

# 基礎研 レポート

## 日本年金機構をめぐる不祥事のほか、 積立金に関する前財相の発言も話題に

「年金」を含むツイートに含まれるリンクの基礎分析(2021年10月)

年金総合リサーチセンター公的年金調査室長・上席研究員 中嶋 邦夫  
(03)3512-1859 nakasima@nli-research.co.jp

### 1 —— 本稿の問題意識と分析対象：「年金」を含むツイートは、何を契機に発信されているか？

ツイートとは、Twitter に投稿されるメッセージ（発言）である。文字数が、基本的に全角で 140 文字（半角で 280 文字）までと制限されており、「つぶやき」とも呼ばれる。近年はマスコミの報道や国会審議などでも取り上げられているが、ツイートには、熟考の上で投稿されたもの、反射的に投稿されたもの、宣伝、広報、プログラムによって投稿されたものなど、様々なものが混在している。

本稿では、前稿<sup>1</sup>で行った 2021 年 9 月中下旬に発信された「年金」を含むツイートの基礎的な投稿の状況やツイートに対する関心の動向分析に続き、テキストマイニングの手法を使ったツイート内容の分析に先立ってツイートに含まれるリンクを分析する。分析の対象は、2021 年 10 月に発信された筆者の研究領域の 1 つである「年金」を含むツイートである（図表 1）。後日実施するツイート内容の分析が最終目的であるため、分析対象とするツイートに単純なリツイート（投稿者自身の発言を含まないもの）と同日の他の対象ツイートと冒頭 50 字が同一なものを含んでいない。言い換えれば、分析対象とするツイートは投稿者自身の何らかの態度が示されている発言、ととらえることも出来る。

図表 1 本稿が分析対象としたツイート

基本的な対象	・ ツイート中に「年金」を含むツイート
除外対象	・ 単純なリツイート（投稿者自身の発言を含まないもの） ・ 同日の他の対象ツイートと冒頭50字が同一なもの
投稿期間	・ 2021年10月1日～31日。
データ取得時期	・ 基本的に発信日翌日の早朝。
分析対象件数	・ 162,249件

<sup>1</sup> 中嶋邦夫(2021)「大臣会見でツイート数と投稿者数が急増もリツイート数はごく少数のツイートが左右：「年金」を含むツイートの基礎的な動向分析（2021年9月中下旬）」『基礎研レポート』2021-11-09号。

以下では、発信傾向の指標として、ツイート数（発言数）、投稿者数、いいね数、リツイート数を確認する。ツイート数は、前述のとおり単純なリツイート等を含んでいない点に注意が必要である。投稿者数は、分析対象としたツイートのユーザー名を日付などの集計区分ごとに名寄せした件数であり、同一の集計区分において同一のユーザー名で複数のツイートを投稿していても1名と数えている。いいね数とリツイート数は、分析対象としたツイートのいいね数とリツイート数を集計区分ごとに合計した件数である。いいね数とリツイート数は時間の経過とともに変化するが、データ取得時点の値を集計している。ツイートによって投稿時点からデータ取得時点までの間隔が異なるため、いいね数とリツイート数はツイート間で厳密には比較できない。ツイートへの関心を大雑把に示す指標として参照している。

## 2 —— 前稿で発見した傾向の検証

### 1 | 本章の問題意識：前月に見られた傾向は、今月にも当てはまるか

前稿では、2021年9月中下旬に発信された「年金」を含むツイートに対する分析から、(1)年金に関するニュースがあるとツイート数だけでなく投稿者数も増える、(2)いいね数とリツイート数は年金に関するニュースとは必ずしも関係がない、(3)いいね数やリツイート数が多い少数のツイートがいいね数全体とリツイート数全体の大部分を占める、などの傾向を発見した。そこで、これらの傾向が他の時期にも当てはまるかを確認するために、2021年10月の投稿に対しても前稿と同じ枠組みの分析を行う。

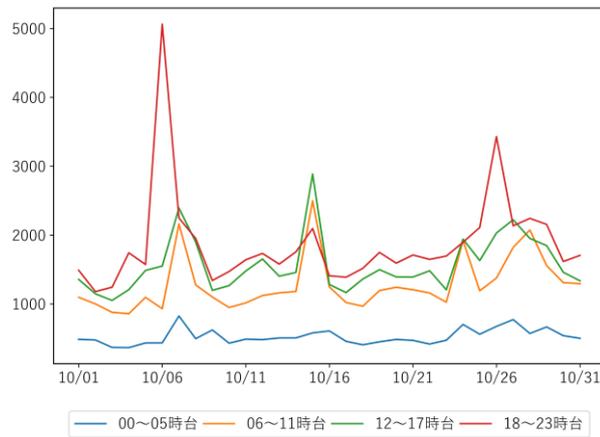
### 2 | 投稿日時ごとの傾向や特徴：前月と同様にニュースへの反応は見られるが、深夜の反応は低調

前述した(1)年金に関するニュースがあるとツイート数だけでなく投稿者数も増えるかと、(2)いいね数とリツイート数は年金に関するニュースとは必ずしも関係がないかを確認するために、前稿と同様に投稿日と投稿時間帯（0時を起点とする6時間単位）の組み合わせごとの傾向や特徴を見たのが図表2である。

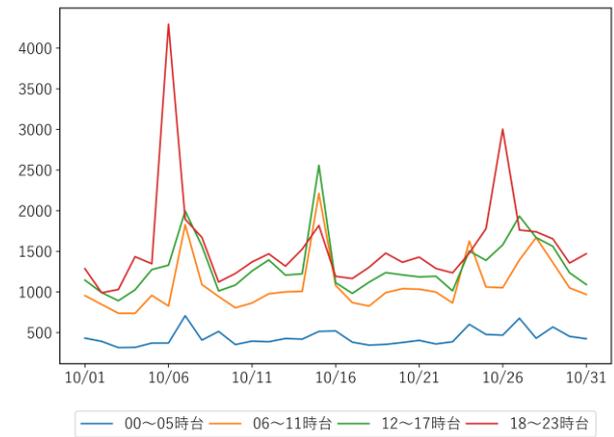
まず、ツイート数と投稿者数（図表2の左上段と右上段）は、10月6～7日辺りと10月15日と10月26日辺りに多くなっていた。10月6～7日辺りには年金振込通知書の誤送付、10月15日は年金の支給日であること、10月24日辺りには年金積立金の運用に関する前財務大臣の街頭演説での発言、10月26日辺りには日本年金機構で活用されていない市町村向け貸出用PCが多数あること、に関するニュースに影響されてツイート数と投稿者数が多くなった可能性がある。このように、(1)年金に関するニュースがあるとツイート数だけでなく投稿者数も増える傾向は、前稿で分析した9月中下旬に投稿されたツイートに関する分析と同じであった。

図表2 投稿日×時間帯ごとの傾向や特徴

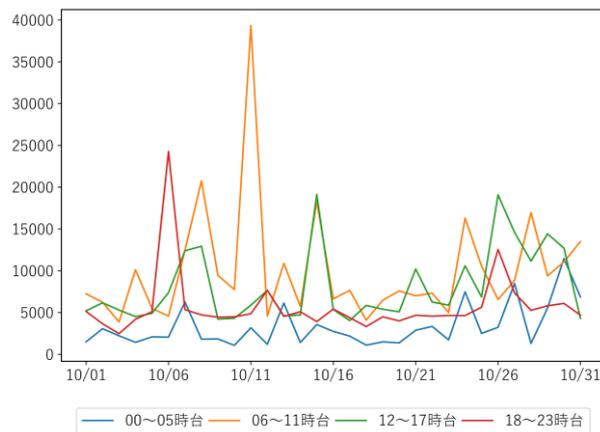
ツイート数



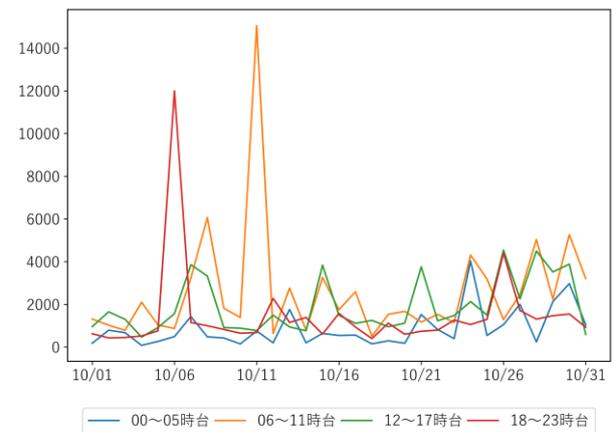
投稿者数



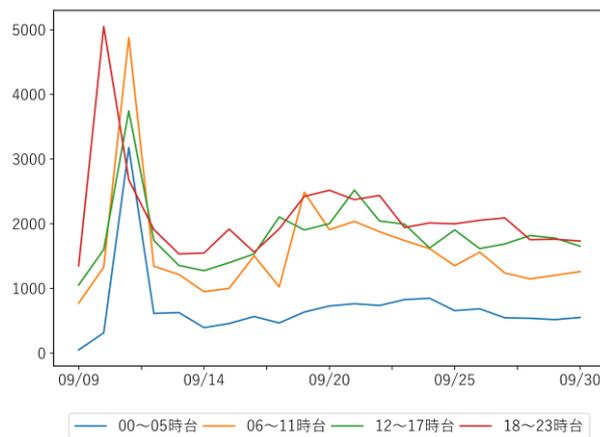
いいね数



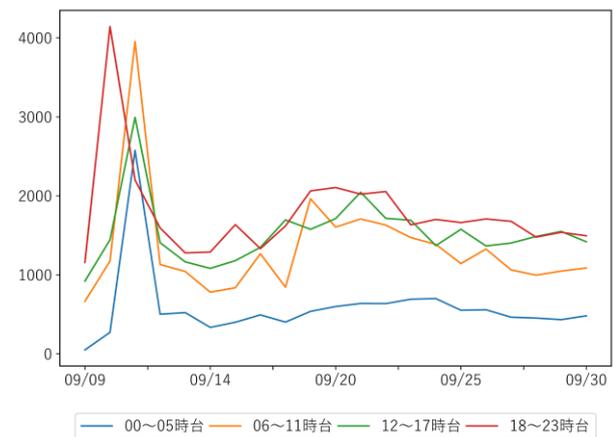
リツイート数



(参考) 9月中下旬のツイート数



(参考) 9月中下旬の投稿者数



(注1) いいね数とリツイート数は、ツイートごとに投稿時点からデータ取得時点までの間隔が異なるため、大雑把な指標。

ただし、前稿で分析した9月中下旬に投稿されたツイートに関しては、9月10～11日に発生した厚生労働大臣が定例記者会見で年金制度改正についての発言に関するツイート数と投稿者数の増加において、9月10日の18～23時台から9月11日の18～23時台にかけて00～05時台も含めてツイート数

と投稿者数が多かった（図表2の左下段と右下段）。これに対して、本稿で分析対象としているツイートに関しては、10月7日の00～05時台において、他の日の00～05時台よりも多い傾向が見られるものの、9月11日の00～05時台ほどの多さは見られなかった。この背景には、2つの問題に対する人々の関心の強さの違いのほか、9月11日は土曜で10月7日は木曜という曜日の違いが影響した可能性も考えられる。

次に、図表2の左中段と右中段にあるいいね数とリツイート数を見ると、ツイート数と投稿者数が多かった時間帯に加えて、10月11日の06～11時台にも多くなっている。このように、(2)いいね数とリツイート数は年金に関するニュースとは必ずしも関係していない傾向は、前稿で分析した9月中下旬に投稿されたツイートに関する分析と同じであった。なお、いいね数とリツイート数は、いいねの付与やリツイートを行った日時ではなく、いいねの付与やリツイートを行った対象のツイートの投稿日時で集計している点に、注意して見る必要がある。

### 3 | 投稿日といいね数・リツイート数区分ごとの傾向や特徴：ごく少数のツイートが全数を左右

前述した(3)いいね数やリツイート数が多い少数のツイートがいいね数全体とリツイート数全体の大部分を占めるかを確認するために、前稿と同様に投稿日と各ツイートのいいね数やリツイート数の区分を組み合わせて傾向や特徴を見たのが、図表3である。

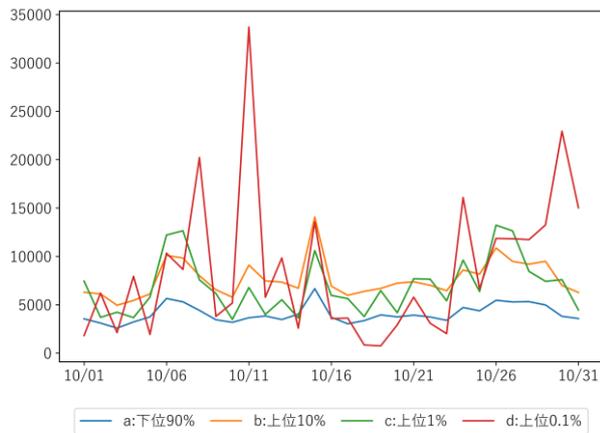
図表3の左上段と右上段を見ると、前述した、ニュースと関連せずがいいね数やリツイート数が多い10月11日を始めとして、それぞれの上位0.1%以上に該当するツイートのいいね数やリツイート数がいいね数やリツイート数の大半を占めている。図表3の左中段と右中段（ともに対数軸）を見ると、それぞれの上位0.1%以上に該当するツイートは各日とも概ね数件（10の0乗＝1件と10の1乗＝10件の間）であり、いいね数やリツイート数が多い10月11日に上位0.1%以上に該当するツイートが多いわけでもない。これらの状況から、(3)いいね数やリツイート数が多い少数のツイートがいいね数全体とリツイート数全体の大部分を占めるという傾向は、前稿で分析した9月中下旬に投稿されたツイートに関する分析と同じであったと言える。

なお、各ツイートのいいね数の区分は、いいね数の分布の上位0.1%以上（各ツイートのいいね数が554件以上）、上位1%以上（各ツイートのいいね数が65件以上）、上位10%以上（各ツイートのいいね数が6件以上）、下位90%未満（各ツイートのいいね数が6件未満）の4区分である。また、各ツイートに対するリツイート数の区分は、分布の上位0.1%以上（各ツイートに対するリツイート数が118件以上）、上位1%以上（各ツイートに対するリツイート数が14件以上）、上位10%以上（各ツイートに対するリツイート数が1件以上）、下位90%未満（各ツイートに対するリツイート数が0件）の4区分である。

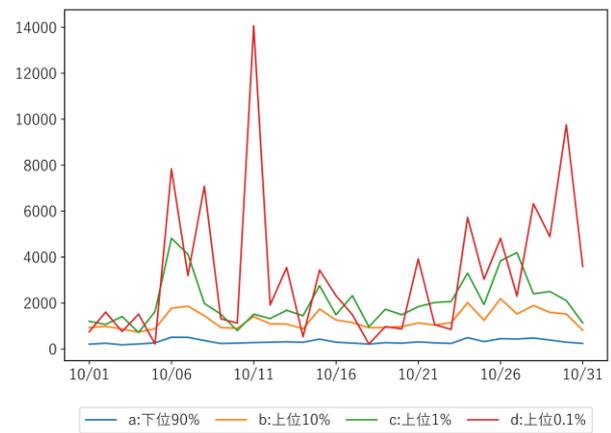
前稿で分析した9月中下旬に投稿されたツイートと比べると、各区分の境界となる件数はすべて減少していた。この傾向は、図表3の左上段と右上段に示した10月のいいね数やリツイート数の縦軸目盛が、図表3の左下段と右下段に示した9月中下旬のいいね数やリツイート数の縦軸目盛の概ね半分の水準になっていることから伺える。

図表3 投稿日×いいね数とリツイート数の区分(それぞれの上位パーセンタイル)ごとの傾向や特徴

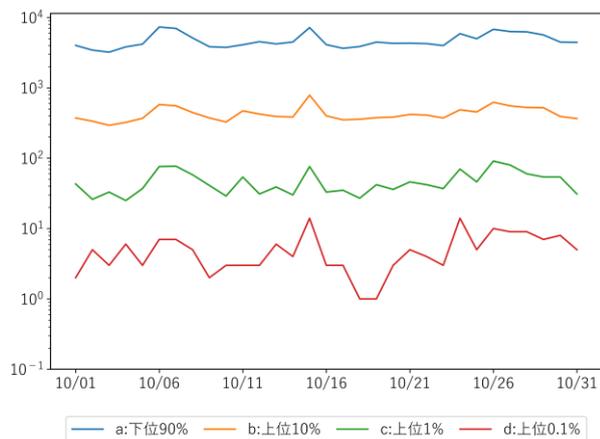
いいね数



リツイート数



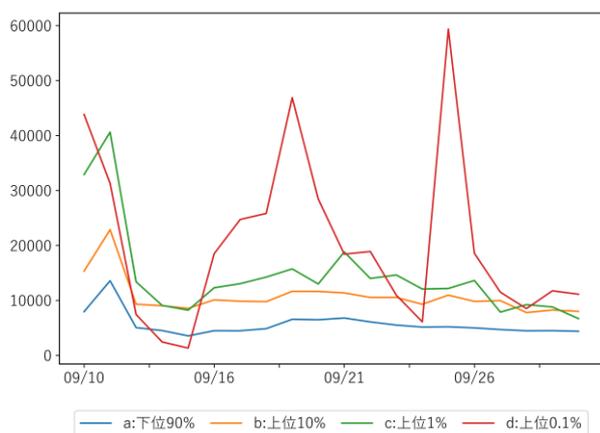
いいね数の区分ごとのツイート数 (対数軸)



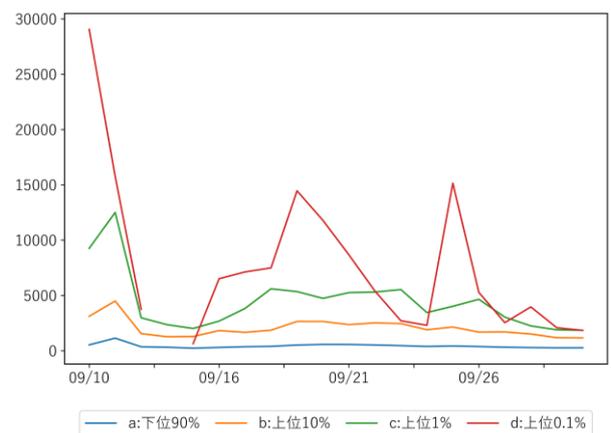
リツイート数の区分ごとのツイート数 (対数軸)



(参考) 9月中下旬のいいね数



(参考) 9月中下旬のリツイート数



- (注1) いいね数とリツイート数は、ツイートごとに投稿時点からデータ取得時点までの間隔が異なるため、大雑把な指標。  
 (注2) いいね数とリツイート数の区分はそれぞれの上位パーセンタイルで区切っているため、境界となる件数はいいね数の区分とリツイート数の区分で異なる。また、9月中下旬と10月でも境界となる件数は異なる。

### 3 —— ツイートに含まれるリンクの傾向や特徴

#### 1 | 本章の問題意識: ニュースがある日にツイートが多いのは、記事ページから投稿しているためか?

前章で確認したとおり、9月中下旬と10月の双方で、年金に関するニュースがあるとツイート数だけでなく投稿者数も増える傾向があった。この要因に関して、**前稿**では「Webで公開されるニュース記事の脇にツイートを投稿するためのボタンが配置されていることが多いため、関心を持ったニュースを拡散するツイートや、記事を参照・引用した上で自分の意見を述べるツイートが散見される」と記載したが、具体的な確認には至らなかった。

そこで本章では、Web上のニュース記事がツイートを投稿する契機になっているかを確認するために、ツイートに含まれるリンク(Webページの参照)の傾向や特徴を分析する。ツイートに含まれるリンクのURLがWeb上のニュース記事のものであれば、Web上のニュース記事がツイートを投稿する契機になっている可能性が高いと言えよう。

#### 2 | ツイートに含まれるリンク数の傾向や特徴: ツイートの75%はリンクなし、25%はリンク1個

本節では、1つのツイートに含まれるリンクの数ごとに傾向や特徴を確認する。

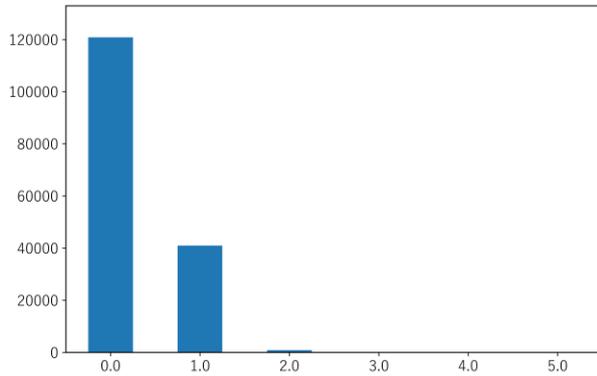
まず、1つのツイートに含まれるリンクの数ごとにツイート数を集計したのが、**図表4**の上段である。本稿の分析対象である約16万件のツイートのうち、リンクを含んでいないもの(0個)が約75%(約12万件)、リンクが1個だけのツイートが約25%(約41,000件)であり、これらでツイートのほとんどを占めていた。少数ではあるが、リンクが複数あるものは合計で約2,700件あり、リンクの数は最大でも5個であった。リンクの数が2個、3個、4個と増えて行くのに従って該当するツイートの数は少なくなっているが、リンクを5個含んでいるツイートはリンクを4個含んでいるツイートよりも多かった。

リンクの数ごとに投稿者数を集計したのが、**図表4**の中段である。リンク数が0~4個についてはツイート数と同じ傾向だが、リンクを5個含んでいるツイートの投稿者数は、ツイート数の傾向とは違い、リンクを4個含んでいるツイートの投稿者数よりも少なかった。

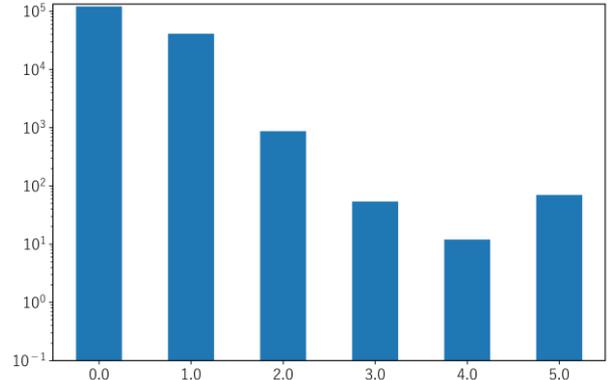
リンクの数ごとにいいね数とリツイート数を集計すると、リンク数が0~3個についてはツイート数と同じ傾向だが、リンク数が4~5個についてはいいね数とリツイート数に大きな差がなく、ツイート数の傾向と投稿者数の傾向の中間的な傾向が見られた。

図表4 1つのツイートに含まれるリンクの数ごとの傾向や特徴

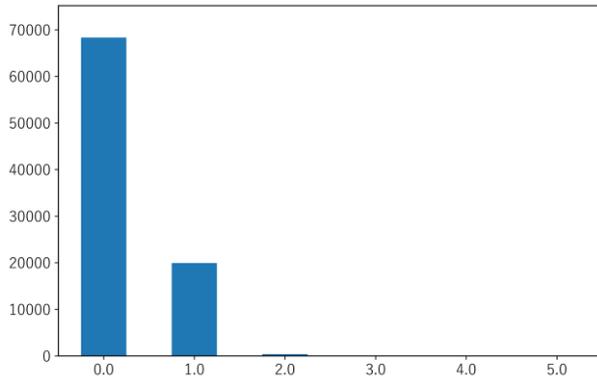
ツイート数



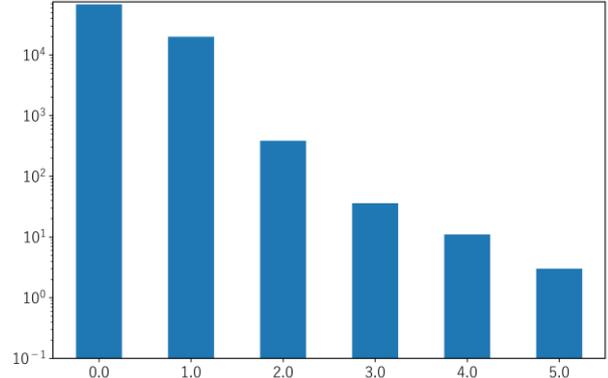
ツイート数 (対数軸)



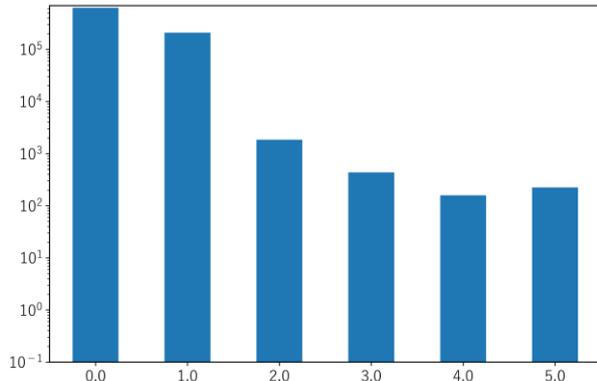
投稿者数



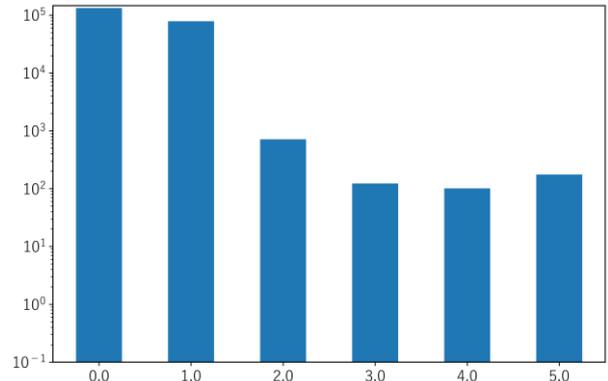
投稿者数 (対数軸)



いいね数 (対数軸)



リツイート数 (対数軸)



(注1) いいね数とリツイート数は、ツイートごとに投稿時点からデータ取得時点までの間隔が異なるため、大雑把な指標。

### 3 | リンク先 Web サイトの傾向や特徴: 引用ツイートや Yahoo!ニュースに加え、Web サイトの宣伝も多数

(1) リンク先 Web サイト (ドメイン) ごとの傾向や特徴: 少数の投稿者による参照が多いサイトあり  
次に、ツイートに含まれるリンクが参照している Web サイト (ドメイン) ごとの傾向や特徴を確認する。前節の分析でツイートに含まれるリンクの数は0個もしくは1個で約100%を占めたため、以

下ではツイートに含まれるリンクのうち1個目だけに注目する。

リンク先の Web サイト(ドメイン)を名寄せして (データとして unique なものを) 集計したところ、分析対象のツイートで参照されている Web サイトは約 2200 か所あった。このうち参照しているツイート数が多い上位 20 件 (上位約 1%) を列挙したのが図表 5 である。

前節で述べたとおり、リンクを含んでいないツイートが約 12 万件 (分析対象ツイートの約 75%) あり、リンク先の Web サイトとして最も多かったのが「twitter.com」で、ツイート数は約 1 万件 (分析対象ツイートの約 6%、リンクを含む分析対象ツイートの約 25%) である。本稿の分析対象には単純なリツイート (投稿者自身の発言を含まないもの) を含めていないため、本稿の分析対象の中でリンク先が「twitter.com」となっているのは、いわゆる引用リツイート (投稿者自身の発言を含むもの) である。投稿者が「twitter.com」のサイトやアプリで表示される情報を一種のニュースととらえ、それに対する自身のコメントを発信しているものと思われる。

次に多かったのが「news.yahoo.co.jp (Yahoo!ニュース)」で、この Web サイトをリンク先としているツイートは約 6500 件 (分析対象ツイートの約 4%、リンクを含む分析対象ツイートの約 16%) であった。このサイトではニュース記事の脇にツイートを投稿するためのリンク (アイコン) が設置されているため、これを利用して、投稿者が印象に残った記事の拡散や、記事に対する自身のコメントの発信を行っているものと思われる。

以上の「リンクなし」「twitter.com」「news.yahoo.co.jp (Yahoo!ニュース)」までで、分析対象ツイートの約 85%、リンクを含む分析対象ツイートの約 40%を占める。

図表 5 リンク先 Web サイトごとの傾向や特徴 (ツイート数上位 20 件)

リンク先ドメイン名	ドメインのトップページのタイトル(一部)	ツイート数	投稿者数	いいね数	リツイート数	1人当たり ツイート数
リンクなし		120,882	68,350	629,794	133,127	1.8
twitter.com	「いま」を見つけよう / Twitter	10,484	7,298	85,563	29,837	1.4
news.yahoo.co.jp	Yahoo!ニュース	6,553	4,668	28,357	11,731	1.4
the-end.killerz13s-d	夢を実現させる-make a dreams com	2,218	4	2,202	227	554.5
youtu.be	YouTube	1,693	748	11,446	3,212	2.3
www.asahi.com	朝日新聞デジタル：朝日新聞社のコ	1,388	1,192	19,180	7,618	1.2
peing.net	Peing(ペイング) - 質問箱 - 匿名で質問	1,235	215	452	22	5.7
www.nikkei.com	日本経済新聞	957	655	4,822	1,259	1.5
bit.ly	URL Shortener - Short URLs & Cust	683	113	576	311	6.0
www.timeticket.jp	タイムチケット   個人の時間を30分単位	599	2	20	0	299.5
is.gd	is.gd - a URL shortener. Mmmm, ta	482	26	439	14	18.5
a.msn.com	MSN Japan - ニュース, 天気, メール (O	448	309	483	163	1.4
www3.nhk.or.jp	NHK	418	325	2,024	1,505	1.3
mainichi.jp	毎日新聞	331	279	1,587	942	1.2
jmtj.jp	ジモティー 無料の広告掲示板	312	3	15	0	104.0
www.rui.jp	るいネット - 探求の輪	295	57	161	106	5.2
ameblo.jp	アメーバブログ (アメブロ)   Amebaで	260	141	963	173	1.8
sr-learn.com	社労士過去問どっとこむ   社会保険労	228	1	14	1	228.0
nordot.app	Nordot   ノアドット株式会社	221	195	2,876	1,834	1.1
www.nenkin.go.jp	日本年金機構	205	146	1,083	451	1.4
www.youtube.com	YouTube	177	99	498	128	1.8

リンク先の Web サイトとして 3 番目に多かった Web サイトでは、この Web サイトをリンク先としているツイートが約 2200 件あるものの、この約 2200 件のツイートはわずか 4 名の投稿者から発信されたものであり、投稿者 1 人当たりの該当ツイート数は 1 か月間で約 550 件に及んだ。このようなツイートは、リンク先の Web サイトを宣伝・周知する目的で投稿されている可能性が高く、ツイート内容の分析を通じて年金に対する人々の態度を知るといふ筆者の最終的な研究目的にそぐわない可能性がある。

そこで、ツイート数が 20 位以下の Web サイトも含めてリンク先の Web サイトごとに集計したところ、その Web サイトをリンク先としているツイートの投稿者が 10 人未満、かつ当該投稿者 1 人当たりのツイート数が 31 件（1 日当たり 1 件）以上のリンク先 Web サイトは、Web サイト数としては 14 件（リンク先 Web サイト総数 2165 件の 0.6%）あり、ツイート数としては 3895 件（リンクありツイート総数 42,004 件の 9%）に及んだ一方、投稿者数は 20 人（リンクありツイートの投稿者数 23,542 人の 0.1%）にとどまった。さらに対象を広げると、投稿者が 10 人未満、かつ当該投稿者 1 人当たりのツイート数が 10 件以上のリンク先 Web サイトは、Web サイト数としては 82 件（リンク先 Web サイト総数の 4%）あり、ツイート数としては 5263 件（リンクありツイート総数の 9%）に及んだ一方、投稿者数は 103 人（リンクありツイートの投稿者数の 0.4%）にとどまった。

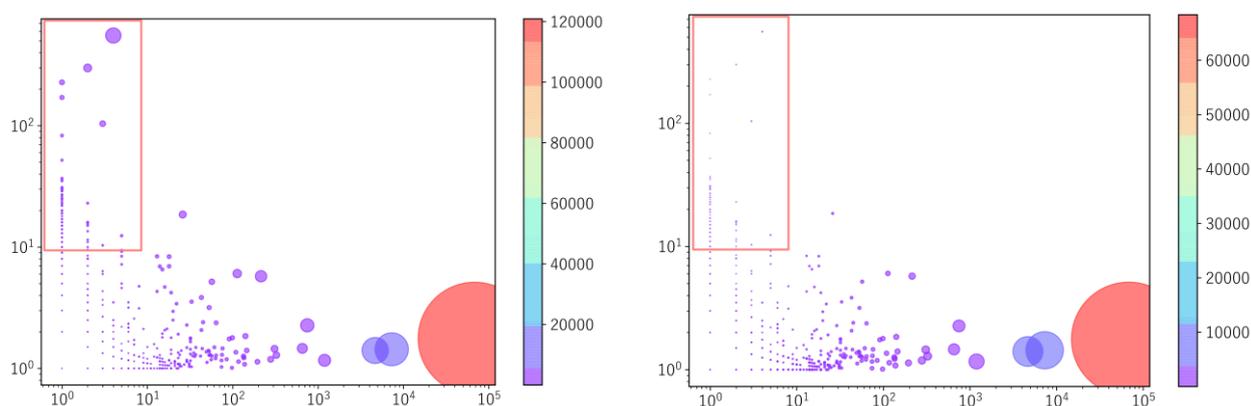
このようなリンク先の Web サイトを宣伝・周知する目的で投稿されている可能性が高いツイートは、投稿者数が少ない一方でツイート数としては一定の影響力を持っているため、ツイート内容を分析する際はこれらのツイートの取り扱い方（除外するかや区分するか）を検討する必要がある。

図表 6 リンク先 Web サイトごとの投稿者数と 1 人当たりのツイート数の傾向や特徴

（横軸＝Web サイトごとの投稿者数、縦軸＝同 1 人当たりのツイート数、ともに対数軸）

ツイート数（円の大きさと色で表現）

投稿者数（円の大きさと色で表現）



（注 1）各点がリンク先の Web サイト。赤枠部分が、少数の投稿者による多数の投稿になっている部分（左図のツイート数は多い（円の大きい）が、右図の投稿者数は少ない（円の小さい））。

また、図表 5 のトップページのタイトルを見ると、報道機関と思われる Web サイトが散見される。前述した Yahoo!ニュースと同様に、報道機関の Web サイトでもニュース記事の脇にツイートを投稿するためのリンク（アイコン）が設置されている傾向があるため、投稿者がこれを利用して印象に残っ

た記事の拡散や記事に対する自身のコメントの発信を行っているものと思われる。

しかし、ツイート数が 20 位以下の Web サイトにも地方紙の Web サイトなどが多数含まれており、約 2200 か所あるリンク先 Web サイトが報道機関か否かを個別に確認することは困難である。そこで、前述の Yahoo!ニュースも含めてトップページのタイトルに「新聞」「ニュース」「News」のいずれかを含む Web サイトを機械的に抽出して集計したところ、Web サイト数としては 204 件（リンク先 Web サイト総数 2165 件の 9%）あり、ツイート数としては 11,802 件（リンクありツイート総数 42,004 件の 28%）、投稿者数は 8689 人（リンクありツイートの投稿者数 23,542 人の 37%）を占めた<sup>2</sup>。

## (2) 投稿日とリンク先 Web サイト区分ごとの傾向や特徴：Web サイトの宣伝はニュースに反応せず

この分析に基づいてリンク先 Web サイトを図表 7 のように区分し、投稿日との関係を確認したのが図表 8 である。

図表 7 リンク先 Web サイトの区分

区分	区分の基準
a:リンクなし	ツイートにリンクを含まない
b:Twitter	リンク先のドメイン名に「twitter.com」を含む
c:新聞等あり	リンク先 Web サイトのトップページのタイトルに「新聞」「ニュース」「News」のいずれかを含む
d:新聞等なし	リンク先 Web サイトのトップページのタイトルに「新聞」「ニュース」「News」のいずれも含まない
e:少数者多投稿	リンク先 Web サイトごとに集計した結果、名寄せ済み投稿者が 10 人未満、かつ投稿者 1 人当たりのツイート数が 10 件以上
f:タイトル不明	リンク先 Web サイトのトップページのタイトルを取得できなかった

(注 1) 区分対象のリンクはツイート中の 1 つ目のリンク。

(注 2) 区分方法の適用順(優先度)は、a→b→e→f→c→d。

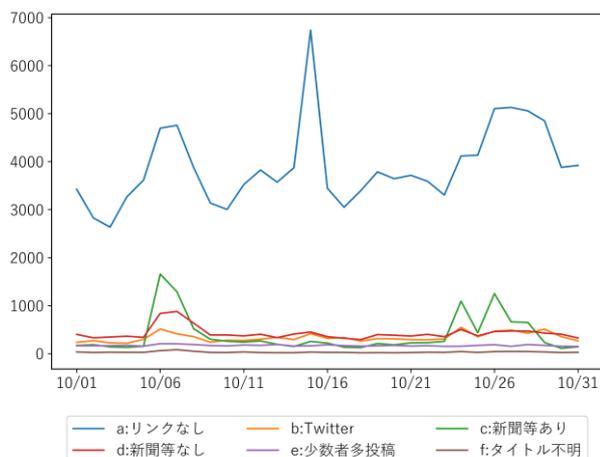
図表 8 の左上段のツイート数を見ると、まず、「a:リンクなし」と他の区分で推移の傾向が異なることが分かる。年金振込通知書の誤送付が話題になった 10 月 6～7 日辺りと、年金積立金の運用に関する前財務大臣の発言が話題になった 10 月 24 日辺りと、日本年金機構で活用されていない市町村向け貸出用 PC が多数あることが話題になったには 10 月 26 日辺りには、ツイート数の大半を占める「a:リンクなし」のほか、「c:新聞等あり」「d:新聞等なし」「b:Twitter」も他の日より多い。これに対して、年金の支給日である 10 月 15 日には「a:リンクなし」が大幅な増加になっているものの、「c:新聞等あり」などでは 10 月 6～7 日辺りや 10 月 26 日辺りほどの増加は見られない。これは、10 月 6～7 日辺りや 10 月 26 日辺りは非日常的なニュースであるため大幅に増加したが、10 月 15 日の年金支給日は定例的な出来事だったためあまり変化しなかったと考えられる。なお、「a:リンクなし」が急増してい

<sup>2</sup> この区分方法では、報道機関のうち一部の放送局や通信社あるいは報道機関の記事を掲載するサイトなどを含まず、逆に「○○新聞」などのタイトルを付けた個人のブログなどを含めてしまう懸念もあるが、客観的に処理するために機械的に割り切った。さらなる精緻化は今後の課題としたい。

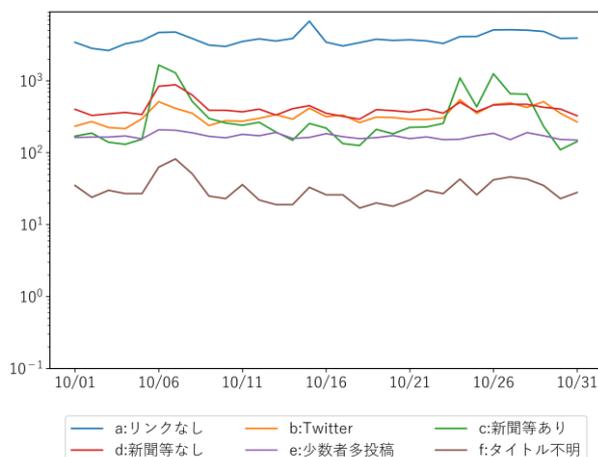
る要因は、ツイート内容の分析で明らかにする必要がある。

図表8 投稿日×リンク先Webサイト区分ごとの傾向や特徴

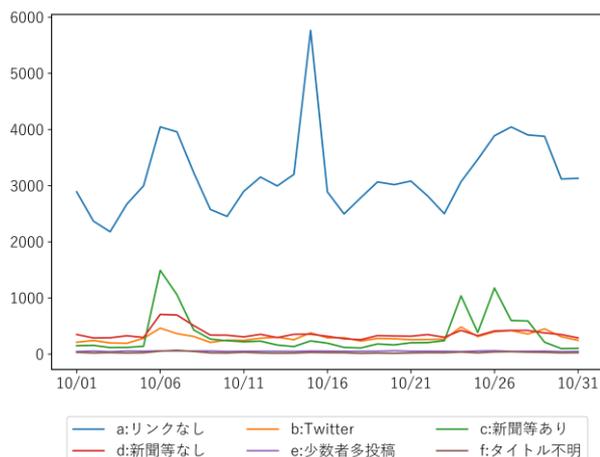
ツイート数



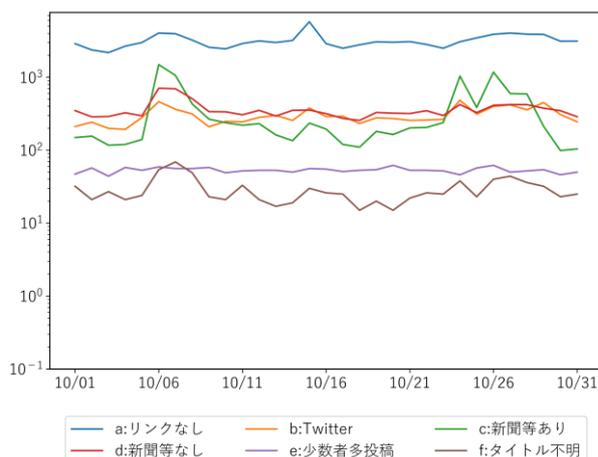
ツイート数 (対数軸)



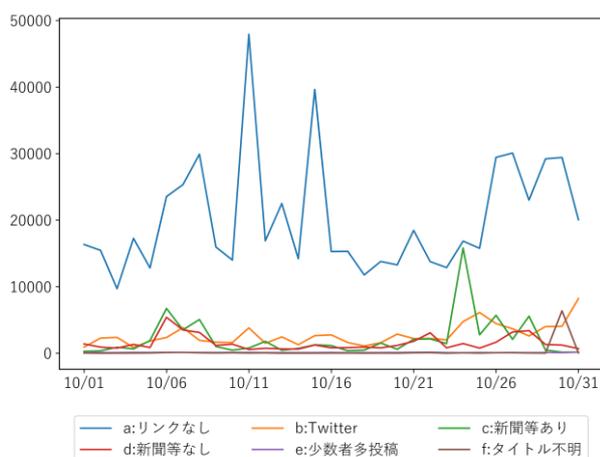
投稿者数



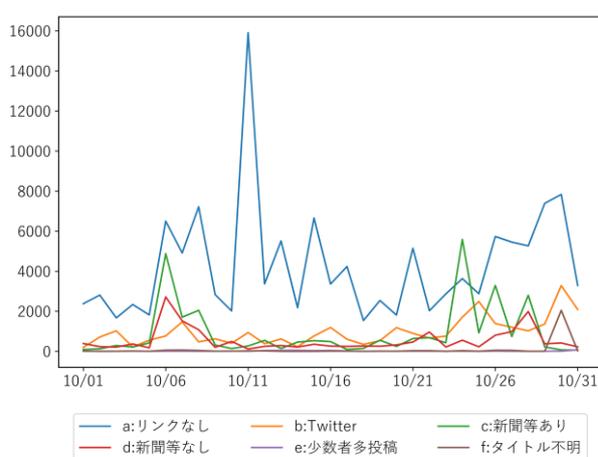
投稿者数 (対数軸)



いいね数



リツイート数



(注1) いいね数とリツイート数は、ツイートごとに投稿時点からデータ取得時点までの間隔が異なるため、大雑把な指標。

次に、図表 8 の右上段のツイート数（対数軸）を見ると、(1)「c:新聞等あり」と「d:新聞等なし」「b:Twitter」の推移の傾向が異なる、(2)「d:新聞等なし」と「b:Twitter」の推移の傾向が似ている、(3)「e:少数者多投稿」は年金に関するニュースとは関係なく常に一定程度で推移している、ということが分かる。

(1)については、年金に関するニュースがある 10 月 6～7 日辺りや 10 月 26 日辺りには「c:新聞等あり」が「d:新聞等なし」や「b:Twitter」を上回っている一方で、他の期間には「c:新聞等あり」が「d:新聞等なし」や「b:Twitter」を下回っている。このことから、「c:新聞等あり」は年金に関するニュースに反応しやすいと言える。

他方で(2)については、「d:新聞等なし」には「c:新聞等あり」に含めきれない報道機関やニュース掲載サイトが含まれ、「b:Twitter」にはニュースに関するツイートが投稿されるために年金に関するニュースがある時期にある程度増え、それと同時に、両者にはニュース以外の情報も掲載されるために年金に関するニュースがない時期にもあまり減少しないと考えられる<sup>3</sup>。

(3)については、前節で確認したとおり「e:少数者多投稿」はリンク先 Web サイトの宣伝や周知である可能性が高いため、年金に関するニュースとは関係なく常に一定程度で推移していると考えられる。

図表 8 左下段のいいね数と右下段のリツイート数を見ると、両者に特徴的な 10 月 11 日の急増は「a:リンクなし」の急増で起きている。他方で 10 月 26 日辺りの増加は、「c:新聞等あり」の増加が一因となった可能性がある。

#### 4 | リンク先の Web ページの傾向や特徴：報道機関記事の参照が多く、年金を含むとニュースに連動

##### (1) リンク先 Web ページごとの傾向や特徴：報道機関記事の参照が多く、話題によっては発信元に偏り

リンク先の Web ページ(URL)を名寄せして（データとして unique なものを）集計したところ、分析対象のツイートで参照されている Web ページ(URL)は約 19,000 件、このうち twitter.com 以外の Web ページは約 12,000 件あった。このうち参照しているツイート数が多い上位 20 件を列挙したのが図表 9 である。

前節で区分した Web サイト区分の列を見ると、上位 20 件のうち 12 件が「c:新聞等あり」で、1 件の「d:新聞等なし」も放送局のサイトであるため、ツイートで参照されている件数が多い Web ページは報道機関や報道機関の記事を掲載する Web サイトが大半を占めていると言える。他方で、上位 20 件のうち 6 件が「e:少数者多投稿」となっており、リンク先の Web サイトを宣伝・周知する目的で投稿されている可能性が高いツイートが上位に食い込んでいる。

リンク先ページのタイトルを見ると、(1)ほぼ同じタイトルが複数回登場している、(2)話題によっては記事の発信元に偏りがある、(3)同じ出来事に対する記事でも用いる単語が違う、(4)「e:少数者多投稿」の Web サイトのページタイトルには年金と直接の関係がないものも含まれる、(5)「c:新聞等あり」

<sup>3</sup> 「c:新聞等あり」に含む報道機関や報道機関の記事を掲載するサイトを精緻化すれば、「d:新聞等なし」が年金に関するニュースに対してより鈍感になる可能性もあるが、「a:リンクなし」が年金に関するニュースにある程度反応していることを踏まえれば、それほど鈍感にならない可能性も考えられる。今後の課題としたい。

の1人当たりのツイート数はほぼ1.0件、という傾向が見られる。

図表9 リンク先 Web ページ(URL)ごとの傾向や特徴 (ツイート数上位20件)

リンク先ページのタイトル (一部)	Webサイト区分	ツイート数	投稿者数	いいね数	リツイート数	1人当たり ツイート数
リンクなし	a:リンクなし	120,882	68,350	629,794	133,127	1.8
「年金増えたのは、運用を株でやったお	c:新聞等あり	637	618	13,699	5,210	1.0
年金通知書の誤送付「97万2千件」(日	c:新聞等あり	571	565	1,313	743	1.0
40代の婚活は人生の瀬戸際!!   タイ	e:少数者多投稿	533	2	17	0	266.5
年金機構 PC1300台活用されず - Yaho	c:新聞等あり	417	411	1,967	1,567	1.0
ずさんなパソコン調達でお蔵入り続出	c:新聞等あり	326	324	553	193	1.0
「年金増えたのは、運用を株でやったお	c:新聞等あり	299	262	1,724	452	1.1
<終活とはなんだ?> 100歳超の日	e:少数者多投稿	282	3	14	0	94.0
単身高齢者の入居困難、20件超拒否さ	c:新聞等あり	244	242	2,828	1,457	1.0
携帯料金実質0円計画好評発動中!	e:少数者多投稿	217	2	202	23	108.5
年金振込通知書 97万件を誤送付 - Ya	c:新聞等あり	199	199	535	391	1.0
リモート、テレワークに最適! お子様のオン	e:少数者多投稿	144	3	145	22	48.0
ずさんなパソコン調達でお蔵入り続出	c:新聞等あり	142	139	320	163	1.0
スマホが100円で使えるキャンペーン好評	e:少数者多投稿	140	3	172	23	46.7
年金通知書 宛名と中身取り違え - Yah	c:新聞等あり	118	117	797	398	1.0
日本年金機構がノートPC471台を二重	c:新聞等あり	116	116	397	189	1.0
「年金振込通知書」約97万件に別人の	d:新聞等なし	113	113	306	305	1.0
小室眞子さんの怒気に満ちた回答...許	c:新聞等あり	107	105	395	92	1.0
年金通知書 再送付に5千万円超 - Yat	c:新聞等あり	97	96	441	251	1.0
バーチャルオフィス東京なら【ユナイテッド	e:少数者多投稿	93	1	1	0	93.0
年金通知書の誤送付 厚労相謝罪 - Ya	c:新聞等あり	88	88	199	94	1.0

(1)は、ほぼ同一の記事について、報道機関の記事を掲載する Web サイトの記事を参照しているツイートと、報道機関の Web サイトの記事を参照しているツイートが存在するためである。また、上位20件には入っていないが、リンク URL の末尾に「?utm\_source=」などの URL パラメータがつくことで、Web ページとしては同一であるものの別の URL として集計しているものも存在する<sup>4</sup>。

(2)は、年金積立金の運用に関する記事は、掲載先が違うものの1つの報道機関による記事が上位20件に入っている一方で、振込通知書の誤送付に関する記事は複数の報道機関による記事が上位20件に入っていることを指している。このことから、普段は特定の報道機関だけが注目されている訳ではないものの、記事のタイトルや内容によっては特定の報道機関だけが注目されやすい可能性を読み取れる。

(3)は、振込通知書の誤送付に関する記事には「年金振込通知書」と書く記事と「年金通知書」と書く記事が混在していることを指している。テキスト分析の際には、同一の物に対して複数の表現が存在することに留意する必要がある。

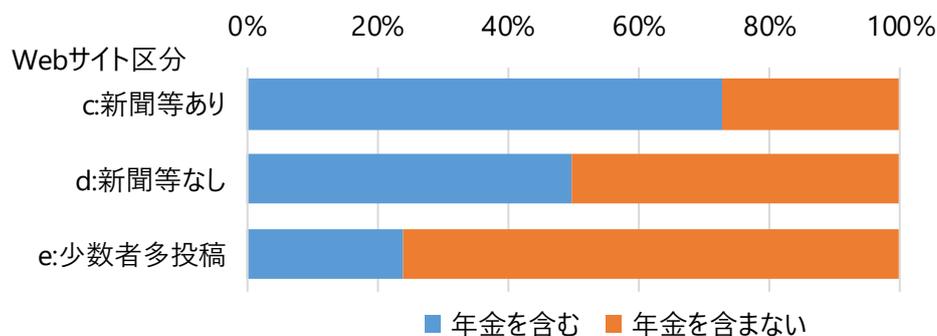
(4)は、端的に言えばリンク先の Web ページのタイトルに「年金」が入っていないものが存在することを指している。「高齢者」や「終活」なども年金と関連する単語ではあるが、機械的に割り切って、

<sup>4</sup> URL パラメータを詳細に分析することでどのようにしてツイートが作成されたかを確認できる可能性はあるが、Web サイトによって URL パラメータの仕様が異なるため本稿では分析していない。

Web ページのタイトルに「年金」が入っているか否かと Web サイトの区分との関係を、見たのが図表 10 である。全体では約半数が「年金」を含んでいるが、「c:新聞等あり」に区分される報道機関等の Web サイトのページが参照される場合には、その 7 割強で Web ページのタイトルに「年金」が含まれていた。他方で「e:少数者多投稿」に区分される Web サイトのページが参照される場合には、その 7 割強で Web ページのタイトルに「年金」が含まれなかった。なお、「d:新聞等なし」に区分される Web サイトのページが参照される場合には、「年金」を含む場合と含まない場合とが半数ずつであった。

(5)は、あるページを参照してツイートを投稿したユーザーは、そのページを再度参照してツイートを投稿しない傾向が強いことを示している。これは、報道機関等の Web サイトに掲載されている記事についてコメントを発信した投稿者は、その記事に対する別のコメントの発信や同じコメントの拡散を行っていないことを示している。このようなツイートを分析することで、記事に対する投稿者の印象を把握できる可能性が示唆される。

図表 10 リンク先 Web サイトの区分と Web ページのタイトルに「年金」を含むか否かの関係  
(ツイート数の比較)



## (2) 投稿日とリンク先 Web ページ区分ごとの傾向や特徴：タイトルに年金を含むとニュースに連動

この分析に基づいて、リンク先が Twitter 以外の Web ページか否かとリンク先 Web ページのタイトルに「年金」を含むか否かで区分し、投稿日との関係を確認したのが図表 11 である（「a:リンクなし」と「b:Twitter」は図表 8 と同じである）。

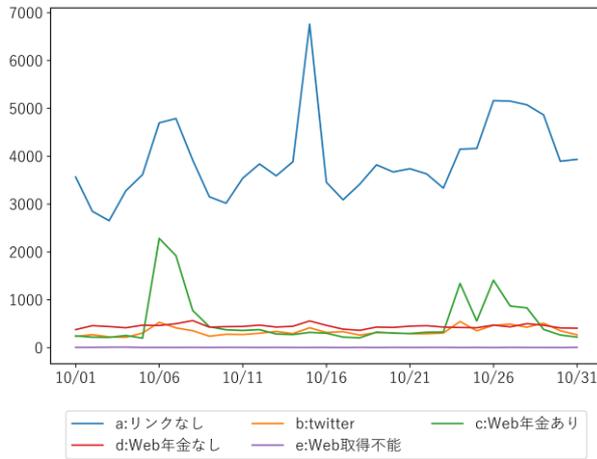
図表 11 の上段や中段を見ると、「d:Web 年金なし」すなわちリンク先 Web ページのタイトルに「年金」を含まないツイートは、年金に関するニュースとは関係なく常に一定程度のツイート数や投稿者数で推移している。リンク先 Web サイトで区分した図表 8 では「d:新聞等なし」でも年金に関するニュースとの関連がある程度は見られたのに対して、図表 11 の「d:Web 年金なし」では関連がほとんど見られない。このことから、リンク先が Twitter 以外の Web ページを参照しているツイートにおいては、リンク先 Web ページのタイトルに「年金」を含まなければ、年金のニュースに関係するツイートではない可能性が高いことが示唆される。

また、図表 11 下段の「c:Web 年金あり」を見ると、図表 8 下段の「c:新聞等あり」と比べて、いいね数やリツイート数の多さがツイート数の多さ、すなわち年金に関するニュースの状況とより連動する形で推移している。このことから、「c:Web 年金あり」に該当するツイートへのいいね数やリツイ

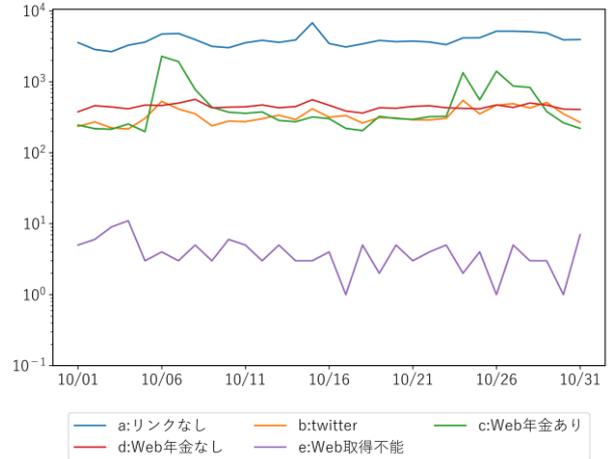
ート数を見れば、年金に関するニュースへの反響が推察できる可能性があることが示唆される。

図表 11 投稿日×リンク先 Web ページ区分ごとの傾向や特徴

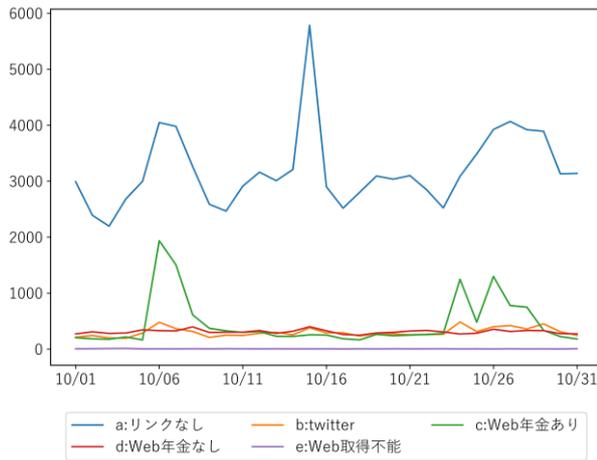
ツイート数



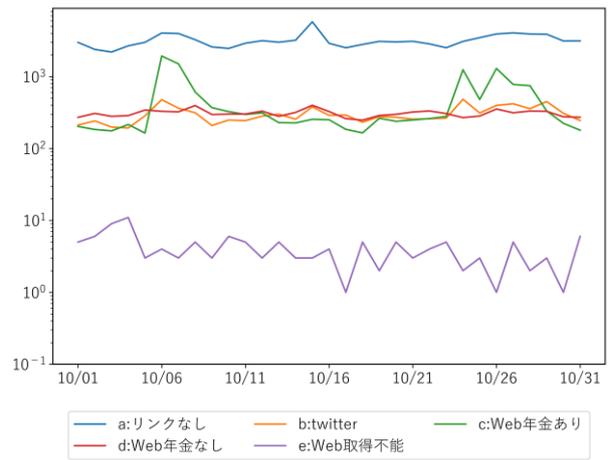
ツイート数 (対数軸)



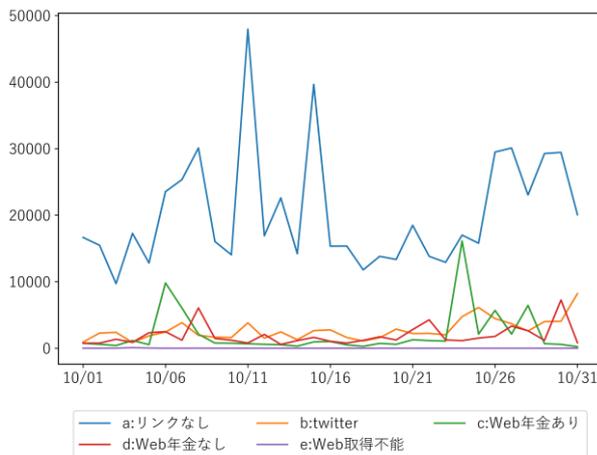
投稿者数



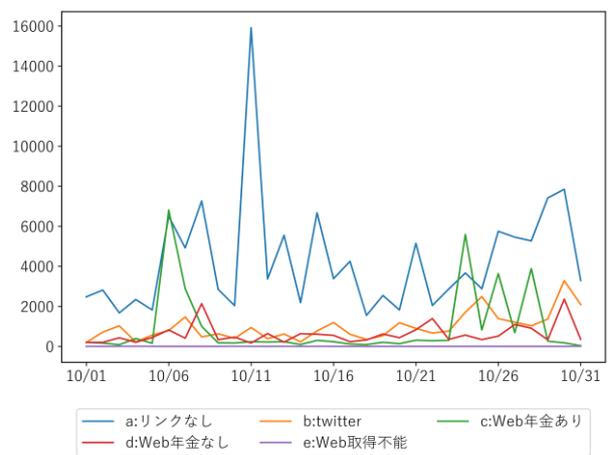
投稿者数 (対数軸)



いいね数



リツイート数



(注1) いいね数とリツイート数は、ツイートごとに投稿時点からデータ取得時点までの間隔が異なるため、大雑把な指標。

## 5 | リンク先ツイートの投稿者の傾向や特徴：一般の Web サイトと比べ、参照先の投稿者が分散

### (1) リンク先ツイート投稿者の傾向や特徴：報道機関のツイートが参照されているが、Web より低調

これまでの 2 節ではリンク先が一般の Web ページとなっているツイート（約 31,000 件）を分析したが、本節ではリンク先がツイートになっている引用ツイート（約 10,000 件）を分析する。リンク先がツイートの場合、一般の Web ページのようにページのタイトルからツイートの概要を入手できないが、リンクの URL から取得できる投稿者の ID とツイートの ID を利用する。

リンク先のツイートの投稿者 ID を名寄せして（データとして unique なものを）集計したところ、分析対象のツイートで参照されているツイートの投稿者 ID は約 4700 件であった。このうち参照しているツイート数が多い上位 20 件（上位約 1%）を列挙したのが図表 12 である。なお、Twitter は著名な組織や個人の投稿者 ID を認証しており、図表 12 には認証を受けているユーザーのみユーザー名を表示した<sup>5</sup>。

図表 12 リンク先ツイートの投稿者ごとの傾向や特徴（参照ツイート数上位 20 件）

リンク先ツイートの投稿者(ユーザー名)	ツイート数	投稿者数	いいね数	リツイート数	1人当たり ツイート数
(ツイートへのリンクなし)	152,050	79,297	755,918	182,478	1.9
日本経済新聞 電子版 (日経電子版)	141	139	449	56	1.0
Yahoo!ニュース	135	131	357	109	1.0
GPIF	100	96	2,305	823	1.0
NHKニュース	99	92	297	101	1.1
朝日新聞(asahi shimbun)	96	86	450	256	1.1
(認証不明のため非表示)	88	13	4	2	6.8
(認証不明のため非表示)	80	77	612	209	1.0
毎日新聞ニュース	72	69	594	238	1.0
ライブドアニュース	71	71	257	76	1.0
共同通信公式	62	58	177	45	1.1
ひろゆき, Hiroyuki Nishimura	61	61	712	124	1.0
(認証なしのため非表示)	59	56	589	291	1.1
(認証なしのため非表示)	58	56	211	46	1.0
時事ドットコム (時事通信ニュース)	54	51	651	310	1.1
(認証なしのため非表示)	47	46	107	23	1.0
朝日新聞デジタル	44	42	68	32	1.0
(認証なしのため非表示)	43	37	2,117	521	1.2
(認証不明のため非表示)	38	36	1,753	278	1.1
毎日新聞	37	29	311	157	1.3
NEWSポストセブン	35	32	76	11	1.1

図表 12 とリンク先 Web サイト(ドメイン)ごとに集計した図表 5 と比べると、図表 12 の各ユーザーの参照ツイート数（例えば、最も多い「日本経済新聞 電子版 (日経電子版)」で 141 件）は、図表 5

<sup>5</sup> 認証不明となっているのは、分析時点(11月中旬)において当該投稿者 ID の情報が取得できなかったものを指している。

の各 Web サイトの参照ツイート数（例えば、一般 Web サイトで最も多い「Yahoo!ニュース」で 6553 件）と比べて、かなり少ない。これは、一般 Web サイトを参照しているツイートが約 31,000 件で、参照されている Web サイト数が名寄せ後で約 2100 件だったのに対して、ツイートを参照しているツイートが約 10,000 件で、参照されている投稿者数が名寄せ後で約 4700 件と、かなり分散しているためである。これは、Twitter では最新のツイートが次々と表示されるため、参照される対象が分散しやすくなったと考えられる。

このように参照されている Web サイト数と比べて参照されている投稿者数が多くなっている一方で、図表 12 では図表 5 と比べて報道機関が上位に入っている。参照されている投稿者は分散しているものの、その中でも報道機関によるツイートは重要な情報源として利用されていることがうかがわれる。また、上位 20 位に含まれる公的機関は、図表 5（Web サイト）では「日本年金機構」だったのに対して、図表 12（ツイート投稿者）では「G P I F」（年金積立金管理運用独立行政法人）となっている。Web サイトを参照するユーザーとツイートを参照するユーザーで、関心の対象が異なる可能性がある。

## (2) 投稿日とリンク先ツイートの投稿者区分ごとの傾向や特徴：認証あり投稿者でもニュースと連動小

この分析に基づいて、リンク先のツイート投稿者が Twitter の認証を得ているか否かで区分し、投稿日との関係を確認したのが図表 13 である。

図表 13 の上段を見ると、「b:Tw 認証あり」すなわち報道機関など Twitter の認証を得ている投稿者のツイートをリンクしているツイート数や投稿者数は、10 月 6 日辺りなどの年金に関するニュースがあった日に増加する傾向があった。ただし、図表 8 下段の「c:新聞等あり」と比べると年金に関するニュースがない日との差が小さく、かつ特段のニュースがない日にも増加する傾向がある。また、図表 13 下段のいいね数やリツイート数の推移も年金に関するニュースがない日にも増加する傾向がある。このことから、「b:Tw 認証あり」と年金に関するニュースとの関係は限定的であると考えられる。

## 6 | リンク先ツイートの傾向や特徴：年金を含むツイートも参照されているが、ニュースとの連動小

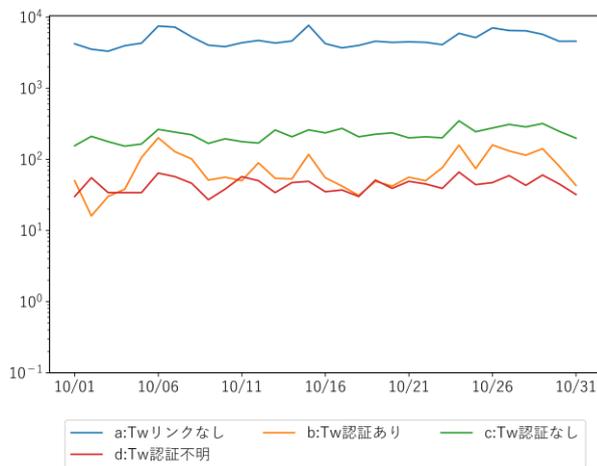
### (1) リンク先ツイートの傾向や特徴：年金を含むツイートも参照されているが、Web に比べて低調

前述したように、リンク先がツイートの場合、一般の Web ページのようにページのタイトルからツイートの概要を入手できないが、「年金」を含むツイートで単純なリツイートでないものは本稿の分析対象としてデータを貯めてあるため、このデータを使ってリンク先のツイートの「年金」が含まれるか否かをある程度確認できる<sup>6</sup>。「ある程度」なのは、本稿の分析対象には単純なリツイートなどを含めていないためである（図表 1）。

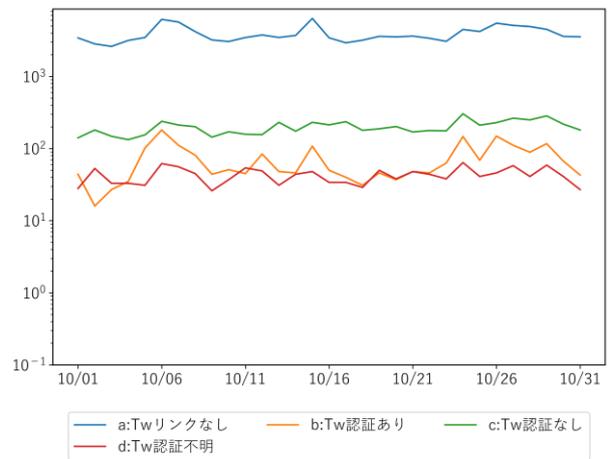
<sup>6</sup> 9 月 27 日投稿分からツイート ID を取得して、10 月投稿分のリンク先ツイートのツイート ID と照合した。

図表 13 投稿日×リンク先ツイート投稿者区分ごとの傾向や特徴

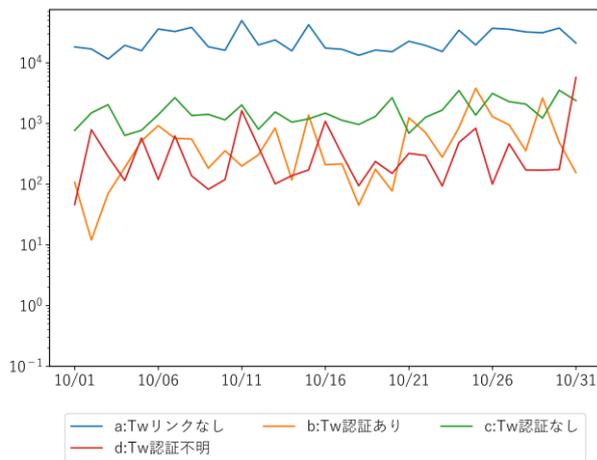
ツイート数（対数軸）



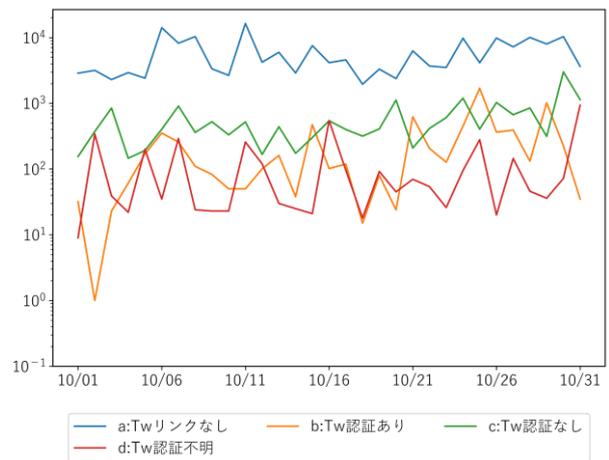
投稿者数（対数軸）



いいね数



リツイート数



(注1) いいね数とリツイート数は、ツイートごとに投稿時点からデータ取得時点までの間隔が異なるため、大雑把な指標。

リンク先ツイートのツイート ID を名寄せして（データとして unique なものを）集計したところ、分析対象のツイートで参照されているツイートの ID は約 4700 件であった。このうち参照しているツイート数が多い上位 20 件（上位約 1%）を列挙したのが図表 14 である。リンク先のツイート ID が本稿の分析対象に含まれていればツイート内容（の一部）を表示しているが、含まれていない場合は「年金なし」と表示している。例えば、参照しているツイート数が多いツイートが「年金なし」となっているが、このツイートは首相が金融所得課税の見直し検討の意向を示したことを伝える内容である。

図表 14 リンク先ツイートの傾向や特徴（参照ツイート数上位 20 件）

リンク先ツイートの内容 (一部)	ツイート数	投稿者数	いいね数	リツイート数	1人当たり ツイート数
リンクなし	152,050	79,297	755,918	182,478	1.9
(年金なし)	69	69	189	18	1.0
「年金増えたのは、運用を株でやったおかげでしょうが」自民・	66	65	415	241	1.0
日本年金機構がノートPC471台を二重調達 6300万円無	59	59	573	232	1.0
麻生氏「年金増えたのは、運用を株でやったおかげでしょうが」	48	45	574	288	1.1
【年金機構 PC1300台活用されず】日本年金機構が市町村	48	47	125	24	1.0
(年金なし)	46	12	0	0	3.8
【正式発表】日本年金機構、約97万件の年金通知書を誤	43	43	143	59	1.0
#GPIF が2001年度に市場運用を開始して以降、2021年6月	42	42	1,067	415	1.0
#GPIF は長期的な観点から年金積立金の運用を行っており、	39	39	1,169	403	1.0
(年金なし)	37	3	2	2	12.3
(年金なし)	25	25	24	7	1.0
(年金なし)	24	24	192	33	1.0
知ってますか？今年金事務所は早くから年金貰ってる人をタ-	23	23	36	12	1.0
(年金なし)	22	1	0	0	22.0
株は売って、利益を確定させて初めて「儲けた」と言えます。東	22	22	639	96	1.0
【NHKニュース速報 20:08】年金振込通知書約9 7 万件に	22	22	43	9	1.0
(年金なし)	20	1	36	29	20.0
知り合いのおばあちゃん、年金5万で介護保険1万もとられる。	19	19	46	31	1.0
(年金なし)	19	19	56	19	1.0
月 6 万円ポッキリ支給で！日本の年金、生活保護、社会保	19	1	32	20	19.0

図表 14 とリンク先 Web ページごとに集計した図表 9 と比べると、リンク先ツイートの投稿者ごとの集計 (図表 12) とリンク先 Web サイト(ドメイン)ごとの集計 (図表 5) を比べた時と同様に、図表 14 の各ツイートの参照ツイート数は、図表 9 の各 Web ページの参照ツイート数 (例えば、一般 Web サイトで最も多いページで 637 件) と比べて、かなり少ない。これは、前述したように、Twitter では最新のツイートが次々と表示されるため、参照される対象が分散しやすくなったものと考えられる。

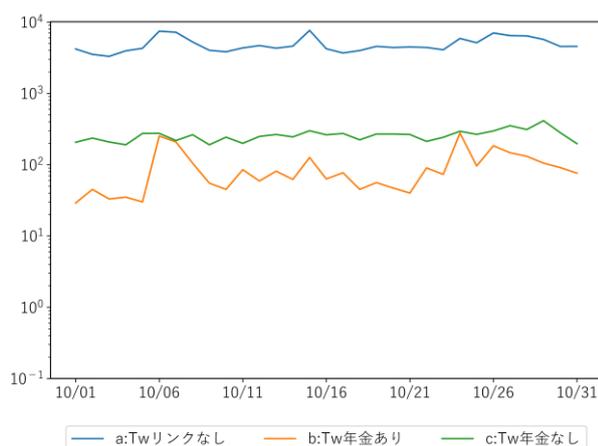
## (2) 投稿日とリンク先ツイート区分ごとの傾向や特徴：年金を含むツイートでもニュースとの連動小

次に、リンク先のツイートを「年金」を含むか否か (厳密には、本稿の分析対象データにあるか否か) で区分し、投稿日との関係を確認したのが図表 15 である。

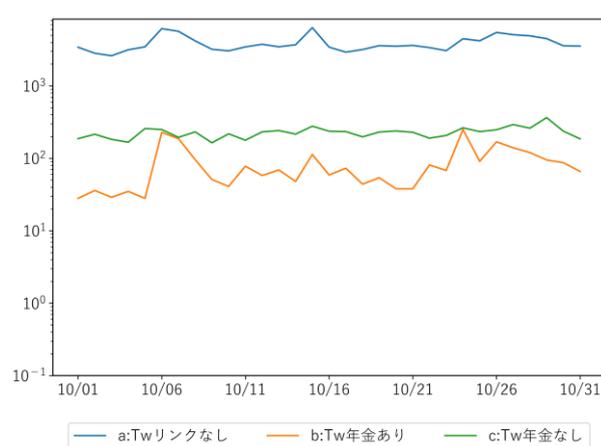
図表 15 の上段を見ると、「b:Tw 年金あり」すなわちリンク先のツイートに「年金」が含まれるツイート数や投稿者数は、10月6日辺りなどの年金に関するニュースがあった日に増加する傾向があった。ただし、図表 8 下段の「c:新聞等あり」と比べると年金に関するニュースがない日との差が小さく、かつ特段のニュースがない日にも増加する傾向がある。また、図表 15 下段のいいね数やリツイート数の推移も年金に関するニュースがない日にも増加する傾向がある。このことから、「b:Tw 認証あり」と年金に関するニュースとの関係は限定的であると考えられる。

図表 15 投稿日×リンク先ツイートの区分ごとの傾向や特徴

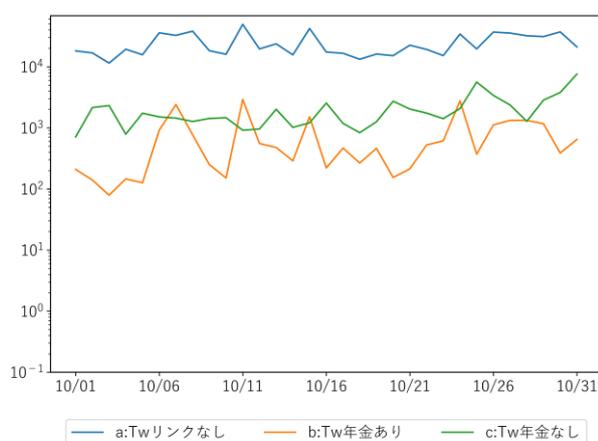
ツイート数（対数軸）



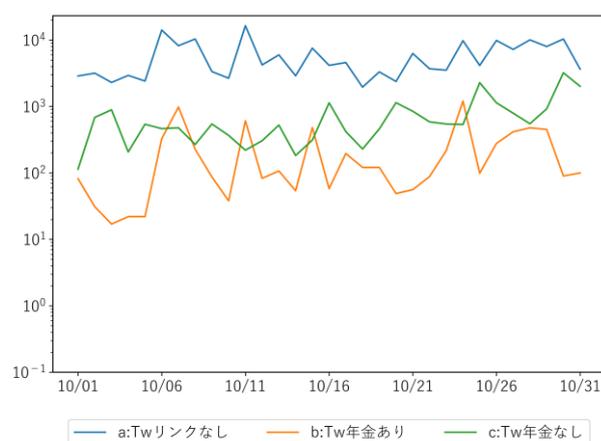
投稿者数（対数軸）



いいね数（対数軸）



リツイート数（対数軸）



（注1） いいね数とリツイート数は、ツイートごとに投稿時点からデータ取得時点までの間隔が異なるため、大雑把な指標。

## 4 ——— まとめ

本稿では、テキストマイニングの手法を使ったツイート内容の分析に先立って、前稿において2021年9月中下旬に投稿されたツイートに対して発見した傾向が同年10月のツイートにも当てはまるかを確認し、「年金」を含むツイートが何を契機に、特に年金に関するニュースを契機として投稿されているかを推察するために、本稿の分析対象ツイート（図表 1）のリンクを分析した。その結果、主に次のような傾向や特徴を発見した。

### ○前稿で発見した傾向の検証

- ・年金に関するニュースがあるとツイート数や投稿者数が増える傾向は、前稿と同じ。ただし、深夜や未明の投稿は曜日に依存する可能性がある。
- ・いいね数とリツイート数が多いごく少数のツイートがいいね数とリツイート数全体を左右する傾向は、前稿と同じ。

## ○分析対象ツイートに含まれるリンクの分析

- ・ツイートの75%はリンクなしで、25%はリンクが1個。複数のリンクがあるツイートは若干だった。
- ・リンク先のWebサイト（ドメイン）を確認したところ、リンクがあるツイートの25%は引用リツイート、16%はYahoo!ニュースで、これ以外にリンク先のトップページのタイトルに「新聞」「ニュース」「News」を含むツイートが12%あった。また、あるWebサイトを宣伝・周知する目的で投稿されている可能性が高いツイート<sup>7</sup>が9%あり、投稿日別に見るとツイート数や投稿者数がほぼ一定水準で推移していた。
- ・リンク先のWebページのタイトルに「年金」を含んでいるツイートが約半数あり、投稿日別に見るとツイート数や投稿者数などが年金に関するニュースとかなり連動していた。
- ・リンク先がツイートとなっているもの（引用リツイート）では、リンク先ツイートの投稿者やツイート当たりの参照数が、リンク先がWebサイトとなっている場合のWebサイトやWebページ当たりの参照数よりもかなり少なかった。また、年金に関するニュースとの連動が小さかった。

以上の結果から、ツイート内容のテキスト分析に向けた準備として、次のことが示唆される。

- ・年金に関する個人の態度を示すツイートに絞り込むために、あるWebサイトを参照している投稿者が少ないなどの状況から宣伝・周知する目的で投稿されているツイートを選別できる可能性が高い。
- ・リンク先のWebページのタイトルに「年金」を含んでいるツイートの投稿状況を把握することで、年金に関するニュースが話題になっている時期を機械的に把握できる可能性が高い。

本稿ではツイートに含まれるリンクを分析したが、リンクが含まれるツイートは分析対象ツイートの約25%にとどまっている。リンクが含まれないツイートの分析についても、検討を進めたい。

---

<sup>7</sup> そのWebサイトを参照している投稿者が10人未満かつ、そのWebサイトを参照しているツイートが当該投稿者1人当たり10件以上となっているWebサイトを、参照しているツイート。