

研究ノート

不動産業況はどのように語られてきたか

—記述式コメントにみる語彙構造と話題傾向—

一般財団法人土地総合研究所 客員主任研究員
麗澤大学 経済学部 教授 大越 利之
おおこし としゆき

1. はじめに

近年、テキストデータ分析の分野では、自然言語処理 (Natural Language Processing, NLP) や生成 AI の進展により、実務・研究の両面で高度な分析が身近なものとなっている。特に自然言語処理は、経済や社会の動向を把握する手法として注目を集めており、たとえばフランス銀行と OECD が共催した 2024 年の国際会議では、「ビッグデータと大規模言語モデルの登場によって、テキスト情報の表現力と分析力が飛躍的に向上し、そこから多様なシグナルや意味を抽出できるようになるという、大きな変革期を迎えている」と報告されている¹。また、金融分野においては、企業報告書や政策文書、ニュース記事などの非構造的データに対し、自然言語処理技術を適用する動きが急速に広がっており、定量的な経済分析を補完する実践的な手法として注目されている。たとえば、Du et al. (2025) は、自然言語処理が金融セクターにおける意思決定や革新の在り方を変えつつあることを、10 の応用領域を通じて包括的にレビューしている。

本研究では、一般財団法人土地総合研究所が四半期ごとに公表する「不動産業況等調査」において収集された記述式設問への回答 (以下、コメント) のうち、2016 年から 2025 年までの直近 10 年間に記述された 1,942 件を対象として、自然言

語処理の手法を用いた語彙構造および話題構造の分析を行う。本調査では、経営状況に関する選択式の設問に加え、各期の特徴的な事項に関してコメントを収集しており、これらは記述形式で提供されている。選択式回答は「不動産業況指数」として指数化され、長期的な推移の分析に用いられてきたが、コメントについては、これまで系統的な分析が行われてこなかった (たとえば、伊藤 2023 ; 大越 2016, 2013)。

この 10 年間には、日本銀行によるマイナス金利政策の導入 (2016 年)、新型コロナウイルス感染症による経済社会の混乱 (2020 年以降)、近年の物価上昇や急速な円安など、不動産市場に大きな影響を与える外的な変化が相次いだ。これらの環境変化に対して、業界関係者がどのような言葉を用いて経営状況を記述していたのかを構造的に明らかにすることで、不動産における認識の変遷や特徴的な言説の出現・変化を把握することを試みる。

次節以降の構成は以下のとおりである。第 2 節では、分析に用いたデータの前処理方法および自然言語処理の手法 (共起ネットワーク、LDA トピックモデル) について概説する。第 3 節では、全期間・全業種のコメントを対象にした語彙と話題の全体傾向を示す。第 4 節では期間を 3 つに区分して時系列的な変化を考察し、第 5 節では 3 業種 (住宅・宅地分譲業、不動産流通業、ビル賃貸業) に分けて業態別の特徴を分析する。最後に、第 6

¹ [Banque de France & OECD \(2024\) Leveraging Natural Language Processing \(NLP\) to Answer Economic Questions](#)

節では、時間軸と業態軸を交差させる視点から、自然言語処理によって得られた知見の総括と考察を行う。

2. 分析方法

本研究は、2016年から2025年までの不動産業者による記述式コメント(1,942件)を対象とし、自然言語処理技術を用いてその語彙的構造および話題傾向を分析するものである。本節では、その分析手順と技術的手法について説明する。

まず、一次データとして用いたのは、「不動産業業況等調査」において、不動産業者が記述形式で提出したコメントである。これらのコメントは文体が不統一であり、話し言葉や箇条書き、省略表現、記号、業界内の略語や語法も含まれていた。このようなデータをそのまま自然言語処理にかけることは困難であるため、事前処理として「正規化(normalization)」を行った。

コメントの正規化には、OpenAIが提供する生成AI(GPT-4o)を活用した。これは、各コメントに含まれる曖昧な語句や口語的表現を意味の通る文として再構成するものであり、従来は人手によって行われていた前処理作業を自動化するものである。コメントはCSV形式でまとめた上で、生成AIに逐次的に処理させ、文意を保ったまま平易な文章に整形した。具体的には、①箇条書きや口語表現を意味の通る文として再構成し、②無意味な接続詞や接頭語(例:「なお」「また」など)を削除し、③文法的に不自然な箇所や句読点の欠落を補正し、④原文の意味を損なわない範囲で、冗長な部分を簡潔に要約した。すなわち、人間が1コメントずつ内容を読み取りながら修正するような作業を、生成AIが自動で代替したことになる。

この処理は、従来であれば研究者自身、または研究補助者などが手作業で行う必要があった作業であり、時間・コスト・再現性の面で大きな制約があった。生成AIによる文脈理解・編集能力を活用することで、数千件規模の記述データを高精度かつ短時間で整形できることが確認された。これにより、自然言語処理の前提となる高品質なテキ

ストデータの確保が大幅に効率化され、生成AI時代における実証研究において、データ整形作業の効率化・高度化を可能にする技術的手段の一例を示したものと見える。

正規化後のコメントに対しては、形態素解析器(MeCab+mecab-ipadic-NEologd)を用いて名詞を中心に語彙を抽出し、語彙の頻出傾向、共起関係、トピック構造を分析した。分析手法の概要は以下のとおりである。

- ① 頻出語分析: 語の出現頻度を集計し、主要な関心語彙を抽出。
- ② 共起ネットワーク分析: 語と語の共出現関係をネットワーク構造として可視化。
- ③ LDA トピックモデル分析: 文書(コメント)に含まれる語彙の出現パターンに基づき、潜在的な話題(トピック)を確率的に抽出。

なお、LDAの可視化にはpyLDAvisを用い、Web上でも動的に閲覧可能なHTMLファイルとして提供している²。視覚化画面では、トピック間の関係性、各語の代表性・特異性をインタラクティブに確認できる。

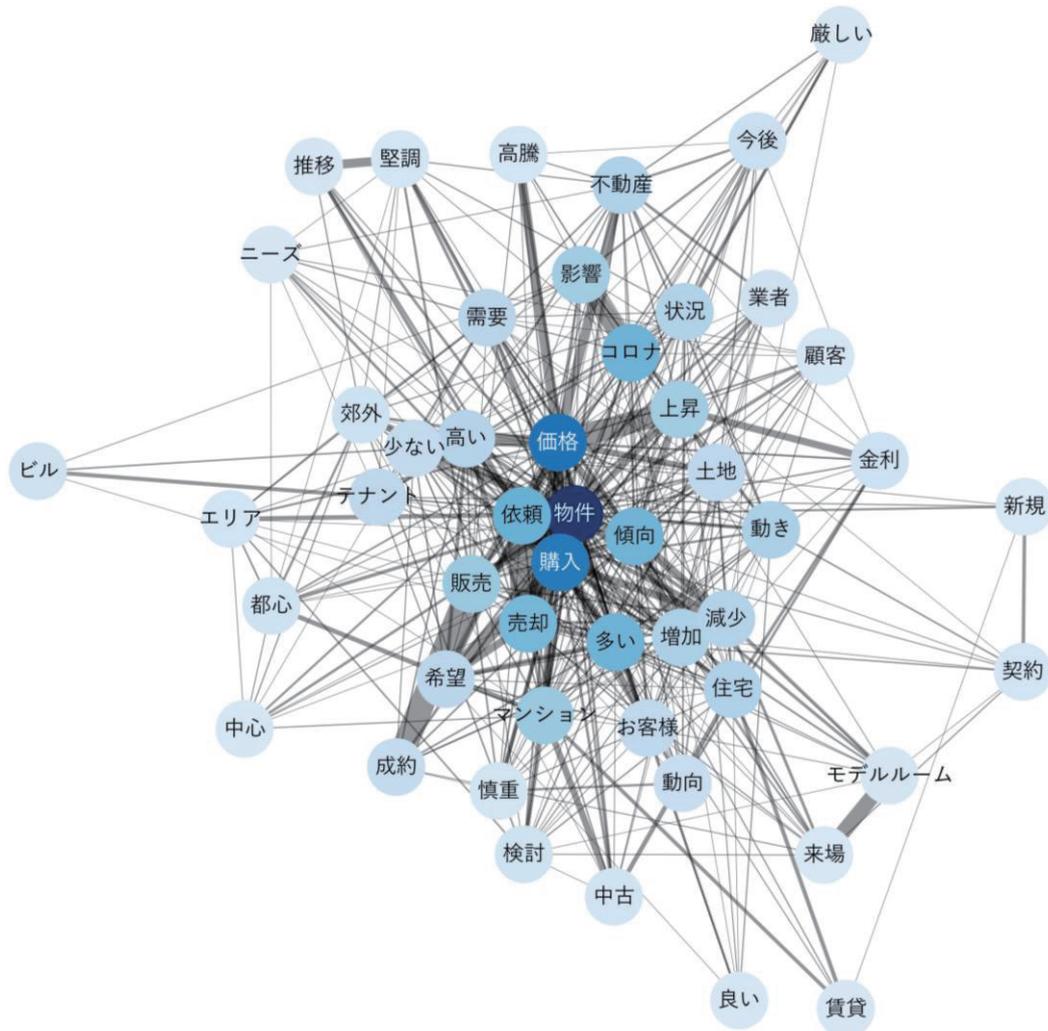
3. 全業種・全期間における記述コメントの語彙構造と話題傾向

本節では、全期間(2016年4月~2025年4月)の不動産業者のコメント全体を対象とし、記述テキストに含まれる語彙構造および話題傾向の全体像を可視化することを目的とする。分析には自然言語処理の技法を用い、語と語の共起関係に注目する共起ネットワーク分析、ならびに語彙の出現分布に基づく話題抽出を行うLDAトピックモデル分析を併用した。

分析の対象は、1,942件のコメントである。事前処理として、全データに対し表記の揺れや助詞などのノイズを除去する正規化を行った上で、形態素解析(Janome)により名詞と形容詞のみを抽出した。さらに、出現頻度の高い汎用語(例:「よ

² [pyLDAvis](#): LDA トピックモデルの結果を可視化するPython用ツール。

図1. 不動産業者の記述式コメントにもとづく共起関係（全期間・全業種）



う」「こと」「ため」「なお」など)に加え、本分析の意義を損なわないと判断された語(例:「当社」「お客様」「新型」など)をストップワードとして除去した。これらの前処理を経て、構造的な語彙関係と話題構造を分析した。

(1) 共起ネットワークによる語彙の関係性

共起ネットワーク分析では、コメント内で頻繁に同時出現する語の組み合わせを抽出し、ノード(語)とエッジ(共起)からなるネットワーク構造として可視化した(図1)。この分析からは、語同士の関係性や意味的なクラスターの存在が明らかとなる。たとえば、「物件」「価格」「購入」などは、ネットワークの中心に位置づけられ、これらの語

が業者コメント全体において、いかなる話題であれ頻繁に参照される共通語彙であることが示唆された。一方で、「テナント」「ビル」「厳しい」などは、ネットワーク上ではやや周縁に位置づけられており、特定の業態(たとえばビル賃貸業)や時期(たとえばコロナ禍)において集中的に使用されている語彙であると考えられる。また、「価格」や「金利」のような語は、複数のクラスターを横断して接続しており、市場全体の状況や判断にまたがる共通の説明変数である可能性を示している。

(2) LDA トピックモデルによる話題抽出

LDA (Latent Dirichlet Allocation) トピックモデル分析では、個別のコメント群が複数の話題

表1. 不動産業者の記述式コメントに関する LDA 分析結果 (全期間・全業種)

	トピック	トピックごとの主要語彙 (出現確率上位)
全体 (ALL)	トピック1	0.022*住宅 + 0.019*コロナ + 0.016*開発 + 0.015*マンション + 0.015*賃貸 + 0.010*金利 + 0.009*影響 + 0.009*増加 + 0.008*事業 + 0.008*経済
	トピック2	0.044*価格 + 0.040*購入 + 0.037*物件 + 0.033*依頼 + 0.031*売却 + 0.024*多い + 0.019*傾向 + 0.018*高い + 0.016*販売 + 0.015*希望
	トピック3	0.027*テナント + 0.022*物件 + 0.017*コロナ + 0.015*新規 + 0.013*動き + 0.012*需要 + 0.011*移転 + 0.011*状況 + 0.010*契約 + 0.010*価格
	トピック4	0.048*物件 + 0.022*傾向 + 0.021*需要 + 0.017*購入 + 0.014*増加 + 0.011*消費 + 0.011*影響 + 0.011*顧客 + 0.010*都心 + 0.010*価格
	トピック5	0.035*ビル + 0.024*影響 + 0.023*コロナ + 0.017*賃料 + 0.016*状況 + 0.012*テナント + 0.009*予定 + 0.009*推移 + 0.008*厳しい + 0.007*動き

※ 各トピックは、LDA モデルにより学習された語彙分布を表しており、「0.035*XX」のような表記は、当該語 (XX) がそのトピック内で占める出現割合 $P(w|topic)$ を示す (表2、3も同様)。

(トピック) の混合から成ると仮定し、それぞれのトピックがどのような語彙で構成されているかを確率的に推定した。全体コメントを5つの代表的な話題 (トピック) に分類した結果、以下のような構造が明らかになった (表1)。各トピックは、語彙の組み合わせによって意味的に異なる特徴を示している。

●トピック1

「住宅」「コロナ」「金利」「事業」など、住宅市場やマクロ政策 (たとえば金利変動) に関する認識が中心である。

●トピック2

「価格」「購入」「売却」「依頼」「販売」など、売買に関する直接的な行動が反映された話題となっている。

●トピック3

「テナント」「移転」「契約」などが代表語となっており、企業の移転や更新を巡る動向に注目したトピックである。

●トピック4

「物件」「傾向」「需要」「消費」「顧客」などが上位にあり、需要の変化や消費者の志向性に関連する話題と読み取れる。

●トピック5

「ビル」「影響」「コロナ」「賃料」「テナント」などの語が多く、ビル賃貸市場における収益構造

やコロナ禍での稼働状況に関する話題で構成されている。

LDA の可視化ツール (pyLDAvis) では、各トピックのコメント集合に対する寄与度が円の大きさを示されている³。図中の青いバーは「全体コーパスにおける語の出現頻度」を、赤いバーは「選択されたトピック内の語の出現頻度」を示す。赤いバーが青いバーの高さに近い語は、そのトピックで頻繁に使われている語といえる。すなわち、赤いバーが相対的に高い語は、そのトピックとの関連性が強い語として解釈される。なお、表1の結果は、pyLDAvis におけるパラメータ λ を1.0に設定し、「各語のトピック内出現割合 $P(w|topic)$ 」に基づいて代表語を抽出したものである。これは、そのトピックにおいて頻繁に用いられる語を重視して選定したものであり、語の「特徴的な重要度」よりも「頻出度」に重きを置いた視点である⁴。

³ 各トピック間の位置関係や、トピックごとに特徴的な語の寄与度は、pyLDAvis により作成したインタラクティブな可視化ページで確認できる。[土地総合研究所ホームページ](#)からファイルをダウンロードし、ローカルに保存したうえでブラウザで開くことで閲覧可能である。

⁴ λ は、語の代表性の定義を調整するパラメータであり、0~1の範囲で設定される。 $\lambda = 1$ では「トピック内の出現頻度」が重視され、頻繁に出現する語が代表語として抽出される。一方、 λ が0に近づくほど、「他のトピックにはあまり出現しない語 (特徴語)」が優先される。

表 2. 不動産業者の記述式コメントに関する LDA 分析結果（期間別・全業種）

期間	トピック	トピックごとの主要語彙（出現確率上位）
P1 (2016-2019)	トピック1	0.037**"物件" + 0.032**"価格" + 0.022**"購入" + 0.015**"傾向" + 0.014**"不動産" + 0.013**"マンション" + 0.012**"上昇" + 0.012**"状況" + 0.012**"動き" + 0.012**"契約"
	トピック2	0.062**"依頼" + 0.054**"売却" + 0.047**"購入" + 0.037**"物件" + 0.025**"増加" + 0.024**"希望" + 0.021**"傾向" + 0.019**"マンション" + 0.017**"中古" + 0.013**"減少"
	トピック3	0.030**"価格" + 0.026**"販売" + 0.022**"物件" + 0.020**"成約" + 0.019**"動向" + 0.018**"傾向" + 0.016**"消費" + 0.016**"購入" + 0.013**"増税" + 0.013**"状況"
	トピック4	0.055**"物件" + 0.018**"ビル" + 0.016**"テナント" + 0.014**"傾向" + 0.014**"価格" + 0.011**"動き" + 0.010**"業者" + 0.009**"推移" + 0.008**"マンション" + 0.008**"状況"
	トピック5	0.020**"業者" + 0.016**"価格" + 0.016**"住宅" + 0.013**"影響" + 0.012**"購入" + 0.011**"ビル" + 0.010**"動き" + 0.010**"消費" + 0.009**"物件" + 0.009**"テナント"
P2 (2020-2022)	トピック1	0.041**"コロナ" + 0.024**"物件" + 0.022**"動き" + 0.019**"影響" + 0.016**"需要" + 0.014**"テナント" + 0.011**"宣言" + 0.010**"緊急" + 0.010**"価格" + 0.009**"事態"
	トピック2	0.046**"コロナ" + 0.027**"影響" + 0.024**"販売" + 0.020**"物件" + 0.013**"状況" + 0.013**"成約" + 0.010**"価格" + 0.010**"ニーズ" + 0.008**"増加" + 0.008**"減少"
	トピック3	0.058**"購入" + 0.055**"依頼" + 0.045**"売却" + 0.032**"価格" + 0.025**"物件" + 0.020**"コロナ" + 0.018**"希望" + 0.017**"不動産" + 0.016**"傾向" + 0.014**"土地"
	トピック4	0.031**"コロナ" + 0.019**"影響" + 0.017**"モデルルーム" + 0.017**"賃料" + 0.016**"来場" + 0.016**"物件" + 0.014**"状況" + 0.013**"増加" + 0.013**"テナント" + 0.012**"検討"
	トピック5	0.031**"住宅" + 0.026**"賃貸" + 0.023**"マンション" + 0.021**"事業" + 0.017**"土地" + 0.015**"物件" + 0.011**"郊外" + 0.011**"開発" + 0.010**"中古" + 0.009**"取扱い"
P3 (2023-2025)	トピック1	0.052**"物件" + 0.037**"価格" + 0.025**"上昇" + 0.023**"購入" + 0.022**"売却" + 0.020**"依頼" + 0.015**"郊外" + 0.014**"傾向" + 0.013**"推移" + 0.013**"都心"
	トピック2	0.030**"価格" + 0.028**"上昇" + 0.028**"購入" + 0.021**"不動産" + 0.017**"金利" + 0.016**"建築" + 0.016**"影響" + 0.013**"物件" + 0.012**"高騰" + 0.012**"動向"
	トピック3	0.023**"傾向" + 0.020**"販売" + 0.017**"購入" + 0.016**"金利" + 0.014**"成約" + 0.013**"増加" + 0.013**"物件" + 0.012**"ローン" + 0.011**"高騰" + 0.011**"価格"
	トピック4	0.024**"価格" + 0.018**"マンション" + 0.016**"傾向" + 0.011**"不動産" + 0.011**"増加" + 0.010**"金利" + 0.008**"物件" + 0.008**"契約" + 0.008**"コロナ" + 0.008**"住宅"
	トピック5	0.025**"価格" + 0.019**"購入" + 0.017**"売却" + 0.015**"不動産" + 0.013**"案件" + 0.012**"希望" + 0.012**"会社" + 0.011**"オフィス" + 0.011**"店舗" + 0.010**"動き"

4. 期間別にみた話題傾向の変化

本節では、不動産業者による記述コメントの話題傾向について、マクロ環境の転換点を踏まえ、以下の3期間に区分して比較分析を行う。

- P1：マイナス金利導入以後
(2016年4月～2019年10月)
- P2：新型コロナウイルスの影響期
(2020年1月～2022年10月)
- P3：コロナ収束・マクロ経済ショック期
(2023年1月～2025年4月)

この区分は、各年の1月・4月・7月・10月時点の景況感について実施された調査結果に基づくものであり、制度的・景気的な外的変化が不動産市場における主観的評価へ与える影響を分析することを目的として設定された。

各期間の記述コメントに対して、LDA トピックモデルを適用し、5つの話題（トピック）を抽出した（表2）。ここでの代表語は、pyLDAvisにおいてパラメータ λ を1.0に設定し、「各語のトピック内出現割合 $P(w|topic)$ 」に基づく頻出語として抽出されたものである。すなわち、「そのトピック

内で頻繁に用いられた語」に焦点を当て、話題の輪郭を浮かび上がらせている。

(1) マイナス金利下での活況と価格への関心 : P1

P1 では、複数のトピックにおいて「物件」「価格」「購入」「依頼」などが高頻度で出現しており、総じて活発な売買市場を反映する話題構造となっていた。トピック1では「価格」「物件」「不動産」「上昇」「マンション」などが頻出しており、住宅価格の上昇とそれに対する反応が中心的話題であったと考えられる。トピック2では「依頼」「売却」「購入」が上位を占めており、住宅流通市場における売り手と買い手の活発な動きが読み取れる。トピック3では、「価格」「販売」「物件」「成約」などの語が上位に並んだほか、「消費」「増税」などの語も含まれており、2019年10月の消費税率引き上げ(8%→10%)を控えた駆け込み需要や消費者心理の変化が、市場の話題として反映されていた可能性がある。

(2) コロナ禍による制約と回復の兆し : P2

P2 では、「コロナ」「影響」など、感染症に起因する制約的要素が多くの特ピックに現れた。特にトピック1では「宣言」「緊急」「事態」などが頻出し、社会全体の行動制限が市場の話題を一色に染めていたことがうかがえる。トピック4では「モデルルーム」「来場」「検討」が並び、分譲販売における対面営業の困難さと顧客の様子見姿勢が示唆される。一方、トピック3には「購入」「依頼」「売却」など、トピック2には「販売」「成約」「ニーズ」などが表れており、中期的な回復の兆しが浮かび上がってくる。このように、P2期はコロナによる制約を主軸としつつも、徐々に市場活動の再開を示唆する語彙も現れ始めており、その過程がトピック構造に反映されていると考えられる。

(3) 経済状況の変化と価格・金利への注目 : P3

P3 では、トピック構造がより金融環境やマクロ経済条件に関する関心へとシフトしていた。トピック1では「物件」「価格」「郊外」「都心」など、

地理的選好(郊外 vs 都心)や価格水準の推移に関する話題が多く含まれており、住宅購入意欲の変化や供給側の動向に関する評価が読み取れる。また、トピック2では「価格」「上昇」「金利」「建築」「高騰」などが挙がっており、住宅ローン金利や建築コストの上昇といったマクロ経済の影響が、実需層や業者の市場認識に変化をもたらしていた可能性がある。

(4) 期間別にみる話題の遷移と構造の変容

P1の段階では、低金利下における価格上昇と需要の活発化が主要な話題となっていたが、P2ではコロナによる外的制約が市場の語彙を大きく支配した。P3に入ると、価格や金利、資材コストの上昇といったマクロ経済ショックを反映する語が目立つようになり、より複雑で構造的な市場認識へと変化していたことがうかがえる。

このように、LDAトピックモデルは、時間軸に沿った話題の変化を自然言語レベルで把握する手段として有効であり、記述コメントに内在する「関心の転換」を可視化する分析手法としての可能性を示している。なお、各期間(P1~P3)におけるLDAトピックモデルの可視化図(pyLDAvisによるHTMLファイル)は、オンライン補足資料として脚注に掲載している⁵。

5. 業種別に見た話題の特徴

本節では、業種別(住宅・宅地分譲業、不動産流通業、ビル賃貸業)の記述コメントについてLDAを適用し、業種ごとの特徴的な話題(トピック)を抽出した。トピック数は各業種とも5とし、語の「トピック内出現割合 $P(w|topic)$ 」に基づいて代表語を抽出した結果を表3に示す。分析には、文書頻度の高い語の出現パターンに基づくLDAモデル($\lambda = 1.0$)を使用している。各業種の分析結果を以下に概説する(可視化図は脚注5を参照)。

⁵ 各期、および各業種のpyLDAvisで作成した可視化ページは[土地総合研究所ホームページ](#)からファイルをダウンロードし、ローカルに保存したうえで閲覧可能である。

表3 不動産業者の記述式コメントに関するLDA分析結果（業種別・全期間）

業種	トピック	トピックごとの主要語彙（出現確率上位）
分譲業	トピック1	0.091*物件 + 0.032*価格 + 0.023*郊外 + 0.021*都心 + 0.021*エリア + 0.018*堅調 + 0.016*顧客 + 0.015*傾向 + 0.014*中心 + 0.013*推移
	トピック2	0.037*モデルルーム + 0.034*来場 + 0.031*コロナ + 0.022*販売 + 0.021*影響 + 0.021*傾向 + 0.02*住宅 + 0.018*価格 + 0.018*購入 + 0.015*減少
	トピック3	0.036*物件 + 0.031*購入 + 0.028*消費 + 0.027*価格 + 0.025*増税 + 0.025*需要 + 0.023*マンション + 0.022*慎重 + 0.017*立地 + 0.016*可能
	トピック4	0.042*住宅 + 0.036*賃貸 + 0.033*価格 + 0.025*マンション + 0.025*上昇 + 0.02*金利 + 0.019*分譲 + 0.015*開発 + 0.014*中古 + 0.014*用地
	トピック5	0.026*必要 + 0.024*開発 + 0.023*対応 + 0.02*普及 + 0.019*コロナ + 0.016*ニーズ + 0.016*商品 + 0.016*生活 + 0.016*柔軟 + 0.016*提案
流通業	トピック1	0.082*売却 + 0.081*依頼 + 0.047*購入 + 0.036*価格 + 0.032*物件 + 0.021*希望 + 0.016*相続 + 0.014*土地 + 0.014*住宅 + 0.014*中古
	トピック2	0.039*購入 + 0.027*販売 + 0.027*価格 + 0.026*マンション + 0.024*傾向 + 0.021*動向 + 0.021*成約 + 0.019*上昇 + 0.018*土地 + 0.018*物件
	トピック3	0.04*物件 + 0.033*価格 + 0.028*不動産 + 0.022*今後 + 0.02*動き + 0.015*状況 + 0.015*上昇 + 0.014*需要 + 0.013*住宅 + 0.009*取引
	トピック4	0.063*購入 + 0.036*依頼 + 0.031*傾向 + 0.03*物件 + 0.028*希望 + 0.023*売却 + 0.021*減少 + 0.021*件数 + 0.017*価格 + 0.016*増加
	トピック5	0.037*コロナ + 0.029*影響 + 0.025*物件 + 0.022*動き + 0.021*不動産 + 0.019*景気 + 0.019*増加 + 0.016*販売 + 0.014*成約 + 0.012*価格
ビル賃貸業	トピック1	0.033*テナント + 0.031*コロナ + 0.029*新規 + 0.024*店舗 + 0.023*オフィス + 0.022*状況 + 0.022*動き + 0.018*影響 + 0.018*契約 + 0.016*ビル
	トピック2	0.032*ビル + 0.026*テナント + 0.017*賃料 + 0.013*オフィス + 0.012*以上 + 0.012*予定 + 0.012*募集 + 0.011*今後 + 0.011*動き + 0.01*10
	トピック3	0.035*ビル + 0.03*物件 + 0.024*賃料 + 0.02*新規 + 0.019*テナント + 0.019*移転 + 0.017*供給 + 0.017*需要 + 0.016*減少 + 0.014*状況
	トピック4	0.042*テナント + 0.033*賃料 + 0.019*影響 + 0.015*ビル + 0.013*コロナ + 0.013*企業 + 0.011*増加 + 0.011*面積 + 0.01*希望 + 0.01*駅前
	トピック5	0.023*コロナ + 0.022*状況 + 0.021*テナント + 0.016*不足 + 0.014*宣言 + 0.013*人手 + 0.013*事態 + 0.011*緊急 + 0.01*安全 + 0.01*感染

(1) 住宅・宅地分譲業

住宅・宅地分譲業では、消費増税や金利動向、価格変動に関する語が多く、住宅購入者側の需要・判断材料に関する話題が中心となっていた。トピック1では「都心」「郊外」「価格」「堅調」など、立地による価格差や需要の偏在が示唆され、都市部と周辺部で異なる販売動向が話題になっていたことがうかがえる。トピック2では「モデルルーム」「来場」「コロナ」など、販促活動や顧客接点に関する話題が抽出されており、コロナ禍の来場減少や需要の変化が影響していると考えられ

る。また、トピック3は「物件」「購入」「消費」「価格」「増税」などが上位語であり、購入判断における価格感度や消費税の影響が強く反映されている。

(2) 不動産流通業

不動産流通業では、トピック2では「マンション」「購入」「価格」「傾向」「動向」「成約」などが上位であり、市場全体の取引動向や需給の変化を示す内容が見られる。一方、トピック1で「売却」「依頼」「価格」「希望」などの語が多く、顧客か

らの売却依頼や価格に関する話題が主に取り上げられていた。「相続」や「中古」の語も含まれており、高齢者の住み替えや相続対応といった実需系の取引が中心にあると考えられる。また、トピック5に含まれる「コロナ」「景気」「販売」「成約」などは、景気全般への影響を反映した広範な話題として解釈できる。

(3) ビル賃貸業

ビル賃貸業では、「ビル」「テナント」「賃料」「オフィス」などの語が全トピックに分布し、入居者動向と賃料に関するテーマが繰り返し語られている。特にトピック1では「コロナ」「新規」「店舗」「契約」などの語が上位に現れ、コロナによる契約動向の変化や空室問題が主要な話題となっていたことがわかる。

また、トピック4の「企業」「影響」「コロナ」「面積」などは、企業の動向やコロナ禍におけるオフィス縮小の傾向を示唆しており、長期的なオフィス需要の見通しに不透明感があったことが読み取れる。

6. おわりに

本節では、これまでの期間別分析（第4節）および業種別分析（第5節）で明らかとなった語彙傾向と話題構造の違いを総合し、不動産業者のコメントに内在する特徴を俯瞰する。

まず、期間別分析では、2020年前後を境としたコロナ禍の影響が明確に読み取れた。P1（2016～2019年）においては、「価格」「消費」「増税」など、市場の需給やマクロ経済環境に関するトピックが優勢であったのに対し、P2（2020～2022年）ではコロナおよびその影響に関連する語彙が複数のトピックにわたって出現していた。P3（2023～2025年）では、「金利」「建築」「高騰」といった新たな制約要因が顕在化しており、コロナからの回復過程と新たな局面への移行が確認できる。これらの変化は、LDAの可視化においても、時期ごとのトピック分布の広がりや中心性の変化として視覚的に捉えることができた。

業種別分析では、各業種が持つ業務特性に起因する語彙傾向の差異が顕著であった。分譲業では「モデルルーム」「来場」「価格」「上昇」といった語が頻出し、販売活動や市場環境の影響を色濃く反映していた。流通業では「依頼」「売却」「相続」など、所有者の売却意思と買主の需要動向の両面に関わる語が多く、需給マッチングに関する視点が強く表れていた。ビル賃貸業では「テナント」「賃料」「契約」などの語が中心となり、賃貸契約や空室動向への関心が優勢であった。これらは、各業態が接する顧客層や契約形態の違いを反映したものと解釈できる。

これらの結果を総合すると、不動産市場に関する業者コメントは、時期や業態によって取り上げられる話題や語彙が大きく異なることが明らかとなった。たとえば、価格や購入といった語彙は全体を通して一定の重要性を持つが、分譲業ではモデルルームとの関係性を通じて語られ、流通業では依頼や成約との関係性に焦点が置かれ、ビル賃貸業では価格ではなく賃料として言及されるなど、語の用いられ方に違いが見られた。加えて、コロナ禍の影響も各業態に異なる形で表れており、分譲業では販売活動の制約として、ビル賃貸業ではテナント撤退や新規契約減少の要因として、それぞれの構造の中で異なる意味を持っていた。

最後に、本研究の限界と今後の課題について言及する。第一に、コメントには「慎重な姿勢」「先行き懸念」など定量化が難しい評価的表現が多く、また「価格が高い」「上昇中」「高騰」といった表現にもばらつきがある。今回は語の出現傾向と共起関係に基づく分析にとどめたが、今後は文単位での意味タグ付けや極性分析（ポジティブ/ネガティブ）の導入が求められる。

第二に、本稿ではLDAトピックモデルおよび共起ネットワークを主たる分析手法として用いたが、トピック数の設定やストップワードの除去に関しては一定の恣意性を伴う。第三に、業種別・期間別を交差させた分析は、サンプル数の制約から行っていない。第四に、生成AI（GPT-4o）を前処理に用いた点特徴的だが、再現性や透明性の担保

という点では、引き続き検討が必要である。

課題は多く残されているものの、本稿では、実務的な文脈に基づく業況コメントを対象として、自然言語処理を用いてその語彙・話題の構造を再現性のある手続きにより構造的かつ定量的に整理・表現する手法を示した。なお、本研究で抽出された話題の多くは、業界関係者にとって直感的に理解可能な内容であったが、それらの認識を定量的に表現できた点で、実務と研究の橋渡しを試みるものといえる。

謝辞

本研究は JSPS 科研費 JP20K04986 の助成を受けたものである。

参考文献

- 1) Du, K., Zhao, Y., Mao, R., Xing, F., & Cambria, E. (2025) "Natural language processing in finance: A survey", *Information Fusion*, 115.
- 2) 伊藤夏生(2023)「コロナ禍が不動産業に与えたインパクト—不動産業業況等調査からの分析—」, 土地総合研究 31 卷(1 号), 土地総合研究所
- 3) 大越利之(2016)「不動産業の景況感の特徴—不動産業業況指数と景気関連指標との関係—」, 土地総合研究, 24 卷(3 号), 土地総合研究所
- 4) 大越利之(2013)「不動産業業況等調査のとらえ方 — 20 年の蓄積データを振り返る —」, 土地総合研究, 21 卷(3 号), 土地総合研究所