進展していますが、その多くは事後における対策を対象としています。特定空き家の認定後の助言・指導では既に相続放棄されている場合や、所有者・相続者との連絡等が取りづらくなり、その間に老朽化が進行し負動産化するケースが増加してきています。2024 年には相続登記義務化が施行されますが、野村総合研究所の予測では、2013 年の空き家率は 13.5%から 2030 年には 30%台に上回ると予測されており、より急務な対策が不可欠になると考えられます。

そこで、本事業では将来空き家になりやすい住宅を事前に特定・優先順位付けをできるようにすることで、空き家の発生予防に向けた相続支援・空き家流通支援施策等の事前対策を支援する体制を構築・提供します。現場職員と密に連携し、現場で提供可能なデータや実運用に耐えうる形式を確認・精査しながら、AI によって持続的に空き家発生の事前対策を支援する体制を構築・提供します。

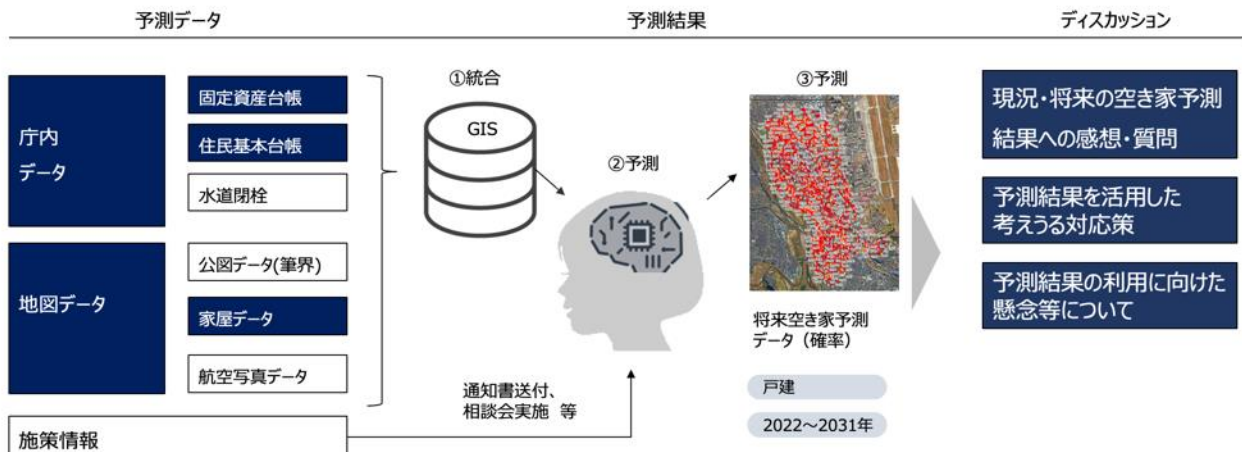


図 1 将来空き家予測の概要

2. 事業の内容

(1) 事業の概要と手順

本取組は、住民基本台帳や相続登記台帳などの庁内において取得可能なデータを活用することで、住戸 1 棟単位において過去から現在までの動向を一貫して追跡し、将来動向まで空き家発生を推察可能にする体制を構築・提供します（図 1）。機械学習によって空き家発生となる転出・死亡等¹のパターンを学習し、現在から中長期の将来までの予測を行います、得られた予測結果から将来の累計空き家発生率や空き家発生数の予測マップを作成し、現場担当者に提供することで、新陳代謝を見越して地域ごとに重点的に対策が行われるべき将来年度や対策内容を検討しやすくなるか試行錯誤を行いました。

(2) 事業の取組詳細

①データ基盤構築

将来空き家予測に必要なデータとして、住民基本台帳、登記台帳（または固定資産台帳）、水道閉栓などの庁内データをご提供いただきました。各市町村によって提供データ項目が異なっていたため、空き家予測を行う内容も異なっています。共通して存在するカラムを同一カラム名としたり、単位や表記の統一等のデータクレンジングを実施しました。空き家予測において考慮できる入力項目が多い方が、予測精度が上がることから、市町村によっては住所情報をご提供いただき、住所をキーに自然言語処理および GIS を用いて住民基本台帳、登記台帳（または固定資産台帳）を統合し、一つのデータベースを構築しました。市の中心部以外では筆界が荒く、住所が重複する場合や一部住所が欠損している場合などが多く確認されたため、重複結合が起きないように結合を実施したところ、郊外部では 40-50%、中心部では 60-70%結合できることを確認しました。空き家予測ではモデル学習時に結合ができたデータはマルチソースで学習し、結合できなかったデータはシングルソースで学習を行いました。

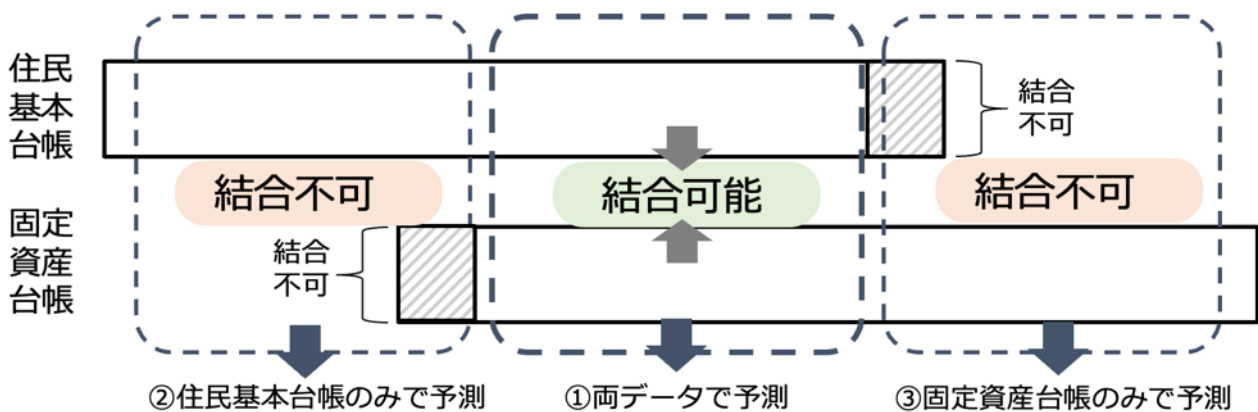


図 2 データ統合の考え方

¹ 試行錯誤の結果、住民基本台帳における異動（転出・死亡等）が最もデータ数が多く、時系列で安定して存在していることから本データを空き家予測の教師データとしました。水道閉栓も候補として上がりましたが、件数が少なく、モデル学習においては不十分と判断しました。

A.複数棟すべて同じ住所の場合

住所だけを用いた結合の限界。住所に対応するGISデータの併用が必要です。



B.住所が不完全なケース

住所が番地までしかないなどの不完全な場合があります。



図3 データ統合に失敗するパターン

②予測・順位づけ

②-1 学習モデルの構築

戸建住宅を対象に空き家になるまでの残存年数・確率を戸単位で予測を行うAIを開発しました。学習・予測の進め方として、過去のデータで学習を行い、過去データから現在を予測し、現在のデータで予測性能を評価しました（図4）。具体的には、2011-2016年までを学習期間とし、2021年までを予測対象とし、2021年までの期間において精度検証を実施しました。予測時には2021年を起点として2022年以降を予測しました。

今回実施した予測では、表1に示すように良好な予測性能を得ることができました。本予測における空き家は死亡と転出等の異動事由を対象としております。実験のなかで死亡事由による空き家のみに限定した方がより精度が高くなることや、東京都以外の市町村における実験において転出においては水道使用量を用いることで高い精度で予測できることがわかりました。しかしながら、今回いただいた4市町村では異動事由項目が提供されていないことや、水道使用量は使用できなかったことから、予測としては異動事由を区別せずに空き家発生予測を行いました。

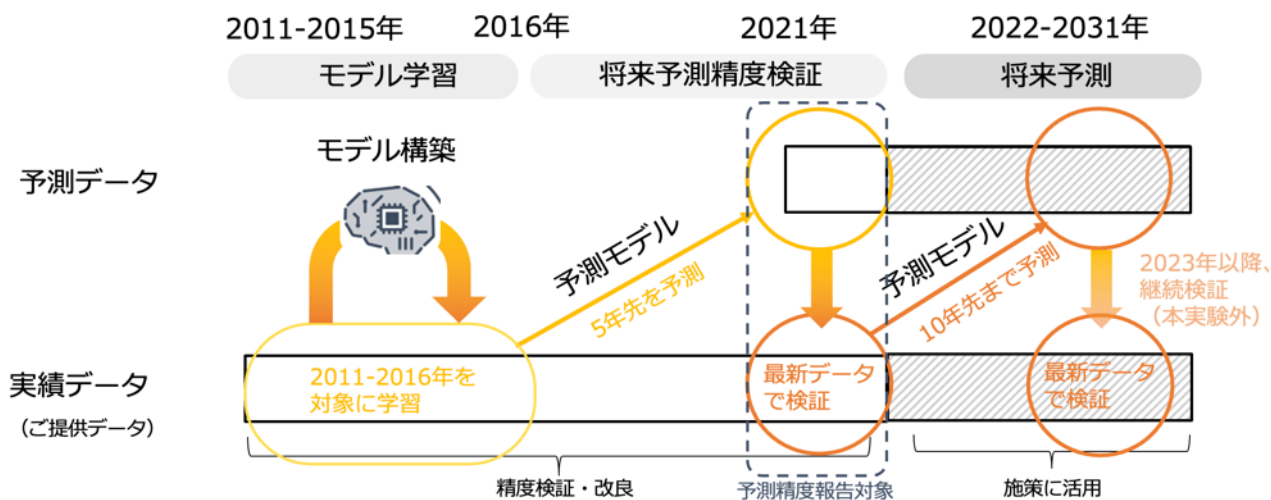


図4 将来予測の全体像

表1 空き家定義別予測精度（5年後空き家になる予測）

項目	精度指標	
	予測精度 Accuracy	適合・再現率 F1 Score
住民基本台帳+固定資産台帳	95.9%	0.66
住民基本台帳のみ	81%	0.65

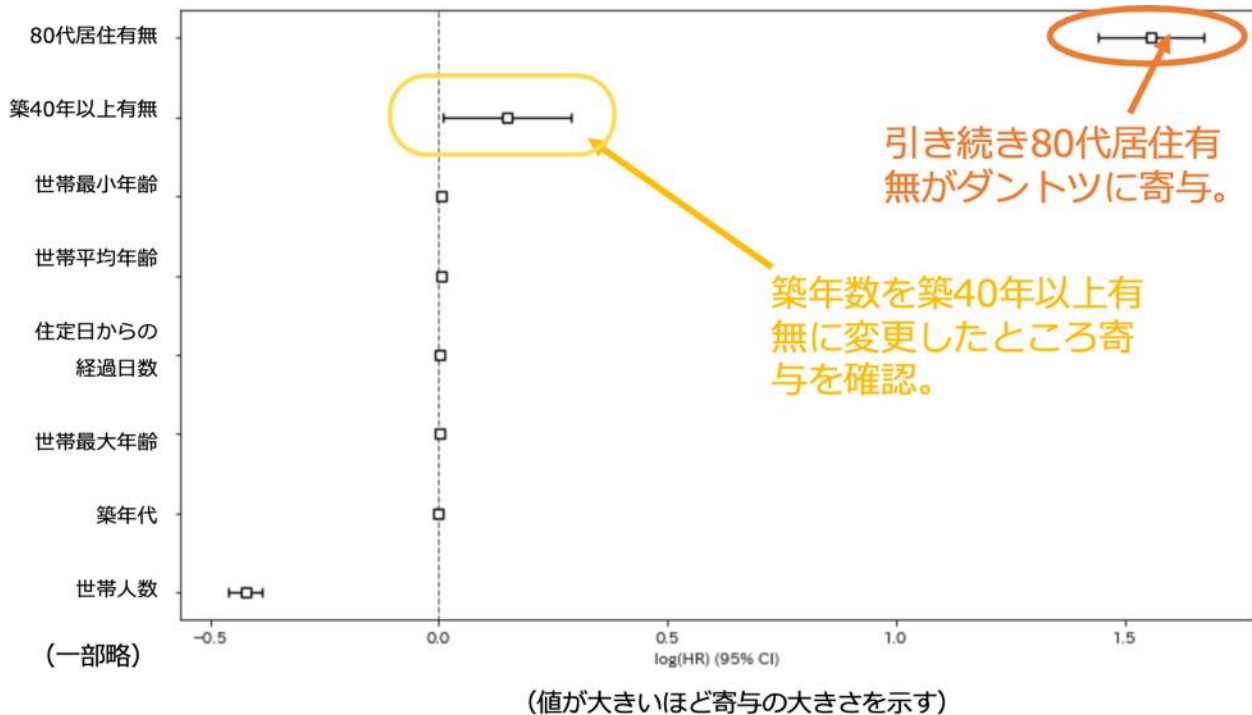


図 5 空き家予測に寄与が確認された項目

なお、本予測では、複数の市町村からご提供いただいたデータを統合してモデル学習・予測を行うことで予測精度の向上を図りました。なかでも町田市からご提供いただいたデータが最も対象年次が長く、項目としても充実していたことから、町田市のデータが学習においてウェイトを占めています。

②-2 空き家予測に寄与した項目および考察

予測への寄与については、住民基本台帳と固定資産台帳（または相続登記）を併用した場合の方が、予測精度が高くなることがわかりました。特に住民基本台帳の方が予測に寄与している結果となりました。これは世帯人員や年齢などの住民基本台帳情報は空き家になる要因と関係が高い居住者の状況がわかるためと考えられます。固定資産台帳（または相続登記）では老朽度の指標となりうる築年数などがあり、空き家との関係が高いことを確認できていましたが（表 2）、学習結果からは大きな寄与は確認されませんでした。この理由としては、対象市区町村では 1990 年代、2000 年代が最も建造が多く、建物全体、および、空き家のボリュームゾーンがこれらの年代に偏ったために、「築年が古いことにより、空き家多くなる」という比例関係が成立しなくなっていました（図 6）。そのため、空き家予測上は築年代別住宅数を均一化してモデルを再学習することで精度向上を確認できました。しかしながら、精度向上は微細なものであり、大きな寄与とはなりません。この理由としては、築年代では 10 年単位にすると築年と空き家率に比例確認が顕著になってきますが、今回は残存データ期間および将来予測精度確認のため、5 年先までしか予測をおこなっていないため、本来築年数が予測するのに適した期間になっていないことが考えられます。20 年以上先の予測をする際に寄与度が高まることが期待され、今回対象としている予測対象期間における機械学習的アプローチには適していないと考えられます。

表 2 築年代別空き家率²

	空き家数	全住宅数	全住宅比率	築年代別 空き家率	備考
1960's	210	914	2%	23%	
1970's	701	5205	13%	13%	
1980's	816	6152	15%	13%	
1990's	1199	9225	22%	13%	
2000's	1056	12627	31%	8%	
2010's	842	6732	16%	13%	
2020's	189	429	1%	44%	竣工後、購入・住定までのギャップがあると考えられます。
合計	5013	41284			

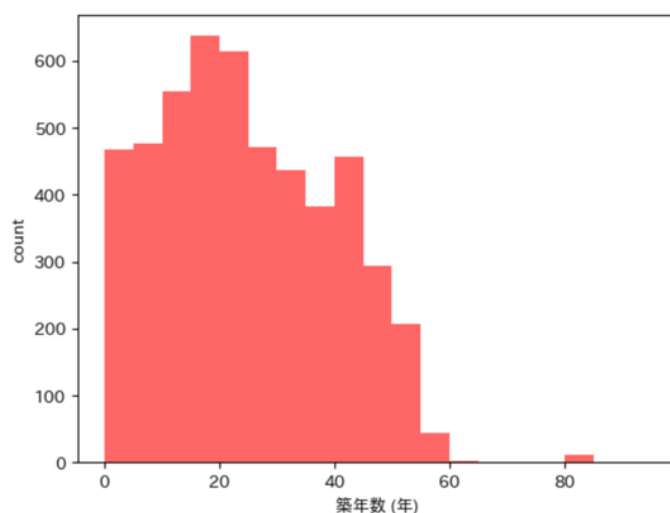


図 6 戸建空き家数と築年代別戸建住宅数の関係（町田市）

また、本学習では水道閉栓日も寄与を確認できましたが、全体としては大きな寄与は確認できませんでした。この理由として、住民基本台帳の異動日と水道閉栓日の関係を確認できたところ、住宅によって住民基本台帳の異動日と水道閉栓日の開きが大きく、はっきりとした傾向が確認されなかったことや、件数自体が多くなかったことが考えられます。後述の順位づけにおいては有用と考えられます。

なお、住民基本台帳の異動日と水道閉栓日の関係としては、以下の3つのパターンを確認できました（図7）。このうち学習ではAとCを考慮しました。

A) 消除と閉栓同じタイミングで発生（赤色）

死亡・転出と同時に退去手続きをとったケース

B) 消除が閉栓より先に発生（黄緑）

死亡・相続などが発生し、現所有者が住宅売却・除却まで水道を利用していたケース

C) 閉栓が消除より先に発生（水色）

高齢者が介護施設・病院へ入居するなど、引越しを行ったものの自宅に居住しなくなったケース

² 町田市を対象に住民基本台帳と固定資産台帳を結合でき、欠損がないデータを対象としております。

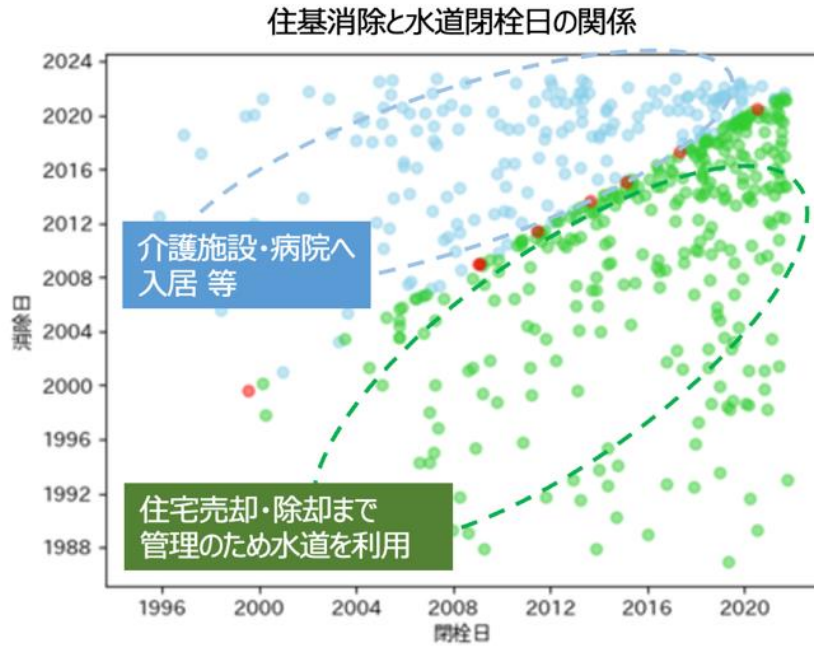


図 7 住民基本台帳の異動日と水道閉栓日の関係

②-3 予測結果および順位付

構築した予測アルゴリズムにより、現在から将来の空き家予測結果データを地図にプロット（GIS ファイル化）しました（図 8）。住宅を将来年度別に空き家有無において示す描画方法と、将来年度別居住確率として示す 2 パターンを示したところ、統計としては前者が、地図描画としては後者の方が解釈しやすいというフィードバックを得られました。



図 8 戸別空き家予測結果（2030 年、福生市）

※個人を特定できないようにイメージを記載

	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031
空家発生件数	31	128	183	71	212	139	185	54	20	35
住宅数 (前年空家除く)	6854	6823	6695	6512	6441	6229	6090	5905	5851	5831
空家発生件率	0%	2%	3%	1%	3%	2%	3%	1%	0%	1%
累積空家率	0%	2%	5%	6%	9%	12%	15%	16%	16%	17%

図 9 優先順位付結果（対象市町村における某町丁字）

さらに、空き家予測結果から町丁字単位における空き家対策の優先順位付を実施しました（図 9）。地域によって後期高齢者のみで構成される世帯など世帯構成が異なり、空き家が発生するピークが異なることから、地域だけでなく、将来年度においても優先順位付を行いました。将来年度別に空き家発生件数を集計することで、空き家が集中的に発生する将来の期間や地域が明らかとなりました。また、累計空き家件数が 10%に到達する年度においても注目していくことで、地域ごとの傾向が可視化され、重点的に対策をとっていくべき地域を明らかにしました。

③ワークショップ

③-1 予測結果および順位付

②の実施結果を報告・議論を行うワークショップを各市町村において開催しました。前半ではマイクロバスより空き家予測精度を報告し、後半からは現場担当者ご関心を持っていただいた各部署のご担当者とのディスカッションというかたちで進めました。

予測結果に対する感想や解釈、課題などをフィードバックいただくとともに、施策への活用案をディスカッションしました。施策に望ましい空き家定義や、予測の解釈をどうすると良いか、予測精度をあげるにはどういう改善をしていくかよいかなど今後につながる複数のフィードバックをいただきました。特に、福生市では都内ではじめて人口減少が起きたというお話や、その原因の一つとなる開発する土地がないという課題感をご共有いただき、町田市や八王子市、あきる野市にはない特有課題を詳細にご共有いただきました。

いただいたフィードバックのなかで、空き家発生だけでなく、期間についても確認できないかという質問をいただきました。空き家になってからの期間が把握できることで、実態調査における空き家定義と近い定義となり、取り扱いやすくなるという声をいただきました。データ上で表現を行おうとすると、空き家発生とは別に、空き家への再入居をモデリング可能となることで、表現することができるが分かりました。一度空き家になってから、同住所に対して住定があった場合に「再入居」と定義し、過去の累計を計算しました。その結果、図 10 のように鉄道沿線沿いで駅近ほど再入居率が高い傾向を確認できました。空き家発生と再入居を地域別に重ね合わせることで、再入居も考慮した新陳代謝の良い地域と悪い地域を区別することができました（図 11）。現場担当者に本結果をご確認いただいたところ、行政として介入すべきかを判断しやすくなったという声をいただきました。空き家発生自体は寿命や就業状況等により防げるものではなく、むしろ、空き家期間を短くし、新陳代謝を悪くしないことが、空き家対策において重要であり、みていくべき指標の一つということを確認しあうことができました。

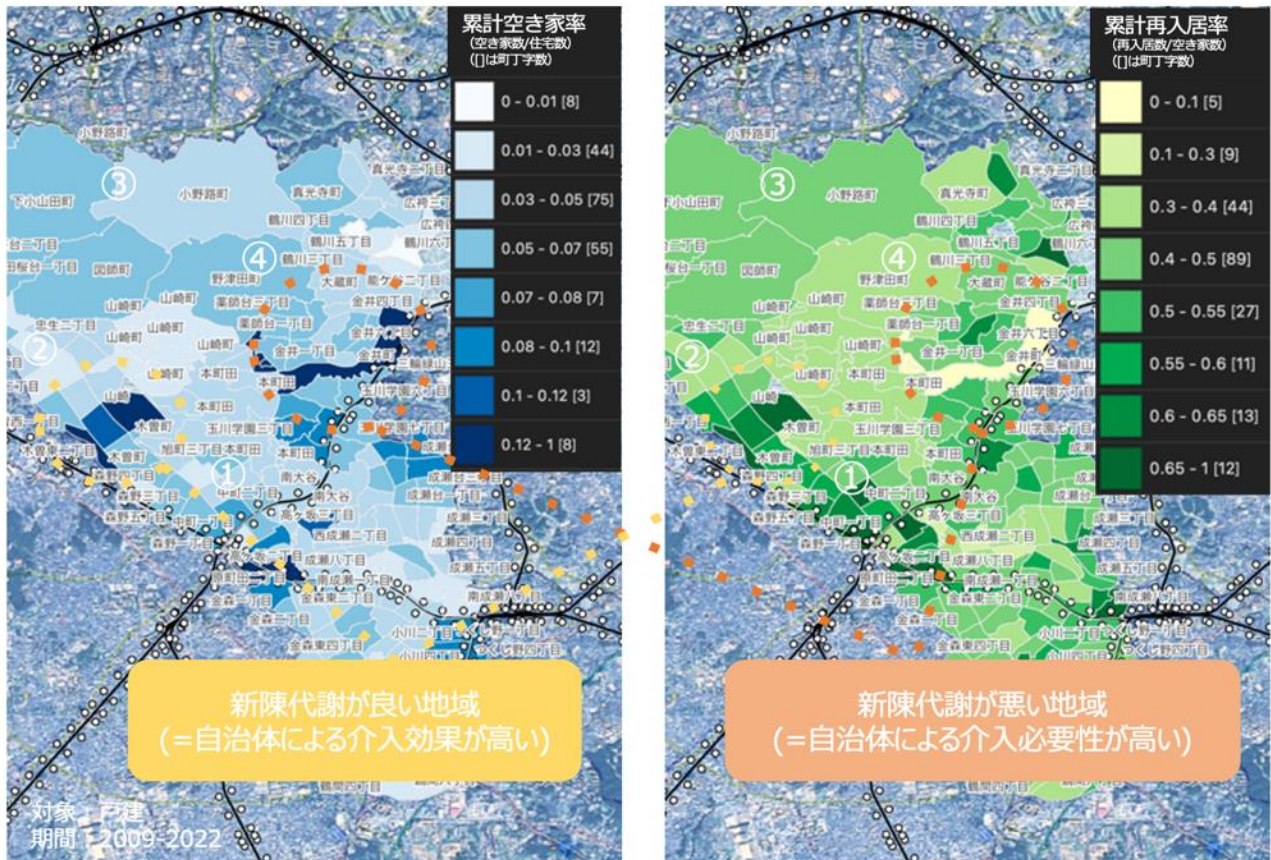


図 10 空き家発生と再入居から見る地域の代謝

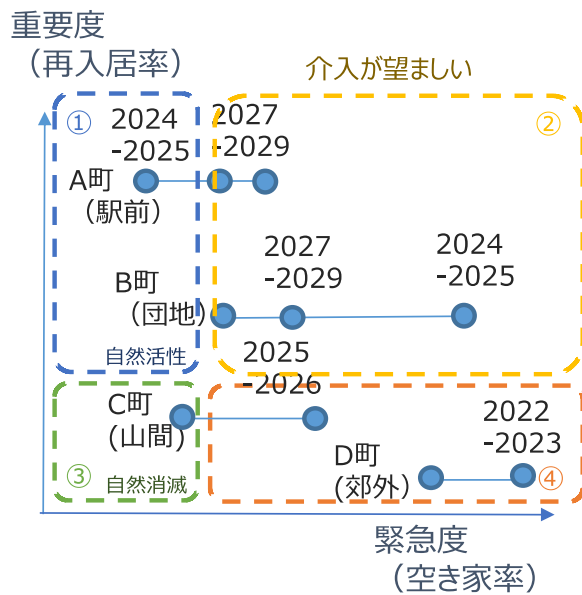


図 11 空き家発生と再入居から見る重点エリアの考え方

再入居は、考慮すべき項目は異なりますが、空き家発生と同様のアルゴリズムにおいて将来予測可能です。将来の新陈代谢を予測・可視化できるようになることが現場において有用であり、来年度以降の課題として認識することができました。

③-2 合同空き家報告会の開催

本取組の成果をご協力いただいた市町村を交えてご報告、議論する場として、空き家合同報告会を開催しました。学術有識者として東京都立大学の饗庭教授、東京都市大学の秋山准教授にもご参加いただき、最先端の学術的な知見・事例をご紹介いただき、議論を深めました。

背景として、今年度ご協力いただいた市町村と議論していくなかで、各市で空き家対策の考え方や課題感から、データ提供や予測データ活用などの考え方まで広く違いがあることがわかりました。また、現場の担当者にとってみると庁外の取り組みはあまり知られておらず、先例がないなか対策を悩んでいらっしゃるお話を伺うことが少なくありませんでした。そこで、合同報告会という場を設けることで、相互の知見を共有しあうとともに、空き家予測結果の活用などにおいても見解を共有しあえることで、今後の判断に活かすことができると考えました。

東京都立大学の饗庭教授からはよる都市計画的観点からの空き家問題の捉え方や空き家の現状解釈、東京都市大学の秋山准教授からはデータを活用した最先端の空き家判定手法についてご講演いただきました。弊社だけでは提供できない知見については専門家のちからをお借りすることで、相互に視野や判断要素を広め合うことができました。



図 12 合同空き家報告会の様子

③-3 システム開発

本事業で予測した結果を提供するにあたり、庁内ではデータを可視化する体制がなく、GISで読み込みを行うにも負担が伴うという声をいただき、ウェブアプリケーション上で予測結果を確認ができるシステムを構築しました。本システムでは該当住所を入力することで、空き家予測結果を表示したり、町丁字単位の予測結果マップを確認ができるものとなっています。来年度以降、本事業を展開していくにあたり、予測アルゴリズムと同時に本システムを提供して参りたいと考えています。



図 13 自動予測システム

(3) 成果

本事業の成果として以下の成果を得ることができました。

- 空き家把握に有用な庁内データを明らかにするとともに、住戸単位において高精度の予測を実現することができました。
- 現場担当者とのワークショップでは、庁内データによって現地確認を行わなくとも居住状況を推察することが可能となり、市民からの通報などが入る前に事前に空き家となりえる住宅を把握することや、相談会などの実施場所の優先順位付に有用であるというフィードバックを得られました。また、新陳代謝まで確認できるようになることで、より実態調査から得られる結果とも整合がとれ、今後の対策検討を行いやすくなることがわかりました。
- ご協力いただいた市町村にて合同報告会を実施し、各市町村の持つ課題や知見、データ提供や予測データ活用への考えなどを共有し合うことができました。

3. 評価と課題

本事業への評価と今後の課題をとして、以下のように考察しております。

- 庁内データの提供は、非常にハードルが高く、本事業が開始可能となるまで多大な日数を要しましたが、今年度協力を依頼した4市町村すべてからご協力をいただけることができました。本事業の意義をご理解いただき、データ提供だけでなく現場担当者との複数回にわたる議論や予測結果へのフィードバックを賜ることができたのは大きな成果であったと考えております。
- 計画当初の想定以上に、提供いただいた庁内データから高精度な予測結果を得られ、方法論を確立することができました。実際にある市町村では来年度に本格的な事業として展開をご検討いただいております。
- 計画当初では空き家予測アルゴリズムを活用した空き家施策効果測定を計画していましたが、ご協力いただいた市町村からは施策実績がないケースや、提供いただくことができず、実現には至りませんでした。
- 東京都立大学の饗庭教授からのアドバイスや現場担当者との議論において、住宅の新陳代謝が健全に行われているかを判断したいというフィードバックをいただきました。本事業では空き家発生のみを対象としており、新陳代謝の把握には、空き家への再入居（建替、中古居住）も対象とする必要があるという気づきを得られました。本事業後半より再入居を予測するため、市町村より追加データを提供いただき、一定の精度を得られましたが、実用までには引き続き試行錯誤が必要で課題が残っています。

4. 今後の展開

以上を振り返り、来年度以降以下について展開して参りたいと考えています。

① 施策効果測定に向けた継続的な開発

2月15日に開催した合同報告会においても、参加市町村から施策による効果をどう測定し、改善していけば良いかという問いをいただきました。限られたリソースのなかで効果的な施策を実施していく必要があり、その大前提として、現状の施策をどう評価することができるかという問題意識に基づくものでした。いわば、住宅政策をEBPM（データに基づく政策立案）によって推進していくことそのものであると言えます。

ウェブ広告などの業界ではA/Bテストに代表される因果推論が当たり前のように行われており、最適化が絶え間なく繰り返されながらUIアップデートや機能追加などの各ウェブ施策が展開されています。これにはGoogle Analyticsのようにアクセス数やPVなどのトラフィックデータを収集・蓄積し、流入からコンバージョンに至るまでの過程を分析・最適化・予測ができる機能があることによって実現可能となっています。住宅政策においても施策効果測定を行なっていくには、同様の基盤が必要であり、その基盤として空き家予測が基礎になると考えています。たとえば、相談会・DM送付や税務に関する緩和施策、補助施策などの各種施策を変数として空き家予測モデル上で考慮し、シミュレーションできるようになることです。また、トラフィックデータに相当するデータとしては、網羅度が高く、かつ更新頻度の高い住民基本

台帳や水道使用量といった日々更新されるデータが基盤となりえます。施策が与えた効果を住民基本台帳や水道使用量などのデータとして観測可能とし、関係性を定量化する体制が、弊社が考える冒頭の問いへの回答です。

本事業ではこうした体制構築に向けて、今年度の成果に加え、引き続き再入居の分析・予測の完成度を高めることで、住宅の新陳代謝をモデリング可能にすることを目指して参ります。そのうえで、施策があった場合とない場合とで新陳代謝に与える効果を定量化できるようになる基盤へと発展させて参ります。それには、観測データとなる住民基本台帳などと同様に、大量の施策データが必要であり、さらに、住宅そのもののデータも必要になってくることが想像されます。継続的な開発を進めていくべく、今年度ご協力いただいた市町村に引き続き協議し、まちづくりコンサルタント会社や不動産事業者などと連携しながら進めて参りたいと考えています。

②事業拡大に向けた体制構築

今年度の取り組みによって、将来空き家予測手法が確立し、良好な予測精度を得ることができました。空き家把握に課題を持っている市町村や重点エリアを定めて展開されようとする市町村に対して、積極的に展開して参りたいと考えています。展開にあたっては、市町村によって提供可能なデータが変わってくるのが考えられ、提供データに応じて実施可能な予測を整理し、提供メニューを構築します。

また、マイナンバーや不動産 ID の進展などにより、より一層のデータ連携が加速するとともに、スマートメーターの普及などさらなるデータ資源の活用可能性が高まっています。一方、漏洩した場合のリスクが大きくなることから、セキュリティ対策の重要度が非常に高まっています。今年度、弊社では ISMS の認証を取得したばかりで体制が発展途上であるため、こうした社会的な動向に対応していけるように早急に対応して参りたいと考えています。

③市町村間のナレッジ共有の仕組み化

上述の 2 月 15 日の合同報告会では市町村間によってお互いの取り組みや知見が共有され、議論が行えることに大いに意味があるという声をいただくことができました。こうした場は先例の取り組みを行う際には非常に有効であり、今年度ご協力いただいた以外の市町村にも声がけをしてナレッジ共有を仕組み化して参りたいと考えております。特に市町村単体ではモデル学習にデータ数が少ない場合にも、他市町村データを併用することで、予測精度を高められることを今年度において確かめることができました。知見に加え、データなども含めた集合知として市町村間で連携する体制を整えることで、より効果的な空き家発生予防が期待されます。